**REGULARITY ESTIMATES FOR SINGULAR PARABOLIC MEASURE DATA PROBLEMS WITH SHARP GROWTH**

JUNG-TAE PARK AND PILSOO SHIN

**Abstract.** We prove global gradient estimates for parabolic $p$-Laplace type equations with measure data, whose model is

$$u_t - \text{div}(|Du|^{p-2}Du) = \mu \quad \text{in } \Omega \times (0, T) \subset \mathbb{R}^n \times \mathbb{R},$$

where $\mu$ is a signed Radon measure with finite total mass. We consider the singular case

$$\frac{2n}{n+1} < p \leq 2 - \frac{1}{n+1}$$

and give possibly minimal conditions on the nonlinearity and the boundary of $\Omega$, which guarantee the regularity results for such measure data problems.

1. Introduction and results

Partial differential equations with measure data allow to take into account a variety of models in the area of applied mathematics: for instance, the flow pattern of blood in the heart [57, 58], surface tension forces concentrated on the interfaces of fluids [47, 56, 70], and state-constrained optimal control theory [19–21, 49].

In this paper, we establish global gradient estimates for solutions of quasilinear parabolic equations with measure data, having the form

$$\begin{cases}
  u_t - \text{div} a(Du, x, t) = \mu & \text{in } \Omega_T, \\
  u = 0 & \text{on } \partial_p \Omega_T.
\end{cases}$$

(1.1)

Here $\Omega_T := \Omega \times (0, T)$ is a cylindrical domain with parabolic boundary $\partial_p \Omega_T := (\partial\Omega \times [0, T]) \cup (\Omega \times \{0\})$, where $\Omega \subset \mathbb{R}^n$ is a bounded domain with nonsmooth boundary $\partial\Omega$, $n \geq 2$ and $T > 0$. The nonhomogeneous term $\mu$ is a signed Radon measure on $\Omega_T$ with finite total mass. From now on we assume that the measure $\mu$ is defined on $\mathbb{R}^{n+1}$ by letting zero outside $\Omega_T$; that is,

$$|\mu|(\Omega_T) = |\mu|(\mathbb{R}^{n+1}) < \infty.$$ 

A typical model of the problem (1.1) is given by the parabolic $p$-Laplace equation; that is,

$$u_t - \text{div} (|Du|^{p-2}Du) = \mu.$$
Throughout the paper, the nonlinearity $a = a(\xi, x, t) : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ is measurable in $(x, t)$ and $C^1$-regular in $\xi$, satisfying
\begin{equation}
\begin{aligned}
|a(\xi, x, t)| + |\xi||D_\xi a(\xi, x, t)| &\leq \Lambda_1 |\xi|^{p-1}, \\
\Lambda_0 |\xi|^{p-2} |\eta|^2 &\leq \langle D_\xi a(\xi, x, t), \eta, \eta \rangle
\end{aligned}
\end{equation}
for almost every $(x, t) \in \mathbb{R}^n \times \mathbb{R}$, for every $\eta \in \mathbb{R}^n$, $\xi \in \mathbb{R}^n \setminus \{0\}$ and for some constants $\Lambda_1 \geq \Lambda_0 > 0$. Note that (1.2) implies $a(0, x, t) = 0$ for $(x, t) \in \mathbb{R}^n \times \mathbb{R}$ and the following monotonicity condition:
\begin{equation}
\langle a(\xi_1, x, t) - a(\xi_2, x, t), \xi_1 - \xi_2 \rangle \geq \hat{\Lambda}_0 \left( |\xi_1|^2 + |\xi_2|^2 \right)^{\frac{p-2}{p}} |\xi_1 - \xi_2|^2
\end{equation}

for all $(x, t) \in \mathbb{R}^n \times \mathbb{R}$ and $\xi_1, \xi_2 \in \mathbb{R}^n$, and for some constant $\hat{\Lambda}_0 = \hat{\Lambda}_0(n, \Lambda_0, p) > 0$. In this paper we shall focus on the singular case
\begin{equation}
2 \leq p < 2 - \frac{1}{n+1},
\end{equation}
since the case $p > 2 - \frac{1}{n+1}$ has been treated in [16]. The lower bound (1.4) is sharp in the sense that it reflects the fundamental solution of the parabolic $p$-Laplace equation (see Section 1.1 below). For more detailed information concerning the lower bound (1.4), we refer to Remark 1.7 below.

### 1.1. Renormalized solutions.

Let us first consider the parabolic $p$-Laplace equation with Dirac measure
\begin{equation}
\Gamma(x, t) = \begin{cases} 
0 & \text{if } t > 0, \\
\hat{\Lambda}_0 \left( \frac{p-2}{p} \right)^{\frac{\theta}{p-2}} \left( |x|^{-\frac{p}{\theta}} \right) & \text{otherwise},
\end{cases}
\end{equation}
where $\theta := \frac{1}{p(n+1)-2n}$. The solution $\Gamma$ is well defined provided
\begin{equation}
\theta > 0 \iff p > \frac{2n}{n+1}.
\end{equation}

We can check by a direct calculation that
\begin{equation}
D\Gamma \in L^q(\mathbb{R}^{n+1}) \quad \text{for all } q < p - \frac{n}{n+1},
\end{equation}
which implies that the solution $\Gamma$ does not belong to the usual energy space. Furthermore, we emphasize $D\Gamma \notin L^1(\mathbb{R}^{n+1})$ if $p \leq 2 - \frac{1}{n+1}$. Under (1.4), we thus need a proper notion of generalized solution as well as its gradient. For this, let us introduce a nonlinear parabolic capacity. For every $p > 1$ and every open subset $Q \subset \Omega_T$, the $p$-parabolic capacity of $Q$ is defined by
\begin{equation}
\text{cap}_p(Q) := \inf \{ \| u \|_W : u \in W, u \geq \chi_Q \text{ a.e. in } \Omega_T \},
\end{equation}
where $\chi_Q$ is the usual characteristic function of $Q$ and
\begin{equation}
W := \left\{ u \in L^p(0, T; V) : u_t \in L^{p'}(0, T; V') \right\}
\end{equation}
endowed with the norm
\[ \|u\|_W := \|u\|_{L^p(0,T;V)} + \|u_t\|_{L^{p'}(0,T;V')} . \]

Here \( p' \) is the Hölder conjugate of \( p \) with \( \frac{1}{p} + \frac{1}{p'} = 1 \), \( V := W^{1,p}_0(\Omega) \cap L^2(\Omega) \) and \( V' \) is the dual space of \( V \). For \( p > \frac{2n}{n+2} \), the embedding \( W^{1,p}_0(\Omega) \subset L^2(\Omega) \) is valid, so that \( V := W^{1,p}_0(\Omega) \).

We say a function \( u \) is \( cap_{p,q} \)-quasi continuous if for each \( \epsilon > 0 \), there exists an open set \( \bar{Q} \subset \Omega_T \) such that \( cap_p(\bar{Q}) < \epsilon \) and \( u \) is continuous on \( \Omega_T \setminus \bar{Q} \). Note that every function in \( W \) has a \( \text{cap}_{p,q} \)-quasi continuous representative. We refer to [2, 30, 36, 52, 65] for further information concerning parabolic capacities.

Let \( \mathcal{M}_b(\Omega_T) \) be the space of all signed Radon measures on \( \Omega_T \) with finite total mass. We denote by \( \mathcal{M}_a(\Omega_T) \) the subspace of \( \mathcal{M}_b(\Omega_T) \), which is absolutely continuous with respect to the \( p \)-parabolic capacity. We also denote by \( \mathcal{M}_r(\Omega_T) \) the space of finite signed Radon measures in \( \Omega_T \) with support on a set of zero \( p \)-parabolic capacity. Then a measure \( \mu \in \mathcal{M}_b(\Omega_T) \) can be uniquely decomposed into the following two components: (see [33, Lemma 2.1])

\[ \mu = \mu_a + \mu_s, \quad \mu_a \in \mathcal{M}_a(\Omega_T), \quad \mu_s \in \mathcal{M}_s(\Omega_T). \]

Also, \( \mu_a \in \mathcal{M}_a(\Omega_T) \) if and only if \( \mu_a \) can be written as sum of the following functions:

\[ \mu_a = f + g_t + \text{div} \, G, \]

where \( f \in L^1(\Omega_T), g \in L^p(0,T;V) \) and \( G \in L^{p'}(\Omega_T) \) (see [30, 41]). We write \( \mu = \mu^+ - \mu^- \), where \( \mu^+ \) and \( \mu^- \) are the positive and negative parts, respectively, of a measure \( \mu \in \mathcal{M}_b(\Omega_T) \) and set \( [\mu] := \mu^+ + \mu^- \).

Let us define the truncation operator

\[ T_k(s) := \max\{-k, \min\{k, s\}\} \quad \text{for any } k > 0 \text{ and } s \in \mathbb{R}. \]

If \( u \) is a measurable function defined in \( \Omega_T \), finite almost everywhere, such that \( T_k(u) \in L^p(0,T;W^{1,p}_0(\Omega)) \) for any \( k > 0 \), then there exists a unique measurable function \( U \) such that \( DT_k(u) = U\chi(|u|<k) \) a.e. in \( \Omega_T \) for all \( k > 0 \). In this case, we denote the spatial gradient \( Du \) of \( u \) by \( Du := U \). If \( u \in L^1(0,T;W^{1,1}(\Omega)) \), then it coincides with the usual weak gradient.

Now we introduce the definition of renormalized solution given in [61].

**Definition 1.1.** Let \( p > 1 \) and let \( \mu = \mu_a + \mu_s \in \mathcal{M}_b(\Omega_T) \) with \( \mu_a \in \mathcal{M}_a(\Omega_T) \) and \( \mu_s \in \mathcal{M}_s(\Omega_T) \). A function \( u \in L^1(\Omega_T) \) is a renormalized solution of the problem (1.1) if \( T_k(u) \in L^p(0,T;W^{1,p}_0(\Omega)) \) for every \( k > 0 \) and the following property holds: for any \( k > 0 \) there exist sequences of nonnegative measures \( \nu_k^+, \nu_k^- \in \mathcal{M}_a(\Omega_T) \) such that

\[ \nu_k^+ \to \mu^+ s, \quad \nu_k^- \to \mu^- s \quad \text{tightly as } k \to \infty \]

and

\[ -\int_{\Omega_T} T_k(u)\varphi_t \, dxdt + \int_{\Omega_T} (a(DT_k(u),x,t),D\varphi) \, dxdt = \int_{\Omega_T} \varphi \, d\mu_k \]

for every \( \varphi \in W \cap L^\infty(\Omega_T) \) with \( \varphi(\cdot, T) = 0 \), where \( \mu_k := \mu_a + \nu_k^+ - \nu_k^- \).

Here we say that a sequence \( \{\mu_k\} \subset \mathcal{M}_b(\Omega_T) \) converges tightly (or in the narrow topology of measures) to \( \mu \in \mathcal{M}_b(\Omega_T) \) if

\[ \lim_{k \to \infty} \int_{\Omega_T} \varphi \, d\mu_k = \int_{\Omega_T} \varphi \, d\mu. \]
for every bounded and continuous function \( \varphi \) on \( \Omega_T \).

**Remark 1.2.** Since \( \varphi \in W \), a test function \( \varphi \) admits a unique cap_{p'}-quasi continuous representative. This and the regularity of \( T_k(u) \) imply that every term of (1.7) is well defined. Furthermore, (1.7) is equivalent to

\[
T_k(u)_t - \text{div} a(DT_k(u), x, t) = \mu_k \quad \text{in } \mathcal{D}'(\Omega_T).
\]

Since \( T_k(u) \) belongs to \( L^p(0, T; W^{1,p}_0(\Omega)) \), we observe from (1.2) that the measure \( \mu_k \) belongs to \( L^p(0, T; W^{-1,p'}(\Omega)) \); hence we can regard \( T_k(u) \) as a kind of weak solution to (1.8). This is necessary as we select a test function to obtain some comparison estimates in Section 3 below.

**Remark 1.3.** A renormalized solution \( u \) of (1.1) becomes a distributional solution; this is, \( u \) satisfies

\[
- \int_{\Omega_T} u \varphi_t \, dxdt + \int_{\Omega_T} \langle a(Du, x, t), D\varphi \rangle \, dxdt = \int_{\Omega_T} \varphi \, d\mu
\]

for any \( \varphi \in C_c^\infty(\Omega_T) \) (see [61, Proposition 3]). Moreover, we note that if \( \mu \in L^{p'}(0, T; W^{-1,p'}(\Omega)) \), then a renormalized solution coincides with a weak solution (see [60, 61]).

The notion of renormalized solution was first introduced by DiPerna and Lions [27, 28] for study of the Boltzmann and transport equations. This notion is adapted to obtain existence results for elliptic \( p \)-Laplace type equations with general measure data (\( \mu \in \mathcal{M}_b(\Omega) \)) by Dal Maso, Murat, Orsina and Prignet [22]. For parabolic \( p \)-Laplace type problems, we refer to [7] for the case of \( L^1 \) data (\( \mu \in L^1(\Omega_T) \)) and [30, 60] for the case of diffuse (or soft) measure data (\( \mu \in \mathcal{M}_a(\Omega_T) \)). For the case with general measure data (\( \mu \in \mathcal{M}_b(\Omega_T) \)), Petitta [59] proved the existence of a renormalized solution when \( p > 2 - \frac{n}{n-1} \) (see [6] for stability results), and Petitta and Porretta [61] later generalize the result for \( p > 1 \). It is also worthwhile to note that there are different notions of solutions for elliptic and parabolic measure data problems: SOLA (Solution Obtained by Limits of Approximations, see [8–10, 23]), entropy solution (see [5, 11, 66]), and superparabolic solution (see [38, 39]). On the other hand, the uniqueness of a renormalized solution for parabolic measure data problems such as (1.1) remains a major open problem except the following special cases: (i) \( \mu \in L^1(\Omega_T) \) (see [7]), (ii) \( \mu \in \mathcal{M}_a(\Omega_T) \) (see [30, 60]), or (iii) the linear case; that is, \( a(\xi, x, t) = a(x, t)\xi \) (see [59, Section 9]).

**1.2. Main results.** The aim of this paper is to establish global gradient estimates for renormalized solutions to the problem (1.1). For this, let us first introduce the regularity assumptions on the nonlinearity \( a \) and the boundary of \( \Omega \) (see Section 2 below for our basic notation).

**Definition 1.4.** Let \( R > 0 \) and \( \delta \in (0, \frac{1}{n}) \). We say \( (a, \Omega) \) is \((\delta, R)\)-vanishing if

\[
(i) \text{ the nonlinearity } a(\xi, x, t) \text{ satisfies}
\]

\[
\sup_{t_1, t_2, r \in \mathbb{R}} \sup_{0 < r \leq R} \sup_{y \in \mathbb{R}^n} \int_{t_1}^{t_2} \int_{B_r(y)} \Theta(a, B_r(y))(x, t) \, dxdt \leq \delta,
\]

where

\[
\Theta(a, B_r(y))(x, t) := \sup_{\xi \in \mathbb{R}^n \setminus \{0\}} \frac{a(\xi, x, t) - \int_{B_r(y)} a(\xi, \hat{x}, t) \, d\hat{x}}{(|\xi|^{p-1})};
\]
(ii) for each \( y_0 \in \partial \Omega \) and each \( r \in (0, R) \), there exists a new coordinate system \( \{y_1, \ldots, y_n\} \) such that in this coordinate system, the origin is \( y_0 \) and \( B_r(0) \cap \{y_n > \delta r\} \subset B_r(0) \cap \Omega \subset B_r(0) \cap \{y_n > -\delta r\} \).

**Remark 1.5.** (i) The first assumption of Definition 1.4 implies that the map \( x \mapsto \frac{\| x \cdotp \|}{|\xi|} \) is of BMO (Bounded Mean Oscillation) such that its BMO semi-norm is less than or equal to \( \delta \), uniformly in \( \xi \) and \( t \).

(ii) If the second condition of Definition 1.4 holds, then we say \( \Omega \) is called a \((\delta, R)\)-Reifenberg flat domain. This domain includes Lipschitz domain with a sufficiently small Lipschitz constant and has the following geometric properties:

\[
\sup_{0 < r \leq R} \sup_{y \in \Omega} \frac{|B_r(y)|}{|\Omega \cap B_r(y)|} \leq \left( \frac{2}{1 - \delta} \right)^n \leq \left( \frac{16}{7} \right)^n,
\]

\[
\inf_{0 < r \leq R} \inf_{y \in \partial \Omega} \frac{|\Omega^c \cap B_r(y)|}{|B_r(y)|} \geq \left( \frac{1 - \delta}{2} \right)^n \geq \left( \frac{7}{16} \right)^n.
\]

For a further discussion on Reifenberg flat domains, see [17, 46, 71] and the references therein.

We are ready to state the first main result of this paper.

**Theorem 1.6.** Let \( \frac{2n}{n+1} < p \leq 2 - \frac{1}{n+1} \) and let \( 0 < q < \infty \). Then there exists a small constant \( \delta = \delta(n, \Lambda_0, \Lambda_1, p, q) > 0 \) such that the following holds: if \((a, \Omega)\) is \((\delta, R)\)-vanishing for some \( R > 0 \), then for any renormalized solution \( u \) of the problem (1.1) we have

\[
\int_{\Omega_T} |Du|^q \, dx \, dt \leq c \left\{ \int_{\Omega_T} |M_1(\mu)|^{\frac{2}{(n+1)p-2n}} \, dx \, dt + ||\mu||(\Omega_T)^{\beta_0} \right\}
\]

for some constant \( c = c(n, \Lambda_0, \Lambda_1, p, q, R, \Omega_T) \geq 1 \), where \( \beta_0 := \min \left\{ 1, \frac{q(n+2)}{(n+1)p-n} \right\} \).

Here the fractional maximal function of order 1 for \( \mu \), denoted by \( M_1(\mu) \), is defined as

\[
M_1(\mu)(x, t) := \sup_{r > 0} \frac{|\mu|(Q_r(x, t))}{r^{n+1}} \quad \text{for} \quad (x, t) \in \mathbb{R}^n \times \mathbb{R}.
\]

Note that

\[
||\mu||(\Omega_T)^{\alpha} \leq c(n, \Omega_T) \int_{\Omega_T} |M_1(\mu)|^{\alpha} \, dx \, dt \quad \text{for all} \, \alpha > 0.
\]

**Remark 1.7.** (i) The lower bound of \( p \) in Theorem 1.6 comes from the following: (1) comparison estimates below \( L^1 \) spaces (see Remark 3.9 later); (2) the exponent \( \frac{2}{(n+1)p-2n} > 0 \) in (1.11).

(ii) We note that both the constant \( c \) and the exponent \( \frac{2}{(n+1)p-2n} \) in (1.11) tend to \( +\infty \) as \( p \searrow \frac{2n}{n+1} \). The exponent \( \frac{2}{(n+1)p-2n} \) reflects the anisotropic structure (a constant multiple of a solution no longer yields another solution) of the problem (1.1) as well as the structure of the fundamental solution (1.5). Specifically, this exponent comes from a geometric difference between standard and intrinsic parabolic cylinders (see [16, Lemma 5.4]). The exponent \( \frac{2}{(n+1)p-2n} \) also appears in parabolic potential estimates (see [29, 42]).
(iii) If $\frac{2n}{n+1} < p \leq 2$, then we have $\frac{n+2}{(n+1)p-n} \leq \frac{2}{(n+1)p-2n}$. From (1.11) and (1.13),
\[
\int_{\Omega_T} |Du|^q \, dx \, dt \leq c \left\{ \int_{\Omega_T} [M_1(\mu)]^{\frac{2n}{(n+1)p-n}} \, dx \, dt + 1 \right\}.
\]

Remark 1.8. We remark that the elliptic counterpart of the estimate (1.11) (under the range $1 < p \leq 2 - \frac{1}{n}$) was proved by Nguyen and Phuc [53, 55]. Also, they have recently obtained pointwise potential estimates for the same elliptic problems under the range $\frac{2n}{n+1} < p \leq 2 - \frac{1}{n}$ (see [54]). On the other hand, we refer to [31, 32, 45, 50, 51, 62–64, 67, 68] for various regularity results for elliptic measure data problems with $p > 2 - \frac{1}{n}$.

If the measure $\mu$ is time-independent or can be decomposed as in (1.14) below, then we derive more sharp gradient estimate than the estimate (1.11):

Theorem 1.9. Let $\frac{2n}{n+1} < p \leq 2 - \frac{1}{n+1}$ and let $p - 1 < q < \infty$. Suppose that the following decomposition holds:
\[
\mu = \mu_0 \otimes f,
\]
where $\mu_0$ is a finite signed Radon measure on $\Omega$ and $f \in L^{\frac{2n}{n+1}}(0, T)$. Then there exists a small constant $\delta = \delta(n, \Lambda_0, \Lambda_1, p, q) > 0$ such that the following holds: if $(a, \Omega)$ is $(\delta, R)$-vanishing for some $R > 0$, then for any renormalized solution $u$ of the problem (1.1) we have
\[
\int_{\Omega_T} |Du|^q \, dx \, dt \leq c \left\{ \int_{\Omega_T} \left[ |(M_1(\mu_0)) f|^{\frac{2n}{\alpha+2}} \right] dx \, dt + \left[ \|\mu_0(\cdot)|\Omega\| \|f\|_{L^1(0, T)} \right]^{\beta_0} \right\}
\]
for some constant $c = c(n, \Lambda_0, \Lambda_1, p, q, R, \Omega_T) \geq 1$, where $\beta_0 := \min \left\{ 1, \frac{q(n+2)}{(n+1)p-n} \right\}$.

Here the (elliptic) fractional maximal function $M_1(\mu_0)$ is given by
\[
M_1(\mu_0)(x) := \sup_{r > 0} \frac{|\mu_0| (B_r(x))}{r^{n+1}} \quad \text{for} \quad x \in \mathbb{R}^n.
\]
Note that
\[
|\mu_0(\cdot)|\Omega|^{\alpha} \leq c(n, \Omega) \int_{\Omega} |M_1(\mu_0)|^{\alpha} \, dx \quad \text{for all} \quad \alpha > 0.
\]

Remark 1.10. (i) Unlike (1.11), the estimate (1.15) has the form of elliptic estimates (cf. [51, 53, 55, 62, 64]). Since $\frac{2}{(n+1)p-2n} > \frac{1}{p-1}$ under (1.4), we observe that (1.15) gives a more sharp result.

(ii) If $1 < p \leq 2$, then we have $\frac{n+2}{(n+1)p-n} \leq \frac{1}{p-1}$. From (1.15) and (1.17), we deduce
\[
\int_{\Omega_T} |Du|^q \, dx \, dt \leq c \left\{ \int_{\Omega_T} \left[ |(M_1(\mu_0)) f|^{\frac{2n}{\alpha+2}} \right] dx \, dt + 1 \right\}.
\]

1.3. Novelty and outline of the paper. There have been many regularity results for parabolic measure data problems: for instance, Calderón-Zygmund type estimates (see [16]), potential estimates (see [42–44]), and Marcinkiewicz estimates (see [3, 4, 13]). These results are based on the fact that the spatial gradient of a solution belongs to at least the $L^1$ space, thereby the assumption $p > 2 - \frac{1}{n+1}$ is essential to obtain such regularity estimates. However, as mentioned earlier in Section 1.1, the fundamental solution (1.5) is indeed valid when $p > \frac{2n}{n+1}$.
The aim of the present paper is to fill this gap by developing the global gradient estimates for the problem (1.1) under $\frac{3}{n+2} < p \leq 2 - \frac{1}{n+1}$ (see Theorems 1.6 and 1.9). The main difficulty in obtaining Theorems 1.6 and 1.9 lies in that the spatial gradient of a renormalized solution $u$ of (1.1) could not belong to the $L^1$ space. To overcome this situation, we construct some comparison estimate below $L^1$. More precisely, we will show that if $w$ is a weak solution of the homogeneous problem $w_t - \text{div} a(Dw, x, t) = 0$, then $|Du - Dw|$ is bounded in $L^p$, the constant $\theta \in (0, 1)$ to be determined later, under in particular the range $\frac{3}{n+2} < p \leq 2 - \frac{1}{n+1}$ (see Lemma 3.1 below).

This paper is structured as follows:

- In Section 2, we collect basic notation and preliminary results used throughout the paper.
- Section 3 commences with the $L^p$-comparison estimate between $Du$ and $Dw$ (Lemma 3.1). We investigate as well a higher integrability for $Dw$ (Lemma 3.3) and regularity results such as Lipschitz continuity for reference problems, with the purpose of obtaining local comparison estimates below $L^1$ (Propositions 3.8 and 3.10).
- In Section 4, we derive decay estimates for the spatial gradient of a renormalized solution $u$ (Proposition 4.9). For this, we apply a covering argument (Lemma 4.2) developed in [15, 16] under intrinsic parabolic cylinders, having the intrinsic (fractional) maximal operators. Then we describe a relationship between intrinsic and standard fractional maximal functions (Lemmas 4.6 and 4.7).
- In Section 5, using the decay estimates obtained in Section 4, we finally prove Theorems 1.6 and 1.9.

2. Preliminaries

Let us first introduce basic notation, which will be used later. We denote by $c$ to mean a universal positive constant that can be computed in terms of known quantities; the exact value denoted by $c$ may be different from line to line. A point $x \in \mathbb{R}^n$ will be written $x = (x_1, \cdots, x_n)$. Let $B_r(x_0)$ denote the open ball in $\mathbb{R}^n$ with center $x_0$ and radius $r > 0$, and let $B^+_r(x_0) := B_r(x_0) \cap \{x \in \mathbb{R}^n : x_n > 0\}$. We denote by

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

the standard parabolic cylinder in $\mathbb{R}^n \times \mathbb{R} =: \mathbb{R}^{n+1}$ with center $(x_0, t_0) \in \mathbb{R}^{n+1}$, radius $r$ and height $2r^2$. With $\lambda > 0$, we also consider the intrinsic parabolic cylinder

$$Q^\lambda_r(x_0, t_0) := B_r(x_0) \times (t_0 - \lambda^2 - p r^2, t_0 + \lambda^2 - p r^2),$$

see [24, 44, 72] for more detailed information of intrinsic geometry related to the intrinsic parabolic cylinder. We also use the following notation:

$$\Omega_r := \Omega \times (0, T), \quad \Omega_\infty := \Omega \times (-\infty, T), \quad \Omega_\infty := \Omega \times (-T, T),$$

$$K^\lambda_r(x_0, t_0) := Q^\lambda_r(x_0, t_0) \cap \Omega_\infty, \quad I^\lambda_r(t_0) := (t_0 - \lambda^2 - p r^2, t_0 + \lambda^2 - p r^2),$$

$$Q^\lambda_{r^+}(x_0, t_0) := B^+_r(x_0) \times I^\lambda_r(t_0),$$

$$T^\lambda_r(x_0, t_0) := (B_r(x_0) \cap \{x \in \mathbb{R}^n : x_n = 0\}) \times I^\lambda_r(t_0).$$
Let us use both the notation \( f_t \) and \( \partial_t f \) to denote the time derivative of a function \( f \). We denote by \( Df \) the spatial gradient of \( f \). Given a real-valued function \( f \), we write
\[
(f)_+ := \max \{ f, 0 \} \quad \text{and} \quad (f)_- := -\min \{ f, 0 \}.
\]
For each set \( Q \subset \mathbb{R}^{n+1} \), \(|Q|\) is the \((n + 1)\)-dimensional Lebesgue measure of \( Q \) and \( \chi_Q \) is the usual characteristic function of \( Q \). For \( f \in L^1_{\text{loc}}(\mathbb{R}^{n+1}) \), \( f_Q \) stands for the integral average of \( f \) over a bounded open set \( Q \subset \mathbb{R}^{n+1} \); that is,
\[
\bar{f}_Q := \frac{1}{|Q|} \int_Q f(x, t) \, dxdt.
\]
For \( f \in L^1_{\text{loc}}(\mathbb{R}^{n+1}) \) and \( \lambda > 0 \), we define the \((\text{intrinsic})\) \( \lambda \)-maximal function of \( f \) as
\[
\mathcal{M}^\lambda f(x, t) := \sup_{r > 0} \int_{Q^\lambda_r(x, t)} |f(y, s)| \, dyds.
\]
We write
\[
(2.1) \quad \mathcal{M}^\lambda_Q f := \mathcal{M}^\lambda (f \chi_Q)
\]
provided \( f \) is defined on a set \( Q \subset \mathbb{R}^{n+1} \). In particular, it coincides the \( \text{classical} \) maximal function \( \mathcal{M} f \) when \( \lambda = 1 \) or \( p = 2 \).

We give weak \((1,1)\)-estimates for the \( \lambda \)-maximal function as follows:

**Lemma 2.1.** Let \( Q \) be an open set in \( \mathbb{R}^{n+1} \). If \( f \in L^1(Q) \), then there exists a constant \( c = c(n) \geq 1 \) such that
\[
(2.2) \quad |\{ (y, s) \in Q : \mathcal{M}^\lambda_Q f(y, s) > \alpha \}| \leq \frac{c}{\alpha} \int_Q |f(x, t)| \, dxdt
\]
for any \( \alpha > 0 \). Moreover, we have
\[
|\{ (y, s) \in Q : \mathcal{M}^\lambda_Q f(y, s) > 2\alpha \}| \leq \frac{c}{\alpha} \int_{\{ (y, s) \in Q : |f| > \alpha \}} |f| \, dxdt
\]
for any \( \alpha > 0 \).

**Proof.** The proof is directly obtained from [16, Lemma 2.12] with \( f \) replaced by \( f \chi_Q \). \( \square \)

We introduce a useful integral property, which can be easily computed by the Fubini theorem.

**Lemma 2.2.** Let \( Q \) be an open set in \( \mathbb{R}^{n+1} \). For any \( q > l \geq 0 \), we have
\[
(2.3) \quad \int_Q T_k(|f|)^{q-l} |f|^l \, dxdt = (q - l) \int_0^k \lambda^{q-l-1} \left[ \int_{Q \cap \{ |f| > \lambda \}} |f|^l \, dxdt \right] d\lambda,
\]
where \( T_k \) is the truncation operator defined in (1.6). Furthermore, if \( f \in L^q(Q) \), then (2.3) also holds for \( k = \infty \).

We also record an embedding theorem for parabolic Sobolev spaces.

**Lemma 2.3** (See [24, Chapter I, Proposition 3.1]). Let \( q, l \geq 1 \) and let \( Q := B \times (t_1, t_2) \subset \mathbb{R}^n \times \mathbb{R} \). Then there is a constant \( c = c(n, q, l) \geq 1 \) such that for every \( f \in L^\infty(t_1, t_2; L^l(B)) \cap L^q(t_1, t_2; W_0^{1,q}(B)) \), we have
\[
\int_Q |f|^{\frac{q}{q-l}} \, dxdt \leq c \left( \int_Q |Df|^q \, dxdt \right) \left( \sup_{t_1 < t < t_2} \int_{B \times \{t\}} |f|^l \, dx \right)^{\frac{q}{n}}.
\]
3. Local comparison estimates below $L^1$ spaces

In this section we derive local comparison estimates for the spatial gradient of a solution $T_k(u) \in L^p(0, T; W^{1,\theta}_0(\Omega))$ to (1.7) in an intrinsic parabolic cylinder. Here we only consider comparison results near a boundary region, since the counterparts in an interior region can be done in the same way. Also, we obtain these estimates below $L^1$ spaces (see Lemma 3.1 and Propositions 3.8 and 3.10 below), as the spatial gradient of a renormalized solution $u$ to the problem (1.1) does not generally belong to $L^1(\Omega_T)$ under the assumption $p \leq 2 - \frac{1}{n+1}$ (see Section 1.1). As noted in Remark 1.2, $T_k(u)$ becomes a weak solution of (1.7) with $\mu_k \in L^p(0, T; W^{-1,\rho'}(\Omega))$. Throughout this section, we replace $T_k(u)$ by $u$ and $\mu_k$ by $\mu$, and we extend $u$ by zero for $t < 0$ (see Remark 4.3 for the reason for this time extension).

Suppose that $(a, \Omega)$ is $(\delta, R)$-vanishing for some $R > 0$, where $\delta \in (0, \frac{1}{2})$ is to be determined later. Fix any $\lambda > 0$, $(x_0, t_0) \in \Omega_\delta$ and $0 < r \leq \frac{R}{8}$ satisfying

$$B^r_8(x_0) \subset B_r(x_0) \cap \Omega \subset B_r(x_0) \cap \{x \in \mathbb{R}^n : x_n > -16\delta r\}.$$  

In this section, we for simplicity omit denoting the center by $K^\lambda_r \equiv K^\lambda_r(x_0, t_0)$, $B_r \equiv B_r(x_0)$ and $\partial_r K^\lambda_{r'}$.

Let $w$ be the unique weak solution to the Cauchy-Dirichlet problem

$$\begin{cases}
  w_t - \text{div} a(Dw, x, t) = 0 & \text{in } K^\lambda_r, \\
  w = u & \text{on } \partial_r K^\lambda_{r'}.
\end{cases}$$

We first give a comparison estimate for the difference of $Du$ and $Dw$, which is a crucial estimate in this paper.

**Lemma 3.1.** Let $\frac{3n+2}{2n+2} < p \leq 2 - \frac{1}{n+1}$, let $u$ be a weak solution of (1.7) and let $w$ as in (3.2) with (3.1). Then there exists a constant $c = c(n, A_0, p, \theta) \geq 1$ such that

$$\left( \int_{K^\lambda_{r'}} |Du - Dw|^\theta \ dxdt \right)^{\frac{1}{\theta}} \leq c \left[ \frac{\mu((K^\lambda_{2r}))}{|K^\lambda_{r'}|} \right]^{\frac{n+2}{(n+1)p - n}} \left( \int_{K^\lambda_{r'}} |Du|^\theta \ dxdt \right)^{\frac{2(p - \theta)(n+1)}{\theta(n+2)}}$$

for any constant $\theta$ such that $\frac{n+2}{2(n+1)} < \theta < p - \frac{n}{n+1} \leq 1$.

**Proof.** To streamline the proof, we will take the test functions in (3.7) and (3.12) below without the use of the so-called Steklov average (see [24] for its definition, properties and standard use). For simplicity of notation, we temporarily write

$$a(\xi, \eta, t) := a(\xi, x, t) \text{ and } a(a, b) := \{(x, t) \in K^\lambda_{2r} : a < (u - w) \pm (x, t) < b \}$$

for any $0 \leq a < b^{\frac{1}{\gamma - 1}} \leq \infty$ and $0 \leq \gamma < 1$. Let us also introduce the vector field $V : \mathbb{R}^n \to \mathbb{R}^n$ defined by $V(\xi) := |\xi|^{\frac{\gamma - 2}{\gamma}} \xi$ for all $\xi \in \mathbb{R}^n$. Note that for any $\xi_1, \xi_2 \in \mathbb{R}^n$, it holds

$$c^{-1} \left( |\xi_1|^2 + |\xi_2|^2 \right)^{\frac{\gamma - 2}{\gamma}} \leq \frac{|V(\xi_1) - V(\xi_2)|^2}{|\xi_1 - \xi_2|^2} \leq c \left( |\xi_1|^2 + |\xi_2|^2 \right)^{\frac{\gamma - 2}{\gamma}}$$

for some constant $c = c(n, p) \geq 1$ (see [35,50] for a further discussion on the vector field $V$).
Step 1. We will first show that

\[ \sup_{t \in I^\lambda_{\Omega_{B_1}}} |u - w| \, dx \leq |\mu|(K^\lambda_{\Omega_{B_1}}) \]

and

\[ \int_{K^\lambda_{\Omega_{B_1}}} |u - w|^{-\gamma} |V(Du) - V(Dw)|^2 \, dx \leq \frac{c \alpha(1-\gamma)(1-\xi)}{(1-\gamma)(\xi - 1)} |\mu|(K^\lambda_{\Omega_{B_1}}) \]

for any \( 0 \leq \gamma < 1, \alpha > 0 \) and \( \xi > 1 \), where \( c = c(n, \Lambda_0, p) \geq 1 \) and \( \Omega_{B_1} := B_{\Omega_{B_1}} \cap \Omega \). For any fixed \( \varepsilon \) and \( \tilde{\varepsilon} \) with \( \varepsilon > \tilde{\varepsilon} \), choose a test function

\[ \varphi_1 = \pm \min \left\{ 1, \max \left\{ \frac{(u - w)_{+}^{-\gamma} - \tilde{\varepsilon}^{1-\gamma}}{\varepsilon - \tilde{\varepsilon}^{1-\gamma}}, 0 \right\} \right\} \tilde{\varepsilon}, \]

where \( \zeta : \mathbb{R} \to [0, 1] \) is a nonincreasing smooth function satisfying \( \zeta(t) = 0 \) for all \( t \geq \tau \) with \( \tau \in I^\lambda_{\Omega_{B_1}} \). We then directly compute

\[ D\varphi_1 = \frac{1 - \gamma}{\varepsilon - \tilde{\varepsilon}^{1-\gamma}} \chi_{(\varepsilon, \tilde{\varepsilon})}(u - w)^{-\gamma} (Du - Dw). \]

Since \( (u - w)^{-\gamma} \leq \tilde{\varepsilon}^{-\gamma} \) on \( \Omega(\varepsilon, \tilde{\varepsilon}) \), we have \( \varphi_1 \in L^p(I^\lambda_{\Omega_{B_1}}; W^{1,p}_0(\Omega_{B_1})) \) with \( |\varphi_1| \leq 1 \) and \( \varphi_1(\cdot, \lambda^{2-p}(8r)^2) = 0 \). Substituting \( \varphi_1 \) into the weak formulation of the subtracted equation of (1.7) and (3.2) and then integrating on \( I^\lambda_{\Omega_{B_1}} \), we obtain

\[ \int_{K^\lambda_{\Omega_{B_1}}} D_t(u - w) \varphi_1 \, dx dt + \int_{K^\lambda_{\Omega_{B_1}}} (a(Du) - a(Dw), D\varphi_1) \, dx dt = -\int_{K^\lambda_{\Omega_{B_1}}} \varphi_1 \, d\mu. \]

To estimate \( I_1 \), we compute

\[ D_t(u - w) \min \left\{ 1, \max \left\{ \frac{(u - w)_{+}^{1-\gamma} - \tilde{\varepsilon}^{1-\gamma}}{\varepsilon - \tilde{\varepsilon}^{1-\gamma}}, 0 \right\} \right\} \]

\[ = \pm D_t \int_{\varepsilon}^{\tilde{\varepsilon}} \min \left\{ 1, \max \left\{ \frac{s^{1-\gamma} - \tilde{\varepsilon}^{1-\gamma}}{\varepsilon - \tilde{\varepsilon}^{1-\gamma}}, 0 \right\} \right\} ds. \]

Then the integration by parts gives

\[ I_1 = \int_{K^\lambda_{\Omega_{B_1}}} \left[ \zeta(u - w)_{+}^{1-\gamma} \min \left\{ 1, \frac{s^{1-\gamma} - \tilde{\varepsilon}^{1-\gamma}}{\varepsilon - \tilde{\varepsilon}^{1-\gamma}} \right\} ds \right] (-\zeta_t) \, dx dt \geq 0, \]

since \( \zeta_t \leq 0 \). Also, we have from (1.3) that

\[ I_2 = \frac{1 - \gamma}{\varepsilon - \tilde{\varepsilon}^{1-\gamma}} \int_{\varepsilon}^{\tilde{\varepsilon}} \zeta(u - w)_{+}^{1-\gamma} (a(Du) - a(Dw), Du - Dw) \, dx dt \geq 0 \]

and from \( |\varphi_1| \leq 1 \) that

\[ |I_3| = \left| \int_{K^\lambda_{\Omega_{B_1}}} \varphi_1 \, d\mu \right| \leq |\mu|(K^\lambda_{\Omega_{B_1}}). \]

Utilizing the three inequalities above in (3.8) and letting \( \tilde{\varepsilon} \to 0 \), we derive

\[ \int_{K^\lambda_{\Omega_{B_1}}} \left[ \zeta(u - w)_{+}^{1-\gamma} \min \left\{ 1, \frac{s^{1-\gamma}}{\varepsilon} \right\} ds \right] (-\zeta_t) \, dx dt \leq |\mu|(K^\lambda_{\Omega_{B_1}}) \]
and
\[ (3.11) \quad \frac{1 - \gamma}{\varepsilon} \int_{A(0, \varepsilon)} \zeta (u - w)^{1-\gamma} (a(Du) - a(Dw), Du - Dw) \, dxdt \leq |\mu|(K^\lambda_{8r}). \]

As \( \varepsilon \to 0 \) in (3.10), Lebesgue’s dominated convergence theorem implies
\[ \int_{K^\lambda_{8r}} |u - w| (-\zeta_t) \, dxdt \leq |\mu|(K^\lambda_{8r}). \]

We then let \( \zeta \) approximate the characteristic function \( \chi_{(-\infty, \tau)} \) to obtain
\[ \int_{\Omega_{8r} \times \{\tau\}} |u - w| \, dx \leq |\mu|(K^\lambda_{8r}) \]
for every \( \tau \in I^\lambda_{8r} \), which implies the estimate (3.5).

To obtain (3.6), we take another test function
\[ (3.12) \quad \varphi_2 = \frac{\varphi_1}{(\alpha^{1-\gamma} + (u - w)^{1-\gamma})^{\xi-1}}, \]
where \( 0 \leq \gamma < 1, \alpha > 0 \) and \( \xi > 1 \) are to be determined later in a universal way.

Testing \( \varphi_2 \) to the subtracted equation of (1.7) and (3.2) and then integrating over \( I^\lambda_{8r} \), we get
\[ (3.13) \quad \int_{K^\lambda_{8r}} \partial_t (u - w) \varphi_2 \, dxdt + \int_{K^\lambda_{8r}} \langle a(Du) - a(Dw), D\varphi_2 \rangle \, dxdt = \int_{K^\lambda_{8r}} \varphi_2 \, d\mu. \]

Since \( \varphi_2 \leq \alpha^{(1-\gamma)(1-\xi)} \varphi_1 \), we employ (3.9) and (3.10) to discover
\[ \left| \lim_{\varepsilon \to 0} \int_{K^\lambda_{8r}} \varphi_2 \, d\mu \right| \leq \alpha^{(1-\gamma)(1-\xi)}|\mu|(K^\lambda_{8r}) \]
and
\[ \lim_{\varepsilon \to 0} \int_{K^\lambda_{8r}} \partial_t (u - w) \varphi_2 \, dxdt \leq \alpha^{(1-\gamma)(1-\xi)}|\mu|(K^\lambda_{8r}). \]

To estimate the second term on the left-hand side of (3.13), we establish
\[
\int_{K^\lambda_{8r}} \langle a(Du) - a(Dw), D\varphi_2 \rangle \, dxdt \\
= \int_{K^\lambda_{8r}} \frac{\langle a(Du) - a(Dw), D\varphi_1 \rangle}{(\alpha^{1-\gamma} + (u - w)^{1-\gamma})^{\xi-1}} \, dxdt \\
+ (1 - \xi) \int_{K^\lambda_{8r}} \frac{\varphi_1 \langle a(Du) - a(Dw), D(u - w)^{1-\gamma} \rangle}{(\alpha^{1-\gamma} + (u - w)^{1-\gamma})^{\xi}} \, dxdt \\
=: I_4 + I_5.
\]

It follows from (3.11) that
\[ \lim_{\varepsilon \to 0} I_4 \leq \alpha^{(1-\gamma)(1-\xi)}|\mu|(K^\lambda_{8r}). \]
As $\varepsilon \to 0$, we have

$$ I_5 \to (1-\varepsilon) \int_{K^\lambda_{8r}} \zeta \min \left\{ 1, \frac{(u-w)^{1-\gamma}}{\varepsilon} \right\} \frac{\langle a(Du) - a(Dw), D(u-w)^{1-\gamma} \rangle}{(\alpha^{1-\gamma} + (u-w)^{1-\gamma})^\xi} \, dx \, dt. $$

We then insert the previous estimates into (3.13) and utilize (1.3) and (3.4), to discover

$$ \int_{K^\lambda_{8r}} |u-w|^{-\gamma} |V(Du) - V(Dw)|^2 \min \left\{ 1, \frac{|u-w|^{1-\gamma}}{\varepsilon} \right\} \, dx \, dt \leq c \left( \frac{\varepsilon^{1-\gamma}(1-\xi)}{(1-\gamma)(\xi-1)} |\mu|(K^\lambda_{8r}) \right) \, dx \, dt. $$

for some constant $c = c(n, \Lambda_0, p) \geq 1$. As $\varepsilon \to 0$, we obtain the estimate (3.6).

**Step 2.** Let $\theta$ be such that $\frac{n+2}{2(n+1)} < \theta < p - \frac{n}{n+1} \leq 1$. For fixed $\varepsilon > 0$, set $\mathcal{B}_\varepsilon := \{(x,t) \in K^\lambda_{8r} : |u-w| > \varepsilon\}$. Let $\beta \in \left[0, \frac{2}{3}\right)$ be the constant satisfying $\frac{\beta n}{p} = \frac{(1-\theta)(\alpha n+1)}{n}$. Define

$$ M_\varepsilon := \frac{\varepsilon}{\beta} \int_{K^\lambda_{8r}} |D|u-w|^{(1-\theta)(\alpha n+1)} |\chi_{\mathcal{B}_\varepsilon}| \, dx \, dt. $$

Indeed, $M_\varepsilon < \infty$ since $|u-w| > \varepsilon$ on $\mathcal{B}_\varepsilon$. We see from Hölder’s inequality that

$$ \int_{K^\lambda_{8r}} |Du - Dw|^\theta |\chi_{\mathcal{B}_\varepsilon}| \, dx \, dt = \int_{K^\lambda_{8r}} \left| |u-w|^{-\frac{(1-\theta)(\alpha n+1)}{n}} |Du - Dw| \right| ^\theta \, dx \, dt \leq M_\varepsilon^{\theta} \left( \int_{K^\lambda_{8r}} |u-w|^{\frac{(\theta n+1)}{n}} |\chi_{\mathcal{B}_\varepsilon}| \, dx \, dt \right)^{1-\theta}. $$

Applying Lemma 2.3 with $f = \left( |u-w|^{(\alpha n+1)} - \varepsilon \right)_+^\theta$, $q = 1$ and $l = \frac{\varepsilon}{\beta}$, we find

$$ \int_{K^\lambda_{8r}} |u-w|^{\frac{\theta(n+1)}{n}} |\chi_{\mathcal{B}_\varepsilon}| \, dx \, dt \leq c \int_{K^\lambda_{8r}} \left( |u-w|^{\frac{\alpha n+1}{p}} - \varepsilon \frac{\alpha n+1}{p} \right)^{\frac{\theta(n+1)}{n}} \, dx \, dt + c\varepsilon^\frac{\theta(n+1)}{n}, $$

by noting that $\frac{\theta(n+1)}{n} = \frac{(n+1)p-\beta n}{np}$. Let us set

$$ \alpha_\varepsilon := \left[ |\mu|(K^\lambda_{8r}) M_\varepsilon^{\frac{1}{n}} \right]^{\frac{1}{(n+1)}} + \varepsilon. $$

Then it follows from (3.5) and (3.15) that

$$ \int_{K^\lambda_{8r}} |u-w|^{\frac{\theta(n+1)}{n}} |\chi_{\mathcal{B}_\varepsilon}| \, dx \, dt \leq c \varepsilon^\frac{\theta(n+1)}{n}. $$
then we have as a consequence of (3.18), we obtain

$$
\int_{K_{8 r}} |Du - Dw|^\theta \chi_{\mathcal{B}_r} \, dx dt \leq c M_\epsilon^\theta \alpha_\epsilon^{-\theta(n+1)}
$$

for some constant $c(n, \Lambda_0, p, \theta) \geq 1$.

Now, we will estimate the quantity $M_\epsilon$. We notice from (3.4) that

$$
|Du - Dw| \leq c |V(Du) - V(Dw)|^{\frac{2}{p}} + c |Du|^{\frac{2\alpha}{p}} |V(Du) - V(Dw)|.
$$

Then we have

$$
M_\epsilon \leq c \int_{K_{8 r}} |u - w|^{-\frac{2}{p}} |V(Du) - V(Dw)|^{\frac{2}{p}} \chi_{\mathcal{B}_r} \, dx dt \tag{3.19}
$$

$$
= J_1 + c \int_{K_{8 r}} |u - w|^{-\frac{2}{p}} |Du|^{\frac{2\alpha}{p}} |V(Du) - V(Dw)| \chi_{\mathcal{B}_r} \, dx dt.
$$

From (3.6) with $\gamma = \beta \in [0, 1)$ and $\alpha = \alpha_\epsilon > 0$, we have

$$
J_1 \leq \int_{K_{8 r}} \left( \frac{|u - w|^{-\beta} |V(Du) - V(Dw)|^{2}}{\left( \alpha_\epsilon^{1-\beta} + |u - w|^{1-\beta} \right)^{\xi}} \right)^{\frac{1}{p}} \left( \alpha_\epsilon^{1-\beta} + |u - w|^{1-\beta} \right)^{\frac{p-1}{p}} \chi_{\mathcal{B}_r} \, dx dt
$$

$$
\leq \left( \int_{K_{8 r}} |u - w|^{-\beta} |V(Du) - V(Dw)|^{2} \chi_{\mathcal{B}_r} \, dx dt \right)^{\frac{1}{p}} \left( \int_{K_{8 r}} \left( \alpha_\epsilon^{1-\beta} + |u - w|^{1-\beta} \right)^{\frac{p-1}{p}} \chi_{\mathcal{B}_r} \, dx dt \right)^{\frac{1}{p}}
$$

$$
\leq c \alpha_\epsilon^{\frac{(1-\beta)\xi}{p}} \left[ \frac{|\mu| (K_{8 r})}{|K_{8 r}|} \right]^{\frac{1}{p}} \left\{ \alpha_\epsilon^{\frac{(1-\beta)\xi}{p}} + \left( \int_{K_{8 r}} |u - w|^{\frac{(1-\beta)\xi}{p-1}} \chi_{\mathcal{B}_r} \, dx dt \right)^{\frac{p-1}{p}} \right\}.
$$

Furthermore since

$$
\theta < p - \frac{n}{n+1} \iff 1 - \beta < \frac{(n+1)p - n\beta}{np} = \frac{\theta(n+1)}{n},
$$

we can choose $\xi > 1$ so that $\frac{(1-\beta)\xi}{p-1} < \frac{\theta(n+1)}{n}$, and then we discover

$$
\left( \int_{K_{8 r}} |u - w|^{\frac{(1-\beta)\xi}{p-1}} \chi_{\mathcal{B}_r} \, dx dt \right)^{\frac{p-1}{p}} \leq \left( \int_{K_{8 r}} |u - w|^{\frac{\theta(n+1)}{n}} \chi_{\mathcal{B}_r} \, dx dt \right)^{\frac{1-\beta\xi}{p}}
$$

as a consequence of (3.17). Therefore we have

$$
J_1 \leq c \alpha_\epsilon^{\frac{1-\beta}{p}} \left[ \frac{|\mu| (K_{8 r})}{|K_{8 r}|} \right]^{\frac{1}{p}}.
$$

(3.20)
To estimate $J_2$, we have from (3.6) with $\gamma = \frac{2\beta}{p} \in [0, 1)$ and $\alpha = \alpha_\varepsilon > 0$ that

$$J_2 \leq \int_{K_\varepsilon} \left( \frac{|u - w|^{-(\gamma - 1)} |V(Du) - V(Dw)|^2}{(\alpha_\varepsilon^{(\frac{1}{\gamma})} + |u - w|^{1-\gamma})^\xi} \right)^{\frac{1}{\xi}} (\alpha_\varepsilon^{(\frac{1}{\gamma})} + |u - w|^{1-\gamma})^\xi |Du|^{\frac{2-p}{\theta}} \chi_{\mathcal{B}_\varepsilon} \, dx dt$$

\[
\leq c\alpha_\varepsilon \frac{((1-\gamma)\xi)^{\frac{\theta}{\theta-2+\rho}}}{(\theta-2+\rho)^{\frac{\theta}{\theta-2+\rho}}} \left( \int_{K_\varepsilon} (\alpha_\varepsilon^{(\frac{1}{\gamma})} + |u - w|^{1-\gamma})^{\frac{\theta}{\theta-2+\rho}} \chi_{\mathcal{B}_\varepsilon} \, dx dt \right)^{\frac{\theta}{\theta-2+\rho}}.
\]

From the fact that

$$\theta > \frac{n + 2}{2(n + 1)} > 2 - p \quad \text{(since $p \geq \frac{3n + 2}{2n + 2}$)},$$

it follows

$$\int_{K_\varepsilon} (\alpha_\varepsilon^{(\frac{1}{\gamma})} + |u - w|^{1-\gamma})^{\xi} |Du|^{2-p} \chi_{\mathcal{B}_\varepsilon} \, dx dt$$

\[\leq \left( \int_{K_\varepsilon} |Du|^\theta \, dx dt \right)^{\frac{2-p}{\theta}} \left( \int_{K_\varepsilon} (\alpha_\varepsilon^{(\frac{1}{\gamma})} + |u - w|^{1-\gamma})^{\frac{\theta}{\theta-2+\rho}} \chi_{\mathcal{B}_\varepsilon} \, dx dt \right)^{\frac{\theta-2+\rho}{\theta}}.
\]

In addition, from the following relation

$$\theta < p - \frac{n}{n + 1} \iff (1-\gamma)\theta \frac{\theta}{\theta-2+\rho} + (n + 1)p - n\beta \leq \theta(n + 1)$$

we can take $\xi > 1$ so that $\frac{(1-\gamma)\theta}{\theta-2+\rho} < \frac{\theta(n + 1)}{n}$, and then we discover

$$\left( \int_{K_\varepsilon} (\alpha_\varepsilon^{(\frac{1}{\gamma})} + |u - w|^{1-\gamma})^{\frac{\theta}{\theta-2+\rho}} \chi_{\mathcal{B}_\varepsilon} \, dx dt \right)^{\frac{\theta-2+\rho}{\theta}}$$

\[\leq c\alpha_\varepsilon (1-\gamma)\xi + c \left( \int_{K_\varepsilon} |u - w|^{\frac{n\theta(n + 1)}{n}} \chi_{\mathcal{B}_\varepsilon} \, dx dt \right)^{(1-\gamma)\frac{n}{\theta(n + 1)}} \leq c\alpha_\varepsilon (1-\gamma)\xi
\]

by using (3.17). Thus we have

$$J_2 \leq c\alpha_\varepsilon \frac{\frac{\theta-2+\rho}{\theta}}{\theta} \left[ \frac{|\mu(K_{\varepsilon})|}{|K_{\varepsilon}|} \right]^{\frac{\theta}{\theta-2+\rho}} \left( \int_{K_\varepsilon} |Du|^\theta \, dx dt \right)^{\frac{2-p}{\theta}}.$$

If $\lim_{\varepsilon \to 0} M_\varepsilon = 0$, then $Du \equiv Dw$ a.e. in $K_{\varepsilon}^\lambda$, and so the proof is done. Thus, we may assume $\inf_{\varepsilon > 0} M_\varepsilon > 0$, and then there exists a constant $\varepsilon_0 > 0$ such that $0 < \varepsilon < \left[ |\mu(K_{\varepsilon})| M_\varepsilon^{\theta(n + 1)} \right]^{\frac{1}{\theta(n + 1)}}$ whenever $0 < \varepsilon < \varepsilon_0$. Consequently, we see from (3.16) that

$$\alpha_\varepsilon < 2 \left[ |\mu(K_{\varepsilon})| M_\varepsilon^n \right]^{\frac{1}{\theta(n + 1)}} \quad \text{for all} \quad 0 < \varepsilon < \varepsilon_0.$$
Inserting (3.20) and (3.21) into (3.19), we employ (3.22) and Young’s inequality to discover
\[
M_{\varepsilon} \leq c \left[ \frac{|\mu|(K^{\lambda}_{K_{\varepsilon}^r})}{|K^{\lambda}_{K_{\varepsilon}^r}|} \right]^{\frac{\theta(n+1)}{(n+1)p-n}} \left[ \frac{|\mu|(K^{\lambda}_{K_{\varepsilon}^r})}{|K^{\lambda}_{K_{\varepsilon}^r}|} \right]^{\frac{1-\beta}{n+1+p-n}}
\]
(3.23)
\[
+ c \left[ \frac{|\mu|(K^{\lambda}_{K_{\varepsilon}^r})}{|K^{\lambda}_{K_{\varepsilon}^r}|} \right]^{\frac{\theta(n+1)}{(n+1)p-n}} \left[ \frac{|\mu|(K^{\lambda}_{K_{\varepsilon}^r})}{|K^{\lambda}_{K_{\varepsilon}^r}|} \right]^{\frac{1-\beta}{n+1+p-n}} \left( \int_{K^{\lambda}_{K_{\varepsilon}^r}} |Du|^\theta \, dx \right)^{\frac{(2-p)(n+1)}{(n+1)+\beta}}
\]
Finally, we combine (3.18), (3.22) and (3.23) to obtain
\[
\left( \int_{K^{\lambda}_{K_{\varepsilon}^r}} |Du - Dw|^\theta \chi_{B_r} \, dx \right)^{\frac{1}{\theta}} \leq c \left[ \frac{|\mu|(K^{\lambda}_{K_{\varepsilon}^r})}{|K^{\lambda}_{K_{\varepsilon}^r}|} \right]^{\frac{n+2}{n+1+p-n}} \left[ \frac{|\mu|(K^{\lambda}_{K_{\varepsilon}^r})}{|K^{\lambda}_{K_{\varepsilon}^r}|} \right]^{\frac{1-\beta}{n+1+p-n}} \left( \int_{K^{\lambda}_{K_{\varepsilon}^r}} |Du|^\theta \, dx \right)^{\frac{(2-p)(n+1)}{(n+1)+\beta}}
\]
whenever $0 < \varepsilon < \varepsilon_0$. Letting $\varepsilon \to 0$, we obtain the desired estimate (3.3). \qed

**Remark 3.2.** The approach used in the proof above is motivated by [42,44,53]. This approach is applicable to comparison estimates for elliptic problems (with $p \leq 2 - \frac{1}{n}$), which indeed gives a different way from that of [53, Lemma 2.2]. On the other hand, for comparison estimates like (3.3) with $p > 2 - \frac{1}{n+1}$, we refer to [44, Lemma 4.1] and [42, Lemma 4.3].

The next lemma gives a boundary self-improving result for $Dw$ (see [37] for an interior version).

**Lemma 3.3.** Let $\frac{2n}{n+\sigma} < p \leq 2$ and let $\frac{(2-p)n}{2} < \theta \leq p$. If $w$ is the weak solution of (3.2) satisfying (3.1) and
\[
\int_{K^{\lambda}_{K_{\varepsilon}^r}} |Dw|^\theta \, dx \leq c_w \lambda^\theta
\]
for some constant $c_w \geq 1$, then there exist two constants $\sigma = \sigma(n, \Lambda_0, \Lambda_1, p, \theta) > 0$ and $c = c(n, \Lambda_0, \Lambda_1, p, \theta, c_w) \geq 1$ such that
\[
\int_{K^{\lambda}_{K_{\varepsilon}^r}} |Dw|^{p(1+\sigma)} \, dx \leq c \lambda^{p(1+\sigma)}.
\]

**Proof.** From [12, Theorem 2.2] and [34, Remark 6.12] (see also [16, Lemma 4.2]), we infer
\[
\int_{K^{\lambda}_{K_{\varepsilon}^r}} |Dw|^{p(1+\sigma)} \, dx \leq c \lambda^{p(1+\sigma)} \left[ \left( \int_{K^{\lambda}_{K_{\varepsilon}^r}} |Dw|^{p_\sigma} \, dx \right)^{\frac{1+\sigma}{1+\sigma+p}} + 1 \right]
\]
for every $s \in \left( \frac{(2-p)n}{2}, 1 \right]$, where $c = c(n, \Lambda_0, \Lambda_1, p, s) \geq 1$ and $\sigma := \frac{2}{(n+2)p-2n}$. By taking $s = \frac{p}{p}$ and using (3.24), we obtain the desired estimate (3.25). \qed

**Remark 3.4.** (i) It is worth noting that the scaling deficit $\theta$ and interpolation inequality give the lower bound of $s$ in (3.26), which determines the range of $\theta$. For elliptic equations with $p$-growth ($p > 1$), on the other hand, the estimate like (3.26) holds for every $s \in (0, 1]$ (see [34, Remark 6.12]).
The proof follows from Lemma 3.5.

**Lemma 3.5.** Let $p > \frac{2n}{n+2}$, let $w$ be the weak solution of (3.2) satisfying (3.1) and (3.24), and let $v$ as in (3.27). Then there is a constant $c = c(n, \Lambda_0, \Lambda_1, p) \geq 1$ such that

$$\int_{K^*_4} |Dw - Dv|^p \, dx \, dt \leq c\delta^\eta \lambda^p,$$

where $\sigma = \sigma_1(n, \Lambda_0, \Lambda_1, p) > 0$.

**Proof.** The proof follows from Lemma 3.3, (1.9), (1.10) and [14, Lemma 3.10].

For interior regularity results (see [24–26]), we see $Dv \in L^\infty_{loc}(Q^4_n)$ in the interior region $(Q^4_n \subset \Omega_T)$. On the other hand, for the boundary case $(Q^4_n \not\subset \Omega_T)$, the $L^\infty$-norm of $Dv$ could not be bounded when $\partial \Omega$ is very irregular. Thus, we need to consider a weak solution $\bar{v}$ to the following problem:

$$\left\{ \begin{array}{ll} \bar{v}_t - \text{div} \bar{a}^+_n(D\bar{v}, t) = 0 & \text{in } Q^4_n, \\ \bar{v} = 0 & \text{on } T^4_n, \end{array} \right.$$  \hspace{1cm} (3.28)

We recall the boundedness of $D\bar{v}$ near the flat boundary, as follows:

**Lemma 3.6** (See [48, Theorem 1.6]). Let $p > \frac{2n}{n+2}$. For any weak solution $\bar{v}$ of (3.28), we have

$$||D\bar{v}||^p_{L^\infty(Q^4_n)} \leq c \int_{Q^4_n} |D\bar{v}|^p \, dx \, dt + c\lambda^p$$

for some constant $c = c(n, \Lambda_0, \Lambda_1, p) \geq 1$.

If the boundary of $\Omega$ is sufficiently flat in the sense of $(\delta, R)$-Reifenberg domain, for some appropriate weak solution $\bar{v}$ of (3.28) we have a comparison estimate between (3.27) and (3.28) as follows:

**Lemma 3.7.** Let $p > \frac{2n}{n+2}$. For any $\varepsilon \in (0, 1)$, there exists a small constant $\delta = \delta(n, \Lambda_0, \Lambda_1, p, \varepsilon) > 0$ such that the following holds: if $\bar{v}$ is the weak solution of (3.27) satisfying (3.1) and

$$\int_{K^*_4} |Dv|^p \, dx \, dt \leq c_v \lambda^p$$

for some given constant $c_v \geq 1$, then there is a weak solution $\bar{v}$ of (3.28) such that

$$\int_{K^*_4} |Dv - D\bar{v}|^p \, dx \, dt \leq \varepsilon^p \lambda^p \quad \text{and} \quad \int_{K^*_4} |D\bar{v}|^p \, dx \, dt \leq c\lambda^p$$  \hspace{1cm} (3.29)
for some constant \( c = c(n, \Lambda_0, \Lambda_1, p, c_v) \geq 1 \), where \( \bar{v} \) is extended by zero from \( Q_{2r}^{\Lambda, +} \) to \( K_{2r}^{\Lambda} \).

**Proof.** The first estimate in (3.29) comes from the compactness argument as in [14, Lemma 3.8]. It follows from this first estimate and (1.10) that the second estimate in (3.29) holds. \( \square \)

Finally, combining all the previous results, we directly obtain the desired local comparison estimate below \( L^1 \) spaces near the boundary of \( \Omega \).

**Proposition 3.8.** Let \( \frac{2n}{n + 1} < p \leq 2 - \frac{1}{n + 1} \) and let \( \max \left\{ \frac{n + 2}{2(n + 1)}, \frac{(2 - p)n}{2} \right\} < \theta < p - \frac{n}{n + 1} \leq 1 \). For any \( \varepsilon \in (0, 1) \), there is a small constant \( \delta = \delta(n, \Lambda_0, \Lambda_1, p, \theta, \varepsilon) > 0 \) such that the following holds: if \( u \) is a weak solution of (1.7) satisfying (3.1),

\[
\int_{K_{2r}^{\Lambda}} |Du|^\theta \, dx \leq \Lambda^\theta \text{ and } \frac{|\mu(K_{2r}^{\Lambda})|}{\rho^{n+1}} \leq \delta \lambda,
\]

then there exists a weak solution \( \bar{v} \) of (3.28) such that

\[
\int_{K_{2r}^{\Lambda}} |Du - D\bar{v}|^\theta \, dx \leq \varepsilon \Lambda^\theta \text{ and } \|D\bar{v}\|_{L^\infty(K_{2r}^{\Lambda})} \leq c \lambda
\]

for some constant \( c = c(n, \Lambda_0, \Lambda_1, p, \theta) \geq 1 \), where \( \bar{v} \) is extended by zero from \( Q_{2r}^{\Lambda, +} \) to \( K_{2r}^{\Lambda} \).

**Remark 3.9.** (i) In view of Lemmas 3.1 and 3.3, the valid range of \( p \) in Proposition 3.8 is \( \frac{2n}{n + 1} < p \leq 2 - \frac{1}{n + 1} \), not \( \max \left\{ \frac{n + 2}{2(n + 1)}, \frac{(2 - p)n}{2} \right\} < p \leq 2 - \frac{1}{n + 1} \), since the constant \( \theta \) exists only when \( p - \frac{n}{n + 1} > \max \left\{ \frac{n + 2}{2(n + 1)}, \frac{(2 - p)n}{2} \right\} \). Note that \( \frac{2n}{n + 1} \geq \max \left\{ \frac{n + 2}{2(n + 1)}, \frac{2n}{n + 2} \right\} \), where the equality holds if and only if \( n = 2 \).

(ii) We note that the value \( \frac{|\mu(K_{2r}^{\Lambda})|}{\rho^{n+1}} \) in (3.30) is related to the intrinsic fractional maximal function \( M_{2r}^{\Lambda}(\mu) \) defined in (4.7), see Section 4 below.

Similarly, we have an interior comparison estimate below \( L^1 \) spaces.

**Proposition 3.10.** Let \( \frac{2n}{n + 1} < p \leq 2 - \frac{1}{n + 1} \) and let \( \max \left\{ \frac{n + 2}{2(n + 1)}, \frac{(2 - p)n}{2} \right\} < \theta < p - \frac{n}{n + 1} \leq 1 \). For any \( \varepsilon \in (0, 1) \), there is a small constant \( \delta = \delta(n, \Lambda_0, \Lambda_1, p, \theta, \varepsilon) > 0 \) such that the following holds: if \( u \) is a weak solution of (1.7) satisfying

\[
\int_{Q_{2r}^{\Lambda}} |Du|^\theta \, dx \leq \Lambda^\theta \text{ and } \frac{|\mu(Q_{2r}^{\Lambda})|}{\rho^{n+1}} \leq \delta \lambda,
\]

then there exists a weak solution \( v \) of (3.27) such that

\[
\int_{Q_{2r}^{\Lambda}} |Du - Dv|^\theta \, dx \leq \varepsilon \Lambda^\theta \text{ and } \|Dv\|_{L^\infty(Q_{2r}^{\Lambda})} \leq c \lambda
\]

for some constant \( c = c(n, \Lambda_0, \Lambda_1, p, \theta) \geq 1 \).
4. λ-COVERING ARGUMENTS

We now consider a renormalized solution \( u \) of the problem \((1.1)\). We denote by \( u_k := T_k(u) \) \((k \in \mathbb{N})\) the truncation of \( u \) and \( \mu_k \in L^p(0, T; W^{-1,p}(\Omega)) \) the corresponding weak solutions of \((1.2)\), \((3.2)\) and \((3.27)\), respectively. The goal of this section is to derive an appropriate decay estimate for the upper level set of the λ-maximal function of \( |Du|^\theta \) (see Proposition 4.9 later).

We note that \( \mu_k = \mu_k^+ - \mu_k^- + \nu_k^+ - \nu_k^- \) for \( k \in \mathbb{N} \). Since \( \mu_k^+ + \nu_k^+ \to \mu_k^+ + \mu_k^+ \) tightly as \( k \to \infty \), we have

\[
\limsup_{k \to \infty} |\mu_k|(K \cap \Omega_T) \leq |\mu|(K \cap \Omega_T)
\]

for every compact set \( K \subset \mathbb{R}^{n+1} \). Having this relation in mind, we start with a standard energy type estimate for \((1.1)\) as follows:

**Lemma 4.1.** Let \( \frac{3n+2}{2n+2} < p \leq 2 - \frac{1}{n+1} \). If \( u \) is a renormalized solution of \((1.1)\), then there exists a constant \( c = c(n, \Lambda_0, p, \theta) \geq 1 \) such that

\[
\left( \int_{\Omega_T} |Du|^p \, dx \right)^{\frac{1}{p}} \leq c \left[ \frac{\mu(\Omega_T)}{\text{vol}(\Omega_T)^{\frac{n+2}{n+2}}} \right]^{\frac{1}{p-\frac{n+2}{n+2}}}
\]

for any constant \( \theta \) such that \( 0 < \theta < p - \frac{n}{n+1} \).

**Proof.** Since \( a(0, x, t) = 0 \), the zero function \( \bar{w} \) solves the Cauchy-Dirichlet problem

\[
\begin{align*}
\bar{w}_t - \text{div} a(D\bar{w}, x, t) &= 0 \quad \text{in } \Omega_T, \\
\bar{w} &= 0 \quad \text{on } \partial_p \Omega_T.
\end{align*}
\]

Replacing \( u, w, \mu \) and \( K^\lambda \) by \( u_k, \bar{w}(\equiv 0), \mu_k \) and \( \Omega_T \), respectively in the proof of Lemma 3.1, we deduce

\[
\left( \int_{\Omega_T} |Du_k|^\theta \, dx \right)^{\frac{1}{\theta}} \leq c \left[ \frac{\mu_k(\Omega_T)}{\text{vol}(\Omega_T)^{\frac{n+2}{n+2}}} \right]^{\frac{1}{p-\frac{n+2}{n+2}}}
\]

whenever \( \frac{n+2}{2(n+1)} < \theta < p - \frac{n}{n+1} \). Also, it follows from \((4.1)\) that

\[
\limsup_{k \to \infty} |\mu_k|(\Omega_T) \leq |\mu|(\Omega_T).
\]

Taking the limit supremum of both sides of \((4.3)\), we see that \((4.2)\) holds for all \( \theta \in \left( \frac{n+2}{2(n+1)}, p - \frac{n}{n+1} \right) \). On the other hand, if \( \theta \in \left( 0, \frac{n+2}{2(n+1)} \right) \), we employ Hölder’s inequality to obtain the desired estimate. \( \square \)

We next introduce a modified version of Vitali’s covering lemma for intrinsic parabolic cylinders, which is an important tool for obtaining Proposition 4.9 later.

**Lemma 4.2** (See [16, Lemma 2.14]). Let \( 0 < \varepsilon < 1 \), let \( \lambda > 0 \), and let \( \Omega_T := \Omega \times (-\infty, T) \), where \( \Omega \) is \((\delta, R)\)-Reifenberg flat. Let \( \mathcal{C} \subset \mathcal{D} \subset \Omega_T \) be two bounded measurable subsets such that

(i) \( |\mathcal{C}| < \varepsilon |Q^{\lambda}_{K/10}| \), and

(ii) for any \( (y, s) \in \Omega_T \) and any \( r \in \left( 0, \frac{K}{10} \right] \) with \( |\mathcal{C} \cap Q^\lambda_r(y, s)| \geq \varepsilon |Q^\lambda_r| \)

\( Q^\lambda_r(y, s) \cap \Omega_T \subset \mathcal{D} \).
Then we have

$$|\mathcal{C}| \leq \left(\frac{10}{1 - \delta}\right)^{n+2} \varepsilon |\mathcal{D}| \leq \left(\frac{80}{7}\right)^{n+2} \varepsilon |\mathcal{D}|.$$

**Remark 4.3.**

(i) It is worth noting that the covering lemma above is obtained under $\mathcal{C} \subset \mathcal{D} \subset \Omega_T$, not $\mathcal{C} \subset \mathcal{D} \subset \Omega_T := \Omega \times (0,T)$; thereby the relation (4.4) is independent of $\lambda$ (see [16, Remark 2.15] for details). For this reason, we considered the comparison estimates in Section 3 on the localized region of $\Omega_T$, not $\Omega_T$.

(ii) For Vitali’s covering lemmas with respect to standard balls or cubes, we refer to for instance [17, 18, 51, 73].

Suppose that $(a, \Omega)$ is $(\delta, R)$-vanishing and that $p$ satisfies (1.4), unless otherwise stated. We write

$$\beta := \frac{n + 2}{(n + 1)p - n} \quad \text{and} \quad d := \frac{2}{(n + 1)p - 2n}.$$

For any fixed $\varepsilon \in (0, 1)$, we set

$$\lambda_0 := \left[\frac{|\mu|(|\Omega_T|)}{|\Omega_T| \frac{n}{n+1}}\right]^\beta \varepsilon |\Omega_T| + \left[\frac{|\mu|(|\Omega_T|)}{\delta T \frac{n}{n+1}}\right]^d + 1,$$

where $\theta$ is a constant such that

$$\max \left\{ \frac{n + 2}{2(n + 1)}, \frac{2 - p}{2}, p - 1 \right\} < \theta < p - \frac{n}{n + 1} \leq 1.$$

We remark that the constant $\beta$ arises in Lemma 4.1, $d$ in Lemma 4.6, and $\theta$ in Proposition 3.8 and Lemma 4.4. We may assume, upon letting $u \equiv 0$ for $t < 0$, that a renormalized solution $u$ is defined in $\Omega_T := \Omega \times (-\infty, T)$. For any fixed $N > 1$ and $\lambda \geq \lambda_0 \geq 1$, we write

$$\mathcal{C} := \left\{ (x, t) \in \Omega_T : \mathcal{M}^\lambda_{\Omega_T}|Du|^\theta (x, t) > (N\lambda)^\theta \right\}$$

and

$$\mathcal{D} := \left\{ (x, t) \in \Omega_T : \mathcal{M}^\lambda_{\Omega_T}|Du|^\theta (x, t) > \lambda^\theta \right\} \cup \left\{ (x, t) \in \Omega_T : \mathcal{M}^1_{\Omega_T}(\mu)(x, t) > \delta \lambda \right\},$$

where $\mathcal{M}^\lambda_{\Omega_T}$ is given by (2.1), and the operator $\mathcal{M}^\lambda_1$ is the intrinsic fractional maximal function of order 1 for $\mu$ defined by

$$\mathcal{M}^\lambda_1(\mu)(x, t) := \sup_{r > 0} \frac{|\mu|(Q^\lambda_r(x, t))}{r^{n+1}} \quad \text{for} \quad (x, t) \in \mathbb{R}^{n+1}.$$

From Lemma 4.1, the $\lambda$-maximal function $\mathcal{M}^\lambda_{\Omega_T}|Du|^\theta (x, t)$ is well defined for all $(x, t) \in \Omega_T$. We note that since the support of $|Du|$ is bounded, both $N\lambda$-upper level set and $\lambda$-upper level set of $\mathcal{M}^\lambda_{\Omega_T}|Du|^\theta$ should be bounded. Combining this fact and Lemma 4.6, we infer that both the upper level sets $\mathcal{C}$ and $\mathcal{D}$ are bounded measurable subsets of $\Omega_T$.

Under these settings, we first prove two assumptions of Lemma 4.2.

**Lemma 4.4.** There exists a constant $N_1 = N_1(n, \lambda_0, p, \theta) > 1$ such that for any fixed $N \geq N_1$ and $\lambda \geq \lambda_0$, we have

$$|\mathcal{C}| < \varepsilon \left|Q^\lambda_{N/10}\right|.$$
Proof. Since $\theta > p - 1$ and $\lambda \geq \lambda_0 \geq 1$, we note that $\lambda^{-\theta} \leq \lambda^{1-p} \leq \lambda^{2-p}\lambda_{\theta}^{-1}$. Then we have from (2.2), Lemma 4.1 and (4.5) that

$$
|\mathcal{C}| \leq \frac{c}{(\lambda N)^\theta} \int_{\Omega_T} |Du|^{\theta} \, dx \, dt \leq \frac{c|\Omega_T|}{(\lambda N)^\theta} \left[ \frac{|\mu(\Omega_T)|}{|\Omega_T|^{\lambda \theta}} \right]^{\beta \theta} \\
< \frac{c\lambda^{2-p\varepsilon} |Q_{n/10}|}{N_1^\theta} \leq \varepsilon \left| Q_{n/10}^\lambda \right|
$$

by selecting $N_1$ large enough.

Lemma 4.5. For any $\varepsilon \in (0, 1)$, there exist $N_2 = N_2(n, \Lambda_0, \Lambda_1, p, \theta) > 1$ and $\delta = \delta(n, \Lambda_0, \Lambda_1, p, \theta, \varepsilon) \in (0, \frac{1}{4})$ such that the following holds: for any fixed $\lambda \geq \lambda_0$, $N \geq N_2$, $r \in \left(0, \frac{1}{16}\right)$ and $(y, s) \in \Omega_T$ with

$$
|\mathcal{C} \cap Q_n^\lambda(y, s)| \geq \varepsilon \left| Q_n^\lambda \right|
$$

we have

$$
K_{r}^\lambda(y, s) \subset \mathcal{D}.
$$

Proof. We argue by a contradiction. Assume $K_{r}^\lambda(y, s) \not\subset \mathcal{D}$. Then there is a point $(\tilde{x}, \tilde{t}) \in K_{r}^\lambda(y, s)$ such that for all $\rho > 0$,

$$
\frac{1}{|Q_{\rho}^n|} \int_{K_{\rho}^\lambda(\tilde{x}, \tilde{t})} |Du|^{\theta} \, dx \, dt \leq \lambda^\theta \quad \text{and} \quad \frac{|\mu(K_{\rho}^\lambda(\tilde{x}, \tilde{t}))|}{\rho^{n+1}} \leq \delta \lambda.
$$

We split the proof into an interior case $(Q_{32\rho}^\lambda(\tilde{x}, \tilde{t}) \subset \Omega_T)$ and a boundary case $(Q_{32\rho}^\lambda(\tilde{x}, \tilde{t}) \not\subset \Omega_T)$. In this proof, we only consider the boundary case. We can similarly prove the interior case by using Proposition 3.10 instead of Proposition 3.8. Note that $[K_{\rho}^\lambda(y, s)] := K_{\rho}^\lambda(y, s) \cup \partial_{\rho} K_{\rho}^\lambda(y, s) \subset K_{\rho+\varepsilon}(\tilde{x}, \tilde{t})$ for any $\rho > r$. Then (1.10) and (4.9) yield

$$
\int_{K_{\rho}^\lambda(y, s)} |Du|^{\theta} \, dx \, dt \leq \frac{c}{|Q_{\rho}^n|} \int_{K_{\rho}^\lambda(y, s)} |Du|^{\theta} \, dx \, dt \\
\leq \frac{c}{|Q_{\rho+\varepsilon}^n|} \int_{K_{\rho+\varepsilon}^\lambda(\tilde{x}, \tilde{t})} |Du|^{\theta} \, dx \, dt \leq c2\lambda^\theta
$$

whenever $\rho \geq 32r$. Since $u_k$ is the truncation of $u$, we have

$$
\int_{K_{\rho}^\lambda(y, s)} |Du_k|^{\theta} \, dx \, dt \leq \int_{K_{\rho}^\lambda(y, s)} |Du|^{\theta} \, dx \, dt \leq c2\lambda^\theta
$$

for any $k \in \mathbb{N}$. Combining (4.1) and (4.9), we also deduce

$$
\frac{|\mu_k(K_{\rho}^\lambda(y, s))|}{\rho^{n+1}} \leq 2|\mu|\left(\frac{|K_{\rho}^\lambda(y, s)|}{\rho^{n+1}}\right) \leq c2|\mu|\left(\frac{|K_{\rho+\varepsilon}^\lambda(\tilde{x}, \tilde{t})|}{(\rho + r)^{n+1}}\right) \leq c2\delta \lambda
$$

whenever $\rho \geq 32r$ and $k$ is sufficiently large. We apply Proposition 3.8 with $u, \mu, (x_0, t_0), \lambda, r$ and $\varepsilon$ replaced by $u_k, \mu_k, (y, s), c2\lambda, 4r$ and $\eta$, respectively. Then for any $\eta \in (0, 1)$, there is a small constant $\delta = \delta(n, \Lambda_0, \Lambda_1, p, \theta, \eta) > 0$ such that the following holds: if $u_k$ is a weak solution of (1.7) satisfying (4.10) and (4.11), then there exists a corresponding weak solution $\tilde{u}_k$ of (3.28) such that

$$
\int_{K_{\rho}^\lambda(y, s)} |Du_k - D\tilde{u}_k|^{\theta} \, dx \, dt \leq \eta(c2\lambda)^\theta
$$
and
\begin{equation}
\|D\tilde{v}_k\|_{L^\infty(K_{\varepsilon r}^\lambda(y,s))} \leq c c_2 \lambda =: c_3 \lambda
\end{equation}
for some constant $c_3 = c_3(n, A_0, A_1, p, \theta) \geq 1$. Using Lemma 4.1 and the absolute continuity of the Lebesgue integral, we find
\[
\int_{K_{\varepsilon r}^\lambda(y,s)} |Du - Du_k|^\theta \, dxdt = \int_{K_{\varepsilon r}^\lambda(y,s)} \chi_{\{|Du| > k\}} |Du|^\theta \, dxdt \leq \eta(c_2 \lambda)^\theta
\]
for sufficiently large $k$. Combining this inequality and (4.12), we discover
\begin{equation}
\int_{K_{\varepsilon r}^\lambda(y,s)} |Du - D\tilde{v}_k|^\theta \, dxdt \leq c_4 \eta \lambda^\theta.
\end{equation}

We next show that
\begin{equation}
\left\{(x,t) \in K_{\varepsilon r}^\lambda(y,s) : M_{\Omega r^2}^\lambda |Du|^\theta > (N\lambda)^\theta \right\}
\subset \left\{(x,t) \in K_{\varepsilon r}^\lambda(y,s) : M_{K_{\varepsilon r}^\lambda(y,s)}^\lambda |Du - D\tilde{v}_k|^\theta > \lambda^\theta \right\}
\end{equation}
provided $N \geq N_2 := \max \{2^{n+2}, 1 + c_3^\theta\}$. To do so, let
\[(\tilde{y}, \tilde{s}) \in \left\{(x,t) \in K_{\varepsilon r}^\lambda(y,s) : M_{K_{\varepsilon r}^\lambda(y,s)}^\lambda |Du - D\tilde{v}_k|^\theta \leq \lambda^\theta \right\}.
\]
Then for any $\tilde{r} > 0$,
\begin{equation}
\frac{1}{|Q_{\tilde{r}}^\lambda|} \int_{K_{\tilde{r}}^\lambda(\tilde{y},\tilde{s}) \cap K_{\varepsilon r}^\lambda(y,s)} |Du - D\tilde{v}_k|^\theta \, dxdt \leq \lambda^\theta.
\end{equation}

If $\tilde{r} \in (0, 2r]$, then $K_{\tilde{r}}^\lambda(\tilde{y}, \tilde{s}) \subset K_{\varepsilon r}^\lambda(y,s)$. It follows from (4.16) and (4.13) that
\[
\frac{1}{|Q_{\tilde{r}}^\lambda|} \int_{K_{\tilde{r}}^\lambda(\tilde{y},\tilde{s})} |Du|^\theta \, dxdt \leq \frac{1}{|Q_{\tilde{r}}^\lambda|} \int_{K_{\tilde{r}}^\lambda(\tilde{y},\tilde{s})} (|Du - D\tilde{v}_k|^\theta + |D\tilde{v}_k|^\theta) \, dxdt \\
\leq (1 + c_3^\theta) \lambda^\theta.
\]

If $\tilde{r} > 2r$, then $K_{\tilde{r}}^\lambda(\tilde{y}, \tilde{s}) \subset K_{2\tilde{r}}^\lambda(\tilde{x}, \tilde{t})$. We use the first inequality of (4.9) to obtain
\[
\frac{1}{|Q_{\tilde{r}}^\lambda|} \int_{K_{\tilde{r}}^\lambda(\tilde{y},\tilde{s})} |Du|^\theta \, dxdt \leq \frac{1}{|Q_{\tilde{r}}^\lambda|} \int_{K_{2\tilde{r}}^\lambda(\tilde{x},\tilde{t})} |Du|^\theta \, dxdt \leq 2^{n+2} \lambda^\theta.
\]
Recalling $N_2 = \max \{2^{n+2}, 1 + c_3^\theta\}$, we obtain
\[(\tilde{y}, \tilde{s}) \in \left\{(x,t) \in K_{\varepsilon r}^\lambda(y,s) : M_{\Omega r^2}^\lambda |Du|^\theta \leq (N\lambda)^\theta \right\},
\]
which implies (4.15).

Finally, we compute from (4.15), (2.2) and (4.14) that
\[
\left| \left\{(x,t) \in K_{\varepsilon r}^\lambda(y,s) : M_{\Omega r^2}^\lambda |Du|^\theta > (N\lambda)^\theta \right\} \right|
\leq \left| \left\{(x,t) \in K_{\varepsilon r}^\lambda(y,s) : M_{K_{\varepsilon r}^\lambda(y,s)}^\lambda |Du - D\tilde{v}_k|^\theta > \lambda^\theta \right\} \right|
\leq \frac{c}{\lambda^\theta} \int_{K_{\varepsilon r}^\lambda(y,s)} |Du - D\tilde{v}_k|^\theta \, dxdt \leq cc_4 \eta |Q_y^\lambda| < \varepsilon |Q_y^\lambda|,
\]
by selecting $\eta$ small enough. This is a contradiction to (4.8), which completes the proof. \qed
Taking $N = \max \{N_1, N_2\}$ from Lemmas 4.4 and 4.5, we can apply Lemma 4.2 to discover
\[
\left| \left\{ (x,t) \in \Omega_T : \mathcal{M}_{\Omega_T}^{\alpha}\|Du\|^{\theta} > (N\lambda)^{\theta} \right\} \right|
\]
(4.17)
\[
\leq \varepsilon_0 \left| \left\{ (x,t) \in \Omega_T : \mathcal{M}_{\Omega_T}^{\alpha}\|Du\|^{\theta} > \lambda^{\theta} \right\} \right|
\]
\[
+ \varepsilon_0 \left| \left\{ (x,t) \in \Omega_T : \mathcal{M}_1^{\frac{1}{d}}(\mu) > \delta\lambda \right\} \right|
\]
where $\varepsilon_0 := \left(\frac{s_0}{T}\right)^{n+2}$.

In the following two lemmas, we investigate the precise rate of the upper level sets of the (standard) fractional maximal function (1.12) and the intrinsic one (4.7) (see [1, 40, 51, 64] for more information about the fractional maximal function).

**Lemma 4.6** (See [16, Lemmas 5.4 and 5.7]). Let $\frac{2n}{n+1} < p \leq 2$ and let $\lambda \geq \lambda_0$. Then we have
\[
\left\{ (x,t) \in \Omega_T : \mathcal{M}_1^{\frac{1}{d}}(\mu) > \delta\lambda \right\} \subset \left\{ (x,t) \in \Omega_T : \left[\mathcal{M}_1^{\frac{1}{d}}(\mu)\right]^{d} > \delta^d\lambda \right\},
\]
where $\Omega_T := \Omega \times (-T,T)$.

**Lemma 4.7** (See [16, Lemma 5.5]). Let $p > 1$ and let $\lambda \geq 1$. Suppose that $\mu = \mu_0 \otimes f$, where $\mu_0$ is a finite signed Radon measure in $\Omega$ and $f$ is a Lebesgue function in $(-\infty,T)$. Then we have
\[
\left\{ (x,t) \in \Omega_T : \mathcal{M}_1^{\frac{1}{d}}(\mu) > \delta\lambda \right\} \subset \left\{ (x,t) \in \Omega_T : \left[2\left(\mathcal{M}_1(\mu_0)(\mathcal{M}f)\right)\right]^{\frac{1}{d-1}} > \delta^{\frac{1}{d-1}}\lambda \right\},
\]
where $\mathcal{M}_1(\mu_0)$ is given by (1.16), and the maximal function $\mathcal{M}f$ is defined by
\[
\mathcal{M}f(t) := \sup_{r > 0} \int_{t-r}^{t+r} |f(s)| \, ds = \sup_{r > 0} \frac{1}{2r} \int_{t-r}^{t+r} |f(s)| \, ds.
\]

**Remark 4.8.** We note that if $\mu = \mu_0 \otimes f$, we can drop the condition $\lambda_0 \geq \left[\frac{|\mu_0(\Omega_T)|}{\delta^d\|f\|_{L^1(0,T)}}\right]^d$ in (4.5); thus we instead take
\[
\lambda_0 := \left[\frac{|\mu_0(\Omega)|\|f\|_{L^1(0,T)}}{|\Omega_T|^\frac{1}{d-1}}\right]^\frac{1}{d-1} \frac{|\Omega_T|}{\varepsilon|Q_{R/10}|} + 1,
\]
see [16, Remark 5.8] for details.

Finally, combining (4.17) and Lemmas 4.6–4.7, we directly obtain the following decay estimate:

**Proposition 4.9.** Let $N = \max \{N_1, N_2\}$ from Lemmas 4.4 and 4.5. Then for any $\varepsilon \in (0,1)$, there exists $\delta = \delta(n,\lambda_0,\Lambda_1,p,\theta,\varepsilon)$ such that if $(a,\Omega)$ is $(\delta,R)$-vanishing for some $R > 0$, then for any renormalized solution $u$ of (1.1) and any $\lambda \geq \lambda_0$, we have
\[
\left| \left\{ (x,t) \in \Omega_T : \mathcal{M}_{\Omega_T}^{\alpha}\|Du\|^{\theta} > (N\lambda)^{\theta} \right\} \right|
\]
\[
\leq \varepsilon_0 \left| \left\{ (x,t) \in \Omega_T : \mathcal{M}_{\Omega_T}^{\alpha}\|Du\|^{\theta} > \lambda^{\theta} \right\} \right|
\]
\[
+ \varepsilon_0 \left| \left\{ (x,t) \in \Omega_T : \left[\mathcal{M}_1^{\frac{1}{d}}(\mu)\right]^{d} > \delta^d\lambda \right\} \right|.
\]
Furthermore, if $\mu = \mu_0 \otimes f$, then we have
\[
\left| \{(x,t) \in \Omega_T : M_{\Omega, n} |Du|^\theta > (N\lambda)^\theta \} \right|
\leq \varepsilon_0 \left| \{(x,t) \in \Omega_T : M_{\Omega, n} |Du|^\theta > \lambda^\theta \} \right|
+ \varepsilon_0 \left| \{(x,t) \in \Omega_T : |2(M_1(\mu_0))(Mf)|^{\frac{1}{p}} > \delta^{\frac{1}{p}} \lambda \} \right|.
\]

5. Proof of global Calderón-Zygmund type estimates

In this section, we give the proof of Theorems 1.6 and 1.9. Throughout this section, we assume that $(\mathbf{a}, \Omega)$ is $(\delta, R)$-vanishing and that $p$ satisfies (1.4). We fix any $\varepsilon \in (0,1)$ and $\theta = \theta(n,p)$ satisfying (4.6) and $\beta_0 > 1$, where $\beta := \frac{n+2}{(n+1)p-n}$. Since $\beta \not\rightarrow \frac{n+2}{n+1}$ as $\theta \not\rightarrow p - \frac{n}{n+1}$, we can choose $\theta$ such that $\beta \theta > 1$. Let $N = N(n, \Lambda_0, \Lambda_1, p, \theta) > 1$ be a given constant in Proposition 4.9 above. We assume that a renormalized solution $u$ of (1.1) is defined in $\Omega_T := \Omega \times (-\infty, T)$, upon letting $u \equiv 0$ for $t < 0$ and $\mu \equiv 0$ for $\mathbb{R}^{n+1} \setminus \Omega_T$. If $\mu = \mu_0 \otimes f$, where $\mu_0$ is a finite signed Radon measure on $\Omega$ and $f \in L^s(0,T)$ for some $s \geq 1$, then we let both $\mu_0$ and $f$ be 0 outside $\Omega$ and $(0,T)$, respectively.

We introduce the following two decay estimates of integral type:

**Lemma 5.1.** Let $\lambda_0 \geq 1$ be a given constant in (4.5). If $u$ is a renormalized solution of (1.1), then for any $\theta_0 \in \left(\theta, p - \frac{n}{n+1}\right)$ and any $\lambda \geq \lambda_0$, there exists a constant $c = c(n, \Lambda_0, \Lambda_1, p, \theta_0) \geq 1$ such that
\[
\int_{\{(x,t) \in \Omega_T : |Du| > N\lambda\}} |Du|^{\theta_0} \, dx \, dt
\leq c\varepsilon \int_{\{(x,t) \in \Omega_T : |Du| > \frac{\lambda}{2}\}} |Du|^{\theta_0} \, dx \, dt
+ \frac{c\varepsilon}{\delta^{\theta_0}} \int_{\{(x,t) \in \Omega_T : |M_1(\mu)|^{d} > \delta^d \lambda\}} [M_1(\mu)]^{d\theta_0} \, dx \, dt,
\]
where $d := \frac{2}{(n+1)p-2n}$, $\Omega_T := \Omega \times (-T, T)$, and $M_1(\mu)$ is given in (1.12).

**Proof.** The proof is almost like that of [16, Lemma 6.1]. \qed

**Lemma 5.2.** Let $\lambda_0 \geq 1$ be a given constant in (4.19). If $u$ is a renormalized solution of (1.1) and $\mu = \mu_0 \otimes f$, then for any $\theta_0 \in \left(\theta, p - \frac{n}{n+1}\right)$ and any $\lambda \geq \lambda_0$, there exists a constant $c = c(n, \Lambda_0, \Lambda_1, p, \theta_0) \geq 1$ such that
\[
\int_{\{(x,t) \in \Omega_T : |Du| > N\lambda\}} |Du|^{\theta_0} \, dx \, dt
\leq c\varepsilon \int_{\{(x,t) \in \Omega_T : |Du| > \frac{\lambda}{2}\}} |Du|^{\theta_0} \, dx \, dt
+ \frac{c\varepsilon}{\delta^{\theta_0}} \int_{\{(x,t) \in \Omega_T : |2(M_1(\mu_0))(Mf)|^{\frac{1}{p}} > \delta^\frac{1}{p} \lambda\}} [(M_1(\mu_0))(Mf)]^{\frac{\theta_0}{p}} \, dx \, dt,
\]
where $M_1(\mu_0)$ and $Mf$ are given in (1.16) and (4.18), respectively.

**Proof.** The proof is almost like that of [16, Lemma 6.2]. \qed
5.1. Proof of Theorem 1.6.

Proof of Theorem 1.6. If $0 < q \leq \theta$, then we have from Lemma 4.1 that

\begin{equation}
\int_{\Omega_T} |Du|^q \, dx \, dt \leq c \, |[\mu](\Omega_T)|^{\beta q} .
\end{equation}

Noting that

\begin{equation}
\beta := \frac{n+2}{(n+1)p-n} < \frac{2}{(n+1)p-2n} =: d \quad \text{for all} \quad \frac{2n}{n+1} < p \leq 2 - \frac{1}{n+1}
\end{equation}

and $\beta_0 := \min\{1, \beta q\}$, we have

\begin{equation}
|[\mu](\Omega_T)|^{\beta q} = |[\mu](\Omega_T)|^{\beta_0} \, |[\mu](\Omega_T)|^{\beta q - \beta_0}
\leq c |[\mu](\Omega_T)|^{\beta_0} \left(|[\mu](\Omega_T)|^{dq - \beta_0} + 1\right)
= c \left(|[\mu](\Omega_T)|^{dq} + |[\mu](\Omega_T)|^{\beta_0}\right).
\end{equation}

Combining (5.1), (5.3) and (1.13) with $\alpha = dq$, we obtain the desired estimates (1.11) for the case $0 < q \leq \theta$.

Now, let us assume $q > \theta$ and fix a constant $\theta_0 = \theta_0(n, p, q)$ arbitrarily so that $\theta < \theta_0 < \min\left\{p - \frac{n}{n+1}, q\right\}$. Recalling the truncation operator (1.6) and Lemma 2.2, we have for any $k > N\lambda_0$,

\begin{equation}
\int_{\Omega_T} T_k(|Du|)^{q-\theta_0} |Du|^\theta_0 \, dx \, dt
= (q - \theta_0) N^{q-\theta_0} \int_0^{\frac{1}{\lambda_0}} \lambda^{q-\theta_0 - 1} \left[\int_{\{(x,t) \in \Omega_T : |Du| > N\lambda\}} |Du|^\theta_0 \, dx \, dt\right] d\lambda
\leq c \int_0^{\frac{1}{\lambda_0}} \lambda^{q-\theta_0 - 1} \left[\int_{\{(x,t) \in \Omega_T : |Du| > N\lambda\}} |Du|^\theta_0 \, dx \, dt\right] d\lambda
=: P_1
+ c \int_0^{\frac{1}{\lambda_0}} \lambda^{q-\theta_0 - 1} \left[\int_{\{(x,t) \in \Omega_T : |Du| > N\lambda\}} |Du|^\theta_0 \, dx \, dt\right] d\lambda
=: P_2
\end{equation}

for some constant $c = c(n, \Lambda_0, \Lambda_1, p, q) \geq 1$, where $\lambda_0$ is given in (4.5). It follows from Lemma 4.1 that

\begin{equation}
P_1 \leq \int_0^{\lambda_0} \lambda^{q-\theta_0 - 1} d\lambda \int_{\Omega_T} |Du|^\theta_0 \, dx \, dt \leq c\lambda_0^{q-\theta_0} |[\mu](\Omega_T)|^{\beta \theta_0}
\end{equation}
for some constant \( c = c(n, \Lambda_0, p, q, \Omega_T) \geq 1 \). To estimate \( P_2 \), we apply Lemma 5.1 and Fubini’s theorem to obtain

\[
P_2 \leq c \varepsilon \int_{\lambda_0}^{\lambda} \lambda^{q-\theta_0-1} \left[ \int_{\{(x,t) \in \Omega \times [0,T] : |Du| > \lambda \}} |Du|^\theta_0 \, dx \, dt \right] \, d\lambda
\]

\[+ \frac{c \varepsilon}{\delta^{\theta_0}} \int_{\lambda_0}^{\lambda} \lambda^{q-\theta_0-1} \left[ \int_{\{(x,t) \in \Omega_T : |M_1(\mu)| > \delta \lambda \}} |M_1(\mu)|^\theta_0 \, dx \, dt \right] \, d\lambda
\]

\[
\leq c \varepsilon \int_{\Omega_T} T_k(|Du|)^{q-\theta_0} |Du|^\theta_0 \, dx \, dt + \frac{c \varepsilon}{\delta^{\theta_0}} \int_{\Omega_T} |M_1(\mu)|^d \, dx \, dt
\]

for some constant \( c = c(n, \Lambda_0, \Lambda_1, p, q) \geq 1 \). Inserting these inequalities into (5.4), we discover

\[
\int_{\Omega_T} T_k(|Du|)^{q-\theta_0} |Du|^\theta_0 \, dx \, dt \leq c_0 \varepsilon \int_{\Omega_T} T_k(|Du|)^{q-\theta_0} |Du|^\theta_0 \, dx \, dt
\]

\[+ \frac{c \varepsilon}{\delta^{\theta_0}} \int_{\Omega_T} |M_1(\mu)|^d \, dx \, dt + c \lambda_0^{q-\theta_0} |\mu|^{\beta \theta_0}
\]

for some \( c_0 = c_0(n, \Lambda_0, \Lambda_1, p, q) \geq 1 \). Now we select \( \varepsilon > 0 \) with \( c_0 \varepsilon < 1 \), and then a corresponding \( \delta = \delta(n, \Lambda_0, \Lambda_1, p, q) > 0 \) is determined. Letting \( k \to \infty \), we obtain

\[
\int_{\Omega_T} |Du|^q \, dx \, dt \leq c \int_{\Omega_T} |M_1(\mu)|^d \, dx \, dt + c \lambda_0^{q-\theta_0} |\mu|^{\beta \theta_0}.
\]

Moreover, it follows from (4.5), (4.6) and (5.2) that

\[
\lambda_0 \leq c \left( |\mu|(|\Omega_T|)^d + 1 \right)
\]

and

\[
|\mu|(|\Omega_T|)^{\beta \theta_0} \leq |\mu|(|\Omega_T|)^{\beta \theta_0} \left( |\mu|(|\Omega_T|)^d + 1 \right)^{\beta \theta_0 - \theta_0}
\]

\[
\leq |\mu|(|\Omega_T|)^{\beta \theta_0} \left( |\mu|(|\Omega_T|)^d + 1 \right)^{\theta_0 - \beta \theta_0},
\]

where we used the fact that \( \beta \theta_0 > \beta \theta > 1 \geq \beta_0 \). Then the above two inequalities yield

\[
\lambda_0^{q-\theta_0} |\mu|(|\Omega_T|)^{\beta \theta_0} \leq c \left( |\mu|(|\Omega_T|)^d + |\mu|(|\Omega_T|)^{\beta \theta_0} \right)
\]

for some constant \( c = c(n, \Lambda_0, \Lambda_1, p, q, R, \Omega_T) \geq 1 \). On the other hand, for each \( (x, t) \in \Omega_T \) we know \( M_1(\mu)(x, t) \leq M_1(\mu)(x, T) \), which implies

\[
\int_{\Omega_T} [M_1(\mu)]^d \, dx \, dt \leq c \int_{\Omega_T} [M_1(\mu)]^d \, dx \, dt.
\]

Combining (1.13) and (5.5)–(5.7), we finally obtain the desired estimate (1.11) for \( q > \theta \). This completes the proof.
5.2. Proof of Theorem 1.9.

Proof of Theorem 1.9. Let \( p-1 < q \leq \theta \). Since \( \beta(p-1) < 1 \) and \( \beta_0 := \min \{1, \beta q\} \), we have

\[
\|\mu(\Omega_T)\|^{\beta q} = \|\mu(\Omega_T)\|^\beta_0 \|\mu(\Omega_T)\|^{\beta q - \beta_0}
\]

\[
\leq c \|\mu(\Omega_T)\|^\beta_0 \left(\|\mu(\Omega_T)\|^{\frac{1}{q^*} - \beta_0} + 1\right)
\]

\[
= c \left(\|\mu(\Omega_T)\|^{\frac{1}{q^*}} + \|\mu(\Omega_T)\|^{\beta_0}\right).
\]

(5.8)

It follows from (5.1), (5.8) and (1.13) with \( \alpha = \frac{q}{p-1} \) that

\[
\int_{\Omega_T} |Du|^q \, dx \, dt \leq c \left\{ \int_{\Omega_T} [(M_1(\mu_0))(Mf)]^{\frac{1}{q^*}} \, dx \, dt + \|\mu(\Omega_T)\|^{\beta_0} \right\}.
\]

Here we used the fact that \( M_1(\mu) \leq (M_1(\mu_0))(Mf) \), where \( M_1(\mu_0) \) and \( Mf \) are given in (1.16) and (4.18), respectively. Applying the strong \( \left(\frac{q}{p-1}, \frac{q}{p-1}\right) \)-estimate for the function \( f \) (see for instance [69, Chapter I, Theorem 1]), we obtain the desired estimates (1.15) for the case \( p-1 < q \leq \theta \).

Let \( q > \theta \) and let \( \theta_0 = \theta_0(n, p, q) \) be an arbitrary constant with \( \theta < \theta_0 < \min \left\{ p - \frac{\alpha}{n+1}, q \right\} \). Proceeding as in the proof of Theorem 1.6 and using Lemma 5.2, we infer

\[
\int_{\Omega_T} T_k(|Du|^{q-\theta_0})|Du|^\theta_0 \, dx \, dt \leq c \lambda_0^{q-\theta_0} \left[ \|\mu_0(\Omega)\|f\|\|_{L^1(\Omega, T)} \right]^{\beta_0} + cS,
\]

(5.9)

where

\[
S := \int_{\lambda_0}^{\lambda_0^{q-\theta_0-1}} \left[ \int_{\Omega_T \cap \{\|2(M_1(\mu_0))(Mf)\|^{\frac{1}{q^*}} > \delta \}} [(M_1(\mu_0))(Mf)]^{\frac{1}{q^*}} \, dx \, dt \right] \, d\lambda.
\]

From Fubini’s theorem and the strong \( \left(\frac{q}{p-1}, \frac{q}{p-1}\right) \)-estimate, we have

\[
S \leq c \int_{\Omega} |M_1(\mu_0)|^{\frac{1}{q^*}} \, dx \int_0^T |f(t)|^{\frac{1}{q^*}} \, dt,
\]

(5.10)

where we used the fact that \( f \equiv 0 \) for \( t \leq 0 \). On the other hand, it follows from (4.19) and (4.6) that

\[
\lambda_0 \leq c \left[ \left(\|\mu_0(\Omega)\|f\|\|_{L^1(\Omega, T)}^{\beta_0} + 1\right) \leq c \left(\|\mu_0(\Omega)\|f\|\|_{L^1(\Omega, T)}\right) + 1\right]
\]

and

\[
\|\mu(\Omega_T)\|^{\beta_0} \leq \|\mu(\Omega_T)\|^{\beta_0} \left(\|\mu(\Omega_T)\|^{\frac{1}{q^*}} + 1\right)^{(p-1)(\beta \theta_0 - \beta_0)}
\]

\[
\leq \|\mu(\Omega_T)\|^{\beta_0} \left(\|\mu(\Omega_T)\|^{\frac{1}{q^*}} + 1\right)^{\theta_0 - \theta_0(p-1)},
\]

where we used the fact that \( \theta_0 > \beta \theta > 1 \geq \beta_0 \) and \( \beta(p-1) < 1 \). Since \( \|\mu(\Omega_T)\| = \|\mu_0(\Omega)\|f\|\|_{L^1(\Omega, T)} \), the above two inequalities imply

\[
\lambda_0^{-\theta_0} \left[ \|\mu_0(\Omega)f\|\|\|_{L^1(\Omega, T)} \right]^{\beta_0}
\]

\[
\leq c \left(\|\mu_0(\Omega)f\|\|\|_{L^1(\Omega, T)}\right) + \left[ \|\mu_0(\Omega)f\|\|\|_{L^1(\Omega, T)}\right]^{\beta_0},
\]

(5.11)
Inserting (5.10) and (5.11) into (5.9) and using (1.17), we discover
\[
\int_{\Omega_T} T_k(|Du|^{q-\theta_0}|Du|^\theta_0) \, dx \, dt \leq c \int_{\Omega_T} \left[ (\mathcal{M}_1(\mu_0)f)^{\frac{q}{p-1}} \right] \, dx \, dt + c \left[ |\mu_0|_L^1(\Omega) \right]^{\beta_0}
\]
Letting \( k \to \infty \), we finally obtain the desired estimate (1.15) for \( q > \theta \). This completes the proof. \( \square \)

ACKNOWLEDGMENTS

The authors thank the anonymous referee for the valuable comments, which improved the exposition and the accuracy of the article.

REFERENCES

[1] D. R. Adams and L. I. Hedberg, Function spaces and potential theory, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 314, Springer-Verlag, Berlin, 1996.
[2] B. Avelin, T. Kuusi, and M. Parviainen, Variational parabolic capacity, Discrete Contin. Dyn. Syst. 35 (2015), no. 12, 5665–5688.
[3] P. Baroni, Marcinikiewicz estimates for degenerate parabolic equations with measure data, J. Funct. Anal. 267 (2014), no. 9, 3397–3426.
[4] , Singular parabolic equations, measures satisfying density conditions, and gradient integrability, Nonlinear Anal. 153 (2017), 89–116.
[5] P. Bénilan, L. Boccardo, T. Gallouët, R. Gariepy, M. Pierre, and J. L. Vázquez, An \( L^1 \)-theory of existence and uniqueness of solutions of nonlinear elliptic equations, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 22 (1995), no. 2, 241–273.
[6] M.-F. Bidaut-Véron and Q.-H. Nguyen, Stability properties for quasilinear parabolic equations with measure data, J. Eur. Math. Soc. (JEMS) 17 (2015), no. 9, 2103–2135.
[7] D. Blanchard and F. Murat, Renormalised solutions of nonlinear parabolic problems with \( L^1 \) data: existence and uniqueness, Proc. Roy. Soc. Edinburgh Sect. A 127 (1997), no. 6, 1137–1152.
[8] L. Boccardo, A. Dall’Aglio, T. Gallouët, and L. Orsina, Nonlinear parabolic equations with measure data, J. Funct. Anal. 147 (1997), no. 1, 237–258.
[9] L. Boccardo and T. Gallouët, Nonlinear elliptic and parabolic equations involving measure data, J. Funct. Anal. 87 (1989), no. 1, 149–169.
[10] , Nonlinear elliptic equations with right-hand side measures, Comm. Partial Differential Equations 17 (1992), no. 3-4, 641–655.
[11] L. Boccardo, T. Gallouët, and L. Orsina, Existence and uniqueness of entropy solutions for nonlinear elliptic equations with measure data, Ann. Inst. H. Poincaré Anal. Non Linéaire 13 (1996), no. 5, 539–551.
[12] V. Bögelein and M. Parviainen, Self-improving property of nonlinear higher order parabolic systems near the boundary, NoDEA Nonlinear Differential Equations Appl. 17 (2010), no. 1, 21–54.
[13] T. A. Bui and X. T. Duong, Global Marcinikiewicz estimates for nonlinear parabolic equations with nonsmooth coefficients, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 18 (2018), no. 3, 881–916.
[14] 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.
[15] S.-S. Byun, D. K. Palagachev, and P. Shin, Optimal regularity estimates for general nonlinear parabolic equations, Manuscripta Math. 162 (2020), no. 1-2, 67–98.
[16] S.-S. Byun, J.-T. Park, and P. Shin, Global regularity for degenerate/singular parabolic equations involving measure data, Calc. Var. Partial Differential Equations 60 (2021), no. 1, Paper No. 18, 32 pp.
[17] S.-S. Byun and L. Wang, Elliptic equations with \( BMO \) coefficients in Reifenberg domains, Comm. Pure Appl. Math. 57 (2004), no. 10, 1283–1310.
[18] L. A. Caffarelli and I. Peral, \( W^{1,p} \) estimates for elliptic equations in divergence form, Comm. Pure Appl. Math. 51 (1998), no. 1, 1–21.
[19] E. Casas, Control of an elliptic problem with pointwise state constraints, SIAM J. Control Optim. 24 (1986), no. 6, 1309–1318.

[20] ______, Boundary control of semilinear elliptic equations with pointwise state constraints, SIAM J. Control Optim. 31 (1993), no. 4, 993–1006.

[21] E. Casas, J. C. de los Reyes, and F. Tröltzsch, Sufficient second-order optimality conditions for semilinear control problems with pointwise state constraints, SIAM J. Optim. 19 (2008), no. 2, 616–643.

[22] G. Dal Maso, F. Murat, L. Orsina, and A. Prignet, Renormalized solutions of elliptic equations with general measure data, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 28 (1999), no. 4, 741–808.

[23] A. Dall’Aglio, Approximated solutions of equations with $L^1$ data. Application to the $H$-convergence of quasi-linear parabolic equations, Ann. Mat. Pura Appl. (4) 170 (1996), 207–240.

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

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

[26] ______, Addendum to: “Hölder estimates for nonlinear degenerate parabolic systems”, J. Reine Angew. Math. 363 (1985), 217–220.

[27] R. J. DiPerna and P.-L. Lions, On the Cauchy problem for Boltzmann equations: global existence and weak stability, Ann. of Math. (2) 130 (1989), no. 2, 321–366.

[28] ______, Ordinary differential equations, transport theory and Sobolev spaces, Invent. Math. 98 (1989), no. 3, 511–547.

[29] J. Kinnunen, R. Korte, T. Kuusi, and M. Parviainen, Nonlinear parabolic capacity and polar sets of superparabolic functions, Math. Ann. 355 (2013), no. 4, 1349–1381.

[30] J. Kinnunen and J. L. Lewis, Higher integrability for parabolic systems of $p$-Laplacian type, Duke Math. J. 102 (2000), no. 2, 253–271.

[31] J. Kinnunen, T. Lukkari, and M. Parviainen, An existence result for superparabolic functions, J. Funct. Anal. 258 (2010), no. 3, 713–728.

[32] ______, Local approximation of superharmonic and superparabolic functions in nonlinear potential theory, J. Fixed Point Theory Appl. 13 (2013), no. 1, 291–307.

[33] J. Kinnunen and E. Saksman, Regularity of the fractional maximal function, J. Fixed Point Theory Appl. 13 (2013), no. 1, 291–307.

[34] T. Klimsiak and A. Rozkosz, On the structure of diffuse measures for parabolic capacities, C. R. Math. Acad. Sci. Paris 357 (2019), no. 5, 443–449.

[35] 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.

[36] ______, Riesz potentials and nonlinear parabolic equations, Arch. Ration. Mech. Anal. 212 (2014), no. 3, 727–780.

[37] The Wolff gradient bound for degenerate parabolic equations, J. Eur. Math. Soc. (JEMS) 16 (2014), no. 4, 835–892.

[38] ______, Vectorial nonlinear potential theory, J. Eur. Math. Soc. (JEMS) 20 (2018), no. 4, 929–1004.
SINGULAR PARABOLIC MEASURE DATA PROBLEMS

[46] A. Lemenant, E. Milakis, and L. V. Spinolo, On the extension property of Reifenberg-flat domains, Ann. Acad. Sci. Fenn. Math. 39 (2014), no. 1, 51–71.
[47] R. J. LeVeque and Z. Li, The immersed interface method for elliptic equations with discontinuous coefficients and singular sources, SIAM J. Numer. Anal. 31 (1994), no. 4, 1019–1044.
[48] G. M. Lieberman, Boundary and initial regularity for solutions of degenerate parabolic equations, Nonlinear Anal. 20 (1993), no. 5, 551–569.
[49] C. Meyer, L. Panizzi, and A. Schiela, Uniqueness criteria for the adjoint equation in state-constrained elliptic optimal control, Numer. Funct. Anal. Optim. 32 (2011), no. 9, 983–1007.
[50] G. Mingione, The Calderón-Zygmund theory for elliptic problems with measure data, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 6 (2007), no. 2, 195–261.
[51] ______, Gradient estimates below the duality exponent, Math. Ann. 346 (2010), no. 3, 571–627.
[52] Q.-H. Nguyen, Potential estimates and quasilinear parabolic equations with measure data, Mem. Amer. Math. Soc., to appear. arXiv:1405.2587.
[53] Q.-H. Nguyen and N. C. Phuc, Good-λ and Muckenhoupt-Wheeden type bounds in quasilinear measure datum problems, with applications, Math. Ann. 374 (2019), no. 1-2, 67–98.
[54] ______, Pointwise gradient estimates for a class of singular quasilinear equations with measure data, J. Funct. Anal. 278 (2020), no. 5, 108391, 35 pp.
[55] ______, Existence and regularity estimates for quasilinear equations with measure data: the case $1 < p < \frac{2n-2}{n-2}$, Anal. PDE, to appear. arXiv:2003.03725.
[56] S. Osher and R. Fedkiw, Level set methods and dynamic implicit surfaces, Applied Mathematical Sciences, vol. 153, Springer-Verlag, New York, 2003.
[57] C. S. Peskin, Numerical analysis of blood flow in the heart, I. Immersed elastic fibers in a viscous incompressible fluid, J. Comput. Phys. 25 (1977), no. 3, 220–252.
[58] C. S. Peskin and D. M. McQueen, A three-dimensional computational method for blood flow in the heart. I. Immersed elastic fibers in a viscous incompressible fluid, J. Comput. Phys. 81 (1989), no. 2, 372–405.
[59] F. Petitta, Renormalized solutions of nonlinear parabolic equations with general measure data, Ann. Mat. Pura Appl. (4) 187 (2008), no. 4, 563–604.
[60] F. Petitta, A. C. Ponce, and A. Porretta, Diffuse measures and nonlinear parabolic equations, J. Evol. Equ. 11 (2011), no. 4, 861–905.
[61] F. Petitta and A. Porretta, On the notion of renormalized solution to nonlinear parabolic equations with general measure data, J. Elliptic Parabol. Equ. 1 (2015), 201–214.
[62] N. C. Phuc, Global integral gradient bounds for quasilinear equations below or near the natural exponent, Ark. Mat. 52 (2014), no. 2, 329–354.
[63] ______, Morrey global bounds and quasilinear Riccati type equations below the natural exponent, J. Math. Pures Appl. (9) 102 (2014), no. 1, 99–123.
[64] ______, Nonlinear Muckenhoupt-Wheeden type bounds on Reifenberg-flat domains, with applications to quasilinear Riccati type equations, Adv. Math. 250 (2014), 387–419.
[65] M. Pierre, Parabolic capacity and Sobolev spaces, SIAM J. Math. Anal. 14 (1983), no. 3, 522–533.
[66] A. Prignet, Existence and uniqueness of “entropy” solutions of parabolic problems with $L^1$ data, Nonlinear Anal. 28 (1997), no. 12, 1943–1954.
[67] C. Scheven, Elliptic obstacle problems with measure data: potentials and low order regularity, Publ. Mat. 56 (2012), no. 2, 327–374.
[68] ______, Gradient potential estimates in non-linear elliptic obstacle problems with measure data, J. Funct. Anal. 262 (2012), no. 6, 2777–2832.
[69] E. M. Stein, Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals, Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.
[70] M. Sussman, P. Smereka, and S. Osher, A level set approach for computing solutions to incompressible two-phase flow, J. Comput. Phys. 114 (1994), no. 1, 146–159.
[71] T. Toro, Doubling and flatness: geometry of measures, Notices Amer. Math. Soc. 44 (1997), no. 9, 1087–1094.
[72] J. M. Urbano, The method of intrinsic scaling, Lecture Notes in Mathematics, vol. 1930, Springer-Verlag, Berlin, 2008. A systematic approach to regularity for degenerate and singular PDEs.
[73] L. Wang, *A geometric approach to the Calderón-Zygmund estimates*, Acta Math. Sin. (Engl. Ser.) 19 (2003), no. 2, 381–396.

J.-T. Park: Korea Institute for Advanced Study, Seoul 02455, Republic of Korea

Email address: ppark00@kias.re.kr

P. Shin: Department of Mathematics, Kyonggi University, Suwon 16227, Republic of Korea

Email address: shinpilsoo.math@kgu.ac.kr