GLOBAL REGULARITY FOR DEGENERATE/SINGULAR PARABOLIC EQUATIONS INVOLVING MEASURE DATA

SUN-SIG BYUN, JUNG-TAE PARK, AND PILSOO SHIN

Abstract. We consider degenerate and singular parabolic equations with \( p \)-Laplacian structure in bounded nonsmooth domains when the right-hand side is a signed Radon measure with finite total mass. We develop a new tool that allows global regularity estimates for the spatial gradient of solutions to such parabolic measure data problems, by introducing the (intrinsic) fractional maximal function of a given measure.

1. Introduction

We study the following Cauchy-Dirichlet problem for a degenerate/singular parabolic equation:

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

Here \( \Omega_T := \Omega \times (0, T), T > 0 \), is a cylindrical domain with parabolic boundary \( \partial_p \Omega_T := (\partial \Omega \times [0, T]) \cup (\Omega \times \{0\}) \), where \( \Omega \subset \mathbb{R}^n, n \geq 2 \), is a bounded domain with nonsmooth boundary \( \partial \Omega \), and we write \( Du := D_x u \). The right-hand side term \( \mu \) is a signed Radon measure on the domain \( \Omega_T \) 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, 

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

The nonlinear operator \( a = a(\xi, x, t) : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \) is assumed to be measurable in \( x \)- and \( t \)-variables and satisfies the following growth and ellipticity conditions:

\[
\begin{aligned}
  |a(\xi, x, t)| + |\xi||D_\xi a(\xi, x, t)| &\leq \Lambda_1 |\xi|^{p-1}, \\
  \Lambda_0 |\xi|^{p-2}|\eta|^2 &\leq \langle D_\xi a(\xi, x, t)\eta, \eta \rangle,
\end{aligned}
\]

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 \( \Lambda_1 \geq \Lambda_0 > 0 \). Here \( D_\xi a(\xi, x, t) \) is the Jacobian matrix of the operator \( a \) with respect to the variable \( \xi \), and \( \langle \cdot, \cdot \rangle \) is the standard inner product in \( \mathbb{R}^n \times \mathbb{R}^n \).

This type of study is modeled after the parabolic \( p \)-Laplace equation

\[ u_t - \text{div} \left( |Du|^{p-2} Du \right) = \mu, \]
as one can take \( a(\xi, x, t) = |\xi|^{p-2}\xi \) in the problem (1.1). In this paper we assume

\[
p > 2 - \frac{1}{n+1},
\]

which ensures that the spatial gradient of our solution belongs to \( L^1(\Omega_T) \) (see Section 2.2 below for more details). Note that the structure conditions (1.2) imply \( a(0, x, t) = 0 \) for \((x, t) \in \mathbb{R}^n \times \mathbb{R}\) and the following monotonicity condition:

\[
\langle a(\xi_1, x, t) - a(\xi_2, x, t), \xi_1 - \xi_2 \rangle \geq \begin{cases} 
\hat{\Lambda}_0 |\xi_1 - \xi_2|^p & \text{if } p \geq 2, \\
\hat{\Lambda}_0 \left( |\xi_1|^2 + |\xi_2|^2 \right)^{\frac{p-2}{2}} |\xi_1 - \xi_2|^2 & \text{if } 1 < p < 2
\end{cases}
\]

for all \((x, t) \in \mathbb{R}^n \times \mathbb{R}\) and \(\xi_1, \xi_2 \in \mathbb{R}^n\), and for some constant \(\hat{\Lambda}_0 = \hat{\Lambda}_0(n, \Lambda_0, p) > 0\).

Measure data problems including the problem (1.1) arise in a variety of models, primarily in physics and biology. For instance, the incompressible Navier-Stokes equations with measure data describe in biology the flow pattern of blood in the heart [41, 42]. This relates to the design of artificial hearts. Measure data problems appear as well in the study of surface tension forces concentrated on the interfaces of fluids [33, 40, 46] and state-constrained optimal control theory [15–17, 35].

Concerning regularity results for the parabolic problem (1.1), the first contribution has been given by Boccardo and Gallouët [7] (see Section 2.2 below for more detailed explanation). Later this results lead to sharp estimates in terms of Marcinkiewicz spaces, see [3, 11] for \( p \geq 2 \) and [4] for \( 2 - \frac{1}{n+1} < p < 2 \). For pointwise potential estimates for the spatial gradient of solutions, finer regularity results have been proven in [21] for \( p = 2 \), [30, 31] for \( p \geq 2 \), and [28] for \( 2 - \frac{1}{n+1} < p \leq 2 \) (see also survey papers [29, 38] and the references given there for an overview of potential estimates). Lastly, Calderón-Zygmund type estimates, namely the spatial gradient estimates of solutions corresponding to given data, have been obtained in [5, 13, 39] only for \( p = 2 \) in the literature of parabolic measure data problems.

The main goal of this article is to prove global Calderón-Zygmund type estimates for the problem (1.1) under (1.2) and (1.3) (see Theorem 2.4 and 2.8 below). Here our proof covers both the cases of \( p \geq 2 \) and of \( 2 - \frac{1}{n+1} < p \leq 2 \). For the regularity estimates, we need to consider a suitable notion of solution (Definition 2.1) as well as find optimal regularity requirements on the operator \( a \) and the boundary \( \partial \Omega \) (Definition 2.2). We also refer to [36, 37, 44] for the Calderón-Zygmund type estimates to the elliptic case of (1.1).

Before discussing how to prove Theorem 2.4 and 2.8, we introduce an effective and systematic approach to obtain Calderón-Zygmund type estimates for quasilinear elliptic and parabolic problems having divergence form: the so-called maximal function technique. This technique by Caffarelli and Peral [14] uses the standard energy estimate, Hardy-Littlewood maximal function and Calderón-Zygmund decomposition. The advantage of this approach lies in the fact that it can avoid the use of singular integrals and explicit kernels. It works for nonlinear elliptic equations and nondegenerate \( (p = 2) \) parabolic equations (see for instance [13, 37, 39, 44]). However, this method does not work for degenerate/singular \( (p \neq 2) \) parabolic equations. This is not only because of the structure of parabolic \( p \)-Laplace type equations (roughly speaking it scales differently in space and time), but also because of the absence of suitable maximal function operators associated with such parabolic equations.
To overcome these difficulties (as mentioned above) and to derive our main estimates (Theorem 2.4 and 2.8), we define suitable intrinsic (fractional) maximal function operators (see Section 2.1 and 2.4 below for definitions and some properties) and verify the precise relationship between the intrinsic (fractional) maximal functions and the classical (fractional) maximal functions (see Section 5 and 6 below). Our approach proposed here would be technically delicate, but it eventually provides a powerful and useful assertion concerning the desired regularity estimates for the problem (1.1). It is also worth pointing out that Acerbi and Mingione [1] have obtained Calderón-Zygmund type estimates for parabolic problems having p-Laplacian structure (without measure data), by using the so-called maximal function free technique. This approach is based upon a stopping-time argument and direct PDE estimates without using maximal functions (see for example [3, 4, 9, 11, 12] for regularity results via this method).

Let us outline the plan of this paper as follows: in Section 2, we introduce some notion, preliminary tools, a suitable solution and regularity assumptions to establish our main theorems (Theorem 2.4 and 2.8). Section 3 concerns the standard energy type estimates for the problem (1.1). In Section 4, we investigate comparison estimates between the problem (1.1) and its limiting problems. Section 5, the main ingredient of this paper, is devoted to verifying the assumptions of the covering argument (Lemma 2.14) under the intrinsic parabolic cylinders alongside our intrinsic (fractional) maximal operator. Finally, in Section 6, we prove the global Calderón-Zygmund type estimates (Theorem 2.4 and 2.8) for the problem (1.1).

2. Preliminaries and main results

2.1. Notation. Let us first introduce general notation, which will be used later. A point \( x \in \mathbb{R}^n \) will be written \( x = (x_1, \ldots, x_n) \). Let \( B_r(x_0) \) denote the open ball in \( \mathbb{R}^n \) with center \( x_0 \) and radius \( r > 0 \), and let \( B_r^+(x_0) := B_r(x_0) \cap \{ x \in \mathbb{R}^n : x_n > 0 \} \). The standard parabolic cylinder in \( \mathbb{R}^n \times \mathbb{R} =: \mathbb{R}^{n+1} \) is denoted \( Q_r(x_0, t_0) := B_r(x_0) \times (t_0 - r^2, t_0 + r^2) \) with center \((x_0, t_0) \in \mathbb{R}^{n+1}\), radius \( r \) and height \( 2r^2 \). With \( \lambda > 0 \), we will often consider the intrinsic parabolic cylinder

\[
Q_r^\lambda(x_0, t_0) := B_r(x_0) \times (t_0 - \lambda^2 - p r^2, t_0 + \lambda^2 - p r^2),
\]

see [18, 31, 48] for the concept of intrinsic geometry related to the intrinsic parabolic cylinder. We also use the shorter notation as follows:

\[
\Omega_T := \Omega \times (0, T), \quad \Omega_{\bar{T}} := \Omega \times (-\infty, T), \quad \Omega_{\bar{T}} := \Omega \times (-T, T),
\]

\[
K_r^\lambda(x_0, t_0) := Q_r^\lambda(x_0, t_0) \cap \Omega_{\bar{T}},
\]

\[
Q_r^{\lambda+}(x_0, t_0) := B_r^+(x_0) \times (t_0 - \lambda^2 - p r^2, t_0 + \lambda^2 - p r^2),
\]

\[
T_r^\lambda(x_0, t_0) := (B_r(x_0) \cap \{ x \in \mathbb{R}^n : x_n = 0 \}) \times (t_0 - \lambda^2 - p r^2, t_0 + \lambda^2 - p r^2).
\]

Given a real-valued function \( f \), we write

\[
(f)^+ := \max \{ f, 0 \} \quad \text{and} \quad (f)^- := -\min \{ f, 0 \}.
\]

For each set \( U \subset \mathbb{R}^{n+1}, \| U \| \) is the \((n + 1)\)-dimensional Lebesgue measure of \( U \), \( \text{diam}(U) \) is the diameter of \( U \), and \( \chi_U \) is the usual characteristic function of \( U \). For \( f \in L^1_{\text{loc}}(\mathbb{R}^{n+1}), f_U \) stands for the integral average of \( f \) over a bounded open set \( U \subset \mathbb{R}^{n+1} \); that is,

\[
\bar{f}_U := \int_U f(x, t) \, dx dt := \frac{1}{|U|} \int_U f(x, t) \, dx dt.
\]
We use both the notation $f_t$ and $\partial_t f$ to denote the time derivative of a function $f$. In what follows, we denote by $c$ to mean any constant that can be computed in terms of known quantities; the exact value denoted by $c$ may be different from line to line.

Now we introduce the fractional maximal function used in this paper. For a deeper discussion of the fractional maximal function, we refer the reader to [2, 27, 37, 44]. For a locally finite measure $\nu$ on $\mathbb{R}^{n+1}$, the fractional maximal function of order 1 for $\nu$, denoted by $M_1(\nu)$, is defined as

$$ M_1(\nu)(x,t) := \sup_{r > 0} \left( \frac{\nu(r(x,t))}{r^{n+1}} \right) \quad \text{for } (x,t) \in \mathbb{R}^{n+1}. $$

When a measure $\nu_0$ is given on $\mathbb{R}^n$, we similarly define $M_1(\nu_0)$, with the standard parabolic cylinder $Q_r(x,t)$ replaced by the ball $B_r(x)$; that is,

$$ M_1(\nu_0)(x) := \sup_{r > 0} \left( \frac{\nu_0(B_r(x))}{r^{n+1}} \right) \quad \text{for } x \in \mathbb{R}^n. $$

Also, with $\lambda > 0$, we define the intrinsic fractional maximal function of order 1 for $\nu$ to be

$$ M_1^\lambda(\nu)(x,t) := \sup_{r > 0} \left( \frac{\nu(Q_r^\lambda(x,t))}{r^{n+1}} \right) \quad \text{for } (x,t) \in \mathbb{R}^{n+1}. $$

2.2. SOLA (Solution Obtained by Limits of Approximations). Let us first introduce a weak solution to the problem (1.1) when $\mu$ on the right-hand side of (1.1) is some integrable function. We say $u \in C(0, T; L^2(\Omega)) \cap L^p(0, T; W_0^{1,p}(\Omega))$ is a weak solution of (1.1) if

$$ \int_{\Omega_T} -u \varphi_t + \langle a(Du, x, t), D\varphi \rangle \ dx \, dt = \int_{\Omega_T} \mu \varphi \ dx \, dt $$

for all testing functions $\varphi \in C_c^\infty(\Omega_T)$. With the concept of the Steklov average (see Section 2.4 below for definition), we can also define the following discrete version of a weak formulation, which is equivalent to (2.5):

$$ \int_{\Omega_T \times \{t\}} \partial_t [u]_{h} \varphi + \langle [a(Du, x, t)]_h, D\varphi \rangle \ dx = \int_{\Omega_T \times \{t\}} [\mu]_h \varphi \ dx $$

for all $0 < t < T - h$ and all $\varphi \in W_0^{1,p}(\Omega)$. However, since the right-hand side measure $\mu$ of (1.1) does not in general belong to the dual space of $C(0, T; L^2(\Omega)) \cap L^p(0, T; W_0^{1,p}(\Omega))$, a solution $u$ of (1.1) does not become a weak solution. For instance, let us consider the parabolic $p$-Laplace problem

$$ u_t - \div (|Du|^{p-2}Du) = \delta_0 \quad \text{in } \mathbb{R}^n \times (0, \infty), $$

where $\delta_0$ is the Dirac measure charging the origin. Then the fundamental solution of (2.7) is

$$ \Gamma(x,t) = \begin{cases} 
-\frac{n}{p-1} \left( c_b - \frac{2}{p} - \frac{2}{p} \theta \frac{|x|}{t^\theta} \right)^{\frac{p-1}{p}} + & \text{if } p \neq 2, \\
(4\pi t)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4t}} & \text{if } p = 2, 
\end{cases} $$

for some $c_b = c_b(n, p) > 0$. Here $\theta := \frac{1}{p(n+1)-2n}$. The solution $\Gamma$ is well defined provided that $\theta > 0$; that is, $p > \frac{2n}{n+1}$. In addition, we see from the similarity
method (see for example [22, Chapter 4.2.2]) that \(|t|^q = |x|, |	au|\) denoting the dimension of the quantity \(\tau\). Then we deduce

\[
\int_0^\infty \int_{\mathbb{R}^n} |D\Gamma|^q \, dx \, dt < \infty \quad \text{for all } q < p - \frac{n}{n + 1},
\]

which implies the solution \(\Gamma \notin L^p(0, \infty; W^{1,p}(\mathbb{R}^n))\); that is, \(\Gamma\) is not a weak solution. The result in (2.8) also gives us that \(\Gamma \in L^1(0, \infty; W^{1,1}(\mathbb{R}^n))\) if and only if

\[
p - \frac{n}{n + 1} > 1 \iff p > 2 - \frac{1}{n + 1} \left( > \frac{2n}{n + 1} \right).
\]

This relation shows that the condition (1.3) is natural and crucial.

For this reason, we need to consider a more general class of solutions beyond the notion of weak solutions.

**Definition 2.1.** \(u \in L^1(0, T; W^{1,1}_0(\Omega))\) is a SOLA to (1.1) if there exists a sequence of weak solutions \(\{u_m\}_{m \geq 1} \subset C(0, T; L^2(\Omega)) \cap L^p(0, T; W^{1,p}_0(\Omega))\) of the regularized problems

\[
\begin{cases}
\partial_t u_m - \div a(Du_m, x, t) = \mu_m & \text{in } \Omega_T, \\
u_m = 0 & \text{on } \partial_o \Omega_T
\end{cases}
\]

such that

\[
u_m \to u \quad \text{in } L^\tilde{d}(0, T; W^{1,\tilde{d}}_0(\Omega)) \quad \text{as } m \to \infty,
\]

where \(\tilde{d} := \max\{1, p - 1\}\). Here \(\mu_m \in L^\infty(\Omega_T)\) converges weakly to \(\mu\) in the sense of measure and satisfies that for each cylinder \(Q := B \times (t_1, t_2) \subset \mathbb{R}^{n+1},\)

\[
\limsup_{m \to \infty} |\mu_m|(Q) \leq |\mu|(\{Q\}),
\]

with \(\mu_m\) defined in \(\mathbb{R}^{n+1}\) by the zero extension of \(\mu_m\) to \(\mathbb{R}^{n+1} \setminus \Omega_T\).

In the right-hand side of (2.11), the symbol \(\{Q\}\) denotes the parabolic closure of \(Q\) defined as

\[
\{Q\} := Q \cup \partial_o Q,
\]

where \(\partial_o Q := (\partial B \times [t_1, t_2]) \cup (B \times \{t_1\})\) is the parabolic boundary of \(Q\). We regard \(\mu_m\) as a suitable convolution of the measure \(\mu\) via mollification (see for instance [26, Lemma 5.1]). Then we obtain \(\mu_m \to \mu\) in the sense of measure satisfying the property (2.11) and the following uniform \(L^1\)-estimate:

\[
\|\mu_m\|_{L^1(\Omega_T)} \leq |\mu|(\Omega_T).
\]

An approximation method in [6, 7] ensures that there exists a SOLA \(u\) of the problem (1.1) such that

\[
u_m \to u \quad \text{in } L^q(0, T; W^{1,q}_0(\Omega)) \quad \text{for all } q < p - \frac{n}{n + 1}.
\]

We note that if the measure \(\mu\) belongs to the dual space \(L^p'(0, T; W^{-1,p'}(\Omega))\), then a SOLA and a weak solution actually coincide each other. On the other hand, the uniqueness of SOLA still remains unsolved except for \(\mu \in L^1(\Omega_T)\) or the linear case \((p = 2)\); that is, \(a(\xi, x, t) = a(x, t)\xi\). We refer to [6, 7, 43] and the references therein for a further discussion.
2.3. Main results. Let us first introduce the main regularity assumptions on the operator $a$ and the boundary of $\Omega$.

**Definition 2.2.** Given $R > 0$ and $\delta \in (0, \frac{1}{4})$, we say $(a, \Omega)$ is $(\delta, R)$-vanishing if the following two conditions hold:

1. The operator $a(\xi, x, t)$ satisfies
   
   $$
   \sup_{t, t' \in \mathbb{R}} \sup_{0 < r \leq R} \int_{B_r(y)} \Theta(a, B_r(y))(x, t) \, dx \, dt \leq \delta,
   $$
   
   where
   
   $$
   \Theta(a, B_r(y))(x, t) := \sup_{\xi \in \mathbb{R}^n \setminus \{0\}} \frac{|a(\xi, x, t) - \bar{a}_{B_r(y)}(\xi, t)|}{|\xi|^{p-1}}
   $$
   
   with $\bar{a}_{B_r(y)}(\xi, t)$ denoting the integral average of $a(\xi, \cdot, t)$ over the ball $B_r(y)$; that is,
   
   $$
   \bar{a}_{B_r(y)}(\xi, t) := \int_{B_r(y)} a(\xi, x, t) \, dx.
   $$

2. The domain $\Omega$ is called $(\delta, R)$-Reifenberg flat if 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
   
   $$
   B_r(0) \cap \{y \in \mathbb{R}^n : y_n > \delta r\} \subset B_r(0) \cap \Omega \subset B_r(0) \cap \{y \in \mathbb{R}^n : y_n > -\delta r\}.
   $$

**Remark 2.3.** (i) The parameter $\delta$ is sufficiently small to be determined in the proofs of Theorem 2.4 and 2.8. This number is invariant under the scaling of the problem (1.1), and the number $R$ is given arbitrary.

(ii) The condition (1) of Definition 2.2 implies that $a(\xi, x, t)$ is just measurable in the $t$-variable and of small BMO (Bounded Mean Oscillation) in the $x$-variable uniformly in $\xi$.

(iii) The boundary of the $(\delta, R)$-Reifenberg flat domain $\Omega$ can be trapped between two hyperplanes at the small scales. This boundary includes Lipschitz boundary with a small Lipschitz constant. Moreover, if $\Omega$ is $(\delta, R)$-Reifenberg flat, then we see that the following measure density conditions hold:

$$
\begin{align*}
\sup_{0 < r \leq R \subseteq \Omega} \frac{|B_r(y)|}{|\Omega \cap B_r(y)|} \leq \left( \frac{2}{1 - \delta} \right)^n \leq \left( \frac{16 \gamma}{7} \right)^n, \\
\inf_{0 < r \leq R \subseteq \partial \Omega} \frac{|\Omega \cap B_r(y)|}{|B_r(y)|} \geq \left( \frac{1 - \delta}{2} \right)^n \geq \left( \frac{7}{16} \right)^n.
\end{align*}
$$

For a further discussion on Reifenberg flat domains, we refer to [32, 47] and the references therein.

We are ready to present the following global Calderón-Zygmund type estimates for our problem (1.1):

**Theorem 2.4.** Let $p > 2 - \frac{1}{n+q}$ and let $u$ be a SOLA of the problem (1.1). Then for any $1 < q < \infty$, there exists a small constant $\delta = \delta(n, \Lambda_0, \Lambda_1, p, q) > 0$ such that if $(a, \Omega)$ is $(\delta, R)$-vanishing for some $R > 0$, then we have the following estimates:

1. For $p \geq 2$,
   
   $$
   \int_{\Omega_T} |Du|^q \, dx \, dt \leq c \left\{ \int_{\Omega_T} [M_1(\mu)]^q \, dx \, dt + [\mu(\Omega_T)]^{(n+2)(p-1)q/(n+1)p-n} + 1 \right\}
   $$
   
   where $M_1(\mu)$ is the Hardy-Littlewood maximal function of $\mu$.
for some constant $c = c(n, \Lambda_0, \Lambda_1, p, q, R, \Omega_T) \geq 1$, where $\mathcal{M}_1(\mu)$ is given in (2.2); for $2 - \frac{1}{n+1} < p \leq 2$,

$$
(2.16) \quad \int_{\Omega_T} |Du|^q \, dx \, dt \leq c \left\{ \int_{\Omega_T} [\mathcal{M}_1(\mu)]^{2n \alpha - 2n} \, dx \, dt + [\mu(\Omega_T)]^{2n \alpha - 2n} + 1 \right\}
$$

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

**Remark 2.5.** (i) Note that both the constants $c$ in (2.15) and (2.16) are stable as $p \to 2$. On the other hand, they blow up as $q \searrow 1$, see Remark 6.3 below. (ii) For $0 < q \leq 1$, gradient estimates like (2.15) and (2.16) directly follow from (2.17) and Lemma 3.1 below. In this case, we also refer to [3, 4, 11] in the setting of more general function frames including Marcinkiewicz spaces.

**Remark 2.6.** Note that $|\mu|(\Omega_T) \leq \text{diam}(\Omega_T) n + 1 \mathcal{M}_1(\mu)$, which implies

$$
(2.17) \quad [\mu(\Omega_T)]^{\alpha q} \leq c(n, \alpha, q, \Omega_T) \left( \int_{\Omega_T} \mathcal{M}_1(\mu) \, dx \, dt \right)^{\alpha q}
$$

for any $\alpha > 0$.

(i) If $p = 2$, then $\frac{(n+2)(p-1)}{(n+1)p-n} = \frac{2}{(n+1)p-n} = 1$; it follows from Theorem 2.4 and (2.17) that

$$
\int_{\Omega_T} |Du|^q \, dx \, dt \leq c \left\{ \int_{\Omega_T} [\mathcal{M}_1(\mu)]^q \, dx \, dt + 1 \right\},
$$

which is the main estimate in [13].

(ii) If $p > 2$, then $\frac{(n+2)(p-1)}{(n+1)p-n} > 1$; we derive from (2.15) and (2.17) that

$$
(2.18) \quad \int_{\Omega_T} |Du|^q \, dx \, dt \leq c \left\{ \left( \int_{\Omega_T} [\mathcal{M}_1(\mu)]^q \, dx \, dt \right)^{\frac{(n+2)(p-1)}{(n+1)p-n} - 1} + 1 \right\}.
$$

(iii) If $2 - \frac{1}{n+1} < p < 2$, then $\frac{2}{(n+1)p-n} > 1$; we have from (2.16) and (2.17) that

$$
(2.19) \quad \int_{\Omega_T} |Du|^q \, dx \, dt \leq c \left\{ \int_{\Omega_T} [\mathcal{M}_1(\mu)]^{2n \alpha - 2n} \, dx \, dt + 1 \right\}.
$$

**Remark 2.7.** The occurrence of the exponents $\frac{(n+2)(p-1)}{(n+1)p-n}$ and $\frac{2}{(n+1)p-n}$ in Theorem 2.4 is closely related to the structure and anisotropic property (a constant multiple of solution no longer becomes another solution) of the problem (1.1). More precisely, the exponent $\frac{(n+2)(p-1)}{(n+1)p-n}$ comes from the standard energy type estimate for the problem (1.1), see Lemma 3.1 and 5.1 below. On the other hand, the exponent $\frac{2}{(n+1)p-n}$ appears owing to the ratio between the standard parabolic cylinder $Q_r(x, t)$ and the intrinsic parabolic cylinder $Q_r^+(x, t)$, see Lemma 5.4 below. Also, there are similar situations in the regularity theory for parabolic $p$-Laplace type problems. We refer to [28, 30, 31] for potential estimates and [1, 9, 12] for Calderón-Zygmund estimates, respectively.

If the measure $\mu$ is time-independent or admits a favorable decomposition, see (2.20) below, then we obtain more sharp gradient estimates than the estimates in Theorem 2.4 as follows:
Theorem 2.8. Let $p > 2 - \frac{1}{n+1}$ and let $\delta := \max\{1, p - 1\} < q < \infty$. Suppose that the following decomposition holds:

\[
\mu = \mu_0 \otimes f,
\]

where $\mu_0$ is a finite signed Radon measure on $\Omega$ and $f \in L^{\frac{q}{n}}(0, T)$. Then there exists a small constant $\delta = \delta(n, \Lambda_0, \Lambda_1, p, q) > 0$ such that the following holds: if $(\mathbf{a}, \Omega)$ is $(\delta, R)$-vanishing for some $R > 0$, then for any SOLA $u$ of the problem (1.1) we have

\[
\int_{\Omega_T} |Du|^q \, dx \, dt \leq c \left\{ \int_{\Omega_T} \left[ \left( \mathcal{M}_1(\mu_0) \right) f \right]^{\frac{n}{n-r}} \, dx \, dt \right. \\
+ \left. \left[ |\mu_0|(\Omega) \|f\|_{L^1(0, T)} \right]^{\frac{n+2(1-p)}{n+2(p-1)}} + 1 \right\},
\]

for some constant $c = c(n, \Lambda_0, \Lambda_1, p, q, \Omega_T) \geq 1$, where $\mathcal{M}_1(\mu_0)$ is given in (2.3).

Remark 2.9. (i) The constant $c$ in (2.21) blows up as $q \searrow \max\{1, p - 1\}$, see Remark 6.3 and (6.13) below.

(ii) For $p > 2 - \frac{1}{n+1}$, we see from (2.21) and (2.17) that

\[
\int_{\Omega_T} |Du|^q \, dx \, dt \leq c \left\{ \left( \int_{\Omega_T} \left[ \left( \mathcal{M}_1(\mu_0) \right) f \right]^{\frac{n}{n-r}} \, dx \, dt \right) \right. \\
+ \left. \left[ \frac{|\mu_0|(\Omega) \|f\|_{L^1(0, T)}}{\max\{1, p - 1\}} \right]^{\frac{n+2(1-p)}{n+2(p-1)}} + 1 \right\}.
\]

Comparing (2.18)–(2.19) with (2.22), we observe that the estimate (2.22) gives a more natural and sharp result.

2.4. Preliminary tools. Let us first recall definition and some properties of Steklov average, see [18] for details. Given $0 < h < T$, the Steklov average $[f]_h$ of a function $f \in L^1(\Omega_T)$ is defined by

\[
[f]_h(\cdot, t) := \begin{cases} \frac{1}{h} \int_t^{t+h} f(\cdot, \tau) \, d\tau & \text{for } t \in (0, T - h], \\ 0 & \text{for } t > T - h. \end{cases}
\]

Lemma 2.10 (See [18, Chapter I, Lemma 3.2]). Let $f \in L^r(0, T; L^q(\Omega))$ for some $q, r \geq 1$. Then $[f]_h \to f$ in $L^r(0, T - \varepsilon; L^q(\Omega))$ as $h \to 0$ for every $\varepsilon \in (0, T)$. If $f \in C(0, T; L^q(\Omega))$, then $[f]_h(\cdot, t) \to f(\cdot, t)$ in $L^q(\Omega)$ for every $t \in (0, T - \varepsilon)$ and every $\varepsilon \in (0, T)$.

We also use a parabolic embedding theorem as follows:

Lemma 2.11 (See [18, Chapter I, Proposition 3.1]). Let $q, s \geq 1$. Then there is a constant $c = c(n, q, s) \geq 1$ such that for every $f \in L^\infty(0, T; L^q(\Omega)) \cap L^q(0, T; W^{1,q}_{0}(\Omega))$, we have

\[
\int_{\Omega_T} |f|^{q+\frac{q}{n+1}} \, dx \, dt \leq c \left( \int_{\Omega_T} |Df|^q \, dx \, dt \right) \left( \sup_{0 \leq t < T} \int_{\Omega} |f|^s \, dx \right)^{\frac{q}{s}}.
\]

We now introduce some analytic and geometric properties which will be crucially used in this paper. Let $f$ be a locally integrable function in $\mathbb{R}^{n+1}$ and let $\lambda > 0$. We define the (intrinsic) $\lambda$-maximal function of $f$ as

\[
\mathcal{M}_\lambda^f(x, t) := \sup_{r > 0} \frac{1}{Q^1_\lambda(x, t)} \int_{Q^1_\lambda(x, t)} |f(y, s)| \, dy \, ds,
\]
where $Q_\lambda(x,t)$ is the intrinsic parabolic cylinder as in (2.1). We also use

\[(2.23)\quad \mathcal{M}_U^\lambda f := \mathcal{M}^\lambda (\chi_U f)\]

if $f$ is defined on a set $U$ with the usual characteristic function $\chi_U$ of $U$. In particular, when $\lambda = 1$ or $p = 2$, it becomes the classical maximal function $\mathcal{M}f$.

We will use the following weak $(1,1)$-estimates for the $\lambda$-maximal function:

**Lemma 2.12.** If $f \in L^1(\mathbb{R}^{n+1})$, then there exists a constant $c = c(n) \geq 1$ such that

\[(2.24)\quad \left| \left\{ (y,s) \in \mathbb{R}^{n+1} : \mathcal{M}^\lambda f(y,s) > \alpha \right\} \right| \leq \frac{c}{\alpha} \int_{\mathbb{R}^{n+1}} |f(x,t)| \, dxdt\]

for any $\alpha > 0$. Moreover, we have

\[(2.25)\quad \left| \left\{ (y,s) \in \mathbb{R}^{n+1} : \mathcal{M}^\lambda f(y,s) > 2\alpha \right\} \right| \leq \frac{c}{\alpha} \int_{\{(y,s)\in\mathbb{R}^{n+1} : |f| > \alpha\}} |f| \, dxdt\]

for any $\alpha > 0$.

**Proof.** Let $f^\lambda(x,t) = f(x,\lambda^2-p\tau)$ and set $s = \lambda^2-p\tau$. Then we find

\[\int_{Q_{r}(y,\tau)} |f^\lambda(x,t)| \, dxdt = \int_{Q_{\lambda}(y,s)} |f(x,t)| \, dxdt\]

for any $r > 0$ and any $(y,\tau) \in \mathbb{R}^{n+1}$, which implies

\[(\mathcal{M}f^\lambda)(y,\tau) = (\mathcal{M}^\lambda f)(y,s)\]

From the weak type $(1,1)$-estimate for the classical maximal function (see for instance [45, Chapter I, Theorem 1]), it follows

\[\left| \left\{ (y,s) \in \mathbb{R}^{n+1} : \mathcal{M}^\lambda f(y,s) > \alpha \right\} \right| = \lambda^{2-p} \left| \left\{ (y,\tau) \in \mathbb{R}^{n+1} : \mathcal{M}f^\lambda(y,\tau) > \alpha \right\} \right|
\]

\[\leq c \frac{\lambda^{2-p}}{\alpha} \int_{\mathbb{R}^{n+1}} |f^\lambda(x,t)| \, dxdt
\]

\[= \frac{c}{\alpha} \int_{\mathbb{R}^{n+1}} |f(x,t)| \, dxdt,
\]

which shows (2.24).

To prove (2.25), let us consider the function $f^\alpha := \chi_{\{|f| > \alpha\}} |f|$. Then it is straightforward to check that for $(y,s) \in \mathbb{R}^{n+1}$,

\[|f(y,s)| \leq f^\alpha(y,s) + \alpha,
\]

and so

\[\mathcal{M}^\lambda f(y,s) \leq \mathcal{M}^\lambda f^\alpha(y,s) + \alpha.
\]

Therefore we see

\[
\{(y,s) \in \mathbb{R}^{n+1} : \mathcal{M}^\lambda f(y,s) > 2\alpha \}\subset \{(y,s) \in \mathbb{R}^{n+1} : \mathcal{M}^\lambda f^\alpha(y,s) > \alpha\}.
\]

Applying (2.24) to $\{(y,s) \in \mathbb{R}^{n+1} : \mathcal{M}^\lambda f^\alpha(y,s) > \alpha\}$, we derive (2.25). \qed

We record some useful property as follows:

**Lemma 2.13.** Let $U$ be an open set in $\mathbb{R}^{n+1}$ and let $q > s \geq 0$. For $g \in L^s(U)$, we have

\[(2.26)\quad \int_{U} |g|^{q-s} |g|^s \, dxdt = (q-s) \int_{0}^{k} \lambda^{q-s-1} \left[ \int_{\{(y,s) \in U : |g(y,s)| > \lambda\}} |g|^s \, dxdt \right] d\lambda,
\]
where the truncated function \( |g|_k := \min\{ |g|, k \} \) for any \( k > 0 \). If \( g \in L^q(U) \), then (2.26) also holds for \( k = \infty \).

**Proof.** By Fubini’s theorem, it is easy to check that Lemma 2.13 holds.

We end this section with the following modified version of Vitali’s covering lemma for the intrinsic parabolic cylinder:

**Lemma 2.14.** Assume that \( \Omega \) is \((\delta, R)\)-Reifenberg flat. Let \( 0 < \varepsilon < 1 \), let \( \lambda > 0 \), and let \( \mathcal{C} \subset \mathcal{D} \) be two bounded measurable subsets of \( \Omega_T := \Omega \times (-\infty, T) \) such that the following two conditions hold:

(i) \( |\mathcal{C}| < \varepsilon |Q^\lambda_{R/10}| \), and

(ii) for any \((x, t) \in \Omega_T\) and any \( r \in (0, \frac{R}{10}] \) with \( |\mathcal{C} \cap Q^\lambda_r(x, t)| \geq \varepsilon |Q^\lambda_r| \),

\( Q^\lambda_r(x, t) \cap \Omega_T \subset \mathcal{D} \).

Then we have

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

**Proof.** According to (i), for almost every \((x, t) \in \mathcal{C}\), there is \( r_{(x, t)} \in (0, \frac{R}{10}) \) such that

\[
|\mathcal{C} \cap Q^\lambda_{r_{(x, t)}}(x, t)| = \varepsilon |Q^\lambda_{r_{(x, t)}}| \quad \text{and} \quad |\mathcal{C} \cap Q^\lambda_r(x, t)| < \varepsilon |Q^\lambda_r|
\]

for all \( r \in (r_{(x, t)}, \frac{R}{10}] \). Since \( \{Q^\lambda_{r_{(x, t)}}(x, t) : (x, t) \in \mathcal{C}\} \) is a covering of \( \mathcal{C} \), the standard Vitali covering lemma (see [8, Theorem C.1] or [23, Theorem 1.24]) implies that there exists a disjoint subcovering \( \{Q^\lambda_{r_{i}}(x_i, t_i) : (x_i, t_i) \in \mathcal{C}\}_{i=1}^\infty \) such that

\[
\mathcal{C} \subset \bigcup_{i \geq 1} Q^\lambda_{r_{i}}(x_i, t_i) \quad \text{and} \quad |\mathcal{C}| \leq 5^{n+2} \sum_{i \geq 1} |Q^\lambda_{r_{i}}|.
\]

On the other hand, we compute for \((x, t) \in \Omega_T\) that

\[
\frac{|Q^\lambda_r(x, t)|}{|\Omega_T \cap Q^\lambda_r(x, t)|} = \frac{2\lambda^2-p r^2 |B_r(x)|}{|(t-\infty, T) \cap (t - \lambda^2-p r^2, t + \lambda^2-p r^2)| |\Omega \cap B_r(x)|}
\]

\[
= \frac{2\lambda^2-p r^2 |B_r(x)|}{|(t-\lambda^2-p r^2, \min\{T, t + \lambda^2-p r^2\})| |\Omega \cap B_r(x)|}
\]

\[
\leq \frac{2 |B_r(x)|}{|\Omega \cap B_r(x)|}.
\]

Since \( \Omega \) is \((\delta, R)\)-Reifenberg flat, (2.29) and (2.14) yield

\[
\sup_{0 < r \leq R} \sup_{(x, t) \in \Omega_T} \frac{|Q^\lambda_r(x, t)|}{|\Omega_T \cap Q^\lambda_r(x, t)|} \leq \sup_{0 < r \leq R} \sup_{x \in \Omega} \frac{2 |B_r(x)|}{|\Omega \cap B_r(x)|} \leq 2 \left( \frac{2}{1-\delta} \right)^{n+2}.
\]
Finally, we have from (2.28), (2.27), (2.30) and (ii) that

\[ |\mathcal{C}| = \left| \bigcup_{i \geq 1} (\mathcal{C} \cap Q_{\delta r_i}(x_i, t_i)) \right| \leq \sum_{i \geq 1} |\mathcal{C} \cap Q_{\delta r_i}^\lambda(x_i, t_i)| < \varepsilon \sum_{i \geq 1} |Q_{\delta r_i}^\lambda(x_i, t_i)| \]

\[ = \varepsilon^{5n+2} \sum_{i \geq 1} |Q_{r_i}^\lambda(x_i, t_i)| \leq \left( \frac{10}{1-\delta} \right)^{n+2} \varepsilon \sum_{i \geq 1} |Q_{r_i}^\lambda(x_i, t_i) \cap \Omega_T| \]

\[ = \left( \frac{10}{1-\delta} \right)^{n+2} \varepsilon \left| \bigcup_{i \geq 1} (Q_{r_i}^\lambda(x_i, t_i) \cap \Omega_T) \right| \leq \left( \frac{10}{1-\delta} \right)^{n+2} \varepsilon |\mathcal{D}|, \]

which completes the proof.

Remark 2.15. If we replace \( \Omega_T \) by \( \Omega_T \) in the lemma above, then we could not have the measure density condition (2.30) independent of \( \lambda \), see (2.29). For this reason we obtain comparison results on the localized region of \( \Omega_T \), not \( \Omega_T \), see Section 4 and 5 below.

Remark 2.16. A shape of the parabolic intrinsic cylinders \( Q_{r_i}^\lambda(x, t) \) is essential to prove Lemma 2.14; if we take the intrinsic cylinders having the top center such as \( B_r(x) \times (t - \lambda^2 - r^2, t) \) instead of \( Q_{r_i}^\lambda(x, t) \), then the standard Vitali covering lemma no longer holds, see [8, Theorem C.1].

3. Standard energy type estimates

In this section we derive standard energy type estimates for the parabolic measure data problem (1.1) as follows:

Lemma 3.1. Let \( u \) be a SOLA of (1.1) with \( p > 2 - \frac{1}{n+1} \) and let \( 0 < \kappa < p - \frac{n}{n+1} \). Then there exists a constant \( c = c(n, \Lambda_0, p, \kappa) \geq 1 \) such that

\[ \left( \int_{\Omega_T} |Du|^\kappa \, dx \, dt \right)^{\frac{1}{\kappa}} \leq c \left[ \frac{\mu|\Omega_T|}{|\Omega_T|^{\frac{n+2}{n+2}}} \right]^{\frac{n+2}{n+\kappa}}. \]  

Proof. The ideas of our proof are developed in [31, Lemma 4.1]. We also refer to [6, Lemma 2.2], [28, Lemma 4.3] and [21, Lemma 4.1]. A main idea of the proof is to take some truncations of a solution \( u \), instead of \( u \) itself, as a test function, see (3.4) and (3.11) below. For the convenience of the reader, we give all the technicalities of the proof.

Step 1. Since \( u \) is a SOLA of (1.1), there exists a sequence of weak solutions \( \{u_m\} \) of the regularized problems (2.9) satisfying (2.10) and (2.13). We will first show that

\[ \sup_{0<t<T} \int_{\Omega \times \{t\}} |u_m| \, dx \leq |\mu|(\Omega_T) \]  

and

\[ \int_{\Omega_T} \frac{|Du_m|^p}{(\alpha + |u_m|)^\xi} \, dx \, dt \leq c \frac{\alpha^{1-\xi}}{\xi-1} |\mu|(\Omega_T) \]
for $\alpha > 0$ and $\xi > 1$, where $c = c(n, \Lambda_0, p) \geq 1$. For any fixed $\varepsilon > 0$, choose a test function
\begin{equation}
\varphi_1(x, t) = \pm \min\left\{1, \frac{[(u_m)_{h}]_{\pm}(x, t)}{\varepsilon}\right\} \zeta(t),
\end{equation}
where $\zeta$ is a nonincreasing smooth function on $t \in \mathbb{R}$ satisfying $0 \leq \zeta \leq 1$ and $\zeta(t) = 0$ for all $t \geq \tau$ and for $\tau \in (0, T)$. Clearly, $\varphi_1(\cdot, t) \in W_0^{1, P}(\Omega)$ with $|\varphi_1| \leq 1$. Substituting $\varphi_1$ into the weak formulation (2.6) and integrating on $(0, T)$, we obtain (3.5)
\begin{equation}
\int_{\Omega_T} \partial_t [(u_m)_{h}]_{\varphi_1} \, dx \, dt + \int_{\Omega_T} \langle [a(Du_m, x, t)]_{h}, D\varphi_1 \rangle \, dx \, dt = \int_{\Omega_T} [(\mu_m)_{h}]_{\varphi_1} \, dx \, dt.
\end{equation}

To estimate $I_1$, we see
\begin{equation}
\partial_t [(u_m)_{h}] \min\left\{1, \frac{[(u_m)_{h}]_{\pm}(x, t)}{\varepsilon}\right\} = \pm \partial_t \int_0^{[(u_m)_{h}]_{\pm}} \min\left\{1, \frac{s}{\varepsilon}\right\} \, ds,
\end{equation}
and then an integration by parts and Lemma 2.10 yield
\begin{equation}
I_1 = \int_{\Omega} \left[ \int_0^{[(u_m)_{h}]_{\pm}} \min\left\{1, \frac{s}{\varepsilon}\right\} \, ds \right] (-\zeta_t) \, dx \, dt
- \int_{\Omega} \left[ \int_0^{[(u_m)_{h}]_{\pm}(x, 0)} \min\left\{1, \frac{s}{\varepsilon}\right\} \, ds \right] \zeta(x, 0) \, dx
\xrightarrow{h \to 0} \int_{\Omega} \left[ \int_0^{(u_m)_{\pm}} \min\left\{1, \frac{s}{\varepsilon}\right\} \, ds \right] (-\zeta_t) \, dx \, dt \geq 0,
\end{equation}
since $\zeta_t \leq 0$. On the other hand, it follows from Lemma 2.10, (1.4) and (2.13) that
\begin{equation}
I_2 \xrightarrow{h \to 0} \frac{1}{\varepsilon} \int_{\Omega} \langle a(Du_m, x, t), Du_m \rangle \chi_{\{0 < (u_m)_{\pm} < \varepsilon\}} \zeta(t) \, dx \, dt \geq 0
\end{equation}
and
\begin{equation}
|I_3| \leq \int_{\Omega} |\mu_m|_h \, dx \xrightarrow{h \to 0} \int_{\Omega} |\mu_m| \, dx \leq |\mu|(\Omega_T).
\end{equation}
Combining (3.5)–(3.9), we find
\begin{equation}
\int_{\Omega} \left[ \int_0^{(u_m)_{\pm}} \min\left\{1, \frac{s}{\varepsilon}\right\} \, ds \right] (-\zeta_t) \, dx \, dt \leq |\mu|(\Omega_T),
\end{equation}
and letting $\varepsilon \to 0$ we see from Lebesgue’s dominated convergence theorem that
\begin{equation}
\int_{\Omega} \{u_m\} (-\zeta_t) \, dx \, dt \leq |\mu|(\Omega_T).
\end{equation}
Approximating $\zeta$ by the mollification of the characteristic function $\chi_{(-\infty, \tau]}$ yields
\begin{equation}
\int_{\Omega \times \{\tau\}} |u_m| \, dx \leq |\mu|(\Omega_T)
\end{equation}
for every $\tau \in (0, T)$, which implies (3.2).

It also follows from (3.5)–(3.9) that
\begin{equation}
\sup_{\varepsilon > 0} \frac{1}{\varepsilon} \int_{\Omega} \langle a(Du_m, x, t), Du_m \rangle \chi_{\{0 < (u_m)_{\pm} < \varepsilon\}} \, dx \, dt \leq |\mu|(\Omega_T).
\end{equation}
We now select the test function
\[
\varphi_2 := \frac{\varphi_1}{(\alpha + ([u_m]_h)_\pm)^{1-\xi}},
\]
where both constants $\xi > 1$ and $\alpha > 0$ are to be determined later. Substituting $\varphi_2$ into (2.6) and integrating over $(0, T)$, we get
\[
\int_{\Omega_T} \partial_t[u_m]_h \varphi_2 \, dx \, dt + \int_{\Omega_T} \langle [a(Du_m, x, t)]_h, D\varphi_2 \rangle \, dx \, dt = \int_{\Omega_T} [\mu_m]_h \varphi_2 \, dx \, dt.
\]
By calculations similar to those in (3.6)–(3.9), we find
\[
\sup_{\varepsilon > 0} \lim_{h \to 0} \int_{\Omega_T} \partial_t[u_m]_h \varphi_2 \, dx \, dt \leq \alpha^{1-\xi} |\mu|(\Omega_T)
\]
and
\[
\lim_{h \to 0} \left| \int_{\Omega_T} [\mu_m]_h \varphi_2 \, dx \, dt \right| \leq \alpha^{1-\xi} |\mu|(\Omega_T).
\]
For the second term on the left-hand side of (3.12), we establish
\[
\int_{\Omega_T} \langle [a(Du_m, x, t)]_h, D\varphi_2 \rangle \, dx \, dt
\]
\[
= \int_{\Omega_T} \langle [a(Du_m, x, t)]_h, D\varphi_1 \rangle \frac{1}{(\alpha + ([u_m]_h)_\pm)^{1-\xi}} \, dx \, dt
\]
\[
+ (1 - \xi) \int_{\Omega_T} \langle [a(Du_m, x, t)]_h, D([u_m]_h)_\pm \rangle \frac{\varphi_1}{(\alpha + ([u_m]_h)_\pm)^{\xi}} \, dx \, dt
\]
\[
=: J_1 + J_2.
\]
It follows from Lemma 2.10, (3.8) and (3.10) that
\[
\sup_{\varepsilon > 0} \lim_{h \to 0} J_1 \leq \alpha^{1-\xi} |\mu|(\Omega_T).
\]
To estimate $J_2$, letting $h \to 0$ gives
\[
\lim_{h \to 0} J_2 = (1 - \xi) \int_{\Omega_T} \langle a(Du_m, x, t), D(u_m)_\pm \rangle \min \left\{ 1, \frac{(u_m)_\pm}{\varepsilon} \right\} \frac{\varphi_1}{(\alpha + (u_m)_\pm)^{\xi}} \, dx \, dt.
\]
Combining (3.12)–(3.17) and (1.4), we discover
\[
\int_{\Omega_T} \frac{|Du_m|^p}{(\alpha + |u_m|)_\pm} \min \left\{ 1, \frac{|u_m|}{\varepsilon} \right\} \, dx \, dt \leq c \frac{\alpha^{1-\xi}}{\xi - 1} |\mu|(\Omega_T)
\]
for some constant $c = c(n, \Lambda_0, p) \geq 1$. As $\varepsilon \to 0$, we obtain (3.3).

**Step 2.** Now recalling that $0 < \kappa < p - \frac{n}{n+1}$, we set
\[
\xi := \frac{n + 1}{n} (p - \kappa) \quad \text{and} \quad \alpha := \left( \int_{\Omega_T} |u_m|^{\frac{n+1}{\kappa}} \, dx \, dt \right)^{\frac{\kappa}{n+1}}.
\]
Clearly $\xi > 1$ and we may assume that $\alpha > 0$. (If $\alpha = 0$, then $u_m = 0$ and (3.1) holds with a SOLA $u$ replaced by $u_m$. When $m \to \infty$, the proof is done.)
We first assume that $1 \leq \kappa < p - \frac{n}{n+1}$. From Lemma 2.11 (with $q = \kappa$ and $s = 1$) and (3.2), we discover
\begin{equation}
\alpha \leq c(n, \kappa) \left( \int_{\Omega_T} |Du_m|^\kappa dt \left( \sup_{0 < t < T} \int_{\Omega \times \{t\}} |u_m| dx \right) \right)^{\frac{\kappa}{\kappa+1}} \tag{3.19}
\end{equation}
\begin{equation}
\leq c |\mu|(\Omega_T)^{\frac{\kappa}{\kappa+1}} \left( \int_{\Omega_T} |Du_m|^\kappa dt \right)^{\frac{\kappa}{\kappa+1}}.
\end{equation}

Then by Hölder’s inequality, (3.3), (3.18) and (3.19), we deduce
\begin{equation}
\int_{\Omega_T} |Du_m|^\kappa dt = \int_{\Omega_T} \left( \frac{|Du_m|^p}{(\alpha + \left| u_m \right|)^\xi} \right)^{\frac{\kappa}{p}} \left( \alpha + \left| u_m \right| \right)^{\frac{\kappa}{\alpha} - \frac{\kappa}{p}} dx dt
\end{equation}
\begin{equation}
\leq \left( \int_{\Omega_T} \left( \frac{|Du_m|^p}{(\alpha + \left| u_m \right|)^\xi} \right)^{\frac{\kappa}{p}} dx dt \right)^{\frac{\kappa}{\kappa+1}} \left( \int_{\Omega_T} \left( \alpha + \left| u_m \right| \right)^{\frac{\kappa}{\alpha} - \frac{\kappa}{p}} dx dt \right)^{\frac{\kappa}{\kappa+1}}
\end{equation}
\begin{equation}
\leq c \left( \frac{|\mu|(\Omega_T)^{\frac{\kappa}{\kappa+1}}}{|\Omega|^\frac{\kappa}{\kappa+1}} \right)^{\frac{\kappa}{\kappa+1}} \left( \int_{\Omega_T} |Du_m|^\kappa dx dt \right)^{\frac{\kappa}{\kappa+1}},
\end{equation}
for some constant $c = c(n, \Lambda, p, \kappa) \geq 1$. Applying Young’s inequality to (3.20) and letting $m \to \infty$, we finally obtain the estimate (3.1) for $1 \leq \kappa < p - \frac{n}{n+1}$.

If $0 < \kappa < 1$, then it follows from Hölder’s inequality that
\begin{equation}
\left( \int_{\Omega_T} |Du|^\kappa dx dt \right)^{\frac{\kappa}{\kappa+1}} \leq \int_{\Omega_T} |Du| dx dt \leq c \left[ \frac{|\mu|(\Omega_T)^{\frac{\kappa}{\kappa+1}}}{|\Omega|^\frac{\kappa}{\kappa+1}} \right]^{\frac{\kappa}{\kappa+1}},
\end{equation}
which we have used the estimate (3.1) with $\kappa = 1$ for the second inequality above. This completes the proof. \hfill \Box

**Remark 3.2.** We point out that the constant $c$ in (3.1) blows up when $\kappa \nearrow p - \frac{n}{n+1}$, since $c$ is proportional to $\frac{1}{1-\kappa}$ and $\xi \searrow 1$ as $\kappa \searrow p - \frac{n}{n+1}$, see (3.20) and (3.18).

4. $L^1$-Comparison Estimates on Intrinsic Parabolic Cylinders

In this section, when considering our problem (1.1) with (1.2) and (1.3), we first extend a solution $u$ by zero for $t < 0$ and assume that
\begin{equation}
\begin{cases}
\mu \in L^1(\Omega_T) \cap L^p(-\infty, T; W^{-1, p'}(\Omega)), \\
\mu \equiv 0 \quad &\text{for } t \leq 0,
\end{cases}
\end{equation}
see Remark 2.15 for the discussion of the time extension. We establish comparison $L^1$-estimates for the spatial gradient of the weak solution $u$ to (1.1) in a localized intrinsic parabolic cylinder near the boundary of $\Omega$. We only treat comparison results in such a boundary region, as the counterparts in an interior region can be derived in the same way.

Suppose that $(a, \Omega)$ is $(\delta, R)$-vanishing for some $R > 0$, where $\delta \in (0, \frac{1}{8})$ is to be determined later. Fix any $\lambda > 0$, $(x_0, t_0) \in \Omega_T$ and $0 < r \leq \frac{R}{\delta}$ satisfying
\begin{equation}
B_{r}(x_0) \subset B_{s}(x_0) \cap \Omega \subset B_{s}(x_0) \cap \{x \in \mathbb{R}^n : x_n > -16\delta r\}.
\end{equation}
Recall from Section 2.1 that $\Omega_T := \Omega \times (-\infty, T)$ and $K^\lambda_r(x_0, t_0) := Q^\lambda_r(x_0, t_0) \cap \Omega_T$. Throughout this section, for simplicity, we omit the center point $(x_0, t_0)$ in the intrinsic parabolic cylinder $K^\lambda_r(x_0, t_0)$. We denote, for a measurable set $E \subset \mathbb{R}^{n+1}$,

\begin{equation}
|\mu|(E) := \int_E |\mu(x, t)| \, dxdt
\end{equation}

and write

\[ \chi_{(p<2)} := \begin{cases} 0 & \text{if } p \geq 2, \\ 1 & \text{if } p < 2. \end{cases} \]

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

\begin{equation}
\begin{cases}
w_t - \text{div} a(Dw, x, t) = 0 \text{ in } K^\lambda_r, \\
w = u \text{ on } \partial_p K^\lambda_r.
\end{cases}
\end{equation}

Now, using the measure density condition (2.14), we extend the comparison estimates from [31, Lemma 4.1] and [28, Lemma 4.3] up to the boundary, as follows:

**Lemma 4.1.** Let $u$ be the weak solution of (1.1) and let $w$ as in (4.3). Then there exists a constant $c = c(n, \Lambda_0, p) \geq 1$ such that

\[ \int_{K^\lambda_r} |Dw - Du| \, dxdt \leq c \left[ \frac{|\mu(K^\lambda_r)|}{|K^\lambda_r|} \right]^{\frac{n+2}{n+1+p-n}} + c\chi_{(p<2)} \left[ \frac{|\mu(K^\lambda_r)|}{|K^\lambda_r|} \right] \left( \int_{K^\lambda_r} |Du| \, dxdt \right)^{\frac{(2-p)\frac{n+1}{n+2}}{n+2}}. \]

We next derive a boundary higher integrability result of the spatial gradient of the weak solution $w$ to the problem (4.3). For the case of an interior higher integrability, we refer to [25].

**Lemma 4.2.** Let $w$ be the weak solution of (4.3). Assume that

\begin{equation}
\int_{K^\lambda_r} |Dw| \, dxdt \leq c_w\lambda
\end{equation}

for some constant $c_w > 0$. Then there exist $\sigma = \sigma(n, \Lambda_0, \Lambda_1, p) > 0$ and $c = c(n, \Lambda_0, \Lambda_1, p, c_w) \geq 1$ such that

\begin{equation}
\int_{K^\lambda_r} |Dw|^{p(1+\sigma)} \, dxdt \leq c\lambda^{p(1+\sigma)}.
\end{equation}

**Proof.** We first set

\[ \tilde{\Omega} = \{ y \in \mathbb{R}^n : ry \in \Omega \} \quad \text{and} \quad \tilde{T} = \frac{T}{\lambda^{2-pf^2}} \]

and consider the rescaled operator $\tilde{a} = \tilde{a}(\xi, \lambda^2, \lambda^{2-pf^2}) : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ defined by

\[ \tilde{a}(\xi, x, t) = a \left( \lambda \xi, \frac{rx}{\lambda}, \frac{\lambda^{2-pf^2}t}{\lambda^{p-1}} \right). \]

We set

\[ \tilde{u}(x, t) = \frac{u(rx, \lambda^{2-pf^2}t)}{\lambda r} \quad \text{and} \quad \tilde{w}(x, t) = \frac{w(rx, \lambda^{2-pf^2}t)}{\lambda r}. \]
for $(x, t) \in \tilde{K}_s := Q_s \cap \left( \tilde{\Omega} \times (-\infty, \tilde{T}) \right)$. It is straightforward to check that $\tilde{a}$ satisfies the structure condition (1.2) with $a$ replaced by $\tilde{a}$ and that $\tilde{w}$ is the weak solution to the following problem

$$
\begin{align*}
\tilde{w}_t - \text{div} \tilde{a}(D\tilde{w}, x, t) &= 0 \quad \text{in} \quad \tilde{K}_s, \\
\tilde{w} &= \tilde{u} \quad \text{on} \quad \partial_p \tilde{K}_s.
\end{align*}
$$

Note that since $\Omega$ is $(\delta, R)$-Reifenberg flat, the measure density condition (2.14) holds, which implies that $\mathbb{R}^n \setminus \Omega$ becomes uniformly $p$-thick, see [10, Section 3.1] for more details. Then we deduce from [10, Theorem 2.2] and [24, Remark 6.12] that

$$
\int_{\tilde{K}_s} |D\tilde{w}|^{(1+\sigma)} \ dx dt \leq c \left( \int_{\tilde{K}_s} |D\tilde{w}|^{ps} \ dx dt \right)^{\frac{1}{p-1}} + c
$$

for every $s \in (0, 1]$, where $c = c(n, \Lambda_0, \Lambda_1, p, s) \geq 1$ and the constant $\tilde{d}$ is given by

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

After taking $s = \frac{1}{p}$, we scale back in (4.6) to discover

$$
\int_{K_{\lambda, r}^+} |Dw|^{(1+\sigma)} \ dx dt \leq c \lambda^{p\sigma(1-\tilde{d})} \left( \int_{K_{\lambda, r}^+} |Dw|^{p} \ dx dt \right)^{\frac{p(1+\sigma)}{2}} + c\lambda^{p(1+\sigma)}.
$$

Finally, by (4.4) and (4.7), we obtain (4.5). \hfill \Box

We next consider a freezing operator $\tilde{a}_{B_r^+} = \tilde{a}_{B_r^+}(\xi, t) : \mathbb{R}^n \times I_{\delta, r}^+ \to \mathbb{R}^n$ given by

$$
\tilde{a}_{B_r^+}(\xi, t) := \int_{B_{r}^+} a(\xi, x, t) \ dx,
$$

where $I_{\delta, r}^+ := (-\lambda^2-\rho(4r)^2, \lambda^2-\rho(4r)^2)$. Then $\tilde{a}_{B_r^+}$ satisfies the structure condition (1.2) with $a(\cdot, \cdot, t)$ replaced by $\tilde{a}_{B_r^+}(\xi, t)$. Recalling (1) of Definition 2.2, we also observe

$$
\int_{Q_{\lambda, r}^+} \Theta(a, B_r^+)(x, t) \ dx dt \leq 4 \int_{Q_{\lambda, r}^+} \Theta(a, B_r^+)(x, t) \ dx dt \leq 4\delta.
$$

Let $v$ be the unique weak solution to the coefficient frozen problem

$$
\begin{align*}
v_t - \text{div} \tilde{a}_{B_r^+}(Dv, t) &= 0 \quad \text{in} \quad K_{\lambda, r}^+, \\
v &= w \quad \text{on} \quad \partial_p K_{\lambda, r}^+.
\end{align*}
$$

From Lemma 4.2, (2.14), (4.8) and [12, Lemma 3.10], we derive the comparison result between (4.3) and (4.9) as follows:

**Lemma 4.3.** Let $w$ be the weak solution of (4.3) satisfying (4.4) and let $v$ as in (4.9). Then there is a constant $c = c(n, \Lambda_0, \Lambda_1, p) \geq 1$ such that

$$
\int_{K_{\lambda, r}^+} |Dw - Dv|^p \ dx dt \leq c\sigma_1^1 \lambda^p,
$$

where $\sigma_1$ is a positive constant depending only on $n$, $\Lambda_0$, $\Lambda_1$ and $p$. 
In the interior region ($K_{2r}^\lambda = Q_{2r}^\lambda$), the spatial gradient of the weak solution $v$ to (4.9) is locally bounded, see [18-20]. However, for the boundary case ($K_{2r}^\lambda \neq Q_{2r}^\lambda$), the $L^\infty$-norm of $Dv$ need not necessarily be bounded when the boundary of $\Omega$ may be extremely irregular. For this reason, we consider a weak solution $\bar{v}$ to the following limiting problem near the flat boundary:

$$\begin{aligned}
\bar{v}_t - \text{div} \tilde{a}_{B^r}(D\bar{v}, t) &= 0 \quad \text{in } Q_{2r}^{\lambda^+, 1} , \\
\bar{v} &= 0 \quad \text{on } T_{2r}^\lambda ,
\end{aligned}$$

where $Q_{2r}^{\lambda^+, 1}$ and $T_{2r}^\lambda$ are given in Section 2.1.

Then from [34, Theorem 1.6] (with $G(\tau) = \tau^p$ and $\varphi = 0$), we obtain the boundedness of the spatial gradient of $\bar{v}$ near the flat boundary, as follows:

**Lemma 4.4.** For any weak solution $\bar{v}$ of (4.10), we have

$$||D\bar{v}||_{L^\infty(Q_{2r}^{\lambda^+, 1})} \leq c \int_{Q_{2r}^{\lambda^+, 1}} |D\bar{v}|^p \, dxdt + c\lambda^p$$

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

We next have the following comparison estimate between (4.9) and (4.10):

**Lemma 4.5.** For any $\varepsilon \in (0, 1)$, there is a small constant $\delta = \delta(n, \Lambda_0, \Lambda_1, p, \varepsilon) > 0$ such that if $v$ is the weak solution of (4.9) satisfying

$$\int_{K_{2r}^\lambda} |Dv|^p \, dxdt \leq c_v \lambda^p$$

for some given constant $c_v \geq 1$, then there exists a weak solution $\bar{v}$ of (4.10) such that

$$\int_{K_{2r}^\lambda} |Dv - D\bar{v}|^p \, dxdt \leq \varepsilon^p \lambda^p \quad \text{and} \quad \int_{K_{2r}^\lambda} |D\bar{v}|^p \, dxdt \leq c\lambda^p$$

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

*Proof.* The first estimate in (4.11) can be derived from the compactness argument as in [12, Lemma 3.8]. From this estimate, (2.14) and the triangle inequality, we directly obtain the second estimate in (4.11). \qed

We finally combine the previous results to obtain the boundary comparison $L^1$-estimate.

**Lemma 4.6.** For any $\varepsilon \in (0, 1)$, there is a small constant $\delta = \delta(n, \Lambda_0, \Lambda_1, p, \varepsilon) > 0$ such that the following holds: if $u, w$ and $v$ are the weak solutions of (1.1), (4.3) and (4.9), respectively, satisfying

$$\int_{K_{2r}^\lambda} |Du| \, dxdt \leq \lambda \quad \text{and} \quad \frac{|\mu|(K_{2r}^\lambda)}{r^{n+1}} \leq \delta \lambda ,$$

then there exists a weak solution $\bar{v}$ of (4.10) such that

$$\int_{K_{2r}^\lambda} |Du - D\bar{v}| \, dxdt \leq \varepsilon \lambda \quad \text{and} \quad \|D\bar{v}\|_{L^\infty(K_{2r}^\lambda)} \leq c\lambda ,$$

for some constant $c = c(n, \Lambda_0, \Lambda_1, p) \geq 1$, where $\bar{v}$ is extended by zero from $Q_{2r}^{\lambda^+, 1}$ to $K_{2r}^\lambda$. 
Proof. First of all, it follows from (4.12) and Lemma 4.1 that
\begin{equation}
\int_{K^{4.1\nu}_{\lambda}} |Du - D\bar{w}| \, dxdt \leq c_\text{min} \min \left\{ 1, \frac{\lambda^2}{(\rho+\delta)^2} \right\} \lambda \quad \text{and} \quad \int_{K^{4.1\nu}_{\lambda}} |Dw| \, dxdt \leq c_w \lambda
\end{equation}
for some constant \( c_w = c_w(n, \Lambda_0, p) \geq 1 \). Combining Lemma 4.2 and 4.3 yields
\begin{equation}
\int_{K^{4.1\nu}_{\lambda}} |Dv^p| \, dxdt \leq 2^{p-1} \left\{ \int_{K^{4.1\nu}_{\lambda}} |Du - D\bar{v}|^p \, dxdt + \int_{K^{4.1\nu}_{\lambda}} |Dv|^p \, dxdt \right\}
\leq c_\delta \min \{ 1, \frac{\lambda^2}{(\rho+\delta)^2} \} \lambda + c_\Lambda \sigma \lambda + \frac{\lambda}{3},
\end{equation}
Applying Lemma 4.5 with \( \varepsilon \) replaced by \( \varepsilon \), we find a weak solution \( \bar{v} \) of (4.10) such that
\begin{equation}
\int_{K^{4.1\nu}_{\lambda}} |Du - D\bar{v}| \, dxdt \leq \left( \int_{K^{4.1\nu}_{\lambda}} |Dv - D\bar{v}|^p \, dxdt \right)^{\frac{1}{p}} \leq \varepsilon \lambda \leq \frac{\varepsilon}{3} \lambda,
\end{equation}
by selecting \( \bar{\varepsilon} \) with \( 0 < \bar{\varepsilon} \leq \frac{\varepsilon}{3} \). Then we employ (2.14) and (4.14)–(4.16) to estimate
\begin{align*}
\int_{K^{4.1\nu}_{\lambda}} |Du - D\bar{v}| \, dxdt & \leq \int_{K^{4.1\nu}_{\lambda}} |Du - D\bar{v}| + |Dv - D\bar{v}| \, dxdt \\
& \leq c_\delta \min \{ 1, \frac{\lambda^2}{(\rho+\delta)^2} \} \lambda + c_\Lambda \sigma \lambda + \frac{\varepsilon}{3} \lambda \leq \varepsilon \lambda,
\end{align*}
by choosing \( \delta \) sufficiently small.

The second estimate in (4.13) follows from Lemma 4.4 and 4.5. \( \Box \)

Remark 4.7. The second assumption in (4.12) is equivalent to the statement
\begin{equation}
\left[ \frac{|\mu|(K_{4.1\nu}_{\lambda})}{\lambda^{2-p-n+1}} \right]^{\frac{1}{p-1}} \leq \delta \frac{1}{p-1} \lambda.
\end{equation}
In view of (4.12), we see that the value \( \frac{|\mu|(K_{4.1\nu}_{\lambda})}{\lambda^{2-p-n+1}} \) is related to the intrinsic fractional maximal function \( M^*_\lambda(\mu) \), see Section 5 below. The function \( M^*_\lambda(\mu) \) also involves the intrinsic Riesz potential \( I^*_\lambda(\mu) \), see [28, 30]. On the other hand, in view of (4.17), we discover that the value \( \left[ \frac{|\mu|(K_{4.1\nu}_{\lambda})}{\lambda^{2-p-n+1}} \right]^{\frac{1}{p-1}} \) is connected with the intrinsic Wolff potential \( W^*_\lambda(\mu) \), see [31].

5. \( \lambda \)-Covering Arguments

We now consider a SOLA \( u \) of (1.1) and the corresponding weak solutions \( u_m \) (\( m \in \mathbb{N} \)) of (2.9). Suppose that \( (a, \Omega) \) is \( (\delta, R) \)-vanishing. Regarding \( \mu_m \) in (2.9) as an approximation of \( \mu \) in (1.1) via mollification backward in time and then using a suitable cutoff function in time, \( \mu_m \in L^\infty(\Omega_T) \) satisfies the properties (2.11), (2.13) and (4.1); then one can apply all the results obtained in Section 4 to \( u = u_m \) and \( \mu = \mu_m \). Also, we denote by \( w_m, v_m \) and \( \bar{v}_m \), the corresponding weak solutions of (4.3), (4.9) and (4.10), respectively.

Fix any \( \varepsilon \in (0, 1) \) and \( N > 1 \), and set
\begin{equation}
\lambda_0 := \left[ \frac{|\mu|(\Omega_T)}{|\Omega_T|^{\frac{2-d}{n}}} \right]^{\frac{\beta d}{\varepsilon |Q_{R/10}|}} + \left[ \frac{|\mu|(\Omega_T)}{\delta T^{\frac{d-2}{n}}} \right]^d + 1,
\end{equation}
where \( \beta := \frac{n+2}{(n+1)p-n} \), \( \tilde{d} := \max\{1, p-1\} \) and

\[
d := \begin{cases} 
1 & \text{if } p \geq 2, \\
\frac{2}{(n+1)p-2n} & \text{if } 2 - \frac{1}{n+1} < p \leq 2.
\end{cases}
\]

Note that the constants \( \beta, \tilde{d} \) and \( d \) come from Lemma 3.1, 5.1 and 5.4, respectively. Recall from Section 2.1 that \( \Omega_T := \Omega \times (-\infty, T) \) and \( K^\lambda(x, t) := Q^\lambda_T(x, t) \cap \Omega_T \).

We may assume, upon letting \( u \equiv 0 \) for \( t < 0 \), that a SOLA \( u \) is defined in \( \Omega_T \). If \( \lambda \geq \lambda_0 \), then we write

\[
\mathcal{E} := \{(x, t) \in \Omega_T : M^\lambda_{\Omega_T} |Du|(x, t) > \lambda N \}
\]

and

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

where \( M^\lambda_{\Omega_T} \) and \( M^1_{\Omega_T} \) are given by (2.23) and (2.4), respectively. Note that we infer from Lemma 5.1, 5.4 and 5.7 that both the upper level sets \( \mathcal{E} \) and \( \mathcal{D} \) are bounded measurable subsets of \( \Omega_T \).

We now verify two assumptions of the covering lemma (Lemma 2.14).

**Lemma 5.1.** There exists a constant \( N_1 = N_1(n, \Lambda_0, p) > 1 \) such that for any fixed \( N \geq N_1 \) and \( \lambda \geq \lambda_0 \), we have

\[
|\mathcal{E}| < \varepsilon \left| Q^\lambda_{R/10} \right|.
\]

**Proof.** Fix any \( N \geq N_1 \) and \( \lambda \geq \lambda_0 > 1 \). We have from Hölder’s inequality, (2.24) and Lemma 3.1 (with \( \kappa = \tilde{d} := \max\{1, p-1\} \)) that

\[
|\mathcal{E}| := \left| \{ (x, t) \in \Omega_T : M^\lambda_{\Omega_T} |Du| > \lambda N \} \right| \leq \left| \{ (x, t) \in \Omega_T : M^\lambda_{\Omega_T} |Du|^\tilde{d} > (\lambda N)^\tilde{d} \} \right|
\]

\[
\leq \frac{c}{(\tilde{d}^\tilde{d})} \int_{\Omega_T} |Du|^\tilde{d} \, dx \, dt = \frac{c}{(\tilde{d}^\tilde{d})} \int_{\Omega_T} |Du|^\tilde{d} \, dx \, dt \leq \frac{c|\Omega_T|}{\tilde{d}^\tilde{d}} \left[ \frac{\mu(\Omega_T)}{|\Omega_T|^{\frac{\tilde{d}}{\tilde{d}+1}}} \right]^\tilde{d}
\]

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

If \( p \geq 2 \), then we compute from (5.1) that

\[
|\mathcal{E}| \leq \frac{c|\Omega_T|}{(\tilde{d}^{\tilde{d}-p})} \left[ \frac{\mu(\Omega_T)}{|\Omega_T|^{\frac{\tilde{d}}{\tilde{d}+1}}} \right]^\tilde{d}
\]

\[
< \frac{c\lambda^{2-p} |Q_{R/10}|}{N_1^{p-1}} \leq 2 \varepsilon \lambda^{2-p} |Q_{R/10}| = \varepsilon \left| Q^\lambda_{R/10} \right|,
\]

by choosing \( N_1 \) sufficiently large.

If \( 2 - \frac{1}{n+1} < p \leq 2 \), then we note that \( \lambda^{-1} < \lambda_0^{-1} < \lambda^{2-p} \lambda_0 \). Therefore

\[
|\mathcal{E}| \leq \frac{c|\Omega_T|}{\lambda^{p-1}} \left[ \frac{\mu(\Omega_T)}{|\Omega_T|^{\frac{p}{p+1}}} \right]^\beta
\]

\[
< \frac{c\lambda^{2-p} \varepsilon |Q_{R/10}|}{N_1} \leq \varepsilon \left| Q^\lambda_{R/10} \right|,
\]

by selecting \( N_1 \) large enough. \( \square \)
Lemma 5.2. There exists $N_2 = N_2(n, \Lambda_0, \Lambda_1, p) > 1$ so that for any $\varepsilon \in (0, 1)$, there is $\delta = \delta(n, \Lambda_0, \Lambda_1, p, \varepsilon) \in (0, \frac{1}{3})$ such that the following holds: for any fixed $\lambda \geq \lambda_0$, $N \geq N_2$, $r \in (0, \frac{1}{4\lambda})$ and $(y, s) \in \Omega_{\Sigma}$ with

\begin{equation}
|\xi \cap Q_\lambda^\lambda(y, s)| \geq \varepsilon |Q_\lambda^\lambda|,
\end{equation}

we have

\[ K_\lambda^\lambda(y, s) \subset \mathcal{D}. \]

Proof. We assume to the contrary that $K_\rho^\lambda(y, s) \not\subset \mathcal{D}$ and derive a contradiction. Suppose that there is a point $(\tilde{x}, \tilde{t}) \in K_\lambda^\lambda(y, s)$ such that for all $\rho > 0$,

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

We recall from (2.12) that $|K_\rho^\lambda(y, s)| < K_{\rho+r}^\lambda(\tilde{x}, \tilde{t})$ for any $\rho > r$. Then it follows from (2.14) and (5.4) that

\[
\int_{K_\rho^\lambda(\tilde{x}, \tilde{t})} |Du| \, dx \, dt \leq 2 \left( \frac{16}{7} \right)^n \frac{1}{|Q_\rho^\lambda|} \int_{K_\rho^\lambda(\tilde{x}, \tilde{t})} |Du| \, dx \, dt 
\leq 2 \left( \frac{16}{7} \right)^n \left( \frac{9}{8} \right)^{n+2} \frac{1}{|Q_{\rho+r}|} \int_{K_{\rho+r}^\lambda(\tilde{x}, \tilde{t})} |Du| \, dx \, dt 
\leq \left( \frac{18}{7} \right)^{n+2} \lambda
\]

for every $\rho \geq 8r$. Then the resulting inequality and (2.10) imply that for any $\tilde{\varepsilon} \in (0, 1)$, we have

\[
\int_{K_\rho^\lambda(\tilde{x}, \tilde{t})} |Du_m| \, dx \, dt \leq \int_{K_\rho^\lambda(\tilde{x}, \tilde{t})} |Du| \, dx \, dt + \int_{K_\rho^\lambda(\tilde{x}, \tilde{t})} |Du - Du_m| \, dx \, dt 
\leq \left( \frac{18}{7} \right)^{n+2} + \tilde{\varepsilon} \leq 2 \left( \frac{18}{7} \right)^{n+2} \lambda =: c_2 \lambda
\]

whenever $m$ is large enough. Likewise, we recall (4.2) with $\mu$ replaced by $\mu_m$, and then combine (5.4) and (2.11) to deduce

\[
\frac{|\mu_m|(K_\rho^\lambda(\tilde{x}, \tilde{t}))}{\rho^{n+1}} \leq \frac{|\mu|(K_\rho^\lambda(\tilde{x}, \tilde{t}))}{\rho^{n+1}} + \frac{\varepsilon}{\rho^{n+1}} 
\leq \left( \frac{9}{8} \right)^{n+1} + \frac{\varepsilon}{\rho^{n+1}} \leq c_2 \delta \lambda
\]

for any $\rho \geq 8r$, by selecting $\varepsilon$ sufficiently small. Now applying Lemma 4.6 with $(x_0, t_0)$, $\lambda$, $r$ and $\varepsilon$ replaced by $(y, s)$, $c_2 \lambda$, $4r$ and $\eta$, respectively, we can find $\delta = \delta(n, \Lambda_0, \Lambda_1, p, \eta) > 0$ such that

\begin{equation}
\int_{K_{\rho+r}^\lambda(\tilde{x}, \tilde{t})} |Du_m - D\tilde{v}_m| \, dx \, dt \leq c_2 \eta \lambda \quad \text{and} \quad \|D\tilde{v}_m\|_{L^\infty(K_{\rho+r}^\lambda(\tilde{x}, \tilde{t}))} \leq cc_2 \lambda =: c_3 \lambda,
\end{equation}

for some constant $c_3 = c_3(n, \Lambda_0, \Lambda_1, p) \geq 1$. Combining (5.5) and (2.10) yields

\begin{equation}
\int_{K_{\rho+r}^\lambda(\tilde{x}, \tilde{t})} |Du - D\tilde{v}_m| \, dx \, dt \leq c_4 \eta \lambda.
\end{equation}
We next show that

\[ \{ (x, t) \in K_r(y, s) : |Du| > N\lambda \} \]

(5.7)

\[ \subset \bigg \{ (x, t) \in K_r(y, s) : M_{\Omega_{1+2}}^r |Du - D\bar{v}_m| > \lambda \} \]

provided \( N \geq N_2 := \max \{ 2^{n+2}, 1 + c_3 \} \). To do so, let

\[ (\tilde{y}, \tilde{s}) \in \bigg \{ (x, t) \in K_r(y, s) : M_{\Omega_{1+2}}^r |Du - D\bar{v}_m| \leq \lambda \} \].

Then for any \( \tilde{r} > 0 \),

\[ \frac{1}{|Q_r^2|} \int_{K_1^r(\tilde{y}, \tilde{s}) \cap K_1^r(y, s)} |Du - D\bar{v}_m| \, dx \, dt \leq \lambda. \]  

(5.8)

If \( \tilde{r} \in (0, 2r] \), then \( K_1^r(\tilde{y}, \tilde{s}) \subset K_{3r}^r(y, s) \), and so we have from (5.8) and (5.5) that

\[ \frac{1}{|Q_r^2|} \int_{K_1^r(\tilde{y}, \tilde{s})} |Du| \, dx \, dt \leq \frac{1}{|Q_r^2|} \int_{K_1^r(\tilde{y}, \tilde{s})} |Du - D\bar{v}_m| + |D\bar{v}_m| \, dx \, dt \leq (1 + c_3) \lambda. \]

On the other hand, if \( \tilde{r} > 2r \), then \( K_1^r(\tilde{y}, \tilde{s}) \subset K_{3r}^r(y, s) \subset K_{2r}(\tilde{x}, \tilde{t}) \): so the first inequality of (5.4) implies

\[ \frac{1}{|Q_r^2|} \int_{K_1^r(\tilde{y}, \tilde{s})} |Du| \, dx \, dt \leq \frac{1}{|Q_r^2|} \int_{K_1^r(\tilde{x}, \tilde{t})} |Du| \, dx \, dt \leq 2^{n+2} \lambda. \]

Taking \( N_2 = \max \{ 2^{n+2}, 1 + c_3 \} \), we obtain

\[ (\tilde{y}, \tilde{s}) \in \bigg \{ (x, t) \in K_r(y, s) : M_{\Omega_{1+2}}^r |Du| \leq N\lambda \} \]

which implies (5.7).

We now employ (5.7), (2.24) and (5.6) to conclude

\[ \bigg | \bigg \{ (x, t) \in K_r(y, s) : M_{\Omega_{1+2}}^r |Du| > N\lambda \bigg \} \bigg | \]

\[ \leq \bigg | \bigg \{ (x, t) \in K_r(y, s) : M_{\Omega_{1+2}}^r |Du - D\bar{v}_m| > \lambda \bigg \} \bigg | \]

\[ \leq \frac{c}{\lambda} \int_{K_1^r(\tilde{x}, \tilde{t})} |Du - D\bar{v}_m| \, dx \, dt \leq cc_4 \eta |Q_r^2| < \epsilon |Q_r^2|, \]

by selecting \( \eta \) that satisfies the last inequality above. This is a contradiction to (5.3).

When taking \( N = \max \{ N_1, N_2 \} \) from Lemma 5.1 and 5.2, we can apply Lemma 2.14 to obtain the following decay estimate:

**Corollary 5.3.** Under the assumptions as in Lemma 5.1 and 5.2, we have

\[ \big | \big \{ (x, t) \in \Omega_T : M_{\Omega_{1+2}}^r |Du| > \lambda N \big \} \big | \]

\[ \leq \varepsilon_0 \left( \left( \frac{1}{\mu} \right)^{n+2} \right) \]

\[ \leq \varepsilon_0 \left( \left( \frac{1}{\mu} \right)^{n+2} \right) \]

\[ \leq \varepsilon_0 \left( \left( \frac{1}{\mu} \right)^{n+2} \right) \]

where \( \varepsilon_0 := \left( \frac{1}{\mu} \right)^{n+2} \).

Recalling (2.2) and (2.4), we obtain the following relation:

**Lemma 5.4.** Let \( \lambda \geq 1 \). Then we have

\[ \big \{ (x, t) \in \Omega_T : M_{\Omega_{1+2}}^r |Du| > \lambda \big \} \subset \big \{ (x, t) \in \Omega_T : |M_1(\mu) - \mu| > \delta \lambda \big \}, \]

(5.9) where the constant \( d \) is given in (5.2).
Proof. Let \( \lambda \geq 1 \). If \( p \geq 2 \), then we have \( Q_\lambda^r(x,t) \subset Q_r(x,t) \). It follows from the definitions (2.2) and (2.4) that \( M^1_\lambda \mu(x,t) \leq M_1 \mu(x,t) \) for all \( (x,t) \in \mathbb{R}^{n+1} \), which implies that (5.9) holds for \( p \geq 2 \).

Assume \( 2 - \frac{1}{n+1} < p \leq 2 \). If \( (y,s) \in \{ (x,t) \in \Omega_T : M^1_\lambda(\mu) > \lambda \} \), then we have

\[
\frac{|\mu|(Q_\lambda^r(y,s))}{r^{n+1}} > \lambda
\]

for any \( r > 0 \). Since \( Q_\lambda^r(y,s) \subset Q_{\lambda^{\frac{2-p}{r}}}^r(y,s) \), we have

\[
\frac{|\mu|(Q_\lambda^r(y,s))}{r^{n+1}} \leq \frac{|\mu|(Q_{\lambda^{\frac{2-p}{r}}}^r(y,s))}{r^{n+1}} = \left[ \frac{|\mu|(Q_{\lambda^{\frac{2-p}{r}}}^r(y,s))}{(\lambda^{\frac{2-p}{r}})^{n+1}} \right]^{\frac{n+1}{2-p}}.
\]

By setting \( r_0 = \lambda^{\frac{2-p}{r}} \), we discover

\[
\frac{|\mu|(Q_\lambda^r(y,s))}{r_0^{n+1}} > \lambda^{\frac{(n+1)(2-p)-2n}{2}},
\]

which implies \([M_1(\mu)]^d(y,s) > \lambda \); that is, \((y,s) \in \{ (x,t) \in \Omega_T : [M_1(\mu)]^d > \lambda \} \).

Assertion (5.9) follows.

If the given measure \( \mu \) in \( \Omega_T \) can be decomposed into a finite signed Radon measure \( \mu_0 \) in \( \Omega \) and a Lebesgue function \( f \) in \((\Omega \times (T,\infty)), \), we write \( \mu = \mu_0 \otimes f \).

**Lemma 5.5.** Let \( \lambda \geq 1 \). If \( \mu = \mu_0 \otimes f \), then we have

\[
\{ (x,t) \in \Omega_T : M_1^\lambda(\mu) > \lambda \}
\]

(5.10)

\[
\subset \{ (x,t) \in \Omega_T : [2(M_1(\mu_0))(Mf)]^{\frac{1}{2-p}} > \lambda \},
\]

where the maximal function \( Mf \) is defined by

\[
Mf(t) := \sup_{r > 0} \frac{1}{r} \int_{t-r}^{t+r} |f(s)| \, ds = \sup_{r > 0} \frac{1}{2r} \int_{t-r}^{t+r} |f(s)| \, ds.
\]

**Proof.** Let \( \lambda \geq 1 \). If \( \mu = \mu_0 \otimes f \), then it follows

\[
\frac{|\mu|(Q_\lambda^r(x,t))}{\lambda^{2-p} r^{n+1}} = 2 \frac{|\mu_0|(B_r(x))}{r^{n-1}} \int_{t-\lambda^{2-p} r^2}^{t+\lambda^{2-p} r^2} |f(s)| \, ds.
\]

Taking the supremum with respect to \( r \), we find

\[
\lambda^{p-2} M_1^\lambda(\mu)(x,t) \leq 2 M_1(\mu_0)(x) Mf(t).
\]

If \( M_1^\lambda(\mu) > \lambda \), then we see that \( 2(M_1(\mu_0))(Mf) > \lambda^{p-1} \), which implies (5.10).

**Remark 5.6.** The difference between the standard parabolic cylinder \( Q_r(x,t) \) and the intrinsic parabolic cylinder \( Q_\lambda^r(x,t) \) causes the appearance of the deficit constant \( d \) as in (5.2). This constant determines the exponent in our main estimates of Theorem 2.4. However if \( \mu = \mu_0 \otimes f \), then we can derive more natural estimates without the deficit constant \( d \), see Theorem 2.8 and Remark 2.9.

Recall from Section 2.1 that \( \Omega_T := \Omega \times (0,T) \) and \( \Omega_T := \Omega \times (-T,T) \).
Lemma 5.7. Let $\lambda \geq \lambda_0$, where $\lambda_0$ be given in (5.1). Then we have
\begin{equation}
(x, t) \in \Omega_T : [M_1(\mu)]^d > \delta^d \lambda \}
\end{equation}
where the constant $d$ is given in (5.2).

Proof. Suppose $(x_0, t_0) \in \{(x, t) \in \Omega_T : [M_1(\mu)]^d > \delta^d \lambda \}$. Recall that the measure $\mu$ vanishes outside the domain $\Omega_T$. If $t_0 \leq -T$, then we see
\[ M_1(\mu)(x_0, t_0) = \sup_{r > \sqrt{T}} \frac{|\mu|(Q_r(x_0, t_0))}{r^{n+1}} \leq \frac{|\mu|(\Omega_T)}{T^{\frac{n}{p-1}}}. \]
Since $[M_1(\mu)(x_0, t_0)]^d > \delta^d \lambda$, we have
\[ \left[ \frac{|\mu|(\Omega_T)}{T^{\frac{n}{p-1}}} \right]^d > \delta^d \lambda, \]
a contradiction to our assumption $\lambda \geq \lambda_0$. Hence $-T < t_0 < T$, and so the assertion (5.12) follows. \hfill $\Box$

Remark 5.8. The measure $\mu$ vanishes outside $\Omega_T$, but $M_1(\mu)$ need not be zero in $\mathbb{R}^{n+1} \setminus \Omega_T$. Thus, the boundedness of the set $\{(x, t) \in \Omega_T : [M_1(\mu)]^d > \delta^d \lambda \}$ is important to derive the estimates in Theorem 2.4, see Section 6 below.

In the case that $\mu = \mu_0 \otimes f$, on the other hand, we do not need to know the boundedness of the set $\{(x, t) \in \Omega_T : 2(M_1(\mu_0))(\mathcal{M}_f)^{\frac{1}{p-1}} > \delta^\frac{1}{p-1} \lambda \}$, see (6.13) below. Therefore, in this case, we can drop the condition $\lambda_0 \geq \left[ \frac{|\mu|(\Omega_T)}{\delta T^\frac{n}{p-1}} \right]^d$ in (5.1); that is, if $\mu = \mu_0 \otimes f$, then we take
\begin{equation}
\lambda_0 := \left[ \frac{|\mu_0|(\Omega)}{\Omega_T^\frac{n+2}{p-1}} \right]^\beta d \frac{|\Omega_T|}{\varepsilon |Q_{R/10}|} + 1,
\end{equation}
where $\beta := \frac{n+2}{(n+1)p-n}$ and $d := \max\{1, p-1\}$.

Combining Corollary 5.3, Lemma 5.4, 5.5 and 5.7, we finally obtain

Lemma 5.9. Let $N = \max \{N_1, N_2\}$ from Lemma 5.1 and 5.2. Then for any $\varepsilon \in (0, 1)$, there exists $\delta = \delta(n, \lambda_0, \lambda_1, p, \varepsilon) \in (0, \frac{1}{2})$ such that if $(a, \Omega)$ is $(\delta, R)$-vanishing for some $R > 0$, then for any SOLA $u$ of (1.1) and any $\lambda \geq \lambda_0$, we have
\[
\left| \left\{ (x, t) \in \Omega_T : M_{\text{\Omega}_T}^\lambda |Du| > N \lambda \right\} \right| \leq \varepsilon_0 \left| \left\{ (x, t) \in \Omega_T : M_{\text{\Omega}_T}^\lambda |Du| > \lambda \right\} \right| \\
+ \varepsilon_0 \left| \left\{ (x, t) \in \Omega_T : [M_1(\mu)]^d > \delta^d \lambda \right\} \right|
\]
where $\varepsilon_0 := \left( \frac{80}{T} \right)^{n+2} \varepsilon$, $\Omega_T := \Omega \times (-T, T)$, and $M_{\text{\Omega}_T}^\lambda$, $M_1(\mu)$, $d$ are given by (2.23), (2.2), (5.2), respectively.

Furthermore, if $\mu = \mu_0 \otimes f$, then we have
\[
\left| \left\{ (x, t) \in \Omega_T : M_{\text{\Omega}_T}^\lambda |Du| > N \lambda \right\} \right| \\
\leq \varepsilon_0 \left| \left\{ (x, t) \in \Omega_T : M_{\text{\Omega}_T}^\lambda |Du| > \lambda \right\} \right| \\
+ \varepsilon_0 \left| \left\{ (x, t) \in \Omega_T : [M_1(\mu_0)](\mathcal{M}_f)^{\frac{1}{p-1}} > \delta^\frac{1}{p-1} \lambda \right\} \right|
\]
where $M_1(\mu_0)$ and $\mathcal{M}_f$ are given by (2.3) and (5.11), respectively.
6. Global Calderón-Zygmund type estimates for parabolic measure data problems

We devote this section to proving Theorem 2.4 and 2.8. Recall from Section 2.1 that \( \Omega_T = \Omega \times (0, T), \Omega_{\tau} = \Omega \times (-\infty, \tau) \) and \( \Omega_{T_{\tau}} = \Omega \times (-\tau, \tau) \). Assume that \((a, \Omega)\) is \((\delta, R)\)-vanishing. Fix any \( \varepsilon \in (0, 1) \). Let \( N = N(n, \Lambda_0, \Lambda_1, p) > 1 \) be a given constant in Lemma 5.9. We suppose, upon letting \( u \equiv 0 \) for \( t < 0 \), that a SOLA \( u \) of (1.1) is defined in \( \Omega_T \). We also assume that the measure \( \mu \) is defined in \( \mathbb{R}^{n+1} \) by considering the zero extension to \( \mathbb{R}^{n+1} \). 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 extend both \( \mu_0 \) and \( f \) to be 0 outside \( \Omega \) and \((0, T)\), respectively.

We first derive the following decay estimates of integral type:

**Lemma 6.1.** Let \( 1 < \kappa_0 < p - \frac{n}{n+1} \) and let \( \lambda \geq \lambda_0 \), where \( \lambda_0 \) be given in (5.1). If \( u \) is a SOLA of (1.1), then there exists a constant \( c = c(n, \Lambda_0, \Lambda_1, p, \kappa_0) \geq 1 \) such that

\[
\int_{\{ (x,t) \in \Omega_T : |Du| > N\lambda \}} |Du|^\kappa_0 \, dx \, dt 
\leq c \varepsilon \int_{\{ (x,t) \in \Omega_T : |Du| > \frac{N}{2} \}} |Du|^\kappa_0 \, dx \, dt 
\]

\[
+ \frac{c \varepsilon}{d \delta \kappa_0} \int_{\{ (x,t) \in \Omega_T : |M_1(\mu)|^d > \delta \lambda \}} |M_1(\mu)|^d \kappa_0 \, dx \, dt,
\]

where \( M_1(\mu) \) and the constant \( d \) are given in (2.2) and (5.2), respectively.

**Proof.** We first recall from Lemma 3.1 that \( Du \in L^{\kappa_0}(\Omega_{\tau}) \). Having in mind that

\[
M_{\Omega_{\tau}}^d |Du|(x,t) \geq |Du|(x,t)
\]

for any \( \tau > 0 \) and \( (x,t) \in \Omega_{\tau} \), we compute

\[
\int_{\{ (x,t) \in \Omega_{\tau} : |Du| > N\lambda \}} |Du|^\kappa_0 \, dx \, dt 
= \kappa_0 \int_{\lambda}^{\infty} \tau^{\kappa_0-1} \left| \{ (x,t) \in \Omega_{\tau} : |Du| > N\tau \} \right| \, d\tau 
+ (N\lambda)^{\kappa_0} \left| \{ (x,t) \in \Omega_{\tau} : |Du| > N\lambda \} \right|
\leq \kappa_0 \int_{\lambda}^{\infty} \tau^{\kappa_0-1} \left| \{ (x,t) \in \Omega_{\tau} : M_{\Omega_{\tau}}^d |Du| > N\tau \} \right| \, d\tau 
\]

\[
= : J_1
\]

\[
+ (N\lambda)^{\kappa_0} \left| \{ (x,t) \in \Omega_{\tau} : M_{\Omega_{\tau}}^d |Du| > N\lambda \} \right|
\leq J_2
\]
Applying Lemma 5.9 and (2.25), we deduce
\[ J_1 \leq \varepsilon_0 \int_{\lambda}^{\infty} \tau^{-\kappa_0 - 1} \left| \left\{ (x, t) \in \Omega_T : M'_{\Omega_T} |Du| > \tau \right\} \right| d\tau \\
+ \varepsilon_0 \int_{\lambda}^{\infty} \tau^{-\kappa_0 - 1} \left| \left\{ (x, t) \in \Omega_T : |M_1(\mu)|^d > \delta^d \tau \right\} \right| d\tau \\
\leq c(n) \varepsilon_0 \int_{\lambda}^{\infty} \tau^{-\kappa_0 - 2} \left[ \int_{\{x, t \in \Omega_T : |Du| > \frac{\tau}{2}\}} |Du| dt dx \right] d\tau \\
+ \varepsilon_0 \int_{\lambda}^{\infty} \tau^{\kappa_0 - 1} \left| \left\{ (x, t) \in \Omega_T : |M_1(\mu)|^d > \delta^d \tau \right\} \right| d\tau, 
\]
where \( \varepsilon_0 = \left( \frac{\mu_0}{\xi} \right)^{n+2} \varepsilon \). Furthermore we see from Fubini’s theorem that
\[
\int_{\lambda}^{\infty} \tau^{\kappa_0 - 2} \left[ \int_{\{x, t \in \Omega_T : |Du| > \frac{\tau}{2}\}} |Du| dt dx \right] d\tau \\
= \int_{\{x, t \in \Omega_T : |Du| > \frac{\tau}{2}\}} \left[ \int_{\lambda}^{\infty} \tau^{\kappa_0 - 2} d\tau \right] |Du|^{\kappa_0} dt dx \\
\leq \frac{2^{\kappa_0 - 1}}{\kappa_0 - 1} \int_{\{x, t \in \Omega_T : |Du| > \frac{\tau}{2}\}} |Du|^{\kappa_0} dt dx,
\]
and that
\[
\int_{\lambda}^{\infty} \tau^{\kappa_0 - 1} \left| \left\{ (x, t) \in \Omega_T : |M_1(\mu)|^d > \delta^d \tau \right\} \right| d\tau \\
\leq \frac{1}{\kappa_0 \delta^{d\kappa_0}} \int_{\{x, t \in \Omega_T : |M_1(\mu)|^d > \delta^d \lambda\}} |M_1(\mu)|^{d\kappa_0} dt dx.
\]
On the other hand, it follows from Hölder’s inequality, Lemma 5.9 and (2.25) that
\[
J_2 \leq \varepsilon_0 \left| \{ (x, t) \in \Omega_T : M'_{\Omega_T} |Du| > \lambda \} \right| + \varepsilon_0 \left| \{ (x, t) \in \Omega_T : |M_1(\mu)|^d > \delta^d \lambda \} \right| \\
\leq \varepsilon_0 \int_{\{x, t \in \Omega_T : |Du| > \frac{\tau}{2}\}} |Du|^{\kappa_0} dt dx \\
+ \varepsilon_0 \int_{\{x, t \in \Omega_T : |M_1(\mu)|^d > \delta^d \lambda\}} |M_1(\mu)|^{d\kappa_0} dt dx.
\]
Combining the inequalities above, we obtain the estimate (6.1).

The following result may be proved in much the same way as Lemma 6.1.

**Lemma 6.2.** Let \( 1 < \kappa_0 < p - \frac{1}{n+1} \) and let \( \lambda \geq \lambda_0 \), where \( \lambda_0 \) be given in (5.13). If \( u \) is a SOLA of (1.1) and \( \mu = \mu_0 \otimes f \), then there exists a constant \( c = c(n, \Lambda_0, \Lambda_1, \beta, \kappa_0) \geq 1 \) such that
\[
\int_{\{x, t \in \Omega_T : |Du| > N\lambda\}} |Du|^{\kappa_0} dt dx \\
\leq c \varepsilon \int_{\{x, t \in \Omega_T : |Du| > \frac{\tau}{2}\}} |Du|^{\kappa_0} dt dx \\
+ \varepsilon_0 \int_{\{x, t \in \Omega_T : |2(M_1(\mu_0))(M_f)|^{\frac{1}{\nu}} > \delta^d \lambda\}} |(M_1(\mu_0))(M_f)|^{\frac{\kappa_0}{\nu_0}} dt dx.
\]
where the operators $M_1(\mu_0)$ and $Mf$ are given in (2.3) and (5.11), respectively.

**Remark 6.3.** Both the constants $c$ in (6.1) and (6.3) blow up as $\kappa_0 \to 1$, see (6.2).

We need a technical assertion in the proof of Theorem 2.4.

**Lemma 6.4.** For each $(x, t) \in \Omega_T$ we have

\[(6.4) \quad M_1(\mu)(x, t) - M_1(\mu) < \frac{|\mu|(Q_{r_0}(x, t))}{r_n^{n+1}} + \epsilon.
\]

**Proof.** Let $(x, t) \in \Omega_T$. From the definition (2.2), for every $\epsilon > 0$ there exists $r_0 > 0$ such that

$$M_1(\mu)(x, t) < \frac{|\mu|(Q_{r_0}(x, t))}{r_n^{n+1}} + \epsilon.$$ 

Since $Q_{r_0}(x, t) \cap \Omega_T \subset Q_{r_0}(x, t) \cap \Omega_T$, we have

$$|\mu|(Q_{r_0}(x, t)) = |\mu|(Q_{r_0}(x, t) \cap \Omega_T) \leq |\mu|(Q_{r_0}(x, t) \cap \Omega_T),$$

which implies

$$M_1(\mu)(x, t) < \frac{|\mu|(Q_{r_0}(x, t) \cap \Omega_T)}{r_n^{n+1}} + \epsilon \leq M_1(\mu)(x, t) + \epsilon.$$ 

Letting $\epsilon \to 0$, we obtain (6.4). \hfill \Box

### 6.1. Proof of Theorem 2.4.

**Proof of Theorem 2.4.** Let $q > 1$ and let $\kappa_0 = \kappa_0(n, p, q)$ be an arbitrary constant with $1 < \kappa_0 < \min \left\{ \frac{n}{n+1}, q \right\}$. Given $k \in \mathbb{N}$, we define the truncation of the function $|Du|$ to be

$$|Du|_k := \min\{|Du|, k\}.$$ 

From Lemma 2.13, we compute for any $k > N\lambda_0$

\begin{align*}
\int_{\Omega_T} |Du|^q - \kappa_0 |Du|^\kappa_0 d\lambda d\tau &\leq (q - \kappa_0) N^{q - \kappa_0} \int_0^\lambda \int_{\{x, t\} \in \Omega_T : |Du| > N\lambda} |Du|^\kappa_0 d\lambda d\tau \\
&\leq c \int_0^\lambda \int_{\{x, t\} : |Du| > N\lambda} |Du|^\kappa_0 d\lambda d\tau + c \int_0^\lambda \int_{\{x, t\} \in \Omega_T : |Du| > N\lambda \cap \Omega_T} |Du|^\kappa_0 d\lambda d\tau.
\end{align*}

for some constant $c = c(n, \Lambda_0, \Lambda_1, p, q) \geq 1$, where $\lambda_0$ be given in (5.1). It follows from Lemma 3.1 that

\[(6.5) \quad P_1 \leq \int_0^{\lambda_0} \lambda^{q - \kappa_0} d\lambda \int_{\Omega_T} |Du|^\kappa_0 d\lambda d\tau \leq \frac{\lambda_0^{q - \kappa_0}}{q - \kappa_0} \int_{\Omega_T} |Du|^\kappa_0 d\lambda d\tau \leq c\lambda_0^{q - \kappa_0} |\mu|(|\Omega_T|)^{\beta\kappa_0}.
\]
for some constant $c = c(n, \Lambda_0, p, q, \Omega_T) \geq 1$, where $\beta := \frac{(n+2)}{(n+2-p)}$. Furthermore we see from Lemma 6.1 that

$$P_2 \leq c \varepsilon \int_{\lambda_0}^{\hat{\lambda}} \lambda^{q-\kappa_0-1} \left[ \int_{\{x,t\in\Omega_T:|Du|>\hat{\lambda}\}} |Du|^{\kappa_0} \, dx \, dt \right] \, d\lambda =: S_1$$

and similarly,

$$S_2 \leq \int_{\Omega_T} \left[ \int_{\{x,t\in\Omega_T:|M_1(\mu)|^d>\hat{\lambda}\}} [M_1(\mu)]^{d\kappa_0} \, dx \, dt \right] \, d\lambda.$$  

(6.6)

Therefore,

$$P_2 \leq c \varepsilon \left\{ \int_{\Omega_T} |Du|^{q-\kappa_0}_k |Du|^{\kappa_0} \, dx \, dt + \frac{1}{\delta^{dq}} \int_{\Omega_T} [M_1(\mu)]^{dq} \, dx \, dt \right\}$$  

(6.7)

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

We employ (6.5) and (6.7) to derive that

$$\int_{\Omega_T} |Du|^{q-\kappa_0}_k |Du|^{\kappa_0} \, dx \, dt \leq c_0 \varepsilon \int_{\Omega_T} |Du|^{q-\kappa_0}_k |Du|^{\kappa_0} \, dx \, dt$$

$$+ \frac{c \varepsilon}{\delta^{dq}} \int_{\Omega_T} [M_1(\mu)]^{dq} \, dx \, dt + c \lambda^{q-\kappa_0}_0 \|\mu\|_{(\Omega_T)}^{\beta \kappa_0}$$

for some $c_0 = c_0(n, \Lambda_0, \Lambda_1, p, q) \geq 1$. Now, we take $\varepsilon > 0$ so small that $c_0 \varepsilon < 1$, and then we can determine a corresponding $\delta = \delta(n, \Lambda_0, \Lambda_1, p, q) > 0$. Letting $k \to \infty$,
we obtain

\[ \int_{\Omega_T} |Du|^q \, dx \, dt = \int_{\Omega_T} |Du|^q \, dx \, dt \]

(6.8)

\[ \leq c \int_{\Omega_T} [\mathcal{M}_1(\mu)]^{dq} \, dx \, dt + c\lambda_0^{\beta - \kappa_0} \left[ \mu(\Omega_T) \right]^{\beta \kappa_0}, \]

where the equality above have used the fact that \( u \equiv 0 \) for \( t \leq 0 \).

Moreover recalling (5.1), we deduce

\[ \lambda_0 \leq \begin{cases} \left[ \mu(\Omega_T) \right]^{\beta(p-1)} + 1 & \text{if } p \geq 2, \\ \left[ \mu(\Omega_T) \right]^{\frac{\beta p}{(n+1)p-2n}} + 1 & \text{if } 2 - \frac{1}{n+1} < p \leq 2 \end{cases} \]

for some constant \( c = c(n, \Lambda_0, \Lambda_1, p, q, R, \Omega_T) \geq 1 \), where we have used the fact that \( \beta(p-1) \geq 1 \) for \( p \geq 2 \), and that \( \beta < \frac{2}{(n+1)p-2n} \) for \( 2 - \frac{1}{n+1} < p \leq 2 \). Consequently,

\[ \lambda_0^{\beta - \kappa_0} \left[ \mu(\Omega_T) \right]^{\beta \kappa_0} \leq \begin{cases} \left[ \mu(\Omega_T) \right]^{\beta(p-1)q} + 1 & \text{if } p \geq 2, \\ \left[ \mu(\Omega_T) \right]^{\frac{2p}{n+1}p - 1} + 1 & \text{if } 2 - \frac{1}{n+1} < p \leq 2 \end{cases} \]

(6.9)

On the other hand, Lemma 6.4 implies

\[ \int_{\Omega_T} [\mathcal{M}_1(\mu)]^{dq} \, dx \, dt \leq 2 \int_{\Omega_T} [\mathcal{M}_1(\mu)]^{dq} \, dx \, dt. \]

(6.10)

Combining (6.8)–(6.10), we finally obtain the desired estimates (2.15) and (2.16). This completes the proof. \( \square \)

6.2. **Proof of Theorem 2.8.**

**Proof of Theorem 2.8.** Let \( q > \max \{1, p-1\} =: \tilde{d} \) and let \( \kappa_0 \) be an arbitrary constant with \( 1 < \kappa_0 < \min \left\{ p - \frac{n}{n+1}, q \right\} \). Proceeding as in the proof of Theorem 2.4 and using Lemma 6.2, we infer

\[ \int_{\Omega_T} |Du_k|^{q-\kappa_0} |Du|^{\kappa_0} \, dx \, dt \leq c \varepsilon \int_{\Omega_T} |Du_k|^{q-\kappa_0} |Du|^{\kappa_0} \, dx \, dt 
+ c\lambda_0^{\beta - \kappa_0} \left[ \mu(\Omega) \right]^{\beta \kappa_0} + \frac{c\varepsilon}{\delta^{\kappa_0}} S_3. \]

(6.11)

Here \( |Du_k| := \min\{|Du|, k\} \), \( \beta := \frac{(n+2)}{(n+1)p-n} \), \( \lambda_0 \) is given in (5.13) and

\[ S_3 := \frac{\sqrt{\beta - \kappa_0}}{\lambda_0} \left[ \int_{\{x,t \in \Omega_T : 2[\mathcal{M}_1(\mu_0)](\mathcal{M}f) \sqrt{\beta - \kappa_0} > \phi \}} \left[ \mathcal{M}_1(\mu_0) \mathcal{M}f \right] \frac{1}{\sqrt{\beta - \kappa_0}} \, dx \, dt \right] \right] d\lambda. \]
where $\mathcal{M}_1(\mu_0)$ and $Mf$ are given in (2.3) and (5.11), respectively. By calculation as in (6.6), we similarly derive

$$S_3 \leq \int_{\Omega_T} \left[ \int_0^1 \frac{1}{(\mathcal{M}_1(\mu_0))(Mf)^{\frac{q}{p-1}}} \, \lambda^{q-\kappa_0-1} \, d\lambda \right] \left[ (\mathcal{M}_1(\mu_0))(Mf)^{\frac{\kappa_0}{q}} \right] dx \, dt$$

(6.12)

$$\leq \frac{c}{\delta^{q-\kappa_0}} \int_{\Omega_T} \left[ (\mathcal{M}_1(\mu_0))(Mf)^{\frac{q}{p-1}} \right] dx \, dt$$

$$= \frac{c}{\delta^{q-\kappa_0}} \int_\Omega [\mathcal{M}_1(\mu_0)]^{\frac{q}{p-1}} dx \int_T^{T,\Omega} (Mf)^{\frac{q}{p-1}} \, dt.$$

Applying the strong $\left( \frac{q}{p-1}, \frac{q}{p-1} \right)$-estimate for the function $f$ (see for instance [45, Chapter I, Theorem 1]), we find

$$\int_{-\infty}^T (Mf)^{\frac{q}{p-1}} \, dt \leq c(n, p, q) \int_{-\infty}^T |f|^{\frac{q}{p-1}} \, dt = c \int_{0}^{T} |f|^{\frac{q}{p-1}} \, dt,$$

where we have used the condition that $q > p - 1$.

We employ (6.11)–(6.13) and then select $\varepsilon > 0$ sufficiently small (with $\delta = \delta(n, \Lambda_0, \Lambda_1, p, q) > 0$ being determined), to discover

$$\int_{\Omega_T} |[Du]_{k}^{q-\kappa_0}| [Du]^{\kappa_0} \, dx \, dt \leq c \int_{\Omega_T} \left[ (\mathcal{M}_1(\mu_0))f \right]^{\frac{q}{p-1}} \, dx \, dt$$

(6.14)

$$+ c\lambda_0^{q-\kappa_0} \left[ [\mu_0](\Omega) \|f\|_{L^1(0, T)} \right]^{\beta_{\kappa_0}}.$$

On the other hand, (5.13) implies

$$\lambda_0^{q-\kappa_0} \left[ [\mu_0](\Omega) \|f\|_{L^1(0, T)} \right]^{\beta_{\kappa_0}} \leq c \left\{ \left[ [\mu_0](\Omega) \|f\|_{L^1(0, T)} \right]^{\beta_{\kappa_0}} \right\}.$$

(6.15)

Finally, combining (6.14)–(6.15) and letting $k \to \infty$, we obtain the desired estimate (2.21). This completes the proof.

\section*{References}

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

[2] D. R. Adams and L. I. Hedberg, \textit{Function spaces and potential theory}, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 314, Springer-Verlag, Berlin, 1996.

[3] P. Baroni, \textit{Marcinkiewicz estimates for degenerate parabolic equations with measure data}, J. Funct. Anal. \textbf{267} (2014), no. 9, 3397–3426.

[4] P. Baroni, \textit{Singular parabolic equations, measures satisfying density conditions, and gradient integrability}, Nonlinear Anal. \textbf{153} (2017), 89–116.

[5] P. Baroni and J. Habermann, \textit{Calderón-Zygmund estimates for parabolic measure data equations}, J. Differential Equations \textbf{252} (2012), no. 1, 412–447.

[6] L. Boccardo, A. Dall’Aglio, T. Gallouët, and L. Orsina, \textit{Nonlinear parabolic equations with measure data}, J. Funct. Anal. \textbf{147} (1997), no. 1, 237–258.

[7] L. Boccardo and T. Gallouët, \textit{Nonlinear elliptic and parabolic equations involving measure data}, J. Funct. Anal. \textbf{87} (1989), no. 1, 149–169.

[8] V. Bögelein, \textit{Regularity results for weak and very weak solutions of higher order parabolic systems}, Ph.D. Thesis, 2007.

[9] V. Bögelein, \textit{Global Calderón-Zygmund theory for nonlinear parabolic systems}, Calc. Var. Partial Differential Equations \textbf{51} (2014), no. 3–4, 555–596.
[10] 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.

[11] T. A. Bui and X. T. Duong, Global Marcinkiewicz estimates for nonlinear parabolic equations with nonsmooth coefficients, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 18 (2018), no. 3, 881–916.

[12] 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.

[13] S.-S. Byun and J.-T. Park, Global weighted Orlicz estimates for parabolic measure data problems: application to estimates in variable exponent spaces, J. Math. Anal. Appl. 467 (2018), no. 2, 1194–1207.

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

[15] E. Casas, Control of an elliptic problem with pointwise state constraints, SIAM J. Control Optim. 24 (1986), no. 6, 1309–1318.

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

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

[18] E. Giusti, Direct methods in the calculus of variations, World Scientific Publishing Co., Inc., River Edge, NJ, 2003.

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

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

[21] J. Kinnunen and E. Saksman, Regularity of the fractional maximal function, Bull. London Math. Soc. 35 (2003), no. 4, 529–535.

[22] T. Kuusi and G. Mingione, Gradient regularity for nonlinear parabolic equations, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 6 (2007), no. 2, 195–261.
PARABOLIC EQUATIONS INVOLVING MEASURE DATA

[37] , Gradient estimates below the duality exponent, Math. Ann. 346 (2010), no. 3, 571–627.
[38] , Nonlinear measure data problems, Milan J. Math. 79 (2011), no. 2, 429–496.
[39] Q.-H. Nguyen, Global estimates for quasilinear parabolic equations on Reifenberg flat domains and its applications to Riccati type parabolic equations with distributional data, Calc. Var. Partial Differential Equations 54 (2015), no. 4, 3927–3948.
[40] S. Osher and R. Fedkiw, Level set methods and dynamic implicit surfaces, Applied Mathematical Sciences, vol. 153, Springer-Verlag, New York, 2003.
[41] C. S. Peskin, Numerical analysis of blood flow in the heart, J. Computational Phys. 25 (1977), no. 3, 220–252.
[42] 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.
[43] F. Petitta, Renormalized solutions of nonlinear parabolic equations with general measure data, Ann. Mat. Pura Appl. (4) 187 (2008), no. 4, 563–604.
[44] N. C. Phuc, Nonlinear Muckenhoupt-Wheeden type bounds on Reifenberg flat domains, with applications to quasilinear Riccati type equations, Adv. Math. 250 (2014), 387–419.
[45] E. M. Stein, Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals, Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, N.J., 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.
[46] 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.
[47] T. Toro, Doubling and flatness: geometry of measures, Notices Amer. Math. Soc. 44 (1997), no. 9, 1087–1094.
[48] 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.

Sun-Sig Byun: Department of Mathematical Sciences and Research Institute of Mathematics, Seoul National University, Seoul 08826, Republic of Korea
E-mail address: byun@snu.ac.kr

Jung-Tae Park: Korea Institute for Advanced Study, Seoul 02455, Republic of Korea
E-mail address: ppark009@kias.re.kr

Pilsoo Shin: Department of Mathematical Sciences, Seoul National University, Seoul 08826, Republic of Korea
E-mail address: shinpilsoo@snu.ac.kr