Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XXI (2020), 535-568

Global weighted Lorentz estimates for parabolic equations with measure via strong fractional maximal functions

THE ANH BUI

Abstract. In this paper, we prove a weighted norm inequality for the gradient of solutions to nonlinear parabolic equations with measure data via strong fractional maximal functions. It is worth noticing that our paper is the first one which studies the gradient estimates of solutions to such equations via strong fractional maximal functions.

Mathematics Subject Classification (2010): 35R06 (primary), 35R05, 35K65, 35B65 (secondary).

1. Introduction

Let Ω be a bounded open domain in $\mathbb{R}^n$, $n \geq 2$. For $p > 2 - \frac{1}{n+1}$, we consider the following parabolic equation with measure data

$$\begin{cases} u_t - \operatorname{div} \mathbf{a}(Du, x, t) = \mu & \text{in} \quad \Omega_T \\ u = 0 & \text{on} \quad \partial_p \Omega_T, \end{cases} \tag{1.1}$$

where $T > 0$ is a given positive constant, $\Omega_T = \Omega \times (0, T)$, $\partial_p \Omega_T = (\partial\Omega \times (0, T)) \cup (\bar{\Omega} \times \{0\})$, and μ is a signed Borel measure with finite total mass. Throughout the paper, we denote $u_t = \frac{\partial u}{\partial t}$ and $Du = D_x u := (D_{x_1}, \dots, D_{x_n})$.

In this paper, we assume that the nonlinearity $\mathbf{a}(\xi, x, t) = (\mathbf{a}^1, \dots, \mathbf{a}^n) : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ in (1.1) is measurable in (x, t) for every ξ, differentiable in ξ for a.e. (x, t), and satisfies the following conditions: there exist $\Lambda_1, \Lambda_2 > 0$ so that

$$|\mathbf{a}(\xi, x, t)| + |\xi| \cdot |D_\xi \mathbf{a}(\xi, x, t)| \leq \Lambda_1 |\xi|^{p-1}, \tag{1.2}$$

and

$$\langle \mathbf{a}(\xi, x, t) - \mathbf{a}(\eta, x, t), \xi - \eta \rangle \geq \Lambda_2 \begin{cases} |\xi - \eta|^p, & p \geq 2, \\ (|\xi| + |\eta|)^{p-2} |\xi - \eta|^2, & 2 - \frac{1}{n+1} < p < 2. \end{cases} \tag{1.3}$$

for a.e $(\xi, \eta) \in \mathbb{R}^n \times \mathbb{R}^n$ and a.e. $(x, t) \in \mathbb{R}^n \times \mathbb{R}$.

Received July 02, 2018; accepted March 20, 2019.
Published online December 2020.

The nonlinearity $\mathbf{a}(\xi, x, t)$ satisfying these conditions is modelled on the prototype of p-Laplacian $\Delta_p u = \operatorname{div}(|Du|^{p-2}Du)$ with respect to $\mathbf{a}(\xi, x, t) = |\xi|^{p-2}\xi$.

Definition 1.1. A function $u \in C(0, T; L^2(\Omega)) \cap L^p(0, T; W_0^{1,p}(\Omega))$ is said to be a weak solution to equation (1.1) if the following holds true

$$-\int_{\Omega_T} u\varphi_t dz + \int_{\Omega_T} \langle \mathbf{a}(Du, x, t), D\varphi\rangle dz = \int_{\Omega_T} \varphi d\mu, \tag{1.4}$$

for every $\varphi \in C_c^\infty(\Omega_T)$.

In general, if μ is a signed Borel measure with finite total mass, it is not clear whether the weak solution to equation (1.1) exists. However, this guarantees the existence of a particular type of solution so called SOLA (Solution Obtained as Limits of Approximation). For the sake of convenience, we sketch the ideas about the SOLA in [9,10]. For each $k \in \mathbb{N}$, we consider the regularized problem

$$\begin{cases} (u_k)_t - \operatorname{div} \mathbf{a}(Du_k, x, t) = \mu_k & \text{in} \quad \Omega_T \\ u_k = 0 & \text{on} \quad \partial_p \Omega_T, \end{cases} \tag{1.5}$$

where $\mu_k \in C^\infty(\Omega_T)$ converges to μ in the weak sense of measure and

$$|\mu_k|(Q_R \cap \Omega_T) \leq |\mu|(Q_R \cap \Omega_T), \quad k \geq 1, R > 0.$$

As a classical result, equation (1.5) admits a weak solution $u_k \in C(0,T;L^2(\Omega)) \cap L^p(0, T; W_0^{1,p}(\Omega))$ for each k. Moreover, it was proved in [11] that there exists u so that $u_k \to u$ in $L^q(0, T; W_0^{1,q}(\Omega))$ for any $q \in [1, p-1+\frac{1}{n+1})$. By this reason, the limit of approximation solution u is refered to SOLA (Solution Obtained as Limits of Approximation). In the general case, the SOLA may not be unique. However, the uniqueness of SOLA is guaranteed if $\mu \in L^1(\Omega_T)$. See for example [18]. For this reason, for the sake of simplicity we assume that $\mu \in L^1(\Omega_T)$ and we will state the main result (see Theorem 1.5) for weak solutions instead of SOLAs; however, needless to say, our resutls still hold for SOLAs.

The nonlinear elliptic and parabolic equations with measure data have received a great deal of attention by many mathematicians. See for example [8–11,22,23, 30,38–41] and the references therein. One of the most interesting problems concerning the SOLAs to equation (1.1) is the gradient estimate for its solutions. More precisely, we look for the conditions on the measure μ, the nonlinearity $\mathbf{a}$ and the domain Ω so that the gradient Du of the solutions to (1.1) lies in some function spaces. Recently, there have been a number of research which investigates this problem under the condition that the measure μ belongs to certain Morrey spaces. Recall that for $0 < \theta \leq n+2$, we say that the measure μ is in the Morrey space $L^{1,\theta}(\Omega_T)$ if the following holds true:

$$\sup_{z\in\Omega_T} \sup_{0<r\leq \operatorname{diam}\Omega_T} \frac{|\mu|(Q_r(z)\cap\Omega_T)}{|Q_r(z)\cap\Omega_T|^{1-\frac{\theta}{n+2}}} < \infty,$$

where $Q_r(z) = B_r(x) \times (t - r^2, t + r^2)$ with $z = (x, t)$ and $B_r(x) = \{y \in \mathbb{R}^n : |x - y| < r\}$. We would like to give a shortlist of research in this direction.

(i) The local Marcinkiewicz type estimates were obtained for the elliptic equations in [39]:

$$\mu \in L^{1,\theta}(\Omega), 2 \leq \theta \leq n \Longrightarrow |Du| \in L_{\mathrm{loc}}^{\frac{\theta(p-1)}{\theta-1},\infty}(\Omega)$$

where $L^{m,\infty}(\Omega_T)$ is the weak-Lebesgue space, or the Marcinkiewicz space, defined by the set of all measurable functions f on Ω_T satisfying

$$\|f\|_{L^{m,\infty}(\Omega_T)} := \sup_{\lambda>0} \lambda |\{z \in \Omega_T : |f(z)| > \lambda\}|^{\frac{1}{m}} < +\infty.$$

It is easy to see that when $\theta = n$, this estimate turns out to be:

$$\mu \in L^{1,n}(\Omega) \Longrightarrow |Du|^{p-1} \in \mathcal{M}_{\mathrm{loc}}^{\frac{n}{n-1}}(\Omega), \quad p < n.$$

See for example [9,11]. The borderline case $p = n$ is much more difficult and was obtained in [21].

(ii) For the parabolic equation, the local version of Marcinkiewicz type estimates for $p = 2$ was obtained in [7] by making use of the maximal function technique.

(iii) The case $p \geq 2$ was proved in [4]. It was proved in [4] that there exists $\tilde{\theta} \in (1, 2)$ so that

$$\mu \in L^{1,\theta}(\Omega_T), \ \ \theta \in (\tilde{\theta}, n+2] \ \ \Longrightarrow |Du| \in \mathcal{M}_{\mathrm{loc}}^{m}(\Omega_T), \ \ m = p-1+\frac{1}{\theta-1}.$$

The number $\tilde{\theta} \in (1, 2)$ is a threshold and has a connection with the exponent in higher integrability estimates of the associated homogeneous equation. Later, the authors in [13] extended to study the global estimates with a more general nonlinearity $\mathbf{a}(\xi, x, t)$. Recently, the case $2 - \frac{1}{n+1} < p < 2$ has been obtained in [5].

Before coming to our main results, we will clarify the assumptions which will be considered in the paper. Apart from (1.2) and (1.3), the nonlinearity $\mathbf{a}$ will be asssumed to satisfy the small BMO norm condition (1.6) below. We set

$$\Theta(\mathbf{a}, B_r(y))(x, t) = \sup_{\xi \in \mathbb{R}^n \setminus \{0\}} \frac{|\mathbf{a}(\xi, x, t) - \overline{\mathbf{a}}_{B_r(y)}(\xi, t)|}{|\xi|^{p-1}},$$

where

$$\overline{\mathbf{a}}_{B_r(y)}(\xi, t) = \fint_{B_r(y)} \mathbf{a}(\xi, x, t) dx.$$

Definition 1.2. Let $R_0, \delta > 0$. The nonlinearity $\mathbf{a}$ is said to satisfy the small (δ, R_0)-BMO condition if

$$[\mathbf{a}]_{2,R_0} := \sup_{(y,s)\in\mathbb{R}^{n+1}} \sup_{0<r\le R_0,\tau>0} \fint_{Q_{(r,\tau)}(y,s)} |\Theta(\mathbf{a}, B_r(y))(x,t)|^2 dz \le \delta^2, \quad (1.6)$$

where $Q_{(r,\tau)}(y,s) = B_r(y) \times (s-\tau, s+\tau)$.

Note that the small (δ, R_0)-BMO condition is satisfied even when $\mathbf{a}(\xi, x, t)$ is discontinuous with respect to x and t.

Concerning the underlying domain Ω, we do not assume any smoothness condition on Ω, but the following flatness condition.

Definition 1.3. Let $\delta, R_0 > 0$. The domain Ω is said to be a (δ, R_0) Reifenberg flat domain if for every $x \in \partial\Omega$ and $0 < r \le R_0$, then there exists a coordinate system depending on x and r, whose variables are denoted by $y = (y_1, \dots, y_n)$ such that in this new coordinate system x is the origin and

$$B_r \cap \{y : y_n > \delta r\} \subset B_r \cap \Omega \subset \{y : y_n > -\delta r\}. \quad (1.7)$$

The condition of (δ, R_0)-Reifenberg flatness was first introduced in [47]. This condition does not require any smoothness on the boundary of Ω, but sufficient flatness in the Reifenberg's sense. The Reifenberg flat domains include domains with rough boundaries of fractal nature, and Lipschitz domains with small Lipschitz constants. For further discussion about the Reifenberg domain, we refer to [19,43,47,51] and the references therein.

Throughout the paper, we always assume that the domain Ω *is a* (δ, R_0) *Reifenberg flat domain, and the nonlinearity* $\mathbf{a}$ *satisfies* (1.2), (1.3) *and the small* (δ, R_0)-*BMO condition* (1.6).

We set

$$\mathcal{Q} = \{Q : Q = B_r(x) \times (t_1, t_2), x \in \mathbb{R}^n, r > 0; t_1, t_2 \in \mathbb{R}\}.$$

Let $1 \le q < \infty$. A nonnegative locally integrable function w belongs to the class A_q^*, say $w \in A_q^*$, if there exists a positive constant C so that

$$[w]_{A_q^*} := \sup_{Q\in\mathcal{Q}} \Big(\fint_Q w(z)dz\Big)\Big(\fint_Q w^{-1/(q-1)}(z)dz\Big)^{q-1} \le C, \quad \text{if } 1<q<\infty, \quad (1.8)$$

and

$$\fint_Q w(z)dz \le C \operatorname*{ess-inf}_{z\in Q} w(z), \quad \text{if } q = 1, \quad (1.9)$$

for all $Q \in \mathcal{Q}$. We say that $w \in A_\infty^*$ if $w \in A_q^*$ for some $q \in [1, \infty)$. We will denote $w(E) := \int_E w(z)dz$ for any measurable set $E \subset \mathbb{R}^{n+1}$.

We note that if we replace the family $\mathcal{Q}$ by the family of parabolic cylinders of the form $Q_r(z) = B_r(x) \times (t-r^2, t+r^2)$ with $z = (x,t) \in \mathbb{R}^{n+1}$ in (1.8) and (1.9), then we have the Muckenhoupt weights A_p. Hence, it is clearly that $A_p^* \subset A_p$ for all $1 \le p < \infty$.

Some similar results to the Muckenhoupt weights follow due to [3].

Lemma 1.4. *Let $w \in A^*_q$ for some $1 \le q < \infty$. There exist $\tau = \tau([w]_{A^*_q})$, and a constant $C = C([w]_{A^*_q})$ such that for any $Q \in \mathcal{Q}$, and any measurable subset $E \subset Q$,*

$$C^{-1}\left(\frac{|E|}{|Q|}\right)^q \le \frac{w(E)}{w(Q)} \le C\left(\frac{|E|}{|Q|}\right)^\tau .$$

Let $w \in A^*_\infty$, $0 < q < \infty$, $0 < r \le \infty$ and let E be a subset of $\mathbb{R}^{n+1}$. The weighted Lorentz space $L^{q,r}_w(E)$ is defined as the set of all measurable functions f on E such that

$$\|f\|_{L^{q,r}_w(E)} := \left\{ q\int_0^\infty \left[t^q w\left(\{z \in E : |f(z)| > t\}\right)\right]^{r/q} \frac{dt}{t}\right\}^{1/r} < \infty .$$

In the particular case $q = r$, the weighted Lorentz space $L^{q,q}_w(E)$ coincides with the weighted Lebesgue space $L^q_w(E)$ which is defined as the space of all measurable functions f on E such that

$$\|f\|_{L^q_w(E)} = \left(\int_E |f(z)|^q w(z) dz\right)^{1/q} < \infty .$$

In order to state our main result, we first recall the concept of the strong fractional maximal function:

$$\mathcal{M}^{\mathrm{s}}_1(\mu)(z) = \sup_{r,\tau>0} \frac{|\mu|(Q_{r,\tau}(z))}{|Q_{r,\tau}(z)|^{\frac{n+1}{n+2}}} \tag{1.10}$$

where the supremum is taken over all cylinders $Q_{r,\tau}(z) = B_r(x) \times (t - \tau^2, t + \tau^2)$ with $z = (x, t) \in \mathbb{R}^{n+1}$.

We note that the strong fractional maximal function defined in (1.10) is a natural variant version of strong maximal functions introduced in [17]. This kind of strong fractional maximal functions was introduced in [50] to study the multi-parameter of Riezs type potentials.

The main aim of this paper is to establish the gradient estimate for solutions to (1.1) in terms of the strong maximal function defined in (1.10). This is contrast with those in [5,7,13,21,30,39] where the gradient esitmates were obtained with Morrey data conditions for μ. Our main result is the following theorem.

Theorem 1.5. *Let $w \in A^*_\infty$, $0 < q < \infty$ and $0 < r \le \infty$. Then there exists a positive constant δ such that the following holds. If u is a weak solution to* (1.1) *with $\mu \in L^1(\Omega_T)$, the domain Ω is a (δ, R_0)-Reifenberg flat domain, and the nonlinearity* **a** *satisfies* (1.2), (1.3) *and the small* (δ, R_0)*-BMO condition* (1.6)*, then we have*

$$\|Du\|_{L^{q,r}_w(\Omega_T)} \le C\left(\left\| \left[\mathcal{M}^{\mathrm{s}}_1(\mu)\right]^{\frac{n+2}{p+(p-1)n}} \right\|_{L^{q,r}_w(\Omega_T)} + 1\right)^{\frac{2}{2-n(2-p)}}, \qquad 2 - \frac{1}{n-1} < p < 2, \tag{1.11}$$

and

$$\|Du\|_{L^{q,r}_w(\Omega_T)} \le C\left(\left\|[\mathcal{M}^{\rm s}_1(\mu)]^{\frac{(p-1)(n+2)}{p+(p-1)n}}\right\|_{L^{q,r}_w(\Omega_T)}+1\right), \quad p\ge 2, \tag{1.12}$$

where C is a constant depending on n, w, Λ_1, Λ_2, δ, R_0, Ω_T.

Needless to say, as mentioned earlier the results in Theorem 1.5 still hold true for SOLAs to (1.1).

Although there have been a number of research dedicated to the improvement of integrability for the gradient of solutions to (1.1), our paper is the first one where investigates the gradient estimates for the solutions to (1.1) via strong maximal function $\mathcal{M}^{\rm s}_1$. This result is even new for unweighted estimate.

Note that in the particular case $p=2$, it was proved in [42] that

$$\|Du\|_{L^{q,r}_w(\Omega_T)} \le C\left\|\mathcal{M}_1(\mu)\right\|_{L^{q,r}_w(\Omega_T)} \tag{1.13}$$

for all $0<q<\infty, 0<r\le\infty$ and $w\in A_\infty$, where A_∞ is a class of Muckenhoupt weights, and $\mathcal{M}_1$ is a fractional maximal function defined by

$$\mathcal{M}_1(\mu)(z)=\sup_{r>0}\frac{|\mu|(Q_r(z))}{|Q_r(z))|^{\frac{n+1}{n+2}}}, \tag{1.14}$$

where the supremum is taken over all parabolic cylinders $Q_r(z)=B_r(x)\times(t-r^2,t+r^2)$ with $z=(x,t)\in\mathbb{R}^{n+1}$. This estimate is clearly sharper than the estimates in Theorem 1.5, since $\mathcal{M}_1\le\mathcal{M}^{\rm s}_1$ and $A^*_\infty\subset A_\infty$. The reason for this lies in the fact that in our setting we have to work with intrinsic cylinder in the form $B_{\lambda^{\frac{p-2}{2}}r}(x)\times(t-r^2,t+r^2)$ as $2-\frac{1}{2}<p<2$, and $B_r(x)\times(t-\lambda^{2-p}r^2,t+\lambda^{2-p}r^2)$ as $p\ge 2$. Hence, the strong fractional maximal function $\mathcal{M}^{\rm s}_1$ and the new class of weights A^*_∞ are natural and reasonable. Meanwhile, as $p=2$, we only deal with parabolic cylinders of the form $B_r(x)\times(t-r^2,t+r^2)$. This explains why in the particular case $p=2$ we can replace $\mathcal{M}^{\rm s}_1$ and the class A^*_∞ by $\mathcal{M}_1$ and the class A_∞, respectively. However, it is worth noticing that our approach still works well in the case $p=2$ to deduce the estimate (1.13).

Some comments on the techniques used in the paper are in order. In the particular case $p=2$, the gradient estimate via maximal functions can be otained by using Hardy-Littlewood maximal function techniques. See for example [42]. However, this harmonic analysis tool does not work well for the case $p\neq 2$ for the following reasons. Firstly, the homogeneity of the parabolic equations is no longer true as $p\neq 2$, even when $\mu\equiv 0$. Secondly, as mentioned earlier, in our setting we work with intrinsic cylinder in the form $B_{\lambda^{\frac{p-2}{2}}r}(x)\times(t-r^2,t+r^2)$ as $2-\frac{1}{2}<p<2$, and $B_r(x)\times(t-\lambda^{2-p}r^2,t+\lambda^{2-p}r^2)$ as $p\ge 2$ instead of parabolic cylinders. Our approach is based on a covering arguments which is an effective tool in studying the general nonlinear parabolic equations. See for example [1,2,4,12,13,15]. Although, this technique is more or less standard in the parabolic setting, a number of non-trivial improvements would be required. See Section 3 and Section 5.

The organization of the paper is as follows. Part 1 will treat problem for the case $2-\frac{1}{n+1}<p<2$. Section 2 gives some comparison estimates in both interior and boundary cases. The proof of Theorem 1.5 corresponding to $2-\frac{1}{n+1}<p<2$ will be represented in Section 3. The case $p>2$ will be considered in Part 2. Section 4 gives briefly some comparison estimates. The proof of Theorem 1.5 with respect to $p\geq 2$ will be given in Section 5.

Throughout the paper, we always use C and c to denote positive constants that are independent of the main parameters involved but whose values may differ from line to line. We will write $A\lesssim B$ if there is a universal constant C so that $A\leq CB$ and $A\sim B$ if $A\lesssim B$ and $B\lesssim A$.

Part 1

Global weighted estimates for the gradient of solutions: the case $2-\frac{1}{n+1}<p<2$

In this part we employ the following notations. For $z=(x,t)$ with $x\in\mathbb{R}^n$, $t>0$ and $\lambda\geq 1$, we denote:

- $B_r^\lambda(x)=B_{\lambda^{\frac{p-2}{2}}r}(x)$, $I_r(t)=(t-r^2,t+r^2)$, $\Omega_r^\lambda(x)=B_r^\lambda(x)\cap\Omega$, $Q_r^\lambda(z)=B_r^\lambda(x)\times I_r(t)$;
- $K_r^\lambda(z)=Q_r^\lambda(z)\cap\Omega_T$, $\partial_w K_r^\lambda(z)=Q_r^\lambda(z)\cap(\partial\Omega\times\mathbb{R})$, $\partial_p K_r^\lambda(z)=\partial K_r^\lambda(z)\backslash(\bar{\Omega}_r^\lambda(x)\times\{t+r^2\})$;
- $(Q_r^\lambda)^+(z)=(B_r^\lambda)^+(x)\times(t-r^2,t+r^2)$ where $(B_r^\lambda)^+(x)=\{y: y\in B_r^\lambda(x), y_n>x_n\}$.

2. Comparison estimates

2.1. Interior estimates

For $z_0=(x_0,t_0)\in\Omega_T$, $0<R<R_0/4$ and $\lambda\geq 1$ satisfying $B_{4R}^\lambda\equiv B_{4R}^\lambda(x_0)\subset\Omega$, we set

$$Q_{4R}^\lambda\equiv Q_{4R}^\lambda(z_0)=B_{4R}^\lambda\times I_{4R}(t_0). \tag{2.1}$$

For the sake of simplicity, we may assume that $I_{4R}(t_0)\subset(0,T)$, or equivalently, $Q_{4R}^\lambda\subset\Omega_T$. The case $I_{4R}^\lambda(t_0)\cap(0,T)^c\neq\emptyset$ can be done in the same manner with minor modifications.

Assume that u is a weak solution to (1.1). It is well-known that there exists a unique weak solution $w\in C(I_{4R}(t_0);L^2(B_{4R}^\lambda(x_0)))\cap L^p(I_{4R}(t_0);W^{1,p}(B_{4R}^\lambda(x_0)))$ to the following equation

$$\begin{cases} w_t-\operatorname{div}\mathbf{a}(Dw,x,t)=0 & \text{in } Q_{4R}^\lambda\\ w=u & \text{on } \partial_p Q_{4R}^\lambda. \end{cases} \tag{2.2}$$

We begin with the following comparison result:

Lemma 2.1. *Let w be a weak solution to problem* (2.2). *Then for every* $1 \le q < p - 1 + \frac{1}{n+1}$, *there exists C so that*

$$\left(\fint_{Q^\lambda_{4R}} |D(u-w)|^q dz\right)^{1/q} \le C\left[\frac{|\mu|(Q^\lambda_{4R})}{|Q^\lambda_{4R}|^{\frac{n+1}{n+2}}}\right]^{\frac{n+2}{p+(p-1)n}} + C\frac{|\mu|(Q^\lambda_{4R})}{|Q^\lambda_{4R}|^{\frac{n+1}{n+2}}}\left(\fint_{Q^\lambda_{4R}} |D(u-w)|^q dz\right)^{\frac{2-p}{q}\frac{n+1}{n+2}}. \tag{2.3}$$

Proof. The proof of this lemma can be found in [28, Lemma 4.1] (see also [5, Lemma 4.1]). □

The result below shows that the estimate for L^p-norm of the gradient Dw can be inherited from its L^1-norm estimate. We have:

Lemma 2.2. *Let w be a weak solution to problem* (2.2). *If*

$$\fint_{Q^\lambda_{4R}} |Dw| dz \le \kappa\lambda, \quad \kappa \ge 1,$$

then

$$\fint_{Q^\lambda_{2R}} |Dw|^p dz \le c\lambda^p$$

where $c = c(n, p, \Lambda_1, \Lambda_2, \kappa)$.

Proof. We refer to [5, Proposition 3.5] for the proof. □

The next estimate is known as a reverse-Hölder's inequality for the solution to (2.2).

Proposition 2.3. *Let w be a weak solution to problem* (2.2). *Assume that*

$$\lambda^p \lesssim \fint_{Q^\lambda_{2R}} |Dw|^p dz \ \text{ and } \ \fint_{Q^\lambda_{4R}} |Dw|^p dz \lesssim \lambda^p. \tag{2.4}$$

Then there exists $\epsilon_0 > 0$ *such that*

$$\left(\fint_{Q^\lambda_{2R}} |Dw|^{p+\epsilon_0} dz\right)^{\frac{1}{p+\epsilon_0}} \le C\fint_{Q^\lambda_{4R}} |Dw| dz,$$

where C depends on n, p, Λ_1, Λ_2 *and* κ.

Proof. The proof of this proposition is quite standard. See for example [26]. □

Let w be a weak solution to (2.2) under the condition (2.4). We now consider the following problem

$$\begin{cases} v_t - \operatorname{div} \overline{\mathbf{a}}_{B_R}(Dv, t) = 0 & \text{in} \quad Q_R^\lambda \\ v = w & \text{on} \quad \partial_p Q_R^\lambda. \end{cases} \tag{2.5}$$

We then obtain the following estimate.

Lemma 2.4. *Let v be a weak solution to* (2.5) *under the condition* (2.4). *Then for any $\epsilon > 0$ there exists $\delta > 0$ so that*

$$\fint_{Q_R^\lambda} |D(w-v)|^p dz \le \epsilon \lambda^p. \tag{2.6}$$

Proof. The proof is similar to that of Lemma 2.3 in [13] using Proposition 2.3. □

The following approximation result will play an important role in the proof of the main result.

Proposition 2.5. *For each $\epsilon > 0$ there exists $\delta > 0$ so that the following holds true. If u is a weak solution to problem* (1.1) *satisfying*

$$\lambda \le \fint_{Q_R^\lambda} |Du| dz \ \text{ and } \fint_{Q_{4R}^\lambda} |Du| dz \le \lambda, \tag{2.7}$$

and

$$\left[\frac{|\mu|(K_{4R}^\lambda)}{|K_{4R}^\lambda|^{(n+1)/(n+2)}} \right]^{\frac{n+2}{p+(p-1)n}} \le \delta \lambda, \tag{2.8}$$

then the weak solution v to problem (2.5) *satisfies*

$$\|Dv\|_{L^\infty(Q_{R/2}^\lambda)} \lesssim \lambda, \tag{2.9}$$

and

$$\fint_{Q_R^\lambda} |D(u-v)| dz \le \epsilon \lambda. \tag{2.10}$$

Proof. From (2.3) and (2.8), we have

$$\begin{aligned} \fint_{Q_{4R}^\lambda} |D(u-w)| dz &\le C \left[\frac{|\mu|(Q_{4R}^\lambda)}{|Q_{4R}^\lambda|^{\frac{n+1}{n+2}}} \right]^{\frac{n+2}{p+(p-1)n}} \\ &\quad + C \frac{|\mu|(Q_{4R}^\lambda)}{|Q_{4R}^\lambda|^{\frac{n+1}{n+2}}} \left(\fint_{Q_{4R}^\lambda} |D(u-w)| dz \right)^{\frac{(2-p)(n+1)}{n+2}} \\ &\le C\delta\lambda + [\delta\lambda]^{\frac{p+(p-1)n}{n+2}} \left(\fint_{Q_{4R}^\lambda} |D(u-w)| dz \right)^{\frac{(2-p)(n+1)}{n+2}}. \end{aligned}$$

This implies that

$$\fint_{Q^\lambda_{4R}} |D(u-w)|dz \le C\delta\lambda.$$

Taking δ to be sufficiently small, from the above inequality and (2.7) we infer that

$$\lambda \lesssim \fint_{Q^\lambda_R} |Dw|dz, \quad \fint_{Q^\lambda_{4R}} |Dw|dz \lesssim \lambda.$$

We now apply Hölder's inequality and Lemma 2.2 to deduce that

$$\lambda^p \lesssim \fint_{Q^\lambda_R} |Dw|^p dz, \quad \fint_{Q^\lambda_{2R}} |Dw|^p dz \lesssim \lambda^p. \tag{2.11}$$

Then by Lemma 2.4, we have that for any $\tilde{\epsilon}$ there exists $\delta > 0$ so that

$$\fint_{Q^\lambda_R} |D(w-v)|^p dz \le \tilde{\epsilon}\lambda^p.$$

Therefore,

$$\begin{aligned} \|Dv\|^p_{L^\infty(Q^\lambda_{R/2})} &\lesssim \fint_{Q^\lambda_R} |Dv|^p dz \\ &\lesssim \fint_{Q^\lambda_R} |Dw|^p dz + \fint_{Q^\lambda_R} |D(w-v)|^p dz \\ &\lesssim \lambda^p, \end{aligned}$$

which proves (2.9), where in the first inequality we used Hölder estimates for Du. See for example [20, Chapter VIII].

Then the inequality (2.10) can be obtained via the following estimates:

$$\begin{aligned} \fint_{Q^\lambda_R} |D(u-v)|dz &\lesssim \fint_{Q^\lambda_R} |D(u-w)|dz + \fint_{Q^\lambda_R} |D(w-v)|dz \\ &\lesssim C(\delta + \tilde{\epsilon})\lambda^p. \end{aligned}$$

By taking δ and $\tilde{\epsilon}$ to be sufficiently small, we obtain (2.10). This completes our proof. □

2.2. Boundary estimates

Fix $t_0 \in (0, T)$ and $x_0 \in \partial\Omega$, we set $z_0 = (x_0, t_0)$. Let $0 < R < R_0/4$ and $\lambda \ge 1$. For the sake of simplicity, we restrict ourself to consider the lateral boundary case with respect to

$$I_{4R}(t_0) \subset (0, T),$$

since the initial boundary case can be done in the same manner.

Before coming to the main comparison estimates, we shall establish some boundary estimates on weak solutions to the homogeneous equations associated to (1.1).

2.2.1. *Some boundary estimates for homogeneous equations*

We now consider the weak solution

$$w \in C(I_{4R}(t_0); L^2(\Omega^\lambda_{4R}(x_0))) \cap L^p(I_{4R}(t_0); W^{1,p}(\Omega^\lambda_{4R}(x_0)))$$

to the following equation

$$\begin{cases} w_t - \operatorname{div} \mathbf{a}(Dw, x, t) = 0 & \text{in} \quad K^\lambda_{4R}(z_0) \\ w = 0 & \text{on} \quad \partial_w K^\lambda_{4R}(z_0). \end{cases} \tag{2.12}$$

Lemma 2.6. *Let w be a weak solution to problem* (2.12). *Let $K^\lambda_{\rho_1}(\bar{z}) \subset K^\lambda_{\rho_2}(\bar{z}) \subset K^\lambda_{4R}(z_0)$ with $\bar{z} = (\bar{x}, \bar{t})$ and $\rho_2 > \rho_1 > 0$. Then there exists $c = c(n, p, \Lambda_1, \Lambda_2)$ so that*

$$\begin{aligned} &\int_{K^\lambda_{\rho_1}(\bar{z})} |Dw|^p dz + \sup_{t \in I_{\rho_1}(\bar{t})} \int_{B^\lambda_{\rho_1}(\bar{x})} |w|^2 dx \\ &\le \frac{1}{(\rho_2^2 - \rho_1^2)} \int_{K^\lambda_{\rho_2}(\bar{z})} |w|^2 dz + \frac{c}{\lambda^{\frac{p(p-2)}{2}} (\rho_2 - \rho_1)^p} \int_{K^\lambda_{\rho_2}(\bar{z})} |w|^p dz. \end{aligned}$$

Proof. The proof of this lemma is quite standard. See for example [26]. Hence, we omit details. □

We now give a useful result which will be used in the sequel.

Lemma 2.7. *Let w be a weak solution to equation* (2.12). *Then for $\theta \in (0, 1)$ and $K^\sigma_r(z_0) \subset K^\lambda_{4R}(z_0)$ with $r, \sigma > 0$ we have*

$$\sup_{K^\sigma_{3r}(z_0)} |w| \le c(r\sigma^{p/2})^{\frac{n(p-2)}{n(p-2)+p}} \left(\fint_{K^\sigma_{4r}(z_0)} |w| dz \right)^{\frac{p}{n(p-2)+p}} + cr\sigma^{p/2}. \tag{2.13}$$

Proof. Recall that a sub-solution is a function such that the left-hand side of the weak formula of (2.12) is negative, for all positive test functions.

Note that since w is a weak solution to (2.12), $|w|$ is a nonnegative subsolution to equation (2.12). See for example Lemma 1.1 in [20].

The estimate (2.13) can be found in Theorem 5.1 in [20, Chapter V] for the interior case. This argument still works well in our situation with a minor modification. Hence, we omit details and leave it to interested readers. □

We now recall a Sobolev-Poincaré's inequality near the Reifenberg domain which is a particular case of that in [25,37].

In the particular case when Ω is a Reifenberg flat domain, we have the following result.

Lemma 2.8. *Let Ω be a (δ, R_0) Reifenberg domain. Suppose that $1 < q < \infty$ and that u is a q-quasicontinuous function in $W^{1,q}(\Omega_r(x_0))$, where $x_0 \in \partial\Omega$ and $0 < r < R_0$. Then*

$$\left(\fint_{\Omega_r(x_0)} |u|^{\kappa q} dx\right)^{\frac{1}{\kappa q}} \le cr \left(\fint_{\Omega_r(x_0)} |\nabla u|^q dx\right)^{1/q}, \tag{2.14}$$

where $c = c(n, q) > 0$ and $\kappa = n/(n-q)$ if $1 < q < n$ and $\kappa = 2$ if $q \ge n$.

In particularly, we have

$$\left(\fint_{\Omega_r(x_0)} |u|^q dx\right)^{\frac{1}{q}} \le cr \left(\fint_{B_r(x_0)} |\nabla \bar{u}|^q dx\right)^{1/q}. \tag{2.15}$$

Note that in the interior case, the L^p-norm estimate for Dw can be obtained from the its L^1-norm estimate. However, it is not clear if this might be true in the boundary case, due to a techical reason which the Sobolev-Poincaré inequality near the boundary (2.14) may not be true as $q = 1$. Hence, in the boundary case, we have a slightly different estimate:

Proposition 2.9. *Let w be a weak solution to problem* (2.12) *satisfying the estimates*

$$\lambda^{1+\sigma_0} \lesssim \fint_{K^\lambda_R(z_0)} |Dw|^{1+\sigma_0} dz \ \text{ and } \fint_{K^\lambda_{4R}(z_0)} |Dw|^{1+\sigma_0} dz \lesssim \lambda^{1+\sigma_0}, \tag{2.16}$$

for $\lambda > 1$ and some $0 < \sigma_0 < p - 1$. Then we have

$$\lambda^p \lesssim \fint_{K^\lambda_R(z_0)} |Dw|^p dz \ \text{ and } \fint_{K^\lambda_{2R}(z_0)} |Dw|^p dz \lesssim \lambda^p. \tag{2.17}$$

Proof. By Hölder's inequality, we have

$$\fint_{K^\lambda_R(z_0)} |Dw|^p dz \ge C\lambda^p.$$

It remains to prove the second inequality in (2.17). Indeed, from Lemma 2.6 we have

$$\fint_{K^\lambda_{2R}(z_0)} |Dw|^p dz \le \frac{c}{R^2} \fint_{K^\lambda_{3R}(z_0)} |w|^2 dz + \frac{c}{\lambda^{\frac{p(p-2)}{2}} R^p} \fint_{K^\lambda_{3R}(z_0)} |w|^p dz. \tag{2.18}$$

To do this, by Lemma 2.7 with $\sigma = \lambda$, $r = 4R$ we find that

$$\sup_{K^\lambda_{3R}(z_0)} |w| \lesssim (R\lambda^{p/2})^{\frac{n(p-2)}{n(p-2)+p}} \left(\fint_{K^\lambda_{4R}(z_0)} |w| dz\right)^{\frac{p}{n(p-2)+p}} + cR\lambda^{p/2}.$$

This along with Sobolev-Poincaré's inequality (2.15), Hölder's inequality and (2.16) implies that

$$\begin{aligned}\sup_{K^\lambda_{3R}(z_0)} |w| &\lesssim (R\lambda^{p/2})^{\frac{n(p-2)}{n(p-2)+p}} (\lambda^{\frac{p-2}{2}} R)^{\frac{p}{n(p-2)+p}} \left(\fint_{K^\lambda_{4R}(z_0)} |Dw|^{1+\sigma_0} dz \right)^{\frac{p}{(1+\sigma_0)[n(p-2)+p]}} \\ &\quad + cR\lambda^{p/2} \\ &\lesssim (R\lambda^{p/2})^{\frac{n(p-2)}{n(p-2)+p}} (\lambda^{\frac{p-2}{2}} R)^{\frac{p}{n(p-2)+p}} \lambda^{\frac{p}{n(p-2)+p}} + cR\lambda^{p/2} \\ &\lesssim R\lambda^{p/2}.\end{aligned}$$

Inserting this into (2.18) we obtain

$$\fint_{K^\lambda_{2R}(z_0)} |Dw|^p dz \lesssim \lambda^p.$$

This completes our proof. □

Similarly to the interior case, a reverse-Hölder's inequality still holds true for the solution to problem (2.12) near the boundary.

Proposition 2.10. *Let w be a weak solution to problem* (2.12). *Assume that*

$$\lambda^p \lesssim \fint_{K^\lambda_R(z_0)} |Dw|^p dz \ \text{ and } \ \fint_{K^\lambda_{2R}(z_0)} |Dw|^p dz \lesssim \lambda^p. \tag{2.19}$$

Then there exists $\epsilon_0 > 0$ so that

$$\left(\fint_{K^\lambda_R(z_0)} |Dw|^{p+\epsilon_0} dz \right)^{\frac{1}{p+\epsilon_0}} \le C \fint_{K^\lambda_{2R}(z_0)} |Dw| dz.$$

Proof. For the proof we refer to [43, Lemma 4.1]. □

We now give some comparison estimates for the weak solutions to (1.1).

2.2.2. *Comparision estimates near boundary*

Assume that u is a weak solution to problem (1.1). We consider the following equation

$$\begin{cases} w_t - \operatorname{div} \mathbf{a}(Dw, x, t) = 0 & \text{in } \ K^\lambda_{4R}(z_0) \\ w = u & \text{on } \ \partial_p K^\lambda_{4R}(z_0). \end{cases} \tag{2.20}$$

Arguing similarly to the proof of [28, Lemma 4.1] (see also [5, Lemma 4.1]), we can prove the similar estimate to that in Lemma 2.2 for the boundary case.

Lemma 2.11. *Let w be a weak solution to problem* (2.20). *Then for every* $1 \le q < p - 1 + \frac{1}{n+1}$, *there exists C so that*

$$\left(\fint_{K^\lambda_{4R}(z_0)} |D(u-w)|^q dz\right)^{1/q} \le C\left[\frac{|\mu|(K^\lambda_{4R}(z_0))}{|K^\lambda_{4R}(z_0)|^{(n+1)/(n+2)}}\right]^{\frac{n+2}{p+(p-1)n}} + C\frac{|\mu|(K^\lambda_{4R})}{|K^\lambda_{4R}|^{\frac{n+1}{n+2}}}\left(\fint_{K^\lambda_{4R}} |D(u-w)|^q dz\right)^{\frac{2-p}{q}\frac{n+1}{n+2}}. \tag{2.21}$$

We now assume that $0 < \delta < 1/50$. Since $x_0 \in \partial\Omega$, there exists a new coordinate system whose variables are still denoted by $(x_1, \dots, x_n)$ such that in this coordinate system the origin is some interior point of Ω, $x_0 = (0, \dots, 0, -\frac{\delta R}{2(1-\delta)})$ and

$$B^+_{R/2} \subset B_{R/2} \cap \Omega \subset B_{R/2} \cap \{x : x_n > -3\delta R\}. \tag{2.22}$$

Note that due to $\delta \in (0, 1/50)$, we further obtain

$$B_{R/8}(x_0) \subset B_{3R/8} \subset B_{R/4}(x_0) \subset B_{R/2} \subset B_R(x_0). \tag{2.23}$$

Let w be a weak solution to (2.20) satisfying

$$\lambda^p \lesssim \fint_{K^\lambda_R(z_0)} |Dw|^p dz \quad\text{and}\quad \fint_{K^\lambda_{2R}(z_0)} |Dw|^p dz \lesssim \lambda^p. \tag{2.24}$$

We now consider the following problem (in the new coordinate system)

$$\begin{cases} h_t - \operatorname{div} \bar{\mathbf{a}}_{B^\lambda_{R/2}}(Dh, t) = 0 & \text{in} \quad K^\lambda_{R/2}(0, t_0) \\ h = w & \text{on} \quad \partial_p K^\lambda_{R/2}(0, t_0). \end{cases} \tag{2.25}$$

Using Proposition 2.10 as a main vehicle and arguing similarly to the proof of [13, Lemma 2.3], we can prove:

Lemma 2.12. *Let h be a weak solution to* (2.25) *under the condition* (2.24). *Then for any $\epsilon > 0$ there exists $\delta > 0$ so that*

$$\fint_{K^\lambda_{R/2}(0,t_0)} |D(w-h)|^p dz \le \epsilon\lambda^p. \tag{2.26}$$

The main different from the interior case is that due to the lack of smoothness condition on the boundary of Ω, we can not expect that the L^∞-norm of Dh is finite near the boundary. To handle this trouble, we consider its associated problem.

$$\begin{cases} v_t - \operatorname{div} \bar{\mathbf{a}}_{B_{R/2}}(Dv, t) = 0 & \text{in} \quad (Q^\lambda_{R/2})^+(0, t_0), \\ v = 0 & \text{on} \quad Q^\lambda_{R/2}(0, t_0) \cap \{z = (x', x_n, t) : x_n = 0\}. \end{cases} \tag{2.27}$$

Proposition 2.13. *For each $\epsilon > 0$ there exists $\delta > 0$ so that the following holds true. If u is a weak solution to problem* (1.1) *satisfying*

$$\lambda^{1+\sigma_0} \lesssim \fint_{K_R^\lambda(z_0)} |Du|^{1+\sigma_0} dz, \quad \fint_{K_{4R}^\lambda(z_0)} |Du|^{1+\sigma_0} dz \lesssim \kappa\lambda^{1+\sigma_0}, \quad \text{for some } 0 < \sigma_0 < p - 2 + \frac{1}{n+1}, \tag{2.28}$$

and

$$\left[\frac{|\mu|(K_{4R}^\lambda(z_0))}{|K_{4R}^\lambda(z_0)|^{(n+1)/(n+2)}}\right]^{\frac{n+2}{p+(p-1)n}} \leq \delta\lambda, \tag{2.29}$$

then there exists a weak solution v to problem (2.27) *satisfying*

$$\|D\hat{v}\|_{L^\infty(Q_{R/8}^\lambda(z_0))} \lesssim \lambda, \tag{2.30}$$

and

$$\fint_{K_{R/4}^\lambda(z_0)} |D(u-\hat{v})|^{1+\sigma_0} dz \leq (\epsilon\lambda)^{1+\sigma_0} \tag{2.31}$$

where $\hat{v}$ is the zero extension of v to $Q_{R/2}^\lambda(0,t_0) \supset Q_{R/4}^\lambda(z_0)$.

Proof. Since $\sigma_0 \in (0, p-2+\frac{1}{n+1})$, we have $1+\sigma_0 < p-1+\frac{1}{n+1}$. Hence, applying (2.21) and (2.29), we have

$$\left(\fint_{K_{4R}^\lambda(z_0)} |D(u-w)|^{1+\sigma_0} dz\right)^{\frac{1}{1+\sigma_0}} \leq C\delta\lambda + [\delta\lambda]^{\frac{p+(p-1)n}{n+2}} \left(\fint_{K_{4R}^\lambda(z_0)} |D(u-w)|^{1+\sigma_0} dz\right)^{\frac{2-p}{1+\sigma_0}\frac{n+1}{n+2}}.$$

As a consequence,

$$\fint_{K_{4R}^\lambda(z_0)} |D(u-w)|^{1+\sigma_0} dz \leq C[\delta\lambda]^{1+\sigma_0}. \tag{2.32}$$

This, along with (2.28), implies that

$$\lambda^{1+\sigma_0} \lesssim \fint_{K_R^\lambda(z_0)} |Dw|^{1+\sigma_0} dz, \quad \fint_{K_{4R}^\lambda(z_0)} |Dw|^{1+\sigma_0} dz \lesssim \lambda^{1+\sigma_0}$$

as long as δ being sufficiently small.

We now apply Proposition 2.9 to find that

$$\lambda^p \lesssim \fint_{K_R^\lambda(z_0)} |Dw|^p dz, \quad \fint_{K_{2R}^\lambda(z_0)} |Dw|^p dz \lesssim \lambda^p. \tag{2.33}$$

Let h be a solution to (2.25). Then from Lemma 2.12 and the fact that $K^\lambda_{R/2}(0,t_0) \subset K^\lambda_R(z_0)$ we have

$$\begin{aligned}\fint_{K^\lambda_{R/2}(0,t_0)} |Dh|^p dz &\lesssim \fint_{K^\lambda_{R/2}(0,t_0)} |D(h-w)|^p dz + \fint_{K^\lambda_{R/2}(0,t_0)} |Dw|^p dz\\ &\lesssim \fint_{K^\lambda_{R/2}(0,t_0)} |D(h-w)|^p dz + \fint_{K^\lambda_{R}(z_0)} |Dw|^p dz\\ &\lesssim \lambda^p.\end{aligned}$$

At this stage, using the similar argument in the proof of [12, Proposition 4.10] we can show that there exists a weak solution v to problem (2.27) such that

$$\|D\hat{v}\|_{L^\infty(Q^\lambda_{R/4}(0,t_0))} \lesssim \lambda, \tag{2.34}$$

and

$$\fint_{K^\lambda_{3R/8}(0,t_0)} |D(h-v)|^p dz \le (\epsilon\lambda)^p \tag{2.35}$$

where $\hat{v}$ is the zero extension of v to $Q^\lambda_{R/2}(0,t_0)$.

With these two estimates in hand, by the fact that $K^\lambda_{R/4}(z_0) \subset K^\lambda_{R/2}(0,t_0) \subset K^\lambda_R(z_0)$, we have

$$\begin{aligned}\fint_{K^\lambda_{R/4}(z_0)} |D(u-v)|^{1+\sigma_0} dz &\lesssim \fint_{K^\lambda_{R/4}(z_0)} |D(u-w)|^{1+\sigma_0} dz\\ &\quad + \fint_{K^\lambda_{R/4}(z_0)} |D(w-h)|^{1+\sigma_0} dz\\ &\quad + \fint_{K^\lambda_{R/4}(z_0)} |D(h-v)|^{1+\sigma_0} dz\\ &\lesssim \fint_{K^\lambda_{R/4}(z_0)} |D(u-w)|^{1+\sigma_0} dz\\ &\quad + \fint_{K^\lambda_{R/2}(0,t_0)} |D(w-h)|^{1+\sigma_0} dz\\ &\quad + \fint_{K^\lambda_{R/2}(0,t_0)} |D(h-v)|^{1+\sigma_0} dz.\end{aligned}$$

At this stage, applying (2.32), (2.26) and (2.35), we implies (2.31).

The estimate (2.30) follows immediately from (2.34) and the Hölder estimates for Dv near the flat boundary

$$\|D\hat{v}\|_{L^\infty(Q^\lambda_{R/8}(z_0))} \le \|D\hat{v}\|_{L^\infty(Q^\lambda_{R/4}(0,t_0))} \text{ (due to (2.23))}.$$

This completes our proof. □

3. Proof of Theorem 1.5: the case $2 - \frac{1}{n+1} < p < 2$

This section is devoted to the proof of Theorem 1.5 for the case $2 - \frac{1}{n+1} < p < 2$.

We assume that $0 < \delta < \frac{1}{50}$ which will be fixed later. Fix $\sigma_0 \in (0, p-2+\frac{1}{n+1}]$ so that $1 + \sigma_0 \in (1, p - 1 + \frac{1}{n+1})$. We set

$$\lambda_0^{1/d_p} := \left[\fint_{\Omega_T} |Du|^{1+\sigma_0} dz\right]^{\frac{1}{1+\sigma_0}} + \frac{1}{\delta}\left[\frac{|\mu|(\Omega_T)}{|\Omega_T|^{\frac{n+1}{n+2}}}\right]^{\frac{n+2}{p+(p-1)n}} + 1, \tag{3.1}$$

where $d_p = \frac{2}{2-n(2-p)}$.

For $\lambda > 0$ the level set $E(\lambda)$ is defined by

$$E(\lambda) = \{z \in \Omega_T : |Du(z)| > \lambda\}.$$

We now fix $w \in A^*_\nu$ for some $\nu \in [1, \infty)$ and take $\tau > 0$ so that

$$\gamma = \nu\tau < q. \tag{3.2}$$

Setting

$$A_0 = \Big(\frac{4^n R_0^{n+2}}{10^{6(n+2)}|\Omega_T|}\Big)^{d_p}, \tag{3.3}$$

then we have the following estimate concerning the level set whose proof will be given after the proof of Theorem 1.5.

Proposition 3.1. *There exists $N_0 > 1$ so that the following holds true. For any $\epsilon > 0$ there exists $\delta > 0$ such that*

$$\begin{aligned} &w(E(N_0\lambda)) \\ &\le \epsilon w(E(\lambda/4)) + \frac{C}{\delta\lambda^\gamma}\int_{\frac{\delta_w^{1/\gamma}\lambda}{4}}^\infty t^{\gamma-1} w(\{z \in \Omega_T : [\mathcal{M}_1^{\mathrm{s}}(\mu)(z)]^{\frac{n+2}{p+(p-1)n}} > t\})dt \end{aligned} \tag{3.4}$$

for any $\lambda > A_0\lambda_0$, where $\delta_w = c_n[w]^{-1}_{A^*_\nu}\delta^\gamma$ with c_n being a constant depending on n only.

We now give the proof of Theorem 1.5 assuming Proposition 3.1.

Proof of Theorem 1.5. We only give the proof for the case $0 < r < \infty$. The remaining case $r = \infty$ can be done in the same manner.

We will prove the theorem under the assumption $|Du| \in L^{q,r}_w(\Omega_T)$. This condition will be removed later. We first have

$$\begin{aligned} \||Du|\|^r_{L^{q,r}_w(\Omega_T)} &= N_0^r \int_0^\infty \big[\lambda^q w(\{z \in \Omega_T : |Du| > N_0\lambda\})\big]^{r/q}\frac{d\lambda}{\lambda} \\ &= N_0^r \int_0^{A_0\lambda_0} \big[\lambda^q w(\{z \in \Omega_T : |Du| > N_0\lambda\})\big]^{r/q}\frac{d\lambda}{\lambda} \\ &\quad + N_0^r \int_{A_0\lambda_0}^\infty \big[\lambda^q w(\{z \in \Omega_T : |Du| > N_0\lambda\})\big]^{r/q}\frac{d\lambda}{\lambda} =: E_1 + E_2. \end{aligned}$$

For the first term E_1 we have

$$E_1 \le [N_0 A_0]^r w(\Omega_T)^{r/q} \lambda_0^r$$
$$\lesssim \left\{ \left[\fint_{\Omega_T} |Du|^{1+\sigma_0} dz \right]^{\frac{1}{1+\sigma_0}} + \left[\frac{1}{\delta} \fint_{\Omega_T} [\mathcal{M}_1^{\mathrm{s}}(\mu)(z)]^{\frac{\tau(n+2)}{p+(p-1)n}} dz \right]^{1/\tau} \right\}^{r d_p}. \quad (3.5)$$

Arguing similarly to the proof of [28, Lemma 4.1], we have

$$\left[\fint_{\Omega_T} |Du|^{1+\sigma_0} dz \right]^{\frac{1}{1+\sigma_0}} \lesssim \left[\frac{|\mu|(\Omega_T)}{|\Omega_T|^{\frac{n+1}{n+2}}} \right]^{\frac{n+2}{p+(p-1)n}}.$$

Moreover,

$$\frac{|\mu|(\Omega_T)}{|\Omega_T|^{\frac{n+1}{n+2}}} \lesssim \inf_{z \in \Omega_T} \mathcal{M}_1^{\mathrm{s}}(\mu)(z).$$

Therefore, we have

$$\left[\fint_{\Omega_T} |Du|^{1+\sigma_0} dz \right]^{\frac{1}{1+\sigma_0}} \lesssim \left[\fint_{\Omega_T} [\mathcal{M}_1^{\mathrm{s}}(\mu)(z)]^{\frac{\tau(n+2)}{p+(p-1)n}} dz \right]^{1/\tau}.$$

Inserting this into (3.5) we obtain

$$E_1 \lesssim \left[\fint_{\Omega_T} [\mathcal{M}_1^{\mathrm{s}}(\mu)(z)]^{\frac{\tau(n+2)}{p+(p-1)n}} dz \right]^{\tau r d_p}.$$

This, along with the embedding $L^\tau(\Omega_T) \hookrightarrow L_w^{q,r}(\Omega_T)$ as $w \in A_\nu^*$ and $\nu\tau < q$, implies that

$$E_1 \lesssim \left\| [\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}} \right\|_{L_w^{q,r}(\Omega_T)}^{r d_p}. \quad (3.6)$$

In order to deal with the second term I_2, applying Proposition 3.1 we get that

$$E_2 \le C\epsilon^{r/q} \int_{A_0\lambda_0}^\infty \left[\lambda^q w(\{z \in \Omega_T : |Du| > \lambda/4\}) \right]^{r/q} \frac{d\lambda}{\lambda}$$
$$+ \frac{C}{\delta^{r/q}} \int_{A_0\lambda_0}^\infty \lambda^{(q-\gamma)r/q} \left[\int_{\frac{\delta_w^{1/\gamma}\lambda}{4}}^\infty t^{\gamma-1} w(\{z \in \Omega_T : [\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}} > t\}) dt \right]^{r/q} \frac{d\lambda}{\lambda}$$
$$\le C_1 \epsilon^{r/q} \| |Du| \|_{L_w^{q,r}(\Omega_T)}^r$$
$$+ \frac{C}{\delta^{r/q}} \int_{A_0\lambda_0}^\infty \lambda^{(q-\gamma)r/q} \left[\int_{\frac{\delta_w^{1/\gamma}\lambda}{4}}^\infty t^{\gamma-1} w(\{z \in \Omega_T : [\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}} > t\}) dt \right]^{r/q} \frac{d\lambda}{\lambda}. \quad (3.7)$$

If $q < r < \infty$, we then apply Hardy's inequality to conclude that

$$\int_{A_0\lambda_0}^{\infty} \lambda^{(q-\gamma)r/q} \left[\int_{\frac{\delta_w^{1/\gamma}\lambda}{4}}^{\infty} t^{\gamma-1} w(\{z \in \Omega_T : [\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}} > t\}) dt \right]^{r/q} \frac{d\lambda}{\lambda}$$
$$\leq C(\delta, w)\epsilon^{r/q} \int_0^{\infty} \lambda^{(q-\gamma)r/q} \lambda^{\gamma r/q} w(\{z \in \Omega_T : [\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}} > \lambda\})^{r/q} \frac{d\lambda}{\lambda}$$
$$= C \left\| [\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}} \right\|_{L_w^{q,r}(\Omega_T)}^r.$$

Hence,

$$E_2 \leq C \left\| [\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}} \right\|_{L_w^{q,r}(\Omega_T)}^r.$$

From the estimates of E_1 and E_2 we have

$$\||Du|\|_{L_w^{q,r}(\Omega_T)}^r \leq C_1\epsilon^{r/q} \||Du|\|_{L_w^{q,r}(\Omega_T)}^r + C \left\| [\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}} \right\|_{L_w^{q,r}(\Omega_T)}^{rd_p}$$
$$+ \left\| [\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}} \right\|_{L_w^{q,r}(\Omega_T)}^r$$
$$\leq C_1\epsilon^{r/q} \||Du|\|_{L_w^{q,r}(\Omega_T)}^r + C \left\| [\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}} \right\|_{L_w^{q,r}(\Omega_T)}^{rd_p}.$$

By choosing ϵ so that $C_1\epsilon^{r/q} < 1$, we then obtain

$$\||Du|\|_{L_w^{q,r}(\Omega_T)}^r \lesssim \left\| [\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}} \right\|_{L_w^{q,r}(\Omega_T)}^{rd_p} + 1$$

as desired.

We now consider the case $0 < r \leq q$. To do this we recall a variant of reverse-Hölder's inequality in [6, Lemma 3.5].

Lemma 3.2. *Let $h : [0, \infty) \to [0, \infty)$ be a non-decreasing, measurable functions and let $1 \leq \alpha < \infty$ and $r > 0$. Then there exists $C > 0$ so that for any $\lambda > 0$ we have*

$$\left[\int_{\lambda}^{\infty} \left(t^r h(t)\right)^{\alpha} \frac{dt}{t} \right]^{1/\alpha} \leq \lambda^r h(\lambda) + C \int_{\lambda}^{\infty} t^r h(t) \frac{dt}{t}.$$

We now turn to the proof for the case $0 < r \leq q$. By Lemma 3.2 we have

$$\left[\int_{\frac{\delta_w^{1/\gamma}\lambda}{4}}^{\infty} t^{\gamma-1} w(\{z \in \Omega_T : [\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}} > t\}) dt \right]^{r/q}$$
$$\leq C\lambda^{\gamma r/q} w(\{z \in \Omega_T : [\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}} > \delta_w^{1/\gamma}\lambda/4\})^{r/q}$$
$$+ C \int_{\frac{\delta_w^{1/\gamma}\lambda}{4}}^{\infty} t^{\gamma r/q} w(\{z \in \Omega_T : [\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}} > t\})^{r/q} \frac{dt}{t}.$$

As a consequence, we have

$$\int_{A_0\lambda_0}^{\infty}\lambda^{(q-\gamma)r/q}\left[\int_{\frac{\delta_w^{1/\gamma}\lambda}{4}}^{\infty}t^{\gamma-1}w(\{z\in\Omega_T:[\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}}>t\})dt\right]^{r/q}\frac{d\lambda}{\lambda}$$
$$\le C(\delta)\int_0^{\infty}\lambda^{(q-\gamma)r/q}\lambda^{\gamma r/q}w(\{z\in\Omega_T:[\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}}>\lambda\})^{r/q}\frac{d\lambda}{\lambda}$$
$$+C\int_0^{\infty}\lambda^{(q-\gamma)r/q}\int_{\frac{\delta_w^{1/\gamma}\lambda}{4}}^{\infty}t^{\gamma r/q}w(\{z\in\Omega_T:[\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}}>t\})^{r/q}\frac{dt}{t}\frac{d\lambda}{\lambda}$$
$$:=F_1+F_2.$$

It is easy to see that

$$F_1\sim\left\|[\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}}\right\|^r_{L_w^{q,r}(\Omega_T)}.$$

Using Fubini's theorem, it is not difficult to show that

$$F_2\lesssim\left\|[\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}}\right\|^r_{L_w^{q,r}(\Omega_T)}.$$

Therefore,

$$\int_{A_0\lambda_0}^{\infty}\lambda^{(q-\gamma)r/q}\left[\int_{\frac{\delta_w^{1/\gamma}\lambda}{4}}^{\infty}t^{\gamma-1}w(\{z\in\Omega_T:[\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}}>t\})dt\right]^{r/q}\frac{d\lambda}{\lambda}$$
$$\lesssim\left\|[\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}}\right\|^r_{L_w^{q,r}(\Omega_T)}.$$

Inserting this into (3.7) we obtain

$$E_2\le C\left\|[\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}}\right\|^r_{L_w^{q,r}(\Omega_T)}.$$

Arguing similarly to the case $q<r<\infty$ one has

$$\||Du|\|^r_{L_w^{q,r}(\Omega_T)}\lesssim\left\|[\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}}\right\|^{rd_p}_{L_w^{q,r}(\Omega_T)}+1.$$

To remove the assumption $|Du|\in L_w^{q,r}(\Omega_T)$ we define

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

and

$$E_k(\lambda)=\{z\in\Omega_T:|Du|_k>\lambda\}$$

for every $k>0$ and $\lambda>0$.

Then we have $|Du|_k \in L_w^{q,r}(\Omega_T)$ and $E_k(\cdot)$ satisfies (3.4) in Proposition 3.1. Therefore, the argument above yields that

$$\| |Du|_k \|^r_{L_w^{q,r}(\Omega_T)} \lesssim \left\| [\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}} \right\|^{rd_p}_{L_w^{q,r}(\Omega_T)} + 1.$$

Then letting $k \to \infty$ we obtain

$$\| |Du| \|^r_{L_w^{q,r}(\Omega_T)} \lesssim \left\| [\mathcal{M}_1^{\mathrm{s}}(\mu)]^{\frac{n+2}{p+(p-1)n}} \right\|^{rd_p}_{L_w^{q,r}(\Omega_T)} + 1.$$

This completes our proof. □

We now prove Proposition 3.1. To do this we need some following technical material.

For $\tilde{z} \in E(\lambda)$, we define

$$G_{\tilde{z}}(r) = \left[\fint_{K_r^\lambda(\tilde{z})} |Du|^{1+\sigma_0} dz \right]^{\frac{1}{1+\sigma_0}} + \frac{1}{\delta} \left[\frac{|\mu|(K_r^\lambda(\tilde{z})}{|K_r^\lambda(\tilde{z})|^{\frac{n+1}{n+2}}} \right]^{\frac{n+2}{p+(p-1)n}}.$$

By Lebesgue's differentiation theorem, we have

$$\lim_{r\to 0} G_{\tilde{z}}(r) = |Du(\tilde{z})| > \lambda. \tag{3.8}$$

Hence, we have, for $10^{-6}\times R_0 < r \leq R_0$,

$$\begin{aligned} G_{\tilde{z}}(r) &= \left[\fint_{K_r^\lambda(\tilde{z})} |Du|^{1+\sigma_0} dz \right]^{\frac{1}{1+\sigma_0}} + \frac{1}{\delta} \left[\frac{|\mu|(K_r^\lambda(\tilde{z})}{|K_r^\lambda(\tilde{z})|^{\frac{n+1}{n+2}}} \right]^{\frac{n+2}{p+(p-1)n}} \\ &\leq \left[\frac{|\Omega_T|}{|K_r^\lambda(\tilde{z})|} \fint_{\Omega_T} |Du|^{1+\sigma_0} dz \right]^{\frac{1}{1+\sigma_0}} + \frac{1}{\delta} \left[\frac{|\Omega_T|}{|K_r^\lambda(\tilde{z})|} \right]^{\frac{n+1}{p+(p-1)n}} \left[\frac{|\mu|(\Omega_T)}{|\Omega_T|^{\frac{n+1}{n+2}}} \right]^{\frac{n+2}{p+(p-1)n}}. \end{aligned}$$

Since $\frac{1}{1+\sigma_0}, \frac{n+1}{p+(p-1)n} < 1$, we have

$$G_{\tilde{z}}(r) \leq \frac{|\Omega_T|}{|K_r^\lambda(\tilde{z})|} \lambda_0^{1/d_p}.$$

On the other hand, from the fact that Ω is a (δ, R_0) Reifenberg domain, we have

$$|K_r^\lambda(\tilde{z})| \geq 4^{-n} r^{n+2} \lambda^{\frac{n(p-2)}{2}}.$$

Hence,

$$G_{\tilde{z}}(r) \leq \frac{10^{6(n+2)} |\Omega_T|}{4^n R_0^{n+2} \lambda^{\frac{n(p-2)}{2}}} \lambda_0^{1/d_p}. \tag{3.9}$$

Then from (3.9), we obtain, for $\lambda > A_0\lambda_0$ with A_0 as in (3.3),

$$G_{\tilde{z}}(r) < \lambda \quad \text{for all } r \in [10^{-6}\times R_0, R_0].$$

This together with (3.8) implies that for each $z \in E(\lambda)$ there exists $0 < r_z < 10^{-6}R_0$ so that

$$G_z(r_z) = \lambda, \quad \text{and } G_z(r) < \lambda \text{ for all } r \in (r_z, 10^{-6}\times R_0).$$

We now apply Vitali's covering lemma to obtain the following result directly.

Lemma 3.3. *There exists a countable disjoint family* $\{K^\lambda_{r_i}(z_i)\}_{i\in\mathcal{I}}$ *with* $r_i < 10^{-6}R_0$ *and* $z_i = (x_i, t_i) \in E(\lambda)$ *such that:*

(a) $E(\lambda) \subset \bigcup_i K^\lambda_{5r_i}(z_i)$;

(b) $G_{z_i}(r_i) = \lambda$, *and* $G_{z_i}(r) < \lambda$ *for all* $r \in (r_i, 10^{-6}R_0)$.

For each i, from Lemma 3.3 we have

$$\lambda = \left[\fint_{K^\lambda_{r_i}(z_i)} |Du|^{1+\sigma_0} dz\right]^{\frac{1}{1+\sigma_0}} + \frac{1}{\delta}\left[\frac{|\mu|(K^\lambda_{r_i}(z_i))}{|K^\lambda_{r_i}(z_i)|^{\frac{n+1}{n+2}}}\right]^{\frac{n+2}{p+(p-1)n}}.$$

This implies that

$$\fint_{K^\lambda_{r_i}(z_i)} |Du|^{1+\sigma_0} dz \geq \frac{\lambda^{1+\sigma_0}}{2^{1+\sigma_0}} \tag{3.10}$$

or

$$\frac{1}{\delta}\left[\frac{|\mu|(K^\lambda_{r_i}(z_i))}{|K^\lambda_{r_i}(z_i)|^{\frac{n+1}{n+2}}}\right]^{\frac{n+2}{p+(p-1)n}} \geq \frac{\lambda}{2}. \tag{3.11}$$

We now set

$$\mathcal{I} = \{i : \ (3.10) \text{ holds true}\}, \quad \mathcal{J} = \{i : \ (3.11) \text{ holds true}\}.$$

We have the following estimate.

Proposition 3.4. *Let* $w \in A^*_\infty$. *For each* $i \in \mathcal{I}$ *we have*

$$w(K^\lambda_{r_i}(z_i)) \lesssim w(K^\lambda_{r_i}(z_i) \cap E(\lambda/4)). \tag{3.12}$$

Proof. We first have

$$\begin{aligned}|K^\lambda_{r_i}(z_i)| &\leq \frac{2^{1+\sigma_0}}{\lambda^{1+\sigma_0}} \int_{K^\lambda_{r_i}(z_i)\setminus E(\lambda/4)} |Du|^{1+\sigma_0} dz + \frac{2^{1+\sigma_0}}{\lambda^{1+\sigma_0}} \int_{K^\lambda_{r_i}(z_i)\cap E(\lambda/4)} |Du|^{1+\sigma_0} dz \\ &\leq \frac{|K^\lambda_{r_i}(z_i)|}{4^{1+\sigma_0}} + \frac{2^{1+\sigma_0}}{\lambda^{1+\sigma_0}} \int_{K^\lambda_{r_i}(z_i)\cap E(\lambda/4)} |Du|^{1+\sigma_0} dz.\end{aligned}$$

This implies

$$|K^\lambda_{r_i}(z_i)| \lesssim \frac{1}{\lambda^{1+\sigma_0}} \int_{K^\lambda_{r_i}(z_i)\cap E(\lambda/4)} |Du|^{1+\sigma_0} dz. \tag{3.13}$$

Note that from the definitions of r_i and the index set $\mathcal{I}$ we have

$$\frac{\lambda^{1+\sigma_0}}{21+\sigma_0} \le \fint_{K^\lambda_{r_i}(z_i)} |Du|^{1+\sigma_0} dz, \quad \fint_{K^\lambda_{4r_i}(z_i)} |Du|^{1+\sigma_0} dz < 3\lambda^{1+\sigma_0}, \tag{3.14}$$

and

$$\left[\frac{|\mu|(K^\lambda_{4r_i}(z_i))}{|K^\lambda_{4r_i}(z_i)|^{\frac{n+1}{n+2}}}\right]^{\frac{n+2}{p+(p-1)n}} \le \delta\lambda. \tag{3.15}$$

By Holder's inequality, for a fixed $\theta > 1$ so that $\theta(1+\sigma_0) \in \left(1, p-1+\frac{1}{n+1}\right)$ we have

$$\begin{aligned} &\left(\frac{1}{|K^\lambda_{r_i}(z_i)|} \int_{K^\lambda_{r_i}(z_i)\cap E(\lambda/4)} |Du|^{1+\sigma_0} dz\right)^{\frac{1}{1+\sigma_0}} \\ &\le \left(\frac{1}{|K^\lambda_{r_i}|} \int_{K^\lambda_{r_i}(z_i)\cap E(\lambda/4)} |Du|^{\theta(1+\sigma_0)} dz\right)^{\frac{1}{\theta(1+\sigma_0)}} \left(\frac{|K^\lambda_{r_i}(z_i)\cap E(\lambda/4)|}{|K^\lambda_{r_i}(z_i)|}\right)^{\frac{\theta-1}{\theta(1+\sigma_0)}}. \end{aligned} \tag{3.16}$$

For each i, consider the following equation

$$\begin{cases} (w^i)_t - \operatorname{div} \mathbf{a}(Dw^i, x, t) = 0 & \text{in} \quad K^\lambda_{4r_i}(z_i) \\ w^i = u & \text{on} \quad \partial_p K^\lambda_{4r_i}(z_i). \end{cases}$$

Then by Lemma 2.11 we have

$$\begin{aligned} &\left(\fint_{K^\lambda_{4r_i}(z_i)} |D(u-w^i)|^{1+\sigma_0} dz\right)^{\frac{1}{1+\sigma_0}} \\ &\le C\left[\frac{|\mu|(K^\lambda_{4R}(z_i))}{|K^\lambda_{4r_i}(z_i)|^{(n+1)/(n+2)}}\right]^{\frac{n+2}{p+(p-1)n}} \\ &\quad + C\frac{|\mu|(K^\lambda_{4R}(z_i))}{|K^\lambda_{4r_i}(z_i)|^{\frac{n+1}{n+2}}} \left(\fint_{K^\lambda_{4r_i}(z_i)} |D(u-w)|^{1+\sigma_0} dz\right)^{\frac{2-p}{1+\sigma_0}\frac{n+1}{n+2}} \\ &\le C\delta\lambda + C[\delta\lambda]^{\frac{p+(p-1)n}{n+2}} \left(\fint_{K^\lambda_{4r_i}(z_i)} |D(u-w^i)|^{1+\sigma_0} dz\right)^{\frac{2-p}{1+\sigma_0}\frac{n+1}{n+2}}. \end{aligned}$$

This implies that

$$\left(\fint_{K^{\lambda}_{4r_i}(z_i)} |D(u-w^i)|^{1+\sigma_0}dz\right)^{\frac{1}{1+\sigma_0}} \le C\delta\lambda.$$

This along with (3.14) yields that

$$\lambda^{1+\sigma_0} \lesssim \fint_{K^{\lambda}_{r_i}(z_i)} |Dw^i|^{1+\sigma_0}dz, \quad \fint_{K^{\lambda}_{4r_i}(z_i)} |Dw^i|^{1+\sigma_0}dz \lesssim \lambda^{1+\sigma_0}, \tag{3.17}$$

provided that δ is sufficiently small.

At this stage, we apply Propositions 2.9 and 2.3 to find that

$$\fint_{K^{\lambda}_{r_i}(z_i)} |Dw^i|^{\theta(1+\sigma_0)}dz \lesssim \lambda^{\theta(1+\sigma_0)},$$

provided that δ is sufficiently small.

Repeating the above argument,

$$\left(\fint_{K^{\lambda}_{4r_i}(z_i)} |D(u-w^i)|^{\theta(1+\sigma_0)}dz\right)^{\frac{\theta}{1+\sigma_0}} \lesssim \delta\lambda.$$

The last two inequalities yield

$$\fint_{K^{\lambda}_{r_i}(z_i)} |Du|^{\theta(1+\sigma_0)}dz \lesssim \lambda^{\theta(1+\sigma_0)},$$

provided that δ is sufficiently small.

Inserting this into (3.16), we get that

$$\left(\frac{1}{|K^{\lambda}_{r_i}(z_i)|}\int_{K^{\lambda}_{r_i}(z_i)\cap E(\lambda/4)} |Du|^{1+\sigma_0}dz\right)^{\frac{1}{1+\sigma_0}} \lesssim \lambda\left(\frac{|K^{\lambda}_{r_i}(z_i)\cap E(\lambda/4)|}{|K^{\lambda}_{r_i}(z_i)|}\right)^{\frac{\theta-1}{\theta(1+\sigma_0)}}.$$

Therefore,

$$\int_{K^{\lambda}_{r_i}(z_i)\cap E(\lambda/4)} |Du|^{1+\sigma_0}dz \lesssim \lambda^{1+\sigma_0}|K^{\lambda}_{r_i}(z_i)|\left(\frac{|K^{\lambda}_{r_i}(z_i)\cap E(\lambda/4)|}{|K^{\lambda}_{r_i}(z_i)|}\right)^{\frac{\theta-1}{\theta}}. \tag{3.18}$$

This, in combination with (3.13), gives that

$$|K^{\lambda}_{r_i}(z)| \lesssim |K^{\lambda}_{r_i}(z_i)|\left(\frac{|K^{\lambda}_{r_i}(z_i)\cap E(\lambda/4)|}{|K^{\lambda}_{r_i}(z_i)|}\right)^{\frac{\theta-1}{\theta}}.$$

As a consequence,

$$|K^{\lambda}_{r_i}(z_i)| \lesssim |K^{\lambda}_{r_i}(z_i) \cap E(\lambda/4)|.$$

This, along with Lemma 1.4, implies

$$w(K^{\lambda}_{r_i}(z_i)) \lesssim w(K^{\lambda}_{r_i}(z_i) \cap E(\lambda/4)). \qquad \square$$

Proposition 3.5. *For each $i \in \mathcal{J}$ we have*

$$w(K^{\lambda}_{r_i}(z_i)) \lesssim \frac{1}{\delta_w \lambda^{\gamma}} \int_{\frac{\delta_w^{1/\gamma}\lambda}{4}}^{\infty} t^{\gamma-1} w\left(\left\{z \in K^{\lambda}_{r_i}(z_i) : [\mathcal{M}^{\mathrm{s}}_1(\mu)(z)]^{\frac{n+2}{p+(p-1)n}} > t\right\}\right) dt, \quad (3.19)$$

where γ is defined by (3.2) *and $\delta_w = c_n [w]^{-1}_{A^*_\nu} \delta^{\gamma}$ with c_n being a constant depending on n only.*

Proof. For $i \in \mathcal{J}$, by Hölder's inequality we have

$$\begin{aligned}
&\left[⨍_{K^{\lambda}_{r_i}(z_i)} [\mathcal{M}^{\mathrm{s}}_1(\mu)(z)]^{\frac{\tau(n+2)}{p+(p-1)n}} dz \right]^{1/\tau} \\
&\leq \left[\frac{1}{w(K^{\lambda}_{r_i}(z_i))} \int_{K^{\lambda}_{r_i}(z_i)} [\mathcal{M}^{\mathrm{s}}_1(\mu)(z)]^{\frac{\nu\tau(n+2)}{p+(p-1)n}} w(z) dz \right]^{\frac{1}{\nu\tau}} \\
&\quad \times \left(⨍_{K^{\lambda}_{r_i}(z_i)} w(z) dz \right)^{\frac{1}{\nu\tau}} \left(⨍_{K^{\lambda}_{r_i}(z_i)} w(z)^{-\nu'/\nu} dz \right)^{\frac{1}{\nu'\tau}}.
\end{aligned}$$

By the definition of $[w]_{A^*_\nu}$ and $\gamma = \nu\tau$ we have

$$\begin{aligned}
&\left[⨍_{K^{\lambda}_{r_i}(z_i)} [\mathcal{M}^{\mathrm{s}}_1(\mu)(z)]^{\frac{\tau(n+2)}{p+(p-1)n}} dz \right]^{1/\tau} \\
&\leq [w]^{\frac{1}{\gamma}}_{A^*_\nu} \left[\frac{1}{w(K^{\lambda}_{r_i}(z_i))} \int_{K^{\lambda}_{r_i}(z_i)} [\mathcal{M}^{\mathrm{s}}_1(\mu)(z)]^{\frac{\gamma(n+2)}{p+(p-1)n}} w(z) dz \right]^{\frac{1}{\gamma}}.
\end{aligned} \quad (3.20)$$

On the other hand, from the definition of $\mathcal{M}^{\mathrm{s}}_1$ it easy to see that

$$\frac{|\mu|(K^{\lambda}_{r_i}(z_i))}{|K^{\lambda}_{r_i}(z_i)|^{\frac{n+1}{n+2}}} \leq c_n \inf_{z \in K^{\lambda}_{r_i}(z_i)} \mathcal{M}^{\mathrm{s}}_1(\mu)(z).$$

Therefore, for each $i \in \mathcal{J}$ we have

$$\left[⨍_{K^{\lambda}_{r_i}(z_i)} [\mathcal{M}^{\mathrm{s}}_1(\mu)(z)]^{\frac{\tau(n+2)}{p+(p-1)n}} dz \right]^{1/\tau} \geq c_n \left[\frac{|\mu|(K^{\lambda}_{r_i}(z_i))}{|K^{\lambda}_{r_i}(z_i)|^{\frac{n+1}{n+2}}} \right]^{\frac{n+2}{p+(p-1)n}} \geq \frac{c_n \delta\lambda}{2}.$$

This, along with (3.20), implies that

$$\frac{1}{w(K^\lambda_{r_i}(z_i))}\int_{K^\lambda_{r_i}(z_i)}[\mathcal{M}^{\mathrm{s}}_1(\mu)(z)]^{\frac{\gamma(n+2)}{p+(p-1)n}}w(z)dz \ge \frac{c_n[w]^{-1}_{A^*_\nu}\delta^\gamma\lambda^\gamma}{2^\gamma} =: \frac{\delta_w\lambda^\gamma}{2^\gamma}. \quad (3.21)$$

Therefore,

$$\begin{aligned} w(K^\lambda_{r_i}(z_i)) &\le \frac{2^\gamma}{\delta_w\lambda^\gamma}\int_{K^\lambda_{r_i}(z_i)}[\mathcal{M}^{\mathrm{s}}_1(\mu)(z)]^{\frac{\gamma(n+2)}{p+(p-1)n}}w(z)dz \\ &= \frac{\gamma 2^\gamma}{\delta_w\lambda^\gamma}\int_0^\infty t^{\gamma-1}w\Big(\{z\in K^\lambda_{r_i}(z_i):[\mathcal{M}^{\mathrm{s}}_1(\mu)(z)]^{\frac{n+2}{p+(p-1)n}}>t\}\Big)dt \\ &= \frac{2^\gamma}{\delta_w\lambda^\gamma}\int_0^{\frac{\delta_w^{1/\gamma}\lambda}{4}} t^{\gamma-1}w\Big(\{z\in K^\lambda_{r_i}(z_i):[\mathcal{M}^{\mathrm{s}}_1(\mu)(z)]^{\frac{n+2}{p+(p-1)n}}>t\}\Big)dt \\ &\quad + \frac{\gamma 2^\gamma}{\delta_w\lambda^\gamma}\int_{\frac{\delta_w^{1/\gamma}\lambda}{4}}^\infty t^{\gamma-1}w\Big(\{z\in K^\lambda_{r_i}(z_i):[\mathcal{M}^{\mathrm{s}}_1(\mu)(z)]^{\frac{n+2}{p+(p-1)n}}>t\}\Big)dt. \end{aligned}$$

This implies that

$$\begin{aligned} w(K^\lambda_{r_i}(z_i)) &\le \frac{1}{2^\gamma}|K^\lambda_{r_i}(z_i)| \\ &\quad + \frac{\gamma 2^\gamma}{\delta_w\lambda^\gamma}\int_{\frac{\delta_w^{1/\gamma}\lambda}{4}}^\infty t^{\gamma-1}w\Big(\{z\in K^\lambda_{r_i}(z_i):[\mathcal{M}^{\mathrm{s}}_1(\mu)(z)]^{\frac{n+2}{p+(p-1)n}}>t\}\Big)dt. \end{aligned}$$

This proves (3.19). □

We are now ready to give the proof of Proposition 3.1.

Proof of Proposition 3.1. We have, by Lemma 3.3,

$$\begin{aligned} |E(N_0\lambda)| &= |\{z\in E(\lambda):|Du|>N_0\lambda\}| \\ &\le \sum_{i\in\mathcal{I}}|\{z\in K^\lambda_{5r_i}(z_i):|Du|>N_0\lambda\}| \\ &\quad + \sum_{i\in\mathcal{J}}|\{z\in K^\lambda_{5r_i}(z_i):|Du|>N_0\lambda\}| =: I_1+I_2. \end{aligned}$$

By Proposition 3.5 we have

$$\begin{aligned} I_2 &\le \sum_{i\in\mathcal{J}}|K^\lambda_{5r_i}(z_i)| \lesssim \sum_{i\in\mathcal{J}}|K^\lambda_{r_i}(z_i)| \\ &\lesssim \sum_{i\in\mathcal{J}}\frac{1}{\delta_w\lambda^\gamma}\int_{\frac{\delta_w^{1/\gamma}\lambda}{4}}^\infty t^{\gamma-1}|\{z\in K^\lambda_{r_i}(z_i):[\mathcal{M}^{\mathrm{s}}_1(\mu)(z)]^{\frac{n+2}{p+(p-1)n}}>t\}|dt \\ &\lesssim \frac{1}{\delta_w\lambda^\gamma}\int_{\frac{\delta_w^{1/\gamma}\lambda}{4}}^\infty t^{\gamma-1}|\{z\in\Omega_T:[\mathcal{M}^{\mathrm{s}}_1(\mu)(z)]^{\frac{n+2}{p+(p-1)n}}>t\}|dt. \end{aligned}$$

For the term I_1, we consider the following two cases.

Case 1: $B^{\lambda}_{40r_i}(x_i) \subset \Omega$**.** In this situation, by Lemma 3.3 and the definition of the index set $\mathcal{I}$, we have

$$\lambda^{1+\sigma_0} \lesssim \fint_{K^{\lambda}_{10r_i}(z_i)} |Du|^{1+\sigma_0} dz, \quad \fint_{K^{\lambda}_{40r_i}(z_i)} |Du|^{1+\sigma_0} dz \lesssim \lambda^{1+\sigma_0}$$

and

$$\left[\frac{|\mu|(K^{\lambda}_{40r_i}(z_i))}{|K^{\lambda}_{40r_i}(z_i)|^{\frac{n+1}{n+2}}} \right]^{\frac{n+2}{p+(p-1)n}} \leq \delta\lambda.$$

Then we apply Proposition 2.5 to find that for any $\tilde{\epsilon}$ there exist $\delta > 0$ and v^i satisfying

$$\|Dv^i\|_{L^\infty(K^{\lambda}_{5r_i})(z_i)} \leq a_1\lambda, \tag{3.22}$$

and

$$\fint_{K^{\lambda}_{5r_i})(z_i)} |D(u - v^i)| dz \leq \tilde{\epsilon}\lambda. \tag{3.23}$$

Therefore, for $N_0 > 2a_1$ we have, by (3.23),

$$\begin{aligned} |\{z \in K^{\lambda}_{5r_i}(z_i) : |Du| > N_0/2\lambda\}| &\leq |\{z \in K^{\lambda}_{5r_i}(z_i) : |D(u - v^i)| > N_0/2\lambda\}| \\ &\quad + |\{z \in K^{\lambda}_{5r_i}(z_i) : |Dv^i| > N_0/2\lambda\}| \\ &\leq |\{z \in K^{\lambda}_{5r_i}(z_i) : |D(u - v^i)| > N_0/2\lambda\}| \\ &\leq \frac{2\lambda}{N_0} \int_{K^{\lambda}_{5r_i}(z_i)} |D(u - v^i)| dz \\ &\leq C\tilde{\epsilon}|K^{\lambda}_{5r_i}(z_i)| \leq C\tilde{\epsilon}|K^{\lambda}_{r_i}(z_i)| \\ &\leq C\tilde{\epsilon}|K^{\lambda}_{r_i}(z_i) \cap E(\lambda/4)|, \end{aligned}$$

where in the last inequality we used Proposition 3.4.

Case 2: $B^{\lambda}_{40r_i}(x_i) \cap \Omega^c \neq \emptyset$**.** Then there exists $x^0_i \in B^{\lambda}_{40r_i}(x_i) \cap \Omega^c$. In this case, we have

$$K^{\lambda}_{5r_i}(z_i) \subset K^{\lambda}_{45r_i}(x^0_i, t_i) \subset K^{\lambda}_{1440r_i}(x^i_0, t_i) \subset K^{\lambda}_{1500r_i}(z_i).$$

This, along with Lemma 3.3 and the definition of the index set $\mathcal{I}$, implies that

$$\lambda^{1+\sigma_0} \lesssim \fint_{K^{\lambda}_{1440r_i}(x^i_0,t_i)} |Du|^{1+\sigma_0} dz, \quad \fint_{K^{\lambda}_{1440r_i}(x^i_0,t_i)} |Du|^{1+\sigma_0} dz \lesssim \lambda^{1+\sigma_0},$$

and

$$\left[\frac{|\mu|(K^\lambda_{1440r_i}(x^i_0,t_i))}{|K^\lambda_{1440r_i}(x^i_0,t_i)|^{\frac{n+1}{n+2}}}\right]^{\frac{n+2}{p+(p-1)n}} \le \delta\lambda.$$

Hence, applying Proposition 2.13 we deduce that for any $\tilde{\epsilon}$ we can find $\delta > 0$ and v^i so that

$$\|Dv^i\|_{L^\infty(K^\lambda_{45r_i}(x^i_0,t_i))} \lesssim \lambda,$$

and

$$\fint_{K^\lambda_{1440r_i}(x^i_0,t_i)} |D(u-v^i)|^{1+\sigma_0}dz \lesssim (\epsilon\lambda)^{1+\sigma_0}.$$

Since $K^\lambda_{5r_i}(z_i) \subset K^\lambda_{1440r_i}(x^i_0,t_i)$, we have

$$\|Dv^i\|_{L^\infty(K^\lambda_{5r_i}(z_i))} \lesssim a_2\lambda, \tag{3.24}$$

and

$$\fint_{K^\lambda_{5r_i}(z_i)} |D(u-v^i)|^{1+\sigma_0}dz \le C(\tilde{\epsilon}\lambda)^{1+\sigma_0}. \tag{3.25}$$

Therefore, for $N_0 > 2a_2$ we have, by (3.23),

$$\begin{aligned}
|\{z \in K^\lambda_{5r_i}(z_i) : |Du| > N_0/2\lambda\}| &\le |\{z \in K^\lambda_{5r_i}(z_i) : |D(u-v^i)| > N_0/2\lambda\}| \\
&\quad + |\{z \in K^\lambda_{5r_i}(z_i) : |Dv^i| > N_0/2\lambda\}| \\
&\le |\{z \in K^\lambda_{5r_i}(z_i) : |D(u-v^i)| > N_0/2\lambda\}| \\
&\le \left(\frac{2\lambda}{N_0}\right)^{1+\sigma_0} \int_{K^\lambda_{5r_i}(z_i)} |D(u-v^i)|^{1+\sigma_0}dz.
\end{aligned}$$

At this stage, arguing similarly to Case 1, we come up with

$$|\{z \in K^\lambda_{5r_i}(z_i) : |Du| > N_0/2\lambda\}| \le C\tilde{\epsilon}|K^\lambda_{r_i}(z_i) \cap E(\lambda/4)|.$$

Therefore, from the estiates in Cases 1 and 2, taking $N = 2(a_1+a_2)+1$ and taking $\tilde{\epsilon}$ to be sufficiently small, we have

$$I_1 \le \sum_{i\in\mathcal{I}} \epsilon|K^\lambda_{r_i}(z_i) \cap E(\lambda/4)| \le \epsilon|E(\lambda/4)|$$

where in the lase inequality we used the fact that $K^\lambda_{r_i}(z_i)$ are pairwise disjoint.

This completes our proof. □

Part 2

Global weighted estimates for the gradient of solutions: the case $p \geq 2$

For $z = (x, t)$ with $x \in \mathbb{R}^n$, $t > 0$ and $\lambda \geq 1$, in this part we use the following notation:

- $I_r^\lambda(t) = (t - \lambda^{2-p} r^2, t + \lambda^{2-p} r^2)$, $\Omega_r(x) = B_r(x) \cap \Omega$, $Q_r^\lambda(z) = B_r(x) \times I_r^\lambda(t)$.
- $K_r^\lambda(z) = Q_r^\lambda(z) \cap \Omega_T$, $\partial_w K_r^\lambda(z) = Q_r^\lambda(z) \cap (\partial\Omega \times \mathbb{R})$, $\partial_p K_r^\lambda(z) = \partial K_r^\lambda(z) \setminus (\bar{\Omega}_r^\lambda(x) \times \{t + r^2\})$.

We note that the main difference with the case $2 - \frac{1}{n+1} < p < 2$ is that we use the intrinsic cylinder of the form $Q_r^\lambda(z) = B_r(x) \times (t - \lambda^{2-p} r^2, t + \lambda^{2-p} r^2)$ instead of $Q_r^\lambda(z) = B_{\lambda^{\frac{p-2}{2}} r}(x) \times (t - r^2, t + r^2)$.

4. Comparison estimates

4.1. Interior estimates

For $z_0 = (x_0, t_0) \in \Omega_T$, $0 < R < R_0/4$ and $\lambda \geq 1$ satisfying $B_{4R} \equiv B_{4R}(x_0) \subset \Omega$, we set

$$Q_{4R}^\lambda \equiv Q_{4R}^\lambda(z_0) = B_{4R}(x_0) \times I_{4R}^\lambda(t_0). \tag{4.1}$$

For the sake of simplicity, we may assume that $I_{4R}^\lambda(t_0) \subset (0, T)$, or equivalently, $Q_{4R}^\lambda \subset \Omega_T$.

Arguing similarly to the proof of Proposition 3.5 in [13] and Proposition 2.5 with some minor modifications, we obtain:

Proposition 4.1. *For each* $\epsilon > 0$ *there exists* $\delta > 0$ *so that the following holds true. If* u *is a weak solution to problem* (1.1) *satisfying*

$$\lambda^{p-1} \leq \fint_{Q_R^\lambda} |Du|^{p-1} dxdt \ \text{ and } \ \fint_{Q_{4R}^\lambda} |Du|^{p-1} dxdt \leq \lambda^{p-1} \tag{4.2}$$

and

$$\left[\frac{|\mu|(Q_{4R}^\lambda)}{|Q_{4R}^\lambda|^{(n+1)/(n+2)}} \right]^{\frac{n+2}{p+(p-1)n}} \leq \delta\lambda, \tag{4.3}$$

then there exists v *satisfying*

$$\|Dv\|_{L^\infty(Q_{R/2}^\lambda)} \lesssim \lambda, \tag{4.4}$$

and

$$\fint_{Q_R^\lambda} |D(u - v)|^{p-1} dxdt \leq (\epsilon\lambda)^{p-1}. \tag{4.5}$$

4.2. Boundary estimates

Fix $t_0 \in (0, T)$ and $x_0 \in \partial\Omega$, we set $z_0 = (x_0, t_0)$. Let $0 < R < R_0/4$ and $\lambda \geq 1$. For the sake of simplicity, we restrict ourself to consider the lateral boundary case with respect to

$$I^{\lambda}_{4R}(t_0) \subset (0, T),$$

since the initial boundary case can be done in the same manner.

Then by an argument used in the proof of [13, Proposition 4.10] and Proposition 2.13 with a minor modification we also have:

Proposition 4.2. *For each $\epsilon > 0$ there exists $\delta > 0$ so that the following holds true. If that u is a weak solution to problem* (1.1) *satisfying*

$$\lambda^{p-1} \leq \fint_{K^{\lambda}_{R}(z_0)} |Du|^{p-1} dxdt, \quad \fint_{K^{\lambda}_{4R}(z_0)} |Du|^{p-1} dxdt \leq \lambda^{p-1} \tag{4.6}$$

and

$$\left[\frac{|\mu|(K^{\lambda}_{4R}(z_0))}{|K^{\lambda}_{4R}(z_0)|^{(n+1)/(n+2)}}\right]^{\frac{n+2}{p+(p-1)n}}, \tag{4.7}$$

then there exists v satisfying

$$\|Dv\|_{L^{\infty}(K^{\lambda}_{R/8}(z_0))} \lesssim \lambda, \tag{4.8}$$

and

$$\fint_{K^{\lambda}_{R/4}(z_0)} |D(u - v)|^{p-1} dxdt \leq (\epsilon\lambda)^{p-1}. \tag{4.9}$$

5. Proof of Theorem 1.5: the case $p \geq 2$

Since the proof in this case is similar to that of the case $2 - \frac{1}{n+1} < p < 2$, we just sketch it.

We now fix $w \in A^*_{\nu}$ for some $\nu \in [1, \infty)$ and take $\tau > 0$ so that

$$\gamma = \nu\tau < q. \tag{5.1}$$

We assume that $0 < \delta < \frac{1}{50}$. We set

$$\lambda_0 := \fint_{\Omega_T} |Du|^{p-1} dz + \frac{1}{\delta}\left[\frac{|\mu|(\Omega_T)}{|\Omega_T|^{\frac{n+1}{n+2}}}\right]^{\frac{(p-1)(n+2)}{p+(p-1)n}} + 1. \tag{5.2}$$

For $\lambda > 0$ the level set $E(\lambda)$ is defined by

$$E(\lambda) = \{z \in \Omega_T : |Du(z)| > \lambda\}.$$

For $\tilde{z} \in E(\lambda)$, we define

$$G_{\tilde{z}}(r) = \fint_{K_r^\lambda(\tilde{z})} |Du|^{p-1} dz + \frac{1}{\delta}\left[\frac{|\mu|(K_r^\lambda(\tilde{z})}{|K_r^\lambda(\tilde{z})|^{\frac{n+1}{n+2}}}\right]^{\frac{(p-1)(n+2)}{p+(p-1)n}}.$$

By Lebesgue's differentiation theorem, we have

$$\lim_{r\to 0} G_{\tilde{z}}(r) = |Du(\tilde{z})|^{p-1} > \lambda^{p-1}. \tag{5.3}$$

Note that $\frac{(p-1)(n+1)}{p+(p-1)n} < 1$. Arguing similarly to (3.9) we have, for $10^{-6}\times R_0 < r \le R_0$,

$$G_{\tilde{z}}(r) \le \frac{10^{6(n+2)}|\Omega_T|}{4^n R_0^{n+2}\lambda^{2-p}}\lambda_0. \tag{5.4}$$

Set

$$B_0 = \frac{4^n R_0^{n+2}}{10^{6(n+2)}|\Omega_T|}. \tag{5.5}$$

Then for $\lambda > B_0\lambda_0$ we have

$$G_{\tilde{z}}(r) < \lambda^{p-1}, \quad \text{for all } r \in [10^{-6}\times R_0, R_0].$$

This together with (5.3)

$$G_z(r_z) = \lambda^{p-1}, \quad \text{and } G_z(r) < \lambda^{p-1} \text{ for all } r \in (r_z, 10^{-6}\times R_0).$$

At this stage, arguing similarly to the proof of Proposition 3.1, we have:

Proposition 5.1. *There exists $N_0 > 1$ so that the following holds true. For any $\epsilon > 0$ there exists $\delta > 0$ such that*

$$w(E(N_0\lambda)) \le \epsilon w(E(\lambda/4)) + \frac{C}{\delta\lambda^\gamma}\int_{\frac{\delta_w^{1/\gamma}\lambda}{4}}^\infty t^{\gamma-1} w(\{z \in \Omega_T : [\mathcal{M}_1^s(\mu)(z)]^{\frac{(p-1)(n+2)}{p+(p-1)n}} > t\})dt$$

for all $\lambda > B_0\lambda_0$, where $\delta_w = c_n[w]_{A_\nu^*}^{-1}\delta^\gamma$ with c_n being a constant depending on n only.

Now the proof of Theorem 1.5 for the case $p \ge 2$ follows immediately by using the similar argument to that of the case $2 - \frac{1}{n+1} < p < 2$. Hence, we would like to leave to the interested reader. This completes the proof.

References

[1] E. ACERBI and G. MINGIONE, *Gradient estimates for the $p(x)$-Laplacean system*, J. Reine Angew. Math. **584** (2005), 117–148.

[2] E. ACERBI and G. MINGIONE, *Gradient estimates for a class of parabolic systems*, Duke Math. J. **136** (2007), 285–320.

[3] R. BAGBY and D. S. KURTZ, *$L(\log L)$ spaces and weights for the strong maximal function*, J. Analyse Math. **44** (1984/85), 21–31.

[4] P. BARONI, *Marcinkiewicz estimates for degenerate parabolic equations with measure data*, J. Funct. Anal. **267** (2014), 3397–3426.

[5] P. BARONI, *Singular parabolic equations, measures satisfying density conditions, and gradient integrability*, preprint. Available at http://cvgmt.sns.it/paper/3217/.

[6] P. BARONI, *Lorentz estimates for degenerate and singular evolutionary systems*, J. Differential Equations **255** (2013), 2927–2951.

[7] P. BARONI and J. HABERMANN, *New gradient estimates for parabolic equations*, Houston J. Math. **38** (3) (2012), 855–914.

[8] 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), 241–273.

[9] L. BOCCARDO and T. GALLOUËT, *Nonlinear elliptic and parabolic equations involving measure data*, J. Funct. Anal. **87** (1989), 149–169.

[10] L. BOCCARDO and T. GALLOUËT, *Nonlinear elliptic equations with right-hand side measures*, Comm. Partial Differential Equations **17** (1992), 641–655.

[11] L. BOCCARDO, A. DALL'AGLIO, T. GALLOUËT and L. ORSINA, *Nonlinear parabolic equations with measure data*, J. Funct. Anal. **147** (1) (1997), 237–258.

[12] T. A. BUI and X. T. DUONG, *Global Marcinkiewicz estimates for nonlinear parabolic equations with nonsmooth coefficients*, to appear in Ann. Sc. Norm. Super. Pisa, Cl. Sci.

[13] T. A. BUI and X. T. DUONG, *Global Lorentz estimates for nonlinear parabolic equations on nonsmooth domains*, to appear in Calc. Var. Partial Differential Equations.

[14] S-S. BYUN and L. WANG, *Parabolic equations in time dependent Reifenberg domains*, Adv. Math. **212** (2007), 797–818.

[15] S-S. BYUN, J. OK and S. RYU, *Global gradient estimates for general nonlinear parabolic equations in nonsmooth domains*, J. Differential Equations **254** (2013), 4290–4326.

[16] L. A. CAFFARELLI and I. PERAL, *On $W^{1,p}$ estimates for elliptic equations in divergence form*, Comm. Pure Appl. Math. **51** (1998), 1–21.

[17] A. CÓRDOBA and R. FEFFERMAN, *A geometric proof of the strong maximal theorem*, Ann. of Math. (2) **102** (1975), 95–100.

[18] 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.

[19] G. DAVID and T. TORO, *A generalization of Reifenberg theorem in $\mathbb{R}^3$*, Geom. Funct. Anal. **18** (4) (2008), 1168–1235.

[20] E. DIBENEDETTO, "Degenerate Parabolic Equations", Universitext, Springer, New York, 1993.

[21] G. DOLZMANN, N. HUNGERBÜHLER and S. MÜLLER, *Uniqueness and maximal regularity for nonlinear elliptic systems of n-Laplace type with measure valued right hand side*, J. Reine Angew. Math. **520** (2000), 1–35.

[22] F. DUZAAR and G. MINGIONE, *Gradient estimates via non-linear potentials*, Amer. J. Math. **133** (2011), 1093–1149.

[23] F. DUZAAR and G. MINGIONE, *Gradient estimates via linear and nonlinear potentials*, J. Funct. Anal. **259** (2010), 2961–2998.

[24] Q. HAN and F. LIN, "Elliptic Partial Differential Equation", Courant Institute of Mathematical Sciences/New York University, New York, 1997.

[25] T. KILPELÄINEN and P. KOSKELA, *Global integrability of the gradients of solutions to partial differential equations*, Nonlinear Anal. **23** (1994), 899–909.

[26] J. KINNUNEN and J. L. LEWIS, *Higher integrability for parabolic systems of p–Laplacian type*, Duke Math. J. **102** (2000), 253–271.

[27] N. V. KRYLOV, *Parabolic and elliptic equations with VMO coefficients*, Comm. Partial Differential Equations **32** (2007), 453–475.

[28] T. KUUSI and G. MINGIONE, *Gradient regularity for nonlinear parabolic equations*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **12** (2013), 755–822.

[29] T. KUUSI and G. MINGIONE, *Potential estimates and gradient boundedness for nonlinear parabolic systems*, Rev. Mat. Iberoam. **28** (2) (2012), 535–576.

[30] T. KUUSI and G. MINGIONE, *The Wolff gradient bound for degenerate parabolic equations*, J. Eur. Math. Soc. (JEMS) **16** (4) (2014), 835–892.

[31] M. GIAQUINTA, "Multiple Integrals in the Calculus of Variations and Nonlinear Elliptic Systems", Annals of Mathematics Studies, Vol. 105, Princeton University Press, Princeton, NJ, 1983.

[32] M. GIAQUINTA and G. MODICA, *Regularity results for some classes of higher order nonlinear elliptic systems*, J. Reine Angew. Math. **311/312** (1979), 145–169.

[33] T. IWANIEC and C. SBORDONE, *Weak minima of variational integrals*, J. Reine Angew. Math. **45** (1994), 143–161.

[34] D. JERISON and C. KENIG, *The inhomogeneous Dirichlet problem in Lipschitz domains*, J. Funct. Anal. **130** (1995), 161–219.

[35] O. A. LADYZHENSKAJA, V. A. SOLONNIKOV and N. N. URALCEVA, "Linear and Quasi-linear Equations of Parabolic Type", translated from the Russian by S. Smith, Transl. Math. Monogr., Vol. 23, Amer. Math. Soc., Providence, RI, 1967.

[36] G. M. LIEBERMAN, *Boundary regularity for solutions of degenerate parabolic equations*, Nonlinear Anal. **14** (1990), 501–524.

[37] V. G. MAZ'YA, "Sobolev Spaces", Springer Series in Soviet Mathematics, Springer, Berlin, 1985.

[38] M. MARCUS and L. VÉRON, "Laurent Nonlinear Second Order Elliptic Equations Involving Measures", De Gruyter Series in Nonlinear Analysis and Applications, Vol. 21, De Gruyter, Berlin, 2014.

[39] G. MINGIONE, *The Calderón–Zygmund theory for elliptic problems with measure data*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **6** (2007), 195–261.

[40] G. MINGIONE, *Gradient estimates below the duality exponent*, Math. Ann. **346** (2010), 571–627.

[41] G. MINGIONE, *Gradient potential estimates*, J. Eur. Math. Soc. (JEMS) **13** (2011), 459–486.

[42] 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), 3927–3948.

[43] M. PARVIAINEN, *Reverse Hölder inequalities for singular parabolic equations near the boundary*, J. Differential Equations **246** (2009), 512–540.

[44] M. PARVIAINEN, *Global gradient estimates for degenerate parabolic equations in nonsmooth domains*, Ann. Mat. Pura Appl. **188** (2009), 333–358.

[45] 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.

[46] N. C. PHUC, *Morrey global bounds and quasilinear Riccati type equations below the natural exponent*, J. Math. Pures Appl. **102** (2014), 99–123.

[47] E. REIFENBERG, *Solutions of the Plateau problem for m-dimensional surfaces of varying topological type*, Acta Math. **104** (1960), 1–92.

[48] D. SARASON, *Functions of vanishing mean oscillation*, Trans. Amer. Math. Soc. **207** (1975), 391–405.

[49] R. E. SHOWALTER, "Monotone Operators in Banach Space and Nonlinear Partial Differential Equations", Math. Surveys Monograph, Vol. 49, American Mathematical Society, Providence, RI, 1997.

[50] T. SJÖDIN, *Weighted L^p-inequalities for multi–parameter Riesz type potentials and strong fractional maximal operators*, Math. Ann. **337** (2007), 317–333.

[51] T. TORO, *Doubling and flatness: geometry of measures*, Notices Amer. Math. Soc. **44** (1997), 1087–1094.

Department of Mathematics and Statistics
Macquarie University
NSW 2109, Australia
the.bui@mq.edu.au
bt_anh80@yahoo.com