INTERIOR CALDERÓN-ZYGUMUND ESTIMATES FOR SOLUTIONS TO GENERAL PARABOLIC EQUATIONS OF $p$-LAPLACIAN TYPE

TRUYEN NGUYEN

Abstract. We study general parabolic equations of the form $u_t = \text{div} A(x,t,u,Du) + \text{div} (|F|^{p-2}F) + f$ whose principal part depends on the solution itself. The vector field $A$ is assumed to have small mean oscillation in $x$, measurable in $t$, Lipschitz continuous in $u$, and its growth in $Du$ is like the $p$-Laplace operator. We establish interior Calderón-Zygmund estimates for locally bounded weak solutions to the equations when $p > 2n/(n+2)$. This is achieved by employing a perturbation method together with developing a two-parameter technique and a new compactness argument. We also make crucial use of the intrinsic geometry method by DiBenedetto $[3]$ and the maximal function free approach by Acerbi and Mingione $[1]$.

1. Introduction

Let $n \geq 2$ and $Q_6 = B_6(0) \times (-36, 36) \subset \mathbb{R}^n \times \mathbb{R}$ be the standard parabolic cylinder centered at the origin. The primary purpose of this paper is to investigate interior spatial gradient estimates of Calderón-Zygmund type for weak solutions to quasilinear parabolic equations of the form

$$u_t = \text{div} A(z,u,Du) + \text{div} (|F|^{p-2}F) + f \quad \text{in} \quad Q_6$$

with $z = (x,t) \in Q_6$, $F : Q_6 \to \mathbb{R}^n$, and $f : Q_6 \to \mathbb{R}$. Let $\mathbb{K} \subset \mathbb{R}$ be an open interval and consider general vector field

$$A = A(z,u,\xi) : Q_6 \times \overline{\mathbb{K}} \times \mathbb{R}^n \to \mathbb{R}^n$$

which is a Carathéodory map, that is, $A(z,u,\xi)$ is measurable in $z$ for every $(u,\xi) \in \overline{\mathbb{K}} \times \mathbb{R}^n$ and continuous in $(u,\xi)$ for a.e. $z \in Q_6$. We assume that there exist constants $\Lambda > 0$ and $1 < p < \infty$ such that $A$ satisfies the following structural conditions for a.e. $z \in Q_6$, all $u \in \mathbb{K}$, and all $\xi, \eta \in \mathbb{R}^n$:

$$A(z,u,\xi) - A(z,u,\eta),\xi - \eta) \geq \begin{cases} \Lambda^{-1} |\xi - \eta|^p & \text{if } p \geq 2, \\ \Lambda^{-1} (1 + |\xi| + |\eta|)^{p-2} |\xi - \eta|^2 & \text{if } 1 < p < 2, \end{cases}$$

$$|A(z,u,\xi)| \leq \Lambda(1 + |\xi|^{p-1}),$$

$$|A(z,u_1,\xi) - A(z,u_2,\xi)| \leq \Lambda|u_1 - u_2|(1 + |\xi|^{p-1}) \quad \forall u_1, u_2 \in \mathbb{K}.$$

The class of equations of the form (1.1) with $A$ satisfying (1.2)–(1.4) contains the well-known parabolic $p$-Laplace equations. More generally, it includes those of the form

$$u_t = \text{div} (a(x,t)|Du|^{p-2}Du) + \text{div} (|F|^{p-2}F) \quad \text{in} \quad Q_6.$$

The regularity theory for weak solutions of (1.5) is well developed $[1,4,6,8,12,14,16,21,22]$. In particular, Calderón-Zygmund-type estimates for (1.5) were derived in $[1,22]$ exploiting the essential fact that the

2010 Mathematics Subject Classification. 35B45, 35B65, 35K65, 35K67, 35K92.

Key words and phrases. parabolic equations, $p$-Laplacian, gradient estimates, Calderón-Zygmund estimates, compactness arguments, intrinsic geometry.

The research of the author is supported in part by a grant from the Simons Foundation (#318995).
equation is invariant with respect to the so-called intrinsic geometry \[6\]. These generalize previous results obtained for elliptic equations of \(p\)-Laplacian type \[5,9,11,13\]. However, there is a great difficulty in studying \((1.5)\) compared to its elliptic counterpart since it scales differently in space and time and as a result there is no natural maximal function associated to \((1.5)\) when \(p \neq 2\). To handle this problem, a new and important maximal function free approach was developed by Acerbi and Mingione \[1\]. These other key ingredients used in \[1\] are the localization method introduced by Kinnunen and Lewis \[12\] and the celebrated \(L^\infty\) estimates due to DiBenedetto and Friedman \[7\] for spatial gradients of solutions to the frozen homogeneous equations. The result and method in \[11\] were extended further in recent articles \[2,3\] to cover equations of the form \(u_t = \text{div} \ A(x,t,Du) + \text{div} (|F|^{p-2}F) + f\).

The aim of this paper is to address Calderón-Zygmund-type estimates for a new class of parabolic equations whose principal parts are allowed to depend on the \(u\) variable. We study general parabolic equations of the form \((1.1)\) which includes equations describing \(p\)-harmonic flows. It is worth pointing out that this class of equations is not invariant with respect to the intrinsic geometry due to the dependence of \(A\) on \(u\). Nevertheless, we are able to establish the following main result about \((1.5)\) estimates for \(Du\). Hereafter, we denote \(Q(z,\theta):= B_{\theta}(\bar{x}) \times (\bar{t} - \theta, \bar{t} + \theta)\) for \(\bar{z} = (\bar{x}, \bar{t})\). Also for a ball \(B \subset \mathbb{R}^n\), \(A_B(t,u,\xi) := \frac{1}{|B|} \int_B A(x,t,u,\xi) \, dx\) is the average of \(A\) with respect to the \(x\) variable.

**Theorem 1.1.** Let \(p > 2n/(n+2)\) and \(A : Q_0 \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n\) be a Carathéodory map such that \(\xi \mapsto A(z, u, \xi)\) is differentiable on \(\mathbb{R}^n \setminus \{0\}\) for a.e. \(z \in Q_0\) and all \(u \in \mathbb{L}\). Assume that \(A(\cdot, \cdot, 0) = 0\) and \(A\) satisfies the following conditions for a.e. \(z \in Q_0\) and all \((u, \xi) \in \mathbb{L} \times (\mathbb{R}^n \setminus \{0\})\):

\[
\begin{aligned}
&\left(\partial_\xi A(z, u, \xi) \eta, \eta\right) \\&\left|\partial_\xi A(z, u, \xi)\right| \\&|A(z, u_1, \xi) - A(z, u_2, \xi)| \\
\end{aligned}
\geq \Lambda^{-1}(\mu^2 + |\xi|^2)^{\frac{p-2}{2}} |\eta|^2, \\
\Lambda (\mu^2 + |\xi|^2)^{\frac{p-2}{2}}, \\
\Lambda |u_1 - u_2|(\mu^2 + |\xi|^2)^{\frac{p-2}{2}} \quad \forall u_1, u_2 \in \mathbb{L}
\end{equation}

for some constants \(\Lambda > 0\) and \(\mu \in [0,1]\). Then for any \(M_0 \in (0, \infty)\) and \(q > 1\), there exists \(\delta > 0\) depending only on \(p, q, n, M_0, \Lambda, \mathbb{L}\), and \(\mathbb{K}\) such that: if

\[
\sup_{z=(x,t) \in Q_0, \ Q_0 \times (\bar{t} - \theta, \bar{t} + \theta)} \int_{Q_0} \sup_{u \in \mathbb{L} \times (\mathbb{R}^n \setminus \{0\})} \left|\frac{A(x,t,u,\xi) - A_{B_\theta}(t,u,\xi)}{1 + |\xi|^{p-1}}\right| \, dx \, dt \leq \delta^p
\]

and \(u\) is a weak solution to \((1.1)\) with \(|u|_{L^\infty(Q_0)} \leq M_0\), we have

\[
\int_{Q_3} |Du|^p q \, dz \leq C \left\{ 1 + \left( \int_{Q_6} (|Du|^p + |F|^p) \, dz \right)^{\frac{p}{p+2(n+2)}} \left( \int_{Q_6} |F|^p \, dz \right)^{\frac{2(n+2)}{p+2(n+2)}} \right\}^{\frac{1}{d(q-1)}}.
\]

Here \(C > 0\) is a constant depending only on \(p, q, n, M_0, \Lambda, \mathbb{L}\), and \(d \geq 1\) and \(\hat{p} > 1\) are the numbers given by

\[
d := \left\{ \begin{array}{ll}
\frac{p}{(n+2)p-2n} & \text{if } p \geq 2, \\
\frac{2n}{n+2} & \text{if } \frac{2n}{n+2} < p < 2
\end{array} \right. \quad \text{and} \quad \hat{p} := \frac{p(n+2) - n}{p(n+1) - n}.
\]

This result generalizes the gradient estimates obtained in \[18\] Theorem 1.6] for the case \(\mathbb{K} = [0,1]\) and \(A(z, u, Du) = (1 + au) a(z) Du\) with \(a > 0\) being a constant. In Theorem 111 \(A\) is only assumed to be measurable in the time variable. As \((1.7)\) is automatically satisfied when \(x \mapsto A(x,t,u,\xi)\) is of vanishing mean oscillation, condition \((1.7)\) allows \(A\) to be discontinuous in \(x\). It is also well known that some smallness
condition in $x$ for $A$ is necessary even in the linear case. On the other hand, we show under merely structural conditions (1.2)–(1.3) for $A$ that spatial gradients of weak solutions to (1.1) enjoy the higher integrability in the sense of Elcrat and Meyers [20] (see Theorem 2.6).

We prove Theorem 1.1 by using a perturbation argument together with the intrinsic geometry method [6]. But as (1.1) is not invariant with respect to this intrinsic geometry, we are led to deal with a rescaling equation which depends on two parameters (see equation (2.3)). Then by employing the localization method [12] and the maximal function free approach [11], we demonstrate that $L^q$ estimates for $Du$ can be derived as long as gradients of solutions to the two-parameter equation can be approximated by bounded gradients in a fashion that is independent of the parameters (Theorem 4.2). The remaining and key part is to prove that there exists such Lipschitz approximation property. We achieve this through a delicate compactness argument involving two scaling parameters, and by using an important gradient bound in [14] which generalizes the fundamental $L^\infty$ gradient estimate by DiBenedetto and Friedman [7]. The compactness procedure consists of two main steps and is employed to compare gradients of solutions of our two-parameter equation to those of the corresponding frozen equation. In the first step, we reduce the problem to the homogeneous case (Lemma 5.2). We then handle the homogeneous equation in the second step (Lemma 5.3) by making use of the higher integrability stated in Theorem 2.6. It is crucial for our purpose that the constants $\delta$ in these two lemmas can be chosen to be independent of the parameters. This two-parameter technique was introduced in our recent paper [18] where parabolic equations whose principal parts are linear in the gradient variable were considered. The technique was further extended in [23] to deal with quasilinear elliptic equations of $p$-Laplacian type. However, the arguments in [18, 23] do not work for the equations under consideration since (1.1) is degenerate/singular and it scales differently in time and space. We overcome this by a different approach in Section 5 which exploits the nature of evolutionary equations and Gronwall type inequality. Another nice feature of this approach is that it works well with highly nonlinear equations and allows us to completely avoid using the Minty-Browder’s technique as in [23]. As a consequence we do not need to impose any condition on $A$ in the time variable except its measurability.

The organization of the paper is as follows. We present some basic properties of our equations in Section 2 where we also state Theorem 2.6 about the higher integrability of gradients. In Section 3 we prove Proposition 3.2 about $L^q$ estimates for a general function under a decay assumption for its distribution function. In Section 4 we formulate a Lipschitz approximation property and show in Theorem 4.2 that this property implies $L^q$ estimates for $Du$ for any $q > p$. We then verify the Lipschitz approximation property in Section 5 by developing a compactness argument involving scaling parameters. The proof of our main result (Theorem 1.1) is given at the end of Subsection 5.1 by combining the mentioned ingredients and employing an important gradient bound from [14]. Section 6 is devoted to the the proof of Theorem 2.6 about the self improving property of gradients.

2. Preliminary results and higher integrability

In this section we derive some elementary estimates which will be used later. We begin with a direct consequence of structural condition (1.2) when $1 < p < 2$.

**Lemma 2.1.** Let $1 < p < 2$ and assume that $A$ satisfies (1.2). Then there exists $C_p > 0$ depending only on $p$ such that: for any $\tau > 0$, we have for a.e. $z \in Q_6$ that

$$\tau^{\frac{1}{p}-1}(1-\tau)|\xi - \eta|^p \leq \tau^\frac{1}{q}(1 + |\xi|^p) + C_p\Lambda(A(z,u,\xi) - A(z,u,\eta),\xi - \eta) \quad \forall \xi, \eta \in \mathbb{R}^n.$$
Proof. Let \( \xi, \eta \in \mathbb{R}^n \). Since \(|\xi| + |\eta| \leq 2|\xi| + |\xi - \eta|\) and \( 1 < p < 2 \), we have from (2.2) that
\[
\langle A(z, u, \xi) - A(z, u, \eta), \xi - \eta \rangle \geq \Lambda^{-1} 2^{p-2} (1 + |\xi| + |\xi - \eta|)^{p-2} |\xi - \eta|^2.
\]
Using Young’s inequality, we obtain
\[
|\xi - \eta|^p = (1 + |\xi| + |\xi - \eta|)^{\frac{2^p}{p-2}} (1 + |\xi| + |\xi - \eta|)^{\frac{p-2}{p-2}} |\xi - \eta|^p
\]
\[
\leq \tau 3^{p-2} (1 + |\xi| + |\xi - \eta|)^p + C_p \tau \tau (1 + |\xi| + |\xi - \eta|)^{p-2} |\xi - \eta|^2.
\]
This together with (2.1) yields the conclusion of the lemma.

Let us next introduce some notations that will be used throughout the paper. For \( \bar{z} = (\bar{x}, \bar{t}) \) and \( r, \theta > 0 \), we define \( Q_{r}(z, \theta) := B_{r}(\bar{x}) \times (\bar{t} - \theta, \bar{t} + \theta) \) and
\[
Q_{r}^{1}(z) := \begin{cases} 
B_{r}(\bar{x}) \times (\bar{t} - \lambda^{2-p} r^2, \bar{t} + \lambda^{2-p} r^2) & \text{if } p \geq 2, \\
B_{\frac{\lambda^{2-p} r^2}{2}}(\bar{x}) \times (\bar{t} - r^2, \bar{t} + r^2) & \text{if } 1 < p < 2.
\end{cases}
\]
Also, \( Q_{r}(\bar{z}) := B_{r}(\bar{x}) \times (\bar{t} - r^2, \bar{t} + r^2) \). For simplicity, we will always write \( B_{r} \) for \( B_{r}(0) \) and \( Q_{r} \) for \( Q_{r}(0) \). The cylinders \( Q_{r}(z) \) and \( Q_{r}^{1}(z) \) shall be called standard parabolic cylinder and intrinsic cylinder, respectively. In addition, \( \partial_{\bar{z}} Q_{r} \) denotes the standard parabolic boundary of \( Q_{r} \).

2.1. Weak solutions.

Definition 2.2 (weak solutions). Let \( A \) satisfy (1.2)-(1.3). Assume that \( \alpha > 0 \), \( Q_{\epsilon}(r, \theta) \subset Q_{0} \), \( f \in L^{1}(Q_{\epsilon}(r, \theta)) \), and \( F \in L^{p}(Q_{\epsilon}(r, \theta)) \). A map
\[
u \in C(\bar{t} - \theta, \bar{t} + \theta; L^{2}(B_{r})) \cap L^{p}(\bar{t} - \theta, \bar{t} + \theta; W^{1,p}(B_{r}))
\]
is called a weak solution to equation
\[
u_{t} = \text{div} A(z, \alpha u, Du) + \text{div} (|F|^{p-2} F) + f \quad \text{in} \quad Q_{\epsilon}(r, \theta)
\]
if \( u(z) \in \frac{1}{\alpha} \bar{u} \) for a.e. \( z \in Q_{\epsilon}(r, \theta) \) and
\[
\int_{Q_{\epsilon}(r, \theta)} \nu \varphi_{t} \, dz = \int_{Q_{\epsilon}(r, \theta)} \langle A(z, \alpha u, Du) + |F|^{p-2} F, D\varphi \rangle \, dz - \int_{Q_{\epsilon}(r, \theta)} f \varphi \, dz \quad \forall \varphi \in C_{0}^{\infty}(Q_{\epsilon}(r, \theta)).
\]

Weak solutions to (2.2) possess a modest degree of regularity in the time variable. In order to work with test functions involving the solution itself, it is therefore convenient to adopt the formulation in terms of the so-called Steklov averages. For \( g \in L^{1}(Q_{\epsilon}(r, \theta)) \) and \( 0 < h < \bar{t} + \theta \), we define the Steklov average \( [g]_{h} \) of \( g \) by
\[
[g]_{h}(x, t) := \begin{cases} 
\frac{1}{h} \int_{0}^{h} g(x, s) \, ds & \text{for } t \in (\bar{t} - \theta, \bar{t} + \theta - h], \\
0 & \text{for } t > \bar{t} + \theta - h.
\end{cases}
\]
Then if \( f \in L^{1}(Q_{\epsilon}(r, \theta)) \) for some \( l > \frac{pn}{p(n+1)-n} \), we have that (2.4) is equivalent to
\[
\int_{B_{r}(\bar{x}) \times (l)} \partial_{t} u_{h} \phi \, dx = - \int_{B_{r}(\bar{x}) \times (l)} \langle A(z, \alpha u, Du) \rangle_{h} + [|F|^{p-2} F]_{h}, D\phi \rangle \, dx + \int_{B_{r}(\bar{x}) \times (l)} [f]_{h} \phi \, dx
\]
for all \( 0 < t < \bar{t} + \theta - h \) and all \( \phi \in W^{1,p}_{0}(B_{r}(\bar{x})) \).
2.2. Scaling properties. The following result displays scaling properties of our equation.

Lemma 2.3. Assume that \( p > 1 \). Suppose \( u \) is a weak solution to equation

\[
(2.7) \quad u_t = \text{div} A(x, t, u, Du) + \text{div} (|F|^{p-2} F) + f \quad \text{in} \quad Q^1_{t_0}(\bar{z}).
\]

For \( p \geq 2 \), we define

\[
\begin{align*}
\tilde{u}(x, t) &:= \frac{u(x, \lambda x, \lambda^2 t)}{\lambda^{\frac{p-2}{2}}} , \quad \tilde{F}(x, t) := \frac{F(x, \lambda x, \lambda^2 t)}{\lambda^{\frac{p-2}{2}}} , \\
\tilde{A}(x, t, u, \xi) &:= \frac{A(x, \lambda x, \lambda^2 t, u, \lambda \xi)}{\lambda^{\frac{p-2}{2}}} , \quad \tilde{f}(x, t) := \frac{f(x, \lambda x, \lambda^2 t)}{\lambda^{\frac{p-2}{2}}} .
\end{align*}
\]

For \( p < 2 \), we define

\[
\begin{align*}
\tilde{u}(x, t) &:= \frac{u(x, \lambda x, \lambda^2 t)}{\lambda^{\frac{2}{p}}} , \quad \tilde{F}(x, t) := \frac{F(x, \lambda x, \lambda^2 t)}{\lambda^{\frac{2}{p}}} , \\
\tilde{A}(x, t, u, \xi) &:= \frac{A(x, \lambda x, \lambda^2 t, u, \lambda \xi)}{\lambda^{\frac{2}{p}}} , \quad \tilde{f}(x, t) := \frac{f(x, \lambda x, \lambda^2 t)}{\lambda^{\frac{2}{p}}} .
\end{align*}
\]

Then \( \tilde{u} \) is a weak solution to equation

\[
(2.8) \quad \tilde{u}_t = \text{div} \tilde{A}(z, \theta \lambda \tilde{u}, D\tilde{u}) + \text{div} (|\tilde{F}|^{p-2} \tilde{F}) + \tilde{f} \quad \text{in} \quad Q_\lambda,
\]

where \( \lambda = \lambda \) if \( p \geq 2 \) and \( \tilde{\lambda} = \lambda^{\frac{2}{p}} \) if \( 1 < p < 2 \).

Proof. This can be easily checked by writing out the weak formulations and making an appropriate change of variables. Let us consider only the case \( p < 2 \). For \( \varphi \in C^0_0(Q^1_{t_0}(\bar{z})) \), let

\[
\tilde{\varphi}(x, t) := \varphi(\bar{x} + \theta \lambda \frac{x}{2}, \bar{t} + \theta^2 t).
\]

Then since \( \tilde{\varphi}_t = \theta^2 \varphi_s, D\tilde{\varphi} = \theta \lambda \frac{2}{p} D_s \varphi, \) and \( D\tilde{u} = \frac{1}{\lambda} D_s u \), we see that the integral

\[
\int_{Q_\lambda} (\tilde{u} \tilde{\varphi}_t)(x, t) \, dx \, dt = \int_{Q_\lambda} \langle \tilde{A}(x, t, \theta \lambda \tilde{u}, D\tilde{u}), \tilde{F}(x, t) \rangle \, dx \, dt - \int_{Q_\lambda} (\tilde{f} \tilde{\varphi})(x, t) \, dx \, dt
\]

is equivalent to

\[
\int_{Q^1_{t_0}(\bar{z})} (u \varphi_s)(y, s) \, dy \, ds = \int_{Q^1_{t_0}(\bar{z})} \langle A(y, s, u, D_s u) + |F|^{p-2} F, D_s \varphi(y, s) \rangle \, dy \, ds - \int_{Q^1_{t_0}(\bar{z})} (f \varphi)(y, s) \, dy \, ds.
\]

Therefore, we infer that \( u \) is a weak solution of (2.7) if and only if \( \tilde{u} \) is a weak solution of (2.8).

2.3. Energy estimates. We see from Lemma 2.3 that our equations are not invariant with respect to the intrinsic geometry. This forces us to deal with equation (2.8) involving two parameters. For simplicity, we set \( \tilde{\alpha} = \theta \lambda \) and consider equation

\[
(2.9) \quad u_t = \text{div} A(z, \tilde{\alpha} u, Du) + \text{div} (|F|^{\tilde{p}-2} F) + f \quad \text{in} \quad Q_4.
\]

We now derive some elementary energy estimates for (2.9). Hereafter, \( d \geq 1 \) and \( \tilde{p} > 1 \) denote the constants given by (1.8). We also use throughout the paper that

\[
\tilde{p} = \frac{p(n+2)}{n} \quad \text{and} \quad \tilde{p}' = \frac{p(n+2)}{p(n+2) - n}.
\]

Notice that \( \tilde{p}' \) is the conjugate of \( \tilde{p} \), i.e., \( \frac{1}{\tilde{p}} + \frac{1}{\tilde{p}'} = 1 \).

Lemma 2.4. Assume that \( \tilde{\alpha} > 0 \), \( \Lambda \) satisfies (1.2)–(1.3) in \( Q_4 \), and \( \Lambda(\cdot, \cdot, 0) = 0 \). Suppose \( u \) is a weak solution of (2.9). There exists a constant \( C > 0 \) depending only on \( p, n, \) and \( \Lambda \) such that
(i) If \( p \geq 2 \), then
\[
\sup_{s \in (-9, 9)} \int_{B_3} u^2(x, s) \, dx + \int_{Q_4} |Du|^p \, dz \leq C \left( \int_{Q_4} (|u| + u^2 + |u|^p + |F|^p) \, dz + \left( \int_{Q_4} |f|^p \, dz \right)^\frac{p}{p-1} \right).
\]

(ii) If \( 1 < p < 2 \), then
\[
\sup_{s \in (-9, 9)} \int_{B_3} u^2(x, s) \, dx + \sigma \int_{Q_3} |Du|^p \, dz \leq C \int_{Q_4} (|u| + u^2) \, dz + C \left( |\sigma| \sigma_0^{\frac{1}{p-1}} \int_{Q_4} |F|^p \, dz + \sigma_0^{\frac{1}{p-1}} \left( \int_{Q_4} |f|^p \, dz \right)^\frac{p}{p-1} \right)
\]
for every \( \sigma > 0 \) small.

**Proof.** Let \( \varphi \in C_0^\infty(Q_4) \) be the standard nonnegative cut-off function which is 1 on \( Q_3 \). Using \( \phi(x) = \varphi(x, t)^p |u|_h(x, t) \) in the weak formulation \((2.6)\) and then integrating in \( t \) we get after letting \( h \to 0^+ \) that
\[
\frac{1}{2} \int_{B_4} (u^2 \varphi^p)(x, t) \, dx - \frac{p}{2} \int_{-16}^t \int_{B_4} u^2 \varphi^{p-1} \varphi_t \, dz = -\int_{-16}^t \int_{B_4} \langle A(z, \alpha u, Du) + |F|^{-2} F, D(\varphi^p u) \rangle \, dz + \int_{-16}^t \int_{B_4} f \varphi^p u \, dz
\]
for each \( s \in (-16, 16) \). Let \( K_s := B_4 \times (-16, s) \). Then it follows from the above identity, the assumption \( A(z, \alpha u, 0) = 0 \), and \((1.3)\) that
\[
\frac{1}{2} \int_{K_s} (u^2 \varphi^p)(x, s) \, dx - \frac{p}{2} \int_{K_s} u^2 \varphi^{p-1} \varphi_t \, dz + \int_{K_s} \langle A(z, \alpha u, Du) - A(z, \alpha u, 0), Du \rangle \varphi^p \, dz
\]
\[= -p \int_{K_s} \langle A(z, \alpha u, Du) + |F|^{-2} F, D \varphi \rangle u \varphi^{p-1} \, dz - \int_{K_s} |F|^{-2} \langle F, Du \rangle \varphi^p \, dz + \int_{K_s} f \varphi^p u \, dz \leq p \int_{K_s} \left[ \bar{A}(1 + |Du|^{p-1}) + |F|^{-2} |u| |D \varphi| \right] \varphi^{p-1} \, dz + \int_{K_s} |F|^{-2} |Du| \varphi^p \, dz + \int_{K_s} |f||u\varphi^p| \, dz.
\]
We next use Hölder’s inequality, the parabolic embedding (see \([6\), Proposition 3.1, page 7]), and Young’s inequality to get
\[
\int_{K_s} |f||u\varphi^p| \, dz \leq \|u\varphi^p\|_{L^p(K_s)} \|f\|_{L^{p'}(K_s)} \leq C(n, p) \left( \int_{K_s} |D(u\varphi^p)|^p \, dz \right)^\frac{1}{p} \left( \sup_{t \in (-16, s)} \int_{B_4} (u\varphi^p)^2(x, t) \, dx \right)^\frac{p}{2p} \|f\|_{L^{p'}(K_s)} \leq \varepsilon \int_{K_s} |D(u\varphi^p)|^p \, dz + \varepsilon \sup_{t \in (-16, s)} \int_{B_4} (u\varphi^p)^2(x, t) \, dx + \frac{C(n, p)}{\varepsilon^{\frac{n+2}{2}}} \|f\|_{L^{p'}(K_s)} \forall \varepsilon > 0.
\]
Applying Young’s inequality to the first two integrals on the right hand side of (2.10) and using (2.11) with a suitable choice of \( \varepsilon \), we then obtain
\[
\frac{1}{2} \int_{B_4} (u^2 \varphi^p)(x, s) \, dx + \frac{1}{2} \int_{K_{s}} |Du|^p \varphi^p \, ds \\
\leq \frac{1}{4} \sup_{s \in (-16, 16)} \int_{B_4} (u^2 \varphi^p)(x, t) \, dx + C \int_{Q_4} (|u| + u^2) \, dz + C \sigma^{1-p} \int_{Q_4} |u|^p \, dz \\
+ C \sigma^{\frac{1}{p-1}} \int_{Q_4} |F|^p \, dz + C \sigma^{\frac{(n+p)}{p(n+1)}} \left( \int_{Q_4} |f|^p \, dz \right)^{\frac{p}{p-1}} \quad \forall \sigma \in (0, 1).
\]
(2.12)

If \( p \geq 2 \), then by using structural condition (1.2) and choosing \( \sigma \) sufficiently small we arrive at
\[
\frac{1}{2} \int_{B_4} (u^2 \varphi^p)(x, s) \, dx + \frac{1}{2} \int_{K_{s}} |Du|^p \varphi^p \, ds \\
\leq \frac{1}{4} \sup_{s \in (-16, 16)} \int_{B_4} (u^2 \varphi^p)(x, t) \, dx + C \left[ \int_{Q_4} (|u| + u^2 + |u|^p + |F|^p) \, dz + \left( \int_{Q_4} |f|^p \, dz \right)^{\frac{p}{p-1}} \right]
\]
for each \( s \in (-16, 16) \). This immediately gives
\[
\sup_{s \in (-16, 16)} \int_{B_4} (u^2 \varphi^p)(x, s) \, dx \leq C \left[ \int_{Q_4} (|u| + u^2 + |u|^p + |F|^p) \, dz + \left( \int_{Q_4} |f|^p \, dz \right)^{\frac{p}{p-1}} \right].
\]
Combining this with (2.13) and using the fact \( \varphi = 1 \) on \( Q_3 \), we obtain (i). In the case \( 1 < p < 2 \), it follows from Lemma 2.1 that
\[
c^2 \tau^{-\frac{p}{2}} |Du|^p - 2c \tau^{-\frac{1}{p}} (1 + |Du|^p) \leq \langle A(z, \alpha u, Du) - A(z, \alpha u, 0), Du \rangle \quad \forall \tau \in (0, 1/2).
\]

This together with (2.12) gives
\[
\frac{1}{2} \int_{B_4} (u^2 \varphi^p)(x, s) \, dx + (c^2 \tau^{-\frac{p}{2}} - \sigma) \int_{K_{s}} |Du|^p \varphi^p \, ds \\
\leq \frac{1}{4} \sup_{s \in (-16, 16)} \int_{B_4} (u^2 \varphi^p)(x, t) \, dx + C \int_{Q_4} (|u| + u^2) \, dz + 2c \tau^{-\frac{1}{p}} \int_{Q_4} (1 + |Du|^p) \, dz \\
+ C \sigma^{1-p} \int_{Q_4} |u|^p \, dz + C \sigma^{-\frac{(n+p)}{p(n+1)}} \left( \int_{Q_4} |f|^p \, dz \right)^{\frac{p}{p-1}} \quad \text{for } s \in (-16, 16).
\]

By taking \( \tau \) such that \( c^2 \tau^{-\frac{p}{2}} = 2\sigma \), we infer as in the case \( p \geq 2 \) that (ii) holds. \( \square \)

The next lemma allows us to estimate the difference between gradients of solutions originating from different equations.

**Lemma 2.5.** Assume that \( \alpha > 0 \), and \( A, \hat{A} \) satisfy (1.2)-(1.3) in \( Q_4 \). Suppose \( u \) is a weak solution of (2.9) and \( v \) is a weak solution of
\[
\begin{align*}
\{ & v_t = \text{div} \hat{A}(z, \alpha v, Dv) \quad \text{in } Q_4, \\
& v = u \quad \text{on } \partial_\nu Q_4.
\end{align*}
\]
Then there exists \( C > 0 \) depending only on \( n, p, \) and \( \Lambda \) such that
\[
\sup_{s \in (-16, 16)} \int_{B_4} |u - v|^2 \, dx + \int_{Q_4} |Du - Dv|^p \, dz \leq C \left[ \int_{Q_4} (1 + |Du|^p + |F|^p) \, dz + \left( \int_{Q_4} |f|^p \, dz \right)^{\frac{p}{p-1}} \right].
\]
Proof. Let \( h = u - v \). Then \( h \) is a weak solution of

\[
h_t = \text{div} [\hat{A}(z, \alpha v, Du) - \hat{A}(z, \alpha v, Dv)] + \text{div} [A(z, au, Du) - \hat{A}(z, \alpha v, Du)] + \text{div} (|F|^{p-2}F) + f \text{ in } Q_4,
\]

with \( h = 0 \) on \( \partial_p Q_4 \). Multiplying the above equation by \( h \) and integrating by parts we obtain for each \( s \in (-16, 16) \) that

\[
\int_{B_4} \frac{h^2(x, s)}{2} \, dx + \int_{-16}^s \int_{B_4} \langle \hat{A}(z, \alpha v, Du) - \hat{A}(z, \alpha v, Dv), Dh \rangle \, dz \\
= - \int_{-16}^s \int_{B_4} \langle A(z, au, Du) - \hat{A}(z, \alpha v, Du) + |F|^{p-2}F, Dh \rangle \, dz + \int_{-16}^s \int_{B_4} fh \, dz.
\]

We deduce from this, structural conditions (2.12)-(2.13), and Lemma 2.1 with \( \tau = 1/2 \) that

\[
\frac{1}{2} \int_{B_4} h^2(x, s) \, dx + c(\Lambda, p) \int_{K_s} |Dh|^p \, dz - c(\Lambda, p) \int_{K_s} (1 + |Du|^p) \, dz \\
\leq \int_{K_s} \left[ 2\Lambda(1 + |Du|^{p-1}) + |F|^{p-1} \right] |Dh| \, dz + \int_{K_s} |f||h| \, dz,
\]

where \( K_s := B_4 \times (-16, s) \). Hence, applying Young’s inequality and collecting like-terms give

\[
|f||h| \, dz \leq \varepsilon \int_{K_s} |Dh|^p \, dz + \varepsilon \sup_{t \in (-16, s)} \int_{B_4} h^2(x, t) \, dx + C(n, p) \sup_{t \in (-16, s)} \|f\|_{L^{2p/(p+1)}(K_s)}^{p+2} \quad \forall \varepsilon > 0.
\]

Hence by taking \( \varepsilon = 1/(2C) \) and substituting the resulting expression into (2.14), we obtain

\[
\int_{B_4} h^2(x, s) \, dx + \frac{1}{2} \int_{K_s} |Dh|^p \, dz \\
\leq \frac{1}{2} \sup_{t \in (-16, 16)} \int_{B_4} h^2(x, t) \, dx + C \left[ \int_{Q_4} (1 + |Du|^p + |F|^p) \, dz + \|f\|_{L^{2p/(p+1)}(Q_4)}^{p+2} \right],
\]

for each \( s \in (-16, 16) \). This implies the conclusion of the lemma. \( \square \)

2.4 Higher integrability of gradients. We next state the higher integrability in the sense of Elcrat and Meyers [20] for spatial gradients of weak solutions to equation (2.2).

Theorem 2.6. Assume that \( \alpha > 0 \) and \( A \) satisfies (2.2)-(2.3). Let \( p > 2n/(n + 2) \) and suppose that \( u \) is a weak solution of (2.9). Then there exist \( \varepsilon_0 > 0 \) small and \( C > 0 \) depending only on \( \Lambda, n, \) and \( p \) such that

\[
\int_{Q_3} |Du|^{p + \varepsilon_0} \, dz \leq C \left\{ 1 + \int_{Q_3} (|Du|^p + |F|^p) \, dz + \left( \int_{Q_3} |f|^p \, dz \right)^{1 + \frac{6n}{p}} \right\}^{1 + \frac{6n}{p}}
\]

\[
(2.15)
\]

In this theorem we do not impose any smallness condition on \( A \) and this self improving property of gradients will be used to perform the perturbation analysis in Section 3. The proof of Theorem 2.6 will be given in Section 6.
3. General arguments without PDEs

In this section, we establish some general results which are independent of the PDEs under consideration.

3.1. A covering argument.

**Lemma 3.1.** Let \( p > 2n/(n+2) \) and \( 0 < R_1 < R_2 \). Assume that \( g \in L^p(Q_{R_2}) \) and \( h_1, h_2 \in L^1(Q_{R_2}) \) are nonnegative functions. Define

\[
\lambda D := \int_{Q_{R_2}} (g^p + h_1 + 1) \, dz + \frac{1}{|Q_{R_2}|} \left( \int_{Q_{R_2}} h_2 \, dz \right)^\frac{1}{\lambda} \quad \text{and} \quad B^p := \left( \frac{10R_2}{R_2 - R_1} \right)^{n+2}.
\]

Then for any \( \lambda \geq B \), there exists a sequence of disjoint intrinsic cylinders \( \{Q^i_{r_i}(z_i)\} \) with \( z_i \in Q_{R_1} \) and \( r_i \in (0, \frac{R_3 - R_1}{10}] \) that satisfies the following properties:

1. \( \int_{Q^i_{r_i}(z_i)} (g^p + h_1) \, dz + \frac{1}{|Q^i_{r_i}(z_i)|} \left( \int_{Q^i_{r_i}(z_i)} h_2 \, dz \right)^\frac{1}{\lambda} = \lambda^p \) for each \( i \).
2. \( \int_{Q^i_{r_i}(z_i)} (g^p + h_1) \, dz + \frac{1}{|Q^i_{r_i}(z_i)|} \left( \int_{Q^i_{r_i}(z_i)} h_2 \, dz \right)^\frac{1}{\lambda} < \lambda^p \) for every \( r \in (r_i, R_2 - R_1] \).
3. \( E := \{ z \in Q_{R_1} : z \text{ is a Lebesgue point of } g \text{ and } g(z) > \lambda \} \subset \bigcup_{i=1}^{\infty} Q^i_{r_i}(z_i) \).

**Proof.** The proof of this lemma can be deduced from the arguments in [11, 12]. For the sake of completeness, we reproduce it here. Observe that due to \( \lambda \geq 1 \) we have \( Q^i_r(z) \subset Q_{R_2} \) for every \( z \in Q_{R_1} \) and every \( r \leq R_2 - R_1 \).

Let \( z \in E \) be arbitrary. On one hand, we have

\[
\int_{Q^i_{r_i}(z)} (g^p + h_1) \, dz + \frac{1}{|Q^i_{r_i}(z)|} \left( \int_{Q^i_{r_i}(z)} h_2 \, dz \right)^\frac{1}{\lambda} \leq \frac{|Q_{R_1}|}{|Q^i_{r_i}(z)|} \left[ \int_{Q_{R_2}} (g^p + h_1) \, dz + \frac{1}{|Q_{R_2}|} \left( \int_{Q_{R_2}} h_2 \, dz \right)^\frac{1}{\lambda} \right]
\]

\[
\leq \lambda^p \left( \frac{R_2}{r} \right)^{n+2} \left( \frac{\lambda}{\lambda} \right)^\frac{1}{\lambda} < \lambda^p \left( \frac{10R_2}{R_2 - R_1} \right)^{n+2} \frac{1}{B^\frac{1}{\lambda}} = \lambda^p
\]

for every \( \frac{R_3 - R_1}{10} < r \leq R_2 - R_1 \). On the other hand, the Lebesgue differentiation theorem gives

\[
\liminf_{r \to 0^+} \left[ \int_{Q^i_{r}(z)} (g^p + h_1) \, dz + \frac{1}{|Q^i_{r}(z)|} \left( \int_{Q^i_{r}(z)} h_2 \, dz \right)^\frac{1}{\lambda} \right] \geq g(z)^p > \lambda^p.
\]

Thus by continuity, for each \( z \in E \) there must exist \( r_z \in (0, \frac{R_3 - R_1}{10}] \) such that

\[
\int_{Q^i_{r}(z)} (g^p + h_1) \, dz + \frac{1}{|Q^i_{r}(z)|} \left( \int_{Q^i_{r}(z)} h_2 \, dz \right)^\frac{1}{\lambda} = \lambda^p
\]

and

\[
\int_{Q^i_{r}(z)} (g^p + h_1) \, dz + \frac{1}{|Q^i_{r}(z)|} \left( \int_{Q^i_{r}(z)} h_2 \, dz \right)^\frac{1}{\lambda} < \lambda^p \quad \forall r \in (r_z, R_2 - R_1].
\]

Hence by the Vitali covering lemma, one can extract a countable subcollection of disjoint *intrinsic cylinders* \( \{Q^i_{r_i}(z_i)\} \) satisfying

\[
\bigcup_{z \in E} Q^i_{r_i}(z) \subset \bigcup_{i=1}^{\infty} Q^i_{r_i}(z_i).
\]

The lemma then follows since \( E \subset \bigcup_{z \in E} Q^i_{r_i}(z) \). \( \square \)
3.2. \textit{L}^1 \textit{estimates under a decay assumption.} For a nonnegative function \(h\) on \(Q_R\) and a number \(\lambda > 0\), we define
\[
E_h(Q_R, \lambda) := \{ z \in Q_R : h(z) > \lambda \}.
\]
In the following result, we derive \(L^1\) estimates for a general function under a decay assumption for its distribution function.

Proposition 3.2. Let \(R > 0\) and \(\lambda_0 > 0\). Let \(f, g, \tilde{g}\) be nonnegative Borel measurable functions on \(Q_{2R}\), and let \(\mu, \nu, \tilde{\nu}\) be nonnegative Borel measures on \(Q_{2R}\). Assume that there exist constants \(N \geq 1\) and \(\alpha > 0\) such that for any \(0 < R_1 < R_2 \leq 2R\) we have
\[
\mu(E_f(Q_{R_1}, 2N\lambda)) \leq \alpha \left[ \mu(E_f(Q_{R_2}, \frac{\lambda}{3})) + \nu(E_g(Q_{R_2}, \frac{\lambda}{3})) + \tilde{\nu}(E_{\tilde{g}}(Q_{R_2}, \frac{\lambda}{3}))^\beta \right]
\]
for all \(\lambda \geq \lambda_0 \left( \frac{10R}{R_2 - R_1} \frac{d(n+2)}{p} \right)\). Then for any \(l > 0\), we obtain
\[
\frac{1}{M^l} \int_{Q_{R_2}} f^l \, d\mu \leq (c_0 \lambda_0)^l \mu(Q_{R_2}) + \left( c_0 \lambda_0 \right)^l \mu(Q_{2R}) + \int_{Q_{2R}} g^l \, d\nu + M^{l-1} \left( \int_{Q_{2R}} \tilde{g}^l \, d\tilde{\nu} \right)^\beta \sum_{j=1}^\infty (\alpha M^j)^l,
\]
where \(M := \max \{ 6N, 2 \frac{d(n+2)}{p} \} \) and \(c_0 := 3^{-1} 2^6 \frac{d(n+2)}{p} \).

Proof. For any \(m \in \{0, 1, 2, \ldots\}\), let \(\rho_m := R \left( 3 - \sum_{k=0}^m \frac{1}{2^k} \right)\). Then \(\rho_0 = 2R\), and \(\rho_m \downarrow R\) as \(m \uparrow \infty\). By using (3.1) for \(R_1 \rightsquigarrow \rho_{m+1}\) and \(R_2 \rightsquigarrow \rho_m\), we have for any \(m \geq 0\) that
\[
\mu(E_f(Q_{\rho_m}, 2N\lambda)) \leq \alpha \left[ \mu(E_f(Q_{\rho_{m+1}}, \frac{\lambda}{3})) + \nu(E_g(Q_{\rho_{m+1}}, \frac{\lambda}{3})) + \tilde{\nu}(E_{\tilde{g}}(Q_{\rho_{m+1}}, \frac{\lambda}{3}))^\beta \right]
\]
for every \(\lambda \geq 2 \frac{d(n+2m+6)}{p} \lambda_0\). It follows that
\[
\mu(E_f(Q_{\rho_{m+1}}, c_0 \lambda_0 6N\lambda')) \leq \alpha \left[ \mu(E_f(Q_{\rho_m}, c_0 \lambda_0 \lambda')) + \nu(E_g(Q_{\rho_m}, c_0 \lambda_0 \lambda')) + \tilde{\nu}(E_{\tilde{g}}(Q_{\rho_m}, c_0 \lambda_0 \lambda'))^\beta \right]
\]
for \(\lambda' \geq 2 \frac{d(n+2m+6)}{p} \lambda_0\). Since \(M^m \geq 2 \frac{d(n+2m)}{p} \) and \(M \geq 6N\), by taking \(\lambda' = M^m\) we thus obtain
\[
\mu(E_f(Q_{\rho_{m+1}}, c_0 \lambda_0 M^{m+1})) \leq \alpha \left[ \mu(E_f(Q_{\rho_m}, c_0 \lambda_0 M^m)) + \nu(E_g(Q_{\rho_m}, c_0 \lambda_0 M^m)) + \tilde{\nu}(E_{\tilde{g}}(Q_{\rho_m}, c_0 \lambda_0 M^m))^\beta \right] \quad \forall m = 0, 1, \ldots
\]
By iterating and using (3.2), we arrive at:
\[
\mu(E_f(Q_{R_1}, c_0 \lambda_0 M^k)) \leq \alpha^k \mu(E_f(Q_{2R}, c_0 \lambda_0)) + \sum_{i=0}^{k-1} \alpha^{k-i} \nu(E_g(Q_{2R}, c_0 \lambda_0 M^i)) + \tilde{\nu}(E_{\tilde{g}}(Q_{2R}, c_0 \lambda_0 M^i))^\beta.
\]
In particular,
\[
\mu(E_f(Q_{R_2}, c_0 \lambda_0 M^k)) \leq \alpha^k \mu(E_f(Q_{2R}, c_0 \lambda_0)) + \sum_{i=0}^{k-1} \alpha^{k-i} I_i \quad \forall k \geq 1,
\]
where \(I_i := \nu(E_g(Q_{2R}, c_0 \lambda_0 M^i)) + \tilde{\nu}(E_{\tilde{g}}(Q_{2R}, c_0 \lambda_0 M^i))^\beta\).
Since
\[ \int_{Q_R} f^l \, d\mu = l \int_0^\infty l^{-1} \mu\{(Q_R : f > t)\} \, dt \]
\[ = l \int_0^{c_0 \lambda_0 M} l^{-1} \mu\{(Q_R : f > t)\} \, dt + \sum_{k=1}^{\infty} l \int_{c_0 \lambda_0 M^k}^{c_0 \lambda_0 M^{k+1}} l^{-1} \mu\{(Q_R : f > t)\} \, dt \]
\[ \leq (c_0 \lambda_0 M)^l \mu(Q_R) + (M^l - 1)(c_0 \lambda_0)^l \sum_{k=1}^{\infty} M^k \mu(E_f(Q_R, c_0 \lambda_0 M^k)), \]
we obtain from (3.3) that
\[ \frac{1}{(M^l - 1)(c_0 \lambda_0)^l} \int_{Q_R} f^l \, d\mu - (c_0 \lambda_0 M)^l \mu(Q_R) \leq \mu(E_f(Q_{2R}, c_0 \lambda_0)) \sum_{k=1}^{\infty} (\alpha M^l)^k + \sum_{k=1}^{\infty} \sum_{i=0}^{k-1} M^k \alpha^{-i} I_i \]
\[ = \mu(E_f(Q_{2R}, c_0 \lambda_0)) \sum_{k=1}^{\infty} (\alpha M^l)^k + \sum_{i=0}^{\infty} M^i I_i \sum_{k=0}^{\infty} (\alpha M^l)^k. \]
Moreover, as \( \hat{p} > 1 \) we have
\[ \sum_{i=0}^{\infty} M^i I_i \leq \sum_{i=0}^{\infty} M^i \nu(E_g(Q_{2R}, c_0 \lambda_0 M^l)) + \left[ \sum_{i=0}^{\infty} M^i \nu(E_g(Q_{2R}, c_0 \lambda_0 M^l)) \right]^\hat{p}. \]
These together with Remark 3.3 below imply that
\[ \int_{Q_R} f^l \, d\mu \leq (c_0 \lambda_0 M)^l \mu(Q_R) \]
\[ + \left( (M^l - 1)(c_0 \lambda_0)^l \mu(E_f(Q_{2R}, c_0 \lambda_0)) + M^l \int_{Q_{2R}} g^l \, dv + M^l \frac{M^l - 1}{(M^l - 1)^{\hat{p}}} \left( \int_{Q_{2R}} g^{\hat{p}} \, dv \right)^\frac{\hat{p}}{\hat{p}} \right) \sum_{j=0}^{\infty} (\alpha M^l)^j. \]
This gives the conclusion of the proposition. \( \square \)

**Remark 3.3.** Assume that \( V \subset \mathbb{R}^n \times \mathbb{R}_+ \) is a nonnegative Borel measure on \( V \), and \( g \in L^1(V) \) for some \( l > 0 \). Then for any \( \alpha_0 > 0 \) and \( M > 1 \), we have
\[ (M^l - 1) \left( \frac{\alpha_0}{M} \right)^j \sum_{i=0}^{\infty} M^i \nu((V : |g| > \alpha_0 M^i)) \leq \int_V |g|^l \, dv. \]
Indeed,
\[ \int_V |g|^l \, dv = l \int_0^\infty t^{-1} \nu((V : |g| > t)) \, dt \geq l \sum_{i=0}^{\infty} \int_{\alpha_0 M^i}^{\alpha_0 M^{i+1}} t^{-1} \nu((V : |g| > t)) \, dt \]
\[ \geq \sum_{i=0}^{\infty} \left[ (\alpha_0 M^i)^l - (\alpha_0 M^{i-1})^l \right] \nu((V : |g| > \alpha_0 M^i)). \]
4. Conditional $L^q$ estimates for spatial gradients

In this section, we formulate a condition guaranteeing $L^q$ estimates for spatial gradients of weak solutions to equation \((1.1)\). The verification of this condition for a large class of vector fields will be carried out in Section 5. For a vector field $\mathbf{G}(x, t, u, \xi)$ and a ball $B \subset \mathbb{R}^n$, we define

$$G_B(t, u, \xi) := \int_B \mathbf{G}(x, t, u, \xi) \, dx.$$ 

**Definition 4.1** (local Lipschitz approximation property). Assume $\mathbf{A}$ satisfies \((1.2)-(1.3)\) and $p > 1$. Given $\bar{z} = (\bar{x}, \bar{t})$, we define

\[
\bar{\mathbf{A}}(x, t, u, \xi) := \begin{cases} 
\mathbf{A}(x + \theta \xi, \bar{t} + \lambda^2 \theta^2 t, u, \lambda \xi) & \text{if } p \geq 2, \\
\mathbf{A}(x + \theta \xi, \bar{t} + \lambda^2 \theta^2 t, u, \lambda \xi) & \text{if } 1 < p < 2.
\end{cases}
\]

We say that $\mathbf{A}$ satisfies the local Lipschitz approximation property with constant $M_0 \in (0, \infty)$ if for any $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon, p, n, M_0, \Lambda, \mathbb{K}) > 0$ such that: if $\lambda \geq 1$, $0 < \theta < 2$, $Q^6_4(\bar{z}) \subset Q_6$,

\[
\int_{Q_4} \left[ \sup_{\mathbf{A} \in \mathbb{K}} \left| \mathbf{A}(x, t, u, \xi) - \bar{\mathbf{A}}_B(t, u, \xi) \right| \right] \, dx \, dt + \int_{Q_4} |\tilde{\mathbf{F}}|^p \, dz + \left( \int_{Q_4} |\tilde{\mathbf{F}}|^p \, dz \right)^{\frac{1}{p}} \leq \delta^p,
\]

and $u$ is a weak solution to

$$\tilde{u}_t = \text{div} \bar{\mathbf{A}}(x, \theta \lambda \tilde{u}, D\tilde{u}) + \text{div} (|\tilde{\mathbf{F}}|^{p-2} \tilde{\mathbf{F}}) + \tilde{f} \quad \text{in} \quad Q_4$$

satisfying $\|\tilde{u}\|_{L^\infty(Q_4)} \leq M_0/\theta \lambda$ and $\int_{Q_4} |D\tilde{u}|^p \, dz \leq 1$, then we have

\[
\int_{Q_2} |D\tilde{u} - \tilde{\mathbf{F}}|^p \, dz \leq \varepsilon^p
\]

for some function $\tilde{\mathbf{F}} \in L^\infty(Q_2; \mathbb{R}^n)$ with $\|\tilde{\mathbf{F}}\|_{L^\infty(Q_2)} \leq N$, where $N \geq 1$ is a constant depending only on $p, n, M_0, \Lambda, \mathbb{K}$. Here $\lambda = \lambda$ if $p \geq 2$ and $\lambda = \lambda^2$ if $1 < p < 2$.

The following main result of the section shows that the Lipschitz approximation property for the vector field $\mathbf{A}$ implies $L^q$ estimates for gradients of weak solutions to the corresponding equation for any $q > p$.

**Theorem 4.2.** Assume $p > 2n/(n+2)$ and $\mathbf{A}$ satisfies the local Lipschitz approximation property with constant $M_0 \in (0, \infty)$. Then for any $q > 1$, there exists $\delta > 0$ depending only on $p, q, n, M_0, \Lambda, \mathbb{K}$ such that: if

\[
\sup_{\bar{z} = (\bar{x}, \bar{t}) \in Q_3, Q_6(\bar{r} \bar{t}) \subset Q_6} \int_{Q_6(\bar{r} \bar{t})} \left[ \sup_{\mathbf{A} \in \mathbb{K}} \left| \mathbf{A}(x, t, u, \xi) - \mathbf{A}_B(t, u, \xi) \right| \right] \, dx \, dt \leq \delta^p
\]

and $u$ is a weak solution to \((1.1)\) with $\|u\|_{L^\infty(Q_4)} \leq M_0$, we have

\[
\int_{Q_3} |Du|^p \, dz \leq C \left\{ 1 + \left( \int_{Q_6} |Du|^p + |\mathbf{F}|^p \, dz \right)^{\frac{1}{p}} \left( \int_{Q_6} |f|^q \, dz \right) \right\}^{1+q(1-d)}
\]

Here $C$ is a positive constant depending only on $p, q, n, M_0, \Lambda, \mathbb{K}$.
Proof. Let $\varepsilon > 0$ be determined later, and let $\delta = \delta(\varepsilon, p, n, M_0, \Lambda, \mathcal{X}) > 0$ be the corresponding constant given by Definition 4.1. Let $0 < R_1 < R_2 \leq 6$,
\[
\lambda_0^\varepsilon := \int_{Q_6} (|D\tilde{u}|^p + \frac{1}{\delta^p} |F|^p + 1) \, dz + \frac{1}{\delta^p} \int_{Q_6} |f|^p \, dz, \quad \text{and} \quad \bar{B}_{\varepsilon} := \left(\frac{10R_2}{R_2 - R_1}\right)^{n+2}.
\]
Also, let us denote $E(Q_{R_1}, \lambda) := \{z \in Q_{R_1} : z \text{ is a Lebesgue point of } |D\tilde{u}| \text{ and } |D\tilde{u}(z)| > \lambda\}$. Then for any $\lambda \geq \bar{B}_{\varepsilon}$, we can apply Lemma 3.1 for $g \sim |D\tilde{u}|$, $h_1 \sim \delta^{-p} |F|^p$, and $h_2 \sim \delta^{-\varepsilon} |f|^p$ to obtain: there exists a sequence of disjoint intrinsic cylinders $\{Q_i^1(zi)\}$ with $z_i = (x_i, t_i) \in Q_{R_1}$ and $r_i \in (0, \frac{R_2 - R_1}{10})$ that satisfies the following properties

1. $E(Q_{R_1}, \lambda) \subset \bigcup_{i=1}^{\infty} Q_i^1(zi)$.
2. $\int_{Q_i^1(zi)} (|D\tilde{u}|^p + \frac{1}{\delta^p} |F|^p) \, dz + \frac{1}{\delta^p} \int_{Q_i^1(zi)} |f|^p \, dz = \lambda^p$ for each $i$.
3. $\int_{Q_i^1(zi)} (|D\tilde{u}|^p + \frac{1}{\delta^p} |F|^p) \, dz + \frac{1}{\delta^p} \int_{Q_i^1(zi)} |f|^p \, dz < \lambda^p$ for every $r \in (r_i, R_2 - R_1]$.

Now let us fix $i$, and note that $Q_i^1(zi) \subset Q_{R_2} \subset Q_6$ as $z_i \in Q_{R_1}$ and $10r_i \leq R_2 - R_1$. Let $\tilde{u}$, $\tilde{F}$, and $\tilde{A}$ be defined as in Lemma 2.3 with $\tilde{x} = x_i$, $\tilde{t} = t_i$, and $\theta = 5r_i/2$. Then by Lemma 2.3, we see that $\tilde{u}$ is a weak solution to the equation
\[
\tilde{u}_t = \text{div} \tilde{A}(z, \theta \tilde{u}, D\tilde{u}) + \text{div} (\tilde{F} \tilde{u}^{-p} \tilde{F}) + \tilde{f} \quad \text{in} \quad Q_4.
\]
Observe that $\|\tilde{u}\|_{L^p(Q_4)} \leq M_0/\theta \lambda$. Using $D\tilde{u} = \frac{1}{\theta} D\tilde{u}(\tilde{x})$ and the definitions of $\tilde{F}$ and $\tilde{f}$, we deduce from property 3) that
\[
\int_{Q_4} |D\tilde{u}(x, t)|^p \, dx dt = \frac{1}{\lambda^p} \int_{Q_1^1(zi)} |D\tilde{u}(z)|^p \, dz < 1,
\]
\[
\int_{Q_4} |\tilde{F}(x, t)|^p \, dx dt = \frac{1}{\lambda^p} \int_{Q_1^1(zi)} |\tilde{F}(z)|^p \, dz < \frac{1}{\lambda^p} \delta^p \lambda^p = \delta^p,
\]
and
\[
\int_{Q_4} |\tilde{f}(x, t)|^p \, dx dt \leq \frac{\lambda^p \delta^p}{\lambda^p (n+1)} \int_{Q_1^1(zi)} |\tilde{f}(z)|^p \, dz = \frac{\lambda^p \delta^p}{|Q_1^1(zi)|} \int_{Q_1^1(zi)} |\tilde{f}(z)|^p \, dz < \delta^p.
\]
Moreover, as $\lambda \geq 1$ it is clear that
\[
\int_{Q_4} \left[ \sup_{\xi \in \mathbb{R}^n} \frac{|\tilde{A}(x, t, u, \xi) - \tilde{A}_B(t, u, \xi)|}{1 + |\xi|^p} \right] \, dx dt \leq \int_{Q_1^1(zi)} \left[ \sup_{\xi \in \mathbb{R}^n} \frac{|\tilde{A}(y, s, u, \eta) - \tilde{A}_B(s, u, \eta)|}{1 + |\eta|^p} \right] \, dy ds \leq \delta^p
\]
thanks to condition (4.3), where the ball $B$ is the projection of $Q_{10r_i}(zi) \subset \mathbb{R}^n \times \mathbb{R}$ onto $\mathbb{R}^n$. Since $A$ satisfies the local Lipschitz approximation property, we conclude that there exists a function $\tilde{\Psi}_i \in L^\infty(Q_2; \mathbb{R}^n)$ such
that

\[ \|\tilde{\psi}_i\|_{L^{\infty}(Q_2)} \leq N \quad \text{and} \quad \int_{Q_2} |D\tilde{u} - \tilde{\psi}_i|^p \, dz \leq \epsilon^p \]

with \( N \geq 1 \) depending only on \( p, n, M_0, \Lambda, \) and \( \mathbb{R} \). Let us rescale back by defining

\[ \psi_i(y, s) := \begin{cases} \lambda \psi_i\left(\frac{y-y_i}{s}, \frac{s}{s}\right) & \text{if } p \geq 2, \\ \lambda \psi_i\left(\frac{y-y_i}{s}, \frac{s}{s}\right) & \text{if } \frac{2n}{n+2} < p < 2. \end{cases} \]

Then we obtain

\[ \|\psi_i\|_{L^{\infty}(Q_{\lambda,\lambda}(z_i)))} \leq \frac{NL\lambda}{\lambda^p} \quad \text{and} \quad \int_{Q_{\lambda,\lambda}(z_i))} |Du - \psi_i|^p \, dz \leq \lambda^p \epsilon^p. \tag{4.4} \]

As a consequence, we get

\[ |Q_{\lambda,\lambda}^4(z_i) \cap E(Q_{R_1}, 2N\lambda)| \leq \left| \{ z \in Q_{\lambda,\lambda}^4(z_i) : |(Du - \psi_i)| > N\lambda \} \right| \leq \frac{1}{N^p \lambda^p} \int_{Q_{\lambda,\lambda}^4(z_i))} |Du - \psi_i|^p \, dz \leq \left( \frac{\epsilon^p \lambda^p}{N} \right)^p |Q_{\lambda,\lambda}^4(z_i)|. \tag{4.5} \]

We next estimate \( |Q_{\lambda,\lambda}^4(z_i)| = 5^{n+2} |Q_{\lambda,\lambda}^4(z_i)| \) on the above right hand side. Setting \( \hat{f}(z) := \delta^{-1} |f(z)|^\frac{p}{p}. \) Then from property 2) and since \( \hat{p} > 1 \), we have

\[ |Q_{\lambda,\lambda}^4(z_i)| = \frac{1}{\lambda^p} \int_{Q_{\lambda,\lambda}^4(z_i))} |Du|^p + \frac{1}{\lambda^p} \int_{Q_{\lambda,\lambda}^4(z_i))} |F|^p \, dz + \frac{1}{\lambda^p} \int_{Q_{\lambda,\lambda}^4(z_i))} |\hat{F}|^p \, dz \]

\[ \leq \frac{1}{\lambda^p} \left[ \int_{\{Q_{\lambda,\lambda}^4(z_i) : |Du| \leq \frac{\lambda^p}{2} \}} |Du|^p \, dz + \frac{1}{\lambda^p} \int_{\{Q_{\lambda,\lambda}^4(z_i) : |F| \leq \frac{\lambda^p}{2} \}} |F|^p \, dz + \left( \int_{\{Q_{\lambda,\lambda}^4(z_i) : |\hat{F}| \leq \frac{\lambda^p}{2} \}} |\hat{F}|^p \, dz \right)^\frac{1}{3} \right] \left( \lambda^p \right)^\frac{1}{3} \left| Q_{\lambda,\lambda}^4(z_i) \right|. \]

It follows that

\[ |Q_{\lambda,\lambda}^4(z_i)| \leq \frac{C_p}{\lambda^p} \left[ \int_{\{Q_{\lambda,\lambda}^4(z_i) : |Du| \leq \frac{\lambda^p}{2} \}} |Du|^p \, dz + \frac{1}{\lambda^p} \int_{\{Q_{\lambda,\lambda}^4(z_i) : |F| \leq \frac{\lambda^p}{2} \}} |F|^p \, dz + \left( \int_{\{Q_{\lambda,\lambda}^4(z_i) : |\hat{F}| \leq \frac{\lambda^p}{2} \}} |\hat{F}|^p \, dz \right)^\frac{1}{3} \right]. \tag{4.6} \]

Using (4.4)–(4.6), we deduce that

\[ \int_{Q_{\lambda,\lambda}^4(z_i) \cap E(Q_{R_1}, 2N\lambda)} |Du|^p \, dz \leq 2^{p-1} \left( \int_{Q_{\lambda,\lambda}^4(z_i) \cap E(Q_{R_1}, 2N\lambda)} |Du - \psi_i|^p \, dz + \int_{Q_{\lambda,\lambda}^4(z_i) \cap E(Q_{R_1}, 2N\lambda)} |\psi_i|^p \, dz \right) \leq 2^{p-1} \left( \lambda^p \epsilon^p |Q_{\lambda,\lambda}^4(z_i)| + (N\lambda)^p |Q_{\lambda,\lambda}^4(z_i) \cap E(Q_{R_1}, 2N\lambda)| \right) \leq 2^p \lambda^p \epsilon^p |Q_{\lambda,\lambda}^4(z_i)| \]

\[ \leq C(n, p) \epsilon^p \left[ \int_{\{Q_{\lambda,\lambda}^4(z_i) : |Du| \leq \frac{\lambda^p}{2} \}} |Du|^p \, dz + \frac{1}{\lambda^p} \int_{\{Q_{\lambda,\lambda}^4(z_i) : |F| \leq \frac{\lambda^p}{2} \}} |F|^p \, dz + \left( \int_{\{Q_{\lambda,\lambda}^4(z_i) : |\hat{F}| \leq \frac{\lambda^p}{2} \}} |\hat{F}|^p \, dz \right)^\frac{1}{3} \right]. \]

Since \( E(Q_{R_1}, 2N\lambda) \subset E(Q_{R_1}, \lambda) \subset \bigcup_{i=1}^{\infty} Q_{\lambda,\lambda}^4(z_i) \) and \( Q_{\lambda,\lambda}^4(z_i) \) is disjoint, by taking the sum over \( i \) we obtain

\[ \int_{E(Q_{R_1}, 2N\lambda)} |Du|^p \, dz \leq C(n, p) \epsilon^p \left[ \int_{\bigcup_{i=1}^{\infty} Q_{\lambda,\lambda}^4(z_i) : |Du| \leq \frac{\lambda^p}{2} \}} |Du|^p \, dz + \frac{1}{\lambda^p} \int_{\bigcup_{i=1}^{\infty} Q_{\lambda,\lambda}^4(z_i) : |F| \leq \frac{\lambda^p}{2} \}} |F|^p \, dz + \left( \int_{\bigcup_{i=1}^{\infty} Q_{\lambda,\lambda}^4(z_i) : |\hat{F}| \leq \frac{\lambda^p}{2} \}} |\hat{F}|^p \, dz \right)^\frac{1}{3} \right]. \]
for all \( \lambda \geq \tilde{B}\lambda_0 \). Therefore, we can apply Proposition 3.2 with \( l := p(q - 1) \), \( \mu(dz) := |Du(z)|^p \, dz \), \( \nu(dz) := |F| d\lambda \), and \( \hat{v}(dz) := \hat{f}^\beta d\lambda \) to conclude that

\[
\begin{align*}
\frac{1}{M} \int_{Q_3}|Du|^l \, d\mu &\leq (c_0\lambda_0)^l \mu(Q_3) \\
&+ \left[ (c_0\lambda_0)^l \mu(Q_6) + \int_{Q_6} \frac{1}{\delta} |Du|^l \, d\nu + \frac{M^l - 1}{(M^l - 1)^\beta} \left( \int_{Q_6} \hat{f}^{\beta} \, d\hat{v} \right)^\beta \right] \sum_{j=1}^\infty \left[ C(n, p) \epsilon \rho M^l \right]^j,
\end{align*}
\]

where \( M := \max \{6N, 2^{\frac{p(n+2)}{p}}\} \) and \( c_0 := 3^{-1} 2^{\frac{p(n+2)}{p}} \). Let us now choose \( \epsilon > 0 \) such that

\[
C(n, p) \epsilon \rho M^l = \frac{1}{2}.
\]

Then with the corresponding \( \delta \), we obtain

\[
\begin{align*}
\frac{1}{M} \int_{Q_3}|Du|^{pq} \, dz &\leq 2(c_0\lambda_0)^l \int_{Q_6}|Du|^p \, dz + \int_{Q_6} \frac{1}{\delta} |Du|^{pq} \, dz + \frac{(M^l - 1)}{(M^l - 1)^\beta \delta^{pq}} \left( \int_{Q_6} |\hat{f}|^{\beta q} \, dz \right)^\beta \\
&+ \frac{M^l - 1}{(M^l - 1)^\beta} \left( \int_{Q_6} \hat{f}^{\beta q} \, dz \right)^\beta.
\end{align*}
\]

This together with the definition of \( \lambda_0 \) yields the conclusion of the theorem. \( \square \)

5. Approximating gradients of solutions

The purpose of this section is to verify the local Lipschitz approximation property for a large class of vector fields and then employ Theorem 4.2 to obtain \( L^q \) estimates for spatial gradients of weak solutions to the corresponding equations. To achieve this and for the first time, the structural condition (1.4) shall be used. Throughout this section, let \( \omega : [0, \infty) \rightarrow [0, \infty) \) be the bounded function defined by

\[
\omega(r) = \begin{cases} 
  r\Lambda & \text{if } 0 \leq r \leq 2, \\
  2\Lambda & \text{if } r > 2.
\end{cases}
\]

Notice that if \( A \) satisfies (1.3) and (1.4), then we obtain from the definition of \( \omega \) that

\[
|A(z, u_1, \xi) - A(z, u_2, \xi)| \leq \omega(|u_1 - u_2|) \left( 1 + |\xi|^{p-1} \right) \quad \forall u_1, u_2 \in \mathbb{R}.
\]

For this reason, (1.4) and (5.1) will be used interchangeably. Our aim is to approximate \( Du \) by a good vector function in \( L^p \) norm, and the following lemma is the starting point for that purpose. Let us define

\[
d_{A, \hat{A}}(z) := \sup_{\xi \in \mathbb{R}} \sup_{u \in \mathbb{R}} \frac{|A(z, u, \xi) - \hat{A}(z, u, \xi)|}{1 + |\xi|^{p-1}}.
\]

Lemma 5.1. Assume that \( \alpha > 0 \) and \( A, \hat{A} \) satisfy (1.2)-(1.3). Assume in addition that \( \hat{A} \) satisfies (1.4). Suppose \( u \) is a weak solution of (2.9) with

\[
\int_{Q_3} |Du|^p \, dz \leq C(\Lambda, p, n) \quad \text{and} \quad \int_{Q_3} |F|^p \, dz + \int_{Q_3} |f|^\beta \, dz \leq 1,
\]

and \( h \) is a weak solution of

\[
\begin{align*}
  h_i &= \text{div} \hat{A}(z, ah, Dh) \quad \text{in} \ Q_3, \\
  h &= u \quad \text{on} \ \partial_p Q_3.
\end{align*}
\]

Let \( m := u - h \). Then there exist positive constants \( C, \epsilon_0 \) depending only on \( p, n, \) and \( \Lambda \) such that
(i) If \( p \geq 2 \), then
\[
\sup_{t \in (-4,4)} \int_{B_2} m(x, t^2) \, dx + \int_{Q_2} |Dm|^p \, dz \leq C \left( \|m\|^p_{L^p(Q_2)} + \|m\|^p_{L^p(Q_2)} + \|m\|^2_{L^2(Q_2)} \right) \\
+ C \left( \left( \int_{Q_2} [\omega(\alpha|m| + d_{A, \lambda}] \, dz \right)^{\eta_0} + \|F\|^p_{L^p(Q_2)} + \left( \int_{Q_2} |f|^p \, dz \right)^q \right).
\]

(ii) If \( 1 < p < 2 \), then
\[
\sup_{t \in (-4,4)} \int_{B_2} m(x, t^2) \, dx + \sup_{t \in (-4,4)} \int_{Q_2} |Dm|^p \, dz \leq C \sigma^{\frac{2}{p-2}} + C \left( \|m\|^p_{L^p(Q_2)} + \|m\|^p_{L^p(Q_2)} + \|m\|^2_{L^2(Q_2)} \right) \\
+ C \sigma^{\frac{2}{p-2}} \left( \left( \int_{Q_2} [\omega(\alpha|m| + d_{A, \lambda}] \, dz \right)^{\eta_0} + \|F\|^p_{L^p(Q_2)} \right) + C \sigma^{\frac{2}{p-2}} \left( \int_{Q_2} |f|^p \, dz \right)^q

for every \( \sigma > 0 \) small.

Proof. It follows from Lemma 2.5 and assumption (5.3) that
\[
\int_{Q_3} |Dh|^p \, dz \leq C \int_{Q_3} (1 + |Du|^p) + \|F\|^p \, dz + C \left( \int_{Q_3} |f|^p \, dz \right)^q \leq C(\Lambda, p, n).
\]

Therefore, we can employ Theorem 2.6 for \( F = 0 \) and \( f = 0 \) to conclude that there exist \( \varepsilon_0 > 0 \) small and \( C > 0 \) depending only on \( \Lambda, n, \) and \( p \) such that
\[
\int_{Q_2} |Dh|^{p+\varepsilon_0} \, dz \leq C.
\]

Let \( \varphi \in C^0_0(Q_2) \) be the standard nonnegative cut-off function which is 1 on \( Q_2 \). Let \( K_s := B_2 \times (-25/4, s) \).

Then by using \( \varphi^p m \) as a test function in the equations for \( u \) and \( h \), we have for each \( s \in (-25/4, 25/4) \) that
\[
\int_{K_s} m \varphi^p m \, dz = \int_{K_s} \left( \hat{A}(z, ah, Dh) - A(z, au, Du) - |F|^{p-2} F, Dm\right) \varphi^p \, dz \\
+ p \int_{K_s} \left( \hat{A}(z, ah, Dh) - A(z, au, Du) - |F|^{p-2} F, D\varphi\right) m \varphi^{p-1} \, dz + \int_{K_s} f m \varphi^p \, dz.
\]

Since
\[
\int_{K_s} m \varphi^p m \, dz = \int_{K_s} \left( \left( \varphi^p m^2 \right) - \frac{P}{2} \varphi^{p-1} \varphi m^2 \right) \, dz = \frac{1}{2} \int_{B_2} (\varphi^p m^2)(x, s) \, dx - \frac{P}{2} \int_{K_s} \varphi^{p-1} \varphi m^2 \, dz,
\]
the above identity gives
\[
\frac{1}{2} \int_{B_2} (\varphi^p m^2)(x, s) \, dx + I_s = \int_{K_s} \left( \hat{A}(z, ah, Dh) - A(z, au, Dh) - |F|^{p-2} F, Dm\right) \varphi^p \, dz \\
+ p \int_{K_s} \left( \hat{A}(z, ah, Dh) - A(z, au, Du) - |F|^{p-2} F, D\varphi\right) m \varphi^{p-1} \, dz + \int_{K_s} f m \varphi^p \, dz + \frac{P}{2} \int_{K_s} \varphi^{p-1} \varphi m^2 \, dz,
\]
where
\[
I_s := \int_{K_s} \left( A(z, au, Du) - A(z, au, Dh), Dm\right) \varphi^p \, dz.
\]
As \( \hat{A}(z, ah, Dh) - A(z, au, Dh) = [\hat{A}(z, ah, Dh) - \hat{A}(z, au, Dh)] - [A(z, au, Dh) - \hat{A}(z, au, Dh)] \), we deduce from this, conditions (1.3), (5.1), and definition (5.2) that

\[
\frac{1}{2} \int_{B_{\frac{3}{2}}} (\varphi^p m^2)(x, s) \, dx + I_s \leq \int_{K_s} \left\{ [\omega(\alpha|m|) + d_{A, A}] (1 + |Dh|^\alpha p^{-1}) + |F|^p \right\} |Dm| \varphi^p \, dz \\
+ C \int_{K_s} (1 + |Dh|^\alpha p^{-1} + |Du|^\alpha p^{-1} + |F|^p) |m| \, dz + \int_{K_s} |f| \, m \varphi^p \, dz + C \int_{K_s} m^2 \, dz.
\]

Using Hölder’s inequality and (5.3)–(5.4), we can bound the above third integral by \( C\|m\|_{L^p(K_s)} \). As a consequence, we obtain

\[
\frac{1}{2} \int_{B_{\frac{3}{2}}} (m^2 \varphi^p)(x, s) \, dx + I_s \leq \int_{K_s} \left\{ [\omega(\alpha|m|) + d_{A, A}] (1 + |Dh|^\alpha p^{-1}) + |F|^p \right\} |Dm| \varphi^p \, dz \\
+ \int_{K_s} |f| \, m \varphi^p \, dz + C \left( \|m\|_{L^p(K_s)} + \|m\|^2_{L^2(K_s)} \right),
\]

(5.6)

where \( C > 0 \) depends only on \( \Lambda, p, \) and \( n \). We next estimate the two integrals on the right hand side of (5.6). Using Young’s inequality, (5.5), and the boundedness of \( d_{A, A} \) and \( \omega \), it follows for any \( \sigma > 0 \) that

\[
\begin{align*}
\int_{K_s} \left\{ [\omega(\alpha|m|) + d_{A, A}] (1 + |Dh|^\alpha p^{-1}) + |F|^p \right\} |Dm| \varphi^p \, dz & \leq \frac{C_p}{\sigma^{\frac{p}{p-1}}} \left\{ \int_{K_s} [\omega(\alpha|m|) + d_{A, A}] \left( 1 + |Dh|^p \right) \, dz + \int_{K_s} |F|^p \, dz \right\} \\
& \leq \frac{C_p}{\sigma^{\frac{p}{p-1}}} \left\{ \left( \int_{K_s} (1 + |Dh|^p)^{\frac{p+\alpha}{p}} \, dz \right)^{\frac{p}{p+\alpha}} \left( \int_{K_s} [\omega(\alpha|m|) + d_{A, A}] \left( 1 + |Dh|^p \right) \, dz \right)^{\frac{\alpha}{p+\alpha}} + \|F\|^p_{L^p(K_s)} \right\} \\
& \leq \frac{C(\Lambda, p, n)}{\sigma^{\frac{p}{p-1}}} \left\{ \left( \int_{K_s} [\omega(\alpha|m|) + d_{A, A}] \, dz \right)^{\frac{\alpha}{p+\alpha}} + \|F\|^p_{L^p(K_s)} \right\}.
\end{align*}
\]

On the other hand, as in (2.11) and by the properties of \( \varphi \) we have

\[
\int_{K_s} |f| \, m \varphi^p \, dz \leq \sigma \int_{K_s} |Dm|^p \varphi^p + |m|^p \, dz + \sigma \sup_{t \in (-\frac{3}{2}, \frac{3}{2})} \int_{B_{\frac{3}{2}}} (m^2 \varphi^p)(x, t) \, dx + \frac{C(n, p)}{\sigma^{\frac{p+1}{p}} \|f\|^p_{L^p(K_s)}}
\]

for all \( \sigma > 0 \). Therefore, we deduce from (5.6) that

\[
\begin{align*}
\frac{1}{2} \int_{B_{\frac{3}{2}}} (m^2 \varphi^p)(x, s) \, dx + I_s & \leq \sigma \int_{K_s} |Dm|^p \varphi^p \, dz + \frac{\sigma}{2} \sup_{t \in (-\frac{3}{2}, \frac{3}{2})} \int_{B_{\frac{3}{2}}} (m^2 \varphi^p)(x, t) \, dx \\
& + C \sigma^{\frac{1}{p-1}} \left\{ \left( \int_{K_s} [\omega(\alpha|m|) + d_{A, A}] \, dz \right)^{\frac{\alpha}{p+\alpha}} + \|F\|^p_{L^p(K_s)} \right\} + C \sigma^{\frac{-(p+\alpha)}{p(p+1)}} \|f\|^p_{L^p(K_s)} \\
& + C \left( \|m\|^p_{L^p(K_s)} + \|m\|_{L^2(K_s)} \right)^2 \quad \forall \sigma > 0.
\end{align*}
\]
Now if \( p \geq 2 \), then (1.2) implies that \( \Lambda^{-1} \int_{K_s} |Dm|^p \psi^p \, dz \leq I_s \). Hence by combining with (5.7) and choosing \( \sigma > 0 \) sufficiently small, we obtain

\[
\frac{1}{2} \int_{B_\frac{p}{2}} (m^2 \varphi^p)(x, s) \, dx + \frac{1}{2 \Lambda} \int_{K_s} |Dm|^p \psi^p \, dz \leq \frac{1}{4} \sup_{x \in (-\frac{25}{4}, \frac{25}{4})} \int_{B_\frac{p}{2}} (m^2 \varphi^p)(x, t) \, dx
\]

\[+ C \left\{ \left( \int_{Q_\frac{p}{4}} [\omega(\alpha|m|) + d_{A, \bar{A}}] \, dz \right) \frac{\epsilon_0}{\sigma \epsilon_0} + ||F||^p_{L^p(Q_\frac{5}{2})} + \left| \frac{1}{2} \right| \left| \frac{\epsilon_0}{\sigma \epsilon_0} \right| \right\}
\]

\[+ C \left( \int_{Q_\frac{p}{4}} [\omega(\alpha|m|) + d_{A, \bar{A}}] \, dz \right) \frac{\epsilon_0}{\sigma \epsilon_0} + ||F||^p_{L^p(Q_\frac{5}{2})} + \left( \int_{Q_\frac{p}{4}} (1 + |Du|^p) \psi^p \, dz \right) \leq I_s
\]

for every \( s \in (-\frac{25}{4}, \frac{25}{4}) \). This implies (i) since \( \varphi = 1 \) on \( Q_2 \). On the other hand, if \( 1 < p < 2 \) then Lemma 2.4 together with (5.3) yields

\[
c \tau \frac{c}{\varphi} \int_{Q_\frac{p}{4}} |Dm|^p \psi^p \, dz - 2 \tau \frac{c}{\varphi} \int_{Q_\frac{p}{4}} |Dm|^p \psi^p \, dz \leq c \tau \frac{c}{\varphi} \int_{Q_\frac{p}{4}} |Dm|^p \psi^p \, dz - 2 \tau \frac{c}{\varphi} \int_{Q_\frac{p}{4}} (1 + |Du|^p) \psi^p \, dz \leq I_s
\]

for all \( \tau > 0 \) small. By combining this with (5.7) and taking \( c \tau \frac{c}{\varphi} = 2 \sigma \), we deduce for \( \sigma > 0 \) small that

\[
\frac{1}{2} \int_{B_\frac{p}{2}} (m^2 \varphi^p)(x, s) \, dx + \sigma \int_{K_s} |Dm|^p \psi^p \, dz \leq C \sigma \frac{\epsilon_0}{\sigma \epsilon_0} + \frac{1}{4} \sup_{x \in (-\frac{25}{4}, \frac{25}{4})} \int_{B_\frac{p}{2}} (m^2 \varphi^p)(x, t) \, dx
\]

\[+ C \left\{ \left( \int_{Q_\frac{p}{4}} [\omega(\alpha|m|) + d_{A, \bar{A}}] \, dz \right) \frac{\epsilon_0}{\sigma \epsilon_0} + ||F||^p_{L^p(Q_\frac{5}{2})} \right\} + C \left( \int_{Q_\frac{p}{4}} (1 + |Du|^p) \psi^p \, dz \right) \leq I_s
\]

for every \( s \in (-\frac{25}{4}, \frac{25}{4}) \). In particular, we get

\[
\frac{1}{4} \sup_{x \in (-\frac{25}{4}, \frac{25}{4})} \int_{B_\frac{p}{2}} (m^2 \varphi^p)(x, t) \, dx \leq C \sigma \frac{\epsilon_0}{\sigma \epsilon_0} + C \sigma \frac{\epsilon_0}{\sigma \epsilon_0} \left\{ \left( \int_{Q_\frac{p}{4}} [\omega(\alpha|m|) + d_{A, \bar{A}}] \, dz \right) \frac{\epsilon_0}{\sigma \epsilon_0} + ||F||^p_{L^p(Q_\frac{5}{2})} \right\}
\]

\[+ C \left( \int_{Q_\frac{p}{4}} (1 + |Du|^p) \psi^p \, dz \right) \leq C \left( \int_{Q_\frac{p}{4}} (1 + |Du|^p) \psi^p \, dz \right) \]

This together with (5.8) gives (ii) as desired. \( \square \)

5.1. **A compactness argument.** In order to verify the local Lipschitz approximation property, we compare gradients of solutions of our equation to those of the corresponding frozen equation. To this end, we employ a compactness argument in two steps. In the first step, we reduce the problem to the homogeneous case (Lemma 5.2). We then handle the homogeneous equation in the second step (Lemma 5.3) by making use of the higher integrability stated in Theorem 2.6. It is crucial that the constants \( \delta \) in these two lemmas can be chosen to be independent of the parameter \( \alpha \).

**Lemma 5.2** (reduction to homogeneous equations). Assume that \( p > 2n/(n + 2) \) and \( M_0 \in (0, \infty) \). Let \( A \) satisfy (1.2)–(1.4), and \( A(\cdot, \cdot, 0) = 0 \). For any \( \varepsilon > 0 \), there exists \( \delta_1 > 0 \) depending only on \( \varepsilon, \Lambda, p, n, \mathbb{K} \), and \( M_0 \) such that: if \( \alpha > 0 \), \( \int_{Q_4} |F|^p \, dz + \left( \int_{Q_4} |f|^p \, dz \right)^\beta \leq \delta_1^\beta \), and \( u \) is a weak solution of

\[
u_t = \text{div} A(z, \alpha, Du) + \text{div} (|F|^p F) + f \quad \text{in} \quad Q_4
\]
which contradicts (5.12). Thus, we conclude that

$$
\|u\|_{L^\infty(Q_4)} \leq \frac{M_0}{\alpha}
$$

and $w$ is a weak solution of

$$
\begin{cases}
    w_i = \text{div } A(z, \alpha w, Dw) & \text{in } Q_\frac{4}{3}, \\
    w = u & \text{on } \partial_p Q_\frac{4}{3},
\end{cases}
$$

then

$$
\int_{Q_3} |Du - Dw|^p \, dz \leq \epsilon_0^p.
$$

**Proof.** We prove (5.9) by contradiction. Suppose that estimate (5.9) is not true. Then there exist $\epsilon_0$, $p$, $\Lambda$, $n$, $\mathbb{K}$, $M_0$, a sequence of positive numbers $(\alpha_k)_{k=1}^\infty$, a sequence $(A^k)_{k=1}^\infty$ satisfying structural conditions (1.2)–(1.4) and $A^k(\cdot, \cdot, 0) = 0$, and sequences of functions $(F_k)_{k=1}^\infty$, $(f_k)_{k=1}^\infty$, $(u^k)_{k=1}^\infty$ such that

$$
\int_{Q_4} |F_k|^p \, dz + \left( \int_{Q_4} |f_k|^p \, dz \right) \leq \frac{1}{k^p},
$$

$u^k$ is a weak solution of

$$
u^k_i = \text{div } A^k(z, \alpha_k u^k, Du^k) + \text{div } (|F_k|^{p-2}F_k) + f_k \quad \text{in } Q_4
$$

with

$$
\|u^k\|_{L^\infty(Q_4)} \leq \frac{M_0}{\alpha_k}
$$

and

$$
\int_{Q_4} |Du^k|^p \, dz \leq 1,
$$

(5.12)

$$
\int_{Q_3} |Du - Dw|^p \, dz > \epsilon_0^p
$$

for all $k$. Here $w^k$ is a weak solution of

$$
\begin{cases}
    w^k_i = \text{div } A^k(z, \alpha_k w^k, Dw^k) & \text{in } Q_\frac{4}{3}, \\
    w^k = u^k & \text{on } \partial_p Q_\frac{4}{3}.
\end{cases}
$$

Using Proposition A.2, Lemma 2.3, and (5.10)–(5.11), we obtain

$$
\|w^k\|_{L^\infty(Q_\frac{4}{3})} \leq \frac{M_0}{\alpha_k}
$$

and

$$
\int_{Q_\frac{4}{3}} |Dw^k|^p \, dz \leq C(\Lambda, p, n).
$$

If the sequence $\{\alpha_k\}$ has a subsequence converging to $+\infty$, then we infer from the fact $\|Du^k - Dw^k\|_{L^p(Q_3)} \leq \|Du^k\|_{L^p(Q_3)} + \|Dw^k\|_{L^p(Q_3)}$, Lemma 2.4, estimates (5.10)–(5.11) and (5.13) that

$$
\liminf_{k \to \infty} \int_{Q_3} |Du - Dw|^p \, dz = 0
$$

which contradicts (5.12). Thus, we conclude that $\{\alpha_k\}$ is bounded and hence there exist a subsequence (still labeled $\{\alpha_k\}$) and a constant $\alpha \in [0, \infty)$ such that $\alpha_k \to \alpha$. Since $\alpha$ could be zero, the sequences $\{u^k\}$ and $\{w^k\}$ might be unbounded in $L^p(Q_\frac{4}{3})$. In spite of that, we claim that: up to a subsequence, there holds

$$
\lim_{k \to \infty} \left[ \|u^k - w^k\|_{L^p(Q_\frac{4}{3})} + \|u^k - w^k\|_{L^2(Q_\frac{4}{3})} \right] = 0.
$$
In order to prove (5.14), we first note that by applying Lemma 2.5 for \( u \rightsquigarrow u^k, F \rightsquigarrow F_k, f \rightsquigarrow f_k, v \rightsquigarrow w^k, \) and using (5.10)–(5.11), we get

\[
\sup_{r \in (\frac{N}{2}, 1)} \int_{B_r^2} |u^k - w^k|^2 dx + \int_{Q_r^2} |Du^k - Dw^k|^p dz \leq C.
\]

This together with the parabolic embedding (see [6, Proposition 3.1, page 7]) gives

\[
(5.15) \quad \int_{Q_r^2} |u^k - w^k|^p dz \leq C(n, p) \left( \int_{Q_r^2} |Du^k - Dw^k|^p dz \right)^{\frac{p}{p - 2}} \left( \sup_t \int_{B_r^2} |u^k - w^k|^2 dx \right)^{\frac{p}{2}} \leq C.
\]

In particular, \( \|u^k - w^k\|_{L^p(Q_r^2)} \leq C. \) Thus there exist subsequences, still denoted by \( \{u^k\} \) and \( \{w^k\} \), and a function \( m(z) \) such that \( u^k - w^k \to m \) strongly in \( L^p(Q_r^2) \) and \( D(u^k - w^k) \to Dm \) weakly in \( L^p(Q_r^2) \). We next show that \( m(z) = 0 \) for a.e. \( z \in Q_r^2 \).

Let \( m^k := u^k - w^k. \) By taking a subsequence if necessary, we can assume that \( m^k(z) \to m(z) \) for a.e. \( z \in Q_r^2. \) For \( \epsilon > 0 \) small, we define the following continuous approximation to the \( \text{sgn}^+ \) function:

\[
(5.16) \quad h_\epsilon(s) := \begin{cases} 
1 & \text{for } s \geq \epsilon, \\
\frac{s}{\epsilon} & \text{for } 0 \leq s < \epsilon \\
0 & \text{for } s < 0.
\end{cases}
\]

By using \( h_\epsilon(m^k) \) as a test function in the equations for \( u^k \) and \( w^k \) and subtracting the resulting expressions, we obtain:

\[
\int_{B_r^2} \left( \int_0^{\min\{m^k(t), 0\}} h_\epsilon(s) ds \right) dx + \int_0^t \int_{B_r^2} \langle A^k(z, \alpha_k u^k, Du^k) - A^k(z, \alpha_k u^k, Dw^k), Dm^k \rangle h_\epsilon(m^k) dz
\]

\[
= \int_0^t \int_{B_r^2} f_k h_\epsilon(m^k) dz + \int_0^t \int_{B_r^2} \langle A^k(z, \alpha_k w^k, Dw^k) - A^k(z, \alpha_k u^k, Dw^k) - |F_k|^{p-2} F_k, Dm^k \rangle h_\epsilon(m^k) dz
\]

for all \( t \in (-49/4, 49/4) \), where \( m^k(x, t) := \max\{m^k(x, t), 0\}. \) Hence, it follows from (1.4) that

\[
\int_{B_r^2} \left( \int_0^{\min\{m^k(t), 0\}} h_\epsilon(s) ds \right) dx + \int_0^t \int_{B_r^2} \langle A^k(z, \alpha_k u^k, Du^k) - A^k(z, \alpha_k u^k, Dw^k), Dm^k \rangle h_\epsilon(m^k) dz
\]

\[
\leq \int_{Q_r^2} |f_k| dz + \int_0^t \int_{B_r^2} \left[ \langle \alpha_k |A| m^k (1 + |Du^k|^p) + |F_k|^{p-1} \rangle |Dm^k| h_\epsilon(m^k) \right] dz.
\]

Let us consider the following two cases.

**Case 1:** \( p \geq 2. \) Then by applying structural condition (1.2) to the second integral in (5.17) and Young’s inequality to the last integral, we obtain after canceling like terms that

\[
\int_{B_r^2} \left( \int_0^{\min\{m^k(t), 0\}} h_\epsilon(s) ds \right) dx \leq \int_{Q_r^2} |f_k| dz + C(A, p) \int_0^t \int_{B_r^2} \left[ \alpha_k \frac{p}{p+1} (m^k)^{\frac{p-1}{p+1}} (1 + |Du^k|^p) + |F_k|^{p-1} \right] h_\epsilon(m^k) dz.
\]
Thanks to the definition of $h_e$ in (5.16) and owing to (5.13), we infer that
\[\int_{B_{\frac{7}{2}}} \left( \int_0^{\theta(x,y)} h_e(s) \, ds \right) \, dx \leq \int_{Q_{7/2}} |f_k| \, dz + C(\Lambda, p, n) \left[ \alpha_k^{\frac{p}{p-1}} \frac{1}{\varepsilon} + \frac{1}{\varepsilon} \int_{Q_{7/2}} |F_k|^p \, dz \right].\]
Hence, by letting $k \to \infty$ and using (5.10) we obtain
\[\int_{B_{\frac{7}{2}}} \left( \int_0^{\theta(x,y)} h_e(s) \, ds \right) \, dx \leq C(\Lambda, p, n) \alpha^{\frac{p}{p-1}}.\]
We next let $\varepsilon \to 0^+$ to get $\int_{B_{\frac{7}{2}}} m_+(x,t) \, dx \leq 0$ for every $t \in (-49/4, 49/4)$. We then conclude that
\[\int_{Q_{7/2}} m_+(x,t) \, dx \, dt = 0,\]
and hence $m(z) \leq 0$ for a.e. $z \in Q_{7/2}$.

**Case 2:** $\frac{2n}{n+2} < p < 2$. Then by applying Lemma 2.1 to the second integral in (5.17) and Young’s inequality to the last integral, we get for any $\tau \in (0, 1/2)$ and any $\sigma > 0$ that
\[\int_{B_{\frac{7}{2}}} \left( \int_0^{\theta(x,y)} h_e(s) \, ds \right) \, dx + \tau \int_0^\tau \int_{B_{\frac{7}{2}}} |Dm^k|^p h'_e(m^k) \, dz \, dt - 2\tau \int_0^\tau \int_{B_{\frac{7}{2}}} (1 + |Du|^p) h'_{e}(m^k) \, dz \, dt \leq \sigma \int_0^\tau \int_{B_{\frac{7}{2}}} |Dm^k|^p h'_e(m^k) \, dz \, dt + \int_{Q_{7/2}} |f_k| \, dz + C(\Lambda, p, n) \sigma^{\frac{1}{p-1}} \int_0^\tau \int_{B_{\frac{7}{2}}} \left[ \alpha_k^{\frac{p}{p-1}} (1 + |Du|^p) + |F_k|^p \right] h'_e(m^k) \, dz \, dt.\]
We now take $\tau > 0$ such that $\sigma = \sigma$, use (5.11) to bound the above third integral and use (5.13) to bound the last integral. As a consequence, we obtain
\[\int_{B_{\frac{7}{2}}} \left( \int_0^{\theta(x,y)} h_e(s) \, ds \right) \, dx \leq C(\Lambda, p) \frac{1}{\varepsilon} \sigma^{\frac{2}{p-1}} + \int_{Q_{7/2}} |f_k| \, dz + C(\Lambda, p, n) \sigma^{\frac{1}{p-1}} \left[ \alpha_k^{\frac{p}{p-1}} \frac{1}{\varepsilon^{p-1}} + \frac{1}{\varepsilon} \int_{Q_{7/2}} |F_k|^p \, dz \right] \forall \sigma > 0 \text{ small.}\]
Therefore, letting $k \to \infty$ and using (5.10) yield
\[\int_{B_{\frac{7}{2}}} \left( \int_0^{\theta(x,y)} h_e(s) \, ds \right) \, dx \leq C \left[ \frac{1}{\varepsilon^{\frac{2}{p-1}}} + \sigma^{\frac{1}{p-1}} \alpha_k^{\frac{p}{p-1}} \frac{1}{\varepsilon^{p-1}} \right] \forall \sigma > 0 \text{ small.}\]
Let us minimize the right hand side by choosing $\sigma = \varepsilon^{2-p}$ to obtain
\[\int_{B_{\frac{7}{2}}} \left( \int_0^{\theta(x,y)} h_e(s) \, ds \right) \, dx \leq C(1 + \alpha^{\frac{p}{p-1}}) \varepsilon \quad \forall \varepsilon > 0 \text{ small.}\]
We next let $\varepsilon \to 0^+$ to get as in Case 1 that $\int_{Q_{7/2}} m_+(x,t) \, dx \, dt = 0$, implying $m(z) \leq 0$ a.e. in $Q_{7/2}$. 


Thus we have shown in both cases that \( m(z) \leq 0 \) a.e. in \( Q_{2} \). By interchanging the role of \( u^{k} \) and \( w^{k} \), we also have \( m(z) \geq 0 \) a.e. in \( Q_{2} \). Thus, \( m = 0 \) in \( Q_{2} \) and so \( u^{k} - w^{k} \to 0 \) in \( L^{p}(Q_{2}) \). This implies claim \((5.14)\) in the case \( p \geq 2 \). For the case \( 2n/(n + 2) < p < 2 \), by taking \( \theta := \frac{n + 2}{n} - \frac{n}{p} \) we have \( \theta \in (0,1) \) and

\[
2 = \theta p + (1 - \theta)\frac{n + 2}{n}.
\]

Therefore, by interpolation and using estimate \((5.15)\) we obtain

\[
\int_{Q_{2}} |u^{k} - w^{k}|^{2} dz \leq \left( \int_{Q_{2}} |u^{k} - w^{k}|^{p} dz \right)^{\theta} \left( \int_{Q_{2}} |u^{k} - w^{k}|^{\frac{n}{p}} dz \right)^{1 - \theta} \leq C \left( \int_{Q_{2}} |u^{k} - w^{k}|^{p} dz \right)\theta.
\]

Thus \( \|u^{k} - w^{k}\|_{L^{2}(Q_{2})} \to 0 \) and claim \((5.14)\) follows in this case as well.

We now use \((5.14)\) to derive a contradiction. By applying Lemma \((5.1)\) for \( \alpha \to \alpha_{k}, u \to u^{k}, h \to w^{k}, F \to f_{k}, f \to f_{k} \) and using \((5.10)\) together with the facts \( \omega(r) \leq \Lambda r \) and \( \{\alpha_{k}\} \) is bounded, we obtain

\[
\int_{Q_{3}} |Dm^{k}|^{p} dz \leq C(\|m^{k}\|_{L^{p}(Q_{2})}^{2} + \|m^{k}\|_{L^{p}(Q_{2})}^{2} + \|m^{k}\|_{L^{2}(Q_{2})}^{2})
\]

\[
\quad + C\left(\|m^{k}\|_{L^{2}(Q_{2})}^{2} + \|F_{k}\|_{L^{n}(Q_{3})}^{p} + \left( \int_{Q_{2}} |f^{k}|^{p} dz \right)\theta \right)
\]

for the case \( p \geq 2 \). Letting \( k \to \infty \) and making use of \((5.10)\) and \((5.14)\), we conclude that

\[
(5.18) \quad \lim_{k \to \infty} \int_{Q_{3}} |Du^{k} - Dw^{k}|^{p} dz = \lim_{k \to \infty} \int_{Q_{3}} |Dm^{k}|^{p} dz = 0.
\]

On the other hand, for the case \( p < 2 \) we get from Lemma \((5.1)\) that

\[
\int_{Q_{3}} |Dm^{k}|^{p} dz \leq C\sigma^{-\frac{p}{\alpha}} + C\sigma^{-1}(\|m^{k}\|_{L^{p}(Q_{2})}^{p} + \|m^{k}\|_{L^{p}(Q_{2})}^{2} + \|m^{k}\|_{L^{2}(Q_{2})}^{2})
\]

\[
\quad + C\sigma^{-p}(\|m^{k}\|_{L^{2}(Q_{2})}^{p} + \|F_{k}\|_{L^{n}(Q_{3})}^{p}) + C\sigma^{-p(\alpha+1)}(\int_{Q_{2}} |f^{k}|^{p} dz)\theta
\]

for all \( \sigma > 0 \) small. By first taking \( k \to \infty \) and then taking \( \sigma \to 0^{+} \), we still arrive at \((5.18)\).

As \((5.18)\) contradicts \((5.12)\), we have produced a contradiction and the lemma is proved. \(\Box\)

In the next result, we only deal with the case \( F = 0 \) and \( f = 0 \), that is, weak solutions of homogeneous equations. The spatial gradients of these weak solutions enjoy the self improving property (Theorem \((2.6)\) which plays an important role in our proof.

**Lemma 5.3** (gradient approximation for homogeneous equations). Assume that \( p > 2n/(n + 2) \) and \( M_{0} \in (0,\infty) \). Let \( A \) satisfy \((1.2) - (1.4)\), and \( A(\cdot,\cdot,0) = 0 \). For any \( \varepsilon > 0 \), there exists \( \delta_{2} > 0 \) depending only on \( \varepsilon, \Lambda, p, n, \mathbb{R}, \) and \( M_{0} \) such that: if \( \alpha > 0 \),

\[
\int_{Q_{1}} \left[ \sup_{u \in \mathbb{R}} \sup_{\xi \in \mathbb{R}^n} \frac{|A(x,t,u,\xi) - A_{B}(t,u,\xi)|}{1 + |\xi|^{p-1}} \right] dx \leq \delta_{2}^{p},
\]
and \( w \) is a weak solution of \( w_t = \text{div} \, A(z, \alpha w, Dw) \) in \( Q_T^+ \) satisfying
\[
\|w\|_{L^\infty(Q_T^+)} \leq \frac{M_0}{\alpha} \quad \text{and} \quad \int_{Q_T^+} |Dw|^p \, dz \leq C(\Lambda, p, n),
\]
and \( v \) is a weak solution of
\[
\begin{cases}
  v_t = \text{div} \, A_B(t, \alpha v, Dv) & \text{in } Q_3, \\
  v = w & \text{on } \partial \rho Q_3,
\end{cases}
\]
then
\[
\int_{Q_3} |Dw - Dv|^p \, dz \leq \varepsilon^0.
\]

Proof. Suppose by contradiction that estimate (5.19) is not true. Then there exist \( \varepsilon_0, p, \Lambda, n, \mathbb{K}, M_0, \) a sequence of positive numbers \( \{\alpha_k\}_{k=1}^{\infty}, \) a sequence \( \{A_k\}_{k=1}^{\infty} \) satisfying structural conditions (1.2)–(1.4) and \( A^k(\cdot, \cdot, 0) = 0, \) and a sequence of functions \( \{w^k\}_{k=1}^{\infty} \) such that
\[
\int_{Q_4} \Theta_k(z) \, dz \leq \frac{1}{k^p} \quad \text{with} \quad \Theta_k(x, t) := d_{A_k^k, A_{B_4}}^k(x, t),
\]
\( w^k \) is a weak solution of \( w^k_t = \text{div} \, A^k(z, \alpha_k w^k, Dw^k) \) in \( Q_T^+ \) with
\[
\|w^k\|_{L^\infty(Q_T^+)} \leq \frac{M_0}{\alpha_k} \quad \text{and} \quad \int_{Q_T^+} |Dw^k|^p \, dz \leq C(\Lambda, p, n),
\]
\[
\int_{Q_3} |Dw^k - Dw|^p \, dz > \varepsilon^0_0 \quad \text{for all } k.
\]
Here \( v^k \) is a weak solution of
\[
\begin{cases}
  v^k_t = \text{div} \, A_{B_4}^k(t, \alpha_k v^k, Dv^k) & \text{in } Q_3, \\
  v^k = w^k & \text{on } \partial \rho Q_3.
\end{cases}
\]
We have from Theorem 2.6 and (5.21) that
\[
\int_{Q_3} |Dw^k|^p \, dz \leq C(\Lambda, n, p).
\]
Also, by using Proposition A.3, Lemma 2.5 for \( F \rightsquigarrow 0, f \rightsquigarrow 0, A \rightsquigarrow A^k, \hat{A} \rightsquigarrow A_{B_4}^k, \) and (5.21), we obtain
\[
\|v^k\|_{L^\infty(Q_3)} \leq \frac{M_0}{\alpha_k} \quad \text{and} \quad \int_{Q_3} |Dv^k|^p \, dz \leq C(\Lambda, p, n).
\]
If the sequence \( \{\alpha_k\} \) has a subsequence converging to \( +\infty, \) then we infer from Lemma 2.4, (5.21), and (5.24) that
\[
\liminf_{k \to \infty} \int_{Q_3} |Dw^k - Dw|^p \, dz = 0
\]
which contradicts (5.22). Thus, we conclude that \( \{\alpha_k\} \) is bounded and hence there exist a subsequence (still labeled \( \{\alpha_k\} \)) and a constant \( \alpha \in [0, \infty) \) such that \( \alpha_k \to \alpha. \) We claim that: up to a subsequence, there holds
\[
\lim_{k \to \infty} \left[ \|w^k - v^k\|_{L^p(Q_3)} + \|w^k - v^k\|_{L^2(Q_3)} \right] = 0.
\]
In order to prove (5.25), we first note that by applying Lemma 2.5 for $u \rightharpoonup w^k$, $F \rightharpoonup 0$, $f \rightharpoonup 0$, $v \rightharpoonup v^k$, and using (5.21), we get

$$\sup_{t \in (-9, 9)} \int_{B_3} |w^k - v^k|^2 dx + \int_{Q_3} |Dw^k - Dv^k|^p dz \leq C.$$  

This together with the parabolic embedding gives

$$(5.26) \quad \int_{Q_3} |w^k - v^k|^p \, dt \leq C.$$  

Thus there exist subsequences, still denoted by $\{w^k\}$ and $\{v^k\}$, and a function $m(z)$ such that $w^k \rightarrow v^k$ strongly in $L^p(Q_3)$ and $D(w^k - v^k) \rightharpoonup Dm$ weakly in $L^p(Q_3)$. We next show that $m(z) = 0$ in $Q_3$.

Let $m^k := w^k - v^k$. By taking a further subsequence, we can assume that $m^k(z) \rightarrow m(z)$ for a.e. $z \in Q_3$. Let $h_k(s)$ be given by (5.16). By using $h_k(m^k)$ as a test function in the equations for $w^k$ and $v^k$ and subtracting the resulting expressions, we obtain:

$$\int_{B_3} \left( \int_0^{m^k_t(x,t)} h_k(s) \, ds \right) dx + \int_0^t \int_{B_3} \langle A_{B_3}(t, \alpha_k v^k, Dw^k) - A_{B_3}(t, \alpha_k v^k, Dv^k), Dm^k \rangle h'_k(m^k) \, dz \leq \int_0^t \int_{B_3} (\Theta_k + \Lambda \alpha_k m^k)(1 + |Dw^k|^p|m^k|) \, dz.$$  

It follows by Young’s inequality to the last integral and canceling like terms that

$$\int_{B_3} \left( \int_0^{m^k_t(x,t)} h_k(s) \, ds \right) dx \leq C(\Lambda, p) \int_0^t \int_{B_3} \left[ \Theta_k^{\frac{p}{p}} + \alpha_k^{\frac{p}{p}} (m^k)^{\frac{p^*}{p}} \right] (1 + |Dw^k|^p) h'_k(m^k) \, dz \leq C(\Lambda, p) \left[ \frac{1}{\varepsilon} \int_{Q_3} \Theta_k^{\frac{p}{p}} (1 + |Dw^k|^p) \, dz + \alpha_k^{\frac{p}{p}} \varepsilon^{\frac{1}{p^*}} \int_{Q_3} (1 + |Dw^k|^p) \, dz \right] \leq C(\Lambda, p) \left[ \frac{1}{\varepsilon} \left( \int_{Q_3} \Theta_k^{\frac{p(p+\varepsilon_0)}{p^*}} \right)^{\frac{p^*}{p_0} - 1} \left( \int_{Q_3} (1 + |Dw^k|^p)^{\frac{p^*}{p_0}} \right)^{\frac{p^*}{p_0}} + \alpha_k^{\frac{p}{p}} \varepsilon^{\frac{1}{p^*}} \right] \int_{Q_3} (1 + |Dw^k|^p) \, dz.$$  

Therefore, we can use (5.23) and the fact $\{\Theta_k\}$ is bounded to deduce that

$$\int_{B_3} \left( \int_0^{m^k_t(x,t)} h_k(s) \, ds \right) dx \leq C \left[ \frac{1}{\varepsilon} \left( \int_{Q_3} \Theta_k \, dz \right)^{\frac{p^*}{p_0}} + \alpha_k^{\frac{p}{p}} \varepsilon^{\frac{1}{p^*}} \right].$$  

Hence, by letting $k \rightarrow \infty$ and making use of (5.20) we obtain

$$\int_{B_3} \left( \int_0^{m_t(x,t)} h_k(s) \, ds \right) dx \leq C \alpha^{\frac{p}{p^*}} \varepsilon^{\frac{1}{p^*}}.$$
We next let \( \varepsilon \to 0^+ \) to get \( \int_{B_3} m_+(x,t) \, dx \leq 0 \) for every \( t \in (-9,9) \). We then conclude that
\[
\int_{Q_3} m_+(x,t) \, dxdt = 0,
\]
and hence \( m(z) \leq 0 \) for a.e. \( z \in Q_3 \). The above arguments can be modified as done in the proof of Lemma 5.2 to get the same conclusion for the case \( 2n/(n+2) < p < 2 \) as well. Now by interchanging the role of \( w^k \) and \( v^k \), we also have \( m(z) \geq 0 \) for a.e. \( z \in Q_3 \). Thus, \( m = 0 \) in \( Q_3 \) and so \( w^k - v^k \to 0 \) in \( L^p(Q_3) \). In the case \( p < 2 \), we can interpolate this \( L^p \) convergence with (5.26) as in the proof of Lemma 5.2 to infer further that \( w^k - v^k \to 0 \) in \( L^2(Q_3) \). Thus, claim (5.25) is proved.

We now use (5.25) to derive a contradiction. By applying Lemma 5.1 for \( \alpha \rightsquigarrow \alpha_k, u \rightsquigarrow w^k, h \rightsquigarrow v^k, F \rightsquigarrow 0, f \rightsquigarrow 0, A \rightsquigarrow A^k, \hat{A} \rightsquigarrow A^k_{B_3} \), and using the facts \( \Theta(r) \leq \Lambda r \) and \( \{\alpha_k\} \) is bounded, we obtain
\[
\int_{Q_2} |Dm^k|^p \, dz \leq C \left[ \|m^k\|_{L^p(Q_3)}^p + \|m^k\|_{L^p(Q_3)}^p + \|m^k\|_{L^2(Q_3)}^2 \right] \quad \text{for the case } p \geq 2.
\]
Letting \( k \to \infty \) and making use of (5.20) and (5.25), we conclude that
\[
\lim_{k \to \infty} \int_{Q_2} |Dw^k - Dv^k|^p \, dz = \lim_{k \to \infty} \int_{Q_2} |Dm^k|^p \, dz = 0.
\]
On the other hand, for the case \( p < 2 \) we get
\[
\int_{Q_2} |Dm^k|^p \, dz \leq C \sigma^{\frac{p}{p-2}} + C \sigma^{-1} \left( \|m^k\|_{L^p(Q_3)}^p + \|m^k\|_{L^p(Q_3)}^p + \|m^k\|_{L^2(Q_3)}^2 \right)
\]
\[
+ C \sigma^{-1} \left( \|m^k\|_{L^1(Q_3)}^p + \|\Theta^k\|_{L^1(Q_3)}^p \right) \frac{\varepsilon^p}{p-2}
\]
for all \( \sigma > 0 \) small. By first taking \( k \to \infty \) and then taking \( \sigma \to 0^+ \), we still arrive at (5.27).

As (5.27) contradicts (5.22), we have produced a contradiction and the lemma is proved. \( \Box \)

5.2. Proof of the main gradient estimate. We need a key \( L^\infty \) gradient estimate from [14] to prove Theorem 1.1. This estimate is a generalization of the fundamental gradient estimate by DiBenedetto and Friedman [7] for the parabolic \( p \)-Laplace system (see also [6] Chapter 8). The statement below can be deduced from [14] Theorem 1.1 and the discussion therein.

Theorem 5.4 (interior Lipschitz estimate, [14]). Assume that \( p > \frac{2n}{n+2} \). Let \( h \) be a weak solution of
\[
h_t - \text{div} \, a(x,t,Dh) = 0 \quad \text{in} \quad Q_3,
\]
with the vector field \( a : Q_3 \times \mathbb{R}^n \to \mathbb{R}^n \) satisfying the assumptions
\[
\begin{align*}
\langle \partial_x a(x, t, \xi), \eta \rangle + |\partial_x a(x, t, \xi)| |\xi|^2 + |\xi|^2 |\eta|^2 & \geq \Lambda^{-1} (s^2 + |\xi|^2)^{\frac{p-1}{p}} |\eta|^2, \\
|a(x, t, \xi)| + |\partial_x a(x, t, \xi)| (s^2 + |\xi|^2)^{\frac{1}{2}} & \leq \Lambda (s^2 + |\xi|^2)^{\frac{p-1}{p}}, \\
|a(x, t, \xi) - a(\bar{x}, t, \xi)| & \leq \Lambda \omega(|x - \bar{x}|) (s^2 + |\xi|^2)^{\frac{p-1}{p}}
\end{align*}
\]
whenever \( \xi, \eta \in \mathbb{R}^n \) and \( (x,t), (\bar{x},t) \in Q_3 \). Here \( s \in [0,1] \) and \( \tilde{\omega} : [0, \infty) \to [0,1] \) is a nondecreasing function satisfying the Dini condition
\[
\int_0^{\infty} \tilde{\omega}(\rho) \frac{d\rho}{\rho} < \infty.
\]
Then \( Dh \in L^\infty_{\text{loc}}(Q_3) \) and there exists a constant \( c \) depending only on \( n, p, \Lambda, \) and \( \tilde{\omega}(\cdot) \) such that

\[
|Dh(x_0, t_0)| \leq c \left( \int_{Q(x_0, t_0)} (|Dh|^p + 1) \, dz \right)^{\frac{1}{p}}
\]

holds whenever \( Q_\epsilon(x_0, t_0) \subset Q_3 \) with \( (x_0, t_0) \) is a Lebesgue point for \( Dh \). The constant \( d \geq 1 \) is the number given by (1.8).

We are now ready to prove our main result.

**PROOF OF THEOREM 1.1** Thanks to Theorem 4.2, it is enough to prove that \( A \) admits the local Lipschitz approximation property with constant \( M_0 \). We first observe that \( A \) satisfies structural conditions (1.2) and (1.3). Indeed, the first condition in (1.6) implies (1.2) (see for example [24, Lemma 1]) and (1.3) follows from the facts

\[
|A(z, u, \xi)| \leq \Lambda (\mu^2 + |\xi|^2)^{\frac{p-2}{2}} \quad \text{if} \quad p \geq 2 \quad \text{and} \quad |A(z, u, \xi)| \leq \frac{\Lambda}{p-1} (\mu^2 + |\xi|^2)^{\frac{p-2}{2}} \quad \text{if} \quad p < 2,
\]

which are consequences of the second condition in (1.6) and the assumption \( A(\cdot, 0) = 0 \).

We next verify the Lipschitz approximation property for \( A \). Let \( \varepsilon > 0 \), and let \( \delta_1 \) and \( \delta_2 \) be the corresponding constants given by Lemma 5.2 and Lemma 5.3 respectively. Let \( \delta := \min \{\delta_1, \delta_2\} \). Now assume that \( \lambda \geq 1, 0 < \theta < 2, Q_\delta(\tilde{z}) \subset Q_6, A \) given by (4.1), \( \tilde{F} \), and \( f \) satisfy

\[
\int_{Q_4} \left[ \sup_{u \in \mathbb{R}} \sup_{t \in \mathbb{R}} \frac{|\tilde{A}(x, t, u, \xi) - \tilde{A}_B(t, u, \xi)|}{1 + |\xi|^{p-1}} \right] \, dxdt + \int_{Q_4} |\tilde{F}|^p \, dz + \left( \int_{Q_4} |\tilde{f}|^p \, dz \right) \leq \delta^p,
\]

and \( \tilde{u} \) is a weak solution to

\[
\tilde{u}_t = \text{div} \, \tilde{A}(z, \theta \hat{\lambda} \tilde{u}, D\tilde{u}) + \text{div} \, (|\tilde{F}|^{p-2} \tilde{F}) + \tilde{f} \quad \text{in} \quad Q_4
\]

with \( ||\tilde{u}||_{L^\infty(Q_4)} \leq M_0/\theta \hat{\lambda} \) and \( \int_{Q_4} |D\tilde{u}|^p \, dz \leq 1 \). We want to show that (4.2) holds true. For this, we note that \( \tilde{A}(\cdot, 0) = 0 \) and \( \tilde{A} \) satisfies conditions (1.2)–(1.4) with the same constant \( \Lambda \) as \( A \). Now let \( \tilde{w} \) be a weak solution of

\[
\begin{cases}
\tilde{w}_t = \text{div} \, \tilde{A}(z, \theta \hat{\lambda} \tilde{w}, D\tilde{w}) & \text{in} \ Q^*_4, \\
\tilde{w} = \tilde{u} & \text{on} \ \partial_\nu Q^*_4
\end{cases}
\]

and \( \tilde{v} \) be a weak solution of the frozen equation

\[
\begin{cases}
\tilde{v}_t = \text{div} \, \tilde{A}_B(t, \theta \hat{\lambda} \tilde{v}, D\tilde{v}) & \text{in} \ Q_3, \\
\tilde{v} = \tilde{w} & \text{on} \ \partial_\nu Q_3.
\end{cases}
\]

The existence of these weak solutions is guaranteed by Remark 4.4. Also, from Proposition 4.2 we have

\[
||\tilde{v}||_{L^\infty(Q_3)} \leq ||\tilde{w}||_{L^\infty(Q^*_4)} \leq ||\tilde{u}||_{L^\infty(Q_4)} \leq M_0/\theta \hat{\lambda}.
\]

In addition, we infer from Lemma 2.5 that

\[
||D\tilde{v}||_{L^p(Q_3)} \leq C(\Lambda, p, n) \left( 1 + ||D\tilde{w}||_{L^p(Q^*_4)} \right)
\]

(5.28)

\[
\leq C(\Lambda, p, n) \left[ 1 + ||D\tilde{u}||_{L^p(Q_4)} + ||\tilde{F}||_{L^p(Q_4)} + \left( \int_{Q_4} |\tilde{f}|^p \, dz \right)^{\frac{1}{p}} \right] \leq C(\Lambda, p, n).
\]

(5.29)
Therefore, by applying Lemma 5.2 and Lemma 5.3 for \( \Lambda \mapsto \tilde{\Lambda} \) and \( \alpha \mapsto \tilde{\alpha} \) we obtain
\[
\int_{Q_1} |D\tilde{u} - D\tilde{v}|^p \, dz \leq \varepsilon^p \quad \text{and} \quad \int_{Q_2} |D\tilde{w} - D\tilde{v}|^p \, dz \leq \varepsilon^p.
\]
Consequently, if we take \( \tilde{\Psi} := D\tilde{v} \) then it follows from the triangle inequality that
\[
\|D\tilde{u} - \tilde{\Psi}\|_{L^p(Q_2)} \leq 2\varepsilon.
\]
Thus it remains to show that there exists \( N > 0 \) depending only on \( p, n, \Lambda, M_0, \) and \( \mathbb{K} \) such that
\[
(5.31) \quad \|D\tilde{v}\|_{L^\infty(Q_2)} \leq N.
\]
To this end, we use the interior Hölder regularity theory (see [6, Theorem 1.1, pages 41] for the case \( p \geq 2 \) and [6, Theorem 1.1, pages 77] for the case \( p < 2 \)) to infer that there exist \( \bar{\alpha} \in (0, 1) \) and \( \gamma > 0 \) depending only on \( n, p, \) and \( \Lambda \) such that
\[
|\tilde{v}(x_1, t) - \tilde{v}(x_2, t)| \leq \tilde{\gamma} |\tilde{v}|_{L^\infty(Q_1)} |x_1 - x_2|^{\bar{\alpha}} \quad \text{for all} \quad (x_1, t), (x_2, t) \in Q_2.
\]
Notice that the presence of \( \tilde{\gamma} \tilde{\lambda} \) in (5.28) does not prevent us from applying the Hölder theory as the equation satisfy all required conditions in [6, Chapter 2, page 16] with structural constants independent of \( \theta \) and \( \tilde{\lambda} \).

Now let \( a(x, t, \xi) := \tilde{\Lambda}(t, \theta \tilde{\lambda} \tilde{v}(x, t), \xi) \). Then (5.28) implies that \( \tilde{v} \) satisfies
\[
\tilde{v}_t = \text{div} a(x, t, D\tilde{v}) \quad \text{in} \quad Q_3
\]
in the weak sense. Let \( s := \mu / \tilde{\lambda} \). Observe that \( \tilde{\Lambda}(\cdot, \cdot, \xi) = \frac{1}{\mu s - \lambda} \tilde{\Lambda}(\cdot, \cdot, \lambda \xi) \) and \( \partial_\xi \tilde{\Lambda}(\cdot, \cdot, \xi) = \frac{1}{\mu s - \lambda} (\partial_\xi \tilde{\Lambda})(\cdot, \cdot, \lambda \xi) \). Thanks to the third condition in (1.6) and Hölder estimate (5.32), the coefficient \( a \) satisfies
\[
|a(x_1, t, \xi) - a(x_2, t, \xi)| \leq \int_{B_4} |\tilde{\Lambda}(x, t, \theta \tilde{\lambda} \tilde{v}(x_1, t), \xi) - \tilde{\Lambda}(x, t, \theta \tilde{\lambda} \tilde{v}(x_2, t), \xi)| \, dx
\]
\[
\leq \tilde{\alpha} \theta \tilde{\lambda} |\tilde{v}(x_1, t) - \tilde{v}(x_2, t)| \frac{1}{\mu s - \lambda} (\mu^2 + |\lambda \xi|^2)^{\frac{\mu}{\mu - 1}}
\]
\[
\leq \tilde{\Lambda} \gamma M_0 |x_1 - x_2|^{\bar{\alpha}} (s^2 + |\xi|^2)^{\frac{\mu}{\mu - 1}}
\]
for any \( (x_1, t), (x_2, t) \in Q_2 \) and any \( \xi \in \mathbb{R}^n \). Moreover, we have
\[
\langle \partial_\xi a(x, t, \xi) \eta, \eta \rangle = \int_{B_4} \langle \partial_\xi \tilde{\Lambda}(y, t, \theta \tilde{\lambda} \tilde{v}(x, t), \xi) \eta, \eta \rangle \, dy
\]
\[
\geq \frac{\tilde{\alpha} \theta \tilde{\lambda}}{\mu s - \lambda} (\mu^2 + |\lambda \xi|^2)^{\frac{\mu}{\mu - 2}} |\eta|^2 = \Lambda^{-1} (s^2 + |\xi|^2)^{\frac{\mu}{\mu - 2}} |\eta|^2,
\]
\[
|\partial_\xi a(x, t, \xi)| \leq \int_{B_4} |\partial_\xi \tilde{\Lambda}(y, t, \theta \tilde{\lambda} \tilde{v}(x, t), \xi)| \, dy \leq \frac{\tilde{\alpha} \theta \tilde{\lambda}}{\mu s - \lambda} (\mu^2 + |\lambda \xi|^2)^{\frac{\mu}{\mu - 2}} = \Lambda (s^2 + |\xi|^2)^{\frac{\mu}{\mu - 2}},
\]
\[
|a(x, t, \xi)| \leq \int_{B_4} |\tilde{\Lambda}(y, t, \theta \tilde{\lambda} \tilde{v}(x, t), \xi)| \, dy \leq \frac{\tilde{\alpha} \theta \tilde{\lambda}}{\mu s - \lambda} (\mu^2 + |\lambda \xi|^2)^{\frac{\mu}{\mu - 2}} = \Lambda (s^2 + |\xi|^2)^{\frac{\mu}{\mu - 2}}
\]
Since \( s = \mu / \tilde{\lambda} \in [0, 1] \), we therefore can conclude from Theorem 5.4 and estimate (5.30) that
\[
\|D\tilde{v}\|_{L^\infty(Q_2)} \leq C(\Lambda, p, n, M_0),
\]
which gives desired estimate (5.31). Thus the proof of Theorem 1.1 is complete. \( \square \)
Remark 5.5. An alternative way of proving estimate (5.31) and working directly with $A$ instead of $\tilde{A}$ is to transform equation (5.28) to its original setting and then employ the intrinsic gradient bound in [14]. Precisely, let us rescale $\tilde{v}$ by defining

$$v(x, t) := \begin{cases} \theta \tilde{v}(\frac{x}{\theta}, \frac{t-\theta}{\theta^2}) & \text{if } p \geq 2, \\ \theta \tilde{v}(\frac{x}{\theta}, \frac{t-\theta}{\theta^2}) & \text{if } \frac{2n}{n+2} < p < 2. \end{cases}$$

Then $v$ is a weak solution of

$$v_t = \text{div} A_B(t, v, Dv) \quad \text{in} \quad Q^1_{2\theta}(\bar{z}),$$

where $B$ is the projection of $Q^1_{2\theta}(\bar{z})$ onto $\mathbb{R}^n$, i.e. $B = B_{4\theta}(\bar{x})$ if $p \geq 2$ and $B = B_{\frac{n-2}{2-n} \theta}(\bar{x})$ if $p < 2$. Moreover, we deduce from (5.29)–(5.30) that

$$||v||_{L^n(Q^1_{2\theta}(\bar{z}))} \leq M_0 \quad \text{and} \quad \left( \int_{Q^1_{2\theta}(\bar{z})} |Dv|^p dz \right)^{\frac{1}{p}} \leq C(\Lambda, p, n, \lambda).$$

Hence if $p \geq 2$, then it follows from the Hölder regularity theory (see [6] Theorem 1.1, pages 41) that there exists $\bar{\alpha} \in (0, 1)$ and $\gamma > 0$ depending only on $n, p,$ and $\Lambda$ such that

$$|v(x_1, t) - v(x_2, t)| \leq \gamma|v||_{L^n(Q^1_{2\theta}(\bar{z}))} \left( \frac{|x_1 - x_2|}{\theta} \right)^{\bar{\alpha}} \leq \gamma M_0 \left( \frac{|x_1 - x_2|}{\theta} \right)^{\bar{\alpha}} \quad \text{for all} \quad (x_1, t), (x_2, t) \in Q^1_{2\theta}(\bar{z}).$$

In the case $p < 2$, we can use [6] Theorem 1.1, pages 77 to obtain:

$$|v(x_1, t) - v(x_2, t)| \leq \gamma|v||_{L^n(Q^1_{2\theta}(\bar{z}))} \left( \frac{|x_1 - x_2|}{\theta} \right)^{\bar{\alpha}} \leq \gamma M_0 \left( \frac{|x_1 - x_2|}{\theta \Lambda^{\frac{n}{2-n}}} \right)^{\bar{\alpha}} \quad \text{for all} \quad (x_1, t), (x_2, t) \in Q^1_{2\theta}(\bar{z}).$$

Therefore, if we let $a(x, t, \xi) := A_B(t, v(x, t), \xi)$ then for any $(x_1, t), (x_2, t) \in Q^1_{2\theta}(\bar{z})$ and any $\xi \in \mathbb{R}^n$ we have

$$|a(x_1, t, \xi) - a(x_2, t, \xi)| \leq \int_B |A(x, t, v(x_1, t), \xi) - A(x, t, v(x_2, t), \xi)| dx$$

$$\leq \Lambda |v(x_1, t) - v(x_2, t)| (\mu^2 + |\xi|^2)^{\bar{\alpha} - 1} \leq \Lambda \tilde{\omega}(|x_1 - x_2|) (\mu^2 + |\xi|^2)^{\bar{\alpha} - 1},$$

where

$$\tilde{\omega}(r) := \begin{cases} \gamma M_0 \left( \frac{r}{\theta} \right)^{\bar{\alpha}} & \text{if } p \geq 2, \\ \gamma M_0 \left( \frac{r}{\theta \Lambda^{\frac{n}{2-n}}} \right)^{\bar{\alpha}} & \text{if } \frac{2n}{n+2} < p < 2. \end{cases}$$

Thus we conclude from [14] Theorem 4.1 and (4.34) that there exist $C_0 = C_0(n, p, \Lambda) > 1$ and $\sigma_0 = \sigma_0(n, p, \Lambda, M_0) > 0$ such that: if $Q^1_{\sigma_0}(z_0) \subset Q^1_{2\theta}(\bar{z})$ with $\sigma \leq \sigma_0$ and $z_0$ is a Lebesgue point of $Dv$, then

$$|Dv(z_0)| \leq C_0 \left[ \int_{Q^1_{\sigma_0}(z_0)} (|Dv| + 1)^p dz \right]^{\frac{1}{p}}.$$

By replacing $\sigma_0$ by $\min\{\sigma_0, \frac{1}{2}\}$ if necessary, we can assume that $\sigma_0 \leq \frac{1}{2}$. Then for any point $z_0 \in Q^1_{2\theta}(\bar{z})$ which is a Lebesgue point of $Dv$, we have $Q^1_{\sigma_0}(z_0) \subset Q^1_{2\theta}(\bar{z})$ and hence

$$|Dv(z_0)| \leq C_0 \left[ \int_{Q^1_{\sigma_0}(z_0)} (|Dv| + 1)^p dz \right]^{\frac{1}{p}}.$$
Using the second estimate in (5.33) to estimate the above right hand side, we deduce that

\[ \|Dv\|_{L^\infty(Q_t^\rho(\bar{x}))} \leq C(\Lambda, p, n, M_0) \lambda. \]

By rescaling back, we obtain \( \|D\tilde{v}\|_{L^\infty(Q_2)} \leq C(\Lambda, p, n, M_0) \) which gives (5.31).

6. Higher integrability of gradients

In this section we prove Theorem 2.6 about the higher integrability of weak solutions to equation (2.9). The proof of this will be given in Subsection 6.2 and is based on the arguments in [12, 21] (see also [10, Section 8.2] and [17, Lemma 12]). The key ingredient is a Caccioppoli type estimate.

6.1. Caccioppoli type estimates. Let \( \eta \in C_0^\infty(B_2(0)) \) be such that \( 0 \leq \eta \leq 1 \), \( \eta = 1 \) in \( B_1(0) \), and \( |D\eta| \leq 2 \) in \( B_2(0) \). For \( \bar{x} \in \mathbb{R}^n \) and \( \rho > 0 \), set \( \eta_{\bar{x},\rho}(x) = \eta(\frac{x - \bar{x}}{\rho}) \). We then define the weighted mean

\[ u_{\bar{x},\rho}^1 = u_{\bar{x},\rho}^1(t) = \frac{\int_{B_\rho(\bar{x})} \eta_{\bar{x},\rho}(x)p u(x, t) \, dx}{\int_{B_\rho(\bar{x})} \eta_{\bar{x},\rho} \, dx}. \]

The next lemma implies the absolute continuity of \( t \mapsto u_{\bar{x},\rho}^1(t) \).

**Lemma 6.1.** Let \( p > 1 \) and suppose that \( u \) is a weak solution of (2.9). Then for any \( t_1, t_2 \in (-16, 16) \) with \( t_1 < t_2 \), we have

\[ |u_{\bar{x},\rho}^1(t_1) - u_{\bar{x},\rho}^1(t_2)| \leq C(\Lambda, n, p) \left[ \frac{1}{\rho^{p+1}} \int_{t_1}^{t_2} \int_{B_\rho(\bar{x})} \left( 1 + |Du|^p + |F|^p \right) dx dt + \frac{1}{\rho^p} \int_{t_1}^{t_2} \int_{B_\rho(\bar{x})} |f| dx dt \right]. \]

**Proof.** For a function \( g(x, t) \) and \( h > 0 \), let \( [g]_h(x, t) \) denote its Steklov average defined as in (2.5). By using \( \phi = \eta_{\bar{x},\rho}(x)p \) in the Steklov formulation (2.6), we get for \( t_1, t_2 \in (-16, 16) \) that

\[ \int_{B_\rho(\bar{x})} [u_{\bar{x},\rho}^1]^p(\cdot, t_2) dx - \int_{B_\rho(\bar{x})} [u_{\bar{x},\rho}^1]^p(\cdot, t_1) dx = \int_{t_1}^{t_2} \int_{B_\rho(\bar{x})} \partial_t [u_{\bar{x},\rho}^1]^p h(\cdot, t) dx dt \]

\[ = -p \int_{t_1}^{t_2} \int_{B_\rho(\bar{x})} \langle [A(\cdot, \alpha, Du)]h(\cdot, t) + [\|F\|^{p-2} F]_h(\cdot, t), Du_{\bar{x},\rho}^1 \rangle \eta_{\bar{x},\rho}^p dx dt + \int_{t_1}^{t_2} \int_{B_\rho(\bar{x})} [f]_h(\cdot, t) \eta_{\bar{x},\rho}^p dx dt. \]

Using the growth condition (1.3) for \( A \) and the choice of the function \( \eta(x) \), we find after passing to the limit \( h \to 0^+ \) that

\[ \left| \int_{B_\rho(\bar{x})} u(\cdot, t_2) \eta_{\bar{x},\rho}^p dx - \int_{B_\rho(\bar{x})} u(\cdot, t_1) \eta_{\bar{x},\rho}^p dx \right| \leq \frac{C(\Lambda, p)}{\rho} \int_{t_1}^{t_2} \int_{B_\rho(\bar{x})} \left( 1 + |Du|^p + |F|^p \right) dx dt + \int_{t_1}^{t_2} \int_{B_\rho(\bar{x})} |f| dx dt \]

for every \( t_1, t_2 \in (-16, 16) \) with \( t_1 < t_2 \). This gives the lemma as desired. \( \square \)

Lemma 6.1 is only used to prove the following Caccioppoli type estimate.
Lemma 6.2 (Caccioppoli type estimate). Let $p > 1$ and suppose that $u$ is a weak solution of (2.9). Then there exists a constant $C > 0$ depending only on $\Lambda$, $n$, and $p$ such that

\[
\sup_{t_1 < t < t_2} \int_{B_\rho(x)} |u(x,t) - u_{x,2p}^p(t)|^2 dx + \int_{Q_{\rho(t_1)}} |Du|^p d\tau \leq C \left( \frac{1}{t_2 - t_1} \int_{Q_{\rho(t_2)}} |u - u_{x,2p}^p(t)|^2 d\tau + \rho^p \int_{Q_{\rho(t_2)}} |u - u_{x,2p}^p(t)|^p d\tau \right)
\]

for all $\rho > 0$ and $0 < t_1 < t_2$ satisfying $Q_{\rho}(2\rho, t_2) \subset Q_4$.

Proof. Let $\sigma \in C^\infty_0((\bar{t} - \tau_2, \bar{t} + \tau_2))$ be such that $0 \leq \sigma \leq 1$, $\sigma = 1$ in $[\bar{t} - \tau_1, \bar{t} + \tau_1]$, and $\|\partial_{\tau} \sigma\| \leq 2/(\tau_2 - \tau_1)$. By using $\sigma(t)\eta_{x,2p}^p(u - u_{x,2p}^p)$ as a test function in (2.9), we obtain

\[
\frac{1}{2} \int_{\bar{t} - \tau_2}^{\bar{t}} \sigma(t) \frac{d}{dt} \left( \int_{B_{2\rho}(\bar{x})} \eta_{x,2p}^p |u - u_{x,2p}^p|^2 dx \right) dt + \int_{\bar{t} - \tau_2}^{\bar{t}} \int_{B_{2\rho}(\bar{x})} \partial_{\tau} u_{x,2p}^p \sigma(t) u_{x,2p}^p (u - u_{x,2p}^p) dx dt
\]

\[
= - \int_{\bar{t} - \tau_2}^{\bar{t}} \int_{B_{2\rho}(\bar{x})} \langle \Lambda(z, au, Du) + |F|^{p-2} F, D(u_{x,2p}^p(u - u_{x,2p}^p)) \rangle \sigma(t) dx
\]

\[
+ \int_{\bar{t} - \tau_2}^{\bar{t}} \int_{B_{2\rho}(\bar{x})} f \eta_{x,2p}^p (u - u_{x,2p}^p) \sigma(t) dx dt.
\]

Now since $t \mapsto u_{x,2p}^p(t)$ is absolutely continuous by Lemma 6.1 and thus $\partial_{\tau} u_{x,2p}^p$ is integrable on $(\bar{t} - \tau_2, \bar{t} + \tau_2)$, we see that the second term on the left hand side is zero. Furthermore, we can apply integration by parts for the first term and decompose $A(z, au, Du)$ as $[\Lambda(z, au, Du) - A(z, au, 0)] + A(z, au, 0)$. Then, by using (1.2)–(1.3) and Lemma 2.1 with $\xi = 0$ and $\tau = 1/2$ we obtain for every $\tau \in (\bar{t} - \tau_2, \bar{t} + \tau_2)$ that

\[
\frac{1}{2} \int_{K_\tau} \eta_{x,2p}(x)^p |u(x, \tau) - u_{x,2p}^p(\tau)|^2 \sigma(t) dx + \int_{K_\tau} |Du|^p \eta_{x,2p}^p \sigma(t) dx
\]

\[
\leq c(A, p) \int_{K_\tau} \eta_{x,2p}^p \sigma(t) dx + \frac{1}{2} \int_{K_\tau} \sigma(t) \eta_{x,2p}^p |u - u_{x,2p}^p|^2 dx
\]

\[
+ p \int_{K_\tau} \langle \Lambda(1 + |Du|^{p-1}) + |F|^{p-1} D |F|^{p-1} \eta_{x,2p}^p |u - u_{x,2p}^p| |F|^{p-1} \eta_{x,2p}^p |u - u_{x,2p}^p| \sigma(t) |u - u_{x,2p}^p| \sigma(t) \rangle dx
\]

\[
+ \int_{K_\tau} \langle \Lambda + |F|^{p-1} |F|^{p-1} \eta_{x,2p}^p |u - u_{x,2p}^p| \sigma(t) \sigma(t) \rangle Du dx + \int_{K_\tau} |f| \eta_{x,2p}^p (u - u_{x,2p}^p) \sigma(t) dx| dz.
\]

where $K_\tau := B_{2\rho}(\bar{x}) \times (\bar{t} - \tau_2, \tau)$. Using Young’s inequality, it follows that

\[
\int_{B_{2\rho}(\bar{x})} \eta_{x,2p}(x)^p |u(x, \tau) - u_{x,2p}^p(\tau)|^2 \sigma(t) dx + \int_{K_\tau} |Du|^p \eta_{x,2p}^p \sigma(t) dx \leq \frac{C}{\tau_2 - \tau_1} \int_{K_\tau} |u - u_{x,2p}^p|^2 dx
\]

\[
+ \frac{C}{\rho^p} \int_{K_\tau} |u - u_{x,2p}^p|^p dx + C \int_{K_\tau} (1 + |F|^{p}) dx + C \int_{K_\tau} |f| \eta_{x,2p}^p (u - u_{x,2p}^p) \sigma(t) dx| dz.
\]

But as in (2.11), the last integral can be estimated as follows

\[
\int_{K_\tau} |f| \eta_{x,2p}^p (u - u_{x,2p}^p) \sigma(t) dx| dz \leq \varepsilon \int_{K_\tau} |D|\eta_{x,2p}^p (u - u_{x,2p}^p) |^p \sigma(t) dx| dz
\]

\[
+ \varepsilon \sup_{t \in (\bar{t} - \tau_2, \tau)} \int_{B_{2\rho}(\bar{x})} \eta_{x,2p}^p |u(x, t) - u_{x,2p}^p(t)|^2 \sigma(t) dx + \frac{C(n, p)}{\varepsilon^\frac{p}{p+1}} \int_{L^p(K_\tau)} \frac{|f|}{|f|_{L^p(K_\tau)}} \langle f, \eta_{x,2p}^p (u - u_{x,2p}^p) \rangle dx \forall \varepsilon > 0.
\]
Therefore, by choosing \( \varepsilon \) suitably and using the definition of \( \hat{p} \) in (1.3), we deduce that

\[
\int_{B_\rho(x)} \eta_{h,2\rho}^p(x) |u(x, \tau) - u_{h,2\rho}^p(\tau)|^2 \sigma(\tau) dx + \frac{1}{2} \int_{K_\tau} |Du|^p \eta_{h,2\rho}^p \sigma(t) dz \\
\leq \frac{1}{2} \sup_{\tau \in (t-\tau_2, t+\tau_2)} \int_{B_\rho(x)} \eta_{h,2\rho}^p |u(x, \tau) - u_{h,2\rho}^p(\tau)|^2 \sigma(\tau) dx + \frac{C}{\tau_2 - \tau_1} \int_{Q_{(2\rho, \tau_2)}} |u - u_{h,2\rho}^p|^2 dz \\
+ \frac{C}{\rho^p} \int_{Q_{(2\rho, \tau_2)}} |u - u_{h,2\rho}^p|^p dz + C \int_{Q_{(2\rho, \tau_2)}} (1 + |F|^p) dz + C \left( \int_{Q_{(2\rho, \tau_2)}} |f|^p dz \right)^\frac{p}{p}.
\]

for every \( \tau \in (t - \tau_2, t + \tau_2) \). From this we infer the conclusion of the lemma. \( \square \)

**Remark 6.3.** We note that the use of the test function in the proof of Lemma 6.2 is justified by making use of Remark 6.3. Notice that the only use of equation (2.9) is to obtain Lemma 6.2. For this reason and for completeness, we choose to present the proof for only the case \( p \geq 2 \). Let \( h(z) := 1 + |F(z)| \) and \( \hat{f}(z) := \frac{|f(z)|^p}{\hat{p}} \). For \( \lambda > 0 \), we denote

\[
E(\lambda) := \{ z \in Q_3 : z \text{ is a Lebesgue point of } |Du| \text{ and } |Du(z)| > \lambda \},
\]

\[
E_h(\lambda) := \{ z \in Q_3 : h > \lambda \}, \quad \text{and} \quad E_{\hat{f}}(\lambda) := \{ z \in Q_3 : \hat{f} > \lambda \}.
\]

Also, define

\[
\lambda^p_0 := \frac{1}{|Q_4|} \int_{Q_4} (|Du|^p + h^p) dz + \frac{1}{|Q_4|} \left( \int_{Q_4} |f|^p dz \right)^\frac{p}{p} \quad \text{and} \quad B^p := 40^{r+2}.
\]

Then for any \( \lambda \geq B\lambda_0 \), by applying a modification of Lemma 3.1, we obtain: there exists a sequence of *intrinsic cylinders* \( \{Q_4(z_i)\} \) with \( z_i \in Q_3 \) and \( r_i \in (0, \frac{1}{20}) \) that satisfies the following properties

\begin{itemize}
  \item[a)] \( \{Q_4(z_i)\} \) is disjoint and \( E(\lambda) \subset \bigcup_{i=1}^n Q_4(z_i) \).
  \item[b)] \( \int_{Q_4(z_i)} (|Du|^p + h^p) dz + \frac{1}{|Q_4(z_i)|} \left( \int_{Q_4(z_i)} |f|^p dz \right)^\frac{p}{p} = \lambda^p \) for each \( i \).
  \item[c)] \( \int_{Q_4(z_i)} (|Du|^p + h^p) dz + \frac{1}{|Q_4(z_i)|} \left( \int_{Q_4(z_i)} |f|^p dz \right)^\frac{p}{p} < \lambda^p \) for every \( r \in (r_1, 1) \).
\end{itemize}

Note that \( Q_{20r_i}(z_i) \subset Q_4 \) as \( z_i \in Q_3 \) and \( 20r_i \leq 1 \).

**Claim:** There exists \( C > 0 \) depending only on \( \Lambda, n, \) and \( p \) such that: for each \( i \), we have

\[
\int_{Q_{20r_i}(z_i)} |Du|^p dz \leq C \left( \int_{Q_{4r_i}(z_i)} |Du|^p dz \right)^\frac{n+2}{n+2} + \int_{Q_{4r_i}(z_i)} h^p dz + \frac{1}{|Q_4(z_i)|} \left( \int_{Q_4(z_i)} |f|^p dz \right)^\frac{p}{p}.
\]
Assume the claim for the moment. Set \( q = np/(n+2) \). Since it follows from property b) that \( \lambda^p \leq C_n \left[ \int_{Q_{2\delta_l}(z)} |Du|^p \, dz + \int_{Q_4(\zeta)} h^p \, dz + \frac{1}{|Q_{1\zeta}(\zeta)|} \left( \int_{Q_{1\zeta}(\zeta)} |f|^\rho \, dz \right)^\beta \right] \), we infer from (6.1) that

\[
\int_{Q_{2\delta_l}(z)} |Du|^p \, dz + \lambda^p \leq C \left[ \left( \int_{Q_{4\zeta}(z)} |Du|^q \, dz \right)^\frac{p}{q} \int_{Q_{4\zeta}(z) \cap E(\eta_l)} |Du|^q \, dz + \int_{Q_{4\zeta}(z) \cap E(\eta_l)} h^p \, dz + \frac{1}{|Q_{4\zeta}(z)|} \left( \int_{Q_{4\zeta}(z) \cap E(\eta_l)} \hat{f}^\rho \, dz \right)^\beta \right]
\]

\[
\leq 3C \eta^p \lambda^p + C \left[ \frac{1}{|Q_{4\zeta}(z)|} \int_{Q_{4\zeta}(z) \cap E(\eta_l)} |Du|^q \, dz \right]^\frac{\beta}{q} + \frac{1}{|Q_{4\zeta}(z)|} \int_{Q_{4\zeta}(z) \cap E(\eta_l)} h^p \, dz + \frac{1}{|Q_{4\zeta}(z)|} \left( \int_{Q_{4\zeta}(z) \cap E(\eta_l)} \hat{f}^\rho \, dz \right)^\beta
\]

for any \( \eta > 0 \). By choosing \( \eta > 0 \) small and applying H"older inequality to the first integral on the right hand side, we then deduce that

\[
\int_{Q_{2\delta_l}(z)} |Du|^p \, dz \leq C \left[ \left( \int_{Q_{4\zeta}(z) \cap E(\eta_l)} |Du|^q \, dz \right)^\frac{p}{q} \int_{Q_{4\zeta}(z) \cap E(\eta_l)} |Du|^q \, dz + \int_{Q_{4\zeta}(z) \cap E(\eta_l)} h^p \, dz + \left( \int_{Q_{4\zeta}(z) \cap E(\eta_l)} \hat{f}^\rho \, dz \right)^\beta \right]
\]

\[
\leq C \left[ \lambda^{p-q} \int_{E(\eta_l)} |Du|^q \, dz + \int_{E(\eta_l)} h^p \, dz + \left( \int_{E(\eta_l)} \hat{f}^\rho \, dz \right)^\beta \right],
\]

where we have used property c) to obtain the last inequality. This together with property a) gives

\[
\int_{E(\eta_l)} |Du|^p \, dz \leq C \sum_i \left[ \lambda^{p-q} \int_{E(\eta_l)} |Du|^q \, dz + \int_{E(\eta_l)} h^p \, dz + \left( \int_{E(\eta_l)} \hat{f}^\rho \, dz \right)^\beta \right]
\]

\[
\leq C \left[ \lambda^{p-q} \int_{E(\eta_l)} |Du|^q \, dz + \int_{E(\eta_l)} h^p \, dz + \left( \int_{E(\eta_l)} \hat{f}^\rho \, dz \right)^\beta \right] \quad \forall \lambda \geq B \lambda_0.
\]

Therefore if we let \( \lambda_1 := \eta^{-1} B \lambda_0 \geq B \lambda_0, \mu(dz) = |Du|^p \, dz, \tilde{\mu}(dz) = |Du|^q \, dz, \nu(dz) = h^p \, dz, \) and \( \sigma(dz) = \hat{f}^\rho \, dz, \) then

\[
(6.2) \quad \int_{E(\eta_1)} |Du|^p \, d\mu - \lambda^{p-q} \mu(E(\eta_0)) = e \int_{\lambda_1} \lambda^{p-q} \mu(E(\lambda)) \, d\lambda
\]

\[
\leq Ce \int_{\lambda_1} \lambda^{p-q+e-1} \tilde{\mu}(E(\eta l)) \, d\lambda + Ce \int_{\lambda_1} \lambda^{p-q-1} \nu(E(\eta l)) \, d\lambda + Ce \int_{\lambda_1} \lambda^{p-q-1} \sigma(E(\eta l)) \, d\lambda
\]

\[
= \frac{Ce}{\eta^{p-q+e}} \int_{\eta_1} t^{p-q+e-1} \tilde{\mu}(E(t)) \, dt + \frac{Ce}{\eta^{p-q+e}} \int_{\eta_1} t^{p-q} \nu(E(t)) \, dt + e \int_{\eta_1} t^{p-q} \sigma(E(t)) \, dt.
\]

Observe that

\[
(p - q + e) \int_{\eta_1} t^{p-q+e-1} \tilde{\mu}(E(t)) \, dt \leq \int_{\tilde{E}(\eta_1)} |Du|^{p-q+e} \, d\tilde{\mu}
\]

\[
\leq \int_{E(\eta_1)} |Du|^{p-q+e} \, d\tilde{\mu} + \lambda^{p-q} \int_{\tilde{E}(\eta_1) \setminus E(\eta_1)} |Du|^{p-q} \, d\tilde{\mu} \leq \int_{E(\eta_1)} |Du|^{p-q+e} \, d\tilde{\mu} + \lambda^{p-q} \int_{E(\eta_1)} |Du|^{p-q} \, d\tilde{\mu}
\]

\[
(6.3) \quad \int_{E(\eta_1)} |Du|^p \, d\tilde{\mu} \leq Ce \int_{\lambda_1} \lambda^{p-q+e} \tilde{\mu}(E(\lambda)) \, d\lambda + Ce \int_{\lambda_1} \lambda^{p-q} \nu(E(\lambda)) \, d\lambda + Ce \int_{\lambda_1} \lambda^{p-q} \sigma(E(\lambda)) \, d\lambda.
\]
and as \( \lambda_1 \geq 2 \) we have
\[
E \int_{\eta \lambda_1}^{\infty} t^{p-1} \sigma(E_j(t))^\beta \, dt \leq \sum_{k=0}^{\infty} E \int_{2^k}^{2^{k+1}} t^{p-1} \sigma(E_j(t))^\beta \, dt \leq (2^e - 1) \sum_{k=0}^{\infty} 2^k \sigma(E_j(2^k))^\beta
\]
\[
\leq (2^e - 1) \left[ \sum_{k=1}^{\infty} 2^k \sigma(E_j(2^k))^\beta \right] \leq \frac{(2^e - 1)2^e}{(2^e - 1)^6} \left( \int_{Q_1} f^\beta \, dx \right)^\beta \leq \frac{4^e}{(2^e - 1)^6} \left( \int_{Q_1} |f|^{\beta(1+\frac{\beta}{p})} \, dz \right)^\beta,
\]
where we have used Remark 3.3 to obtain the inequality right before the last one. Therefore, it follows from (6.2) that
\[
\int_{E(A_1)} |Du|^{p+e} \, dz \leq \frac{C_E}{(p - q + \varepsilon)\eta^{p-q+\varepsilon}} \left[ \int_{E(A_1)} |Du|^{p+e} \, dz + \lambda_1^e \int_{E(\eta A_1)} |Du|^{p} \, dz \right]
\]
\[
+ C \eta^{-e} \left[ \int_{Q_1} h^{p+e} \, dz + \frac{4^e}{(2^e - 1)^6} \left( \int_{Q_3} |f|^{\beta(1+\frac{\beta}{p})} \, dz \right)^\beta \right] + \lambda_1^e \int_{E(A_1)} |Du|^{p} \, dz.
\]
Choosing \( \varepsilon > 0 \) small we can absorb the integral involving \( |Du|^{p+e} \) into the left hand side and get:
\[
\int_{E(A_1)} |Du|^{p+e} \, dz \leq 2\lambda_1^e \int_{E(\eta A_1)} |Du|^{p} \, dz + C \left[ \int_{Q_3} h^{p+e} \, dz + \left( \int_{Q_3} |f|^{\beta(1+\frac{\beta}{p})} \, dz \right)^\beta \right].
\]
Notice that there is a difficulty in moving the term to the left side since it may be infinite. However, this can be handled by using truncation as in the proof of [12 Proposition 4.1] and [2, page 591–594]. The main observation there is that (6.2) still holds true if \( E(A_1) \) is replaced by \( E^*(A_1) \) and \( E(\eta A_1) \) is replaced by \( E^k(\eta A_1) \), where \( E^k(\lambda) := \{ z \in Q_3 : |Du_k(z) > \lambda \} \) with \( |Du_k| := \min \{|Du|, k\} \). As a consequence of (6.4), we obtain
\[
\int_{Q_3} |Du|^{p+e} \, dz \leq \int_{E(A_1)} |Du|^{p+e} \, dz + \lambda_1^e \int_{E(\eta A_1)} |Du|^{p} \, dz
\]
\[
\leq 3\lambda_1^e \int_{Q_3} |Du|^{p} \, dz + C \left[ \int_{Q_3} h^{p+e} \, dz + \left( \int_{Q_3} |f|^{\beta(1+\frac{\beta}{p})} \, dz \right)^\beta \right].
\]
This together with the definitions of \( \lambda_1 \) and \( h(z) \) gives the desired estimate (6.15).

It remains to prove the claim. Thanks to properties b) and c), estimate (6.1) will follow if we can show that
\[
\int_{Q_i^e(z_i)} |Du|^{p} \, dz \leq C \left[ \left( \int_{Q_i^e(z_i)} |Du|^{p+e} \, dz \right)^{\frac{\beta+2}{\beta}} + \int_{Q_i^e(z_i)} h^{p} \, dz + \frac{1}{|Q_i^e(z_i)|} \left( \int_{Q_i^e(z_i)} |f|^{\beta(1+\frac{\beta}{p})} \, dz \right)^\beta \right].
\]
Let us set \( \bar{z} = z_i \) and \( r := r_i \). Then by applying Lemma 6.2 for \( \rho = r, \tau_1 = \lambda_2^{-p}r^2 \) and \( \tau_2 = 2\lambda_2^{-p}r^2 \), we obtain
\[
\int_{Q_i(r_1,2r^2)} |Du|^{p} \, dz \leq \frac{C}{\lambda_2^{-p}r^2} \int_{Q_i(2r,2r_1^2,r^2)} |u - u_{\lambda_2^p}(t)|^2 \, dz + \frac{C}{r^p} \int_{Q_i(2r,2r^2)} |u - u^p_{\lambda_2^p}(t)|^p \, dz
\]
\[
+ C \left[ \int_{Q_i(2r,2r_1^2,r^2)} h^{p} \, dz + \left( \int_{Q_i(2r,2r^2)} |f|^{\beta(1+\frac{\beta}{p})} \, dz \right)^\beta \right].
\]
Moreover, Young’s inequality and property b) give for any $\epsilon > 0$ that
\[
\frac{C}{\lambda^{2-p}r^p} \int_{Q_{(r,2\lambda^2-r^2)}^3} |u - u^{\eta}_{x,2r}(t)|^2 \, dz \leq \epsilon \lambda^{p-1} |Q_{(r,\lambda^2-r^2)}^3| + \frac{2\epsilon}{r^p} \int_{Q_{(r,2\lambda^2-r^2)}^3} |u - u^{\eta}_{x,2r}(t)|^p \, dz
\]
\[
= \epsilon \int_{Q_{(r,2\lambda^2-r^2)}^3} (|Du|^p + h^p) \, dz + \left( \int_{Q_{(r,2\lambda^2-r^2)}^3} |f|^\beta \, dz \right) + \frac{2\epsilon}{r^p} \int_{Q_{(r,2\lambda^2-r^2)}^3} |u - u^{\eta}_{x,2r}(t)|^p \, dz.
\]
Therefore, we deduce that
\[
\int_{Q_{(r,2\lambda^2-r^2)}^3} |Du|^p \, dz \leq \frac{C}{r^p} \int_{Q_{(r,2\lambda^2-r^2)}^3} |u - u^{\eta}_{x,2r}(t)|^p \, dz
\]
\[
+ C \left[ \int_{Q_{(r,2\lambda^2-r^2)}^3} h^p \, dz + \left( \int_{Q_{(r,2\lambda^2-r^2)}^3} |f|^\beta \, dz \right) \right]
\]
\[
\leq \frac{C}{r^p} \left( \int_{Q_{(r,4\lambda^2-r^2)}^3} |Du|^p \, dz \right) \left( \sup_{|r-\bar{r}|<\lambda^2-r^2} \int_{B_{\eta}(\bar{x})} |u - u^{\eta}_{x,\eta^2}(t)|^2 \, dx \right) \frac{1}{\lambda^{p-1}} \frac{1}{r^p} \frac{1}{\lambda^{p-1}} \frac{1}{r^p}
\]
\[
+ C \left[ \int_{Q_{(r,2\lambda^2-r^2)}^3} h^p \, dz + \left( \int_{Q_{(r,2\lambda^2-r^2)}^3} |f|^\beta \, dz \right) \right],
\]
where the last inequality follows from \cite{12} Lemma 3.3. We next estimate the supremum on the right hand side. For this, we apply Lemma \ref{6.2} for $\rho = 4r$, $\tau_1 = 4\lambda^2-r^2$ and $\tau_2 = 8\lambda^2-r^2$ to get
\[
\sup_{|r-\bar{r}|<\lambda^2-r^2} \int_{B_{\eta}(\bar{x})} |u(x,t) - u^{\eta}_{x,\eta^2}(t)|^2 \, dx \leq \frac{C}{\lambda^{2-p}r^p} \int_{Q_{(8\lambda^2,8\lambda^2-r^2)}^3} |u - u^{\eta}_{x,\eta^2}(t)|^2 \, dz
\]
\[
+ \frac{C}{r^p} \int_{Q_{(8\lambda^2,8\lambda^2-r^2)}^3} |u - u^{\eta}_{x,\eta^2}(t)|^p \, dz + C \left[ \int_{Q_{(8\lambda^2,8\lambda^2-r^2)}^3} h^p \, dz + \left( \int_{Q_{(8\lambda^2,8\lambda^2-r^2)}^3} |f|^\beta \, dz \right) \right].
\]
Using Poincaré inequality for functions in Sobolev spaces $W^{1,q}(B_{\eta}(\bar{x})) (q = 2, p)$ together with the fact
\[
\left( \int_{B_{\eta}(\bar{x})} |u - u^{\eta}_{x,\eta^2}(t)|^2 \, dx \right)^{\frac{1}{2}} \leq \left( \int_{B_{\eta}(\bar{x})} |u - u^{\eta}_{x,\eta^2}(t)|^2 \, dx \right)^{\frac{1}{2}} + |B_{\eta}(\bar{x})|^{\frac{1}{2}} |u^{\eta}_{x,2r}(t) - u^{\eta}_{x,\eta^2}(t)|
\]
\[
= \left( \int_{B_{\eta}(\bar{x})} |u - u^{\eta}_{x,\eta^2}(t)|^2 \, dx \right)^{\frac{1}{2}} + \frac{|B_{\eta}(\bar{x})|^{\frac{1}{2}}}{B_{\eta}(\bar{x})} \int_{B_{\eta}(\bar{x})} |u(x,t) - u^{\eta}_{x,\eta^2}(t)| \eta^{\beta}_{x,2r} \, dx
\]
\[
\leq C \left( \int_{B_{\eta}(\bar{x})} |u - u^{\eta}_{x,\eta^2}(t)|^2 \, dx \right)^{\frac{1}{2}},
\]
we infer that
\[
\sup_{|r-\bar{r}|<\lambda^2-r^2} \int_{B_{\eta}(\bar{x})} |u(x,t) - u^{\eta}_{x,\eta^2}(t)|^2 \, dx
\]
\[
\leq C\lambda^{p-2} \int_{Q_{(8\lambda^2,8\lambda^2-r^2)}^3} |Du|^2 \, dz + C \left[ \int_{Q_{(8\lambda^2,8\lambda^2-r^2)}^3} (|Du|^p + h^p) \, dz + \left( \int_{Q_{(8\lambda^2,8\lambda^2-r^2)}^3} |f|^\beta \, dz \right) \right]
\]
\[
\leq C\lambda^{p-2} |Q_{\eta}(\bar{x})|^{\frac{1}{p}} \left( \int_{Q_{\eta}(\bar{x})} |Du|^p \, dz \right)^{\frac{1}{p}} + C \left( \int_{Q_{\eta}(\bar{x})} |Du|^p \, dz + \left( \int_{Q_{\eta}(\bar{x})} |f|^\beta \, dz \right) \right]
\]
\[
\leq C \lambda^p |Q_{\eta}(\bar{x})| = C \lambda^2 r^{n+2}.
\]
This together with (6.6) yields
\[
\int_{Q^1_T} |D u|^p \, dz \leq C \lambda^{\frac{2p}{n}} \int_{Q^1_\lambda} |D u|^p \, dz + C \left[ \int_{Q^1_\lambda} h^p \, dz + \frac{1}{|Q^1_\lambda(\xi)|} \left( \int_{Q^1_\lambda} |f|^p \, dz \right)^{\frac{1}{p}} \right]
\]
\[
\leq \frac{1}{2} \lambda^p + C \left( \int_{Q^1_\lambda} |D u|^p \, dz \right)^{\frac{1}{p}} + C \left[ \int_{Q^1_\lambda} h^p \, dz + \frac{1}{|Q^1_\lambda(\xi)|} \left( \int_{Q^1_\lambda} |f|^p \, dz \right)^{\frac{1}{p}} \right].
\]
It follows by using property b) that
\[
\int_{Q^1_T} |D u|^p \, dz \leq C \left( \left( \int_{Q^1_\lambda} |D u|^p \, dz \right)^{\frac{1}{p}} + \int_{Q^1_\lambda} h^p \, dz + \frac{1}{|Q^1_\lambda(\xi)|} \left( \int_{Q^1_\lambda} |f|^p \, dz \right)^{\frac{1}{p}} \right).
\]
Hence (6.5) is proved and the proof of Theorem 2.6 is complete. \(\square\)

APPENDIX A. A COMPARISON PRINCIPLE

Let \( \Omega_T := \Omega \times (0, T) \) with \( T > 0 \) and \( \Omega \) being a bounded domain in \( \mathbb{R}^n \), \( n \geq 2 \). Let \( \mathbb{K} \subset \mathbb{R} \) be an open interval, \( A : \Omega_T \times \mathbb{K} \times \mathbb{R}^n \rightarrow \mathbb{R}^n \), \( \Phi : \Omega_T \times \mathbb{K} \rightarrow \mathbb{R} \) and \( g, b : \Omega_T \times \mathbb{K} \rightarrow \mathbb{R} \) be Carathéodory maps. We assume that there exist constants \( \Lambda > 0 \) and \( 1 < p < \infty \) such that the following conditions are satisfied for a.e. \( z \in \Omega_T \) and all \( \xi, \eta \in \mathbb{R}^n \):

(A.1) \[ \langle A(z, u, \xi) - A(z, u, \eta), \xi - \eta \rangle \geq 0 \quad \forall u \in \mathbb{K}, \]
(A.2) \[ u \in \mathbb{K} \mapsto g(z, u) \text{ is monotone nondecreasing.} \]

Also, there exist functions \( K \in L^p(\Omega_T) \) and \( k \in L^1(0, T) \) such that

(A.3) \[ |A(z, u_1, \xi) - A(z, u_2, \xi)| \leq |u_1 - u_2| \langle \Lambda|\xi|^{p-1} + K(z) \rangle, \]
(A.4) \[ |\Phi(z, u_1) - \Phi(z, u_2)| \leq |u_1 - u_2| K(z), \quad \text{and} \quad |b(z, u_1) - b(z, u_2)| \leq |u_1 - u_2| |k(t)| \]

for all \( u_1, u_2 \in \mathbb{K} \) with \( |u_1 - u_2| \) sufficiently small.

**Definition A.1.** A map \( u \in C((0, T); L^2(\Omega)) \cap L^p(0, T; W^{1-\rho, \rho}(\Omega)) \) is called a weak solution to

(A.5) \[ u_t = \text{div} A(z, u, Du) + \text{div} \Phi(z, u) - g(z, u) + b(z, u) \quad \text{in} \quad \Omega_T. \]

if \( u(z) \in \mathbb{K} \) for a.e. \( z \in \Omega_T \) and

\[
\int_{\Omega_T} u \varphi_t \, dz = \int_{\Omega_T} \langle A(z, u, Du), D\varphi \rangle \, dz + \int_{\Omega_T} \langle \Phi(z, u), D\varphi \rangle \, dz + \int_{\Omega_T} \langle g(z, u) - b(z, u), \varphi \rangle \, dz \]

for every \( \varphi \in C^0_0(\Omega_T) \).

The following result shows that equation (A.5) admits a comparison principle.

**Proposition A.2** (comparison principle). Assume that \( A, \Phi, g \) and \( b \) satisfy conditions (A.1)–(A.4). Let \( u \) and \( v \) be weak solutions to (A.5) such that \( u \leq v \) on \( \partial_\rho \Omega_T \). Then

\[ u \leq v \quad \text{a.e. in} \quad \Omega_T. \]

**Remark A.3.** By inspecting the arguments below, one sees that we in fact only need to assume that \( u \) is a weak subsolution and \( v \) is a weak supersolution.
Proof. For $\varepsilon > 0$ small, we define $h_\varepsilon(s)$ as in (5.16). Let us denote $\Omega_t = \Omega \times (0, t)$ and by using $h_\varepsilon(u - v)$ as a test function in the equations for $u$ and $v$ and subtracting the resulting expressions, we obtain:

$$
\int_\Omega \left( \int_0^\infty h_\varepsilon(s) \, ds \right) \, dx + \int_\Omega h_\varepsilon'(u) (A(z, u, Du) - A(z, u, Dv), Du - Dv) \, dz \\
+ \int_\Omega [g(z, u) - g(z, v)]h_\varepsilon(u - v) \, dz \\
= \int_\Omega h_\varepsilon'(u - v)(A(z, v, Dv) - A(z, u, Dv), Du - Dv) \, dz \\
- \int_\Omega h_\varepsilon'(u - v)(\Phi(z, u) - \Phi(z, v), Du - Dv) \, dz + \int_\Omega [b(z, u) - b(z, v)]h_\varepsilon(u - v) \, dz
$$

for all $t \in (0, T)$. Since the second and third terms on the left-hand side are nonnegative thanks to (A.1)–(A.2), we deduce that

$$
\int_\Omega \left( \int_0^\infty h_\varepsilon(s) \, ds \right) \, dx \\
\leq \frac{1}{\varepsilon} \int_0^\infty \int_{\Omega \cap \{0 < u < \varepsilon\}} (u - v)(A|Dv|^{p-1} + 2K)|Du - Dv| \, dz + \int_\Omega k(\tau)|u - v|h_\varepsilon(u - v) \, dx d\tau \\
\leq \int_0^\infty \int_{\Omega \cap \{0 < u < \varepsilon\}} (A|Dv|^{p-1} + 2K)|Du - Dv| \, dz + \int_\Omega k(\tau)|u - v|h_\varepsilon(u - v) \, dx d\tau.
$$

As $\varepsilon \to 0^+$, we have

$$
\int_\Omega \left( \int_0^\infty h_\varepsilon(s) \, ds \right) \, dx \longrightarrow \int_\Omega (u - v)_+(x, t) \, dx.
$$

Moreover, the first term on the right-hand side tends to zero and the last term tends to

$$
\int_\Omega k(\tau)|u - v| \text{sgn}^+(u - v) \, dx d\tau = \int_0^\infty \int_\Omega k(\tau)(u - v)_+ \, dx d\tau.
$$

Thus if we denote $m(\tau) := \int_\Omega (u - v)_+(x, \tau) \, dx$, then by letting $\varepsilon \to 0^+$ in (A.6) we obtain

$$
m(\tau) \leq \int_0^\tau k(\tau)m(\tau) \, d\tau \quad \forall t \in (0, T).
$$

Therefore, it follows from the Grönwall’s inequality that $m(\tau) \leq 0$ for every $\tau \in (0, T)$. We then conclude that $\int_\Omega (u - v)_+(x, t) \, dx dt = 0$, and hence $u \leq v$ for a.e. in $\Omega_T$. The proof is complete. ∎

Remark A.4. We note that the comparison principle in Proposition A.2 together with the standard method for proving existence using Galerkin approximation (see [16] pages 466–475 and [19] pages 28) ensures that: for any $u \in L^\infty(Q_3) \cap L^1(-9; W^{1,p}(\Omega))$ satisfying $u(z) \in \mathbb{R}$ for a.e. $z \in Q_3$, the Dirichlet problem

$$
\left\{ \begin{array}{ll}
\nu \tau &= \text{div} \mathbf{A}(z, v, Dv) & \text{in } Q_3, \\
v &= u & \text{on } \partial_\tau Q_3
\end{array} \right.
$$

has a weak solution in the sense of Definition A.7.
References

[1] E. Acerbi and G. Mingione, Gradient estimates for a class of parabolic systems. Duke Math. J. 136 (2007), no. 2, 285–320.

[2] V. Bögelein, Global Calderón-Zygmund theory for nonlinear parabolic systems. Calc. Var. Partial Differential Equations 51 (2014), no. 3-4, 555–596.

[3] S.-S. Byun, J. Ok, and S. Ryu, Global gradient estimates for general nonlinear parabolic equations in nonsmooth domains. J. Differential Equations 254 (2013), no. 11, 4290–4326.

[4] S.-S. Byun and L. Wang, Parabolic equations in Reifenberg domains. Arch. Ration. Mech. Anal. 176 (2005), no. 2, 271–301.

[5] L. Caffarelli and I. Peral, $W^{1,p}$ estimates for elliptic equations in divergence form. Comm. Pure Appl. Math. 51 (1998), no. 1, 1–21.

[6] E. DiBenedetto, Degenerate parabolic equations. Universitext. Springer-Verlag, New York, 1993.

[7] E. DiBenedetto and A. Friedman, Regularity of solutions of nonlinear degenerate parabolic systems. J. Reine Angew. Math. 349 (1984), 83–128.

[8] E. DiBenedetto and A. Friedman, Hölder estimates for nonlinear degenerate parabolic systems. J. Reine Angew. Math. 357 (1985), 1–22.

[9] E. DiBenedetto and J. Manfredi, On the higher integrability of the gradient of weak solutions of certain degenerate elliptic systems. Amer. J. Math. 115 (1993), no. 5, 1107–1134.

[10] T. Iwaniec, Projections onto gradient fields and $L^p$-estimates for degenerated elliptic operators. Studia Math. 75 (1983), no. 3, 293–312.

[11] J. Kinnunen and S. Zhou, A local estimate for nonlinear equations with discontinuous coefficients. Comm. Partial Differential Equations 24 (1999), no. 11-12, 2043–2068.

[12] T. Kuusi and G. Mingione, New perturbation methods for nonlinear parabolic problems. J. Math. Pures Appl. (9) 98 (2012), no. 4, 390–427.

[13] T. Kuusi and G. Mingione, Gradient regularity for nonlinear parabolic equations. Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 12 (2013), no. 4, 755–822.

[14] O. Ladyzenskaja, V. Solonnikov, and N. Ural’tceva, Linear and quasilinear equations of parabolic type. Translations of Mathematical Monographs, Vol. 23, American Mathematical Society, 1968.

[15] C. Leone, M. Misawa, and A. Verde, The regularity for nonlinear parabolic systems of p-Laplacian type with critical growth. J. Differential Equations 256 (2014), no. 8, 2807–2845.

[16] N. Meyer and A. Elcrat, Some results on regularity for solutions of non-linear elliptic systems and quasi-regular functions. Duke Math. J. 42 (1975), 121–136.

[17] M. Misawa, Partial regularity results for evolutional $p$-Laplacian systems with natural growth. Manuscripta Math. 109 (2002), no. 4, 419–454.

[18] M. Misawa, $L^q$ estimates of gradients for evolutional $p$-Laplacian systems. J. Differential Equations 219 (2005), no. 2, 390–420.

[19] J.-L. Lions. Quelques méthodes de résolution des problèmes aux limites non linéaires. Dunod, 1969.

[20] N. Meyers and A. Elcrat, Some results on regularity for solutions of non-linear elliptic systems and quasi-regular functions. Duke Math. J. 42 (1975), 121–136.

[21] M. Misawa, Partial regularity results for evolutional $p$-Laplacian systems with natural growth. Manuscripta Math. 109 (2002), no. 4, 419–454.

[22] M. Misawa, $L^q$ estimates of gradients for evolutional $p$-Laplacian systems. J. Differential Equations 219 (2005), no. 2, 390–420.

[23] T. Nguyen and T. Phan, Interior gradient estimates for quasilinear elliptic equations. Calc. Var. (2016) 55: 59. doi:10.1007/s00526-016-0996-5.

[24] P. Tolksdorf, Regularity for a more general class of quasilinear elliptic equations. J. Differential Equations 51 (1984), no. 1, 126–150.

[25] W. Zou and J. Li, Existence and uniqueness of bounded weak solutions for some nonlinear parabolic problems. Bound. Value Probl. 2015, 2015:69.