Gradient weighted estimates at the natural exponent for Quasilinear Parabolic equations.

Karthik Adimurthi[☆a,*], Sun-Sig Byun[☆☆a,b]

[a] *Department of Mathematical Sciences, Seoul National University, Seoul 08826, Korea.*
[b] *Research Institute of Mathematics, Seoul National University, Seoul 08826, Korea.*

Abstract

In this paper, we obtain weighted norm inequalities for the spatial gradients of weak solutions to quasilinear parabolic equations with weights in the Muckenhoupt class $A_{\frac{q}{p}}(\mathbb{R}^{n+1})$ for $q \geq p$ on non-smooth domains. Here the quasilinear nonlinearity is modelled after the standard p-Laplacian operator. Until now, all the weighted estimates for the gradient were obtained only for exponents $q > p$. The results for exponents $q > p$ used the full complicated machinery of the Calderón-Zygmund theory developed over the past few decades, but the constants blow up as $q \to p$ (essentially because the Maximal function is not bounded on L^1).

In order to prove the weighted estimates for the gradient at the natural exponent, i.e., $q = p$, we need to obtain improved a priori estimates below the natural exponent. To this end, we develop the technique of Lipschitz truncation based on [3, 26] and obtain significantly improved estimates below the natural exponent. Along the way, we also obtain improved, unweighted Calderón-Zygmund type estimates below the natural exponent which is new even for the linear equations.

Keywords: Quasilinear parabolic equations, Muckenhoupt weights, Lipschitz truncation
2010 MSC: 35K10, 99-00, 35K59, 35K65, 35K67.

Contents

[☆] Supported by the National Research Foundation of Korea grant NRF-2017R1A2B2003877.
[☆☆] Supported by the National Research Foundation of Korea grant NRF-2015R1A4A1041675.
[*] Corresponding author
Email addresses: karthikaditi@gmail.com and kadimurthi@snu.ac.kr (Karthik Adimurthi[☆]), byun@snu.ac.kr (Sun-Sig Byun[☆☆])

1. Introduction

In this paper, we are interested in obtaining Calderón-Zygmund type regularity estimates in weighted Lebesgue spaces for equations of the form

$$\begin{cases} u_t - \operatorname{div} \mathcal{A}(x,t,\nabla u) = \operatorname{div} |\mathbf{f}|^{p-2}\mathbf{f} & \text{in } \Omega \times (-T,T), \\ u = 0 & \text{on } \partial_p\left(\Omega \times (-T,T)\right), \end{cases} \tag{1.1}$$

where the nonlinearity $\mathcal{A}(x,t,\nabla u)$ is modelled after the well studied p-Laplacian operator given by $|\nabla u|^{p-2}\nabla u$ in a bounded domain $\Omega \subset \mathbb{R}^n$ with $n \geq 2$, potentially with non-smooth boundary $\partial\Omega$. The parabolic boundary is given by $\partial_p\left(\Omega \times (-T,T)\right) := \partial\Omega \times (-T,T) \bigcup \Omega \times \{-T\}$.

Over the past decades, there have been a plethora of a priori estimates of the Calderón-Zygmund type obtained for (1.1). We shall point out that all the estimates discussed in the introduction are quantitative, but in order to highlight the novelty of the results in this paper, we shall only mention the qualitative nature of the estimates existing in the literature.

The first extension of the Calderón-Zygmund theory for (1.1) with $\mathcal{A}(x,t,\nabla u) = |\nabla u|^{p-2}\nabla u$ for $p > \dfrac{2n}{n+2}$ (note that this restriction is natural for parabolic problems, see [19, Chapter 5]) was obtained in [1], where they proved

$$|\mathbf{f}| \in L^q_{loc} \quad \Longrightarrow \quad |\nabla u| \in L^q_{loc} \qquad \text{for all} \qquad q \geq p.$$

Since then, many extensions were obtained which generalized the estimates in [1] to more general nonlinearities, function spaces and up to the boundary (see [6, 7, 12, 13, 22, 30] and the references therein). *In this paper, the first result we will prove is an improved global a priori estimate of the form*

$$|\mathbf{f}| \in L^q(\Omega_T) \quad \Longrightarrow \quad |\nabla u| \in L^q(\Omega_T) \qquad \text{for all} \qquad q \in [p-\beta_0, p],$$

where β_0 is a sufficiently small universal exponent. The improvement is two fold, firstly this estimate is obtained below the natural exponent and secondly, the estimate assumes no regularity of the coefficients and hence is non-perturbative. *As a consequence, this result is new even for linear equations.*

The second result that we are interested in obtaining is global estimates in weighted Lebesgue spaces with the weight in Muckenhoupt class. For general nonlinear structures with linear growth, i.e., $\mathcal{A}(x,t,\nabla u) \approx \nabla u$ with the coefficients satisfying a small bounded mean oscillation restriction, the following global weighted estimates was obtained in [15]:

$$|\mathbf{f}| \in L^q_\omega(\Omega_T) \quad \Longrightarrow \quad |\nabla u| \in L^q_\omega(\Omega_T) \qquad \text{for all} \qquad q > 2 \qquad \text{and} \qquad \omega \in A_{\frac{q}{2}}(\mathbb{R}^{n+1}).$$

Note that in particular, they cannot consider $q = 2$ in [15].

Subsequently, in [16], they were able to prove analogous results for nonlinearities of the form $\mathcal{A}(x,t,\nabla u) \approx |\nabla u|^{p-2}\nabla u$ with $\frac{2n}{n+2} < p < \infty$ and more general Weighted Orlicz spaces, in particular, they prove

$$|\mathbf{f}| \in L^q_\omega(\Omega_T) \quad \Longrightarrow \quad |\nabla u| \in L^q_\omega(\Omega_T) \qquad \text{for all} \qquad q > p \qquad \text{and} \qquad \omega \in A_{\frac{q}{p}}(\mathbb{R}^{n+1}).$$

Note that in particular, they cannot consider $q = p$ in [16].

The main obstacle in proving weighted estimates at $q = p$ is due to the failure of strong $L^1 - L^1$ bounds for the Hardy-Littlewood Maximal function. Therefore to reach the natural exponent, a different approach is needed. In this paper, we achieve this result by showing the weighted estimate holds with $q = p$, i.e., (1.2) holds. To overcome this difficulty, we construct a suitable test function based on a modification of the techniques developed in [3, 26] and obtain estimates below the natural exponent, i.e., under suitable restrictions on the $\mathcal{A}(x,t,\nabla u)$ and Ω (similar to those in [16]), we prove

$$|\mathbf{f}| \in L^q_\omega(\Omega_T) \quad \Longrightarrow \quad |\nabla u| \in L^q_\omega(\Omega_T) \qquad \text{for all} \qquad q \geq p \qquad \text{and} \qquad \omega \in A_{\frac{q}{p}}(\mathbb{R}^{n+1}). \tag{1.2}$$

There are a few remarks to be made; firstly the estimate (1.2) represents an end point weighted estimate for quasilinear parabolic equations; secondly, the optimal weight class in the elliptic case is conjectured to be $A_{\frac{q}{p-1}}$ (see [5, Theorem 1.9] for more on this and the elliptic Iwaniec conjecture) and in the parabolic case too, the optimal result is expected to be of the form

$$|\mathbf{f}| \in L^q_\omega(\Omega_T) \quad \Longrightarrow \quad |\nabla u| \in L^q_\omega(\Omega_T) \qquad \text{for all} \qquad q > p-1 \qquad \text{and} \qquad \omega \in A_{\frac{q}{p-1}}(\mathbb{R}^{n+1}),$$

but this result seems to be far out of reach of current methods.

The plan of the paper is as follows: In Section 2, we collect all the assumptions on the domain, nonlinear structure and the weight class along with some preliminary well known results, in Section 3, we will describe the main theorem that will be proved, in Section 4, we will develop a general Lipschitz truncation technique and construct a suitable test function, in Section 5, we will define useful perturbations of (1.1) and prove crucial difference estimates below the natural exponent, in Section 6, we will prove Theorem 3.1, in Section 7, we will use standard covering arguments to prove the parabolic analogue of a good-λ estimate and finally use that in Section 8 to prove Theorem 3.3.

Acknowledgement

The authors would like to thank the organisers of the conference *Recent developments in Nonlinear Partial Differential Equations and Applications - NPDE2017* held at TIFR-CAM, Bangalore where part of this work was done.

2. Preliminaries

The following restriction on the exponent p will always be enforced:

$$\frac{2n}{n+2} < p < \infty. \tag{2.1}$$

Remark 2.1. *The restriction in (2.1) is necessary when dealing with parabolic problems because, we invariably have to deal with the L^2-norm of the solution which comes from the time-derivative. On the other hand, the following Sobolev embedding $W^{1,p} \hookrightarrow L^2$ is true provided (2.1) holds. On the other hand, if we assume $u \in L^r(\Omega_T)$ for some $r \geq 1$ such that $\Lambda_r := n(p-2) + rp > 0$ (see [19, Chapter 5] for more on this), then we can obtain analogous result as to Theorem 3.3. This extension of Theorem 3.3 to the case $1 < p \leq \frac{2n}{n+2}$ requires only a technical modification provided $\Lambda_r > 0$ and will be omitted.*

2.1. Assumptions on the Nonlinear structure

We shall now collect the assumptions on the nonlinear structure in (1.1). We assume that $\mathcal{A}(x,t,\nabla u)$ is a Carathéodory function, i.e., we have $(x,t) \mapsto \mathcal{A}(x,t,\zeta)$ is measurable for every $\zeta \in \mathbb{R}^n$ and $\zeta \mapsto \mathcal{A}(x,t,\zeta)$ is continuous for almost every $(x,t) \in \Omega \times (-T,T)$. We also assume $\mathcal{A}(x,t,0) = 0$ and $\mathcal{A}(x,t,\zeta)$ is differentiable in ζ away from the origin, i.e., $d_\zeta \mathcal{A}(x,t,\zeta)$ exists for a.e. $(x,t) \in \mathbb{R}^{n+1}$.

We further assume that for a.e. $(x,t) \in \Omega \times (-T,T)$ and for any $\eta, \zeta \in \mathbb{R}^n$, there exists two given positive constants Λ_0, Λ_1 such that the following bounds are satisfied by the nonlinear structures :

$$\langle \mathcal{A}(x,t,\zeta - \mathcal{A}(x,t,\eta)\ , \zeta - \eta \rangle \geq \Lambda_0 \left(|\zeta|^2 + |\eta|^2\right)^{\frac{p-2}{2}} |\zeta - \eta|^2, \tag{2.2}$$

$$|\mathcal{A}(x,t,\zeta) - \mathcal{A}(x,t,\eta)| \leq \Lambda_1 |\zeta - \eta| \left(|\zeta|^2 + |\eta|^2\right)^{\frac{p-2}{2}}. \tag{2.3}$$

Note that from the assumption $\mathcal{A}(x,t,0)=0$, we get for a.e. $(x,t)\in\mathbb{R}^{n+1}$, there holds

$$|\mathcal{A}(x,t,\zeta)| \leq \Lambda_1|\zeta|^{p-1}.$$

2.2. Structure of Ω

The domain that we consider may be non-smooth but should satisfy some regularity condition. This condition would essentially say that at each boundary point and every scale, we require the boundary of the domain to be between two hyperplanes separated by a distance proportional to the scale.

Definition 2.2. *Given any $\gamma\in(0,1]$ and $S_0>0$, we say that Ω is (γ,S_0)-Reifenberg flat domain if for every $x_0\in\partial\Omega$ and every $r\in(0,S_0]$, there exists a system of coordinates $\{y_1,y_2,\ldots,y_n\}$ (possibly depending on x_0 and r) such that in this coordinate system, $x_0=0$ and*

$$B_r(0)\cap\{y_n>\gamma r\}\subset B_r(0)\subset B_r(0)\cap\{y_n>-\gamma r\}.$$

The class of Reifenberg flat domains is standard in obtaining Calderón-Zygmund type estimates, in the elliptic case, see [5, 11, 14, 17] and the references therein whereas for the parabolic case, see [10, 12, 13, 30] and the references therein.

From the definition of (γ,S_0)-Reifenberg flat domains, it is easy to see that the following property holds:

Lemma 2.3. *Let $\gamma>0$ and $S_0>0$ be given and suppose that Ω is a (γ,S_0)-Reifenberg flat domain, then there exists an $m_e=m_e(\gamma,S_0,n)\in(0,1)$ such that for every $x\in\Omega$ and every $r>0$, there holds*

$$|\Omega^c\cap B_r(x)|\geq m_e|B_r(x)|. \tag{2.4}$$

2.3. Smallness Assumption

In order to prove the main results, we need to assume a smallness condition satisfied by $(\mathcal{A},\Omega)$.

Definition 2.4. *Let $\gamma\in(0,1)$ and $S_0>0$ be given, we then say $(\mathcal{A},\Omega)$ is (γ,S_0)-vanishing if the following assumptions hold:*

(i) Assumption on $\mathcal{A}$: *For any parabolic cylinder $Q_{\rho,s}(\mathfrak{z})$ centered at $\mathfrak{z}:=(\mathfrak{x},\mathfrak{t})\in\mathbb{R}^{n+1}$, let us define the following:*

$$\Theta(\mathcal{A},Q_{\rho,s}(\mathfrak{z}))(x,t):=\sup_{\zeta\in\mathbb{R}^n\setminus\{0\}}\frac{\left|\mathcal{A}(x,t,\zeta)-\overline{\mathcal{A}}_{B_\rho(\mathfrak{x})}(t,\zeta)\right|}{|\zeta|^{p-1}},$$

where we have used the notation

$$\overline{\mathcal{A}}_{B_\rho(\mathfrak{x})}(t,\zeta):=\fint_{B_\rho(\mathfrak{x})}\mathcal{A}(x,t,\zeta)\ dx. \tag{2.5}$$

Then $\mathcal{A}$ is said to be (γ,S_0) vanishing if for some $\tau\in[1,\infty)$, there holds

$$[\mathcal{A}]_{\tau,S_0}:=\sup_{\substack{0<\rho\leq S_0\\ 0<s\leq S_0^2}}\sup_{\mathfrak{z}\in\mathbb{R}^{n+1}}\fint_{Q_{\rho,s}(\mathfrak{z})}|\Theta(\mathcal{A},B_\rho(\mathfrak{x}))(z)|^{\tau}\ dz\leq\gamma^{\tau}. \tag{2.6}$$

Here we have used the notation $z:=(x,t)\subset\mathbb{R}^{n+1}$.

(ii) Assumption on $\partial\Omega$: *We ask that Ω is a (γ,S_0)-Reifenberg flat in the sense of Definition 2.2.*

Remark 2.5. *From (2.2), we see that $|\Theta(\mathcal{A},Q_{\rho,s}(\mathfrak{z}))(x,t)|\leq 2\Lambda_1$, thus combining this with the assumption (2.6), we see from standard interpolation inequality that for any $1\leq\mathfrak{t}<\infty$, there holds*

$$\fint_{Q_{\rho,s}(\mathfrak{z})}|\Theta(\mathcal{A},Q_{\rho,s}(\mathfrak{z}))(z)|^{\mathfrak{t}}\ dz\leq C(\gamma,\Lambda_1),$$

with $C(\gamma,\Lambda_1)\to 0$ whenever $\gamma\to 0$.

2.4. Muckenhoupt weights

In this subsection, let us collect all the properties of the weights that will be considered in the paper. See [24, Chapter 9] for the details concerning this subsection.

Definition 2.6 (Strong Muckenhoupt Weight)**.** *A non negative, locally integrable function ω is a* strong weight *in $A_q(\mathbb{R}^{n+1})$ for some $1 < q < \infty$ if*

$$\sup_{\mathfrak{z}\in\mathbb{R}^{n+1}} \sup_{\substack{0<\rho<\infty,\\ 0<s<\infty}} \left(\fint\!\!\fint_{Q_{\rho,s}(\mathfrak{z})} \omega(z)\ dz \right) \left(\fint\!\!\fint_{Q_{\rho,s}(\mathfrak{z})} \omega^{\frac{-1}{q-1}}(z)\ dz \right)^{q-1} =: [\omega]_q < \infty.$$

In the case $q = 1$, we define the strong $A_1(\mathbb{R}^{n+1})$ weight *to be the class of non negative, locally integrable function $\omega \in A_1(\mathbb{R}^{n+1})$ satisfying*

$$\sup_{z\in\mathbb{R}^{n+1}} \sup_{\substack{0<\rho<\infty,\\ 0<s<\infty}} \left(\fint\!\!\fint_{Q_{\rho,s}(\mathfrak{z})} \omega(z)\ dz \right) \left\| \omega^{-1} \right\|_{L^\infty(Q_{\rho,s}(\mathfrak{z}))} =: [\omega]_1 < \infty.$$

The quantity $[w]_q$ for $1 \le q < \infty$ will be called as the A_q constant of the weight ω.

We will need the following important characterization of Muckenhoupt weights:

Lemma 2.7. *A parabolic weight $w \in A_q$ for $1 < q < \infty$ if and only if*

$$\left(\frac{1}{|Q|} \iint_Q f(x,t)\ dx\ dt \right)^q \le \frac{c}{w(Q)} \iint_Q |f(x,t)|^q w(x,t)\ dx\ dt,$$

holds for all non-negative, locally integrable functions f and all cylinders $Q = Q_{\rho,s}(x,t)$.

As a direct consequence of Lemma 2.7, the following Lemma holds:

Lemma 2.8. *Let $\omega \in A_q(\mathbb{R}^{n+1})$ for some $1 < q < \infty$, then there exists positive constants $c = c(n, q, [\omega]_q)$ and $\tau = \tau(n, q, [\omega]_q) \in (0, 1)$ such that*

$$\frac{1}{c} \left(\frac{|E|}{|Q|} \right)^q \le \frac{\omega(E)}{\omega(Q)} \le c \left(\frac{|E|}{|Q|} \right)^\tau,$$

for all $E \subset Q$ and all parabolic cylinders $Q_{\rho,s}(\mathfrak{z})$.

Another important result regarding the strong Muckenhoupt weights that will be needed is the following self-improvement property:

Lemma 2.9. *Let $1 < q < \infty$ and suppose $\omega \in A_q$ be a given weight, then there exists an $\varepsilon_0 = \varepsilon_0(n, q, [\omega]_q) > 0$ such that $\omega \in A_{q-\varepsilon_0}$ with the estimate $[\omega]_{q-\varepsilon_0} \le C[\omega]_q$ where $C = C(q, n, [\omega]_q)$.*

We will now define the A_∞ class as follows:

Definition 2.10. *A weight $\omega \in A_\infty$ if and only if there are constants $\tau_0, \tau_1 > 0$ such that for every parabolic cylinder $Q = Q_{\rho,s} \subset \mathbb{R}^{n+1}$ and every measurable $E \subset Q$, there holds*

$$\omega(E) \le \tau_0 \left(\frac{|E|}{|Q|} \right)^{\tau_1} \omega(Q).$$

Moreover, if ω is an A_q weight with $[\omega]_q \le \overline{\omega}$, then the constants τ_0 and τ_1 can be chosen such that $\max\{\tau_0, 1/\tau_1\} \le c(\overline{\omega}, n)$.

From the general theory of Muckenhoupt weights, we see that $A_\infty = \bigcup_{1\le q<\infty} A_q$.

Remark 2.11. *The weight class considered in Definition 2.6 is called Strong Muckenhoupt class because the cylinders are decoupled in space and time, i.e., ρ and s are not related when considering cylinders $Q_{\rho,s}$. When considering linear equations (i.e., $p = 2$), the weight class is defined with respect to cylinders of the form Q_{ρ,ρ^2}. This is possible because in the case $p = 2$, there is an invariance property under normalization, which does not exist if $p \neq 2$. It is an open question if one can obtain the results of this paper for Muckenhoupt weights defined with respect to cylinders belonging to a more restricted class (see the very nice thesis [32] for some results concerning the weights arising in doubly nonlinear quasilinear equations).*

2.5. Function Spaces

Let $1 \le \vartheta < \infty$, then $W_0^{1,\vartheta}(\Omega)$ denotes the standard Sobolev space which is the completion of $C_c^\infty(\Omega)$ under the $\|\cdot\|_{W^{1,\vartheta}}$ norm.

The parabolic space $L^{\vartheta}(-T,T;W^{1,\vartheta}(\Omega))$ for any $\vartheta \in [1,\infty)$ is the collection of measurable functions $\phi(x,t)$ such that for almost every $t \in (-T,T)$, the function $x \mapsto \phi(x,t)$ belongs to $W^{1,\vartheta}(\Omega)$ with the following norm being finite:

$$\|\phi\|_{L^{\vartheta}(-T,T;W^{1,\vartheta}(\Omega))} := \left(\int_{-T}^{T} \|\phi(\cdot,t)\|_{W^{1,\vartheta}(\Omega)}^{\vartheta}\ dt\right)^{\frac{1}{\vartheta}} < \infty.$$

Analogously, the parabolic space $L^{\vartheta}(-T,T;W_0^{1,\vartheta}(\Omega))$ is the collection of measurable functions $\phi(x,t)$ such that for almost every $t \in (-T,T)$, the function $x \mapsto \phi(x,t)$ belongs to $W_0^{1,\vartheta}(\Omega)$.

Given a weight $\omega \in A_{\vartheta}$ for some $\vartheta \in [1,\infty)$, the weighted Lebesgue space $L^{\vartheta}(-T,T;L^{\vartheta}_{\omega}(\Omega))$ is the set of all measurable functions $\phi : \Omega_T \mapsto \mathbb{R}$ satisfying

$$\int_{-T}^{T}\left(\int_{\Omega} |\phi(x,t)|^{\vartheta}\omega(x,t)\ dx\right) dt < \infty.$$

Let us recall the following important characterization of Lebesgue spaces:

Lemma 2.12. *Let Ω be a bounded domain in $\mathbb{R}^n$ and let $w \in L^1(\Omega_T)$ be any non-negative function, then for all $\beta > \alpha > 1$ and any non-negative measurable function $g(x,t) : \Omega_T \mapsto \mathbb{R}$, there holds*

$$\iint_{\Omega_T} g^{\beta} w(z)\ dz = \beta \int_0^{\infty} \lambda^{\beta-1} w(\{z \in \Omega_T : g(z) > \lambda\})\ d\lambda = (\beta-\alpha)\int_0^{\infty} \lambda^{\beta-\alpha-1}\left(\iint_{\{z\in\Omega_T:g(z)>\lambda\}} g^{\alpha} w(z)\ dz\right) d\lambda.$$

Before we conclude this subsection, let us now recall the well known Poincaré's inequality (see [2, Corollary 8.2.7] for the proof):

Theorem 2.13. *Let $1 \le \vartheta < \infty$ and let $f \in W^{1,\vartheta}(\tilde{\Omega})$ for some bounded domain $\tilde{\Omega}$ and suppose that the following measure density condition holds:*

$$\left|\{x \in \tilde{\Omega} : f(x) = 0\}\right| \ge m_e > 0,$$

then there holds

$$\int_{\tilde{\Omega}} \left|\frac{f}{\operatorname{diam}(\tilde{\Omega})}\right|^{\vartheta}\ dx \le C_{(n,\vartheta,m_e)} \int_{\tilde{\Omega}} |\nabla f|^{\vartheta}\ dx.$$

2.6. Parabolic metric

Let us define the Parabolic metric on $\mathbb{R}^{n+1}$ that will be used throughout the paper:

Definition 2.14. *We define the parabolic metric d_p on $\mathbb{R}^{n+1}$ as follows: Let $z_1 = (x_1,t_1)$ and $z_2 = (x_2,t_2)$ be any two points on $\mathbb{R}^{n+1}$, then*

$$d_p(z_1,z_2) := \max\left\{|x_1-x_2|, \sqrt{|t_1-t_2|}\right\}.$$

2.7. Maximal Function

For any $f \in L^1(\mathbb{R}^{n+1})$, let us now define the strong maximal function in $\mathbb{R}^{n+1}$ as follows:

$$\mathcal{M}(|f|)(x,t) := \sup_{\tilde{Q}\ni(x,t)} \fint\fint_{\tilde{Q}} |f(y,s)|\ dy\ ds, \tag{2.7}$$

where the supremum is taken over all parabolic cylinders $\tilde{Q}_{a,b}$ with $a,b \in \mathbb{R}^+$ such that $(x,t) \in \tilde{Q}_{a,b}$. An application of the Hardy-Littlewood maximal theorem in $x-$ and $t-$ directions shows that the Hardy-Littlewood maximal theorem still holds for this type of maximal function (see [28, Lemma 7.9] for details):

Lemma 2.15. *If $f \in L^1(\mathbb{R}^{n+1})$, then for any $\alpha > 0$, there holds*

$$|\{z \in \mathbb{R}^{n+1} : \mathcal{M}(|f|)(z) > \alpha\}| \le \frac{5^{n+2}}{\alpha}\|f\|_{L^1(\mathbb{R}^{n+1})},$$

and if $f \in L^{\vartheta}(\mathbb{R}^{n+1})$ for some $1 < \vartheta \le \infty$, then there holds

$$\|\mathcal{M}(|f|)\|_{L^{\vartheta}(\mathbb{R}^{n+1})} \le C_{(n,\vartheta)}\|f\|_{L^{\vartheta}(\mathbb{R}^{n+1})}.$$

2.8. Notation

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

(i) We shall use ∇ to denote derivatives with respect the space variable x.

(ii) We shall sometimes alternate between using $\frac{df}{dt}$, $\partial_t f$ and f' to denote the time derivative of a function f.

(iii) We shall use D to denote the derivative with respect to both the space variable x and time variable t in $\mathbb{R}^{n+1}$.

(iv) Let $z_0 = (x_0, t_0) \in \mathbb{R}^{n+1}$ be a point and $\rho, s > 0$ be two given parameters and let $\lambda \in (0, \infty)$. We shall use the following notation to denote the following regions:

$$\begin{cases} Q_\rho^\lambda(z_0) & := & Q_{\rho,\lambda^{2-p}\rho^2}(z_0) & \text{for } p \geq 2, \\ Q_\rho^\lambda(z_0) & := & Q_{\lambda^{\frac{p-2}{2}}\rho,\rho^2}(z_0) & \text{for } p \leq 2, \end{cases}$$

$$\begin{aligned} &I_s(t_0) := (t_0 - s, t_0 + s) \subset \mathbb{R}, && Q_{\rho,s}(z_0) := B_\rho(x_0) \times I_s(t_0) \subset \mathbb{R}^{n+1}, \\ &\alpha Q_{\rho,s}(z_0) := B_{\alpha\rho}(x_0) \times I_{\alpha^2 s}(t_0) \subset \mathbb{R}^{n+1}, && Q_\rho(z_0) := B_\rho(x_0) \times I_{\rho^2}(t_0) \subset \mathbb{R}^{n+1}, \\ &Q_\rho^{\lambda,+}(z_0) := Q_\rho^\lambda(z_0) \cap \{(x,t) : x_n > 0\}, && Q_\rho^+(z_0) := Q_\rho(z_0) \cap \{(x,t) : x_n > 0\}, \\ &K_\rho^\lambda(z_0) := Q_\rho^\lambda(z) \cap \Omega_T, && K_\rho(z_0) := Q_\rho(z) \cap \Omega_T. \end{aligned} \tag{2.8}$$

(v) We shall use $\int$ to denote the integral with respect to either space variable or time variable and use $\iint$ to denote the integral with respect to both space and time variables simultaneously.

Analogously, we will use $⨍$ and $⨍⨍$ to denote the average integrals as defined below: for any set $A \times B \subset \mathbb{R}^n \times \mathbb{R}$, we define

$$(f)_A := ⨍_A f(x)\ dx = \frac{1}{|A|}\int_A f(x)\ dx,$$

$$(f)_{A\times B} := ⨍⨍_{A\times B} f(x,t)\ dx\ dt = \frac{1}{|A \times B|}\iint_{A\times B} f(x,t)\ dx\ dt.$$

(vi) Given any positive function μ, we shall denote $(f)_\mu := \int f \frac{\mu}{\|\mu\|_{L^1}} dm$ where the domain of integration is the domain of definition of μ and dm denotes the associated measure.

2.9. Weak solutions

For this subsection, let us consider the following general problem:

$$\begin{cases} \phi_t - \operatorname{div} \mathcal{A}(z, \nabla\phi) & = & -\operatorname{div} |\vec{f}|^{p-2}\vec{f} & \text{in } \tilde{\Omega}_T, \\ \phi & = & f & \text{on } \partial\tilde{\Omega} \times (-T,T), \\ \phi(\cdot, -T) & = & f_0 & \text{on } \tilde{\Omega}. \end{cases} \tag{2.9}$$

Now let us define the Steklov average as follows: let $h \in (0, 2T)$ be any positive number, then we define

$$\phi_h(\cdot, t) := \begin{cases} ⨍_t^{t+h} \phi(\cdot, \tau)\ d\tau & t \in (-T, T-h), \\ 0 & \text{else}. \end{cases} \tag{2.10}$$

Definition 2.16 (Weak solution)**.** *We then say $\phi \in L^2(-T,T;L^2(\Omega)) \cap L^p(-T,T;W_0^{1,p}(\Omega))$ is a weak solution of (1.1) if the following holds for any $\psi \in W_0^{1,p}(\Omega) \cap L^\infty(\Omega)$:*

$$\int_{\Omega\times\{t\}} \frac{d[\phi]_h}{dt}\psi + \langle [\mathcal{A}(x,t,\nabla\phi)]_h\ , \nabla\psi\rangle\ dx = \int_{\Omega\times\{t\}} \langle |\vec{f}|^{p-2}\vec{f}\ , \nabla\psi\rangle\ dx \qquad \text{for a.e.} \qquad -T < t < T-h. \tag{2.11}$$

Moreover, the initial datum is taken in the sense of $L^2(\Omega)$, i.e.,

$$\int_\Omega |\phi_h(x,-T) - f_0(x)|^2\ dx \xrightarrow{h \searrow 0} 0.$$

We have the following well known existence result (for example, see [33, Chapter III, Section 6] for the details):

Proposition 2.17. *Let Ω be any bounded domain satisfying a uniform measure density condition, i.e., there exists a constant $m_e > 0$ such that $|B_r(y) \cap \Omega| \geq m_e |B_r(y)|$ holds for every $r > 0$ and $y \in \partial\Omega$ and suppose that $\vec{f} \in L^p(\Omega_T)$, $\nabla f \in L^p(\Omega_T)$ with $\frac{df}{dt} \in \left(W^{1,p}(\Omega_T)\right)'$ and $f_0 \in L^2(\Omega)$ are given. Then there exists a unique weak solution $\phi \in C^0\left(-T,T;L^2(\Omega)\right) \cap L^p\left(-T,T;W^{1,p}(\Omega)\right)$ solving (2.9).*

Moreover if $f = 0$, then we have the following energy estimate

$$\sup_{-T\leq t\leq T} \|\phi(\cdot,t)\|^2_{L^2(\Omega)} + \iint_{\Omega_T} |\nabla\phi|^p \, dz \leq C_{(n,p,\Lambda_0,\Lambda_1)} \left(\iint_{\Omega_T} |\vec{f}|^p \, dz + \|f_0\|^2_{L^2(\Omega)} \right).$$

2.10. Gradient higher integrability estimates

In this subsection, let us collect a few important higher integrability results that will be used throughout the paper. In order to state the general theorems, let $\phi \in L^2(-T,T;L^2(\Omega)) \cap L^p\left(-T,T;W_0^{1,p}(\Omega)\right)$ be a weak solution of

$$\left\{ \begin{array}{rcll} \phi_t - \operatorname{div} \mathcal{A}(x,t,\nabla\phi) & = & -\operatorname{div}(|\vec{f}|^{p-2}\vec{f}) & \text{in } \Omega \times (-T,T), \\ \phi & = & 0 & \text{on } \partial\Omega \times (-T,T), \end{array} \right. \tag{2.12}$$

where the nonlinearity is assumed to satisfy (2.2) and (2.3). Here the domain is assumed to satisfy a uniform measure density condition with constant m_e as in Lemma 2.3

The first one is the higher integrability *above the natural exponent*. In the interior case, this was proved in [25] whereas in the boundary case, using the measure density condition satisfied by Ω, the result was proved in [29, 31].

Lemma 2.18 ([29, 31])**.** *Let $\tilde{\sigma} > 0$ be given, then there exists a $\tilde{\beta}_1 = \tilde{\beta}_1(n,p,\Lambda_0,\Lambda_1,m_e) \in (0,\tilde{\sigma}]$ such that if $\vec{f} \in L^{p(1+\tilde{\sigma})}(\Omega_T)$ and $\phi \in L^p\left(-T,T;W_0^{1,p}(\Omega)\right)$ is a weak solution to (2.12), then $|\nabla\phi| \in L^{p(1+\beta)}(\Omega_T)$ for all $\beta \in (0,\tilde{\beta}_1]$. Moreover, for any $\mathfrak{z} \in \overline{\Omega} \times (-T,T)$, there holds*

$$\fint\fint_{K_\rho(\mathfrak{z})} |\nabla\phi|^{p+\beta} \, dz \leq_{(n,p,\Lambda_0,\Lambda_1,m_e)} \left(\fint\fint_{K_{2\rho}(\mathfrak{z})} \left(|\nabla\phi| + |\vec{f}|\right)^p \, dz \right)^{1+\beta\tilde{\vartheta}_1} + \fint\fint_{K_{2\rho}(\mathfrak{z})} \left(|\vec{f}|+1\right)^{p(1+\beta)} \, dz.$$

Here the constant

$$\tilde{\vartheta}_1 := \left\{ \begin{array}{lll} \frac{p}{2} & \text{if} & p \geq 2, \\ \frac{2p}{p(n+2)-2n} & \text{if} & \frac{2n}{n+2} < p < 2. \end{array} \right.$$

We will also need an improved higher integrability result below the natural exponent. The following theorem was proved for a weaker class of solutions called *very weak solutions*, but also holds true for *weak solutions* as considered in this paper. The interior higher integrability result was proved in the seminal paper [26] whereas the boundary analogue was proved in [3].

Lemma 2.19 ([26, 3])**.** *Let $\vec{f} \in L^p(\Omega_T)$ and $\phi \in L^p\left(-T,T;W_0^{1,p}(\Omega)\right)$ be the unique weak solution to (2.12). There exists $\tilde{\beta}_2 = \tilde{\beta}_2(n,\Lambda_0,\Lambda_1,p,m_e) \in (0,1/4)$ such that for any $\mathfrak{z} \in \overline{\Omega} \times (-T,T)$, there holds*

$$\fint\fint_{K_\rho(\mathfrak{z})} |\nabla\phi|^p \, dz \leq_{(n,\Lambda_0,\Lambda_1,p,m_e)} \left(\fint\fint_{K_{2\rho}(\mathfrak{z})} \left(|\nabla\phi| + |\vec{f}|\right)^{p-\beta} \, dz \right)^{1+\beta\tilde{\vartheta}_2} + \fint\fint_{K_{2\rho}(\mathfrak{z})} \left(|\vec{f}|+1\right)^p \, dz.$$

Here the constant

$$\tilde{\vartheta}_2 := \left\{ \begin{array}{lll} 2-\beta & \text{if} & p \geq 2, \\ p - \beta - \frac{(2-p)n}{2} & \text{if} & \frac{2n}{n+2} < p < 2. \end{array} \right.$$

3. Main Results

In this section, let us describe the main theorem that will be proved. The first is unweighted a priori estimates below the natural exponent.

Theorem 3.1. *Let Ω be a bounded domain satisfying (2.4), then there exists an $\beta_0 = \beta_0(p,n,\Lambda_0,\Lambda_1) \in (0,1)$ such for any $\beta \in (0,\beta_0)$, the following holds: For any $\mathbf{f} \in L^p(\Omega_T)$, let $u \in C^0(-T,T;L^2(\Omega)) \cap L^p(-T,T;W_0^{1,p}(\Omega))$ be the unique* weak solution *of (1.1), then there holds*

$$\iint_{\Omega_T} |\nabla u|^{p-\beta}\ dz \lesssim_{(n,p,\beta,\Lambda_0,\Lambda_1)} \iint_{\Omega_T} |\mathbf{f}|^{p-\beta}\ dz.$$

Remark 3.2. *As a corollary, we can extend the results of [21, Theorem 1.6] to obtain Lorentz space estimates below the natural exponent. The techniques that we develop to prove Theorem 3.1 can be used to obtain the parabolic analogue of [4, Theorem 1.2] for* weak solutions*. In a forthcoming paper, we obtain these results for more general solutions called* very weak solutions

The second theorem we will prove is the end point weighted estimate. As mentioned in the introduction, the main contribution is the case $q = p$.

Theorem 3.3. *Let $q \in [p,\infty)$ and $w \in A_{\frac{q}{p}}$ be a Muckenhoupt weight, then there exists a positive constants $\vartheta_0 = \vartheta_0(\Lambda_0,\Lambda_1,n,p,\Omega)$ and $\gamma = \gamma(n,\Lambda_0,\Lambda_1,p,q)$ such that the following holds: Suppose $(\mathcal{A},\Omega)$ is (γ,S_0) vanishing for some fixed $S_0 > 0$, then the problem (1.1) has a unique weak solution u satisfying the estimate*

$$\iint_{\Omega_T} |\nabla u|^q w(z)\ dz \lesssim_{(n,\Lambda_0,\Lambda_1,p,q,[w]_{\frac{q}{p}},\Omega)} \left(\iint_{\Omega_T} |\mathbf{f}|^q w(z)\ dz + 1 \right)^{\vartheta_0}.$$

4. Construction of test function via Lipschitz truncation

In this section, we will consider the following two problems: Let $\vec{f} \in L^p(\Omega_T)$ be given and suppose that $\varphi \in L^2(-T,T;L^2(\Omega)) \cap L^p(-T,T;W_0^{1,p}(\Omega))$ is a weak solution of

$$\left\{ \ \varphi_t - \operatorname{div} \mathcal{A}(x,t,\nabla\varphi) = \operatorname{div} |\vec{f}|^{p-2}\vec{f} \quad \text{in } \Omega \times (-T,T). \right. \tag{4.1}$$

We will extend $\varphi = 0$ on $\Omega^c \times (-T,T)$, then for any fixed cylinder $Q_{\rho,s}(\mathfrak{z}) \subset \mathbb{R}^n \times (-T,T)$, we see from Proposition 2.17 that for any $\vec{g} \in L^p(Q_{\rho,s}(\mathfrak{z}))$, there exists a unique weak solution $\phi \in L^p(I_s(\mathfrak{t});W^{1,p}(B_\rho(\mathfrak{x})))$ solving

$$\begin{cases} \phi_t - \operatorname{div} \mathcal{A}(x,t,\nabla\phi) = \operatorname{div} |\vec{g}|^{p-2}\vec{g} & \text{in } Q_{\rho,s}(\mathfrak{z}), \\ \phi = \varphi & \text{on } \partial_p Q_{\rho,s}(\mathfrak{z}). \end{cases} \tag{4.2}$$

From (4.1), we see that the condition $\phi = \varphi$ on $\partial_p Q_{\rho,s}(\mathfrak{z})$ makes sense.

In Section 5, we obtain difference estimates below the natural exponent between equations of the form (4.1) and (4.2). In order to do this, we need to construct a suitable test function which will be done in this section.

4.1. A few well known lemmas

We shall recall the following well known lemmas that will be used throughout this section. The first one is a standard lemma regarding integral averages (for a proof in this setting, see for example [8, Chapter 8.2] for the details).

Lemma 4.1. *Let $\lambda > 0$ be any fixed number and suppose $[\psi]_h(x,t) := \fint_{t-\lambda h^2}^{t+\lambda h^2} \psi(x,\tau)\ d\tau$ for some $\psi \in L^1_{loc}$. Then we have the following properties:*

(i) $[\psi]_h \to \psi$ *a.e* $(x,t) \in \mathbb{R}^{n+1}$ *as* $h \searrow 0$.

(ii) $[\psi]_h(x,\cdot)$ *is continuous and bounded in time for a.e.* $x \in \mathbb{R}^n$.

(iii) For any cylinder $Q_{r,\lambda r^2} \subset \mathbb{R}^{n+1}$ with $r > 0$, there holds

$$\iint_{Q_{r,\gamma r^2}} [\psi]_h(x,t)\ dx\ dt \lesssim_n \iint_{Q_{r,\lambda(r+h)^2}} \psi(x,t)\ dx\ dt.$$

(iv) The function $[\psi]_h(x,t)$ is differentiable with respect to $t \in \mathbb{R}$, moreover $[\psi]_h(x,\cdot) \in C^1(\mathbb{R})$ for a.e. $x \in \mathbb{R}^n$.

Let us now prove a time localized version of the Parabolic Poincaré inequality.

Lemma 4.2. *Let $\psi \in L^{\vartheta}(-T,T;W^{1,\vartheta}(\Omega))$ with $\vartheta \in [1,\infty)$ and suppose that $B_r \Subset \Omega$ be compactly contained ball of radius $r>0$. Let $I \subset (-T,T)$ be a time interval and $\rho(x,t) \in L^1(B_r \times I)$ be any positive function such that*

$$\|\rho\|_{L^\infty(B_r\times I)} \lesssim_n \frac{|B_r \times I|}{\|\rho\|_{L^1(B_r\times I)}},$$

and $\mu(x) \in C_c^\infty(B_r)$ be such that $\int_{B_r} \mu(x)\,dx = 1$ with $|\mu| \lesssim \frac{1}{r^n}$ and $|\nabla\mu| \lesssim \frac{1}{r^{n+1}}$, then there holds:

$$\iint_{B_r\times I} \left|\frac{\psi(z)\chi_J - \left(\psi\chi_J\right)_\rho}{r}\right|^{\vartheta} dz \quad \lesssim_{(n,\vartheta)} \iint_{B_r\times I} |\nabla\psi|^{\vartheta}\chi_J\,dz + \sup_{t_1,t_2\in I}\left|\frac{\left(\psi\chi_J\right)_\mu(t_2) - \left(\psi\chi_J\right)_\mu(t_1)}{r}\right|^{\vartheta},$$

where $(\psi)_\rho := \int_{B_r\times I} \psi(z)\frac{\rho(z)}{\|\rho\|_{L^1(B_r\times I)}}\chi_J\,dz$, $\left(\psi\chi_J\right)_\mu(t_i) := \int_{B_r} \psi(x,t_i)\mu(x)\chi_J\,dx$ and $J \Subset (-\infty,\infty)$ be some fixed time-interval.

Proof. Let us first consider the case of $\rho(x,t) = \mu(x)\chi_I(t)$. In this case, we get

$$\begin{aligned}
\iint_{B_r\times I}\left|\frac{\psi(z)\chi_J - \left(\psi\chi_J\right)_{\mu\times\chi_I}}{r}\right|^{\vartheta} dz &\lesssim \iint_{B_r\times I}\left|\frac{\psi(z)\chi_J - \left(\psi\chi_J\right)_\mu(t)}{r}\right|^{\vartheta} + \left|\frac{\left(\psi\chi_J\right)_\mu(t) - \left(\psi\chi_J\right)_{\mu\times I}}{r}\right|^{\vartheta} dz\\
&\overset{(a)}{\lesssim} \iint_{B_r\times I}|\nabla\psi|^{\vartheta}\chi_J\,dz + \sup_{t_1,t_2\in I}\left|\frac{\left(\psi\chi_J\right)_\mu(t_2) - \left(\psi\chi_J\right)_\mu(t_1)}{r}\right|^{\vartheta}\\
&= \iint_{B_r\times I}|\nabla\psi|^{\vartheta}\chi_J\,dz + \sup_{t_1,t_2\in I\cap J}\left|\frac{(\psi)_\mu(t_2) - (\psi)_\mu(t_1)}{r}\right|^{\vartheta}.
\end{aligned}$$

To obtain (a) above, we made us of the standard Poincaré's inequality in the spatial direction which only needs to be applied over a.e. $t \in I\cap J$. Note that the derivative is only in the spatial direction and hence the term χ_J does not cause any problem when applying Poincaré's inequality.

For the general case, we observe that

$$\iint_{B_r\times I}\left|\frac{\psi\chi_J - \left(\psi\chi_J\right)_\rho}{r}\right|^{\vartheta} dz \lesssim \iint_{B_r\times I}\left|\frac{\psi\chi_J - \left(\psi\chi_J\right)_{\mu\times\chi_I}}{r}\right|^{\vartheta} dz + \iint_{B_r\times I}\left|\frac{\left(\psi\chi_J\right)_\rho - \left(\psi\chi_J\right)_{\mu\times\chi_I}}{r}\right|^{\vartheta} dz. \tag{4.3}$$

The first term of (4.3) can be controlled as in (4.1) and to control the second term, we observe that

$$\left|\left(\psi\chi_J\right)_\rho - \left(\psi\chi_J\right)_{\mu\times\chi_I}\right| \le \frac{\|\rho\|_{L^\infty(B_r\times I)}}{\|\rho\|_{L^1(B_r\times I)}}\iint_{B_r\times I}\left|\psi\chi_J - \left(\psi\chi_J\right)_{\mu\times\chi_I}\right| dz \lesssim \iint_{B_r\times I}\left|\psi\chi_J - \left(\psi\chi_J\right)_{\mu\times\chi_I}\right| dz.$$

This completes the proof of the Lemma. ☐

Remark 4.3. *In Lemma 4.2, we can take any bounded region $\tilde{\Omega}$ instead of B_r such that $\tilde{\Omega}$ admits the ϑ-Poincaré inequality. For example, if $\tilde{\Omega}$ satisfies the measure density condition as defined in Definition 2.3 for some $m_e > 0$, then Lemma 4.2 is applicable.*

We will use the following result which can be found in [23, Theorem 3.1] (see also [18]) for proving the Lipschitz regularity for the constructed test function. This very useful simplification of the original technique from [26] first appeared in [9, Chapter 3].

Lemma 4.4. *Let $\gamma > 0$ and $\mathcal{D} \subset \mathbb{R}^{n+1}$ be given. For any $z \in \mathcal{D}$ and $r > 0$, let $Q_{r,\gamma r^2}(z)$ be the parabolic cylinder centered at z with radius r. Suppose there exists a constant $C > 0$ independent of z and r such that the following bound holds:*

$$\frac{1}{|Q_{r,\gamma r^2}(z)\cap\mathcal{D}|}\iint_{Q_{r,\gamma r^2}(z)\cap\mathcal{D}}\left|\frac{f(x,t) - (f)_{Q_{r,\gamma r^2}(z)\cap\mathcal{D}}}{r}\right| dx\,dt \le C \qquad \forall\ z \in \mathcal{D} \text{ and } r > 0,$$

then f is Lipschitz with respect to the metric $d(z_1,z_2) := \max\{|x_1 - x_2|, \sqrt{\gamma^{-1}|t_1 - t_2|}\}$.

4.2. Construction of test function

Let us denote the following functions:

$$v(z) := \varphi(z) - \phi(z) \qquad \text{and} \qquad v_h(z) := [\varphi - \phi]_h(z).$$

where $[\cdot]_h$ denotes the usual Steklov average. From Lemma 4.1, we see that $v_h \xrightarrow{h\searrow 0} v$. It is easy to see from (4.2) that $v(z) = 0$ for $z \in \partial_p Q_{\rho,s}(\mathfrak{z})$.

Let us fix the following exponents for this Section:

$$1 < q < p - 2\beta < p - \beta < p, \tag{4.4}$$

for some $\beta \in (0,1)$. Note that eventually we will obtain a $\beta_0 = \beta_0(n, p, \Lambda_0, \Lambda_1, m_e)$ such that all the estimates hold for any $\beta \in (0, \beta_0)$.

Let us now define the following function:

$$g(z) := \mathcal{M}\left(\left[|\nabla v|^q + |\nabla\varphi|^q + |\nabla\phi|^q + |\vec{f}|^q + |\vec{g}|^q\right]\chi_{Q_{\rho,s}(\mathfrak{z})}\right)^{\frac{1}{q}}(z),$$

where $\mathcal{M}$ is as defined in (2.7).

For a fixed $\lambda > 0$, let us define the *good set* by

$$E_\lambda := \{(x,t) \in \mathbb{R}^{n+1} : g(x,t) \leq \lambda\}.$$

We now have the following parabolic Whitney type decomposition of E_λ^c (see [20, Lemma 3.1] or [9, Chapter 3] for details):

Lemma 4.5. *Let $\kappa := \lambda^{2-p}$, then there exists an κ-parabolic Whitney covering $\{Q_i(z_i)\}$ of E_λ^c in the following sense:*

(W1) $Q_j(z_j) = B_j(x_j) \times I_j(t_j)$ *where* $B_j(x_j) = B_{r_j}(x_j)$ *and* $I_j(t_j) = (t_j - \kappa r_j^2, t_j + \kappa r_j^2)$.

(W2) *we have* $d_\lambda(z_j, E_\lambda) = 16 r_j$.

(W3) $\bigcup_j \frac{1}{2} Q_j(z_j) = E_\lambda^c$.

(W4) *for all* $j \in \mathbb{N}$, *we have* $8Q_j \subset E_\lambda^c$ *and* $16 Q_j \cap E_\lambda \neq \emptyset$.

(W5) *if* $Q_j \cap Q_k \neq \emptyset$, *then* $\frac{1}{2} r_k \leq r_j \leq 2 r_k$.

(W6) $\frac{1}{4} Q_j \cap \frac{1}{4} Q_k = \emptyset$ *for all* $j \neq k$.

(W7) $\sum_j \chi_{4Q_j}(z) \leq c(n)$ *for all* $z \in E_\lambda^c$.

Subject to this Whitney covering, we have an associated partition of unity denoted by $\{\Psi_j\} \in C_c^\infty(\mathbb{R}^{n+1})$ such that the following holds:

(W8) $\chi_{\frac{1}{2}Q_j} \leq \Psi_j \leq \chi_{\frac{3}{4}Q_j}$.

(W9) $\|\Psi_j\|_\infty + r_j \|\nabla \Psi_j\|_\infty + r_j^2 \|\nabla^2 \Psi_j\|_\infty + \lambda r_j^2 \|\partial_t \Psi_j\|_\infty \leq C$.

For a fixed $k \in \mathbb{N}$, let us define

$$A_k := \left\{ j \in \mathbb{N} : \frac{3}{4} Q_k \cap \frac{3}{4} Q_j \neq \emptyset \right\},$$

then we have

(W10) *Let $i \in \mathbb{N}$ be given, then* $\sum_{j \in A_i} \Psi_j(z) = 1$ *for all* $z \in \frac{3}{4} Q_i$.

(W11) *Let $i \in \mathbb{N}$ be given and let $j \in A_i$, then* $\max\{|Q_j|, |Q_i|\} \leq C(n) |Q_j \cap Q_i|$.

(W12) *Let $i \in \mathbb{N}$ be given and let $j \in A_i$, then* $\max\{|Q_j|, |Q_i|\} \leq \left|\frac{3}{4} Q_j \cap \frac{3}{4} Q_i\right|$.

(W13) *For any $i \in \mathbb{N}$, we have $\# A_i \leq c(n)$.*

(W14) *Let $i \in \mathbb{N}$ be given, then for any $j \in A_i$, we have $\frac{3}{4}Q_j \subset 4Q_i$.*

Now we define the following Lipschitz extension function as follows:

$$v_{\lambda,h}(z) := v_h(z) - \sum_i \Psi_i(z)(v_h(z) - v_h^i), \tag{4.5}$$

where

$$v_h^i := \begin{cases} \dfrac{1}{\|\Psi_i\|_{L^1(\frac{3}{4}Q_i)}} \iint_{\frac{3}{4}Q_i} v_h(z)\Psi_i(z)\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\, dz & \text{if } \dfrac{3}{4}Q_i \subset B_\rho(\mathfrak{x}) \times [\mathfrak{t}-s,\infty), \\ 0 & \text{else.} \end{cases} \tag{4.6}$$

Since $\varphi - \phi = 0$ on $\partial B_\rho(\mathfrak{x}) \times [\mathfrak{t}-s, \mathfrak{t}+s]$, we can switch between $\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}$ and $\chi_{Q_{\rho,s}(\mathfrak{z})}$ without affecting the calculations.

Remark 4.6. *Note that even though $v_h(x,\mathfrak{t}-s) \neq 0$ in general, nevertheless the following initial boundary values are satisfied:*

- *The initial condition $(\varphi - \phi)(x, \mathfrak{t}-s) = 0$ is to be understood in the sense*
$$[\varphi - \phi]_h(\cdot, \mathfrak{t}-s) \xrightarrow{h \searrow 0} 0 \text{ in } L^2(B_\rho(\mathfrak{x})).$$
- *For $(x, \mathfrak{t}-s) \in E_\lambda$, we have $v_{\lambda,h}(x,\mathfrak{t}-s) = v_h(x,\mathfrak{t}-s)$.*
- *For $(x, \mathfrak{t}-s) \notin E_\lambda$, we have $v_{\lambda,h}(x,\mathfrak{t}-s) = 0$ by using (4.6).*

Remark 4.7. *From Lemma 4.1, we see that $v_{\lambda,h}(z) \xrightarrow{h \searrow 0} v_\lambda(z)$ almost everywhere.*

We now have the following useful lemma that can be proved just by using the definition of the weak formulation (see for example [3, Lemma 3.5] for details):

Lemma 4.8. *Let $\varphi, \phi, \vec{f}, \vec{g}$ be as in (4.1) and (4.2) and $h \in (0, 2s)$. Let $\alpha(x) \in C_c^\infty(B_\rho(\mathfrak{x}))$ and $\beta(t) \in C^\infty(\mathfrak{t}-s, \mathfrak{t}+s)$ with $\beta(\mathfrak{t}-s) = 0$ be a non-negative function and $[\cdot]_h$ be the Steklov average as defined in (4.1). Then the following estimate holds for any time interval $(t_1, t_2) \subset (\mathfrak{t}-s, \mathfrak{t}+s)$:*

$$\begin{aligned} |(v_h\beta)_\alpha(t_2) - (v_h\beta)_\alpha(t_1)| \quad \leq \quad & C(\Lambda_1, p)\|\nabla\alpha\|_{L^\infty(B_\rho(\mathfrak{x}))}\|\beta\|_{L^\infty(t_1,t_2)} \iint_{B_\rho(\mathfrak{x})\times(t_1,t_2)} [|\nabla\phi|^{p-1} + |\nabla\varphi|^{p-1}]_h\, dz \\ & + \|\nabla\alpha\|_{L^\infty(B_\rho(\mathfrak{x}))}\|\beta\|_{L^\infty(t_1,t_2)} \iint_{B_\rho(\mathfrak{x})\times(t_1,t_2)} [|\vec{f}|^{p-1} + |\vec{g}|^{p-1}]_h\, dz \\ & + \|\phi\|_{L^\infty(B_\rho(\mathfrak{x}))}\|\varphi'\|_{L^\infty(t_1,t_2)} \iint_{B_\rho(\mathfrak{x})\times(t_1,t_2)} |[\varphi - \phi]_h|\, dz. \end{aligned}$$

4.3. Properties of the test function

Lemma 4.9. *For any $z \in E_\lambda^c$, we have*

$$|v_{\lambda,h}(z)| \lesssim_{(n,p,q,\Lambda_0,\Lambda_1,b_0)} \rho\lambda. \tag{4.7}$$

Proof. By construction of the extension in (4.5), for $z \in E_\lambda^c$, we see that $v_{\lambda,h}(z) = \sum_j \Psi_j(z) v_h^j$ with $v_h^j = 0$ whenever $\frac{3}{4}Q_j \nsubseteq B_\rho(\mathfrak{x}) \times [\mathfrak{t}-s,\infty)$.

In order to prove the Lemma, making use of (**W8**), we see that (4.7) follows if the following holds:

$$|v_h^j| \lesssim_{(n,p,q,\Lambda_0,\Lambda_1,b_0)} \rho\lambda. \tag{4.8}$$

We shall now proceed with proving (4.8). Since we only have to consider the case $\frac{3}{4}Q_j \subset B_\rho(\mathfrak{x}) \times [\mathfrak{t}-s,\infty)$, which automatically implies $r_j \lesssim \rho$. We now proceed as follows:

Case $r_j \geq \rho$: In this case, we observe that $B_\rho(\mathfrak{x}) \subset 2B_j$ which gives the following sequence of estimates:

$$
\begin{aligned}
|v_h^j| &\lesssim r_j \iint_{\frac{3}{2}Q_j} \left| \frac{[\varphi-\phi]_h(z)}{r_j} \right| \chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\, dz \\
&\overset{(a)}{\lesssim} \rho \frac{1}{|16I_j|} \int_{16I_j\cap[\mathfrak{t}-s,\mathfrak{t}+s]} \left(\fint_{16B_j} \left| \frac{[\varphi-\phi]_h(x,t)}{r_j} \right|^q dx \right)^{\frac{1}{q}} dt \\
&\overset{(b)}{\lesssim} \rho \frac{1}{|16I_j|} \int_{16I_j\cap[\mathfrak{t}-s,\mathfrak{t}+s]} \left(\fint_{16B_j} |\nabla v_h(x,t)|^q \chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\, dx \right)^{\frac{1}{q}} dt \\
&\overset{(c)}{\lesssim} \rho\lambda.
\end{aligned}
$$

To obtain (a), we used the fact that $r_j \lesssim \rho$ along with Hölder's inequality, to obtain (b), we made use of Poincaré's inequality and finally to obtain (c), we made use of (**W4**).

Case $\frac{3}{4}r_j \leq \rho$: In this case, we gradually enlarge $\frac{3}{4}Q_i$ until it goes outside $B_\rho(\mathfrak{x}) \times [-s,\infty)$. As a consequence, we have to further consider two subcases, the first where $2^{\tilde{k}_1}Q_j$ crosses the lateral boundary first, and the second when $2^{\tilde{k}_2}Q_j$ crosses the initial boundary first.

Let us define the following constant $k_0 := \min\{\tilde{k}_1, \tilde{k}_2\}$ where $\tilde{k}_1$ and $\tilde{k}_2$ satisfy

$$
\begin{gathered}
2^{\tilde{k}_1-1}r_j < \rho \leq 2^{\tilde{k}_1}r_j, \\
2^{\tilde{k}_2-1}Q_j \subset B_\rho(\mathfrak{x}) \times [\mathfrak{t}-s,\infty) \ \text{ but } \ 2^{\tilde{k}_2}Q_j \nsubseteq B_\rho(\mathfrak{x}) \times [\mathfrak{t}-s,\infty).
\end{gathered} \tag{4.9}
$$

Note that k_0 denotes the first scaling exponent under which either we end up in the situation $r_j \geq 2^{k_0}\rho$ or $2^{k_0}Q_j$ goes outside $B_\rho(\mathfrak{x}) \times [\mathfrak{t}-s,\infty)$.

Since we only consider the case $\frac{3}{4}Q_i \subset B_\rho(\mathfrak{x}) \times [\mathfrak{t}-s,\infty)$, using triangle inequality, we get

$$
\begin{aligned}
|v_h^j| &\lesssim \sum_{m=0}^{k_0-2} \left(\left([\varphi-\phi]_h \chi_{Q_{\rho,s}(\mathfrak{z})} \right)_{2^m Q_j} - \left([\varphi-\phi]_h \chi_{Q_{\rho,s}(\mathfrak{z})} \right)_{2^{m+1}Q_j} \right) + \left([\varphi-\phi]_h \chi_{Q_{\rho,s}(\mathfrak{z})} \right)_{2^{k_0-1}Q_j} \\
&:= \sum_{m=0}^{k_0-2} S_1^m + S_2.
\end{aligned} \tag{4.10}
$$

We shall estimate S_1^m and S_2 separately as follows:

Estimate for S_1^m: In this case, we see that $2^{m+1}Q_j \subset B_\rho(\mathfrak{x}) \times [\mathfrak{t}-s,\infty)$. Thus applying Lemma 4.2 for any $\mu \in C_c^\infty(B_{2^{m+1}r_j}(x_j))$ satisfying $|\mu(x)| \leq \frac{C(n)}{(2^{m+1}r_j)^n}$ and $|\nabla\mu(x)| \leq \frac{C(n)}{(2^{m+1}r_j)^{n+1}}$, we get

$$
\begin{aligned}
S_1^m &\lesssim (2^{m+1}r_j) \left(\iint_{2^{m+1}Q_j} |[\nabla(\varphi-\phi)]_h|^q \chi_{Q_{\rho,s}(\mathfrak{z})}\, dz \right)^{\frac{1}{q}} \\
&\quad + (2^{m+1}r_j) \left(\sup_{t_1,t_2\in 2^{m+1}I_j\cap[\mathfrak{t}-s,\mathfrak{t}+s]} \left| \frac{([\varphi-\phi]_h)_\mu(t_2) - ([\varphi-\phi]_h)_\mu(t_1)}{2^{m+1}r_j} \right|^q \right)^{\frac{1}{q}} \\
&\overset{(\mathbf{W4})}{\lesssim} (2^{m+1}r_j)\lambda + (2^{m+1}r_j) \left(\sup_{t_1,t_2\in 2^{m+1}I_j\cap[\mathfrak{t}-s,\mathfrak{t}+s]} \left| \frac{([\varphi-\phi]_h)_\mu(t_2) - ([\varphi-\phi]_h)_\mu(t_1)}{2^{m+1}r_j} \right|^q \right)^{\frac{1}{q}}.
\end{aligned} \tag{4.11}
$$

To estimate the second term on the right of (4.11), using $B_{2^{m+1}r_j}(x_j) \subset B_\rho(\mathfrak{x}) \times [\mathfrak{t}-s,\infty)$, we can apply Lemma 4.8 with the test function $\alpha(x) = \mu(x)$ and $\beta(t) = 1$, which gives for any $t_1, t_2 \in \frac{3}{4}I_j \cap [\mathfrak{t}-s,\mathfrak{t}+s]$, the estimate

$$
\left| ([\varphi-\phi]_h)_\mu(t_2) - ([\varphi-\phi]_h)_\mu(t_1) \right| \overset{(a)}{\lesssim} 2^{m+1}r_j \left(\kappa\lambda^{p-1}\right) = 2^{m+1}r_j\lambda. \tag{4.12}
$$

To obtain (a), we first applied Lemma 4.8 along with (**W1**), (**W4**) and the definition $\kappa = \lambda^{2-p}$.
Substituting (4.12) into (4.11), we get

$$
S_1^m \lesssim 2^{m+1}r_j\lambda. \tag{4.13}
$$

Estimate for S_2: For this term, we know that $2^{k_0-1}Q_j \notin B_\rho(\mathfrak{x}) \times [\mathfrak{t}-s,\infty)$, which implies $2^{k_0-1}Q_j$ crosses either the lateral boundary $\partial B_\rho(\mathfrak{x}) \times [\mathfrak{t}-s,\infty)$ or crosses the initial boundary $B_\rho(\mathfrak{x}) \times \{\mathfrak{t}-s\}$ first. We will consider both the cases separately and estimate S_2 as follows:

In the case $2^{k_0-1}Q_j$ crosses the lateral boundary $\partial B_\rho(\mathfrak{x}) \times [\mathfrak{t}-s,\infty)$ first, we can directly apply Theorem 2.13 to obtain

$$\iint_{2^{k_0-1}Q_j} [\varphi-\phi]_h \chi_{Q_{\rho,s}(\mathfrak{z})}\, dz \lesssim (2^{k_0}r_j)\left(\iint_{2^{k_0}Q_j} |\nabla[\varphi-\phi]_h|^q \chi_{Q_{\rho,s}(\mathfrak{z})}\, dz\right)^{1/q} \overset{(a)}{\lesssim} \rho\lambda. \tag{4.14}$$

To obtain (a), we made use of (**W4**) along with $2^{k_0-2}r_j \le \rho$ given by (4.9).

In the case $2^{k_0}Q_j$ crosses the initial boundary $B_\rho(\mathfrak{x}) \times \{\mathfrak{t}-s\}$ first, by enlarging the cylinder to $2^{k_1+1}Q_j$, we can find a cut-off function $\theta(x,t)$ such that $\operatorname{spt}\theta(x,t) \subset 2^{k_1+1}Q_j \cap \mathbb{R}^n \times (-\infty,\mathfrak{t}-s)$, which combined with the fact $v_h(z)\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} = 0$ on $\mathbb{R}^n \times (-\infty,\mathfrak{t}-s)$, we get $\left(v_h\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\right)_\theta = 0$. Thus applying Lemma 4.2, we get

$$\begin{aligned}
\iint_{2^{k_0+1}Q_j} |v_h(z)|\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\, dz &= \iint_{2^{k_0+1}Q_j} \left|v_h(z)\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} - \left(v_h\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\right)_\theta\right| dz\\
&\lesssim (2^{k_0+1}r_j)\left(\iint_{2^{k_0+1}Q_j} |[\nabla(\varphi-\phi)]_h|^q \chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\, dz\right)^{\frac{1}{q}}\\
&\quad +(2^{k_0+1}r_j)\left(\sup_{t_1,t_2\in 2^{k_0+1}I_j\cap[\mathfrak{t}-s,\mathfrak{t}+s]} \left|\frac{([\varphi-\phi]_h)_\mu(t_2) - ([\varphi-\phi]_h)_\mu(t_1)}{2^{k_0+1}r_j}\right|^q\right)^{\frac{1}{q}}\\
&\overset{(a)}{\lesssim} 2^{k_0+1}r_j\lambda \overset{(b)}{\lesssim} \rho\lambda.
\end{aligned} \tag{4.15}$$

To obtain (a), we made use of (**W1**),(**W4**) along with an application of Lemma 4.8 and to obtain (b), we used (4.9).

Combining (4.14) and (4.15), we get

$$S_2 \lesssim \rho\lambda. \tag{4.16}$$

Thus combining (4.13) and (4.16) into (4.10), we get

$$|v_h^j| \le \sum_{m=0}^{k_0-2} S_1^m + S_2 \lesssim \lambda\left(\sum_{m=0}^{k_0-2} 2^{m+1}r_j + \rho\right) \overset{(4.9)}{\lesssim} \rho\lambda.$$

This completes the proof of the Lemma. ☐

Now we prove a sharper estimate.

Lemma 4.10. *For any $j \in A_i$, there holds*

$$|v_h^i - v_h^j| \lesssim_{(n,p,q,\Lambda_0,\Lambda_1,m_e)} \min\{\rho, r_i\}\lambda.$$

Proof. We only have to consider the case $r_i \le \rho$ because if $\rho \le r_i$, we can directly use Lemma 4.9 to get the required conclusion.

If either $v_h^i = 0$ or $v_h^j = 0$, then $\frac{3}{4}Q_i$ must necessarily intersect the lateral or initial boundary.

Initial Boundary Case $\frac{3}{4}Q_i \subset B_\rho(\mathfrak{x}) \times \mathbb{R}$: Without loss of generality, we can assume $2Q_i \subset B_\rho(\mathfrak{x}) \times \mathbb{R}$. We now pick $\theta(x,t) \in C_c^\infty(\mathbb{R}^{n+1})$ such that $\operatorname{spt}(\theta) \subset 2B_i \times (-\infty,\mathfrak{t}-s)$. Since $\varphi-\phi = 0$ on $2B_i \times (-\infty,\mathfrak{t}-s)$, we see that $\left(v_h\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\right)_\theta = \left([\varphi-\phi]_h\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\right)_\theta = 0$. Thus we get

$$\begin{aligned}
|v_h^i| &\lesssim \iint_{2Q_i} \left|[\varphi-\phi]_h\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} - \left([\varphi-\phi]_h\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\right)_\theta\right| dz\\
&\overset{(a)}{\lesssim} r_i\left(\iint_{2Q_i} |\nabla v_h|^q \chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\, dz + \sup_{t_1,t_2\in 2I_i\cap[\mathfrak{t}-s,\mathfrak{t}+s]} \left|\frac{\left(v_h\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\right)_\mu(t_2) - \left(v_h\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\right)_\mu(t_1)}{r_i}\right|^q\right)^{\frac{1}{q}}\\
&\overset{(b)}{\lesssim} r_i\lambda.
\end{aligned}$$

To obtain (a), we made use of Lemma 4.2 and to obtain (b), we proceed similarly to how (4.12) was estimated.

Lateral Boundary Case $\frac{3}{4}Q_i \cap (B_\rho(\mathfrak{x}) \times \mathbb{R})^c \neq \emptyset$: In this case, using Theorem 2.13 and (**W4**), we get

$$|v_h^i| \lesssim r_i \left(\fint\fint_{2Q_i} \left| \frac{[\varphi - \phi]_h \chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}}{r_i} \right|^q \, dz \right)^{\frac{1}{q}} \lesssim r_i \left(\fint\fint_{2Q_i} |\nabla[\varphi - \phi]_h|^q \chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} \, dz \right)^{\frac{1}{q}} \lesssim r_i \lambda. \tag{4.17}$$

From (4.17) and (4.8), we see that the lemma is proved provided $v_h^j = 0$.

Now let us consider the case $v_h^i \neq 0$ and $v_h^j \neq 0$, which implies $\frac{3}{4}Q_i \subset B_\rho(\mathfrak{x}) \times [-s, \infty)$ and $\frac{3}{4}Q_j \subset B \times [\mathfrak{t} - s, \infty)$. From the definition of v_h^i in (4.6), triangle inequality and (**W12**), we get

$$\begin{aligned} |v_h^i - v_h^j| &\lesssim \frac{|\frac{3}{4}Q_i|}{|\frac{3}{4}Q_i \cap \frac{3}{4}Q_j|} \fint\fint_{\frac{3}{4}Q_i} \left| v_h(z)\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} - v_h^i \right| \, dz + \frac{|\frac{3}{4}Q_j|}{|\frac{3}{4}Q_i \cap \frac{3}{4}Q_j|} \fint\fint_{\frac{3}{4}Q_j} \left| v_h(z)\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} - v_h^j \right| \, dz \\ &\lesssim \fint\fint_{\frac{3}{4}Q_i} \left| v_h(z)\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} - v_h^i \right| \, dz + \fint\fint_{\frac{3}{4}Q_j} \left| v_h(z)\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} - v_h^j \right| \, dz. \end{aligned} \tag{4.18}$$

Let us now estimate each of the terms in (4.18) as follows: we apply Hölder's inequality followed by Lemma 4.2 with $\alpha \in C_c^\infty\left(\frac{3}{4}B_i\right)$ with $|\alpha(x)| \lesssim \frac{1}{r_i^n}$ and $|\nabla\alpha(x)| \lesssim \frac{1}{r_i^{n+1}}$ to get

$$\begin{aligned} \fint\fint_{\frac{3}{4}Q_i} |v_h(z)\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} - v_h^i| \, dz &= r_i \left(\fint\fint_{\frac{3}{4}Q_i} |\nabla v_h|^q \chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} \, dz \right)^{\frac{1}{q}} \\ &\quad + r_i \left(\sup_{t_1,t_2 \in \frac{3}{4}I_i \cap [\mathfrak{t}-s,\mathfrak{t}+s]} \left| \frac{\left([\varphi-\phi]_h\right)_\mu(t_2) - \left([\varphi-\phi]_h\right)_\mu(t_1)}{r_i} \right|^q \right)^{\frac{1}{q}}. \end{aligned} \tag{4.19}$$

The first term on the right of (4.19) can be controlled using (**W4**) and the second term can be controlled similarly as (4.12). Thus we get

$$\fint\fint_{\frac{3}{4}Q_i} |v_h(z)\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} - v_h^i| \, dz \lesssim r_i \lambda.$$

This completes the proof of the Lemma. ∎

Once we have the bounds in Lemma 4.9 and Lemma 4.10, we can obtain the following important estimates:

Lemma 4.11. *Given any $z \in E_\lambda^c$, we have $z \in \frac{3}{4}Q_i$ for some $i \in \mathbb{N}$. Then there holds*

$$|\nabla v_{\lambda,h}(z)| \leq C_{(n,p,q,\Lambda_0,\Lambda_1,m_e)} \lambda. \tag{4.20}$$

Proof. We observe that $\sum_j \Psi_j(z) = \sum_{j:j\in A_i} \Psi_j(z) = 1$ for any $z \in E_\lambda^c$, which implies $\sum_j \nabla\Psi_j(z) = 0$ for all $z \in E_\lambda^c$. Thus using (4.5) along with (**W9**), (**W13**) and Lemma 4.10, we get

$$|\nabla v_{\lambda,h}(z)| \leq \sum_{j:j\in A_i} |\nabla\Psi_j(z)| \left| v_h^j - v_h^i \right| \lesssim \lambda.$$

This completes the proof of the Lemma. ∎

4.4. Estimates on the derivative of $v_{\lambda,h}$

We will now mention some improved estimates which can be proved using Hölder's inequality along with the techniques from Lemma 4.11.

Lemma 4.12. *Let $z \in E_\lambda^c$ and $\varepsilon \in (0,1]$ be any number, then $z \in \frac{3}{4}Q_i$ for some $i \in \mathbb{N}$ from* (**W1**). *There exists a constant $C = C_{(n,p,q,\Lambda_0,\Lambda_1,m_e)}$ such that the following holds:*

$$|v_{\lambda,h}(z)| \leq C \fint\fint_{4Q_i} |v_h(\tilde{z})| \chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} \, d\tilde{z} \leq \frac{C r_i \lambda}{\varepsilon} + \frac{C\varepsilon}{\lambda r_i} \fint\fint_{4Q_i} |v_h(\tilde{z})|^2 \chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} \, d\tilde{z},$$

$$|\nabla v_{\lambda,h}(z)| \leq C \frac{1}{r_i} \fint\fint_{4Q_i} |v_h(\tilde{z})| \chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} \, d\tilde{z} \leq \frac{C\lambda}{\varepsilon} + \frac{C\varepsilon}{\lambda r_i^2} \fint\fint_{4Q_i} |v_h(\tilde{z})|^2 \chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} \, d\tilde{z}.$$

Lemma 4.13. *Let $z \in E_\lambda^c$ and $\varepsilon \in (0,1]$ be any number, then $z \in \frac{3}{4}Q_i$ for some $i \in \mathbb{N}$ from (**W1**). There exists a constant $C = C_{(n,p,q,\Lambda_0,\Lambda_1,m_e)}$ such that the following holds:*

$$|v_{\lambda,h}(z)| \leq C\left(\min\{\rho, r_i\}\lambda + |v_h^i|\right) \leq C\left(\frac{r_i\lambda}{\varepsilon} + \frac{\varepsilon}{r_i\lambda}|v_h^i|^2\right), \tag{4.21}$$

$$|\nabla v_{\lambda,h}(z)| \leq C\frac{\lambda}{\varepsilon}, \tag{4.22}$$

$$|\partial_t v_{\lambda,h}(z)| \leq C\frac{1}{\lambda^{2-p}r_i^2}\fint\fint_{4Q_i} |v_h(\tilde{z})|\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\, d\tilde{z},$$

$$|\partial_t v_{\lambda,h}(z)| \leq C\frac{1}{\lambda^{2-p}r_i^2}\min\{r_i, \rho\}\lambda. \tag{4.23}$$

4.5. Some more properties of $v_{\lambda,h}$

Lemma 4.14. *For any $\vartheta \geq 1$, we have the following bound:*

$$\iint_{Q_{\rho,s}(\mathfrak{z})\setminus E_\lambda} |v_{\lambda,h}(z)|^\vartheta\, dz \lesssim_{(n,p,q,\Lambda_0,\Lambda_1,m_e)} \iint_{Q_{\rho,s}(\mathfrak{z})\setminus E_\lambda} |v_h(z)|^\vartheta \chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\, dz.$$

Proof. Since E_λ^c is covered by Whitney cylinders (see Lemma 4.5), let us pick some $i \in \mathbb{N}$ and consider the corresponding parabolic Whitney cylinder. Using the construction from (4.5) along with (**W5**), (**W9**) and (**W13**), we get

$$\iint_{\frac{3}{4}Q_i} |v_{\lambda,h}(z)|^\vartheta\, dz \lesssim \sum_{j:j\in A_i}\iint_{\frac{3}{4}Q_i} \Psi_j(z)^\vartheta |v_h^j|^\vartheta\, dz \lesssim \iint_{4Q_i} |v_h(z)|^\vartheta \chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\, dz. \tag{4.24}$$

Summing (4.24) over all $i \in \mathbb{N}$ and making use of (**W4**) and (**W7**), we get

$$\iint_{Q_{\rho,s}(\mathfrak{z})\setminus E_\lambda} |v_{\lambda,h}(z)|^\vartheta\, dz \lesssim \sum_i \iint_{4Q_i} |v_h(z)|^\vartheta \chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\, dz \lesssim \iint_{Q_{\rho,s}(\mathfrak{z})\setminus E_\lambda} |v_h(z)|^\vartheta \chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\, dz.$$

This proves the Lemma. □

Lemma 4.15. *For any $0 < \vartheta \leq q$ with q defined as in 4.4, there holds*

$$\iint_{Q_{\rho,s}(\mathfrak{z})\setminus E_\lambda} |\partial_t v_{\lambda,h}(z)(v_{\lambda,h}(z) - v_h(z))|^\vartheta\, dz \lesssim_{(n,p,q,\Lambda_0,\Lambda_1,m_e,\vartheta)} \lambda^{\vartheta p}|\mathbb{R}^{n+1}\setminus E_\lambda|.$$

Proof. From (**W3**), we see that $Q_{\rho,s}(\mathfrak{z}) \setminus E_\lambda \subset \bigcup_{i\in\mathbb{Z}} 4Q_i$, thus, for a given $i \in \mathbb{N}$, let us define the following:

$$J_i := \iint_{\frac{3}{4}Q_i} |\partial_t v_{\lambda,h}(z)(v_{\lambda,h}(z) - v_h(z))|^\vartheta \chi_{Q_{\rho,s}(\mathfrak{z})}\, dz.$$

Making use of (4.23) and Hölder's inequality (recall $\gamma = \lambda^{2-p}$), we get

$$\begin{aligned} J_i &\lesssim \left(\frac{1}{\lambda^{2-p}r_i^2}r_i\lambda\right)^\vartheta \iint_{\frac{3}{4}Q_i} |v_{\lambda,h}(z)\chi_{Q_{\rho,s}(\mathfrak{z})} - v_h(z)\chi_{Q_{\rho,s}(\mathfrak{z})}|^\vartheta\, dz \\ &\overset{(a)}{\lesssim} \left(\frac{1}{\lambda^{2-p}r_i^2}r_i\lambda\right)^\vartheta \sum_{j\in A_i}\iint_{\frac{3}{4}Q_i} |v_h(z)\chi_{Q_{\rho,s}(\mathfrak{z})} - v_h^j|^\vartheta\, dz \\ &\overset{(b)}{\lesssim} \lambda^{\vartheta p}\, |\tfrac{3}{4}Q_i|. \end{aligned} \tag{4.25}$$

To obtain (a), we made use of (4.5), (**W9**) and (**W10**) and to obtain (b), we applied Theorem 2.13 along with (**W4**).

Summing (4.25) over all $i \in \mathbb{N}$ and making use of (**W7**) completes the proof of the lemma. □

4.6. Proof of the Lipschitz continuity of $v_{\lambda,h}$

We shall now prove the Lipschitz continuity of $v_{\lambda,h}$ on $\mathcal{H} := \mathbb{R}^n \times [\mathfrak{t}-s, \mathfrak{t}+s]$.

Lemma 4.16. *The function $v_{\lambda,h}$ from (4.5) is $C^{0,1}(\mathcal{H})$ with respect to the parabolic metric given in Definition (2.14).*

Proof. Let us consider a parabolic cylinder $Q_r(z) := Q_{r,\kappa r^2}(z) := Q$ for some $z \in \mathcal{H}$ and $r > 0$ (recall $\kappa = \lambda^{2-p}$). To prove the Lemma, we make use of Lemma 4.4 and prove the following bound:

$$I_r(z) := \iint_{Q\cap\mathcal{H}} \left| \frac{v_{\lambda,h}(\tilde{z}) - \left(v_{\lambda,h}\right)_{Q\cap\mathcal{H}}}{r} \right| \, d\tilde{z} \leq o(1),$$

where $o(1)$ denotes a constant independent of $z \in \mathcal{H}$ and $r > 0$ only. We will split the proof into several subcases and proceed as follows:

Case $2Q \subset E_\lambda^c$: In this case, from (**W3**), we see that $z \in \frac{3}{4}Q_i$ for some $i \in \mathbb{N}$. From the construction in (4.5), we see that $v_{\lambda,h} \in C^\infty(E_\lambda^c)$ which combined with the mean value theorem gives

$$I_r(z) \lesssim \frac{1}{r} \iint_{Q\cap\mathcal{H}} \iint_{Q\cap\mathcal{H}} |v_{\lambda,h}(\tilde{z}_1) - v_{\lambda,h}(\tilde{z}_2)| \, d\tilde{z}_1 \, d\tilde{z}_2 \lesssim \sup_{\tilde{z}\in Q\cap\mathcal{H}} \left(|\nabla v_{\lambda,h}(\tilde{z})| + \lambda^{2-p} r |\partial_t v_{\lambda,h}(\tilde{z})| \right).$$

Let us pick some $\tilde{z}_0 \in 2Q \subset E_\lambda^c$, then $\tilde{z}_0 \in Q_j$ for some $j \in \mathbb{N}$. Thus we can make use of (4.20) and (4.23) to get

$$|\nabla v_{\lambda,h}(\tilde{z}_0)| + \lambda^{2-p} r |\partial_t v_{\lambda,h}(\tilde{z}_0)| \lesssim \lambda + \lambda^{2-p} r \frac{1}{\lambda^{2-p} r_j^2} r_j \lambda. \tag{4.26}$$

In (4.26), we need to understand the relation between r_j and r. To this end, from $2Q \subset E_\lambda^c$, we see that

$$r \leq d_\lambda(\tilde{z}_0, E_\lambda) \leq d_\lambda(\tilde{z}_0, z_j) + d_\lambda(z_j, E_\lambda) \leq r_j + 16r_j = 17r_i. \tag{4.27}$$

Combining (4.26) and (4.27), we get

$$|\nabla v_{\lambda,h}(\tilde{z}_0)| + \lambda^{2-p} r |\partial_t v_{\lambda,h}(\tilde{z}_0)| \lesssim \lambda.$$

Case $2Q \nsubseteq E_\lambda^c$: In this case, we shall split the proof into three subcases:

Subcase $2Q \subset \mathbb{R}^n \times (-\infty, \mathfrak{t} + s]$ or $2Q \subset \mathbb{R}^n \times [\mathfrak{t} - s, \infty)$: In this situation, it is easy to see that the following holds:

$$|Q \cap \mathcal{H}| \gtrsim |Q|. \tag{4.28}$$

We apply triangle inequality and estimate $I_r(z)$ by

$$\begin{aligned} I_r(z) &\leq \iint_{Q\cap\mathcal{H}} \left| \frac{v_{\lambda,h}(\tilde{z}) - v_h(\tilde{z})}{r} \right| + \left| \frac{v_h(\tilde{z}) - (v_h)_{Q\cap\mathcal{H}}}{r} \right| + \left| \frac{(v_h)_{Q\cap\mathcal{H}} - \left(v_{\lambda,h}\right)_{Q\cap\mathcal{H}}}{r} \right| \, d\tilde{z} \\ &\leq 2J_1 + J_2, \end{aligned} \tag{4.29}$$

where we have set

$$J_1 := \iint_{Q\cap\mathcal{H}} \left| \frac{v_{\lambda,h}(\tilde{z}) - v_h(\tilde{z})}{r} \right| \, d\tilde{z} \qquad \text{and} \qquad J_2 := \iint_{Q\cap\mathcal{H}} \left| \frac{v_h(\tilde{z}) - (v_h)_{Q\cap\mathcal{H}}}{r} \right| \, d\tilde{z}. \tag{4.30}$$

We now estimate each of the terms of (4.30) as follows:

Estimate for J_1: From (4.5), we get

$$J_1 \lesssim \sum_{i\in\mathbb{N}} \frac{1}{|Q\cap\mathcal{H}|} \iint_{Q\cap\mathcal{H}\cap\frac{3}{4}Q_i} \left| \frac{v_h(\tilde{z})\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} - v_h^i}{r} \right| \, d\tilde{z}. \tag{4.31}$$

Let us fix an $i \in \mathbb{N}$ and take two points $\tilde{z}_1 \in Q \cap \frac{3}{4}Q_i$ and $\tilde{z}_2 \in E_\lambda \cap 2Q$. Let z_i denote the center of $\frac{3}{4}Q_i$, making use of (**W2**) along with the trivial bound $d_\lambda(\tilde{z}_1, \tilde{z}_2) \leq 4r$ and $d_\lambda(z_i, \tilde{z}_1) \leq 2r_i$, we get

$$16r_i = d_\lambda(z_i, E_\lambda) \leq d_\lambda(z_i, \tilde{z}_1) + d_\lambda(\tilde{z}_1, \tilde{z}_2) \leq 2r_i + 4r \implies 2r_i \leq r. \tag{4.32}$$

Note that (4.28) holds and thus summing over all $i \in \mathbb{N}$ such that $Q \cap \mathcal{H} \cap \frac{3}{4}Q_i \neq \emptyset$ in (4.31) and making use of (4.32), we get

$$\begin{aligned} J_1 &\lesssim \sum_{\substack{i\in\mathbb{N} \\ Q\cap\mathcal{H}\cap\frac{3}{4}Q_i\neq\emptyset}} \frac{|\frac{3}{4}Q_i|}{|Q\cap\mathcal{H}|} \iint_{\frac{3}{4}Q_i} \left| \frac{v_h(\tilde{z})\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} - v_h^i}{r} \right| \, d\tilde{z} \\ &\overset{(a)}{\lesssim} \sum_{i\in\mathbb{N}} \fint\fint_{\frac{3}{4}Q_i} \left| \frac{v_h(\tilde{z})\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} - v_h^i}{r_i} \right| \, d\tilde{z} \\ &\overset{(b)}{\lesssim} \lambda. \end{aligned}$$

To obtain (a), we made use of (4.28) and (4.32), to obtain (b), we follow the calculation from bounding (4.19).

Estimate for J_2: Note that $Q\cap\mathcal{H}$ is another cylinder. *If $Q \subset B_\rho(\mathfrak{x})\times\mathbb{R}$,* then choose a cut-off function $\mu \in C_c^\infty(B_\rho(\mathfrak{x}))$ and apply Lemma 4.2 to get

$$J_2 \lesssim \left(\fint\fint_{Q\cap\mathcal{H}} |\nabla v_h|^q \chi_{Q_{\rho,s}(\mathfrak{z})} + \sup_{t_1,t_2\in[\mathfrak{t}-s,\mathfrak{t}+s]\cap Q} \left| \frac{\left(v_h\chi_{Q_{\rho,s}(\mathfrak{z})}\right)_\mu(t_1) - \left(v_h\chi_{Q_{\rho,s}(\mathfrak{z})}\right)_\mu(t_1)}{r} \right|^q \right)^{\frac{1}{q}}.$$

Recall that we are in the case $2Q\cap E_\lambda \neq \emptyset$ and $2Q\cap E_\lambda^c \neq \emptyset$. Further applying Lemma 4.8 and proceeding as in (4.11), we get

$$J_2 \lesssim \lambda. \tag{4.33}$$

On the other hand, if $Q \not\subseteq B_\rho(\mathfrak{x})\times\mathbb{R}$, then we can apply Poincaré's inequality from Theorem 2.13 directly and make use of the fact that $2Q\cap E_\lambda \neq \emptyset$ to get

$$J_2 \lesssim \left(\fint\fint_{Q\cap\mathcal{H}} \left|\nabla v_h(\tilde{z})\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]}\right|^q \, d\tilde{z} \right)^{\frac{1}{q}} \lesssim \lambda.$$

Subcase $2Q\cap\mathbb{R}^n\times(-\infty,\mathfrak{t}-s]\neq\emptyset$ and $2Q\cap\mathbb{R}^n\times[\mathfrak{t}+s,\infty)\neq\emptyset$ AND $\kappa r^2\leq s$: In this case, we see that

$$|Q\cap\mathcal{H}| = |B_1|r^n \times 2s.$$

We apply triangle inequality and estimate $I_r(z)$ as we did in (4.29) to get

$$I_r(z) \quad \leq 2J_1 + J_2,$$

where we have set

$$J_1 := \fint\fint_{Q\cap\mathcal{H}} \left| \frac{v_{\lambda,h}(\tilde{z}) - v_h(\tilde{z})}{r} \right| \, d\tilde{z} \qquad \text{and} \qquad J_2 := \fint\fint_{Q\cap\mathcal{H}} \left| \frac{v_h(\tilde{z}) - (v_h)_{Q\cap\mathcal{H}}}{r} \right| \, d\tilde{z}.$$

We estimate J_1 as follows

$$\begin{aligned} J_1 &\lesssim \sum_{i\in\mathbb{N}} \frac{|\frac{3}{4}Q_i|}{|Q\cap\mathcal{H}|} \iint_{\frac{3}{4}Q_i} \left| \frac{v_h(\tilde{z})\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} - v_h^i}{r} \right| \, d\tilde{z} \\ &\overset{(4.32)}{\lesssim} \frac{r_i^{n+2}\kappa}{r^n s} \sum_{i\in\mathbb{N}} \fint\fint_{\frac{3}{4}Q_i} \left| \frac{v_h(\tilde{z})\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} - v_h^i}{r_i} \right| \, d\tilde{z} \\ &\overset{(4.32)}{\lesssim} \frac{r^{n+2}\kappa}{r^n s} \sum_{i\in\mathbb{N}} \fint\fint_{\frac{3}{4}Q_i} \left| \frac{v_h(\tilde{z})\chi_{[\mathfrak{t}-s,\mathfrak{t}+s]} - v_h^i}{r_i} \right| \, d\tilde{z} \\ &\overset{(a)}{\lesssim} \frac{r^2\kappa}{s}\lambda \\ &\overset{(b)}{\lesssim} \lambda. \end{aligned}$$

To obtain (a), we proceed similarly to (4.19) and to obtain (b), we made use of $\kappa r^2 \leq s$.

The estimate for J_2 is already obtained in (4.33) which shows

$$J_2 \lesssim \lambda.$$

Subcase $2Q\cap\mathbb{R}^n\times(-\infty,\mathfrak{t}-s]\neq\emptyset$ and $2Q\cap\mathbb{R}^n\times[\mathfrak{t}+s,\infty)\neq\emptyset$ AND $\kappa r^2 > s$: Using triangle inequality and the bound $|Q \cap \mathcal{H}| = |B_1| r^n \times 2s$, we get

$$\begin{aligned}
\fint\!\!\!\fint_{Q\cap\mathcal{H}} \left| \frac{v_{\lambda,h}(\tilde{z}) - \left(v_{\lambda,h}\right)_{Q\cap\mathcal{H}}}{r} \right| d\tilde{z} &\lesssim \frac{1}{|Q\cap\mathcal{H}|} \iint_{Q\cap\mathcal{H}} |v_{\lambda,h}(\tilde{z})| \, d\tilde{z} \\
&\lesssim \frac{1}{|Q\cap\mathcal{H}|} \iint_{Q\cap\mathcal{H}\cap E_\lambda} |v_{\lambda,h}(\tilde{z})| \, d\tilde{z} + \frac{1}{|Q\cap\mathcal{H}|} \iint_{Q\cap\mathcal{H}\setminus E_\lambda} |v_{\lambda,h}(\tilde{z})| \, d\tilde{z}.
\end{aligned}$$

By construction of $v_{\lambda,h}$ in (4.5), we have $v_{\lambda,h} = v_h$ on E_λ. On $Q_{\rho,s}(\mathfrak{z}) \setminus E_\lambda$, we can apply Lemma 4.9 to obtain the following bound:

$$\begin{aligned}
\fint\!\!\!\fint_{Q\cap\mathcal{H}} \left| \frac{v_{\lambda,h}(\tilde{z}) - (Q\cap\mathcal{H})_{v_{\lambda,h}}}{r} \right| d\tilde{z} &\lesssim \frac{1}{r^n s} \iint_{Q_{\rho,s}(\mathfrak{z})} |v_h(\tilde{z})| \, d\tilde{z} + \frac{1}{|Q\cap\mathcal{H}|} \iint_{Q\cap\mathcal{H}\setminus E_\lambda} \rho\lambda \, d\tilde{z} \\
&\lesssim \left(\frac{\kappa}{s}\right)^{\frac{n}{2}} \frac{1}{s} \|v_h\|_{L^1(Q_{\rho,s}(\mathfrak{z}))} + \rho\lambda \\
&\lesssim o(1).
\end{aligned}$$

This completes the proof of the Lipschitz regularity. □

4.7. Two crucial estimates

We shall now prove the first crucial estimate which holds on each time slice.

Lemma 4.17. *For any $i \in \mathbb{N}$ and any $0 < \varepsilon \leq 1$, there exists a positive constant $C(n, p, q, \Lambda_0, \Lambda_1, m_e)$ such that for almost every $t \in [\mathfrak{t} - s, \mathfrak{t} + s]$, there holds*

$$\left| \int_{B_\rho(\mathfrak{x})} (v(x,t) - v^i) v_\lambda(x,t) \Psi_i(x,t) \, dx \right| \leq C \left(\frac{\lambda^p}{\varepsilon} |4Q_i| + \varepsilon |4B_i| |v^i|^2 \right). \tag{4.34}$$

Proof. Let us fix any $t \in [\mathfrak{t} - s, \mathfrak{t} + s]$, $i \in \mathbb{N}$ and take $\Psi_i(y,\tau) v_{\lambda,h}(y,\tau)$ as a test function in (4.1) and (4.2). Further integrating the resulting expression over $\left(t_i - \kappa\left(\frac{3}{4} r_i\right)^2, t\right)$ or $(\mathfrak{t} - s, t)$ depending on the location of $\frac{3}{4} Q_i$, along with making use of the fact that $\Psi_i(y, t_i - \kappa(3r_i/4)^2) = 0$ or $v_{\lambda,h}(y, \mathfrak{t} - s) = 0$, we get for any $a \in \mathbb{R}$, the equality

$$\begin{aligned}
\int_{B_\rho(\mathfrak{x})} \left((v_h - a)\Psi_i v_{\lambda,h}\right)(y,t) \, dy &= \int_{\max\left\{t_i - \kappa\left(\frac{3}{4} r_i\right)^2, \mathfrak{t} - s\right\}}^{t} \int_{B_\rho(\mathfrak{x})} \partial_t \left((v_h - a)\Psi_i v_{\lambda,h}\right)(y,\tau) \, dy \, d\tau \\
&= \int_{\max\left\{t_i - \kappa\left(\frac{3}{4} r_i\right)^2, \mathfrak{t} - s\right\}}^{t} \int_{B_\rho(\mathfrak{x})} \partial_t \left([\varphi - \phi]_h \Psi_i v_{\lambda,h} - a\Psi_i v_{\lambda,h}\right)(y,\tau) \, dy \, d\tau \\
&= \int_{\max\left\{t_i - \kappa\left(\frac{3}{4} r_i\right)^2, \mathfrak{t} - s\right\}}^{t} \int_{B_\rho(\mathfrak{x})} \langle [\mathcal{A}(y,\tau,\nabla\phi)]_h - [\mathcal{A}(y,\tau,\nabla\varphi)]_h \,, \nabla(\Psi_i v_{\lambda,h}) \rangle \, dy \, d\tau \\
&\quad + \int_{\max\left\{t_i - \kappa\left(\frac{3}{4} r_i\right)^2, \mathfrak{t} - s\right\}}^{t} \int_{B_\rho(\mathfrak{x})} \langle [|\vec{f}|^{p-2}\vec{f} + |\vec{g}|^{p-2}\vec{g}]_h \,, \nabla(\Psi_i v_{\lambda,h}) \rangle \, dy \, d\tau \\
&\quad - \int_{\max\left\{t_i - \kappa\left(\frac{3}{4} r_i\right)^2, \mathfrak{t} - s\right\}}^{t} \int_{B_\rho(\mathfrak{x})} a \partial_t \left(\Psi_i v_{\lambda,h}\right) \, dy \, d\tau.
\end{aligned} \tag{4.35}$$

We can estimate $|\nabla(\Psi_i v_\lambda)|$ using the chain rule and (**W9**), to get

$$|\nabla(\Psi_i v_{\lambda,h})| \lesssim \frac{1}{r_i} |v_\lambda| + |\nabla v_\lambda|. \tag{4.36}$$

Similarly, we can estimate $\left|\partial_t \left(\Psi_i v_\lambda\right)\right|$ using the chain rule and (**W9**), to get

$$\left|\partial_t \left(\Psi_i v_\lambda\right)\right| \lesssim \frac{1}{\kappa r_i^2} |v_\lambda| + |\partial_t v_\lambda|. \tag{4.37}$$

Let us take $a = v_h^i$ in the (4.35) followed by letting $h \searrow 0$ and making use of (4.36) and (2.2), we get

$$\left| \int_{B_\rho(\mathfrak{x})} \left((v - v^i)\Psi_i v_\lambda\right)(y,t) \, dy \right| \lesssim J_1 + J_2 + J_3, \tag{4.38}$$

where we have set

$$J_1 := \frac{1}{r_i} \iint_{Q_{\rho,s}(\mathfrak{z})} \left(|\nabla\varphi|^{p-1} + |\nabla\phi|^{p-1} + |\vec{f}|^{p-1} + |\vec{g}|^{p-1} \right) |v_\lambda| \chi_{\frac{3}{4}Q_i \cap Q_{\rho,s}(\mathfrak{z})} \, dy \, d\tau,$$
$$J_2 := \iint_{Q_{\rho,s}(\mathfrak{z})} \left(|\nabla\varphi|^{p-1} + |\nabla\phi|^{p-1} + |\vec{f}|^{p-1} + |\vec{g}|^{p-1} \right) |\nabla v_\lambda| \chi_{\frac{3}{4}Q_i \cap Q_{\rho,s}(\mathfrak{z})} \, dy \, d\tau,$$
$$J_3 := \iint_{Q_{\rho,s}(\mathfrak{z})} |v - v^i| |\partial_t(\Psi_i v_\lambda)| \chi_{\frac{3}{4}Q_i \cap Q_{\rho,s}(\mathfrak{z})} \, dy \, d\tau. \tag{4.39}$$

Let us now estimate each of the terms as follows:

Bound for J_1: If $\rho \leq r_i$, we can directly use Hölder's inequality, Lemma 4.9 and (**W4**), to find that for any $\varepsilon \in (0, 1]$, there holds

$$J_1 \lesssim \lambda |Q_i| \left(\fint\!\!\!\fint_{16Q_i} \left(|\nabla\varphi|^q + |\nabla\phi|^q + |\vec{f}|^q + |\vec{g}|^q \right) \chi_{Q_{\rho,s}(\mathfrak{z})} \, dy \, d\tau \right)^{\frac{p-1}{q}} \lesssim \frac{\lambda^p}{\varepsilon} |4Q_i|. \tag{4.40}$$

In the case $r_i \leq \rho$, we make use of (4.21), (**W4**) along with the fact $|Q_i| = |B_i| \times 2\lambda^{2-p} r_i^2$, to get

$$\begin{aligned} J_1 &\lesssim \frac{1}{r_i} \left(\frac{r_i \lambda}{\varepsilon} + \frac{\varepsilon}{\lambda r_i} |v^i|^2 \right) |4Q_i| \left(\fint\!\!\!\fint_{4Q_i} \left(|\nabla\varphi|^q + |\nabla\phi|^q + |\vec{f}|^q + |\vec{g}|^q \right) \chi_{Q_{\rho,s}(\mathfrak{z})} \, dy \, d\tau \right)^{\frac{p-1}{q}} \\ &\lesssim \frac{1}{r_i} \left(\frac{r_i \lambda}{\varepsilon} + \frac{\varepsilon}{\lambda r_i} |v^i|^2 \right) |4Q_i| \lambda^{p-1} \lesssim \frac{\lambda^p}{\varepsilon} |4Q_i| + \varepsilon |4B_i| |v^i|^2. \end{aligned} \tag{4.41}$$

Thus combining (4.41) and (4.40), we get

$$J_1 \lesssim \frac{\lambda^p}{\varepsilon} |4Q_i| + \chi_{r_i \leq \rho} \varepsilon |4B_i| |v^i|^2, \tag{4.42}$$

where we have set $\chi_{r_i \leq \rho} = 1$ if $r_i \leq \rho$ and $\chi_{r_i \leq \rho} = 0$ else.

Bound for J_2: In this case, we can directly use Lemma 4.11 and (**W4**) to get for any $\varepsilon \in (0, 1]$, the bound

$$J_2 \lesssim \frac{\lambda |4Q_i|}{\varepsilon} \left(\fint\!\!\!\fint_{4Q_i} \left(|\nabla\varphi|^q + |\nabla\phi|^q + |\vec{f}|^q + |\vec{g}|^q \right) \chi_{Q_{\rho,s}(\mathfrak{z})} \, dy \, d\tau \right)^{\frac{p-1}{q}} \lesssim \frac{\lambda^p}{\varepsilon} |4Q_i|. \tag{4.43}$$

Bound for J_3: Substituting (4.22), (4.23) and (**W9**) into (4.37), for any $\varepsilon \in (0, 1]$, there holds

$$|\partial_t(\Psi_i v_\lambda)(z)| \lesssim \frac{1}{\kappa r_i^2} \left(\frac{r_i \lambda}{\varepsilon} + \frac{\varepsilon}{r_i \lambda} |v^i|^2 \right) + \frac{1}{\kappa r_i^2} \min\{r_i, \rho\} \lambda \approx \frac{1}{\kappa r_i^2} \left(\frac{r_i \lambda}{\varepsilon} + \frac{\varepsilon}{r_i \lambda} |v^i|^2 \right). \tag{4.44}$$

Making use of (4.44) in the expression for J_3 in (4.39), we get

$$J_3 \lesssim \frac{1}{\kappa r_i^2} \left(\frac{r_i \lambda}{\varepsilon} + \frac{\varepsilon}{r_i \lambda} |v^i|^2 \right) \iint_{\frac{3}{4}Q_i} |v - v^i| \chi_{Q_{\rho,s}(\mathfrak{z})} \, dy \, d\tau.$$

We can now proceed similarly to (4.19) to get

$$J_3 \lesssim \frac{1}{\kappa r_i^2} \left(\frac{r_i \lambda}{\varepsilon} + \frac{\varepsilon}{r_i \lambda} |v^i|^2 \right) r_i \lambda |Q_i| \lesssim \frac{\lambda^p}{\varepsilon} |4Q_i| + \varepsilon |4B_i| |v^i|^2. \tag{4.45}$$

Substituting the estimates (4.42), (4.43) and (4.45) into (4.38) gives the proof of (4.34). $\square$

We now come to essentially the most important estimate which will be needed to prove the difference estimate:

Lemma 4.18. *There exists a positive constant $C(n, p, q, \Lambda_0, \Lambda_1, m_e)$ such that the following estimate holds for every $t \in [-s, s]$:*

$$\int_{B_\rho(\mathfrak{x}) \setminus E_\lambda^t} (|v|^2 - |v - v_\lambda|^2)(x, t) \, dx \geq -C \lambda^p |\mathbb{R}^{n+1} \setminus E_\lambda|. \tag{4.46}$$

Proof. Let us fix any $t \in [\mathfrak{t} - s, \mathfrak{t} + s]$ and any point $x \in B_\rho(\mathfrak{x}) \setminus E_\lambda^t$. Now define

$$\Upsilon := \left\{ i \in \mathbb{N} : \operatorname{spt}(\Psi_i) \cap B_\rho(\mathfrak{x}) \times \{t\} \neq \emptyset \quad \text{and} \quad |v| + |v_\lambda| \neq 0 \quad \text{on } \operatorname{spt}(\Psi_i) \cap (B_\rho(\mathfrak{x}) \times \{t\}) \right\}.$$

Hence we only need to consider $i \in \Upsilon$. Note that $\sum_{i \in \Upsilon} \Psi_i(\cdot, t) \equiv 1$ on $\mathbb{R}^n \cap E_\lambda^t$, we can rewrite the left-hand side of (4.46) as

$$\begin{aligned}\int_{B_\rho(\mathfrak{x})\setminus E^t_\lambda}(|v|^2-|v-v_\lambda|^2)(x,t)\ dx &= \sum_{i\in\Upsilon}\int_{B_\rho(\mathfrak{x})}\Psi_i(|v|^2-|v-v_\lambda|^2)\ dx\\ &= \sum_{i\in\Upsilon}\int_{B_\rho(\mathfrak{x})}\Psi_i(z)\left(|v^i|^2+2v_\lambda(v-v^i)\right)\ dx-\sum_{i\in\Upsilon}\int_{B_\rho(\mathfrak{x})}\Psi_i(z)|v_\lambda-v^i|^2\ dx\\ &:= J_1+J_2.\end{aligned}$$

Estimate of J_1: Using (4.34), we get

$$J_1\gtrsim\sum_{i\in\Upsilon}\int_{B_\rho(\mathfrak{x})}\Psi_i(z)|v^i|^2\ dz-\varepsilon\sum_{i\in\Upsilon}|4B_i||v^i|^2-\sum_{i\in\Upsilon}\frac{\lambda^p}{\varepsilon}|4Q_i|. \tag{4.47}$$

From (4.6), we have $v^i=0$ whenever $\operatorname{spt}(\Psi_i)\cap B_\rho(\mathfrak{x})^c\neq\emptyset$. Hence we only have to sum over all those $i\in\Upsilon$ for which $\operatorname{spt}(\Psi_i)\subset B_\rho(\mathfrak{x})\times[\mathfrak{t}-s,\infty)$. In this case, we make use of a suitable choice for $\varepsilon\in(0,1]$, and use (**W7**) along with (**W8**), to estimate (4.47) from below to get

$$J_1\gtrsim-\lambda^p|\mathbb{R}^{n+1}\setminus E_\lambda|. \tag{4.48}$$

Estimate of J_2: For any $x\in B_\rho(\mathfrak{x})\setminus E^t_\lambda$, we have from (**W10**) that $\sum_j\Psi_j(x,t)=1$, which gives

$$\begin{aligned}\Psi_i(z)|v_\lambda(z)-v^i|^2 &\lesssim\Psi_i(z)\sum_{j\in A_i}|\Psi_j(z)|^2\left(v^j-v^i\right)^2\\ &\overset{(a)}{\lesssim}\min\{\rho,r_i\}^2\lambda^2.\end{aligned} \tag{4.49}$$

To obtain (a) above, we made use of Lemma 4.10 along with (**W13**). Substituting (4.49) into the expression for J_2 and using $|Q_i|=|B_i|\times2\kappa r_i^2$, we get

$$J_2\lesssim\sum_{i\in\Upsilon_1}|B_i|\,\frac{\kappa r_i^2}{\kappa}\lambda^2\lesssim\lambda^p|\mathbb{R}^{n+1}\setminus E_\lambda|. \tag{4.50}$$

Substituting (4.48) and (4.50) into (4.7), we get

$$J_2\gtrsim-\lambda^p|\mathbb{R}^{n+1}\setminus E_\lambda|.$$

This completes the proof of the lemma. □

5. Comparison estimates

Before we state the main difference estimates, let us define the the approximations that we will make and recall some useful results in existing literature. Let us fix the point

$$\mathfrak{z}=(\mathfrak{x},\mathfrak{t})\in\overline{\Omega}\times(-T,T).$$

5.1. *Approximations*

Let u be a weak solution of (1.1) and consider the unique weak solution $w\in C^0\left(I_{4\rho}(\mathfrak{t});L^2(\Omega_{4\rho}(\mathfrak{x})\right)\cap L^p\left(I_{4\rho}(\mathfrak{t});W^{1,p}(\Omega_{4\rho}(\mathfrak{x})\right)$ solving

$$\left\{\begin{array}{rcll} w_t-\operatorname{div}\mathcal{A}(x,t,\nabla w) &=& 0 & \text{in } K_{4\rho}(\mathfrak{z}),\\ w &=& u & \text{on } \partial_p K_{4\rho}(\mathfrak{z}).\end{array}\right. \tag{5.1}$$

This is possible, since (1.1) shows $u\in L^p\left(I_{4\rho}(\mathfrak{t});W^{1,p}(\Omega_{4\rho}(\mathfrak{x})\right)$ and $\dfrac{du}{dt}\in\left(W^{1,p}(K_{4\rho}(\mathfrak{z}))\right)'$ in the sense of distribution.

Recalling the notation from (2.5), we will need to make another approximation to (5.1):

$$\left\{\begin{array}{rcll} v_t-\operatorname{div}\overline{\mathcal{A}}_{B_{3\rho}(\mathfrak{x})}(\nabla v,t) &=& 0 & \text{in } K_{3\rho}(\mathfrak{z}),\\ v &=& w & \text{on } \partial_p K_{3\rho}(\mathfrak{z}),\end{array}\right. \tag{5.2}$$

which admits a unique weak solution $v\in C^0\left(I_{3\rho}(\mathfrak{t});L^2(\Omega_{3\rho}(\mathfrak{x})\right)\cap L^p\left(I_{3\rho}(\mathfrak{t});W^{1,p}(\Omega_{3\rho}(\mathfrak{x})\right)$ since Proposition 2.17 is applicable.

5.2. Interior Lipschitz regularity

In the case $K_{3\rho}(\mathfrak{z}) = Q_{3\rho}(\mathfrak{z})$, i.e., we are in the interior case, then we have the following interior Lipschitz regularity for (5.2) (see [19, Theorem 5.1 and Theorem 5.2]):

Lemma 5.1. *There exists a weak solution $v \in C^0\left(I_{3\rho}(\mathfrak{t}); L^2(B_{3\rho}(\mathfrak{x})\right) \cap L^p\left(I_{2\rho}(\mathfrak{t}); W^{1,p}(\Omega^+_{2\rho}(\mathfrak{x})\right)$ solving (5.2). Furthermore, there holds*

$$\sup_{Q_\rho(\mathfrak{z})} |\nabla v| \leq C_{(n,p,\Lambda_0,\Lambda_1)} \left(\fint\fint_{Q_{2\rho}(\mathfrak{z})} |\nabla v|^p \, dz \right)^{\frac{1}{p}}. \tag{5.3}$$

5.3. Boundary Lipschitz regularity

In the boundary case, we may not have Lipschitz regularity for solutions of (5.2) up to the boundary in general. In order to overcome this difficulty, we need to make one further approximation in which we consider a weak solution $\overline{V} \in C^0\left(I_{2\rho}(\mathfrak{t}); L^2(\Omega^+_{2\rho}(\mathfrak{x})\right) \cap L^p\left(I_{2\rho}(\mathfrak{t}); W^{1,p}(\Omega^+_{2\rho}(\mathfrak{x})\right)$ solving

$$\begin{cases} \overline{V}_t - \operatorname{div} \overline{\mathcal{A}}_{B_{3\rho}(\mathfrak{x})}(\nabla \overline{V}, t) &= 0 \quad \text{in } Q^+_{2\rho}(\mathfrak{z}), \\ \overline{V} &= 0 \quad \text{on } T_{2\rho}(\mathfrak{z}). \end{cases}$$

From [27, Theorem 1.6], the following important lemma holds:

Lemma 5.2. *There exists a weak solution $\overline{V} \in C^0\left(I_{2\rho}(\mathfrak{t}); L^2(\Omega^+_{2\rho}(\mathfrak{x})\right) \cap L^p\left(I_{2\rho}(\mathfrak{t}); W^{1,p}(\Omega^+_{2\rho}(\mathfrak{x})\right)$ solving (5.3). Furthermore, there holds*

$$\sup_{Q^+_\rho(\mathfrak{z})} |\nabla \overline{V}| \leq C(n,p,\Lambda_0,\Lambda_1) \left(\fint\fint_{Q^+_{2\rho}(\mathfrak{z})} |\nabla \overline{V}|^p \, dz \right)^{\frac{1}{p}}.$$

5.4. First comparison estimate

In this subsection, we will prove a improved difference estimate between solutions of (1.1) and (5.1).

Theorem 5.3. *Let $\delta > 0$ be given and Let u be a weak solution of (1.1) and w be the unique weak solution of (5.1), then there exists an $\beta_1 = \beta_1(\Lambda_0, \Lambda_1, p, n, m_e, \delta) \in (0,1)$ such that for any $\beta \in (0, \beta_1)$, the following estimate holds:*

$$\fint\fint_{K_{4\rho}(\mathfrak{z})} |\nabla u - \nabla w|^{p-\beta} \, dz \leq \delta \fint\fint_{K_{4\rho}(\mathfrak{z})} |\nabla u|^{p-\beta} \, dz + C(n,p,\beta,\Lambda_0,\Lambda_1,\delta) \fint\fint_{K_{4\rho}(\mathfrak{z})} |\mathbf{f}|^{p-\beta}.$$

Proof. Consider the following cut-off function $\zeta_\varepsilon \in C^\infty(\mathfrak{t} - (4\rho)^2, \infty)$ such that $0 \leq \zeta_\varepsilon(t) \leq 1$ and

$$\zeta_\varepsilon(t) = \begin{cases} 1 & \text{for } t \in (\mathfrak{t} - (4\rho)^2 + \varepsilon, \mathfrak{t} + (4\rho)^2 - \varepsilon), \\ 0 & \text{for } t \in (-\infty, \mathfrak{t} - (4\rho)^2) \cup (\mathfrak{t} + (4\rho)^2, \infty). \end{cases}$$

It is easy to see that

$$\begin{aligned} \zeta'_\varepsilon(t) = 0 \qquad &\text{for} \qquad t \in (-\infty, \mathfrak{t} - (4\rho)^2) \cup (\mathfrak{t} - (4\rho)^2 + \varepsilon, \mathfrak{t} + (4\rho)^2 - \varepsilon) \cup (\mathfrak{t} + (4\rho)^2, \infty), \\ |\zeta'_\varepsilon(t)| \leq \frac{c}{\varepsilon} \qquad &\text{for} \qquad t \in (\mathfrak{t} - (4\rho)^2, \mathfrak{t} - (4\rho)^2 + \varepsilon) \cup (\mathfrak{t} + (4\rho)^2 - \varepsilon, \mathfrak{t} - (4\rho)^2). \end{aligned}$$

Without loss of generality, we shall always take $2h \leq \varepsilon$, since we will take limits in the following order $\lim_{\varepsilon \to 0} \lim_{h \to 0}$.

Let us apply the results of Section 4 with $\varphi = u$, $\phi = w$, $\vec{f} = \mathbf{f}$ and $\vec{g} = 0$ over $Q_{\rho,s}(\mathfrak{z}) = K_{4\rho}(\mathfrak{z})$ to get a Lipschitz test function $v_{\lambda,h}$ satisfying Lemma 4.18. From Lemma 4.16, we have $v_{\lambda,h} \in C^{0,1}(K_{4\rho}(\mathfrak{z}))$ and thus we shall use $v_{\lambda,h}(z)\zeta_\varepsilon(t)$ as a test function to get

$$\begin{aligned} L_1 + L_2 := \iint_{K_{4\rho}(\mathfrak{z})} \frac{d[u-w]_h}{dt} v_{\lambda,h} \zeta_\varepsilon \, dx \, dt + \iint_{K_{4\rho}(\mathfrak{z})} \langle [\mathcal{A}(x,t,\nabla u) - \mathcal{A}(x,t,\nabla w)]_h \ , \nabla v_{\lambda,h} \rangle \zeta_\varepsilon \, dx \, dt \\ = \iint_{K_{4\rho}(\mathfrak{z})} \langle [|\mathbf{f}|^{p-2}\mathbf{f}]_h \ , \nabla v_{\lambda,h} \rangle \zeta_\varepsilon \, dx \, dt =: L_3. \end{aligned}$$

Let us recall from Section 4 the following: for a fixed $1 < q < p - 2\beta$, we have

$$g(z) := \mathcal{M}\left([|\nabla u - \nabla w|^q + |\nabla u|^q + |\nabla w|^q + |\mathbf{f}|^q] \chi_{K_{4\rho}(\mathfrak{z})} \right)^{\frac{1}{q}},$$

where $\mathcal{M}$ is as defined in (2.7) and $E_\lambda = \{z \in \mathbb{R}^{n+1} : g(z) \leq \lambda\}$. *Note that at this point, we have not really made any choice of β.*

From the strong Maximal function estimates (see [28, Lemma 7.9] for the proof), we have

$$\begin{aligned}\|g\|_{L^{p-\beta}(\mathbb{R}^{n+1})} &\lesssim \|\nabla u\|_{L^{p-\beta}(K_{4\rho}(\mathfrak{z}))} + \|\nabla w\|_{L^{p-\beta}(K_{4\rho}(\mathfrak{z}))} + \|\mathbf{f}\|_{L^{p-\beta}(K_{4\rho}(\mathfrak{z}))} + \|\nabla u - \nabla w\|_{L^{p-\beta}(K_{4\rho}(\mathfrak{z}))} \\ &\lesssim \|\nabla u\|_{L^{p-\beta}(K_{4\rho}(\mathfrak{z}))} + \|\mathbf{f}\|_{L^{p-\beta}(K_{4\rho}(\mathfrak{z}))} + \|\nabla u - \nabla w\|_{L^{p-\beta}(K_{4\rho}(\mathfrak{z}))}.\end{aligned} \tag{5.4}$$

Estimate for L_1:

$$\begin{aligned} L_1 &= \int_{\mathfrak{t}-(4\rho)^2}^{\mathfrak{t}+(4\rho)^2} \int_{\Omega_{4\rho}(\mathfrak{x})} \frac{dv_h(y,s)}{ds} v_{\lambda,h}(y,s)\zeta_\varepsilon(s)\ dy\ ds \\ &\quad + \int_{\mathfrak{t}-(4\rho)^2}^{\mathfrak{t}+(4\rho)^2} \int_{\Omega_{4\rho}(\mathfrak{x})} \frac{d\left(\left[(v_h)^2 - (v_{\lambda,h} - v_h)^2\right]\zeta_\varepsilon(s)\right)}{ds}\ dy\ ds \\ &\quad - \int_0^{\mathfrak{t}+(4\rho)^2} \int_{\Omega_{4\rho}(\mathfrak{x})} \frac{d\zeta_\varepsilon}{ds}\left(v_h^2 - (v_{\lambda,h} - v_h)^2\right)\ dy\ ds \\ &:= J_2 + J_1(\mathfrak{t}+(4\rho)^2) - J_1(\mathfrak{t}-(4\rho)^2) - J_3, \end{aligned} \tag{5.5}$$

where we have set

$$J_1(s) := \frac{1}{2}\int_{\Omega_{4\rho}(\mathfrak{x})} ((v_h)^2 - (v_{\lambda,h} - v_h)^2)(y,s)\zeta_\varepsilon(s)\ dy.$$

Note that $J_1(\mathfrak{t}-(4\rho)^2) = J_1(\mathfrak{t}+(4\rho)^2) = 0$ since $\zeta_\varepsilon(\mathfrak{t}-(4\rho)^2) = \zeta_\varepsilon(\mathfrak{t}+(4\rho)^2) = 0$.

Form Lemma 4.15 applied with $\vartheta = 1$, we have the bound

$$|J_2| \lesssim \iint_{K_{4\rho}(\mathfrak{z})\setminus E_\lambda} \left|\frac{dv_{\lambda,h}}{ds}(v_{\lambda,h} - v_h)\right|\ dy\ ds \lesssim \lambda^p|\mathbb{R}^{n+1}\setminus E_\lambda|. \tag{5.6}$$

Estimate for L_2: We split L_2 and make use of the fact that $v_{\lambda,h}(z) = v_h(z)$ for all $z \in E_\lambda \cap K_{4\rho}(\mathfrak{z})$.

$$\begin{aligned} L_2 &= \iint_{K_{4\rho}(\mathfrak{z})\cap E_\lambda} \langle[\mathcal{A}(x,t,\nabla u) - \mathcal{A}(x,t,\nabla w)]_h\ ,\nabla v_{\lambda,h}\rangle\zeta_\varepsilon\ dz \\ &\quad + \iint_{K_{4\rho}(\mathfrak{z})\setminus E_\lambda} \langle[\mathcal{A}(x,t,\nabla u) - \mathcal{A}(x,t,\nabla w)]_h\ ,\nabla v_{\lambda,h}\rangle\zeta_\varepsilon\ dz \\ &= \iint_{K_{4\rho}(\mathfrak{z})\cap E_\lambda} \langle[\mathcal{A}(x,t,\nabla u) - \mathcal{A}(x,t,\nabla w)]_h\ ,\nabla[u-w]_h\rangle\zeta_\varepsilon\ dz \\ &\quad + \iint_{K_{4\rho}(\mathfrak{z})\setminus E_\lambda} \langle[\mathcal{A}(x,t,\nabla u) - \mathcal{A}(x,t,\nabla w)]_h\ ,\nabla v_{\lambda,h}\rangle\zeta_\varepsilon\ dz \\ &= L_2^1 + L_2^2. \end{aligned} \tag{5.7}$$

Estimate for L_2^1: Using ellipticity, we get

$$\begin{aligned} L_2^1 &= \iint_{K_{4\rho}(\mathfrak{z})\cap E_\lambda} \langle[\mathcal{A}(x,t,\nabla u) - \mathcal{A}(x,t,\nabla w)]_h\ ,\nabla[u-w]_h\rangle\zeta_\varepsilon\ dz \\ &\gtrsim \iint_{K_{4\rho}(\mathfrak{z})\cap E_\lambda} \left[|\nabla u - w|^2\left(|\nabla u|^2 + |\nabla w|^2\right)\right]_h^{\frac{p-2}{2}}\zeta_\varepsilon\ dz. \end{aligned} \tag{5.8}$$

Estimate for L_2^2: Using the bound from Lemma 4.11, we get

$$\begin{aligned} L_2^2 &\lesssim \iint_{K_{4\rho}(\mathfrak{z})\setminus E_\lambda} |[\mathcal{A}(x,t,\nabla u) - \mathcal{A}(x,t,\nabla w)]_h|\,|\nabla v_{\lambda,h}|\ dz \\ &\lesssim \lambda \iint_{K_{4\rho}(\mathfrak{z})\setminus E_\lambda} |\nabla[u]_h|^{p-1} + |\nabla[w]_h|^{p-1}\ dz. \end{aligned} \tag{5.9}$$

Estimate for L_3: Analogous to estimate for L_2, we split L_3 into integrals over E_λ and E_λ^c followed by making use of (4.5) and Lemma 4.11 to get

$$L_3 \lesssim \iint_{K_{4\rho}(\mathfrak{z})\cap E_\lambda} [|\mathbf{f}|^{p-1}]_h|\nabla[u-w]_h|\ dz + \lambda\iint_{K_{4\rho}(\mathfrak{z})\setminus E_\lambda} |[\mathbf{f}]_h|^{p-1}\ dz. \tag{5.10}$$

Combining (5.6) into (5.5) followed by (5.8) and (5.9) into (5.7) and making use of (5.5), (5.6) and (5.10), we get

$$
\begin{aligned}
-\int_{\mathfrak{t}-(4\rho)^2}^{\mathfrak{t}+(4\rho)^2}\int_{\Omega_{4\rho}(\mathfrak{x})}\frac{d\zeta_\varepsilon}{ds}\left(v_h^2-(v_{\lambda,h}-v_h)^2\right)\,dy\,ds+\iint_{K_{4\rho}(\mathfrak{z})\cap E_\lambda}|\nabla[u-w]_h|^2\left(|\nabla[u]_h|^2+|\nabla[w]_h|^2\right)^{\frac{p-2}{2}}\zeta_\varepsilon\,dz&\\
\lesssim\iint_{K_{4\rho}(\mathfrak{z})\cap E_\lambda}[|\mathbf{f}|^{p-1}]_h|\nabla[u-w]_h|\,dz+\lambda\iint_{K_{4\rho}(\mathfrak{z})\setminus E_\lambda}|\nabla[u]_h|^{p-1}+|\nabla[w]_h|^{p-1}+|[\mathbf{f}]_h|^{p-1}\,dz&\\
+\lambda^p|\mathbb{R}^{n+1}\setminus E_\lambda|.&
\end{aligned}
\tag{5.11}
$$

In order to estimate $-\int_{\mathfrak{t}-(4\rho)^2}^{\mathfrak{t}+(4\rho)^2}\int_{\Omega_{4\rho}(\mathfrak{x})}\frac{d\zeta_\varepsilon}{ds}\left(v_h^2-(v_{\lambda,h}-v_h)^2\right)\,dy\,ds$, we take limits first in $h\searrow 0$ followed by $\varepsilon\searrow 0$ to get

$$
\begin{aligned}
-\int_{\mathfrak{t}-(4\rho)^2}^{\mathfrak{t}+(4\rho)^2}\int_{\Omega_{4\rho}(\mathfrak{x})}\frac{d\zeta_\varepsilon}{ds}\left(v_h^2-(v_{\lambda,h}-v_h)^2\right)\,dy\,ds\xrightarrow{\lim_{\varepsilon\searrow0}\lim_{h\searrow0}}\;&\int_{\Omega_{4\rho}(\mathfrak{x})}(v^2-(v_\lambda-v)^2)(x,\mathfrak{t}+(4\rho)^2)\,dx\\
&-\int_{\Omega_{4\rho}(\mathfrak{x})}(v^2-(v_\lambda-v)^2)(x,\mathfrak{t}-(4\rho)^2)\,dx.
\end{aligned}
\tag{5.12}
$$

For the second term on the right of (5.12), we observe that on E_λ, we have $v_\lambda=v$ and on E_λ^c and we also have $v_\lambda(\cdot,\mathfrak{t}-(4\rho)^2)=v(\cdot,\mathfrak{t}-(4\rho)^2)=0$. Thus, the second term vanishes because on E_λ, we can use the initial boundary condition and on E_λ^c, it is zero by construction. Thus we get

$$
-\int_{\mathfrak{t}-(4\rho)^2}^{\mathfrak{t}+(4\rho)^2}\int_{\Omega_{4\rho}(\mathfrak{x})}\frac{d\zeta_\varepsilon}{ds}\left(v_h^2-(v_{\lambda,h}-v_h)^2\right)\,dy\,ds\xrightarrow{\lim_{\varepsilon\searrow0}\lim_{h\searrow0}}\int_{\Omega_{4\rho}(\mathfrak{x})}(v^2-(v_\lambda-v)^2)(x,\mathfrak{t}+(4\rho)^2)\,dx.
\tag{5.13}
$$

Thus using (5.13) into (5.11) gives

$$
\begin{aligned}
\int_{\Omega_{4\rho}(\mathfrak{x})}(v^2-(v_\lambda-v)^2)(x,\mathfrak{t}+(4\rho)^2)\,dx+\iint_{K_{4\rho}(\mathfrak{z})\cap E_\lambda}|\nabla[u-w]_h|^2\left(|\nabla[u]_h|^2+|\nabla[w]_h|^2\right)^{\frac{p-2}{2}}\zeta_\varepsilon\,dz&\\
\lesssim\iint_{K_{4\rho}(\mathfrak{z})\cap E_\lambda}[|\mathbf{f}|^{p-1}]_h|\nabla[u-w]_h|\,dz+\lambda\iint_{K_{4\rho}(\mathfrak{z})\setminus E_\lambda}|\nabla[u]_h|^{p-1}+|\nabla[w]_h|^{p-1}+|[\mathbf{f}]_h|^{p-1}\,dz&\\
+\lambda^p|\mathbb{R}^{n+1}\setminus E_\lambda|.&
\end{aligned}
\tag{5.14}
$$

In fact, if we consider a cut-off function $\zeta_\varepsilon^{t_0}(\cdot)$ for some $t_0\in(\mathfrak{t}-(4\rho)^2,\mathfrak{t}+(4\rho)^2)$, where

$$
\zeta_\varepsilon^{t_0}(t)=\begin{cases}1 & \text{for } t\in(-t_0+\varepsilon,t_0-\varepsilon),\\ 0 & \text{for } t\in(-\infty,-t_0)\cup(t_0,\infty).\end{cases}
$$

we get the following analogue of (5.14)

$$
\begin{aligned}
\int_{\Omega_{4\rho}(\mathfrak{x})}(v^2-(v_\lambda-v)^2)(x,t_0)\,dx+\int_{-t_0}^{t_0}\int_{\Omega_{4\rho}(\mathfrak{x})\cap E_\lambda^t}|\nabla(u-w)|^2\left(|\nabla u|^2+|\nabla w|^2\right)^{\frac{p-2}{2}}\,dz&\\
\lesssim\iint_{K_{4\rho}(\mathfrak{z})\cap E_\lambda}|\mathbf{f}|^{p-1}|\nabla(u-w)|\,dz+\lambda\iint_{K_{4\rho}(\mathfrak{z})\setminus E_\lambda}|\nabla u|^{p-1}+|\nabla w|^{p-1}+|\mathbf{f}|^{p-1}\,dz&\\
+\lambda^p|\mathbb{R}^{n+1}\setminus E_\lambda|.&
\end{aligned}
\tag{5.15}
$$

Using Lemma 4.18, we get for any $t\in(\mathfrak{t}-(4\rho)^2,\mathfrak{t}+(4\rho)^2)$, the estimate

$$
\int_{\Omega_{4\rho}(\mathfrak{x})}|(v)^2-(v_\lambda-v)^2|(y,t)\,dy\;\gtrsim\int_{E_\lambda^t}|v(x,t)|^2\,dx-\lambda^p|\mathbb{R}^{n+1}\setminus E_\lambda|.
\tag{5.16}
$$

Since $\int_{E_\lambda^t}|v(x,t)|^2\,dx$ occurs on the left hand side and is positive, we can ignore this term. Thus combining (5.16) with (5.15), we get

$$
\begin{aligned}
\iint_{K_{4\rho}(\mathfrak{z})\cap E_\lambda}|\nabla(u-w)|^2\left(|\nabla u|^2+|\nabla w|^2\right)^{\frac{p-2}{2}}\,dx\,dt\;\lesssim\;&\iint_{K_{4\rho}(\mathfrak{z})\cap E_\lambda}|\mathbf{f}|^{p-1}|\nabla(u-w)|\,dz\\
&+\lambda\iint_{K_{4\rho}(\mathfrak{z})\setminus E_\lambda}|\nabla u|^{p-1}+|\nabla w|^{p-1}+|\mathbf{f}|^{p-1}\,dz\\
&+\lambda^p|\mathbb{R}^{n+1}\setminus E_\lambda|.
\end{aligned}
\tag{5.17}
$$

Let us now multiply (5.17) with $\lambda^{-1-\beta}$ and integrating over $(0,\infty)$ with respect to λ, we get

$$
K_1+K_2\lesssim K_3+K_4,
\tag{5.18}
$$

where we have set

$$K_1 := \int_0^\infty \lambda^{-1-\beta} \iint_{K_{4\rho}(\mathfrak{z})\cap E_\lambda} |\nabla(u-w)|^2 \left(|\nabla u|^2 + |\nabla u|^2\right)^{\frac{p-2}{2}} \ dz \ d\lambda,$$
$$K_2 := \int_0^\infty \lambda^{-1-\beta} \iint_{K_{4\rho}(\mathfrak{z})\cap E_\lambda} |\mathbf{f}|^{p-1}|\nabla(u-w)| \ dz \ d\lambda,$$
$$K_3 := \int_0^\infty \lambda^{-\beta} \iint_{K_{4\rho}(\mathfrak{z})\setminus E_\lambda} |\nabla u|^{p-1} + |\nabla w|^{p-1} + |\mathbf{f}|^{p-1} \ dz \ d\lambda,$$
$$K_4 := \int_0^\infty \lambda^{-1-\beta}\lambda^p |\mathbb{R}^{n+1}\setminus E_\lambda| \ d\lambda.$$

Estimate for K_1: Applying Fubini, we get

$$K_1 \gtrsim \frac{1}{\beta} \iint_{K_{4\rho}(\mathfrak{z})} g(z)^{-\beta} |\nabla(u-w)|^2 \left(|\nabla u|^2 + |\nabla u|^2\right)^{\frac{p-2}{2}} \ dz.$$

Using Young's inequality along with (2.2) and (5.4), we get for any $\epsilon_1 > 0$, the estimate

$$\begin{aligned} \iint_{K_{4\rho}(\mathfrak{z})} |\nabla u - \nabla w|^{p-\beta} \ dz \quad &\lesssim \quad C(\epsilon_1)\beta K_1 + \epsilon_1 \iint_{K_{4\rho}(\mathfrak{z})} |\nabla u - \nabla w|^{p-\beta} + |\nabla u|^{p-\beta} \ dz \\ &\quad + C_3(\epsilon_1) \iint_{K_{4\rho}(\mathfrak{z})} |\mathbf{f}|^{p-\beta}\, dz. \end{aligned} \tag{5.19}$$

Estimate for K_2: Again by Fubini, we get

$$K_2 = \frac{1}{\beta} \iint_{K_{4\rho}(\mathfrak{z})} g(z)^{-\beta} \langle |\mathbf{f}|^{p-2}\mathbf{f} \ , \nabla u - \nabla w\rangle \ dz.$$

From the definition of $g(z)$, we see that for $z \in K_{4\rho}(\mathfrak{z})$, we have $g(z) \geq |\nabla u - \nabla w|(z)$, which implies $g(z)^{-\beta} \leq |\nabla u - \nabla w|^{-\beta}(z)$. Now we apply Young's inequality, for any $\epsilon_2 > 0$, we get

$$\begin{aligned} K_2 \quad &\leq \frac{1}{\beta} \iint_{K_{4\rho}(\mathfrak{z})} |\nabla u - \nabla w|^{1-\beta} |\mathbf{f}|^{p-1} \ dz \\ &\lesssim \frac{C(\epsilon_2)}{\beta} \iint_{K_{4\rho}(\mathfrak{z})} |\mathbf{f}|^{p-\beta} \ dz + \frac{\epsilon_2}{\beta} \iint_{K_{4\rho}(\mathfrak{z})} |\nabla u - \nabla w|^{p-\beta} \ dz. \end{aligned} \tag{5.20}$$

Estimate for K_3: Again applying Fubini, we get

$$K_3 = \frac{1}{p-\beta} \iint_{K_{4\rho}(\mathfrak{z})} g(z)^{1-\beta} \left(|\nabla u|^{p-1} + |\nabla w|^{p-1} + |\mathbf{f}|^{p-1}\right) \ dz.$$

Applying Young's inequality followed by making use of (5.4), we get

$$K_3 \quad \lesssim \iint_{K_{4\rho}(\mathfrak{z})} |\nabla u - \nabla w|^{p-\beta} + |\nabla u|^{p-\beta} + |\mathbf{f}|^{p-\beta} \ dz. \tag{5.21}$$

Estimate for K_4: Applying the layer cake representation followed by using (5.4), we get

$$\begin{aligned} K_4 \quad &= \frac{1}{p-\beta} \iint_{\mathbb{R}^{n+1}} g(z)^{p-\beta} \ dz \\ &\lesssim \iint_{K_{4\rho}(\mathfrak{z})} |\nabla u - \nabla w|^{p-\beta} + |\nabla u|^{p-\beta} + |\mathbf{f}|^{p-\beta} \ dz. \end{aligned} \tag{5.22}$$

We now combine (5.19), (5.20), (5.21) and (5.22) into (5.18), we get

$$\begin{aligned} \iint_{K_{4\rho}(\mathfrak{z})} |\nabla u - \nabla w|^{p-\beta} \ dz \quad &\lesssim \quad [\epsilon_1 + C(\epsilon_1)(\epsilon_2 + \beta)] \iint_{K_{4\rho}(\mathfrak{z})} |\nabla u - \nabla w|^{p-\beta} \ dz + C(\epsilon_1, \epsilon_2, \beta) \iint_{K_{4\rho}(\mathfrak{z})} |\mathbf{f}|^{p-\beta}\, dz \\ &\quad + [\epsilon_1 + C(\epsilon_1)\beta] \iint_{K_{4\rho}(\mathfrak{z})} |\nabla u|^{p-\beta} \ dz. \end{aligned}$$

Choosing ϵ_1 small followed by ϵ_2 and β, for any $\delta > 0$, we get a $\beta_1 = \beta_1(n, p, \Lambda_0, \Lambda_1, \delta)$ such that for any $\beta \in (0, \beta_1)$, there holds

$$⨌_{K_{4\rho}(\mathfrak{z})} |\nabla u - \nabla w|^{p-\beta} \ dz \leq \delta ⨌_{K_{4\rho}(\mathfrak{z})} |\nabla u|^{p-\beta} \ dz + C(n, p, \beta, \Lambda_0, \Lambda_1, \delta) ⨌_{K_{4\rho}(\mathfrak{z})} |\mathbf{f}|^{p-\beta}.$$

This completes the proof of the theorem. □

5.5. Second comparison estimate

In this subsection, we will prove an improved comparison estimate between solutions of (5.1) and (5.2).

Theorem 5.4. *Let w be a weak solution of (5.1) and v be the unique weak solution of (5.2), then there exists an $\beta_2 = \beta_2(\Lambda_0, \Lambda_1, p, n, m_e, \varepsilon) \in (0,1)$ such that for any $\beta \in (0, \beta_2)$ and $\varepsilon > 0$, there exists a $C = C(n, \Lambda_0, \Lambda_1, p, \varepsilon) > 0$ and $\sigma_1 = \sigma_1(\Lambda_0, \Lambda_1, p)$ such that the following estimate holds:*

$$\left(\iint_{K_{3\rho}(\mathfrak{z})} |\nabla w - \nabla v|^{p-\beta}\, dz\right)^{\frac{p}{p-\beta}} \leq \varepsilon \left(\iint_{K_{4\rho}(\mathfrak{z})} |\nabla w|^{p-\beta}\, dz\right)^{1+\beta\tilde{\vartheta}_1} + C[\mathcal{A}]_{2,R_0}^{\sigma_1} \left(\iint_{K_{4\rho}(\mathfrak{z})} |\nabla w|^{p-\beta}\, dz\right)^{\frac{(1+\beta\tilde{\vartheta}_1)(1+\beta\tilde{\vartheta}_2)p}{p+\beta}}.$$

Here $\tilde{\vartheta}_1$ and $\tilde{\vartheta}_2$ are from Lemma 2.18 and Lemma 2.19.

Proof. Since we are in the setting of weak solutions, from [21, Lemma 2.8], we have the following estimate: for any $\varepsilon > 0$, there exists a $C = C(n, \Lambda_0, \Lambda_1, p, \varepsilon) > 0$ and $\sigma_1 = \sigma_1(\Lambda_0, \Lambda_1, p)$ such that

$$\iint_{K_{3\rho}(\mathfrak{z})} |\nabla w - \nabla v|^{p}\, dz \leq \varepsilon \iint_{K_{3\rho}(\mathfrak{z})} |\nabla w|^{p}\, dz + C[\mathcal{A}]_{2,R_0}^{\sigma_1} \left(\iint_{K_{3\rho}(\mathfrak{z})} |\nabla w|^{p+\beta}\, dz\right)^{\frac{p}{p+\beta}}. \tag{5.23}$$

Since w solves the homogeneous equation (5.1), we can control the right hand side of (5.23) by using Lemma 2.18 and Lemma 2.19. For $\beta_2 := \min\{\tilde{\beta}_1, \tilde{\beta}_2\}$, for any $\beta \in (0, \beta_2)$, there holds

$$\iint_{K_{3\rho}(\mathfrak{z})} |\nabla w|^{p+\beta}\, dz \lesssim \left(\iint_{K_{\frac{7}{2}\rho}(\mathfrak{z})} |\nabla w|^{p}\, dz\right)^{1+\beta\tilde{\vartheta}_1} \lesssim \left(\iint_{K_{4\rho}(\mathfrak{z})} |\nabla w|^{p-\beta}\, dz\right)^{(1+\beta\tilde{\vartheta}_2)(1+\beta\tilde{\vartheta}_1)}, \tag{5.24}$$

where $\tilde{\vartheta}_1$ and $\tilde{\vartheta}_2$ are from Lemma 2.18 and Lemma 2.19 respectively.

We now combine (5.24) and (5.23) to get

$$\iint_{K_{3\rho}(\mathfrak{z})} |\nabla w - \nabla v|^{p}\, dz \leq \varepsilon \left(\iint_{K_{4\rho}(\mathfrak{z})} |\nabla w|^{p-\beta}\, dz\right)^{1+\beta\tilde{\vartheta}_1} + C[\mathcal{A}]_{2,R_0}^{\sigma_1} \left(\iint_{K_{4\rho}(\mathfrak{z})} |\nabla w|^{p-\beta}\, dz\right)^{\frac{(1+\beta\tilde{\vartheta}_1)(1+\beta\tilde{\vartheta}_2)p}{p+\beta}}.$$

A simple application of Hölder's inequality now gives

$$\left(\iint_{K_{3\rho}(\mathfrak{z})} |\nabla w - \nabla v|^{p-\beta}\, dz\right)^{\frac{p}{p-\beta}} \leq \varepsilon \left(\iint_{K_{4\rho}(\mathfrak{z})} |\nabla w|^{p-\beta}\, dz\right)^{1+\beta\tilde{\vartheta}_1} + C[\mathcal{A}]_{2,R_0}^{\sigma_1} \left(\iint_{K_{4\rho}(\mathfrak{z})} |\nabla w|^{p-\beta}\, dz\right)^{\frac{(1+\beta\tilde{\vartheta}_1)(1+\beta\tilde{\vartheta}_2)p}{p+\beta}}.$$

□

5.6. Interior approximation estimate

In this subsection, we will prove the interior approximation lemma:

Lemma 5.5. *Let $\beta \in (0, \beta_0)$ for $\beta_0 = \min\{\beta_1, \beta_2\}$ where β_1 is from Theorem 5.3 and β_2 is from Theorem 5.4. For each $\varepsilon > 0$, there exists a $\delta > 0$ (possibly depending on ε) such that the following holds true: Assume u is a weak solution of (1.1) satisfying*

$$\iint_{K_{4\rho}(\mathfrak{z})} |\nabla u|^{p-\beta}\, dz \leq 1, \tag{5.25}$$

then under the condition

$$\iint_{K_{4\rho}(\mathfrak{z})} |\mathbf{f}|^{p-\beta}\, dz \leq \delta^{p-\beta}, \tag{5.26}$$

there exists a weak solution v to (5.2) satisfying

$$\|\nabla v\|_{L^\infty(K_{2\rho}(\mathfrak{z}))} \lesssim 1, \qquad \text{and} \qquad \iint_{K_{2\rho}(\mathfrak{z})} |\nabla u - \nabla v|^{p-\beta}\, dz \leq \varepsilon^{p-\beta}. \tag{5.27}$$

Proof. Let us prove each of the assertions of (5.27) as follows:

First estimate in (5.27): From Lemma 5.1, we have existence of a weak solution v solving (5.2) satisfying the estimate

$$\|\nabla v\|_{L^\infty(K_{2\rho}(\mathfrak{z}))}^p \leq C(n, p, \Lambda_0, \Lambda_1) \iint_{Q_{3\rho}(\mathfrak{z})} |\nabla v|^p\, dz.$$

We now estimate the right hand side as follows:

$$\begin{aligned}
\iint_{Q_{3\rho}(\mathfrak{z})} |\nabla v|^p \ dz &\overset{(a)}{\lesssim} \iint_{Q_{3\rho}(\mathfrak{z})} |\nabla v - \nabla w|^p \ dz + \iint_{Q_{3\rho}(\mathfrak{z})} |\nabla w|^p \ dz \\
&\overset{(b)}{\lesssim} (1+\varepsilon) \left(\iint_{K_{4\rho}(\mathfrak{z})} |\nabla w|^{p-\beta} \ dz \right)^{1+\beta\tilde{\vartheta}_1} + C[\mathcal{A}]_{2,S_0}^{\sigma_1} \left(\iint_{K_{4\rho}(\mathfrak{z})} |\nabla w|^{p-\beta} \ dz \right)^{\frac{(1+\beta\tilde{\vartheta}_1)(1+\beta\tilde{\vartheta}_2)p}{p+\beta}} \\
&\overset{(c)}{\lesssim} (1+\varepsilon) \left(\iint_{K_{4\rho}(\mathfrak{z})} |\nabla w - \nabla u|^{p-\beta} \ dz \right)^{1+\beta\tilde{\vartheta}_1} \\
&\quad + C[\mathcal{A}]_{2,S_0}^{\sigma_1} \left(\iint_{K_{4\rho}(\mathfrak{z})} |\nabla w - \nabla u|^{p-\beta} \ dz \right)^{\frac{(1+\beta\tilde{\vartheta}_1)(1+\beta\tilde{\vartheta}_2)p}{p+\beta}} \\
&\quad + (1+\varepsilon) \left(\iint_{K_{4\rho}(\mathfrak{z})} |\nabla u|^{p-\beta} \ dz \right)^{1+\beta\tilde{\vartheta}_1} + C[\mathcal{A}]_{2,S_0}^{\sigma_1} \left(\iint_{K_{4\rho}(\mathfrak{z})} |\nabla u|^{p-\beta} \ dz \right)^{\frac{(1+\beta\tilde{\vartheta}_1)(1+\beta\tilde{\vartheta}_2)p}{p+\beta}}.
\end{aligned}$$

In order to obtain (a), we made use of triangle inequality, to obtain (b), we made use of Theorem 5.4 along with Lemma 2.19 and finally to obtain (c), we applied triangle inequality.

We can control $\iint_{K_{4\rho}(\mathfrak{z})} |\nabla w - \nabla u|^{p-\beta} \ dz$ using Theorem 5.3 along with making use of (5.25) and (5.26) and observing that $[\mathcal{A}]_{2,S_0} \leq C(p, \Lambda_0, \Lambda_1)$, we get

$$\iint_{Q_{3\rho}(\mathfrak{z})} |\nabla v|^p \ dz \leq C(\Lambda_0, \Lambda_1, n, p, \delta).$$

This proves the first assertion of (5.27).

Second estimate in (5.27): Using triangle inequality, we get

$$\iint_{K_{2\rho}(\mathfrak{z})} |\nabla u - \nabla v|^{p-\beta} \ dz \lesssim \iint_{K_{2\rho}(\mathfrak{z})} |\nabla u - \nabla w|^{p-\beta} \ dz + \iint_{K_{2\rho}(\mathfrak{z})} |\nabla w - \nabla v|^{p-\beta} \ dz.$$

Each of the above terms can be controlled using Theorem 5.3 and Theorem 5.4 along with (5.25) followed by choosing δ sufficiently small (depending on ε) and γ sufficiently small such that $(\mathcal{A}, \Omega)$ is (γ, S_0)-vanishing to get the desired conclusion.

□

5.7. *Boundary approximation estimate*

In this subsection, we will prove the boundary approximation lemma:

Lemma 5.6. *Let $\beta \in (0, \beta_0)$ be fixed and let w be a weak solution of (5.1) satisfying*

$$\iint_{K_{3\rho}(\mathfrak{z})} |\nabla w|^{p-\beta} \ dz \leq 1,$$

then for any $\varepsilon > 0$, there exists a small $\gamma = \gamma(\Lambda_0, \Lambda_1, n, p, \varepsilon) > 0$ such that if $(\mathcal{A}, \Omega)$ is (γ, S_0) vanishing, then there exists a weak solution $\overline{V}$ of (5.3) whose zero extension to $Q_{2\rho}(\mathfrak{z})$ satisfies

$$\iint_{Q^+_{2\rho}(\mathfrak{z})} |\nabla \overline{V}|^p \ dz \leq 1, \qquad \textit{and} \qquad \iint_{K_{\rho}(\mathfrak{z})} |\nabla w - \nabla \overline{V}|^p \ dz \leq \varepsilon^p. \tag{5.28}$$

Proof. From Lemma 2.19, we see that the

$$\iint_{K_{3\rho}(\mathfrak{z})} |\nabla w|^{p-\beta} \ dz \leq 1 \quad \Rightarrow \quad \iint_{K_{3\rho}(\mathfrak{z})} |\nabla w|^p \ dz \lesssim 1.$$

Hence we can apply [13, Lemma 3.8] to get an $\gamma = \gamma(\Lambda_0, \Lambda_1, n, p, \varepsilon) > 0$ such that if $(\mathcal{A}, \Omega)$ is (γ, S_0) vanishing, then there exists a weak solution $\overline{V}$ of (5.3) whose zero extension to $Q_{2\rho}(\mathfrak{z})$ satisfies

$$\iint_{Q^+_{2\rho}(\mathfrak{z})} |\nabla \overline{V}|^p \ dz \leq 1, \qquad \text{and} \qquad \iint_{K_{\rho}(\mathfrak{z})} |\nabla w - \nabla \overline{V}|^p \ dz \leq \varepsilon^p.$$

This completes the proof of the lemma. □

Corollary 5.7. *Let $\beta \in (0, \beta_0)$ be fixed and let w be a weak solution of (5.1) satisfying*

$$\fint\fint_{K_{3\rho}(\mathfrak{z})} |\nabla w|^{p-\beta}\, dz \leq 1,$$

then for any $\varepsilon > 0$, there exists a small $\gamma = \gamma(\Lambda_0, \Lambda_1, n, p, \varepsilon) > 0$ such that if $(\mathcal{A}, \Omega)$ is (γ, S_0) vanishing, then there exists a weak solution $\overline{V}$ of (5.3) whose zero extension to $Q_{2\rho}(\mathfrak{z})$ satisfies

$$\|\nabla \overline{V}\|_{L^\infty(Q_\rho(\mathfrak{z}))} \leq C(n, p, \Lambda_0, \Lambda_1), \qquad \text{and} \qquad \fint\fint_{K_\rho(\mathfrak{z})} |\nabla w - \nabla \overline{V}|^{p-\beta}\, dz \leq \varepsilon^{p-\beta}.$$

Proof. All the hypothesis of Lemma 5.6 is satisfied. Thus the first conclusion follows directly by combining (5.28) along with Lemma 5.1 and the second conclusion follows by a simple application of Hölder's inequality to (5.28). □

6. Proof of Theorem 3.1.

Consider the following cut-off function $\zeta_\varepsilon \in C^\infty(-T, \infty)$ such that $0 \leq \zeta_\varepsilon(t) \leq 1$ and

$$\zeta_\varepsilon(t) = \begin{cases} 1 & \text{for } t \in (-T+\varepsilon, T-\varepsilon), \\ 0 & \text{for } t \in (-\infty, -T) \cup (T, \infty). \end{cases}$$

It is easy to see that

$$\begin{aligned} \zeta_\varepsilon'(t) = 0 \qquad & \text{for} \qquad t \in (-\infty, -T) \cup (-T+\varepsilon, T-\varepsilon) \cup (T, \infty), \\ |\zeta_\varepsilon'(t)| \leq \frac{c}{\varepsilon} \qquad & \text{for} \qquad t \in (-T, -T+\varepsilon) \cup (T-\varepsilon, -T). \end{aligned}$$

Without loss of generality, we shall always take $2h \leq \varepsilon$, since we will take limits in the following order $\lim_{\varepsilon\to 0} \lim_{h\to 0}$.

Since $u = 0$ on $\partial\Omega \times (-T, T)$, we can apply the results of Section 4 with $\varphi = u$, $\phi = 0$, $\vec{f} = \mathbf{f}$ and $\vec{g} = 0$ over $Q_{\rho,s} = \Omega \times (-T, T)$ to get a Lipschitz test function $v_{\lambda,h}$ satisfying Lemma 4.18. Thus we shall use $v_{\lambda,h}(z)\zeta_\varepsilon(t)$ as a test function in (1.1) to get

$$\iint_{\Omega_T} \frac{d[u]_h}{dt} v_{\lambda,h}\zeta_\varepsilon\, dx\, dt + \iint_{\Omega_T} \langle [\mathcal{A}(x,t,\nabla u)]_h\ , \nabla v_{\lambda,h}\rangle \zeta_\varepsilon\, dx\, dt = \iint_{\Omega_T} \langle [|\mathbf{f}|^{p-2}\mathbf{f}]_h\ , \nabla v_{\lambda,h}\rangle \zeta_\varepsilon\, dx\, dt,$$

which we write as $L_1 + L_2 = L_3$.

Let us recall from Section 4 the following: for a fixed $1 < q < p - 2\beta$, we have

$$g(z) := \mathcal{M}\left([|\nabla u|^q + |\mathbf{f}|^q]\, \chi_{\Omega_T}\right)^{\frac{1}{q}},$$

where $\mathcal{M}$ is as defined in (2.7) and $E_\lambda = \{z \in \mathbb{R}^{n+1} : g(z) \leq \lambda\}$. *Note that at this point in the proof, we have not really made any choice of β.*

From the strong Maximal function estimates (see [28, Lemma 7.9] for the proof), we have

$$\|g\|_{L^{p-\beta}(\mathbb{R}^{n+1})} \quad \leq \quad C(n) \left(\||\nabla u|\chi_{\Omega_T}\|_{L^{p-\beta}(\mathbb{R}^{n+1})} + \||\mathbf{f}|\chi_{\Omega_T}\|_{L^{p-\beta}(\mathbb{R}^{n+1})}\right). \tag{6.1}$$

Estimate for L_1:

$$\begin{aligned} L_1 \quad &= \int_{-T}^{T}\int_\Omega \frac{du_h(y,s)}{ds} v_{\lambda,h}(y,s)\zeta_\varepsilon(s)\, dy\, ds \\ &= \int_{-T}^{T}\int_{\Omega\setminus E_\lambda^s} \frac{dv_{\lambda,h}}{ds}(v_{\lambda,h} - u_h)\zeta_\varepsilon(s)\, dy\, ds + \int_{-T}^{T}\int_\Omega \frac{d\left(\left[(u_h)^2 - (v_{\lambda,h} - u_h)^2\right]\zeta_\varepsilon(s)\right)}{ds}\, dy\, ds \\ &\quad - \int_0^T\int_\Omega \frac{d\zeta_\varepsilon}{ds}\left(u_h^2 - (v_{\lambda,h} - u_h)^2\right)\, dy\, ds \\ &:= J_2 + J_1(T) - J_1(-T) - J_3, \end{aligned} \tag{6.2}$$

where we have set

$$J_1(s) := \frac{1}{2}\int_\Omega ((u_h)^2 - (v_{\lambda,h} - u_h)^2)(y,s)\zeta_\varepsilon(s)\, dy.$$

Note that $J_1(-T) = J_1(T) = 0$ since $\zeta_\varepsilon(-T) = \zeta_\varepsilon(T) = 0$.

Form Lemma 4.15 applied with $\vartheta = 1$, we have the bound

$$|J_2| \quad \lesssim \iint_{\Omega_T\setminus E_\lambda} \left|\frac{dv_{\lambda,h}}{ds}(v_{\lambda,h} - u_h)\right|\, dy\, ds \lesssim \lambda^p |\mathbb{R}^{n+1}\setminus E_\lambda|. \tag{6.3}$$

Estimate for L_2: We split L_2 and make use of the fact that $v_{\lambda,h}(z) = u_h(z)$ for all $z \in E_\lambda \cap \Omega_T$.

$$\begin{aligned} L_2 &= \iint_{\Omega_T \cap E_\lambda} \langle [\mathcal{A}(x,t,\nabla u)]_h \ , \nabla v_{\lambda,h} \rangle \zeta_\varepsilon \ dz + \iint_{\Omega_T \setminus E_\lambda} \langle [\mathcal{A}(x,t,\nabla u)]_h \ , \nabla v_{\lambda,h} \rangle \zeta_\varepsilon \ dz \\ &= \iint_{\Omega_T \cap E_\lambda} \langle [\mathcal{A}(x,t,\nabla u)]_h \ , \nabla [u]_h \rangle \zeta_\varepsilon \ dz + \iint_{\Omega_T \setminus E_\lambda} \langle [\mathcal{A}(x,t,\nabla u)]_h \ , \nabla v_{\lambda,h} \rangle \zeta_\varepsilon \ dz \\ &= L_2^1 + L_2^2. \end{aligned} \tag{6.4}$$

Estimate for L_2^1: Using ellipticity from (2.2), we get

$$L_2^1 \ \gtrsim \iint_{\Omega_T \cap E_\lambda} [|\nabla u|^p]_h \, \zeta_\varepsilon \ dz. \tag{6.5}$$

Estimate for L_2^2: Using the bound from Lemma 4.11, we get

$$L_2^2 \ \lesssim \lambda \iint_{\Omega_T \setminus E_\lambda} |\nabla [u]_h|^{p-1} \ dz. \tag{6.6}$$

Estimate for L_3: Analogous to estimate for L_2, we split L_3 into integrals over E_λ and E_λ^c followed by making use of (4.5) and Lemma 4.11 to get

$$L_3 \lesssim \iint_{\Omega_T \cap E_\lambda} [|\mathbf{f}|^{p-1}]_h |\nabla [u]_h| \ dz + \lambda \iint_{\Omega_T \setminus E_\lambda} |[\mathbf{f}]_h|^{p-1} \ dz. \tag{6.7}$$

Combining (6.3) into (6.2) followed by (6.5) and (6.6) into (6.4) and making use of (6.2), (6.3) and (6.7), we get

$$\begin{aligned} -\int_{-T}^{T} \int_\Omega \frac{d\zeta_\varepsilon}{ds} \left(u_h^2 - (v_{\lambda,h} - u_h)^2 \right) \ dy \ ds + \iint_{\Omega_T \cap E_\lambda} |\nabla [u]_h|^p \zeta_\varepsilon \ dz \ \ &\lesssim \ \ \iint_{\Omega_T \cap E_\lambda} [|\mathbf{f}|^{p-1}]_h |\nabla [u]_h| \ dz \\ &\quad + \lambda \iint_{\Omega_T \setminus E_\lambda} |[\mathbf{f}]_h|^{p-1} \ dz \\ &\quad + \lambda^p |\mathbb{R}^{n+1} \setminus E_\lambda|. \end{aligned} \tag{6.8}$$

In order to estimate $-\int_{-T}^{T} \int_\Omega \frac{d\zeta_\varepsilon}{ds} \left(u_h^2 - (v_{\lambda,h} - u_h)^2 \right) \ dy \ ds$, we take limits first in $h \searrow 0$ followed by $\varepsilon \searrow 0$ to get

$$\begin{aligned} -\int_{-T}^{T} \int_\Omega \frac{d\zeta_\varepsilon}{ds} \left(u_h^2 - (v_{\lambda,h} - u_h)^2 \right) \ dy \ ds \xrightarrow{\lim\limits_{\varepsilon \searrow 0} \lim\limits_{h \searrow 0}} \ \ &\int_\Omega (u^2 - (v_\lambda - u)^2)(x,T) \ dx \\ &- \int_\Omega (u^2 - (v_\lambda - u)^2)(x,-T) \ dx. \end{aligned} \tag{6.9}$$

For the second term on the right of (6.9), we observe that on E_λ, we have $v_\lambda = u$ and on E_λ^c and we also have $v_\lambda(\cdot,-T) = u(\cdot,-T) = 0$ using the initial condition. Thus, the second term on the right of (6.9) vanishes from which we get

$$-\int_{-T}^{T} \int_\Omega \frac{d\zeta_\varepsilon}{ds} \left(u_h^2 - (v_{\lambda,h} - u_h)^2 \right) \ dy \ ds \xrightarrow{\lim\limits_{\varepsilon \searrow 0} \lim\limits_{h \searrow 0}} \int_\Omega (u^2 - (v_\lambda - u)^2)(x,T) \ dx.$$

Thus using (5.13) into (6.8) gives

$$\begin{aligned} \int_\Omega (u^2 - (v_\lambda - u)^2)(x,T) \ dx + \iint_{\Omega_T \cap E_\lambda} |\nabla u|^p \ dz \ \ &\lesssim \ \ \iint_{\Omega_T \cap E_\lambda} |\mathbf{f}|^{p-1} |\nabla u| \ dz + \lambda \iint_{\Omega_T \setminus E_\lambda} |\mathbf{f}|^{p-1} \ dz \\ &\quad + \lambda^p |\mathbb{R}^{n+1} \setminus E_\lambda|. \end{aligned} \tag{6.10}$$

In fact, if we consider a cut-off function $\zeta_\varepsilon^{t_0}(\cdot)$ for some $t_0 \in (-T,T)$, where

$$\zeta_\varepsilon^{t_0}(t) = \begin{cases} 1 & \text{for } t \in (-t_0 + \varepsilon, t_0 - \varepsilon), \\ 0 & \text{for } t \in (-\infty, -t_0) \cup (t_0, \infty). \end{cases}$$

we get the following analogue of (6.10)

$$\begin{aligned} \int_\Omega (v^2 - (v_\lambda - v)^2)(x,t_0) \ dx + \int_{-t_0}^{t_0} \int_{\Omega \cap E_\lambda^t} |\nabla u)|^p \ dz \ \ &\lesssim \ \ \iint_{\Omega_T \cap E_\lambda} |\mathbf{f}|^{p-1} |\nabla u| \ dz + \lambda \iint_{\Omega_T \setminus E_\lambda} |\mathbf{f}|^{p-1} \ dz \\ &\quad + \lambda^p |\mathbb{R}^{n+1} \setminus E_\lambda|. \end{aligned} \tag{6.11}$$

Using Lemma 4.18, we get for any $t \in (-T,T)$, the estimate

$$\int_{\Omega} |(u)^2 - (v_\lambda - u)^2|(y,t)\, dy \quad \gtrsim \int_{E_\lambda^t} |u(x,t)|^2\, dx - \lambda^p |\mathbb{R}^{n+1} \setminus E_\lambda|. \tag{6.12}$$

Since $\int_{E_\lambda^t} |u(x,t)|^2\, dx$ occurs on the left hand side and is positive, we can ignore this term. Thus combining (6.12) with (6.11), we get

$$\iint_{\Omega_T \cap E_\lambda} |\nabla u|^p\, dz \quad \lesssim \iint_{\Omega_T \cap E_\lambda} |\mathbf{f}|^{p-1} |\nabla u|\, dz + \lambda \iint_{\Omega_T \setminus E_\lambda} |\mathbf{f}|^{p-1}\, dz + \lambda^p |\mathbb{R}^{n+1} \setminus E_\lambda|. \tag{6.13}$$

Let us now multiply (6.13) with $\lambda^{-1-\beta}$ and integrating over $(0,\infty)$ with respect to λ, we get

$$K_1 \lesssim K_2 + K_3 + K_4, \tag{6.14}$$

where we have set

$$\begin{aligned}
K_1 &:= \int_0^\infty \lambda^{-1-\beta} \iint_{\Omega_T \cap E_\lambda} |\nabla(u-w)|^2 \left(|\nabla u|^2 + |\nabla u|^2\right)^{\frac{p-2}{2}}\, dz\, d\lambda,\\
K_2 &:= \int_0^\infty \lambda^{-1-\beta} \iint_{\Omega_T \cap E_\lambda} |\mathbf{f}|^{p-1} |\nabla(u-w)|\, dz\, d\lambda,\\
K_3 &:= \int_0^\infty \lambda^{-\beta} \iint_{\Omega_T \setminus E_\lambda} |\nabla u|^{p-1} + |\nabla w|^{p-1} + |\mathbf{f}|^{p-1}\, dz\, d\lambda,\\
K_4 &:= \int_0^\infty \lambda^{-1-\beta} \lambda^p |\mathbb{R}^{n+1} \setminus E_\lambda|\, d\lambda.
\end{aligned}$$

Estimate for K_1: Applying Fubini, we get

$$K_1 \gtrsim \frac{1}{\beta} \iint_{\Omega_T} g(z)^{-\beta} |\nabla u|^p\, dz.$$

Using Young's inequality along with (6.1), we get for any $\epsilon_1 > 0$, the estimate

$$\iint_{\Omega_T} |\nabla u|^{p-\beta}\, dz \lesssim C(\epsilon_1)\beta K_1 + \epsilon_1 \iint_{\Omega_T} |\nabla u|^{p-\beta}\, dz + C_3(\epsilon_1) \iint_{\Omega_T} |\mathbf{f}|^{p-\beta}\, dz. \tag{6.15}$$

Estimate for K_2: Again by Fubini, we get

$$K_2 = \frac{1}{\beta} \iint_{\Omega_T} g(z)^{-\beta} |\mathbf{f}|^{p-1} |\nabla u|\, dz.$$

From the definition of $g(z)$, we see that for $z \in \Omega_T$, we have $g(z) \geq |\nabla u|(z)$ which implies $g(z)^{-\beta} \leq |\nabla u|^{-\beta}(z)$. Now we apply Young's inequality, for any $\epsilon_2 > 0$, we get

$$K_2 \lesssim \frac{C(\epsilon_2)}{\beta} \iint_{\Omega_T} |\mathbf{f}|^{p-\beta}\, dz + \frac{\epsilon_2}{\beta} \iint_{\Omega_T} |\nabla u|^{p-\beta}\, dz. \tag{6.16}$$

Estimate for K_3: Again applying Fubini, we get

$$K_3 = \frac{1}{p-\beta} \iint_{\Omega_T} g(z)^{1-\beta} |\mathbf{f}|^{p-1}\, dz \overset{(a)}{\lesssim} \iint_{\Omega_T} |\nabla u|^{p-\beta} + |\mathbf{f}|^{p-\beta}\, dz. \tag{6.17}$$

To obtain (a), we made use of Young's inequality followed by (6.1).

Estimate for K_4: Applying the layer cake representation followed by using (6.1), we get

$$K_4 = \frac{1}{p-\beta} \iint_{\mathbb{R}^{n+1}} g(z)^{p-\beta}\, dz \lesssim \iint_{\Omega_T} |\nabla u|^{p-\beta} + |\mathbf{f}|^{p-\beta}\, dz. \tag{6.18}$$

We now combine (6.15), (6.16), (6.17) and (6.18) into (6.14), we get

$$\begin{aligned}
\iint_{\Omega_T} |\nabla u|^{p-\beta}\, dz \quad &\lesssim \quad [\epsilon_1 + C(\epsilon_1)(\epsilon_2 + \beta)] \iint_{\Omega_T} |\nabla u|^{p-\beta}\, dz + C(\epsilon_1, \epsilon_2, \beta) \iint_{\Omega_T} |\mathbf{f}|^{p-\beta}\, dz\\
&\qquad + [\epsilon_1 + C(\epsilon_1)\beta] \iint_{\Omega_T} |\nabla u|^{p-\beta}\, dz.
\end{aligned}$$

Choosing ϵ_1 small followed by ϵ_2 and β, we get a $\beta_3 = \beta_3(n,p,\Lambda_0,\Lambda_1,\varepsilon)$ such that for any $\beta \in (0,\beta_3)$, there holds

$$\iint_{\Omega_T} |\nabla u|^{p-\beta}\, dz \lesssim_{(n,p,\beta,\Lambda_0,\Lambda_1)} \iint_{\Omega_T} |\mathbf{f}|^{p-\beta}\, dz.$$

This completes the proof of the theorem.

7. Covering arguments

Once we have the estimates in Section 5 and Section 6, the covering arguments can be proved in the standard way. We will only provide a brief sketch of the estimates.

Remark 7.1. *In this section and following section, let β_0 be that such that for all $2\beta \in (0, \beta_0]$, the results in Section 5 and Section 6 are applicable. We will now fix an β with $2\beta \leq \beta_0$.*

Let us define

$$\alpha_0^{\frac{p-\beta}{d}} := \fint\fint_{\Omega_T} \left[|\nabla u|^{p-\beta} + \left(\frac{|\mathbf{f}|}{\gamma} \right)^{p-\beta} \right] \, dx \, dt. \tag{7.1}$$

where d is defined to be

$$d := \begin{cases} \dfrac{p-\beta}{2-\beta} & \text{if } p \geq 2, \\ \dfrac{2(p-\beta)}{2p-2\beta+np-2n} & \text{if } \dfrac{2n}{n+2} < p < 2. \end{cases} \tag{7.2}$$

Furthermore, define $c_e := \left[\left(\frac{16}{7} \right)^n \frac{|\Omega_T|}{|B_1| S_0^{n+2}} \right]^{\frac{d}{p-\beta}}$, and let

$$\lambda \geq c_e \alpha_0. \tag{7.3}$$

We will also need to consider the following superlevel set:

$$E_\lambda := \{ z \in \Omega_T : |\nabla u(z)| > \lambda \}.$$

The first lemma that we need is the following:

Lemma 7.2. *Let $\gamma \in (0,1)$ be any constant, then for any λ satisfying (7.3), there exists a family of disjoint cylinders $\{K^\lambda_{r_i}(z_i)\}_{i \in \mathbb{N}}$ with $z_i \in E_\lambda$ and $r_i \in (0, S_0)$ such that*

$$\fint\fint_{K^\lambda_{r_i}(z_i)} \left[|\nabla u|^{p-\beta} + \left(\frac{|\mathbf{f}|}{\gamma} \right)^{p-\beta} \right] \, dx \, dt = \lambda^{p-\beta},$$

$$\fint\fint_{K^\lambda_{r_i}(z_i)} \left[|\nabla u|^{p-\beta} + \left(\frac{|\mathbf{f}|}{\gamma} \right)^{p-\beta} \right] \, dx \, dt < \lambda^{p-\beta} \qquad \text{for every } r > r_i,$$

$$E_\lambda \subset \bigcup_{i \in \mathbb{N}} K^\lambda_{5r_i}(z_i).$$

The proof following using standard techniques from Measure theory and Lemma 2.3 (see [13, Pages 4311-4313] for the details).

Using Lemma 2.7 the following lemma follows:

Lemma 7.3. *Let $w \in A_s$ for any $s \geq \dfrac{p-\beta}{p-2\beta}$, then it is automatically in A_1. Then there exists a constant $c^* = c^*([w]_1, n, p) > 0)$ such that there holds*

$$w(K^\lambda_{r_i}(z_i)) \leq \frac{C}{\lambda^{p-\beta}} \left[\iint_{K^\lambda_{r_i}(z_i) \cap \{|\nabla u| > \frac{\lambda}{4c^*}\}} |\nabla u|^{p-\beta} w(z) \, dz + \iint_{K^\lambda_{r_i}(z_i) \cap \{|\mathbf{f} u| > \frac{\gamma\lambda}{4c^*}\}} \left| \frac{\mathbf{f}}{\gamma} \right|^{p-\beta} w(z) \, dz \right].$$

The proof of the above Lemma is standard and we refer to [16, Page 4114 - (3.8)] for the necessary details.

Now making use of the a priori estimates in Section 5 and Section 6 and combining with the techniques of [13, Lemma 4.3], the following Lemma holds:

Lemma 7.4. *There exists a constant $N = N(\Lambda_0, \Lambda_1, n, p) > 1$ such that for any $\varepsilon \in (0,1)$, there exists a small $\gamma = \gamma(\Lambda_0, \Lambda_1, \varepsilon, n, p)$ such that if $(\mathcal{A}, \Omega)$ is (γ, S_0) vanishing for such a small γ and some fixed $S_0 > 0$, then there holds*

$$\frac{\left| \{ z \in K^\lambda_{5r_i}(z_i) : |\nabla u(z)| > 2N\lambda \} \right|}{|K^\lambda_{r_i}(z_i)|} \leq c_{(\Lambda_0, \Lambda_1, n, p)} \varepsilon^{p-\beta}.$$

We can now combine Lemma 7.3 and Lemma 7.4 to prove the following weighted estimate on the level sets (again the proof follows exactly as in [16, STEP 4 on Page 4115] and will be omitted).

Lemma 7.5. *Let β be as in Remark 7.1, then there holds*

$$w(E(2N\lambda)) \lesssim \frac{\varepsilon^{(p-2\beta)\tau_1}}{\lambda^{p-\beta}} \left[\iint\limits_{\Omega_T \cap \{|\nabla u| > \frac{\lambda}{4c^*}\}} |\nabla u|^{p-\beta} w(z)\ dz + \iint\limits_{\Omega_T \cap \{|\mathbf{f}u| > \frac{\gamma\lambda}{4c^*}\}} \left|\frac{\mathbf{f}}{\gamma}\right|^{p-\beta} w(z)\ dz \right].$$

Here τ_1 is defined as in Definition 2.10.

8. Proof of Theorem 3.3.

From Lemma 2.12, we get

$$\begin{aligned}\iint_{\Omega_T} |\nabla u|^q w(z)\ dz &= q\int_0^{c_e\alpha_0} (2N\lambda)^{q-1} w(\{z\in\Omega_T : |\nabla u| > 2N\lambda\}) d(2N\lambda) \\ &\quad + q\int_{c_e\alpha_0}^{\infty} (2N\lambda)^{q-1} w(\{z\in\Omega_T : |\nabla u| > 2N\lambda\}) d(2N\lambda) \\ &=: II_1 + II_2.\end{aligned} \tag{8.1}$$

Estimate for II_1: With d as defined in (7.2), we get

$$\begin{aligned} II_1 &\lesssim w(\Omega_T)\alpha_0^q \overset{(7.1)}{=} w(\Omega_T)\left(\fint\!\!\fint_{\Omega_T} \left[|\nabla u|^{p-\beta} + \left(\frac{|\mathbf{f}|}{\gamma}\right)^{p-\beta}\right] dx\ dt\right)^{\frac{qd}{p-\beta}} \\ &\overset{\text{Theorem 3.1}}{\lesssim} w(\Omega_T)\left(\fint\!\!\fint_{\Omega_T} |\mathbf{f}|^{p-\beta}\ dx\ dt\right)^{\frac{qd}{p-\beta}} \\ &\overset{\text{Lemma 2.7}}{\lesssim} w(\Omega_T)\left(\frac{1}{w(\Omega_T)}\iint_{\Omega_T} |\mathbf{f}|^q w(x,t)\ dx\ dt\right)^d \\ &\lesssim \left(\iint_{\Omega_T} |\mathbf{f}|^q w(x,t)\ dx\ dt\right)^d. \end{aligned} \tag{8.2}$$

Estimate for II_2: Since we have $q \geq p > p-\beta > p-2\beta$, we proceed as follows:

$$\begin{aligned} II_2 &\overset{\text{Lemma 7.5}}{\lesssim} \varepsilon^{(p-2\beta)\tau_1}\int_1^\infty \frac{\lambda^{q-1}}{\lambda^{p-\beta}} \iint\limits_{\Omega_T\cap\{|\nabla u| > \frac{\lambda}{4c^*}\}} |\nabla u|^{p-\beta} w(z)\ dz\ d\lambda \\ &\quad + \frac{\varepsilon^{(p-2\beta)\tau_1}}{\gamma^{(p-2\beta)\tau_1}}\int_1^\infty \frac{\lambda^{q-1}}{\lambda^{p-\beta}} \iint\limits_{\Omega_T\cap\{|\mathbf{f}| > \frac{\gamma\lambda}{4c^*}\}} |\mathbf{f}|^{p-\beta} w(z)\ dz\ d\lambda \\ &\overset{\text{Lemma 2.12}}{\lesssim} \varepsilon^{p-2\beta}\iint_{\Omega_T} |\nabla u|^q w(z)\ dz + C(\gamma,\varepsilon)\iint_{\Omega_T} |\mathbf{f}|^q w(z)\ dz. \end{aligned} \tag{8.3}$$

Combining (8.2) and (8.3) into (8.1) following by choosing ε sufficiently small, the proof follows.

References

[1] Emilio Acerbi and Giuseppe Mingione. Gradient estimates for a class of parabolic systems. *Duke Mathematical Journal*, 136(2):285–320, 2007.

[2] David R Adams and Lars I Hedberg. *Function Spaces and Potential Theory*, volume 314. Grundlehren der Mathematischen Wissenschaften, Springer, Berlin, 1996.

[3] Karthik Adimurthi and Sun-Sig Byun. Boundary higher intebrality for very weak solutions of quasilinear parabolic equations. *Journal de Mathmatiques Pures et Appliques*, 2017. To appear.

[4] Karthik Adimurthi and Nguyen Cong Phuc. Global Lorentz and Lorentz-Morrey estimates below the natural exponent for quasilinear equations. *Calculus of Variations and Partial Differential Equations*, 54(3):3107–3139, 2015.

[5] Karthik Adimurthi and Nguyen Cong Phuc. An end-point global gradient weighted estimate for quasilinear equations in non-smooth domains. *Manuscripta Mathematica*, 150(1-2):111–135, 2016.

[6] Paolo Baroni. Lorentz estimates for degenerate and singular evolutionary systems. *J. Differential Equations*, 255(9):2927–2951, 2013.

[7] Paolo Baroni. Lorentz estimates for obstacle parabolic problems. *Nonlinear Anal.*, 96:167–188, 2014.

[8] Verena Bögelein. *Regularity results for weak and very weak solutions of higher order parabolic systems*. PhD thesis, 2007.

[9] Verena Bögelein, Frank Duzaar, and Giuseppe Mingione. *The regularity of general parabolic systems with degenerate diffusion*, volume 221. American Mathematical Society, 2013.

[10] The Anh Bui and Xuan Thinh Duong. Global Lorentz estimates for nonlinear parabolic equations on nonsmooth domains. *arXiv preprint arXiv:1702.06202*, 2017.

[11] Sun-Sig Byun and Jihoon Ok. On $W^{1,q(\cdot)}$-estimates for elliptic equations of $p(x)$-laplacian type. *Journal de Mathématiques Pures et Appliquées*, 106(3):512–545, 2016.

[12] Sun-Sig Byun, Jihoon Ok, Dian K. Palagachev, and Lubomira G. Softova. Parabolic systems with measurable coefficients in weighted Orlicz spaces. *Commun. Contemp. Math.*, 18(2):1550018, 19, 2016.

[13] Sun-Sig Byun, Jihoon Ok, and Seungjin Ryu. Global gradient estimates for general nonlinear parabolic equations in nonsmooth domains. *Journal of Differential Equations*, 254(11):4290–4326, 2013.

[14] Sun-Sig Byun, Jihoon Ok, and Seungjin Ryu. Global gradient estimates for elliptic equations of $p(x)$-laplacian type with BMO nonlinearity. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2016(715):1–38, 2016.

[15] Sun-Sig Byun, Dian K Palagachev, and Seungjin Ryu. Weighted $W^{1,p}$ estimates for solutions of non-linear parabolic equations over non-smooth domains. *Bulletin of the London Mathematical Society*, 45(4):765–778, 2013.

[16] Sun-Sig Byun and Seungjin Ryu. Weighted orlicz estimates for general nonlinear parabolic equations over nonsmooth domains. *Journal of Functional Analysis*, 272(10):4103–4121, 2017.

[17] Sun-Sig Byun and Lihe Wang. Elliptic equations with BMO coefficients in Reifenberg domains. *Communications on pure and applied mathematics*, 57(10):1283–1310, 2004.

[18] Guiseppe Da Prato. $\mathcal{L}^{(p,\theta)}(\omega,\delta)$ e loro proprietá. *Annali di Matematica Pura ed Applicata*, 69(1):383–392, 1965.

[19] Emmanuele DiBenedetto. *Degenerate Parabolic Equations*. Universitext, Springer-Verlag, New York, 1993.

[20] Lars Diening, Michael Růžička, and Jörg Wolf. Existence of weak solutions for unsteady motions of generalized Newtonian fluids. 2010.

[21] Xuan Thinh Duong and The Anh Bui. Global lorentz estimates for nonlinear parabolic equations on nonsmooth domains. *Calculus of Variations and Partial Differential Equations*, 56(2):47, 2017.

[22] Frank Duzaar, Giuseppe Mingione, and Klaus Steffen. Parabolic systems with polynomial growth and regularity. *Mem. Amer. Math. Soc.*, 214(1005):x+118, 2011.

[23] Przemysław Górka. Campanato theorem on metric measure spaces. *Ann. Acad. Sci. Fenn. Math*, 34(2):523–528, 2009. http://www.acadsci.fi/mathematica/Vol34/vol34pp523-528.pdf.

[24] Loukas Grafakos. *Classical and modern Fourier analysis*. Prentice Hall, 2004.

[25] Juha Kinnunen and John L Lewis. Higher integrability for parabolic systems of p-laplacian type. *Duke Mathematical Journal*, 102(2):253–272, 2000.

[26] Juha Kinnunen and John L Lewis. Very weak solutions of parabolic systems of p-laplacian type. *Arkiv för Matematik*, 40(1):105–132, 2002.

[27] Gary M Lieberman. Boundary and initial regularity for solutions of degenerate parabolic equations. *Nonlinear Analysis: Theory, Methods & Applications*, 20(5):551–569, 1993.

[28] Gary M Lieberman. *Second order parabolic differential equations*. World scientific, 1996.

[29] Pasi Mikkonen. On the Wolff potential and quasilinear elliptic equations involving measures. *Ann. Acad. Sci. Fenn. Math. Diss*, 104, 1996.

[30] Dian K. Palagachev and Lubomira G. Softova. Quasilinear divergence form parabolic equations in Reifenberg flat domains. *Discrete Contin. Dyn. Syst.*, 31(4):1397–1410, 2011.

[31] Mikko Parviainen. Reverse Hölder inequalities for singular parabolic equations near the boundary. *Journal of Differential Equations*, 246(2):512–540, 2009.

[32] Olli Saari. *Weights arising from parabolic partial differential equations*. PhD thesis, 2016. https://aaltodoc.aalto.fi/bitstream/handle/123456789/19814/isbn9789526066783.pdf.

[33] Ralph Edwin Showalter. *Monotone operators in Banach space and nonlinear partial differential equations*, volume 49. American Mathematical Soc., 2013.