# Weighted $L^{p}$-type regularity estimates for nonlinear parabolic equations with Orlicz growth 

Fengping Yao ${ }^{\boxtimes}$<br>Department of Mathematics, Shanghai University, Shanghai 200444, China

Received 9 October 2021, appeared 27 April 2022
Communicated by Roberto Livrea

Abstract. In this paper we obtain the following weighted $L^{p}$-type regularity estimates

$$
B(|\mathbf{f}|) \in L^{q}\left(v, v+T ; L_{w}^{q}(\Omega)\right) \text { locally } \Rightarrow B(|\nabla u|) \in L^{q}\left(v, v+T ; L_{w}^{q}(\Omega)\right) \text { locally }
$$

for any $q>1$ of weak solutions for non-homogeneous nonlinear parabolic equations with Orlicz growth

$$
u_{t}-\operatorname{div}\left(a\left((A \nabla u \cdot \nabla u)^{\frac{1}{2}}\right) A \nabla u\right)=\operatorname{div}(a(|\mathbf{f}|) \mathbf{f})
$$

under some proper assumptions on the functions $a, w, A$ and $\mathbf{f}$, where $B(t)=$ $\int_{0}^{t} \tau a(\tau) d \tau$ for $t \geq 0$. Moreover, we remark that two natural examples of functions $a(t)$ are

$$
a(t)=t^{p-2} \quad\left(p \text {-Laplace equation) } \quad \text { and } \quad a(t)=t^{p-2} \log ^{\alpha}(1+t) \quad \text { for } \alpha>0\right.
$$

Moreover, our results improve the known results for such equations.
Keywords: weighted, $L^{p}$-type, regularity, gradient, quasilinear, parabolic, nonhomogeneous.
2020 Mathematics Subject Classification: 35K55, 35K65.

## 1 Introduction

This paper is concerned with the local weighted $L^{p}$-type gradient estimates for weak solutions of the following non-homogeneous nonlinear parabolic equations with Orlicz growth

$$
\begin{equation*}
u_{t}-\operatorname{div}\left(a\left((A \nabla u \cdot \nabla u)^{\frac{1}{2}}\right) A \nabla u\right)=\operatorname{div}(a(|\mathbf{f}|) \mathbf{f}) \quad \text { in } \Omega_{T}:=\Omega \times(v, v+T) \tag{1.1}
\end{equation*}
$$

where $v \in \mathbb{R}, \Omega$ is an open bounded domain in $\mathbb{R}^{n}$, the vector valued function $\mathbf{f}=\left(f_{1}, \ldots, f_{n}\right)$ and $a:(0,+\infty) \rightarrow(0,+\infty) \in C^{1}(0,+\infty)$ satisfies

$$
\begin{equation*}
0 \leq i_{a}:=\inf _{t>0} \frac{t a^{\prime}(t)}{a(t)} \leq \sup _{t>0} \frac{t a^{\prime}(t)}{a(t)}=: s_{a}<+\infty \tag{1.2}
\end{equation*}
$$

[^0]Here, $A(x, t)=\left\{a_{i j}(x, t)\right\}_{n \times n}$ is a symmetric matrix with measurable coefficients satisfying the uniformly parabolic condition

$$
\begin{equation*}
\Lambda^{-1}|\xi|^{2} \leq A(x, t) \xi \cdot \xi \leq \Lambda|\xi|^{2} \tag{1.3}
\end{equation*}
$$

for all $\xi \in \mathbb{R}^{n}$, almost every $(x, t) \in \mathbb{R}^{n} \times \mathbb{R}$ and some positive constant $\Lambda$. Especially when $a(t)=t^{p-2}$, then $2+i_{a}=2+s_{a}=p$ and (1.1) is reduced to the classical parabolic $p$-Laplace equations

$$
\begin{equation*}
u_{t}-\operatorname{div}\left((A \nabla u \cdot \nabla u)^{\frac{p-2}{2}} A \nabla u\right)=\operatorname{div}\left(|\mathbf{f}|^{p-2} \mathbf{f}\right) . \tag{1.4}
\end{equation*}
$$

Define

$$
\begin{equation*}
g(t)=t a(t) \quad \text { and } \quad B(t)=\int_{0}^{t} g(\tau) d \tau=\int_{0}^{t} \tau a(\tau) d \tau \quad \text { for } t \geq 0 \tag{1.5}
\end{equation*}
$$

Then (1.2) implies that

$$
\begin{equation*}
g(t) \text { is strictly increasing and continuous over }[0,+\infty), \tag{1.6}
\end{equation*}
$$

and

$$
\begin{equation*}
B(t) \text { is increasing over }[0,+\infty) \text { and strictly convex with } B(0)=0 \text {. } \tag{1.7}
\end{equation*}
$$

There are two simple examples satisfying the given condition (1.2)

$$
\begin{equation*}
a(t)=t^{p-2} \quad \text { and } \quad a(t)=t^{p-2} \log ^{\alpha}(1+t) \quad \text { for any } p \geq 2 \text { and any } \alpha>0 . \tag{1.8}
\end{equation*}
$$

Additionally, another general and interesting example satisfying (1.2) is related to ( $p, q$ )growth condition which is given by appropriate gluing of the monomials (see page 600 in [7] and page 314 in [37]).

Different from the elliptic $p$-Laplace equation

$$
\operatorname{div}\left((A \nabla u \cdot \nabla u)^{\frac{p-2}{2}} A \nabla u\right)=\operatorname{div}\left(|\mathbf{f}|^{p-2} \mathbf{f}\right) \quad \text { in } \Omega,
$$

the solution in the $p$-parabolic setting (1.4) is no longer invariant under multiplication by a constant, which is one of the most difficulties (see [12]). More precisely, it is slightly difficult to use maximal operators, which are typically used in the elliptic cases (see [22]). First of all, Kinnunen and Lewis [36] established the following Gehring's reverse Hölder inequality

$$
\nabla u \in L_{l o c}^{p+\epsilon}\left(\Omega_{T}\right) \quad \text { for some small } \epsilon>0
$$

for weak solutions of (1.4), which implies the local higher integrability of the gradient. In their article they overcome the difficulties in using normalization and scaling methods by choosing the irregular cylinders whose side lengths depend on the function. Meanwhile, Misawa [42] obtained $L^{q}(q \geq p)$ estimates for gradients for evolutional $p$-Laplacian equations/systems (1.4) with discontinuous coefficients and external force given by the divergence of BMO-functions. Subsequently, Acerbi \& Mingione [1] invented a new covering/iteration argument to prove the sharp local $L^{q}(q \geq p)$ estimates

$$
|\mathbf{f}|^{p} \in L_{l o c}^{q}\left(\Omega_{T}\right) \Rightarrow|\nabla u|^{p} \in L_{l o c}^{q}\left(\Omega_{T}\right) \text { for any } q \geq 1
$$

with

$$
\begin{equation*}
\int_{Q_{r}}|\nabla u|^{p q} d z \leq C\left[\left(\int_{Q_{2 r}}|\nabla u|^{p} d z\right)^{q}+\int_{Q_{2 r}}|\mathbf{f}|^{p q}+1 d z\right]^{p / 2}, \tag{1.9}
\end{equation*}
$$

where the parabolic cylinder $Q_{2 r}=B_{2 r} \times\left(-4 r^{2}, 4 r^{2}\right] \subset \Omega_{T}$, for weak solutions of the parabolic $p$-Laplace equations/systems (1.4) with small BMO coefficients. Furthermore, Byun, Ok \& Ryu [17] obtained the global $L^{q}(q \geq p)$ estimates

$$
\mathbf{f} \in L^{q}\left(\Omega_{T}\right) \Rightarrow \nabla u \in L^{q}\left(\Omega_{T}\right) \quad \text { for any } q \geq p
$$

for weak solutions of the following general parabolic $p$-Laplace equations

$$
\begin{equation*}
u_{t}-\operatorname{div} \mathrm{a}(D u, x, t)=\operatorname{div}\left(|\mathbf{f}|^{p-2} \mathbf{f}\right) . \tag{1.10}
\end{equation*}
$$

The corresponding Hölder estimates for (1.4) and (1.10) can be found in the book [29]. On the other hand, the corresponding $L^{p}$-type estimates and Hölder estimates for the parabolic $p(x, t)$-Laplacian equations are also developed by [6,10]. In addition, some authors [5,13] have researched the Calderón-Zygmund estimates in the setting of Lorentz spaces

$$
\begin{equation*}
|\mathbf{f}|^{p} \in L^{\gamma, q} \Rightarrow|\nabla u|^{p} \in L^{\gamma, q} \quad \text { for any } 1<\gamma<\infty \text { and } 0<q \leq \infty \tag{1.11}
\end{equation*}
$$

for degenerate parabolic equations/systems (1.4). Most recently, there are also many research results $[9,11,16,19,20,25,33,35]$ on the study of various kinds of regularity estimates for weak solutions of the parabolic equations of $p$-Laplacian type.

The following non-homogeneous nonlinear elliptic equations with Orlicz growth which is first introduced by Lieberman [37]

$$
\begin{equation*}
\operatorname{div}(a(|\nabla u|) \nabla u)=f \quad \text { in } \Omega \tag{1.12}
\end{equation*}
$$

can be seen as the most natural generalization of the elliptic $p$-Laplace equations. Afterward, Cianchi \& Maz'ya [26-28] have investigated the global Lipschitz regularity and sharp estimates for (1.12) with the condition (1.2). Moreover, Diening, Stroffolini \& Verde [32] obtained the $\phi$-harmonic approximation lemma for the gradient of solutions to (1.12) with the condition (1.2) and $f=0$. Lately, we [48] established the following local $L^{q}$ estimates

$$
B(|\mathbf{f}|) \in L_{l o c}^{q} \Rightarrow B(|\nabla u|) \in L_{l o c}^{q} \text { for any } q>1
$$

for weak solutions of

$$
\begin{equation*}
\operatorname{div}(a(|\nabla u|) \nabla u)=\operatorname{div}(a(|f|) \mathbf{f}) \quad \text { in } \Omega . \tag{1.13}
\end{equation*}
$$

Additionally, the global gradient estimates in Orlicz spaces for weak solutions of (1.13) in a Reifenberg domain have been developed by Byun and Cho [15]. Recently, Beck \& Mingione [8] also proved the local Lipschitz regularity of weak solutions for nonuniformly elliptic variational problems, which satisfies the condition (1.2) or the fast, exponential-type growth conditions. Meanwhile, Baasandorj, Byun \& Oh [4] discussed the Calderón-Zygmund type estimates for non-uniformly elliptic equations of generalized double phase type in divergence form. A more detailed research progress on (1.12) can be found in the paper [41]. Just like the difference between the elliptic and parabolic quasilinear $p$-Laplace equations, it is much more difficult to deal with the parabolic case (1.1) than the corresponding elliptic case (1.13).

Recently, Diening, Scharle and Schwarzacher [31] obtained the interior Lipschitz regularity for weak solutions of

$$
\begin{equation*}
u_{t}-\operatorname{div}(a(|\nabla u|) \nabla u)=\operatorname{div}(a(|\mathbf{f}|) \mathbf{f}) . \tag{1.14}
\end{equation*}
$$

Subsequently, we [46,47] established the local Calderón-Zygmund estimates in the setting of Sobolev spaces and Lorentz spaces for weak solutions of (1.14). Moreover, many authors $[14,38]$ also studied the regularity estimates of weak solutions for the parabolic case (1.14). In recent years, there are many research activities on the weighted $L^{p}$-type estimates for the nonlinear elliptic equations [ $21,23,39,40,45$ ] and the nonlinear parabolic cases $[3,18,49]$. The purpose of this paper is to investigate the local weighted $L^{p}$-type regularity estimates for weak solutions of (1.1).

Let $Q(\theta, \rho)=B_{\rho} \times(-\theta, \theta]$ be a centered parabolic cylinder. Especially when $\theta=\rho^{2}$, we denote $Q_{\rho} \equiv Q\left(\rho^{2}, \rho\right)$. Throughout this paper we assume that the coefficients of $A=\left\{a_{i j}\right\}$ are in parabolic BMO spaces and their parabolic semi-norm are small enough. More precisely, we have the following definition.

Definition 1.1 (Small BMO coefficients). We say that the matrix $A$ of coefficients is $(\delta, R)$ vanishing if

$$
\sup _{Q_{z}(\theta, p) \subset \mathbb{R}^{n} \times \mathbb{R}} f_{Q_{z}(\theta, \rho)}\left|A(\zeta)-\bar{A}_{Q_{z}(\theta, p)}\right| d \zeta \leq \delta
$$

for $\rho \leq R$ and $\theta \leq R^{2}$, where $z=(x, t)$ and $\zeta=(y, s) \in \mathbb{R}^{n} \times \mathbb{R}$.
As usual, the solutions of (1.1) are taken in a weak sense. We now state the definition of weak solutions.
Definition 1.2. Assume that $\mathbf{f} \in L_{l o c}^{B}\left(\Omega_{T}\right)$ (see Definition 2.4). A function $u \in L_{l o c}^{\infty}(v, v+$ $\left.T ; L_{l o c}^{2}(\Omega)\right) \cap L_{l o c}^{B}\left(v, v+T ; W_{l o c}^{1, B}(\Omega)\right)$ is a local weak solution of (1.1) if for any compact subset $\mathcal{K}$ of $\Omega$ and for any subinterval $\left[t_{1}, t_{2}\right]$ of $(v, v+T)$ we have
$\left.\int_{\mathcal{K}} u \varphi d x\right|_{t_{1}} ^{t_{2}}+\int_{t_{1}}^{t_{2}} \int_{\mathcal{K}}\left\{-u \varphi_{t}+a\left((A \nabla u \cdot \nabla u)^{\frac{1}{2}}\right) A \nabla u \cdot \nabla \varphi\right\} d x d t=-\int_{t_{1}}^{t_{2}} \int_{\mathcal{K}} a(|\mathbf{f}|) \mathbf{f} \cdot \nabla \varphi d x d t$ for any $\varphi \in L_{\text {loc }}^{\infty}\left(v, v+T ; L^{2}(\mathcal{K})\right) \cap L_{\text {loc }}^{B}\left(v, v+T ; W_{0}^{1, B}(\mathcal{K})\right)$.

For convenience of the readers, we shall now give some definitions and properties on the weighted Lebesgue spaces (see $[43,44]$ ).

Definition 1.3. We call the positive function $w(x) \in L_{l o c}^{1}\left(\mathbb{R}^{n}\right)$ belongs to the class of the reverse Hölder weights $A_{q}$ for some $q>1$ if

$$
\left(f_{B_{r}} w(x) d x\right)\left(f_{B_{r}} w(x)^{\frac{-1}{q-1}} d x\right)^{q-1} \leq C
$$

for any ball $B_{r} \subset \mathbb{R}^{n}$ and some constant $C>0$. Moreover, we denote

$$
A_{\infty}:=\bigcup_{1<q<\infty} A_{q} \text { and } w\left(B_{r}\right):=\int_{B_{r}} w(x) d x .
$$

Furthermore, the corresponding weighted Lebesgue space $L_{w}^{p}\left(B_{r}\right)$ for any $p \geq 1$ consists of all functions $h$ which satisfy

$$
\|h\|_{L_{w} p\left(B_{r}\right)}:=\left(\int_{B_{r}}|h|^{p} w(x) d x\right)^{1 / p}<\infty .
$$

Now we are in a position to state the main result of this paper. Here we remark that just like in [49], the occurrence of the assumption $B(|\nabla u|) \in L^{1+\epsilon}\left(v, v+T ; L_{w}^{1+\epsilon}(\Omega)\right)$ locally is essentially due to the presence of the weight function $w$.

Theorem 1.4. Assume that $w(x) \in A_{q}$ for some $q>1$. Let $u$ be a local weak solution of (1.1) in $\Omega_{T}$ with $Q_{2} \subset \Omega_{T}$. Then for every $\epsilon \in(0, q-1)$ there exists a small $\delta=\delta\left(n, w, i_{a}, s_{a}, \epsilon, R, \Lambda\right)>0$ so that for each uniformly parabolic and $(\delta, R)$-vanishing $A$, and for all $\mathbf{f}$ with $B(|\mathbf{f}|) \in L^{q}\left(v, v+T ; L_{w}^{q}(\Omega)\right)$ locally, if $B(|\nabla u|) \in L^{1+\epsilon}\left(v, v+T ; L_{w}^{1+\epsilon}(\Omega)\right)$ locally, then we have $B(|\nabla u|) \in L^{q}\left(v, v+T ; L_{w}^{q}(\Omega)\right)$ locally with the following estimate

$$
\begin{aligned}
& \left(\int_{Q_{1}}[B(|\nabla u|)]^{q} w(x) d z\right)^{\frac{1}{q}} \\
& \quad \leq C\left[\left\{\int_{Q_{2}}[B(|\nabla u|)]^{1+\epsilon} w(x) d z\right\}^{\frac{1}{1+\epsilon}}+\left\{\int_{Q_{2}}[B(|\mathbf{f}|)]^{q} w(x) d z\right\}^{\frac{1}{q}}+1\right]^{\frac{q\left(2+s_{a}\right)}{2}},
\end{aligned}
$$

where the constant $C$ is independent of $u$ and $\mathbf{f}$.

## 2 Proof of the main result

This section is devoted to the proof of the main result stated in Theorem 1.4. For convenience of the readers, we recall some definitions and conclusions on the properties of the general Orlicz spaces.

Definition 2.1. A function $B:[0,+\infty) \rightarrow[0,+\infty)$ is said to be a Young function if it is convex and $B(0)=0$. Then a Young function $B$ is called an $N$-function if

$$
\begin{equation*}
0<B(t)<\infty \text { for } t>0 \text { and } \lim _{t \rightarrow+\infty} \frac{B(t)}{t}=\lim _{t \rightarrow 0+} \frac{t}{B(t)}=+\infty \tag{2.1}
\end{equation*}
$$

Moreover, we call a Young function $B$ satisfies the global $\Delta_{2}$ condition, denoted by $B \in \Delta_{2}$, if there exists a positive constant $K$ such that

$$
\begin{equation*}
B(2 t) \leq K B(t) \quad \text { for every } t>0 . \tag{2.2}
\end{equation*}
$$

Furthermore, an $N$-function $B$ is said to satisfy the global $\nabla_{2}$ condition, denoted by $B \in \nabla_{2}$, if there exists a number $\theta>1$ such that

$$
\begin{equation*}
B(t) \leq \frac{B(\theta t)}{2 \theta} \quad \text { for every } t>0 \tag{2.3}
\end{equation*}
$$

Remark 2.2. For example,

1. $G_{1}(t)=(1+t) \log (1+t)-t \in \Delta_{2}$, but $G_{1}(t) \notin \nabla_{2}$.
2. $G_{2}(t)=e^{t}-t-1 \in \nabla_{2}$, but $G_{2}(t) \notin \Delta_{2}$.
3. $G_{3}(t)=t^{p} \log (1+t) \in \Delta_{2} \cap \nabla_{2}$ for $p>1$.

Actually, it is easy to check that a Young function $B \in \Delta_{2} \cap \nabla_{2}$ if and only if

$$
\begin{equation*}
A_{1}\left(\frac{s}{t}\right)^{\beta_{2}} \leq \frac{B(s)}{B(t)} \leq A_{2}\left(\frac{s}{t}\right)^{\beta_{1}} \tag{2.4}
\end{equation*}
$$

for some constants $A_{2} \geq A_{1}>0, \beta_{1} \geq \beta_{2}>1$ and any $0<t \leq s$.

Lemma 2.3. If $B$ is an $N$-function, then $B$ satisfies the following Young's inequality

$$
\begin{equation*}
s t \leq \widetilde{B}(s)+B(t) \quad \text { for any } s, t \geq 0 \tag{2.5}
\end{equation*}
$$

Additionally, if $B \in \nabla_{2} \cap \Delta_{2}$, then we obtain the following Young's inequality with $\epsilon$

$$
\begin{equation*}
s t \leq \epsilon \widetilde{B}(s)+C(\epsilon) B(t) \quad \text { for any } s, t \geq 0 \quad \text { and } \quad \epsilon>0 . \tag{2.6}
\end{equation*}
$$

Definition 2.4. If $B$ is an $N$-function, then the Orlicz class $K^{B}(\Omega)$ consists of all measurable functions $g: \Omega \rightarrow \mathbb{R}$ satisfying

$$
\int_{\Omega} B(|g|) d x<\infty
$$

Also, the Orlicz space $L^{B}(\Omega)$ is the linear hull of $K^{B}(\Omega)$ endowed with the Luxemburg norm

$$
\|g\|_{L^{B}(\Omega)}:=\inf \left\{k>0: \int_{\Omega} B\left(\frac{|g(x)|}{k}\right) d x \leq 1\right\}
$$

On the other hand, the Orlicz-Sobolev space $W^{1, B}(\Omega):=\left\{g \in L^{B}(\Omega) \mid \nabla g \in L^{B}(\Omega)\right\}$, endowed with the norm $\|g\|_{W^{1, B}(\Omega)}:=\|g\|_{L^{B}(\Omega)}+\|\nabla g\|_{L^{B}(\Omega)}$.

Remark 2.5. In general, $K^{B}(\Omega) \subset L^{B}(\Omega)$ (see [2, Chapter 8]). But when $B \in \Delta_{2}, K^{B}(\Omega)=L^{B}(\Omega)$.
We first state the following properties on the functions $a(t)$ and $B(t)$ described above.
Lemma 2.6 (see [26, Proposition 2.9] and [48, Lemma 1.9]). If $a(t)$ satisfies (1.2) and $B(t)$ is defined in (1.5), then we have

1. $B(t)$ is strictly convex $N$-function and

$$
\widetilde{B}(b(t)) \leq C_{0} B(t) \text { for } t \geq 0 \text { and some constant } C_{0}>0
$$

2. $B(t) \in \Delta_{2} \cap \nabla_{2}$ with the estimate

$$
\begin{equation*}
A_{1}\left(\frac{s}{t}\right)^{2+i_{a}} \leq \frac{B(s)}{B(t)} \leq A_{2}\left(\frac{s}{t}\right)^{2+s_{a}} \quad \text { for any } 0<t \leq s \tag{2.7}
\end{equation*}
$$

3. $a(t) \theta^{i_{a}} \leq a(\theta t) \leq a(t) \theta^{s_{a}}$ for any $t>0$ and $\theta \geq 1$.

We continue to require certain properties of the functions $a(t)$ and $B(t)$, whose proofs can be found in Lemma 3 of [30] and Lemma 2.4 and Remark 2.5 of [48].

Lemma 2.7. If $a(t)$ satisfies (1.2), $A(x, t)$ satisfies (1.3) and $B(t)$ is defined in (1.5), then for any $\xi, \eta \in \mathbb{R}^{n}$ we have

$$
\begin{equation*}
a\left((A \xi \cdot \xi)^{\frac{1}{2}}\right) A \xi \cdot \xi \geq C\left(i_{a}, s_{a}, \Lambda\right) B(|\xi|) \tag{2.8}
\end{equation*}
$$

and

$$
\begin{equation*}
\left[a\left((A \xi \cdot \xi)^{\frac{1}{2}}\right) A \xi-a\left((A \eta \cdot \eta)^{\frac{1}{2}}\right) A \eta\right] \cdot(\xi-\eta) \geq C\left(i_{a}, s_{a}, \Lambda\right) B(|\xi-\eta|) \tag{2.9}
\end{equation*}
$$

Moreover, we first give the following $L^{1}$ estimate of weak solutions.

Lemma 2.8. Assume that $u$ is a local weak solution of (1.1) in $\Omega_{T}$ with $Q_{2} \subset \Omega_{T}$ and (1.2). Then we have

$$
\begin{equation*}
\int_{Q_{1}} B(|\nabla u|) d z \leq C \int_{Q_{2}} B(|u|)+B(|\mathbf{f}|)+1 d z . \tag{2.10}
\end{equation*}
$$

Proof. We may as well select the test function $\varphi=\zeta u$, which may be justified via Steklov average just like Remark 2.2 in [17], where $\zeta \in C_{0}^{\infty}\left(\mathbb{R}^{n+1}\right)$ is a cut-off function satisfying

$$
0 \leq \zeta \leq 1, \quad \zeta \equiv 1 \text { in } Q_{1} \quad \text { and } \quad \zeta \equiv 0 \text { in } \mathbb{R}^{n+1} / Q_{2}
$$

Then by Definition 1.2, after a direct calculation we show the resulting expression as

$$
I_{1}+I_{2}=I_{3}+I_{4}
$$

where

$$
\begin{aligned}
& I_{1}=\left.\frac{1}{2} \int_{B_{2}}|u(x, t)|^{2} \zeta(x, t) d x\right|_{t=-4} ^{t=4}=0, \\
& I_{2}=\int_{Q_{2}} \zeta a\left((A \nabla u \cdot \nabla u)^{\frac{1}{2}}\right) A \nabla u \cdot \nabla u d z, \\
& I_{3}=\frac{1}{2} \int_{Q_{2}} \zeta_{t} u^{2} d z-\int_{Q_{2}} u a\left((A \nabla u \cdot \nabla u)^{\frac{1}{2}}\right) A \nabla u \cdot \nabla \zeta d z, \\
& I_{4}=-\int_{Q_{2}} \zeta a(|\mathbf{f}|) \mathbf{f} \cdot \nabla u+u a(|\mathbf{f}|) \mathbf{f} \cdot \nabla \zeta d z .
\end{aligned}
$$

Estimate of $I_{2}$. It follows from (2.8) and the definition of $\zeta$ that

$$
I_{2} \geq C \int_{Q_{2}} \zeta B(|\nabla u|) d z \geq C \int_{Q_{1}} B(|\nabla u|) d z .
$$

Estimate of $I_{3}$. According to Lemma 2.3 and Lemma 2.6 (1), we conclude that

$$
\begin{aligned}
\left|I_{3}\right| & \leq C \int_{Q_{2}} a(|\nabla u|)|\nabla u||u|+|u|^{2} d z \\
& \leq \frac{\tau}{C_{0}} \int_{Q_{2}} \widetilde{B}(a(|\nabla u|)|\nabla u|) d z+C(\tau) \int_{Q_{2}} B(|u|)+|u|^{2} d z \\
& \leq \tau \int_{Q_{2}} B(|\nabla u|) d z+C(\tau) \int_{Q_{2}} B(|u|)+1 d z \quad \text { for any } \tau>0,
\end{aligned}
$$

where we have used the following inequality

$$
\begin{equation*}
B(\lambda) \geq A_{1} \lambda^{2+i_{a}} B(1) \geq A_{1} B(1) \lambda^{2} \text { for } \lambda \geq 1 \tag{2.11}
\end{equation*}
$$

by Lemma 2.6 (2) and the fact that $i_{a} \geq 0$.
Estimate of $I_{4}$. Similarly to the estimate of $I_{3}$, we have

$$
\begin{aligned}
\left|I_{4}\right| & \leq \int_{Q_{2}} a(|\mathbf{f}|)|\mathbf{f}||\nabla u|+a(|\mathbf{f}|)|\mathbf{f}||u| d z \\
& \leq \tau \int_{Q_{2}} B(|\nabla u|) d z+C(\tau) \int_{Q_{2}} B(|\mathbf{f}|)+B(|u|) d z \quad \text { for any } \tau>0 .
\end{aligned}
$$

Now we combine all the estimates of $I_{i}(1 \leq i \leq 4)$ to deduce that

$$
C \int_{Q_{1}} B(|\nabla u|) d z \leq 2 \tau \int_{Q_{2}} B(|\nabla u|) d z+C(\tau) \int_{Q_{2}} B(|\mathbf{f}|)+B(|u|)+1 d z .
$$

Selecting $\tau$ small enough and removing the above right-hand side first integral by a covering and iteration argument (see Lemma 4.1 of Chapter 2 in [24] or Lemma 2.1 of Chapter 3 in [34]), we finish the proof.

Next, we shall give some lemmas on the properties of $A_{q}$ weight.
Lemma 2.9 (see $[18,39,44]$ ). Assume that $w \in A_{q}$ for some $q>1$. Then there exists a small positive constant $\alpha \in(0,1)$ such that

$$
C_{1}\left(\frac{\left|B_{r}\right|}{\left|B_{R}\right|}\right)^{q} \leq \frac{w\left(B_{r}\right)}{w\left(B_{R}\right)} \leq C_{2}\left(\frac{\left|B_{r}\right|}{\left|B_{R}\right|}\right)^{\alpha}
$$

for any balls $B_{r} \subset B_{R} \subset \mathbb{R}^{n}$ and some constants $C_{1}, C_{2}>0$.
Remark 2.10. We remark that $A_{p_{1}} \subset A_{p}$ for any $1<p_{1} \leq p<\infty$ (see page 195 in [43]).
Now we recall the following self-improved property and reverse Hölder's inequality.
Lemma 2.11 (see $[43,44])$. Assume $w \in A_{q}$ for some $q>1$. Then there exist two constants $q_{1} \in(1, q)$ and $\epsilon_{0}>0$, depending only on $n, q$ and $w$ such that

$$
w \in A_{q_{1}} \quad \text { and } \quad\left\{f_{B_{r}} w^{1+\epsilon_{0}}(x) d x\right\}^{\frac{1}{1+\epsilon_{0}}} \leq C f_{B_{r}} w(x) d x
$$

The following is the measure theory of the weighted Lebesgue spaces.
Lemma 2.12. If $w(x) \in A_{q}$ for some $q>1$ and $g \in L_{w}^{q}\left(\mathbb{R}^{n+1}\right)$, then for any $q>\beta>1$ we have

$$
\begin{aligned}
\int_{\mathbb{R}^{n+1}}|g(z)|^{q} w(x) d x d t & =q \int_{0}^{+\infty} \lambda^{q-1} \int_{\left\{z \in \mathbb{R}^{n+1}:|g|>\lambda\right\}} w(x) d x d t d \lambda \\
& =(q-\beta) \int_{0}^{\infty} \lambda^{q-\beta-1} \int_{\left\{z \in \mathbb{R}^{n+1}:|g|>\lambda\right\}}|g|^{\beta} w(x) d x d t d \lambda
\end{aligned}
$$

Proof. Using Fubini's lemma, we conclude that

$$
\begin{aligned}
\int_{\mathbb{R}^{n+1}}|g(z)|^{q} w(x) d x d t & =\int_{\mathbb{R}^{n+1}}\left[q \int_{0}^{|g(z)|} \lambda^{q-1} d \lambda\right] w(x) d x d t \\
& =\int_{\mathbb{R}^{n+1}}\left[q \int_{0}^{+\infty} \lambda^{q-1} \chi_{\left\{z \in \mathbb{R}^{n+1}:|g|>\lambda\right\}} d \lambda\right] w(x) d x d t \\
& =q \int_{0}^{+\infty} \lambda^{q-1} \int_{\mathbb{R}^{n+1}} \chi_{\left\{z \in \mathbb{R}^{n+1}:|g|>\lambda\right\}} w(x) d x d t d \lambda \\
& =q \int_{0}^{+\infty} \lambda^{q-1} \int_{\left\{z \in \mathbb{R}^{n+1}:|g|>\lambda\right\}} w(x) d x d t d \lambda
\end{aligned}
$$

On the other hand, we apply Fubini's lemma to obtain that

$$
\begin{aligned}
(q- & \beta) \int_{0}^{+\infty} \lambda^{q-\beta-1} \int_{\left\{z \in \mathbb{R}^{n+1}:|g|>\lambda\right\}}|g|^{\beta} w(x) d x d t d \lambda \\
& =(q-\beta) \int_{0}^{+\infty} \lambda^{q-\beta-1} \int_{\mathbb{R}^{n+1}}|g|^{\beta} \chi_{\left\{z \in \mathbb{R}^{n+1}:|g|>\lambda\right\}} w(x) d x d t d \lambda \\
& =(q-\beta) \int_{\mathbb{R}^{n+1}} \int_{0}^{+\infty} \lambda^{q-\beta-1}|g|^{\beta} \chi_{\left\{z \in \mathbb{R}^{n+1}:|g|>\lambda\right\}} w(x) d \lambda d x d t \\
& =\int_{\mathbb{R}^{n+1}}|g|^{\beta}\left[(q-\beta) \int_{0}^{|g|} \lambda^{q-\beta-1} d \lambda\right] w(x) d x d t \\
& =\int_{\mathbb{R}^{n+1}}|g|^{\beta}|g|^{q-\beta} w(x) d x d t=\int_{\mathbb{R}^{n+1}}|g|^{q} w(x) d x d t
\end{aligned}
$$

Thus, we finish the proof.

Next, we shall need the following important iteration-covering lemma, in which we will divide the domain into several small parts.

Lemma 2.13. Given $\lambda \geq \lambda_{*}:=\frac{5^{n q+2}}{\min \left\{\sqrt{A_{1} B(1)}, 1\right\}} \lambda_{0}$, where

$$
\begin{equation*}
\lambda_{0}^{2}:=\left\{\frac{1}{w\left(B_{2}\right)} \int_{Q_{2}}[B(|\nabla u|)]^{q_{0}} w(x) d z\right\}^{\frac{1}{q_{0}}}+\frac{1}{\delta}\left\{\frac{1}{w\left(B_{2}\right)} \int_{Q_{2}}[B(|\mathbf{f}|)]^{q_{0}} w(x) d z\right\}^{\frac{1}{q_{0}}}+1 \tag{2.12}
\end{equation*}
$$

and $q_{0}=1+\epsilon$ with $\epsilon \in(0, q-1)$, there exist a family of disjoint cylinders $\left\{Q_{i}^{0}\right\}_{i \in \mathbb{N}}$ with

$$
Q_{i}^{0}:=Q_{z_{i}}\left(\frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}, \rho_{i}\right)
$$

$z_{i}=\left(x_{i}, t_{i}\right) \in E\left(Q_{1}, \lambda\right):=\left\{z \in Q_{1}: B(|\nabla u|)>B(\lambda)\right\}$ and $0<\rho_{i}=\rho\left(z_{i}\right) \leq 1 / 10$ such that
1.

$$
J\left[Q_{i}^{0}\right]=J\left[Q_{z_{i}}\left(\frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}, \rho_{i}\right)\right]=B(\lambda)
$$

where

$$
\begin{aligned}
J[Q(\theta, \rho)]:= & \left\{\frac{1}{2 \theta w\left(B_{\rho}\right)} \int_{Q(\theta, \rho)}[B(|\nabla u|)]^{q_{0}} w(x) d z\right\}^{\frac{1}{q_{0}}} \\
& +\frac{1}{\delta}\left\{\frac{1}{2 \theta w\left(B_{\rho}\right)} \int_{Q(\theta, \rho)}[B(|\mathbf{f}|)]^{q_{0}} w(x) d z\right\}^{\frac{1}{q_{0}}} .
\end{aligned}
$$

2. 

$$
J\left[Q_{i}^{j}\right]=J\left[5 j Q_{i}^{0}\right]=J\left[Q_{z_{i}}\left(25 j^{2} \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}, 5 j \rho_{i}\right)\right]<B(\lambda) \quad \text { for } j=1,2
$$

where $Q_{i}^{j}=5 j Q_{i}^{0}=Q_{z_{i}}\left(25 j^{2} \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}, 5 j \rho_{i}\right)$.
3. $E\left(Q_{1}, \lambda\right) \subset \bigcup_{i \in \mathbb{N}} Q_{i}^{1} \cup$ negligible set.

Proof. Let $\lambda \geq \lambda_{*}$. From Lemma 2.9 we have

$$
\begin{aligned}
\left\{\frac{\int_{Q_{2}}[B(|\nabla u|)]^{q_{0}} w(x) d z}{\int_{Q_{z_{0}}\left(\frac{\lambda^{2}}{B(\lambda)} \rho^{2}, \rho\right)} w(x) d x d t}\right\}^{\frac{1}{q_{0}}} & =\left\{\frac{\int_{Q_{2}}[B(|\nabla u|)]^{q_{0}} w(x) d z}{2 w\left(B_{\rho}\left(x_{0}\right)\right) \frac{\lambda^{2}}{B(\lambda)} \rho^{2}}\right\}^{\frac{1}{q_{0}}} \\
& \leq\left\{\frac{w\left(B_{2}\right)}{2 w\left(B_{\rho}\left(x_{0}\right)\right) \frac{\lambda^{2}}{B(\lambda)} \rho^{2}} \frac{1}{w\left(B_{2}\right)} \int_{Q_{2}}[B(|\nabla u|)]^{q_{0}} w(x) d z\right\}^{\frac{1}{q_{0}}} \\
& \leq\left\{50 \cdot 20^{n q} \frac{B(\lambda)}{\lambda^{2}} \frac{1}{w\left(B_{2}\right)} \int_{Q_{2}}[B(|\nabla u|)]^{q_{0}} w(x) d z\right\}^{\frac{1}{\eta_{0}}}
\end{aligned}
$$

for any $z_{0}=\left(x_{0}, t_{0}\right) \in Q_{1}$ and $1 / 10 \leq \rho \leq 1$, and then

$$
\left\{\frac{\int_{Q_{2}}[B(|\mathbf{f}|)]^{q_{0}} w(x) d z}{\left.\int_{Q_{z_{0}}\left(\frac{\lambda^{2}}{B(\lambda)}\right.} \rho^{2}, \rho\right)} w\right\}^{\frac{1}{q_{0}}} \leq\left\{50 \cdot 20^{n q} \frac{B(\lambda) d x d t}{\lambda^{2}} \frac{1}{w\left(B_{2}\right)} \int_{Q_{2}}[B(|\mathbf{f}|)]^{q_{0}} w(x) d z\right\}^{\frac{1}{q_{0}}}
$$

So, (2.11) implies that

$$
\begin{aligned}
J\left[Q_{z_{0}}\left(\frac{\lambda^{2}}{B(\lambda)} \rho^{2}, \rho\right)\right] & \leq\left[50 \cdot 20^{n q} A_{1} B(1) \frac{B(\lambda)}{A_{1} B(1) \lambda^{2}}\right]^{\frac{1}{q_{0}}} \lambda_{0}^{2} \\
& \leq\left[25^{n q+2} \cdot A_{1} B(1)\right]^{\frac{1}{q_{0}}} \frac{B(\lambda)}{A_{1} B(1) \lambda^{2}} \lambda_{0}^{2} \\
& \leq 25^{n q+2} \max \left\{1, A_{1} B(1)\right\} \frac{B(\lambda)}{A_{1} B(1) \lambda^{2}} \lambda_{0}^{2} \\
& =\lambda_{*}^{2} \frac{B(\lambda)}{\lambda^{2}} \leq B(\lambda)
\end{aligned}
$$

for any $z_{0} \in Q_{1}$ and $\frac{1}{10} \leq \rho \leq 1$. Therefore, we deduce that

$$
\begin{equation*}
\sup _{z=(x, t) \in Q_{1}} \sup _{\frac{1}{10} \leq \rho \leq 1} J\left[Q_{z}\left(\frac{\lambda^{2}}{B(\lambda)} \rho^{2}, \rho\right)\right] \leq B(\lambda) \tag{2.13}
\end{equation*}
$$

Using Lebesgue's differentiation theorem, for a.e. $z \in E\left(Q_{1}, \lambda\right)$ we conclude that

$$
\lim _{\rho \rightarrow 0} J\left[Q_{z}\left(\frac{\lambda^{2}}{B(\lambda)} \rho^{2}, \rho\right)\right]>B(\lambda)
$$

which implies that there exists some $\rho_{0} \in(0,1]$ satisfying

$$
\begin{equation*}
J\left[Q_{z}\left(\frac{\lambda^{2}}{B(\lambda)} \rho_{0}^{2}, \rho_{0}\right)\right]>B(\lambda) \tag{2.14}
\end{equation*}
$$

Therefore, from (2.13) and (2.14) we can select a radius $\rho_{z} \in(0,1 / 10]$ such that

$$
\rho_{z}=: \max \left\{\rho \left\lvert\, J\left[Q_{z}\left(\frac{\lambda^{2}}{B(\lambda)} \rho^{2}, \rho\right)\right]=B(\lambda)\right., \quad 0<\rho \leq 1 / 10\right\}
$$

which implies that

$$
J\left[Q_{z}\left(\frac{\lambda^{2}}{B(\lambda)} \rho_{z}^{2}, \rho_{z}\right)\right]=B(\lambda) \quad \text { and } \quad J\left[Q_{z}\left(\frac{\lambda^{2}}{B(\lambda)} \rho^{2}, \rho\right)\right]<B(\lambda) \quad \text { for } \rho_{z}<\rho \leq 1
$$

Thus, by using Vitali's covering lemma, we can find a family of disjoint cylinders $\left\{Q_{i}^{0}\right\}_{i \in \mathbb{N}}=$ $\left\{Q_{z_{i}}\left(\frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}, \rho_{i}\right)\right\}_{i \in \mathbb{N}}$ with $z_{i}=\left(x_{i}, t_{i}\right) \in E\left(Q_{1}, \lambda\right)$ and $\rho_{i}=\rho\left(z_{i}\right) \leq 1 / 10$ such that (1)-(3) are true.

Now we employ the above lemma to derive careful estimates on the small parabolic cylinders $\left\{Q_{i}^{0}\right\}_{i \in \mathbb{N}}$, whose precise structure will be needed below.

Lemma 2.14. Under the same hypothetical conditions as the lemma above, for $\lambda \geq \lambda_{*}$ we obtain

$$
\begin{aligned}
w\left(B_{\rho_{i}}\left(x_{i}\right)\right) \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}< & \frac{C}{[B(\lambda)]^{q_{0}}} \int_{\left\{z \in Q_{i}^{0}: B(|\nabla u|)>\frac{B(\lambda)}{4}\right\}}[B(|\nabla u|)]^{q_{0}} w(x) d z \\
& +\frac{C}{\delta q_{0}[B(\lambda)]^{q_{0}}} \int_{\left\{z \in Q_{i}^{0}: B(|\mathbf{f}|)>\frac{B(\lambda)}{40}\right\}}[B(|\mathbf{f}|)]^{q_{0}} w(x) d z
\end{aligned}
$$

Proof. From Lemma 2.13 (1), we find that

$$
\begin{aligned}
& \left\{\frac{1}{2 w\left(B_{\rho_{i}}\left(x_{i}\right)\right) \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}} \int_{Q_{i}^{0}}[B(|\nabla u|)]^{q_{0}} w(x) d z\right\}^{\frac{1}{q_{0}}} \\
& \quad+\frac{1}{\delta}\left\{\frac{1}{2 w\left(B_{\rho_{i}}\left(x_{i}\right)\right) \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}} \int_{Q_{i}^{0}}[B(|\mathbf{f}|)]^{q_{0}} w(x) d z\right\}^{\frac{1}{q_{0}}}=B(\lambda) .
\end{aligned}
$$

Therefore, one of the following inequalities must be true

$$
\begin{equation*}
\left\{\frac{1}{2 w\left(B_{\rho_{i}}\left(x_{i}\right)\right) \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}} \int_{Q_{i}^{0}}[B(|\nabla u|)]^{q_{0}} w(x) d z\right\}^{\frac{1}{q_{0}}}>\frac{B(\lambda)}{2} \tag{2.15}
\end{equation*}
$$

and

$$
\begin{equation*}
\frac{1}{\delta}\left\{\frac{1}{2 w\left(B_{\rho_{i}}\left(x_{i}\right)\right) \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}} \int_{Q_{i}^{0}}[B(|\mathbf{f}|)]^{q_{0}} w(x) d z\right\}^{\frac{1}{q_{0}}}>\frac{B(\lambda)}{2} \tag{2.16}
\end{equation*}
$$

If (2.15) is true, then we have

$$
\begin{aligned}
2 w\left(B_{\rho_{i}}\left(x_{i}\right)\right) \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}< & {\left[\frac{2}{B(\lambda)}\right]^{q_{0}} \int_{Q_{i}^{0}}[B(|\nabla u|)]^{q_{0}} w(x) d z } \\
\leq & {\left[\frac{2}{B(\lambda)}\right]^{q_{0}} \int_{\left\{z \in Q_{i}^{0}: B(|\nabla u|)>\frac{B(\lambda)}{4}\right\}}[B(|\nabla u|)]^{q_{0}} w(x) d z } \\
& +\frac{1}{2^{q_{0}-1}} w\left(B_{\rho_{i}}\left(x_{i}\right)\right) \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2} \quad \text { for } \lambda \geq \lambda_{*}
\end{aligned}
$$

which implies that

$$
w\left(B_{\rho_{i}}\left(x_{i}\right)\right) \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}<\frac{C}{[B(\lambda)]^{q_{0}}} \int_{\left\{z \in Q_{i}^{0}: B(|\nabla u|)>\frac{B(\lambda)}{4}\right\}}[B(|\nabla u|)]^{q_{0}} w(x) d z \quad \text { for } \lambda \geq \lambda_{*}
$$

Similarly, if (2.16) is true, then we have

$$
w\left(B_{\rho_{i}}\left(x_{i}\right)\right) \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}<\frac{C}{\delta^{q_{0}}[B(\lambda)]^{q_{0}}} \int_{\left\{z \in Q_{i}^{0}: B(|\mathbf{f}|)>\frac{B(\lambda)}{4 \delta}\right\}}[B(|\mathbf{f}|)]^{q_{0}} w(x) d z \quad \text { for } \lambda \geq \lambda_{*} .
$$

Finally, we combine the two estimates above to find that

$$
\begin{aligned}
w\left(B_{\rho_{i}}\left(x_{i}\right)\right) \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}< & \frac{C}{[B(\lambda)]^{q_{0}}} \int_{\left\{z \in Q_{i}^{0}: B(|\nabla u|)>\frac{B(\lambda)}{4}\right\}}[B(|\nabla u|)]^{q_{0}} w(x) d z \\
& +\frac{C}{\delta^{q_{0}}[B(\lambda)]^{q_{0}}} \int_{\left\{z \in Q_{i}^{0}: B(|\mathbf{f}|)>\frac{B(\lambda)}{4 \delta}\right\}}[B(|\mathbf{f}|)]^{q_{0}} w(x) d z
\end{aligned}
$$

for $\lambda \geq \lambda_{*}$. Thus, this finishes our proof.

Lemma 2.15. Under the same hypotheses and results as those in Lemma 2.13, we have

$$
\begin{equation*}
\left(f_{Q_{i}^{j}}[B(|\nabla u|)]^{q_{2}} d z\right)^{\frac{1}{q_{2}}}<C B(\lambda) \text { and }\left(f_{Q_{i}^{j}}[B(|\mathbf{f}|)]^{q_{2}} d z\right)^{\frac{1}{q_{2}}}<C \delta B(\lambda) \tag{2.17}
\end{equation*}
$$

for $j=1,2, \lambda \geq \lambda_{*}$ and some constant $q_{2} \in\left(1, q_{0}\right)$.
Proof. It follows from Lemma 2.13 (2) that

$$
\left(\frac{1}{w\left(B_{5 j \rho_{i}}\left(x_{i}\right)\right) \frac{\lambda^{2}}{B(\lambda)} 50 j^{2} \rho_{i}^{2}} \int_{Q_{i}^{j}}[B(|\nabla u|)]^{q_{0}} w(x) d x d t\right)^{\frac{1}{q_{0}}}<B(\lambda)
$$

and

$$
\left(\frac{1}{w\left(B_{5 j \rho_{i}}\left(x_{i}\right)\right) \frac{\lambda^{2}}{B(\lambda)} 50 j^{2} \rho_{i}^{2}} \int_{Q_{i}^{j}}[B(|\mathbf{f}|)]^{q_{0}} w(x) d x d t\right)^{\frac{1}{q_{0}}}<\delta B(\lambda)
$$

for $j=1,2$. Since $w \in A_{q_{0}}$ by Remark 2.10, we find that $w \in A_{q_{1}}$ for some $q_{1} \in\left(1, q_{0}\right)$ in view of Lemma 2.11. Let $q_{2}=\frac{q_{0}}{q_{1}} \in\left(1, q_{0}\right)$. Then by using Hölder's inequality and the two inequalities above, we see

$$
\begin{aligned}
& \left(f_{Q_{i}^{j}}[B(|\nabla u|)]^{q_{2}} d z\right)^{\frac{1}{q_{2}}} \\
& =\left(f_{Q_{i}^{j}}[B(|\nabla u|)]^{q_{2}} w(x)^{\frac{1}{q_{1}}} w(x)^{-\frac{1}{q_{1}}} d z\right)^{\frac{1}{q_{2}}} \\
& \leq\left(\frac{w\left(B_{5 j \rho_{i}}\left(x_{i}\right)\right)}{\left|B_{5 j \rho_{i}}\left(x_{i}\right)\right|} \frac{1}{w\left(B_{5 j \rho_{i}}\left(x_{i}\right)\right) \frac{\lambda^{2}}{B(\lambda)} 50 j^{2} \rho_{i}^{2}} \int_{Q_{i}^{j}}[B(|\nabla u|)]^{q_{0}} w(x) d z\right)^{\frac{1}{q_{0}}}\left(f_{Q_{i}^{j}} w(x)^{-\frac{1}{q_{1}-1}} d z\right)^{\frac{q_{1}-1}{q_{0}}} \\
& <B(\lambda)\left[f_{B_{5 j p_{i}}\left(x_{i}\right)} w(x) d x\left(f_{B_{5 j p_{i}}\left(x_{i}\right)} w(x)^{-\frac{1}{q_{1}-1}} d x\right)^{q_{1}-1}\right]^{\frac{1}{q_{0}}} \\
& <C B(\lambda) \quad \text { for } j=1,2 .
\end{aligned}
$$

Similarly, we have

$$
\left(f_{Q_{i}^{j}}[B(|\mathbf{f}|)]^{q_{2}} d z\right)^{\frac{1}{q_{2}}}<C \delta B(\lambda) \quad \text { for } j=1,2
$$

which finishes our proof.
Moreover, we can obtain the following comparison result and interior Lipschitz regularity.
Lemma 2.16. Assume that $u$ is a local weak solution of (1.1) in $\Omega_{T}$ with (1.2). If $v$ is the weak solution of

$$
\begin{equation*}
v_{t}-\operatorname{div}\left(a\left(\left(\bar{A}_{Q_{2}} \nabla v \cdot \nabla v\right)^{\frac{1}{2}}\right) \bar{A}_{Q_{2}} \nabla v\right)=0 \quad \text { in } Q_{2} \subset \Omega_{T} \tag{2.18}
\end{equation*}
$$

with $v=u$ on $\partial_{p} Q_{2}$, then we have

$$
\begin{equation*}
\int_{Q_{2}} B(|\nabla v|) d z \leq C \int_{Q_{2}} B(|\nabla u|)+B(|\mathbf{f}|) d z \tag{2.19}
\end{equation*}
$$

and

$$
\begin{equation*}
\sup _{Q_{1}} B(|\nabla v|) \leq C\left[f_{Q_{2}} B(|\nabla u|)+B(|\mathbf{f}|)+1 d z\right]^{\frac{2+s_{a}}{2}}, \tag{2.20}
\end{equation*}
$$

where the constant $C=C\left(i_{a}, s_{a}, \Lambda\right)>0$.
Proof. Noting that $u$ and $v$ are the weak solutions of (1.1) and (2.18) respectively, we may as well select the test function $\varphi=v-u$ since $v=u$ on $\partial_{p} Q_{2}$, which is possible modulo Steklov average. Then a direct calculation shows the resulting expression as

$$
I_{1}+I_{2}=I_{3}+I_{4}+I_{5}
$$

where

$$
\begin{aligned}
& I_{1}=\frac{1}{2} \int_{B_{2}}|v(x, 4)-u(x, 4)|^{2} d x \geq 0, \\
& I_{2}=\int_{Q_{2}} a\left(\left(\bar{A}_{Q_{2}} \nabla v \cdot \nabla v\right)^{\frac{1}{2}}\right) \bar{A}_{Q_{2}} \nabla v \cdot \nabla v d z, \\
& I_{3}=\int_{Q_{2}} a\left(\left(\bar{A}_{Q_{2}} \nabla v \cdot \nabla v\right)^{\frac{1}{2}}\right) \bar{A}_{Q_{2}} \nabla v \cdot \nabla u d z, \\
& \left.I_{4}=\int_{Q_{2}} a\left((A \nabla u \cdot \nabla u)^{\frac{1}{2}}\right) A \nabla u \cdot \nabla v-a\left((A \nabla u \cdot \nabla u)^{\frac{1}{2}}\right) A \nabla u \cdot \nabla u\right] d z, \\
& I_{5}=\int_{Q_{2}} a(|\mathbf{f}|) \mathbf{f} \cdot \nabla v-a(|\mathbf{f}|) \mathbf{f} \cdot \nabla u d z .
\end{aligned}
$$

Estimate of $I_{2}$. Owing to Lemma 2.7 we thereby discover

$$
I_{2} \geq C \int_{Q_{2}} B(|\nabla v|) d z .
$$

Estimate of $I_{3}$. According to Lemma 2.3 and Lemma 2.6 (1), we see

$$
\begin{aligned}
\left|I_{3}\right| & \leq \frac{\tau}{C_{0}} \int_{Q_{2}} \widetilde{B}(a(|\nabla v|)|\nabla v|) d z+C(\tau) \int_{Q_{2}} B(|\nabla u|) d z \\
& \leq \tau \int_{Q_{2}} B(|\nabla v|) d z+C(\tau) \int_{Q_{2}} B(|\nabla u|) d z \quad \text { for any } \tau>0 .
\end{aligned}
$$

Estimates of $I_{i}(4 \leq i \leq 5)$. As in the proof of the estimate of $I_{3}$, we compute

$$
\begin{aligned}
& \left|I_{4}\right| \leq \tau \int_{Q_{2}} B(|\nabla v|) d z+C(\tau) \int_{Q_{2}} B(|\nabla u|) d z, \\
& \left|I_{5}\right| \leq \tau \int_{Q_{2}} B(|\nabla v|) d z+C(\tau) \int_{Q_{2}} B(|\nabla u|)+B(|\mathbf{f}|) d z \quad \text { for any } \tau>0 .
\end{aligned}
$$

Therefore, by selecting $\tau>0$ small enough we obtain the desired result (2.19) from the estimates of $I_{i}(1 \leq i \leq 5)$. Rather, from Theorem 2.2 in [31] we have

$$
\min \left\{\sup _{Q_{1}} \rho(|\nabla v|), \sup _{Q_{1}}|\nabla v|^{2}\right\} \leq C f_{Q_{2}}|\nabla v|^{2}+B(|\nabla v|) d z,
$$

where $C=\left(i_{a}, s_{a}, \Lambda\right)$ and

$$
\rho(t)=(B(t))^{\frac{n}{2}} t^{2-n} \geq C t^{n} t^{2-n}=C t^{2} \quad \text { for any } t \geq 1
$$

by (2.11), which implies that

$$
\begin{equation*}
\sup _{Q_{1}}|\nabla v|^{2} \leq C \min \left\{\sup _{Q_{1}} \rho(|\nabla v|), \sup _{Q_{1}}|\nabla v|^{2}\right\}+C \leq C f_{Q_{2}} B(|\nabla v|)+1 d z . \tag{2.21}
\end{equation*}
$$

Furthermore, we deduce from (2.7) that

$$
\sup _{Q_{1}} B(|\nabla v|) \leq C\left[f_{Q_{2}} B(|\nabla v|)+1 d z\right]^{\frac{2+s_{a}}{2}},
$$

which implies that (2.20) is valid by (2.19). Thus, we finish the proof.
Moreover, we shall give the following essential estimate on the level set.
Lemma 2.17. For any $\epsilon>0$, there exists a small $\delta=\delta(\epsilon)>0$ such that if $u$ is a local weak solution of (1.1) in $\Omega_{T}$ with (1.2) and $Q_{2} \subset \Omega_{T}$,

$$
\begin{gather*}
f_{Q_{2}}\left|A-\bar{A}_{Q_{2}}\right| d z \leq \delta,  \tag{2.22}\\
\left\{f_{Q_{2}}[B(|\nabla u|)]^{q_{2}} d z\right\}^{\frac{1}{q_{2}}} \leq 1 \text { and }\left\{f_{Q_{2}}[B(|\mathbf{f}|)]^{q_{2}} d z\right\}^{\frac{1}{q_{2}}} \leq \delta, \tag{2.23}
\end{gather*}
$$

then there exists a constant $N_{1}>1$ such that

$$
\left|\left\{z \in Q_{1}: B(|\nabla u|)>N_{1}\right\}\right| \leq C \epsilon\left|Q_{1}\right| .
$$

Proof. If $v$ is the weak solution of (2.18) in $Q_{2}$ with $v=u$ on $\partial_{p} Q_{2}$, then by selecting the test function $\varphi=u-v$, which is possible modulo Steklov average, after a direct calculation we show the resulting expression as

$$
I_{1}+I_{2}=I_{3}+I_{4}
$$

where

$$
\begin{aligned}
& I_{1}=\frac{1}{2} \int_{B_{2}}|u(x, 4)-v(x, 4)|^{2} d x \geq 0, \\
& I_{2}=\int_{Q_{2}}\left[a\left(\left(\bar{A}_{Q_{2}} \nabla u \cdot \nabla u\right)^{\frac{1}{2}}\right) \bar{A}_{Q_{2}} \nabla u-a\left(\left(\bar{A}_{Q_{2}} \nabla v \cdot \nabla v\right)^{\frac{1}{2}}\right) \bar{A}_{Q_{2}} \nabla v\right] \cdot(\nabla u-\nabla v) d z, \\
& I_{3}=-\int_{Q_{2}}\left[a\left((A \nabla u \cdot \nabla u)^{\frac{1}{2}}\right) A \nabla u-a\left(\left(\bar{A}_{Q_{2}} \nabla u \cdot \nabla u\right)^{\frac{1}{2}}\right) \bar{A}_{Q_{2}} \nabla u\right] \cdot(\nabla u-\nabla v) d z, \\
& I_{4}=\int_{Q_{2}} a(|\mathbf{f}|) \mathbf{f} \cdot \nabla v d z-\int_{Q_{2}} a(|\mathbf{f}|) \mathbf{f} \cdot \nabla u d z .
\end{aligned}
$$

Estimate of $I_{2}$. Using Lemma 2.7, we observe that

$$
I_{2} \geq C \int_{Q_{2}} B(|\nabla u-\nabla v|) d z .
$$

Estimate of $I_{3}$. First of all, we discover

$$
\begin{aligned}
\left|I_{3}\right| \leq & \int_{Q_{2}} a\left((A \nabla u \cdot \nabla u)^{\frac{1}{2}}\right)\left|A-\bar{A}_{Q_{2}}\right||\nabla u||\nabla u-\nabla v| d z \\
& +\int_{Q_{2}}\left|a\left((A \nabla u \cdot \nabla u)^{\frac{1}{2}}\right)-a\left(\left(\bar{A}_{Q_{2}} \nabla u \cdot \nabla u\right)^{\frac{1}{2}}\right)\right|\left|\bar{A}_{Q_{2}} \nabla u\right||\nabla u-\nabla v| d z \\
= & I_{31}+I_{32} .
\end{aligned}
$$

Estimate of $I_{31}$. From (1.3), Lemma 2.6, Young's inequality and Hölder's inequality we find that

$$
\begin{aligned}
\left|I_{31}\right| & \leq C \int_{Q_{2}} a(|\nabla u|)|\nabla u|\left|A-\bar{A}_{Q_{2}}\right||\nabla u-\nabla v| d z \\
& \leq \frac{\epsilon}{2 \Lambda} \int_{Q_{2}} B(|\nabla u-\nabla v|)\left|A-\bar{A}_{Q_{2}}\right| d z+C(\epsilon) \int_{Q_{2}} \widetilde{B}(a(|\nabla u|)|\nabla u|)\left|A-\bar{A}_{Q_{2}}\right| d z \\
& \leq \epsilon \int_{Q_{2}} B(|\nabla u-\nabla v|) d z+C(\epsilon) \int_{Q_{2}} B(|\nabla u|)\left|A-\bar{A}_{Q_{2}}\right| d z \\
& \leq \epsilon \int_{Q_{2}} B(|\nabla u-\nabla v|) d z+C(\epsilon)\left\{\int_{Q_{2}}[B(|\nabla u|)]^{q_{2}} d z\right\}^{\frac{1}{q_{2}}}\left[\int_{Q_{2}}\left|A-\bar{A}_{Q_{2}}\right|^{\frac{q_{2}}{q_{2}-1}} d z\right]^{\frac{q_{2}-1}{q_{2}}}
\end{aligned}
$$

for any $\epsilon>0$, which implies that

$$
\begin{aligned}
\left|I_{31}\right| & \leq \epsilon \int_{Q_{2}} B(|\nabla u-\nabla v|) d z+C(\epsilon)\left[\int_{Q_{2}}\left|A-\bar{A}_{Q_{2}}\right| d z\right]^{\frac{q_{2}-1}{q_{2}}} \\
& \leq \epsilon \int_{Q_{2}} B(|\nabla u-\nabla v|) d z+C(\epsilon) \delta^{\frac{q_{2}-1}{q_{2}}}
\end{aligned}
$$

where we used the given conditions (2.22)-(2.23).
Estimate of $I_{32}$. (1.2), (1.3), Lemma 2.6 and Lagrange's mean value theorem yield the bound

$$
\left|I_{32}\right| \leq C \int_{Q_{2}} a(|\nabla u|)|\nabla u|\left|A-\bar{A}_{Q_{2}}\right||\nabla u-\nabla v| d z
$$

and so

$$
\left|I_{32}\right| \leq \epsilon \int_{Q_{2}} B(|\nabla u-\nabla v|) d z+C(\epsilon) \delta^{\frac{q_{2}-1}{q_{2}}} \quad \text { for any } \epsilon>0
$$

whose proof is totally similar to that of $I_{31}$.
Estimate of $I_{4}$. Lemma 2.3, Lemma 2.6 (1), Hölder's inequality and (2.23) assert

$$
\begin{aligned}
\left|I_{4}\right| & \leq \int_{Q_{2}} a(|\mathbf{f}|)|\mathbf{f}||\nabla u-\nabla v| d z \\
& \leq \epsilon \int_{Q_{2}} B(|\nabla u-\nabla v|) d z+C(\epsilon) \int_{Q_{2}} \widetilde{B}(a(|\mathbf{f}|)|\mathbf{f}|) d z \\
& \leq \epsilon \int_{Q_{2}} B(|\nabla u-\nabla v|) d z+C(\epsilon) \int_{Q_{2}} B(|\mathbf{f}|) d z \\
& \leq \epsilon \int_{Q_{2}} B(|\nabla u-\nabla v|) d z+C(\epsilon)\left[f_{Q_{2}}(B(|\mathbf{f}|))^{q_{2}} d z\right]^{\frac{1}{q_{2}}} \\
& \leq \epsilon \int_{Q_{2}} B(|\nabla u-\nabla v|) d z+C(\epsilon) \delta
\end{aligned}
$$

for any $\epsilon>0$. So, by selecting $\epsilon>0$ small enough and combining the estimates of $I_{i}(1 \leq i \leq$ 4) we conclude that

$$
\begin{equation*}
\int_{Q_{2}} B(|\nabla u-\nabla v|) d z \leq C(\epsilon) \delta+C(\epsilon) \delta^{\frac{q_{2}-1}{q_{2}}} \leq \epsilon \tag{2.24}
\end{equation*}
$$

where we have chosen a small constant $\delta>0$ satisfying the above last inequality. Thus, it follows from Lemma 2.16, Hölder's inequality and the assumed condition (2.23) that

$$
\begin{align*}
\sup _{Q_{1}} B(|\nabla v|) & \leq C\left[f_{Q_{2}} B(|\nabla u|)+B(|\mathbf{f}|)+1 d z\right]^{\frac{2+s_{a}}{2}} \\
& \leq C\left\{\left[f_{Q_{2}}(B(|\nabla u|))^{q_{2}} d z\right]^{\frac{1}{q_{2}}}+\left[f_{Q_{2}}(B(|\mathbf{f}|))^{q_{2}} d z\right]^{\frac{1}{q_{2}}}+1\right\}^{\frac{2+s_{a}}{2}} \\
& \leq N_{0} \tag{2.25}
\end{align*}
$$

for some constant $N_{0}>1$. Finally, from (2.24), (2.25) and the fact that $B \in \Delta_{2} \cap \nabla_{2}$ is convex we have

$$
\begin{aligned}
\mid\{z & \left.\in Q_{1}: B(|\nabla u|)>2 C_{*} N_{0}\right\} \mid \\
& \leq\left|\left\{z \in Q_{1}: B(|\nabla(u-v)|)>N_{0}\right\}\right|+\left|\left\{z \in Q_{1}: B(|\nabla v|)>N_{0}\right\}\right| \\
& =\left|\left\{z \in Q_{1}: B(|\nabla(u-v)|)>N_{0}\right\}\right| \\
& \leq \frac{1}{N_{0}} \int_{Q_{2}} B(|\nabla(u-v)|) d z \leq C \epsilon\left|Q_{2}\right| \leq C \epsilon\left|Q_{1}\right|,
\end{aligned}
$$

where we have used the following inequality

$$
B(a+b) \leq \frac{1}{2} B(2 a)+\frac{1}{2} B(2 b) \leq C_{*} B(a)+C_{*} B(b) \quad \text { for any } a, b \geq 0 .
$$

This completes our proof by choosing $N_{1}=2 C_{*} N_{0}>1$.
Furthermore, we shall give the following result.
Lemma 2.18. Assume that $\lambda \geq \lambda_{*}$. For any $\epsilon>0$, there exists a small $\delta=\delta(\epsilon)>0$ such that if $u$ is a local weak solution of (1.1) in $\Omega_{T}$ with $Q_{2} \subset \Omega_{T}$, then we have

$$
\begin{aligned}
& \int_{\left\{z \in \mathrm{Q}_{1}: B(|\nabla u|)>N_{1} B(\lambda)\right\}} w(x) d x d t \leq \frac{C \epsilon^{\alpha}}{[B(\lambda)]^{q_{0}}} \int_{\left\{z \in \mathrm{Q}_{2}: B(|\nabla u|)>\frac{B(\lambda)}{4}\right\}}[B(|\nabla u|)]^{q_{0}} w(x) d x d t \\
&+\frac{C \epsilon^{\alpha}}{\delta q_{0}}[B(\lambda)]^{q_{0}} \\
& \int_{\left\{z \in Q_{2}: B(| || | \mid)>\frac{B(\lambda)}{40}\right\}}[B(|\mathbf{f}|)]^{q_{0}} w(x) d x d t .
\end{aligned}
$$

Proof. 1. We first claim that

$$
\begin{equation*}
\left|\left\{z \in Q_{i}^{1}: B(|\nabla u|)>N_{1} B(\lambda)\right\}\right| \leq C \epsilon\left|Q_{i}^{1}\right| . \tag{2.26}
\end{equation*}
$$

To prove this, for each $\lambda \geq 1$ we use the normalization and scaling methods by defining

$$
\begin{aligned}
& u_{\lambda}^{i}(x, t):=\frac{u\left(5 \rho_{i}\left(x+x_{i}\right), \frac{\lambda^{2}}{B(\lambda)}\left(5 \rho_{i}\right)^{2}\left(t+t_{i}\right)\right)}{\lambda 5 \rho_{i}}, \\
& \mathbf{f}_{\lambda}^{i}(x, t):=\frac{\mathbf{f}\left(5 \rho_{i}\left(x+x_{i}\right), \frac{\lambda^{2}}{B(\lambda)}\left(5 \rho_{i}\right)^{2}\left(t+t_{i}\right)\right)}{\lambda},
\end{aligned}
$$

$$
\begin{gathered}
A_{\lambda}^{i}(x, t):=A\left(5 \rho_{i}\left(x+x_{i}\right), \frac{\lambda^{2}}{B(\lambda)}\left(5 \rho_{i}\right)^{2}\left(t+t_{i}\right)\right), \\
a_{\lambda}(t):=\frac{a(\lambda t)}{\frac{B(\lambda)}{\lambda^{2}}}, \quad b_{\lambda}(t):=t a_{\lambda}(t) \quad \text { and } \quad B_{\lambda}(t):=\int_{0}^{t} b_{\lambda}(\tau) d \tau=\frac{B(\lambda t)}{B(\lambda)},
\end{gathered}
$$

which implies that

$$
B_{\lambda}(1)=\frac{B(\lambda)}{B(\lambda)}=1 \quad \text { and } \quad B_{\lambda}(t) \quad \text { satisfies (2.7). }
$$

From the definitions of $u_{\lambda}^{i}, \mathbf{f}_{\lambda}^{i}, A_{\lambda}^{i}$ and $a_{\lambda}$, we find that $u_{\lambda}^{i}$ is a local weak solution of

$$
\left(u_{\lambda}^{i}\right)_{t}-\operatorname{div}\left(a_{\lambda}\left(\left(A_{\lambda}^{i}(x, t) \nabla u_{\lambda}^{i} \cdot \nabla u_{\lambda}^{i}\right)^{\frac{1}{2}}\right) A_{\lambda}^{i}(x, t) \nabla u_{\lambda}^{i}\right)=\operatorname{div}\left(a_{\lambda}\left(\left|\mathbf{f}_{\lambda}^{i}\right|\right) \mathbf{f}_{\lambda}^{i}\right) \quad \text { in } Q_{2} .
$$

Without loss of generality we may as well assume that $R=2$ in Definition 1.1 by a scaling argument. Consequently, from Definition 1.1 and Lemma 2.15 we conclude that

$$
\left\{f_{Q_{2}}\left[B_{\lambda}\left(\left|\nabla u_{\lambda}^{i}\right|\right)\right]^{q_{2}} d z\right\}^{\frac{1}{q_{2}}} \leq C, \quad\left\{f_{Q_{2}}\left[B_{\lambda}\left(\left|\mathbf{f}_{\lambda}^{i}\right|\right)\right]^{q_{2}} d z\right\}^{\frac{1}{q_{2}}} \leq C \delta
$$

and

$$
f_{Q_{2}}\left|A_{\lambda}^{i}(x, t)-\overline{A_{\lambda Q_{2}}^{i}}\right| d z \leq \delta
$$

for any $j=1,2$ and $\lambda \geq \lambda_{*}$. Then according to Lemma 2.17, we conclude that

$$
\left|\left\{z \in Q_{1}: B_{\lambda}\left(\left|\nabla u_{\lambda}^{i}\right|\right)>N_{1}\right\}\right| \leq C \epsilon\left|Q_{1}\right| .
$$

Then by way of changing variables, we recover the claim.
2. Now we find that

$$
\begin{aligned}
& \int_{\left\{z \in Q_{i}^{1}: B(|\nabla u|)>N_{1} B(\lambda)\right\}} w(x) d x d t \\
& \quad \leq \int_{t_{i}-25 \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}}^{t_{i}+25 \frac{\lambda^{2}}{(2)} \rho_{i}^{2}} \int_{E\left(B_{5 \rho_{i}}\left(x_{i}\right), t\right)} w(x) d x d t=\int_{t_{i}-25 \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}}^{t_{i}+25 \frac{\lambda^{2}}{(\lambda)} \rho_{i}^{2}} w\left(E\left(B_{5 \rho_{i}}\left(x_{i}\right), t\right)\right) d t,
\end{aligned}
$$

where $E\left(B_{5 \rho_{i}}\left(x_{i}\right), t\right):=\left\{x \in B_{5 \rho_{i}}\left(x_{i}\right): B(|\nabla u(x, t)|)>N_{1} B(\lambda)\right\}$, which implies that

$$
\begin{aligned}
& \frac{\int_{\left\{z \in Q_{i}^{1}: B(|\nabla u|)>N_{1} B(\lambda)\right\}} w(x) d x d t}{50 w\left(B_{5 \rho_{i}}\left(x_{i}\right)\right) \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}} \\
& \leq \frac{1}{50 \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}} \int_{t_{i}-25 \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}}^{t_{i}+25 \frac{\lambda^{2}}{B(\lambda)} \rho_{2}^{2}} \frac{w\left(E\left(B_{5 \rho_{i}}\left(x_{i}\right), t\right)\right)}{w\left(B_{5 \rho_{i}}\left(x_{i}\right)\right)} d t \\
& \leq \frac{1}{50 \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}} \int_{t_{i}-25 \frac{\lambda^{2}}{t_{i}(\lambda)} \rho_{i}^{2}}^{t_{i}+25 \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}}\left(\frac{E\left(B_{5 \rho_{i}}\left(x_{i}\right), t\right)}{\left|B_{5 \rho_{i}}\left(x_{i}\right)\right|}\right)^{\alpha} d t \\
& \leq \frac{1}{\left[50 \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}\right]^{\alpha}}\left[\int_{t_{i}-25 \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}}^{t_{i}+25 \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2}} \frac{E\left(B_{5 \rho_{i}}\left(x_{i}\right), t\right)}{\left|B_{5 \rho_{i}}\left(x_{i}\right)\right|} d t\right]^{\alpha}=\left[\frac{\left|\left\{z \in Q_{i}^{1}: B(|\nabla u|)>N_{1} B(\lambda)\right\}\right|}{\left|Q_{i}^{1}\right|} \leq C \epsilon^{\alpha},\right.
\end{aligned}
$$

where we have used Lemma 2.9, Hölder's inequality and (2.26). Therefore, we conclude that

$$
\int_{\left\{z \in Q_{i}^{1}: B(|\nabla u|)>N_{1} B(\lambda)\right\}} w(x) d x d t \leq C \epsilon^{\alpha} w\left(B_{5 \rho_{i}}\left(x_{i}\right)\right) \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2} \leq C \epsilon^{\alpha} w\left(B_{\rho_{i}}\left(x_{i}\right)\right) \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2} .
$$

So, we can deduce the following result from Lemma 2.13 (3), the fact that the cylinders $\left\{Q_{i}^{0}\right\}$ are disjoint and Lemma 2.14

$$
\begin{aligned}
&\left.\int_{\left\{z \in Q_{1}\right.}: B(|\nabla u|)>N_{1} B(\lambda)\right\} \\
& \leq \sum_{i \in \mathbb{N}} \int_{\left\{z \in Q_{i}^{1}: B(|\nabla u|)>N_{1} B(\lambda)\right\}} w(x) d x d t \\
& \leq C \epsilon^{\alpha} \sum_{i \in \mathbb{N}} w\left(B_{\rho_{i}}\left(x_{i}\right)\right) \frac{\lambda^{2}}{B(\lambda)} \rho_{i}^{2} \\
& \leq \sum_{i \in \mathbb{N}} \frac{C \epsilon^{\alpha}}{[B(\lambda)]^{q_{0}}} \int_{\left\{z \in Q_{i}^{0}: B(|\nabla u|)>\frac{B(\lambda)}{4}\right\}}[B(|\nabla u|)]^{q_{0}} w(x) d x d t \\
&+\sum_{i \in \mathbb{N}} \frac{C \epsilon^{\alpha}}{\delta q_{0}}[B(\lambda)]^{q_{0}} \\
& \int_{\left\{z \in Q_{i}^{0}: B(| || |)>\frac{B(\lambda)}{4 \delta}\right\}}[B(|f|)]^{q_{0}} w(x) d x d t \\
& \leq \frac{C \epsilon^{\alpha}}{[B(\lambda)]^{q_{0}}} \int_{\left\{z \in Q_{2}: B(|\nabla u|)>\frac{B(\lambda)}{4}\right\}}[B(|\nabla u|)]^{q_{0}} w(x) d x d t \\
&+\frac{C \epsilon^{\alpha}}{\delta q_{0}}[B(\lambda)]^{q_{0}} \\
& \int_{\left\{z \in Q_{2}: B(| || |)>\frac{B(\lambda)}{40}\right\}}[B(|\mathbf{f}|)]^{q_{0}} w(x) d x d t .
\end{aligned}
$$

Thus, we finish the proof.
In the following it is sufficient to consider the proof of Theorem 1.4 as an a priori estimate, therefore assuming a priori that $B(|\nabla u|) \in L^{q}\left(0, T ; L_{w}^{q}(\Omega)\right)$ locally. This assumption can be removed by an approximation argument like the one in [1,21]. Now we are ready to prove the main result, Theorem 1.4.

Proof. In light of Lemma 2.12, we find

$$
\begin{aligned}
\int_{Q_{1}} & {[B(|\nabla u|)]^{q} w(x) d x d t } \\
= & q N_{1}^{q} \int_{0}^{+\infty}[B(\lambda)]^{q-1} \int_{\left\{z \in Q_{1}: B(|\nabla u|)>N_{1} B(\lambda)\right\}} w(x) d x d t d B(\lambda) \\
= & q N_{1}^{q} \int_{0}^{\lambda_{*}}[B(\lambda)]^{q-1} \int_{\left\{z \in Q_{1}: B(|\nabla u|)>N_{1} B(\lambda)\right\}} w(x) d x d t d B(\lambda) \\
& +q N_{1}^{q} \int_{\lambda_{*}}^{+\infty}[B(\lambda)]^{q-1} \int_{\left\{z \in Q_{1}: B(|\nabla u|)>N_{1} B(\lambda)\right\}} w(x) d x d t d B(\lambda) \\
= & J_{1}+J_{2} .
\end{aligned}
$$

Estimate of $J_{1}$. Recalling the definitions of $\lambda_{*}$ and $\lambda_{0}$, we estimate

$$
J_{1} \leq C\left\{B\left(\left[\left\{\int_{Q_{2}}[B(|\nabla u|)]^{q_{0}} w(x) d z\right\}^{\frac{1}{q_{0}}}+\frac{1}{\delta}\left\{\int_{Q_{2}}[B(|\mathbf{f}|)]^{q_{0}} w(x) d z\right\}^{\frac{1}{q_{0}}}+1\right]^{\frac{1}{2}}\right)\right\}^{q}
$$

$$
\begin{aligned}
& \leq C\left[\left\{\int_{Q_{2}}[B(|\nabla u|)]^{q_{0}} w(x) d z\right\}^{\frac{1}{q_{0}}}+\frac{1}{\delta}\left\{\int_{Q_{2}}[B(|\mathbf{f}|)]^{q_{0}} w(x) d z\right\}^{\frac{1}{q_{0}}}+1\right]^{\frac{q\left(2+s_{a}\right)}{2}} \\
& \leq C\left[\left\{\int_{Q_{2}}[B(|\nabla u|)]^{q_{0}} w(x) d z\right\}^{\frac{1}{q_{0}}}+\left\{\int_{Q_{2}}[B(|\mathbf{f}|)]^{q} w(x) d z\right\}^{\frac{1}{q}}+1\right]^{\frac{q\left(2+s_{a}\right)}{2}}
\end{aligned}
$$

where $C=C\left(n, i_{a}, s_{a}, q, \delta, w\right)$, since

$$
\int_{Q_{2}}[B(|\mathbf{f}|)]^{q_{0}} w(x) d z=\int_{Q_{2}}[B(|\mathbf{f}|)]^{1+\epsilon} w(x)^{\frac{1+\varepsilon}{q}} w(x)^{\frac{q-1-\epsilon}{q}} d z \leq C\left(\int_{Q_{2}}[B(|\mathbf{f}|)]^{q} w(x) d z\right)^{\frac{1+\varepsilon}{q}}
$$

by using Hölder's inequality and the fact that $q_{0}=1+\epsilon$.
Estimate of $\mathrm{J}_{2}$. Now we apply Lemma 2.12 and Lemma 2.18 to find that

$$
\begin{aligned}
J_{2} \leq & C \epsilon^{\alpha}\left\{\int_{0}^{\infty}[B(\lambda)]^{q-q_{0}-1} \int_{\left\{z \in Q_{2}: B(|\nabla u|)>\frac{B(\lambda)}{4}\right\}}[B(|\nabla u|)]^{q_{0}} w(x) d x d t d B(\lambda)\right. \\
& \left.+\frac{1}{\delta^{q_{0}}} \int_{0}^{\infty}[B(\lambda)]^{q-q_{0}-1} \int_{\left\{z \in Q_{2}: B(|f|)>\frac{B(\lambda)}{40}\right\}}[B(|\mathbf{f}|)]^{q_{0}} w(x) d x d t d B(\lambda)\right\} \\
\leq & C_{1} \epsilon^{\alpha} \int_{Q_{2}}[B(|\nabla u|)]^{q} w(x) d z+C_{2} \int_{Q_{2}}[B(|\mathbf{f}|)]^{q} w(x) d z
\end{aligned}
$$

where $C_{1}=C_{1}\left(n, i_{a}, s_{a}, q, \Lambda\right)$ and $C_{2}=C_{2}\left(n, i_{a}, s_{a}, q, \Lambda, \epsilon, \delta\right)$. Therefore, we combine the estimates of $J_{1}$ and $J_{2}$ to obtain

$$
\begin{aligned}
& \int_{Q_{1}}[B(|\nabla u|)]^{q} w(x) d z \\
& \leq C_{3}\left[\left\{\int_{Q_{2}}[B(|\nabla u|)]^{q_{0}} w(x) d z\right\}^{\frac{1}{q_{0}}}+\left\{\int_{Q_{2}}[B(|\mathbf{f}|)]^{q} w(x) d z\right\}^{\frac{1}{q}}+1\right]^{\frac{q\left(2+s_{a}\right)}{2}} \\
&+C_{1} \epsilon^{\alpha} \int_{Q_{2}}[B(|\nabla u|)]^{q} w(x) d z
\end{aligned}
$$

where $C_{3}=C_{3}\left(n, i_{a}, s_{a}, q, \Lambda, \epsilon, \delta, w\right)$. Selecting proper $\epsilon>0$ small enough and using a covering and iteration argument, we can finish the proof of Theorem 1.4.

## Statements

All data generated or analysed during this study are included in this article. Moreover, the author states that there is no conflict of interest.

## References

[1] E. Acerbi, G. Mingione, Gradient estimates for a class of parabolic systems, Duke Math. J. 136(2007), No. 2, 285-320. https://doi.org/10.1215/S0012-7094-07-13623-8; MR2286632
[2] R. A. Adams, J. J. F. Fournier, Sobolev spaces (2nd edition), Elsevier/Academic Press, Amsterdam, 2003. MR2424078
[3] K. Adimurthi, S. Byun, Gradient weighted estimates at the natural exponent for quasilinear parabolic equations, Adv. Math. 348(2019), 456-511. https://doi.org/10.1016/j. aim.2019.03.015; MR3928653
[4] S. Baasandorj, S. Byun, J. Oh, Calderón-Zygmund estimates for generalized double phase problems, J. Funct. Anal. 279(2020), No. 7, 108670, 57 pp. https://doi.org/10. 1016/j.jfa.2020.108670; MR4107816
[5] P. Baroni, Lorentz estimates for degenerate and singular evolutionary systems, J. Differential Equations 255(2013), No. 9, 2927-2951. https://doi.org/10.1016/j. jde.2013.07. 024; MR3090083
[6] P. Baroni, V. Bögelein, Calderón-Zygmund estimates for parabolic $p(x, t)$-Laplacian systems, Rev. Mat. Iberoam. 30(2014), No. 4, 1355-1386. https://doi.org/10.4171/RMI/ 817; MR3293436
[7] P. Baroni, C. Lindfors, The Cauchy-Dirichlet problem for a general class of parabolic equations, Ann. Inst. H. Poincaré Anal. Non Linéaire 34(2017), No. 3, 593-624. https:// doi.org/10.1016/j.anihpc.2016.03.003; MR3633737
[8] L. Beck, G. Mingione, Lipschitz bounds and nonuniform ellipticity, Comm. Pure Appl. Math. 73(2020), No. 5, 944-1034. https://doi.org/10.1002/cpa.21880; MR4078712
[9] V. Bögelein, Global Calderón-Zygmund theory for nonlinear parabolic systems, Calc. Var. Partial Differential Equations 51(2014), No. 3-4, 555-596. https://doi.org/10.1007/ s00526-013-0687-4; MR3268863
[10] V. Bögelein, F. Duzaar, Hölder estimates for parabolic $p(x, t)$-Laplacian systems, Math. Ann. 354(2012), No. 3, 907-938. https://doi.org/10.1007/s00208-011-0750-4; MR2983073
[11] V. Bögelein, F. Duzaar, G. Mingione, Degenerate problems with irregular obstacles, J. Reine Angew. Math. 650(2011), 107-160. https://doi.org/10.1515/CRELLE.2011.006; MR2770559
[12] V. Bögelein, F. Duzaar, G. Mingione, The regularity of general parabolic systems with degenerate diffusion, Mem. Amer. Math. Soc. 221(2013), No. 1041, vi+143 pp. https:// doi.org/10.1090/S0065-9266-2012-00664-2; MR3059060
[13] T. Bui, X. Duong, Global Lorentz estimates for nonlinear parabolic equations on nonsmooth domains, Calc. Var. Partial Differential Equations 56(2017), No. 2, Art. 47, 24 pp. https://doi.org/10.1007/s00526-017-1130-z; MR3625811
[14] J. Burczak, P. Kaplický, Interior regularity of space derivatives to an evolutionary, symmetric $\varphi$-Laplacian, Monatsh. Math. 183(2017), No. 1, 71-101. https://doi .org/10.1007/ s00605-016-1005-y; MR3634347
[15] S. Byun, Y. Сно, Nonlinear gradient estimates for generalized elliptic equations with nonstandard growth in nonsmooth domains, Nonlinear Anal. 140(2016), 145-165. https: //doi.org/10.1016/j.na.2016.03.016; MR3492733
[16] S. Byun, W. Kim, M. Lim, Higher integrability near the initial boundary for nonhomogeneous parabolic systems of $p$-Laplacian type, Forum Math. 32(2020), No. 6, 1539-1559. https://doi.org/10.1515/forum-2020-0068; MR4168703
[17] S. Byun, J. Ок, S. Ryu, Global gradient estimates for general nonlinear parabolic equations in nonsmooth domains, J. Differential Equations 254(2013), No. 11, 4290-4326. https:// doi.org/10.1016/j.jde.2013.03.004; MR3035434
[18] S. Byun, D. K. Palagachev, S. Ryu, Weighted $W^{1, p}$ estimates for solutions of non-linear parabolic equations over non-smooth domains, Bull. Lond. Math. Soc. 45(2013), No. 4, 765-778. https://doi.org/10.1112/blms/bdt011; MR3081545
[19] S. Byun, D. K. Palagachev, P. Shin, Optimal regularity estimates for general nonlinear parabolic equations, Manuscripta Math. 162(2020), No. 1-2, 67-98. https://doi.org/10. 1007/s00229-019-01127-8; MR4079798
[20] S. Byun, J. Park, P. Shin, Global regularity for degenerate/singular parabolic equations involving measure data, Calc. Var. Partial Differential Equations 60(2021), No. 1, Paper No. 18, 32 pp. https://doi.org/10.1007/s00526-020-01906-2; MR4201641
[21] S. Byun, S. Ryu, Global weighted estimates for the gradient of solutions to nonlinear elliptic equations, Ann. Inst. H. Poincaré Anal. Non Linéaire 30(2013), No. 2, 291-313. https : //doi.org/10.1016/j.anihpc.2012.08.001; MR3035978
[22] L. A. Caffarelli, I. Peral, On $W^{1, p}$ estimates for elliptic equations in divergence form, Comm. Pure Appl. Math. 51(1998), No. 1, 1-21. https://doi.org/10.1002/(SICI) 1097-0312(199801)51:1<1::AID-CPA1>3.3.CO;2-N; MR1486629
[23] D. Cao, T. Mengesha, T. Phan, Weighted- $W^{1, p}$ estimates for weak solutions of degenerate and singular elliptic equations, Indiana Univ. Math. J. 67(2018), No. 6, 2225-2277. https://doi.org/10.1512/iumj.2018.67.7533; MR3900368
[24] Y. Chen, L. Wu, Second order elliptic equations and elliptic systems, American Mathematical Society, Providence, RI, 1998. https://doi.org/10.1090/mmono/174; MR1616087
[25] I. Chlebicka, A. Zatorska-Goldstein, Existence to nonlinear parabolic problems with unbounded weights, J. Evol. Equ. 19(2019), No. 1, 1-19. https://doi.org/10.1007/ s00028-018-0465-z; MR3918514
[26] A. Cianchi, V. Maz'ya, Global Lipschitz regularity for a class of quasilinear elliptic equations, Comm. Partial Differential Equations 36(2011), No. 1, 100-133. https://doi. org/10. 1080/03605301003657843; MR2763349
[27] A. Cianchi, V. Maz'ya, Global boundedness of the gradient for a class of nonlinear elliptic systems, Arch. Ration. Mech. Anal. 212(2014), No. 1, 129-177. https://doi.org/ 10.1007/s00205-013-0705-x; MR3162475
[28] A. Cianchi, V. Maz'ya, Gradient regularity via rearrangements for $p$-Laplacian type elliptic boundary value problems, J. Eur. Math. Soc. 16(2014), No. 3, 571-595. https: //doi.org/10.4171/JEMS/440; MR3165732
[29] E. DiBenedetto, Degenerate parabolic equations, Universitext. Springer-Verlag, New York, 1993. https://doi.org/10.1007/978-1-4612-0895-2; MR1230384
[30] L. Diening, F. Ettwein, Fractional estimates for non-differentiable elliptic systems with general growth, Forum Math. 20(2008), No. 3, 523-556. https ://doi .org/10.1515/FORUM. 2008.027; MR2418205
[31] L. Diening, T. Scharle, S. Schwarzacher, Regularity for parabolic systems of Uhlenbeck type with Orlicz growth, J. Math. Anal. Appl. 472(2019), No. 1, 46-60. https: //doi.org/10.1016/j.jmaa.2018.10.055; MR3906361
[32] L. Diening, B. Stroffolini, A. Verde, The $\phi$-harmonic approximation and the regularity of $\phi$-harmonic maps, J. Differential Equations 253(2012), No. 7, 1943-1958. https://doi. org/10.1016/j.jde.2012.06.010; MR2946961
[33] H. Dong, T. Phan, Regularity for parabolic equations with singular or degenerate coefficients, Calc. Var. Partial Differential Equations 60(2021), No. 1, Paper No. 44, 39 pp. https://doi.org/10.1007/s00526-020-01876-5; MR4204570
[34] M. Giaquinta, Multiple integrals in the calculus of variations and nonlinear elliptic systems, Princeton University Press, 1983. MR717034
[35] R. Giri, J. Kinnunen, K. Moring, Supercaloric functions for the parabolic $p$-Laplace equation in the fast diffusion case, NoDEA Nonlinear Differential Equations Appl. 28(2021), No. 3, Paper No. 33, 21 pp. https://doi.org/10.1007/s00030-021-00694-8; MR4244932
[36] J. Kinnunen, J. L. Lewis, Higher integrability for parabolic systems of $p$-Laplacian type, Duke Math. J. 102(2000), No. 2, 253-271. https://doi.org/10.1215/S0012-7094-00-10223-2; MR1749438
[37] G. M. Lieberman, The natural generalization of the natural conditions of Ladyzhenskaya and Ural'tseva for elliptic equations, Comm. Partial Differential Equations 16(1991), No. 2-3, 311-361. https://doi.org/10.1080/03605309108820761; MR1104103
[38] C. Lindfors, Obstacle problem for a class of parabolic equations of generalized $p$ Laplacian type, J. Differential Equations 261(2016), No. 10, 5499-5540. https://doi.org/ 10.1016/j.jde.2016.08.015; MR3548260
[39] T. Mengesha, N. Phuc, Weighted and regularity estimates for nonlinear equations on Reifenberg flat domains, J. Differential Equations 250(2011), No. 5, 2485-2507. https:// doi.org/10.1016/j.jde.2010.11.009; MR2756073
[40] T. Mengesha, N. Phuc, Global estimates for quasilinear elliptic equations on Reifenberg flat domains, Arch. Ration. Mech. Anal. 203(2012), No. 1, 189-216. https://doi.org/10. 1007/s00205-011-0446-7; MR2864410
[41] G. Mingione, V. Rǎdulescu, Recent developments in problems with nonstandard growth and nonuniform ellipticity, J. Math. Anal. Appl. 501(2021), No. 1, 125197, 41 pp. https: //doi.org/10.1016/j.jmaa.2021.125197; MR4258810
[42] M. Misawa, $L^{q}$ estimates of gradients for evolutional $p$-Laplacian systems, J. Differential Equations 219(2006), No. 2, 390-420. https://doi.org/10.1016/j.jde.2004.11.003; MR2183266
[43] E. Stein, Harmonic analysis, Princeton University Press, Princeton, NJ, 1993. MR1232192
[44] A. Torchinsкy, Real-variable methods in harmonic analysis, Pure and Applied Mathematics, Vol. 123, Academic Press, Inc., Orlando, FL, 1986. MR869816
[45] S. Yang, D. Chang, D. Yang, W. Yuan, Weighted gradient estimates for elliptic problems with Neumann boundary conditions in Lipschitz and (semi-)convex domains, J. Differential Equations 268(2020), No. 6, 2510-2550. https://doi.org/10.1016/j.jde.2019.09. 036; MR4047962
[46] F. Yao, Calderón-Zygmund estimates for a class of quasilinear parabolic equations, Nonlinear Anal. 176(2018), 84-99. https://doi.org/10.1016/j.na.2018.06.008; MR3856719
[47] F. Yao, Lorentz estimates for a class of nonlinear parabolic systems, J. Differential Equations 266(2019), No. 4, 2078-2099. https://doi.org/10.1016/j.jde.2018.08.021; MR3906241
[48] F. Yao, S. Zhou, Calderón-Zygmund estimates for a class of quasilinear elliptic equations, J. Funct. Anal. 272(2017), No. 4, 1524-1552. https://doi.org/10.1016/j.jfa.2016.11. 008; MR3590245
[49] C. Zhang, Global weighted estimates for the nonlinear parabolic equations with nonstandard growth, Calc. Var. Partial Differential Equations 55(2016), No. 5, Art. 109, 27 pp. https://doi.org/10.1007/s00526-016-1066-8; MR3549067


[^0]:    ${ }^{\boxtimes}$ Email: yfp@shu.edu.cn

