Self-improving property of degenerate parabolic equations of porous medium-type

Ugo Gianazza, Sebastian Schwarzacher

American Journal of Mathematics, Volume 141, Number 2, April 2019, pp. 399-446 (Article)

Published by Johns Hopkins University Press

DOI: https://doi.org/10.1353/ajm.2019.0009

For additional information about this article
https://muse.jhu.edu/article/718877/summary
SELF-IMPROVING PROPERTY OF DEGENERATE PARABOLIC EQUATIONS OF POROUS MEDIUM-TYPE

By Ugo Gianazza and Sebastian Schwarzacher

Abstract. We show that the gradient of solutions to degenerate parabolic equations of porous medium-type satisfies a reverse Hölder inequality in suitable intrinsic cylinders. We modify the by-now classical Gehring lemma by introducing an intrinsic Calderón-Zygmund covering argument, and we are able to prove local higher integrability of the gradient of a proper power of the solution $u$.

1. Introduction and main result. The aim of this paper is to study regularity properties of the gradient of non-negative solutions to non-linear, parabolic, partial differential equations, whose prototype is the porous medium equation

\[
(PME) \quad u_t - \Delta u^m = u_t - \text{div} \left( mu^{m-1} Du \right) = 0 \quad m > 0.
\]

When $m = 1$, the non-linear behavior disappears and (PME) reduces to the standard heat equation. When $m \neq 1$, the equation is quasi-linear and its modulus of ellipticity is $u^{m-1}$. When $m > 1$, this quantity vanishes on the set $[u = 0]$, the time evolution dominates over the diffusion process (slow diffusion case) and the equation is said to be degenerate. When $0 < m < 1$ the equation becomes singular, since the modulus of ellipticity blows up as $u \to 0$; in this case we have the so-called fast diffusion equation.

Equations of this form arise from applications, for example in modelling the flow of an isotropic gas in a porous medium or in studying the heat radiation of plasmas. For $m > 1$ the non-linear heat transfer has a finite speed of propagation, i.e., if the initial datum has a finite support, so does the solution for any positive time. Therefore, it naturally appears in numerous physical models, whenever the assumption of constancy of the thermal conductivity (respectively diffusivity) cannot be sustained. Besides the previously mentioned examples, this is also the case of models of population dynamics (where $u$ describes the concentration of the species) and of the theory of lubrication and boundary layers.

From the mathematical point of view, the understanding of local behavior of solutions to such equations plays a role in the $C^{1,\alpha}$-regularity theory for systems of $p$-Laplacian type, see for example [DiB93]. As can be seen from the so-called Barenbatt fundamental solution (we refer for example to [Bare52]), in general solutions to porous medium equations are considerably less regular than solutions to
the parabolic $p$-Laplacian, particularly with respect to gradient estimates, as it shall be discussed with more details below.

In this paper we will deal only with the degenerate situation $m > 1$, and we will study a general class of equations which have the same structure as (PME).

Given a bounded, open set $E \subset \mathbb{R}^n$ with $n \geq 2$, and $T > 0$, let $E_T \equiv E \times (0, T)$.

For $m > 1$ and a positive right-hand side $f \in L^{m+1}_{\text{loc}}(E_T)$, we will consider non-negative solutions to

\begin{equation}
\tag{1.1}
    u_t - \text{div} A(x, t, u, Du) = f \quad \text{weakly in } E_T.
\end{equation}

The vector field

\[ A : E_T \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n \]

is only assumed to be measurable, and we suppose there exist constants $0 < \nu \leq L < \infty$ such that

\begin{equation}
\tag{1.2}
    A(x, t, u, Du) \cdot Du \geq \nu m u^{m-1} |Du|^2 \quad \text{for a.e. } (x, t) \in E_T
\end{equation}

As discussed in [DGV, Chapter 3, Section 5], the structure conditions (1.2) are not sufficient to characterize parabolic partial differential equations. The partial differential equation in (1.1) is parabolic if it satisfies (1.2) and in addition, for every weak, local sub(super)-solution $u$ (see the precise definitions below), for all $k \in \mathbb{R}$, the truncations $(u - k)_+$ and $-(u - k)_-$ are weak, local sub(super)-solutions to (1.1), in the sense of (1.4)–(1.6) below, with $A(x, t, u, Du)$ replaced by

\[ A(x, t, k \pm (u - k)\pm, \pm D(u - k)\pm). \]

In [DGV, Chapter 3, Section 5] the following result is stated and proved.

**Lemma 1.1.** Assume that for all $(x, t, u) \in E_T \times \mathbb{R}$

\begin{equation}
\tag{1.3}
    A(x, t, u, \eta) \cdot \eta \geq 0 \quad \text{for all } \eta \in \mathbb{R}^n.
\end{equation}

Then (1.1)–(1.2) is parabolic.

Henceforth we will assume that the principal part $A(x, t, u, Du)$ satisfies (1.3) too, so that (1.1)–(1.2) is parabolic.

As we have already mentioned, the model problem (PME) corresponds to the case $\nu = L = 1$ and $f \equiv 0$.

1.1. **Weak solutions and sub(super)-solutions.** A function

\begin{equation}
\tag{1.4}
    u \in C^0_{\text{loc}}(0, T; L^2_{\text{loc}}(E)) \quad \text{with} \quad u^{m+1}_{m+1} \in L^2_{\text{loc}}\left(0, T; W^{1,2}_{\text{loc}}(E)\right)
\end{equation}
is a local, weak sub(super)-solution to (1.1)–(1.2) if satisfies the integral identity

$$\int_{E_T} -u \varphi_t + A(x, t, u, Du) \cdot D\varphi \, dx \, dt \leq \left( \geq \right) \int_{E_T} f \varphi \, dx \, dt$$

for all possible choices of non-negative test functions $\varphi \in C_0^\infty(E_T)$. This guarantees that all the integrals in (1.5) are convergent.

A local, weak solution to (1.1)–(1.2) is both a sub- and a super-solution, i.e., it satisfies the integral identity

$$\int_{E_T} -u \varphi_t + A(x, t, u, Du) \cdot D\varphi \, dx \, dt = \int_{E_T} f \varphi \, dx \, dt$$

for all possible choices of test functions $\varphi \in C_0^\infty(E_T)$. Moreover, we talk of homogeneous equations, whenever $f \equiv 0$.

By a standard mollification argument, it is possible to use the solution $u$ as test function. Let $\zeta : \mathbb{R} \to \mathbb{R}$,

$$\zeta(s) \equiv \begin{cases} 
C \exp \left( \frac{1}{|s|^2 - 1} \right) & |s| < 1 \\
0 & |s| \geq 1 
\end{cases}$$

be the standard mollifier ($C$ is chosen in order to have $\| \zeta \|_{L^1(\mathbb{R})} = 1$) and define the family

$$\zeta^\varepsilon(s) = \frac{1}{\varepsilon} \zeta \left( \frac{s}{\varepsilon} \right), \quad \varepsilon > 0.$$ 

Since we need a time regularization, given $\varphi \in C_0^\infty(E_T)$, we consider the family of mollifiers $\{ \zeta^\varepsilon \}$, with

$$\varepsilon < \text{dist} (\text{spt} \varphi, E_T),$$

and we set

$$\varphi^\varepsilon(x, t) = (\varphi \ast \zeta^\varepsilon)(x, t) = \int_{\mathbb{R}} \varphi(x, t - s) \zeta^\varepsilon(s) \, ds.$$ 

We insert $\varphi^\varepsilon$ as test function in (1.6), change variables and apply Fubini's theorem to obtain

$$\int_{E_T} -u^\varepsilon \varphi_t + A^\varepsilon(x, t, u, Du) \cdot D\varphi \, dx \, dt = \int_{E_T} f^\varepsilon \varphi \, dx \, dt,$$

where the subscript in $u^\varepsilon, f^\varepsilon$, and $A^\varepsilon$ denotes the mollification with respect to time.

We conclude this introductory section with our main result.
THEOREM 1.2. (Local higher integrability) Let $u \geq 0$ be a local, weak solution to (1.1)–(1.2) in $E_T$ for $m > 1$. Then, there exists $\varepsilon_0 > 0$, depending only on $n$, $m$, $\nu$, and $L$ of (1.2), such that

$$u^{\frac{m+1}{2}} \in L^{2+\varepsilon}_{loc}(0,T;W^{1,2+\varepsilon}_{loc}(E)) \quad \forall \varepsilon \in (0,\varepsilon_0].$$

Theorem 1.2 is a straightforward consequence of local quantitative estimates. We provide two different versions, a first one for standard parabolic cylinders $B_r(x_o) \times (t_o - r^2, t_o]$ (see Theorem 5.10), and a second version on the so-called intrinsic cylinders (see below), which inherit the natural scaling properties of the solution (see Theorem 5.9).

1.2. Novelty and significance. As apparent from the statement of Theorem 1.2, we are interested in the order of integrability of $|Du|^{\frac{m+1}{2}}$. For elliptic equations and systems, Meyers & Elcrat [ME75] showed that the gradients of solutions locally belong to a slightly higher Sobolev space than expected a priori. The main tools are a reverse Hölder inequality for $|Du|$ and an application of Gehring’s lemma (see the original paper [Geh73] and also [GM79, Str80]). The method works for equations with $p$-growth, hence degenerate and singular elliptic equations of $p$-Laplacian type are allowed.

Giaquinta & Struwe [GS82] extended the elliptic, local, higher integrability result to parabolic equations. However, in their work, in order to derive the reverse Hölder inequality, the diffusion term $A \approx Du$, i.e., it is forced to have a linear growth with respect to $|Du|$, so that degenerate and singular equations are ruled out.

The main obstruction to the extension to the degenerate/singular setting is given by the lack of homogeneity in the energy estimates. This problem can be overcome by using the so-called intrinsic parabolic geometry, that is a scaling, which depends on the solution itself. Under a more physical point of view, the diffusion process evolves at a time scale which depends instant by instant on $u$ itself; the homogeneity is recovered, once the time variable is rescaled by a factor that depends on the solution in a suitable way. This approach was first developed by DiBenedetto & Friedman [DBF85, DiB93] in the context of the parabolic $p$-Laplace equation

$$u \in C^0_{loc}(0,T;L^2_{loc}(E)) \cap L^p_{loc}(0,T;W^{1,p}_{loc}(E)) \quad p > 1.$$

(1.8)

Localisation with respect to intrinsically scaled cylinders, was a key tool to prove the Hölder continuity of the gradients, for smooth right-hand sides $f$.

Later on, by rephrasing these ideas in the context of intrinsic Calderón-Zygmund coverings, Kinnunen & Lewis [KL00] showed that gradients of solutions
to equations with the same structure as (1.8) enjoy a higher integrability property, namely

\[ Du \in L_{\text{loc}}^{p+\varepsilon}(E_T), \quad \text{for some } \varepsilon > 0. \]

This result holds under very general structural assumptions on the operator, and minimal conditions on the right-hand side. The values of \( p \) cover the full degenerate range \( p > 2 \), but are restricted to the super-critical singular range \( \frac{2n}{n+2} < p < 2 \). This restriction on \( p \) in the singular range is a recurrent feature, as discussed at length, for example in [DGV, Appendix B], or in [AM07]. It is noteworthy that, based on the local higher integrability result of the parabolic \( p \)-Laplacian, many applications follow. These include (without any ambition of completion) the full \( L^q \) theory and beyond [AM07, Sch13], as well as partial regularity results [BZM13], or pointwise estimates via potential theory for equations [KM14/2]. Summarizing, many different ways to show regularity for the gradient of solutions to the \( p \)-Laplacian are available which, among other benefits, has a natural impact on the regularity of the time derivative, as it was recently shown in [FSch15].

Taking into account the large amount of results, it might not seem too surprising that the adaption of the non-linear methods developed for the \( p \)-Laplace to the porous medium equation turns out to be more delicate than expected; indeed, the higher integrability result for the porous medium equations has been an open problem for some years.

To our knowledge, so far, the only existing, related contribution is due to Bénilan [Ben83]; he established an abstract result whose application to non-negative solutions of the porous medium equation yields \( \frac{\partial^2 u^m}{\partial x_i \partial x_j} \in L_{\text{loc}}^p(E_T) \) for \( 1 < p < 1 + 1/m \). Since solutions are bounded, by the Sobolev embedding theorem one obtains that

\[ Du^m \in L_{\text{loc}}^{\tilde{p}}(E_T), \quad \text{for } \tilde{p} = \frac{np}{n - p}. \]

Benilán’s solutions satisfy \( Du^m(\cdot, t) \in L^2(E) \) for almost every \( t \in (0, T) \), and therefore, the previous result amounts to a higher integrability estimate for

\[ n < \frac{2m + 2}{m - 1}. \]

The method makes no use of reverse Hölder inequalities or of Gehring’s Lemma, and does not apply to equations with the same generality as considered here: indeed, Theorem 1.2 covers a wide class of equations relying only on (1.2), and gives a quantitative knowledge (at least theoretically) of the higher integrability of the gradient of proper powers of local solutions. Moreover, the estimates developed here are of quantitative and local nature.

What are the main difficulties in the proof of the higher integrability for solutions to (1.1)–(1.2)? Why does it turn out that it is far from being a straightforward
extension of techniques already used when dealing with the analogous result for the $p$-Laplace equation? We conclude this section by shortly discussing these difficulties and the related technical novelties which allow us to overcome them.

The first step to higher integrability is a reverse Hölder estimate, which has always been the backbone to gain Gehring type higher integrability properties. The reverse Hölder estimate, which might be of independent interest, is stated in Proposition 3.4 for general parabolic cylinders.

However, this estimate alone does not allow us to conclude the desired higher integrability, in contrast to the $p$-Laplacian situation. Indeed, if $c$ is a positive constant, and $u$ is a local solution to the $p$-Laplace equation (1.8), then $u - c$ is a solution as well. Clearly, this is not the case for the porous medium equation and this simple difference makes it impossible to apply to the porous medium equation the approach known for the $p$-Laplacian. Indeed, the latter is based on (scaling invariant) Sobolev-Poincaré inequalities of Gagliardo-Nirenberg type in space time, which are invariant by a constant (i.e., the mean value).

We overcome this difficulty by splitting the problem into two cases: degenerate and non-degenerate regimes. This is a very common approach, going back to DeGiorgi. It was used to get estimates for solutions of PDEs in many different ways. For the $p$-Laplacian, besides the references already given and without pretending to list all the relevant contributions, we refer for example to [BZM13, BCDKS15, DLSV11, KM14/1, Sch13]. For solutions to (PME) it is a key tool to derive Harnack inequalities, as well as to prove Hölder continuity: just as an example, see [DGV]. However, as far as we can say, its use in the context of gradient estimates for the porous medium equation is a novelty.

For each case, namely the degenerate and the non-degenerate regime, we prove estimates on intrinsic cylinders, as stated respectively in Proposition 4.3 and 4.9. These are invariant under subtraction of constants, and therefore suitably tailored to our purposes. The proofs employ tools which are different in the two cases.

In the degenerate regime, roughly speaking, we use the fact that we have a control on the amount of oscillation which solutions can have. Conversely, in the non-degenerate case, we rely on the expansion of positivity for super-solutions, as stated in Theorem 4.5. Note that the use of Theorem 4.5 is the only (but crucial) place where we need the assumption of parabolicity, as well as the positivity of the right-hand side.

In Section 5 we prove the higher integrability via an argument of Gehring type. Note that it is not possible to use any standard reference in a straightforward manner. As a matter of fact, we need to establish a Calderón-Zygmund covering, using cylinders which are intrinsically scaled with respect to what seems to be the natural quantity here, namely $u^{m-1}$, where $u$ is the solution. On the other hand, since we want to estimate proper powers of $\left|Du^{m+1}\right|$, we have to develop something like
an intrinsic metric with respect to $u$, which can then be used for the Calderón-Zygmund analysis, independently of the PDE. In turn, this heavily relies on the possibility of adapting the construction of [Sch13] to the porous medium equation.

We think that this tool might be of independent interest in the analysis of degenerate partial differential equations, as it translates the intrinsic scaling into a sort of intrinsic distance.

Acknowledgments. S. Schwarzacher thanks program PRVOUK P47, financed by the Charles University in Prague. Both authors acknowledge the warm hospitality of the Institut Mittag-Leffler, where this research project started, during the program “Evolutionary problems” in the Fall 2013; the authors are also very grateful to the anonymous referee, for all the comments and suggestions, which greatly helped in improving the quality of the manuscript.

2. Preliminaries.

2.1. Notation. Consider a point $z_o = (t_o, x_o) \in \mathbb{R}^{n+1}$ and two parameters $\rho, \tau > 0$. The open ball with radius $\rho$ and center $x_o$ will be denoted by

$$B_\rho(x_o) \equiv \{ x \in E : |x - x_o| < \rho \}.$$

We define the time-space cylinder by

$$Q_{\tau,\rho}(z_o) \equiv (t_o - \tau, t_o + \tau) \times B_\rho(x_o).$$

As we prove local estimates, the reference point is never of importance, and we often omit it by writing $B_\rho$ and $Q_{\tau,\rho}$.

The symbol $|\cdot|$ stands for the Lebesgue measure, either in $\mathbb{R}^n$ or $\mathbb{R}^{n+1}$, and the dimension will be clear from the context.

When writing $a \sim b$ we mean that the ratio $a/b$ is bounded above and below by constants that depend only on the data.

Finally, whenever writing sup, we always mean the essential supremum.

2.2. Constants and data. As usual, the letter $c$ is reserved to positive constants, whose value may change from line to line, or even in the same formula. We say that a generic constant $c$ depends on the data, if $c = c(n, m, \nu, L)$, where $\nu$ and $L$ are the quantities introduced in (1.2).

Let $\eta \in L^\infty(E)$: we denote by

$$(g)_E^\eta = \frac{1}{\|\eta\|_{L^1(E)}} \int_E g \eta \, dx.$$
In the special case of \( \eta \equiv 1 \), we write

\[
(g)^1_E =: (g)_E =: \int_E g \, dx.
\]

We will frequently use what we will refer to in the following as the best constant property. For positive \( \eta \) we have, for any \( c \in \mathbb{R} \) and \( q \in [1, \infty) \)

\[
(1) \left( \frac{1}{\| \eta \|_{L^1(E)}} \int_E |g - (g)_E|^q \eta \, dx \right)^\frac{1}{q} \leq 2 \left( \frac{1}{\| \eta \|_{L^1(E)}} \int_E |g - c|^q \eta \, dx \right)^\frac{1}{q}.
\]

Moreover, if \( 0 \leq \eta \leq 1 \), by (2.1) one obviously obtains that

\[
(2) \left( \frac{1}{\| \eta \|_{L^1(E)}} \int_E |g - (g)_E|^q \, dx \right)^\frac{1}{q} \leq 2 \left( \frac{1}{\| \eta \|_{L^1(E)}} \int_E |g - (g)_E|^q \, dx \right)^\frac{1}{q},
\]

and as a consequence, that

\[
|(g)^1_E - (g)_E| \leq \left( \frac{1}{\| \eta \|_{L^1(E)}} \int_E |g - (g)_E|^q \eta \, dx \right)^\frac{1}{q}
\]

\[
\leq 2 \left( \frac{1}{\| \eta \|_{L^1(E)}} \int_E |g - (g)_E|^q \, dx \right)^\frac{1}{q}. \tag{2.3}
\]

We will also use the following estimate that was first proved in \[DKS11, \text{Lemma 6.2}\]: for \( q \geq 2 \) we have

\[
\int_E |g - (g)_E|^q \, dx \leq c_0 \int_E \left| g^\frac{q}{2} - (g)^\frac{q}{2}_E \right|^2 \, dx \leq c_1 \int_E \left| g^\frac{q}{2} - (g^\frac{q}{2})_E \right|^2 \, dx, \tag{2.4}
\]

where \( c_0 \) and \( c_1 \) are positive constants that depend only on the data.

3. **Reverse Hölder inequalities in general cylinders.** The main result of this section is Proposition 3.4, an estimate of reverse Hölder type on general cylinders: the second and third term on its right-hand side, characterized by the multiplying factor \( \delta \), can be seen as error terms. In § 4, relying on this first result, and working in a suitably scaled intrinsic cylinder (see (4.1) for the definition of such an object), we will prove proper reverse Hölder estimates without error terms.

Proposition 3.4 holds in general cylinders of the form \( Q_{\theta \rho^2, \rho} \), which we assume to be “centered” at the origin for the sake of simplicity, i.e., \( Q_{\theta \rho^2, \rho} = (-\theta \rho^2, 0) \times B_\rho \). It results from the combination of the energy estimate of Lemma 3.1, and the purely analytic estimate of Gagliardo-Nirenberg type of Lemma 3.3.

**Lemma 3.1.** Let \( u \geq 0 \) be a local, weak solution to (1.1)–(1.2) with \( m > 1 \). For \( \rho, \theta > 0 \), suppose \( Q_{2\theta \rho^2, 2\rho} \subset E_T \). Then, there exists a constant \( c > 1 \) depending
only on the data, such that

\[
\sup_{t \in (-\theta \rho^2, 0)} \int_{B_{\theta \rho^2}} \frac{|u - (u(t))_{B_{\theta \rho^2}}|^2}{\theta \rho^2} dx + \int_{-\theta \rho^2 B_{\theta \rho^2}} u^{m-1}|Du|^2 dx dt \\
\leq c \int_{-2\theta \rho^2 B_{2\rho}} \int \left[ \left( u^{m-1} + \frac{1}{\theta} \right) \frac{|u - (u(t))_{B_{2\rho}}|^2}{\rho^2} + \frac{|u - (u(t))_{B_{2\rho}}|^{m+1}}{\rho^2} + \rho^m f^m \right] dx dt.
\]

Moreover, for \(1 \leq a < b \leq 2\), we have

\[
\sup_{t \in (-a \theta \rho^2, 0)} \int_{B_{a \theta \rho^2}} \frac{|u - (u(t))_{B_{a \theta \rho^2}}|^2}{a \theta \rho^2} dx + \int_{-a \theta \rho^2 B_{a \theta \rho^2}} u^{m-1}|Du|^2 dx dt \\
\leq c \int_{-b \theta \rho^2 B_{b \rho}} \int \left[ \left( u^{m-1} + \frac{1}{\theta} \right) \frac{|u - (u(t))_{B_{b \rho}}|^2}{(b-a)^2 \rho^2} + \frac{|u - (u(t))_{B_{b \rho}}|^{m+1}}{\rho^2} + \rho^m f^m \right] dx dt.
\]

**Proof.** The first estimate is a consequence of the second one. Therefore, let \(1 \leq a < b \leq 2\). Fix \(s \in (-a \theta \rho^2, 0]\) and \(\eta \in C^{0,1}_0(B_{b \rho})\), with

\[|D\eta| \leq \frac{c}{(b-a)\rho}, \quad \chi_{B_{a \rho}} \leq \eta \leq \chi_{B_{b \rho}}.\]

Then, take the test function \(\varphi(x, t) = (u(x, t) - (u(t))_{B_{b \rho}}^2) [\eta(x)]^2 \left( \frac{t + b \theta \rho^2}{s + b \theta \rho^2} \right)\), where

\[(u(t))_{B_{b \rho}}^2 := \frac{1}{\int_{B_{b \rho}} [\eta(x)]^2 dx} \int_{B_{b \rho}} u(x, t) [\eta(x)]^2 dx.\]

We insert the test function in the weak formulation of (1.1) and find

\[
\int_{-b \theta \rho^2 B_{b \rho}} \int \left[ \partial_t u \varphi + A(x, t, u, Du) \cdot D\varphi - f \varphi \right] dx dt = 0.
\]

We estimate each integrand separately. We start with the first one. Notice that for any measurable function \(g : (-2\theta \rho^2, 0] \to \mathbb{R}\), by a formal computation which can be made rigorous by a standard Steklov average, we find that

\[
\int_{B_{b \rho}} \partial_t u (u - (u(t))_{B_{b \rho}}^2) \eta^2 = \int_{B_{b \rho}} (\partial_t u - g(t)) (u - (u(t))_{B_{b \rho}}^2) \eta^2 dx,
\]
which implies by the right Steklov approximation, that

$$
\int_{\{t\} \times B_{b\rho}} \partial_t u(u - (u(t))^2_{B_{b\rho}}) \eta^2 dx = \int_{\{t\} \times B_{b\rho}} \partial_t \left( \frac{u - (u(t))^2_{B_{b\rho}}}{2} \right) \eta^2 dx.
$$

Therefore, taking into account the \( \left( \frac{t+b\theta \rho^2}{s+b\theta \rho^2} \right) \) factor in the test function \( \phi \), and using the best constant property, the time-term can be transformed and estimated in the following way:

$$
\int_{-b\theta \rho^2}^{s} \int_{B_{b\rho}} \partial_t u \varphi \, dx \, dt
$$

$$
= \int_{\{s\} \times B_{b\rho}} \left| \frac{u - (u(s))^2_{B_{b\rho}}}{2} \right| \eta^2 \, dx - \int_{-b\theta \rho^2}^{s} \int_{B_{b\rho}} \left| \frac{u - (u(t))^2_{B_{b\rho}}}{2} \right| \eta^2 \, dx \, dt
$$

$$
\geq \int_{\{s\} \times B_{b\rho}} \left| u - (u(s))^2_{B_{b\rho}} \right| \eta^2 \, dx - 2 \int_{-b\theta \rho^2}^{s} \int_{B_{b\rho}} \left| \frac{u - (u(t))^2_{B_{b\rho}}}{2} \right| \eta^2 \, dx \, dt
$$

$$
\geq \frac{1}{4} \int_{\{s\} \times B_{b\rho}} \left| u - (u(s))^2_{B_{b\rho}} \right| \eta^2 \, dx - 2 \int_{-b\theta \rho^2}^{s} \int_{B_{b\rho}} \left| \frac{u - (u(t))^2_{B_{b\rho}}}{2} \right| \eta^2 \, dx \, dt.
$$

To analyze the second integrand, we find by (1.2) and Young’s inequality for every \( t \), that

$$
\int_{B_{b\rho}} A(x, t, u, Du) \cdot D\varphi \, dx
$$

$$
\geq \nu \int_{B_{b\rho}} \left| Du^{m+1} \right|^2 \eta^2 \left( \frac{t+b\theta \rho^2}{s+b\theta \rho^2} \right) \, dx
$$

$$
- (m-1)L \int_{B_{b\rho}} \left| Du^{m+1} \right| \eta^2 \left( \frac{t+b\theta \rho^2}{s+b\theta \rho^2} \right) \, dx
$$

$$
\geq \frac{\nu}{2} \int_{B_{b\rho}} \left| Du^{m+1} \right|^2 \eta^2 \left( \frac{t+b\theta \rho^2}{s+b\theta \rho^2} \right) \, dx - c(L, \nu, m) \int_{B_{b\rho}} u^{m-1} \left| \frac{u - (u(t))^2_{B_{b\rho}}}{(b-a)^2 \rho^2} \right|^2 \, dx.
$$
Finally, the last term on the right-hand side is estimated by Young’s inequality for every \( t \), as

\[
\left| \int_{B_{\rho}} f \varphi \, dx \right| \leq c \int_{B_{\rho}} \frac{\rho^{\frac{2}{m}} f^{\frac{m+1}{m}}}{\rho^2} \, dx + c \int_{B_{\rho}} \frac{|u - (u(t))_{B_{\rho}}|^2}{\rho^2} \, dx.
\]

Inserting all these estimates in (3.2), yields (3.1). \( \square \)

By a completely analogous argument we have the following:

**Lemma 3.2.** Let \( u \geq 0 \) be a local, weak solution to (1.1)-(1.2) with \( m > 1 \). For \( \rho, \theta > 0 \), suppose \( Q_{2\theta \rho^2, 2\rho} \subset E_T \). Then, there exists a constant \( c > 1 \) depending only on the data, such that

\[
\sup_{t \in (-\theta \rho^2, 0]} \int_{\{t\} \times B_{\rho}} \frac{u^2}{\theta \rho^2} \, dx + \int_{-2\theta \rho^2 B_{\rho}}^{0} u^{m-1} |Du|^2 \, dx \, dt \leq c \int_{-2\theta \rho^2 B_{\rho}}^{0} \left( u^{m-1} + \frac{1}{\theta} \right) \frac{u^2}{\rho^2} + \rho^m \, dx \, dt.
\]

**Proof.** It is exactly the same as before; the only difference is in the test function, where the mean values are now discarded. \( \square \)

The second ingredient is the next lemma, which is purely analytical, independent of any partial differential equation.

For the sake of simplicity, from here on we let \( (u)_{Q_{\theta \rho^2, \rho}} = a \) and \( (u(t))_{B_{\rho}} = a(t) \).

**Lemma 3.3.** For any function \( u \in L^\infty(-\theta \rho^2, 0; L^2(B_{\rho})) \), such that \( Du^{\frac{m+1}{2}} \in L^2(Q_{\theta \rho^2, \rho}) \), there exist \( \gamma \in (0, 1) \) and \( q_0 \in (0, \infty) \), such that for every \( \delta \in (0, 1) \) there holds

\[
\int_{Q_{\theta \rho^2, \rho}} \frac{|u - (u(t))_{B_{\rho}}|^2}{\rho^{2\gamma}} \, dx \, dt \leq \delta \sup_{t \in (-\theta \rho^2, 0]} \int_{B_{\rho}} \frac{|u - (u(t))_{B_{\rho}}|^2}{\theta \rho^2} \, dx + c_\delta \left( \theta (u)_{Q_{\theta \rho^2, \rho}}^{m-1} \right)^{q_0} \left( \int_{Q_{\theta \rho^2, \rho}} |Du|^{\frac{m+1}{2}} \, dx \, dt \right)^{\frac{1}{2\gamma}},
\]
where \( c_\delta \) depends on \( \delta \) and the data only.

Proof. The proof is done by interpolation. For \( b > d > 1 \) arbitrary and \( \beta = \frac{2}{d(m+1)} \) we can find \( \sigma \in (0,1) \), such that \( \frac{c}{b} + \frac{(1-\sigma)d}{b} = 1 \). Therefore,

\[
\left( \int_{B_\rho} |u - a(t)|^{(m+1)d} \right)^{\frac{1}{d}} \leq \left( \int_{B_\rho} |u - a(t)|^2 \right)^{\frac{\sigma(m+1)}{d}} \left( \int_{B_\rho} |u - a(t)|^{(m+1)b} \right)^{\frac{(1-\sigma)d}{b}}.
\]

We choose \( \alpha \in (0,1) \) such that \( \alpha(m+1) + \frac{1-\alpha}{d} = 1 \), and find

\[
\int_{Q_{\theta^{\rho^2},\rho}} |u - a(t)|^{m+1} \, dx \, dt
\]

\[
\leq \left( \int_{Q_{\theta^{\rho^2},\rho}} |u - a(t)|^2 \, dx \right)^{\alpha(m+1)} \left( \int_{Q_{\theta^{\rho^2},\rho}} |u - a(t)|^{(m+1)d} \, dx \, dt \right)^{\frac{(1-\alpha)}{d}}
\]

\[
\leq c\alpha^{(m+1)} \sup_{t \in (-\theta^{\rho^2},0]} \left( \int_{B_\rho} |u - a(t)|^2 \, dx \right)^{\frac{\sigma(m+1)(1-\alpha)}{d}} \times \rho^{2(1-\sigma)(1-\alpha)} \left( \int_{-\theta^{\rho^2}}^0 \left( \int_{B_\rho} \frac{|u - a(t)|^{(m+1)b}}{\rho^{2b}} \, dx \right)^{\frac{(1-\sigma)d}{b}} \, dt \right)^{\frac{1-\alpha}{d}}.
\]

By (2.4) and Poincaré’s inequality we find that

\[
\int_{Q_{\theta^{\rho^2},\rho}} |u - a(t)|^{m+1} \, dx \, dt
\]

\[
\leq c\alpha^{(m+1)} \rho^{2(1-\sigma)(1-\alpha)} \sup_{t \in (-\theta^{\rho^2},0]} \left( \int_{B_\rho} |u - a(t)|^2 \, dx \right)^{\frac{\sigma(m+1)(1-\alpha)}{d}} \times \left( \int_{Q_{\theta^{\rho^2},\rho}} \left| Du \frac{m+1}{2} \right|^{2\gamma} \, dx \, dt \right)^{\frac{1-\sigma(1-\alpha)}{\gamma}},
\]
provided that we choose $\gamma \in (0, 1)$ such that $\frac{(1-\sigma)d}{\gamma} \leq 1$. Hölder’s inequality for $(1 - \sigma)(1 - \alpha) + \sigma(1 - \alpha) + \alpha = 1$ gives

$$\int_{Q_{\theta \rho^2, \rho}} \frac{|u - a(t)|^{m+1}}{\rho^2} \, dx \, dt$$

$$\leq c_\rho^{-2(\sigma(1-\alpha)+\alpha)} \theta^{-\alpha \frac{m+1}{m-1}} \sup_{t \in (-\theta \rho^2, 0)} \left( \int_{B_\rho} |u - a(t)|^2 \, dx \right)^{\frac{\sigma(m+1)(1-\alpha)}{2}}$$

$$\times \left( \theta a^{m-1} \right)^{\frac{\alpha}{m-1}} \left( \int_{Q_{\theta \rho^2, \rho}} |Du|^{m+1} \, dx \right)^{\frac{2\gamma}{m+1}}$$

$$\leq \delta \rho^{-2} \theta^{-\left(\frac{m+1}{m} - \sigma(1-\alpha)+\alpha\right)} \left( \int_{B_\rho} |u - a(t)|^2 \, dx \right)^{\frac{\alpha}{1-\sigma(1-\alpha)+\alpha}}$$

$$+ c_\delta \left( \theta a^{m-1} \right)^{\frac{\alpha}{1-\sigma(1-\alpha)+\alpha}} \left( \int_{Q_{\theta \rho^2, \rho}} |Du|^{m+1} \, dx \right)^{\frac{2\gamma}{m+1}}.$$  

Therefore, we need that $\frac{\sigma(1-\alpha)}{\sigma(1-\alpha)+\alpha} = \frac{2}{m+1}$ and $\frac{\alpha}{\sigma(1-\alpha)+\alpha} = \frac{m-1}{m+1}$, which can be realized, provided $b$ and $d$ are properly chosen. Finally $q_0 = \frac{\alpha}{1-\sigma(1-\alpha)+\alpha}$. □

Combining the last two lemmas yields the first estimate of reverse Hölder type.

**Proposition 3.4.** Let $u \geq 0$ be a local, weak solution to (1.1)–(1.2) with $m > 1$. For $\rho, \theta > 0$, suppose $Q_{2\rho^2, 2\rho} \subset E_T$. Then, there exist $\gamma \in (0, 1)$ and $q_0 > 1$, such that for any $\tilde{\delta}$ there is a constant $c > 1$ depending only on $\tilde{\delta}$, $\gamma$ and the data, such that

$$\sup_{t \in (-\theta \rho^2, 0)} \int_{\{t\} \times B_\rho} \frac{|u - u(t)|^{m+1}}{\theta \rho^2} \, dx + \int_{Q_{\theta \rho^2, \rho}} |Du|^{m+1} \, dx \, dt$$

$$\leq c \left( \theta a^{m-1} \right)^{q_0} \left( \int_{Q_{2\rho^2, 2\rho}} |Du|^{m+1} \, dx \right)^{\frac{1}{\gamma}}$$

$$+ \tilde{\delta} \int_{Q_{2\rho^2, 2\rho}} \frac{u^{m+1}}{\rho^2} \, dx \, dt + \frac{\tilde{\delta}}{\rho^2 \theta^{m-1}} + c_\delta \int_{Q_{2\rho^2, 2\rho}} \rho^m \, f^{m+1} \, dx \, dt.$$
Proof. Take $1 \leq a < b \leq 2$ and estimate the right-hand side of (3.1) by Young’s inequality and (2.3). We have

$$
\sup_{t \in (-a \theta \rho^2, 0]} \int_{B_{a \rho}} \frac{|u - (u(t)) B_{a \rho}|^2}{\theta \rho^2} dx + \int_{Q_{a \theta \rho^2, a \rho}} u^{m-1} |Du|^2 dx dt \\
\leq \frac{c}{(b-a)^2} \int_{Q_{b \theta \rho^2, b \rho}} \left( u^{m-1} + \frac{1}{\theta} \right) \frac{|u - (u(t)) B_{b \rho}|^2}{\rho^2} dx dt \\
+ c \int_{Q_{b \theta \rho^2, b \rho}} \left[ \frac{|u - (u(t)) B_{b \rho}|^{m+1}}{\rho^2} + \rho^{\frac{2}{m}} f^{\frac{m+1}{m}} \right] dx dt \\
\leq \frac{\delta}{\rho^2 (b-a)^{\frac{m+1}{m}}} \left( \frac{1}{\theta^{\frac{m+1}{m}}} + \int_{Q_{b \theta \rho^2, b \rho}} u^{m+1} dx dt \right) + c \int_{Q_{b \theta \rho^2, b \rho}} \rho^{\frac{2}{m}} f^{\frac{m+1}{m}} dx dt \\
+ \delta \sup_{t \in (-b \theta \rho^2, 0]} \int_{B_{b \rho}} \frac{|u - (u(t)) B_{b \rho}|^2}{\rho^2} dx \\
+ c(\delta, \delta) \left( \theta (u^{m-1})_{Q_{2 \theta \rho^2, 2 \rho}} \right)^{\gamma} \left( \int_{Q_{2 \theta \rho^2, 2 \rho}} |Du|^\gamma dx dt \right) \frac{1}{\gamma} .
$$

Now the interpolation Lemma 6.1 of [Giu03] yields the result. 

\[ \Box \]

4. Intrinsic reverse H"older inequalities. We will call $Q_{\theta \rho^2, \rho}$ a $K$-intrinsic cylinder, if

$$
\frac{1}{K \theta} \leq (u^{m+1})_{Q_{\theta \rho^2, \rho}}^{\frac{m+1}{m}} \leq \frac{K}{\theta} .
$$
We call $Q_{θρ^2,ρ}$ a $K$-sub-intrinsic cylinder, if only the estimate from above holds. In the following, we will avoid any reference to the constant $K$, meaning that the cylinders will be either intrinsic or sub-intrinsic for some $K > 1$. In the next two subsections we will prove reverse Hölder inequalities in intrinsic cylinders. We will have to distinguish two different conditions, the so-called Degenerate and Non-Degenerate Regimes.

4.1. The degenerate regime. Here and in the following we will consider an intrinsic cylinder $Q_{2θρ^2,2ρ} ⊂ E_T$, where we assume that

$$\left(\int_{Q_{2θρ^2,2ρ}} |u - (u)_{Q_{2θρ^2,2ρ}}|^{m+1} \, dx \, dt\right)^{1/(m+1)} \geq \varepsilon \left(\int_{Q_{2θρ^2,2ρ}} |u^{m+1}| \, dx \, dt\right)^{1/(m+1)}$$

for some $\varepsilon > 0$. We denote this condition as the Degenerate Regime. We need the following lemma, which is proved using the weak time-derivative of the solution.

Lemma 4.1. Let $u \geq 0$ be a local, weak solution to (1.1)–(1.2) for $m > 1$. If $Q_{τ,2ρ}(t_o,x_o) ⊂ E_T$, then there exists $c = c(data)$ such that, for almost all $t_1, t_2 \in (t_o - τ, t_o + τ)$, we have

$$\left|(u(t_1))^{η}_{B_ρ} - (u(t_2))^{η}_{B_ρ}\right| \leq c \frac{τ}{ρ} \int_{Q_{τ,2ρ}(t_o,x_o)} |Du^m| \, dx \, dt + cτ \int_{Q_{τ,2ρ}(t_o,x_o)} f \, dx \, dt,$$

for all $η ∈ C^0_{o,1}(B_2ρ)$, with $|Dη| \leq \frac{ε}{ρ}$ and $χ_{B_ρ} \leq η \leq χ_{B_2ρ}$. Moreover, by Hölder’s inequality we find

$$\left|(u(t_1))^{η}_{B_ρ} - (u(t_2))^{η}_{B_ρ}\right| \leq cτ \int_{Q_{τ,2ρ}(t_o,x_o)} f \, dx \, dt \quad \text{for } \int_{Q_{τ,2ρ}(t_o,x_o)} f \, dx \, dt$$

Proof. We estimate formally (but things can be made rigorous by standard arguments):

$$\left|(u(t_1))^{η}_{B_ρ} - (u(t_2))^{η}_{B_ρ}\right| \leq \frac{1}{\|η\|_{L^1(B_ρ)}} \int_{t_1}^{t_2} \frac{d}{dt} \int_{B_ρ} uη \, dx \, dt$$
\[
= \frac{1}{\|\eta\|_{L^1(B_\rho)}} \int_{t_1}^{t_2} \langle \partial_t u, \eta \rangle \, dt
\]
\[
= \frac{1}{\|\eta\|_{L^1(B_\rho)}} \int_{t_1}^{t_2} \langle A(x,t,u,Du) - f, \eta \rangle \, dt
\]
\[
\leq \frac{c_T}{\rho} \int_{Q_{T,2\rho}(t_0,x_0)} |D\eta|^m \, dx \, dt + c_T \int_{Q_{T,2\rho}(t_0,x_0)} f \, dx \, dt. \quad \square
\]

We also need the following analytical remark.

**Lemma 4.2.** Let \( Q_{s,\rho} \subset \mathbb{R}^{n+1} \). If for a positive function \( g \in L^q(Q_{2s,2\rho}) \) we have

\[
\int_{Q_{s,\rho}} g^q \, dx \, dt \leq K \int_{Q_{s,\rho}} \left| g - (g)_{Q_{s,\rho}} \right|^q \, dx \, dt,
\]

then \( \forall a \in (1,2] \)

\[
\int_{Q_{as,\rho}} g^q \, dx \, dt \leq c(q) \left[ a^{n+1}(2K+1) + 1 \right] \int_{Q_{as,\rho}} \left| g - (g)_{Q_{as,\rho}} \right|^q \, dx \, dt.
\]

**Proof.** By the best constant property we find

\[
\int_{Q_{as,\rho}} g^q \, dx \, dt
\]
\[
= \int_{Q_{as,\rho}} \left| g - (g)_{Q_{as,\rho}} + (g)_{Q_{as,\rho}} - (g)_{Q_{s,\rho}} + (g)_{Q_{s,\rho}} \right|^q \, dx \, dt
\]
\[
\leq c(q) \left[ \int_{Q_{as,\rho}} \left| g - (g)_{Q_{as,\rho}} \right|^q \, dx \, dt + \left| (g)_{Q_{as,\rho}} - (g)_{Q_{s,\rho}} \right|^q \right] + \left( \int_{Q_{s,\rho}} g \, dx \, dt \right)^q
\]
\[
\leq c(q) \left[ \int_{Q_{as,\rho}} \left| g - (g)_{Q_{as,\rho}} \right|^q \, dx \, dt + \int_{Q_{s,\rho}} \left| g - (g)_{Q_{as,\rho}} \right|^q \, dx \, dt
\]
\[
+ K \int_{Q_{s,\rho}} \left| g - (g)_{Q_{s,\rho}} \right|^q \, dx \, dt\right]
\]
\[
\leq c(q) \left[ (2K + 1) \int_{Q_{s,\rho}} \left| g - (g)_{Q_{as,\rho}} \right|^q \, dx \, dt + \int_{Q_{as,\rho}} \left| g - (g)_{Q_{as,\rho}} \right|^q \, dx \, dt \right]
\]
if there exist then let 

(4.4) \[ u \in \mathbb{R} \quad \text{satisfy the degenerate alternative in the intrinsic cylinder } Q_{a_0 \rho^2, \rho} \]

**Proof.** We apply (3.3) for 2

\[ \int_{Q_{as, \rho}} |g - (g)_{Q_{as, \rho}}|^q \, dx \, dt. \]

Therefore, by making \( \varepsilon \) a bit smaller, from (4.2), we deduce that for all \( a \in [2, 4] \)

(4.3) \[ \left( \int_{Q_{a_0 \rho^2, \rho}} |u - (u)_{Q_{a_0 \rho^2, \rho}}|^{m+1} \, dx \right)^{\frac{1}{m+1}} \geq \varepsilon \left( \int_{Q_{a_0 \rho^2, \rho}} u^{m+1} \, dx \right)^{\frac{1}{m+1}}. \]

**Proposition 4.3.** Let \( u \geq 0 \) be a local, weak solution to (1.1)-(1.2) for \( m > 1 \), and let \( u \) satisfy the degenerate alternative in the intrinsic cylinder \( Q_{2 \theta \rho^2, 2 \rho} \). Then there exist \( q \in (0, 1) \) and a constant \( c > 0 \) depending on \( \varepsilon \), and the data, such that, if

\[ \theta(u^m)^{\frac{m-1}{m} Q_{4 \theta \rho^2, 4 \rho}} \leq cK, \]

then

(4.4) \[ \sup_{t \in (-2 \theta \rho^2, 0]} \int_{B_{2 \rho}} \left| \frac{u - (u(t))_{B_{2 \rho}}}{\theta \rho^2} \right|^2 \, dx + \int_{Q_{2 \theta \rho^2, 2 \rho}} \left| Du \right|^{\frac{m+1}{2}} \, dx \, dt + \frac{1}{\rho^2} \int_{Q_{a_0 \rho^2, \rho}} u^{m+1} \, dx \, dt \]

\[ \leq c(\varepsilon, K) \left( \int_{Q_{4 \theta \rho^2, 4 \rho}} \left| Du \right|^{\frac{m+1}{2}} \, dx \, dt \right)^{\frac{1}{q}} + \int_{Q_{4 \theta \rho^2, 4 \rho}} \rho^2 \frac{m+1}{m} \, dx \, dt. \]

**Proof.** We apply (3.3) for 2 \( \leq a < b \leq 4 \) with \( \tilde{\delta} = 1 \) to obtain

\[ \sup_{t \in (-a \theta \rho^2, 0]} \int_{\{t\} \times B_{a \rho}} \left| \frac{u - (u(t))_{B_{a \rho}}}{\theta \rho^2} \right|^2 \, dx + \int_{Q_{a_0 \rho^2, \rho}} u^{m-1} \, |Du|^2 \, dx \, dt \]

\[ \leq \frac{1}{\rho^2 (b-a)^{\frac{m+1}{m}}} \left( \frac{1}{\theta^{\frac{m+1}{m}}} + \int_{Q_{b \theta \rho^2, b \rho}} u^{m+1} \, dx \, dt \right) + c \int_{Q_{b \theta \rho^2, b \rho}} \rho^2 \frac{m+1}{m} \, dx \, dt \]

\[ + \Delta \sup_{t \in (-b \theta \rho^2, 0]} \int_{B_{b \rho}} \left| \frac{u - (u(t))_{B_{b \rho}}}{\theta \rho^2} \right|^2 \, dx \]

\[ + c(\delta)(\theta(u)^{\frac{m-1}{m} Q_{4 \theta \rho^2, 4 \rho}})^{\frac{q_0}{2}} \left( \int_{Q_{4 \theta \rho^2, 4 \rho}} \left| Du \right|^{\frac{m+1}{2}} \, dx \, dt \right)^{\frac{1}{q}}. \]
By the intrinsic nature of the cylinder, and enlarging the ball,
\[
\sup_{t \in (-a\theta^2,0]} \int_{B_{a\theta^2}} \frac{|u-y(t)|_{B_{a\rho}}^2}{\theta \rho^2} \, dx + \int_{Q_{a\theta^2,2\rho}} u^{m-1} |Du|^2 \, dxdt \\
\leq \frac{1 + cK_{m+1}^{m+1}}{\rho^2 (b-a)^{m+1}} \int_{Q_{b\theta^2,2\rho}} u^{m+1} \, dxdt + c \int_{Q_{2b\theta^2,2\rho}} \rho \int_{Q_{2b\theta^2,2\rho}} u^{m+1} \, dxdt \\
+ \delta \sup_{t \in (-b\theta^2,0]} \int_{B_{b\rho}} \frac{|u-y(t)|_{B_{b\rho}}^2}{\theta \rho^2} \, dx \\
+ c(\delta) \left( \frac{\theta(u)^{m-1}_{Q_{4b\theta^2,4\rho}}}{q_o} \left( \int_{Q_{4b\theta^2,4\rho}} |Du|^{m+1} \, dx \right)^{\frac{1}{\gamma}} \right).
\]

We are left with the estimate of \( \frac{1}{\rho^2} \int_{Q_{b\theta^2,2\rho}} u^{m+1} \, dxdt \), which also justifies the extra term in (4.4), as we make no use of the smallness of \( \tilde{\eta} \) in this proof. Once more, we define \( (u(t))_{B_{b\rho}} = a(t) \) and \( a = (u)_{Q_{b\theta^2,2\rho}} \). Using the degenerate condition (4.3), the best constant property (2.1), and relabeling with \( \varepsilon \) the quantity \( \varepsilon^{m+1} \), we find that
\[
\frac{1}{\rho^2} \int_{Q_{b\theta^2,2\rho}} u^{m+1} \, dxdt \leq \frac{1}{\varepsilon \rho^2} \int_{Q_{b\theta^2,2\rho}} |u-a|^{m+1} \, dxdt \\
\leq \frac{c}{\varepsilon \rho^2} \int_{Q_{b\theta^2,2\rho}} |u-(u)_{Q_{b\theta^2,2\rho}}|^{m+1} \, dxdt \\
\leq \frac{c}{\varepsilon \rho^2} \int_{Q_{b\theta^2,2\rho}} |u-(u(t))_{B_{b\theta^2,2\rho}}|^{m+1} \, dxdt \\
+ \frac{c}{\varepsilon \rho^2} \int_{-b\theta^2}^0 \left( (u(t))_{B_{b\theta^2,2\rho}} - (u)_{Q_{b\theta^2,2\rho}} \right)^{m+1} \, dt,
\]
where \( \eta \in C^m_0(B_{2\rho}) \) is chosen as in Lemma 4.1; notice that \( (u)_{Q_{b\theta^2,2\rho}} = (u)_{Q_{2b\theta^2,2\rho}} \). By Jensen’s inequality and Lemma 4.1,
\[
\int_{-b\theta^2}^0 \left( (u(t))_{B_{b\theta^2,2\rho}} - (u)_{Q_{b\theta^2,2\rho}} \right)^{m+1} \, dt \\
\leq c(u^m)_{b\theta^2,2\rho} \sup_{t \in (-b\theta^2,0]} \left| (u(t))_{B_{b\theta^2,2\rho}} - (u)_{Q_{b\theta^2,2\rho}} \right| 
\]
\[
\leq c(u^m)_{Q_{b\theta \rho^2, b \rho}} \theta \rho \left( \int \frac{|Du_m^{m+1}|^{2(m+1)}}{m+3} \ dx \ dt \right)^{\frac{m+3}{2(m+1)}} \left( \int \frac{u^{m+1}}{Q_{b\theta \rho^2, b \rho}} \ dx \ dt \right)^{\frac{m-1}{2(m+1)}} \\
+ c(u^m)_{Q_{b\theta \rho^2, b \rho}} \theta \rho^2 \int f \ dx \ dt \\
\leq c K(u^m)^{\frac{1}{m}}_{Q_{b\theta \rho^2, b \rho}} \rho \left( \int \frac{|Du_m^{m+1}|^{2(m+1)}}{m+3} \ dx \ dt \right)^{\frac{m+3}{2(m+1)}} \left( \int \frac{u^{m+1}}{Q_{b\theta \rho^2, b \rho}} \ dx \ dt \right)^{\frac{1}{q}} \\
+ c K(u^m)^{\frac{1}{m}}_{Q_{b\theta \rho^2, b \rho}} \rho^2 \int f \ dx \ dt \\
\leq c(\delta^*, K) \rho^2 \int \left( |Du_m^{m+1}|^{2q} \ dx \ dt \right)^{\frac{1}{q}} + \delta^* \int \frac{u^{m+1}}{Q_{b\theta \rho^2, b \rho}} \ dx \ dt \\
+ c(\delta^*, K) \rho \int \frac{|f|^{m+1}}{m} \ dx \ dt,
\]

where we can pick \( q \in \left[ \frac{m+1}{m+3}, 1 \right) \). Choose \( \delta^* = \frac{1}{2} \); together with (2.2), this yields

\[
\frac{1}{\rho^2} \int \frac{u^{m+1}}{Q_{b\theta \rho^2, b \rho}} \ dx \ dt \leq \frac{2}{\varepsilon \rho^2} \int \frac{|u - a(t)|^{m+1}}{Q_{b\theta \rho^2, b \rho}} \ dx \ dt \\
+ \frac{c(K)}{\varepsilon} \left( \int \frac{|Du_m^{m+1}|^{2q}}{Q_{b\theta \rho^2, b \rho}} \ dx \ dt \right)^{\frac{1}{q}} \\
+ \frac{c(K)}{\varepsilon} \rho \int \frac{|f|^{m+1}}{m} \ dx \ dt.
\]
We apply Lemma 3.3 with the choice $\delta = \frac{\epsilon (b-a)^{2m+1}}{4(2+cK^{m+1})}$, and obtain

$$\frac{1}{\rho^2} \int_{Q_{b\theta^2, b\rho}} u^{m+1} dx dt$$

$$\leq \frac{(b-a)^{2m+1}}{2(2+cK^{m+1})} \sup_{t \in (-b\theta^2, 0)} \int_{B_{b\rho}} \frac{|u - (u(t))_{B_{b\rho}}|^2}{\theta^2} dx$$

$$+ \frac{c(\epsilon, K)}{(b-a)^{p_1}} \left( \left( \int_{Q_{b\theta^2, b\rho}} |Du|^{m+1} dx dt \right)^{\frac{1}{q}} + \rho^2 \int \frac{m+1}{m} dx dt, \right)$$

where the exponent $p_o$ comes from Lemma 3.3. Collecting all the terms, for $p_1 = \max \{2, p_o\}$ we find

$$\sup_{t \in (-\theta a^2, 0)} \int_{B_{a\rho}} \frac{|u - (u(t))_{B_{a\rho}}|^2}{\rho^2} dx + \int_{Q_{2\theta a^2, 2\rho}} |Du|^{m+1} dx dt + \frac{1}{\rho^2} \int_{Q_{b\theta^2, b\rho}} u^{m+1} dx dt$$

$$\leq \frac{c(\delta, \epsilon, K)}{(b-a)^{p_1}} \left( \left( \int_{Q_{b\theta^2, b\rho}} |Du|^{m+1} dx dt \right)^{\frac{1}{q}} + \rho^2 \int \frac{m+1}{m} dx dt, \right)$$

$$+ \left( \frac{1}{2} + \delta \right) \sup_{t \in (-\theta b^2, 0)} \int_{B_{b\rho}} \frac{|u - (u(t))_{B_{b\rho}}|^2}{\theta^2} dx.$$

Now for $\delta < \frac{1}{2}$ the interpolation Lemma 6.1 of [Giu03] concludes the proof. \qed

4.2. The non-degenerate regime. In the following we assume that the opposite of (4.2) holds. Namely, we consider a cylinder $Q_{2\theta a^2, 2\rho} \subset E_T$, where we assume that

$$\left( \int_{Q_{2\theta a^2, 2\rho}} |u - (u(t))_{Q_{2\theta a^2, 2\rho}}|^{m+1} dx dt \right)^{\frac{1}{m+1}} \leq \epsilon \left( \int_{Q_{2\theta a^2, 2\rho}} u^{m+1} dx dt \right)^{\frac{1}{m+1}}.$$  

We denote this condition the Non-Degenerate Regime; heuristically, it implies that $u$ is close to a solution of the linear heat equation.

First of all, we need the following estimate, which was originally derived in [Sch13].

**Lemma 4.4.** Let $N \in \mathbb{N}$, $\sigma \in (0, 1)$ and consider a non-negative function $u \in L^{m+1}(Q_{s,r})$, where $Q_{s,r} = (-s, 0] \times B_r$. If for $\epsilon \in (0, \frac{1}{(N + 1)^{m+1}})$, $u$ satisfies the
non-degenerate condition, i.e.,

\[
\left( \int_{Q_{s,r}} \left| u - (u)_{Q_{s,r}} \right|^{m+1} \, dx \, dt \right)^{\frac{1}{m+1}} \leq \varepsilon (u^{m+1})^{\frac{1}{Q_{s,r}}}
\]

then

\[
(u^{m+1})^{\frac{1}{Q_{s,r}}} \leq \frac{1}{1 - \left( \frac{N}{\sigma} \right)^{m+1} + 1} \varepsilon \left( \int_{-s+k+1}^{-s+k} u \, dx \, dt \right)
\]

\[
\leq \frac{1}{1 - \left( \frac{N}{\sigma} \right)^{m+1} + 1} \varepsilon \left( \int_{-s+k+1}^{-s+k} u^{m+1} \, dx \, dt \right)^{\frac{1}{m+1}}
\]

\[
\leq \frac{1}{1 - \left( \frac{N}{\sigma} \right)^{m+1} + 1} \varepsilon (u^{m+1})^{\frac{1}{Q_{s,r}}}
\]

for every \( k \in \{0, \ldots, N - 1 \} \) and \( r_1 \in [\sigma r, r] \).

**Proof.** We use the triangular and the Jensen inequalities to find the estimate from below, namely

\[
(u^{m+1})^{\frac{1}{Q_{s,r}}} \leq \left( \int_{Q_{s,r}} \left| u - (u)_{Q_{s,r}} \right|^{m+1} \, dx \, dt \right)^{\frac{1}{m+1}}
\]

\[
+ \left| (u)_{Q_{s,r}} - (u)_{(s-\frac{k}{N}s, s-\frac{k+1}{N}s) \times B_{r_1}} \right| + \left| \frac{s+\frac{k+1}{N}s}{-s+\frac{k}{N}s} \right| B_{r_1} \int_{-s+\frac{k+1}{N}s}^{-s+\frac{k}{N}s} u \, dx \, dt \]

\[
\leq \varepsilon (u^{m+1})^{\frac{1}{Q_{s,r}}} + \left| \frac{s+\frac{k+1}{N}s}{-s+\frac{k}{N}s} \right| B_{r_1} \int_{-s+\frac{k+1}{N}s}^{-s+\frac{k}{N}s} u \, dx \, dt + \left| \frac{s+\frac{k+1}{N}s}{-s+\frac{k}{N}s} \right| B_{r_1} \int_{-s+\frac{k+1}{N}s}^{-s+\frac{k}{N}s} u \, dx \, dt
\]

\[
\leq \varepsilon (u^{m+1})^{\frac{1}{Q_{s,r}}} + \left| \frac{s+\frac{k+1}{N}s}{-s+\frac{k}{N}s} \right| B_{r_1} \int_{-s+\frac{k+1}{N}s}^{-s+\frac{k}{N}s} u \, dx \, dt + \left| \frac{s+\frac{k+1}{N}s}{-s+\frac{k}{N}s} \right| B_{r_1} \int_{-s+\frac{k+1}{N}s}^{-s+\frac{k}{N}s} u \, dx \, dt
\]
\[
\leq \epsilon (u^{m+1})^{\frac{1}{m+1}} + \left( \int_{-s + \frac{k+1}{N}}^{s} \int_{B_{r_1}} |u - (u)_{Q_{s,r}}|^{m+1} \, dx \, dt \right)^{\frac{1}{m+1}} \\
+ \int_{-s + \frac{k+1}{N}}^{s} \int_{B_{r_1}} u \, dx \, dt \\
\leq \epsilon \left( 1 + \left( \frac{N r^n}{r_1^n} \right)^{\frac{1}{m+1}} \right) (u^{m+1})^{\frac{1}{m+1}} + \int_{-s + \frac{k+1}{N}}^{s} \int_{B_{r_1}} u \, dx \, dt.
\]

By absorption, we conclude that

\[
(u^{m+1})^{\frac{1}{m+1}} \leq \frac{1}{1 - (\frac{N}{\sigma n})^{\frac{1}{m+1}} + \epsilon} \int_{-s + \frac{k+1}{N}}^{s} \int_{B_{r_1}} u \, dx \, dt.
\]

As for the estimates from above, a further application of Jensen’s inequality, and enlarging the set of integration yields

\[
\int_{-s + \frac{k+1}{N}}^{s} \int_{B_{r_1}} u \, dx \, dt \leq \left( \int_{-s + \frac{k+1}{N}}^{s} \int_{B_{r_1}} u^{m+1} \, dx \, dt \right)^{\frac{1}{m+1}} \\
\leq \left( \frac{N r^n}{r_1^n} \right)^{\frac{1}{m+1}} \left( \int_{-s + \frac{k+1}{N}}^{s} \int_{-B_{r}} u^{m+1} \, dx \, dt \right)^{\frac{1}{m+1}}.
\]

A classical property of diffusion equations, both linear and non-linear, is the so-called expansion of positivity. The following theorem presents such a property for positive super-solutions to (1.1)-(1.2). Its proof can be found in [DGV, Proposition 7.1, Chapter 4].

**Theorem 4.5.** (Expansion of Positivity) Assume that \( u \geq 0 \) is a super-solution to (1.1)-(1.2) in \( (s,s + T] \times B_{2\rho} \) with \( f = 0 \), and \( T > 0 \) sufficiently large. If

\[
|\{u(\cdot,s) > a\} \cap B_{\rho}| \geq \gamma |B_{\rho}|,
\]

for \( a > 0 \) and \( \gamma \in (0,1) \), there exist \( b > 1 \) and \( \tilde{\eta} \in (0,1) \), depending only on the data and \( \gamma \), such that

\[
u > \tilde{\eta} a \quad \text{in} \quad (s + \frac{b}{2a m - 1} \rho^2, s + \frac{b}{a m - 1} \rho^2] \times B_{2\rho}.
\]
The requirement on $T$ to be sufficiently large is needed to guarantee that
\[
\left(s + \frac{b}{2a^{m-1}}\rho^2, s + \frac{b}{a^{m-1}}\rho^2 \right] \times B_{2\rho} \subset (s, s+T] \times B_{2\rho}.
\]

Since $a > 0$ is a priori known, and $b > 1$ depends only on the data and $\gamma$, such a condition can easily be checked. In order to be able to use Theorem 4.5, we need some basic consequences of (4.5).

**Lemma 4.6.** For every $\alpha, \gamma \in (0, 1)$, there exists $\varepsilon \in (0, \frac{1}{2})$, such that if (4.5) holds for this $\varepsilon$, then

(4.8) \quad (u)_{Q_{2\theta\rho^2, 2\rho}} \geq (1 - \varepsilon)(u)^{m+1} \frac{1}{2^{m+1}}_{Q_{2\theta\rho^2, 2\rho}}

(4.9) \quad (u)_{Q_{2\theta\rho^2, 2\rho}} \leq \frac{1 - \varepsilon}{\gamma} \int_{Q_{2\theta\rho^2, 2\rho}} u \chi_{\{u \geq \alpha(u)_{Q_{2\theta\rho^2, 2\rho}}\}} \, dx dt

(4.10) \quad \left|Q_{2\theta\rho^2, 2\rho} \cap \{u \geq \alpha(u)_{Q_{2\theta\rho^2, 2\rho}}\}\right| \geq \gamma \frac{m+1}{m} \left|Q_{2\theta\rho^2, 2\rho}\right|.

**Proof.** The first statement follows directly by (4.5) and the triangular inequality. To get (4.9) we use that on the set $\{u < \alpha(u)_{Q_{2\theta\rho^2, 2\rho}}\}$ we have

\[
u < (u)_{Q_{2\theta\rho^2, 2\rho}} \leq \frac{1}{1 - \alpha} \left|(u)_{Q_{2\theta\rho^2, 2\rho}} - u\right|.
\]

Combining (4.8) and the previous estimate yields

\[
\int_{Q_{2\theta\rho^2, 2\rho}} u \, dx dt \leq \int_{Q_{2\theta\rho^2, 2\rho}} u \chi_{\{u \geq \alpha(u)_{Q_{2\theta\rho^2, 2\rho}}\}} \, dx dt
\]

\[
+ \frac{1}{1 - \alpha} \left(\int_{Q_{2\theta\rho^2, 2\rho}} \left|(u)_{Q_{2\theta\rho^2, 2\rho}} - u\right| \, dx dt \right)^{\frac{1}{m+1}}
\]

\[
\leq \int_{Q_{2\theta\rho^2, 2\rho}} u \chi_{\{u \geq \alpha(u)_{Q_{2\theta\rho^2, 2\rho}}\}} \, dx dt + \frac{1}{1 - \alpha} \frac{\varepsilon}{1 - \varepsilon} \int_{Q_{2\theta\rho^2, 2\rho}} u \, dx dt.
\]

If $\varepsilon$ is small enough, then the last term can be absorbed and also $1 - \frac{1}{1 - \alpha} \frac{\varepsilon}{1 - \varepsilon} \geq \frac{\gamma}{1 - \varepsilon}$, which implies (4.9). For the third estimate, by (4.8), (4.9), and Hölder’s inequality
we find that
\[
\int_{Q_{2\theta^2,2\rho}} u \, dx \, dt \leq \frac{1-\varepsilon}{\gamma} \int_{Q_{2\theta^2,2\rho}} u \chi_{\{u \geq \alpha(u)Q_{2\theta^2,2\rho}\}} \, dx \, dt
\]
\[
\leq \frac{1-\varepsilon}{\gamma} \left( \frac{\{u \geq \alpha(u)Q_{2\theta^2,2\rho}\}}{|Q_{2\theta^2,2\rho}|} \right) \left( \int_{Q_{2\theta^2,2\rho}} u^{m+1} \, dx \, dt \right)^{\frac{1}{m+1}}
\]
\[
\leq \frac{1}{\gamma} \left( \frac{\{u \geq \alpha(u)Q_{2\theta^2,2\rho}\}}{|Q_{2\theta^2,2\rho}|} \right) \int_{Q_{2\theta^2,2\rho}} u \, dx \, dt.
\]
This implies that
\[
|Q_{2\theta^2,2\rho}| \leq \frac{1}{\gamma} \left| \left\{ u \geq \alpha(u)Q_{2\theta^2,2\rho} \right\} \right|.
\]

Now we can apply Theorem 4.5 to our solution and find the following result.

**Proposition 4.7.** Let \( u \geq 0 \) be a local, weak solution to (1.1)-(1.2) for \( m > 1 \), and let \( u \) satisfy the non-degenerate alternative in the intrinsic cylinder \( Q_{2\theta^2,2\rho} \). Then for \( \alpha \in (0, \frac{1}{\theta^{m-1} (u)Q_{2\theta^2,2\rho}} \) and \( \varepsilon = \varepsilon(\alpha) \) small enough, there exists \( \eta \in (0,1) \), which depends only on the data and \( \alpha \), such that
\[
\{ u \geq \alpha(u)Q_{2\theta^2,2\rho} \}
\]
\[
\{ u \geq \eta(u)Q_{2\theta^2,2\rho} \text{ for all } (t,x) \in (-\theta^2,0] \times B_{2\rho}. \}
\]

**Proof.** Since \( f \) is positive, \( u \) is a super-solution to (1.1)-(1.2) with a vanishing right-hand side, and Theorem 4.5 can be applied. As before, we let \( (u)Q_{2\theta^2,2\rho} = a \).

We choose \( K_\rho \in \left[ \frac{1}{K}, K \right] \), such that \( \theta = \frac{K}{\alpha^{m-1}} \). Let \( \alpha \in (0, \frac{1}{K^{m-1}}] \), \( \gamma \in (0,1) \) and \( b > 1 \) accordingly chosen, by Theorem 4.5. Next, we fix \( \sigma \in (0,1] \), such that
\[
\frac{\sigma^2 b}{2\alpha^{m-1} K_\rho} = \frac{1}{4}.
\]

Now, let \( B_{\sigma \rho} \) be an arbitrary sub-ball of \( B_\rho \). Lemma 4.4 with \( N = 8 \) implies that for \( s \in [-2\theta^2, -\frac{3\theta}{4} \rho^2] \), the cylinder \( Q(s) := (s, s + \frac{\theta}{4} \rho^2] \times B_{\sigma \rho} \) is non-degenerate. Indeed, it is a matter of straightforward computations to check that the analogous of (4.5) holds for \( Q(s) \), with
\[
\tilde{\varepsilon} = \frac{\varepsilon \left( \frac{8 \rho^2}{\sigma^2} \right)^{m+1}}{1 - \varepsilon \left( 1 + \left( \frac{8 \rho^2}{\sigma^2} \right)^{m+1} \right)}
\]
where we also take into account that the ratio of the two radii is \( \frac{\sigma}{\tilde{\sigma}} \).
The construction implies that we can choose \( \varepsilon \) small enough with respect to \( \sigma, m, \) and \( n, \) such that \( \bar{\varepsilon} \) is arbitrarily small. In particular, once \( \alpha \) and \( \gamma \) have been chosen, we can select \( \varepsilon, \) such that \( \bar{\varepsilon} \) satisfies the assumptions of Lemma 4.6. Therefore, by (4.10), we have that for every \( s, \)

\[
|Q(s) \cap \{ u \geq \alpha a \}| \geq \gamma \frac{m+1}{m} |Q(s)|;
\]

in turn, this implies that there exists at least one \( t \in (s, s + \frac{\theta}{4} \rho^2], \) such that

\[
|B_{\sigma \rho} \cap \{ u(\cdot, t) \geq \alpha a \}| \geq \gamma \frac{m+1}{m} |B_{\sigma \rho}|.
\]

By Theorem 4.5 we conclude that \( u \geq \tilde{\eta}(\alpha a) \) in \( (t + \frac{\theta}{4} \rho^2, t + \frac{2\theta}{4} \rho^2] \times B_{2\sigma \rho}. \) By a further application of Theorem 4.5 we have

\[
u \geq \tilde{\eta}^2 \alpha a \quad \text{in} \quad \left(-\frac{3\theta}{2} \rho^2, 0\right] \times B_{2\sigma \rho},
\]

as \( s \) was arbitrarily chosen in \([-2\theta \rho^2, -\frac{3\theta}{4} \rho^2]. \) The proof is concluded, since we can cover \( Q_{2\theta \rho^2, 2\rho} \) with proper cylinders having cross section \( B_{\sigma \rho} \) and height \( \frac{\theta}{4} \rho^2. \)

**Remark 4.8.** As it is apparent from the proof, the functional dependence of \( \eta \) on \( \alpha \) is given by

\[
\eta = \tilde{\eta}^2 \alpha.
\]

Moreover, in the intrinsic cylinder we are working, the quantity \( \alpha \) is independent of \( u. \)

**Proposition 4.9.** Let \( u \geq 0 \) be a local, weak solution to (1.1)-(1.2) for \( m > 1, \)
and let \( u \) satisfy the non-degenerate condition in the intrinsic cylinder \( Q_{2\theta \rho^2, 2\rho}. \) Then there exist \( q \in (0,1) \) and \( c > 1 \) depending on \( \varepsilon \) and the data, such that

\[
\int_{Q_{2\theta \rho^2, 2\rho}} |Du|^{m+1}_{-\frac{1}{q}} \leq c \left( \int_{Q_{\theta \rho^2, \rho}} |Du|^{2q} dxdt \right)^{\frac{1}{q}} + c \int_{Q_{\theta \rho^2, \rho}} \rho^{\frac{m+1}{2m}} f dxdt.
\]

**Proof.** We estimate the right-hand side of (3.1). However, since we wish to use the expansion of positivity, on the right-hand side we choose to work in the
time interval \((-\theta \rho^2, 0]\). We estimate the different terms appearing in the right-hand side of (3.1) for \(\frac{1}{2} \leq a' < b \leq 1\). We begin with the most difficult part. We define \(a(t) = (u(t))_{B_{b \rho}}^\eta\) and \(a = (u)_{b \theta \rho^2, b \rho}\). By the use of (2.3), we find that

\[
\int_{B_{b \rho}} u^{m-1} \left| \frac{u - (u(t))_{B_{b \rho}}^\eta}{\rho^2} \right|^2 \, dx
\]

\[
\leq 2^{m-1} a^{m-1} \int_{B_{b \rho}} \chi_{|u - a| \leq \theta} \left| \frac{u - (u(t))_{B_{b \rho}}^\eta}{\rho^2} \right|^2 \, dx + 2^m \left| a - (u(t))_{B_{b \rho}}^\eta \right|^{m-1} \int_{B_{b \rho}} \left| \frac{u - (u(t))_{B_{b \rho}}^\eta}{\rho^2} \right|^2 \, dx
\]

\[
\leq (2a)^{m-1} \int_{B_{b \rho}} \left| \frac{u - (u(t))_{B_{b \rho}}^\eta}{\rho^2} \right|^2 \, dx + 2^m \int_{B_{b \rho}} \left| u - (u(t))_{B_{b \rho}}^\eta \right|^{m+1} \, dx
\]

\[
\leq (4a)^{m-1} \int_{B_{b \rho}} \left| \frac{u - a(t)}{\rho^2} \right|^2 \, dx + 2^m \int_{B_{b \rho}} \left| u - a(t) \right|^{m+1} \, dx
\]

\[
+ 2^m |a(t) - a|^{m-1} \int_{B_{b \rho}} \left| \frac{u - (u(t))_{B_{b \rho}}^\eta}{\rho^2} \right|^2 \, dx
\]

\[
+ 2^m |a(t) - (u(t))_{B_{b \rho}}^\eta|^{m-1} \int_{B_{b \rho}} \left| \frac{u - (u(t))_{B_{b \rho}}^\eta}{\rho^2} \right|^2 \, dx
\]

\[
\leq (4a)^{m-1} \int_{B_{b \rho}} \left| \frac{u - a(t)}{\rho^2} \right|^2 \, dx + 2^m \int_{B_{b \rho}} \left| u - a(t) \right|^{m+1} \, dx
\]

\[
+ 2^m |a - a(t)|^{m-1} \int_{B_{b \rho}} \left| \frac{u - a(t)}{\rho^2} \right|^2 \, dx
\]

\[
+ c \int_{B_{b \rho}} \left| \frac{u - a(t)}{\rho^2} \right|^2 \, dx \int_{B_{b \rho}} \left| u - a(t) \right|^{m-1} \, dx
\]

\[
\leq (4a)^{m-1} \int_{B_{b \rho}} \left| \frac{u - a(t)}{\rho^2} \right|^2 \, dx + c \int_{B_{b \rho}} \left| u - a(t) \right|^{m+1} \, dx
\]

\[
+ 2^m |a - a(t)|^{m-1} \int_{B_{b \rho}} \left| \frac{u - a(t)}{\rho^2} \right|^2 \, dx
\]

\[
=: (I) + (II) + (III).
\]
We integrate with respect to time, average over the domain of integration, and
estimate the three terms; by Poincaré’s and Young’s inequalities (I) gives

\[ a^{m-1} \int_{Q_{b\theta \rho^2, b\rho}} \frac{|u - a(t)|^2}{\rho^2} \, dx \, dt \]

\leq a^{m-1} \sup_{t \in (-b\theta \rho^2, 0]} \left( \frac{\int_{B_{b\rho}} |u - a(t)|^2}{\rho^2} \, dx \right)^{1-q} \left( \frac{\int_{-b\theta \rho^2} |u - a(t)|^2}{\rho^2} \, dx \right)^q dt

\leq \delta a^{m-1} \sup_{t \in (-b\theta \rho^2, 0]} \int_{B_{b\rho}} \frac{|u - a(t)|^2}{\rho^2} \, dx + c(\delta) a^{m-1} \left( \int_{Q_{b\theta \rho^2, b\rho}} |Du|^{2q} \, dx \, dt \right)^{\frac{1}{q}},

for any \( q \in \left[ \frac{n}{n+2}, 1 \right) \).

Relying on Lemma 3.3, the term coming from (II) yields

\[ (4.11) \]

\[ \int_{Q_{b\theta \rho^2, b\rho}} \frac{|u - a(t)|^{m+1}}{\rho^2} \, dx \, dt \leq \delta \sup_{t \in (-b\theta \rho^2, 0]} \int_{B_{b\rho}} \frac{|u - (u(t))_{B_{b\rho}}|^2}{\theta \rho^2} \, dx \]

\[ + c(\delta) \left( \theta(u)_{Q_{b\theta \rho^2, b\rho}}^{m-1} \right)^{q_0} \left( \int_{Q_{b\theta \rho^2, b\rho}} |Du|^{\frac{m+1}{2}} \, dx \, dt \right)^{\frac{1}{q}}, \]

We estimate (III) by using the non-degenerate condition, Lemma 4.4 with \( N = 1 \) and \( r_1 = b\rho \), and Hölder’s inequality.

\[ \int_{-b\theta \rho^2} |a - a(t)|^{m-1} \, dt \int_{B_{b\rho}} \frac{|u - a(t)|^2}{\rho^2} \, dx \]

\leq \left( \int \frac{0}{-b\theta \rho^2} |a - a(t)|^{m+1} \, dt \right)^{\frac{m-1}{m+1}} \left( \int_{-b\theta \rho^2} \left( \frac{\int_{B_{b\rho}} |u - a(t)|^2}{\rho^2} \, dx \right)^{\frac{m+1}{2}} \, dt \right)^{\frac{1}{m+1}}

\leq c \varepsilon a^{m-1} \left( \int_{-b\theta \rho^2} \left( \frac{\int_{B_{b\rho}} |u - a(t)|^2}{\rho^2} \, dx \right)^{\frac{m+1}{2}} \, dt \right)^{\frac{1}{m+1}},

where we relabel \( \varepsilon^{m-1} \) with \( \varepsilon \).
We choose \( q \in \left[ \frac{n}{n+2}, 1 \right) \), estimate by Poincaré’s and Young’s inequalities, use that \( a^{m-1} \leq \frac{c}{\theta} \) by Lemma 4.4, rely on the intrinsic nature of \( Q_{2\theta \rho^2, \rho} \), and take into account Proposition 4.7, to obtain

\[
(III) \leq c\varepsilon a^{m-1} \sup_{t \in (-b\theta \rho^2, 0]} \left( \int_{B_{\rho^2}} \frac{|u - a(t)|^2}{\rho^2} dx \right)^{\frac{m+1-2q}{m+1}} \times \left( \int_{-b\theta \rho^2}^{0} \left( \int_{B_{\rho^2}} \frac{|u - a(t)|^2}{\rho^2} dx \right) dt \right)^{\frac{2}{m+1}} \leq \delta (b - a')^2 a^{m-1} \sup_{t \in (-b\theta \rho^2, 0]} \int_{B_{\rho^2}} \frac{|u - a(t)|^2}{\rho^2} dx \]

\[
+ c(\delta, \varepsilon) a^{m-1} (b - a')^{-(m+1-2q)} \left( \int_{Q_{b\theta \rho^2, \rho}} |Du|^{2q} dx dt \right)^{\frac{1}{q}} \leq \delta \frac{(b - a')^2}{\theta} \sup_{t \in (-b\theta \rho^2, 0]} \int_{B_{\rho^2}} \frac{|u - a(t)|^2}{\rho^2} dx \]

\[
+ c(\delta, \varepsilon) (b - a')^{-(m+1-2q)} \left( \int_{Q_{b\theta \rho^2, \rho}} |Du|^{m+1} dx dt \right)^{\frac{1}{q}},
\]

where we have relabeled \( \delta \), in order to take into account an extra constant \( c \).

Therefore, using (2.3) and the last estimates, we find that for a proper exponent \( q_2 = q_2(q) \)

\[
\int_{Q_{b\theta \rho^2, 2\rho}} u^{m-1} \left( \frac{|u - (u(t))|^{q_2}}{\rho^2} \right)^2 dx dt
\]
\[
\leq c(\delta) (b - a')^{-q} \left[ 1 + \left( \theta(u)_{Q_{\theta p^2, \rho}}^{m-1} \right)^q \right] \left( \int_{Q_{b \theta p^2, \rho}} |Du|_{\frac{m+1}{2}}^{2q} \, dx \, dt \right)^{\frac{1}{q}} + \frac{\delta(b - a')^2}{\theta} \sup_{t \in (-b \theta p^2, 0)} \int_{B_{b p}} \frac{|u - a(t)|^2}{\rho^2} \, dx,
\]

where \( q \in (0, 1) \) is the maximum between \( \frac{n}{n+2} \) and \( \gamma \) given in (4.11).

Proposition 4.7 implies
\[
\int_{B_{b \rho}} \frac{1}{\rho^2} \left| u - (u(t))_{B_{b \rho}} \right|^2 \, dx \leq c \int_{B_{b \rho}} \frac{u^{m-1}}{\rho^2} \left| u - (u(t))_{B_{b \rho}} \right|^2 \, dx;
\]
therefore, we are left with the estimate of the term involving \( f \). However, it can be estimated just as in Proposition 3.4. Hence, as in (3.3) we find that
\[
\leq \delta \sup_{t \in (-b \theta p^2, 0)} \int_{B_{b \rho}} \frac{|u - (u(t))_{B_{b \rho}}|^2}{\theta \rho^2} \, dx + \int_{Q_{\theta_p^2, a^\prime \rho}} u^{m-1} |Du|^2 \, dx \, dt
\]

(4.12)
\[
+ c(b - a')^{-q_2} \left[ 1 + \left( \theta(u)_{Q_{\theta p^2, \rho}}^{m-1} \right)^q \right] \left( \int_{Q_{\theta_p^2, \rho}} |Du|_{\frac{m+1}{2}}^{2q} \, dx \, dt \right)^{\frac{1}{q}} + c \int_{Q_{\theta p^2, \rho}} \rho^{\frac{2}{m}} \int_{m}^{m+1} \, dx \, dt.
\]

Since \( Q_{\theta_p^2, \rho} \) is sub-intrinsic with constant \( \tilde{K} = 2^{n+1}K \), we conclude by removing the \( \delta \) term via the interpolation Lemma 6.1 of [Giu03].

5. Higher Integrability.

5.1. Covering. We now construct sub-intrinsic cylinders with properties convenient for our purposes.

**Lemma 5.1.** Let \( m \geq 1 \), \( Q_{S,R} \subset E_T \), \( u \in L^{m+1}(Q_{S,R}) \), and \( \hat{b} \in (0, 2) \). For every \( 0 < r < \rho \leq R \) there exist \( s(r) \), \( \theta_r \), and a cylinder \( Q_{s(r),r} \) with the same center as \( Q_{S,R} \), such that the following properties hold:

1. \( 0 \leq s(r) \leq S \) and \( s(r) = \theta_r r^2 \). In particular, \( Q_{s(r),r} \subset E_T \).
We will also need the following result, which is stated and proved in [Sch13, Lemma 4.1]. In particular, \(Q_{s(r),r} \subset Q_{s(r),r}\) for all \(r \in [r, \rho]\).

(3) \(\int_{Q_{s(r),r}} u^m dx dt \leq \frac{1}{\theta_{\sigma r}^{m+1}}, \text{i.e., } Q_{s(r),r} \) is sub-intrinsic.

(4) If \(s(r) < \left( \frac{r}{\rho} \right)^{b} s(\rho)\), then there exists \(r_1 \in [r, \rho]\) such that \(\int_{Q_{s(r_1),r_1}} u^m dx dt = \frac{1}{\theta_{\sigma r_1}^{m+1}}\).

(5) If for all \(r \in (r_1, \rho)\), one has \(\int_{Q_{s(r),r}} u^m dx dt < \frac{1}{\theta_{\sigma r}^{m+1}}\), then \(\frac{1}{\theta_{\sigma r}} \leq \left( \frac{r}{\rho} \right)^{2b} \frac{1}{\theta_{\sigma r}} \) for all \(r \in [r_1, \rho]\).

(6) For \(\sigma \in (0,1]\), \(\sigma^{2b} \leq \frac{1}{\theta_{\sigma r}} \leq \frac{c}{\sigma^{2b} \theta_{\sigma r}}\).

(7) For \(\sigma \in (0,1], \left| Q_{s(\sigma r),\sigma r} \right|^{-1} \leq c\sigma^{-(n+2)(1+\frac{m-1}{2})} \left| Q_{s(r),r} \right|^{-1}\).

(8) For \(\sigma \in (0,1]\), we have \(Q_{s(\sigma r),\sigma r} \subset Q_{\sigma^{b} s(r),\sigma^{b} r}\). For \(a > 1\) we have \(Q_{as(r),ar} \subset Q_{s(\sigma r),\sigma r}\), for \(\bar{a} = \max\{a^{1/b}, a\}\).

The positive constant \(c\) in (6) and (7) depends only on the data.

The proof of the lemma is done in [Sch13, Lemma 4.1] for the parabolic \(p\)-laplacian. It can be adapted to the porous medium equation almost verbatim, if one replaces \(p\) by \(m+1\) and \(|Du|\) by \(u\). For pure technical reasons, we will use a slightly modified construction. Indeed, we will use intrinsic cylinders of the form

\[
Q_{\theta_{r},r^2,r}(t,x) := \left( t - \frac{\theta_{r}}{2} r^2, t + \frac{\theta_{r}}{2} r^2 \right) \times B_r(x).
\]

We will also need the following result, which is stated and proved in [Sch13, Lemma A.1].

**Lemma 5.2.** Let \(Q_1 \subset Q\) be two cylinders and \(f \in L^q(Q)\) for some \(q \in [1,\infty)\). If for some \(\varepsilon \in (0,1)\) we have

\[
\left| (f)_{Q_1} \right| \leq \varepsilon \left[ \left( |f|^q \right)_{Q} \right]^{\frac{1}{q}},
\]

then

\[
\left| (f)_{Q_1} \right| \leq \varepsilon \left[ \left( |f|^q \right)_{Q} \right]^{\frac{1}{q}} \leq \frac{\varepsilon}{1 - \varepsilon} \left( 1 + \left( \frac{|Q|}{|Q_1|} \right)^{\frac{1}{q}} \right) \left( \int_{Q} |f - (f)_{Q_1}|^q dx dt \right)^{\frac{1}{q}}.
\]

We can finally come to the core of our argument.
Lemma 5.3. Let \( m \geq 1 \), \( Q_{2S,2R} \subset E_T \), and \( \hat{b} \in (0,2) \). For every \( z \equiv (t,x) \in Q_{S,R} \) and \( 0 < r \leq R \) there exist \( s(r,z) \), \( \theta_{r,z} \), and a sub-intrinsic cylinder \( Q_{s(r,z),r}(z) = (t - \theta_{r,z}r^2, t + \theta_{r,z}r^2) \times B_r(x) \), such that all properties of Lemma 5.1 hold.

Moreover, if we assume that

\[
(5.2) \quad \left( \int_{Q_{2S,2R}} u^{m+1} \, dx \, dt \right)^{\frac{m-1}{m+1}} \leq \frac{R^2}{S} =: \frac{C}{\theta_o},
\]

with \( C > 1 \), then

1. for all \( z \equiv (t,x) \in Q_{S,R} \), we have \( \frac{1}{\theta_o} \leq \frac{1}{\theta_{r,z}} \leq C^2 \frac{2^{2(m+1)+(m-1)n}}{m-1} \frac{1}{\theta_o} \).

2. There is a constant \( c_1 > 1 \), depending on \( n \), \( m \), \( \hat{b} \) only, such that if \( Q_{s(r,z),r}(z) \cap Q_{s(r,y),r}(y) \neq \emptyset \), then

\[
Q_{s(r,z),r}(z) \subset Q_{s(c_1r,y),c_1r}(y) \quad \text{and} \quad Q_{s(r,y),r}(y) \subset Q_{s(c_1r,z),c_1r}(z),
\]

for each \( r \leq \frac{R}{c_o} \), with \( c_o = \max\{3^{1/\hat{b}}, 3\} \).

Proof. We directly assume (5.2). For \( z = (t,x) \in Q_{S,R} \), we will construct proper sub-intrinsic cylinders. We have to fix the initial cylinder \( Q_{S(z),R(z)} \), required by Lemma 5.1. We do this by defining

\[
\frac{S}{2^{2(m+1)+(m-1)n}} \leq S(z) \leq S,
\]

such that

\[
\left( \int_{t-S(z)}^{t+S(z)} \int_{B_R(x)} u^{m+1} \, dx \, dt \right)^{m-1} S(z)^2 \leq R^{2(m+1)} |B_R|^{m-1},
\]

which is possible, since

\[
\left( \int_{t-S}^{t+S} \int_{B_R(x)} u^{m+1} \, dx \, dt \right)^{m-1} S^2 \leq \left( \int_{-2S}^{2S} \int_{B_{2R}} u^{m+1} \, dx \, dt \right)^{m-1} S^2 \leq C^{m-1} (2R)^{2(m+1)} |B_{2R}|^{m-1}.
\]

Moreover, this implies that

\[
\left( \int_{Q_{S(z),R}(z)} u^{m+1} \, dx \right)^{\frac{m-1}{m+1}} \leq \frac{R^2}{2^{m+1}S(z)} \leq \frac{R^2}{S(z)} =: \frac{1}{\theta_{r,z}}.
\]
For \( r \in (0, R] \) we define
\[
\tilde{s}(r, z) = \max \left\{ s \in (0, S(z)] \mid \left( \int_{t-s}^{t+s} \int_{B_r(x)} u^{m+1} \, dx \, dt \right)^{m-1} s^2 \leq r^{2(m+1)} |B_r(x)|^{m-1} \right\},
\]
and for \( \hat{b} \in (0, 2) \)
\[
s(r, z) = \min_{r \leq a \leq R} \left( \frac{r}{a} \right)^{\hat{b}} \tilde{s}(a, z).
\]
(5.3)

Furthermore, we define \( \theta_{r,z} = \frac{s(r,z)}{r^2} \). By the same proof of [Sch13, Lemma 4.1], the functions \( s(r, z) \) and \( \theta_{r,z} \) satisfy all properties of Lemma 5.1 stated for \( s(r) \) and \( \theta_r \).

Let us now come to (2). We take \( r \leq \frac{R}{c_o} \), where \( c_o = \max\{\frac{31}{\hat{b}}, 3\} \). This implies that
\[
s(r, y), s(r, z) \leq \frac{S}{3}.
\]
Without loss of generality, we may assume that \( s(r, z) \geq s(r, y) \). By (8) of Lemma 5.1 this implies that
\[
Q_{s(r,y),3r(y)} \subset Q_{3s(r,z),3r(z)} \subset Q_{s(c_or,z),c_or(z)}.
\]
By its definition, there is a \( \rho \in [r, R] \) such that \( s(r, y) = \left( \frac{r}{\rho} \right)^{\hat{b}} \tilde{s}(\rho, y) \). Moreover, since \( s(r, z) \leq \left( \frac{r}{\rho} \right)^{\hat{b}} \tilde{s}(\rho, z) \) by (2) of Lemma 5.1, we have that \( \tilde{s}(\rho, y) \leq s(\rho, z) \).

Now we let \( y = (t, x) \) and \( z = (t_1, x_1) \) and estimate.

If \( \rho \in [\frac{R}{c_o}, R] \), then
\[
\left( \frac{R}{c_o} \right)^{2(m+1)} |B_\rho|^{m-1} \leq \rho^{2(m+1)} |B_\rho|^{m-1} \leq \left( \int_{t-s(\rho,y)}^{t+s(\rho,y)} \int_{B_\rho(x)} u^{m+1} \, dx \, dt \right)^{m-1} \tilde{s}(\rho, y)^2
\]
\[
= \left( \int_{Q_{2R,2S}} u^{m+1} \, dx \, dt \right)^{m-1} \tilde{s}(\rho, y)^2
\]
\[
\leq \left( \int_{Q_{2R,2S}} u^{m+1} \, dx \, dt \right)^{m-1} s(\rho, z)^2 \\
\leq \left( \int_{Q_{2R,2S}} u^{m+1} \, dx \, dt \right)^{m-1} s^2 \\
\leq (2R)^{2(m+1)} |B_{2R}|^{m-1},
\]

which implies that \( s(\rho, z) \leq \tilde{c}_1 \tilde{s}(\rho, y) \), for \( \tilde{c}_1 = (2c_o) \frac{2^{(m+1)+n(m-1)}}{2} \); moreover,

\[
s(r, z) \leq \left( \frac{r}{\rho} \right)^{\hat{b}} s(\rho, z) \leq \tilde{c}_1 \left( \frac{r}{\rho} \right)^{\hat{b}} \tilde{s}(\rho, y) = \tilde{c}_1 s(r, y).
\]

On the other hand, if \( \rho \in [r, \frac{R}{c_o}] \), we similarly have

\[
\rho^{2(m+1)} |B_{\rho}|^{m-1} = \left( \int_{t+\tilde{s}(\rho, y)}^{t-s(\rho, y)} \int_{B_{\rho}(x)} u^{m+1} \, dx \, dt \right)^{m-1} \tilde{s}(\rho, y)^2 \\
\leq \left( \int_{t_1+s(\rho, z)}^{t_1-3s(\rho, z)} \int_{B_{\rho}(x_1)} u^{m+1} \, dx \, dt \right)^{m-1} \tilde{s}(\rho, y)^2 \\
\leq \left( \int_{t_1-s(3\rho, z)}^{t_1+s(3\rho, z)} \int_{B_{3\rho}(x_1)} u^{m+1} \, dx \, dt \right)^{m-1} \tilde{s}(\rho, y)^2 \\
\leq \left( \int_{t_1-s(c_o\rho, z)}^{t_1+s(c_o\rho, z)} \int_{B_{c_o\rho}(x_1)} u^{m+1} \, dx \, dt \right)^{m-1} s(c_o\rho, z)^2 \\
\leq (c_o\rho)^{2(m+1)} |B_{c_o\rho}|^{m-1},
\]

where the last inequality follows by (8) of Lemma 5.1 and the construction. This implies (again) that

\[
s(r, z) \leq \left( \frac{r}{\rho} \right)^{\hat{b}} s(\rho, z) \leq \tilde{c}_1 \left( \frac{r}{\rho} \right)^{\hat{b}} \tilde{s}(\rho, y) = \tilde{c}_1 s(r, y).
\]

Therefore, for \( r \in (0, \frac{R}{c_o}] \), we find that

\[
Q_{s(r, z), r}(z) \subset Q_{\tilde{c}_1 s(r, y), \tilde{c}_1 r}(y) \subset Q_{s(c_1 r, y), c_1 r}(y),
\]

with \( c_1 = c_o \tilde{c}_1 \), which finishes the proof. \( \square \)

The following version of Vitali’s covering can be applied to the cylinders built in Lemma 5.1. It is inspired by [Ste93, Chapter 1, Lemma 1 and Lemma 2]. See also [EV92, Paragraph 1.5, Theorem 1].
Lemma 5.4. Let $\Omega \subset \mathbb{R}^M$ and $R \in \mathbb{R}$. Let there be given a two-parameter family $\mathcal{F}$ of nonempty and open sets

$$\{U(x, r) \mid x \in \Omega, \ r \in (0, R]\},$$

which satisfy two conditions:

1. They are nested, that is,

   $$\text{(5.4)} \quad \text{for any } x \in \Omega, \text{ and } 0 < s < r \leq R, \ U(x, s) \subset U(x, r);$$

2. There exists a constant $c_1 > 1$, such that

   $$\text{(5.5)} \quad U(x, r) \cap U(y, r) \neq \emptyset \implies U(x, r) \subset U(y, c_1 r).$$

Then we can find a disjoint subfamily $\{U_i\}_{i \in \mathbb{N}} = \{U(x_i, r_i)\}_{i \in \mathbb{N}}$, such that

$$\bigcup_{\{x \in \Omega, \ r \in (0, \frac{R}{c_1}]\}} U(x, r) \subset \bigcup_{i \in \mathbb{N}} \tilde{U}_i,$$

with $\tilde{U}_i = U(x_i, 2c_1 r_i)$.

Proof. Without loss of generality, we may assume that $R = c_1$. Now we define the classes

$$Q_k := \left\{ U(x, r) \mid x \in \Omega, \ r \in [2^{-k}, 2^{1-k}] \right\} \quad \text{for } k \in \mathbb{N}.$$

Next we take $\mathcal{Q}_1$ as a maximal disjoint subfamily of $Q_1$: by the fact that $U(x, r)$ are open and $\mathcal{Q}_1$ is a family of disjoint sets, we conclude that $\mathcal{Q}_1$ possesses at most countable many members. Indeed, let $\Omega_1 := \bigcup_{\{x \in \Omega, \ r \in (\frac{1}{2}, 1]\}} U(x, r)$, and $\Omega_1 := \{x \in \Omega_1 : x_i \in \mathbb{Q} \text{ for any } i = 1, \ldots, M\}$. The set $\Omega_1$ is countable, and dense in $\Omega_1$. It is straightforward to see, that for any $x \in \tilde{\Omega}_1$, there exists a unique $U(y, r) \in \mathcal{Q}_1$ which contains $x$, and each $U(y, r) \in \mathcal{Q}_1$ contains at least an element of $\tilde{\Omega}_1$.

Next, we proceed inductively. Assuming that $\mathcal{Q}_1, \mathcal{Q}_2, \ldots, \mathcal{Q}_{k-1}$ have already been selected, we choose $\mathcal{Q}_k$ to be a maximal disjoint subfamily of

$$\text{(5.6)} \quad \left\{ U(x, r) \in Q_k : U(x, r) \cap U(y, \rho) = \emptyset \text{ for all } U(y, \rho) \in \bigcup_{l=1}^{k-1} \mathcal{Q}_l \right\}.$$

Now we enumerate the members of $\bigcup_{l \in \mathbb{N}} \mathcal{Q}_k$ and let

$$\{U_j\}_{j \in \mathbb{N}} = \{U(x_j, r_j)\}_{j \in \mathbb{N}} := \bigcup_{l \in \mathbb{N}} \mathcal{Q}_k.$$

Take any $U(x, r) \in \mathcal{F}$. There exists $k \in \mathbb{N}$ such that $U(x, r) \in Q_k$. By (5.6) and the maximality of $\mathcal{Q}_k$, there exists $i \in \mathbb{N}$, such that $r \leq 2r_i$ and $U(x, 2r_i) \cap U(x_i, 2r_i) \neq \emptyset$; we conclude by (5.5). □
From Lemma 5.4, we immediately deduce the corollary below, which might be of use also in future applications.

**COROLLARY 5.5.** Let \( \Omega \subset \mathbb{R}^M \) and \( R \in \mathbb{R} \). Let there be given a two-parameter family \( \mathcal{F} \) of nonempty and open sets

\[
\{ U(x,r) \mid x \in \Omega, \ r \in (0,R]\},
\]

which satisfy (5.4)–(5.5), and the following third condition:

1. There exists a constant \( a > 1 \) such that, for all \( r \in (0,R]\),

\[
0 < |U(x,2r)| \leq a|U(x,r)| < \infty.
\]

Then we can find a disjoint subfamily \( \{ U_i \}_{i \in \mathbb{N}} = \{ U(x_i,r_i) \}_{i \in \mathbb{N}} \), such that

\[
\bigcup_{\{x \in \Omega, \ r \in (0,R]\}} U(x,r) \subset \bigcup_{i \in \mathbb{N}} \tilde{U}_i,
\]

with \( \tilde{U}_i = U(x_i,2c_1r_i) \), \( |U_i| \sim |\tilde{U}_i| \) and

\[
|\Omega| \leq c \sum_i |U_i|,
\]

where the constant \( c > 1 \) depends only on \( c_1, a \), and the dimension \( M \).

In the following we will build a proper covering with respect to \( u \) for the level sets of the function \( |Du|^{m+1} \).

We assume that we have a sub-intrinsic cylinder

\[
\frac{C}{\theta_0} \geq \left( \int_{Q_{2\theta_0 R^2,2R}} u^{m+1} \, dx \, dt \right)^{\frac{m-1}{m+1}}.
\]

We start by scaling everything to the cylinder \( Q_{2,2} \). This can be done by introducing \( \gamma = \theta_0^{-\frac{1}{M-1}} \). Then we define \( \tilde{u}(y,s) = \gamma u(Ry,\gamma^{m-1}R^2s) \). For this scaled solution we find

\[
C \geq \gamma^{m+1} \int_{Q_{2\gamma^{m-1} R^2,2R}} u^{m+1} \, dx \, dt = \int_{Q_{2,2}} \tilde{u}^{m+1} \, dy \, ds.
\]

Now \( \tilde{u} \) is a weak solution in \( Q_{2,2} \) to

\[
\tilde{u}_s - \text{div} \tilde{A}(y,s,\tilde{u},D\tilde{u}) = \tilde{f}
\]

with right-hand side \( \tilde{f}(y,s) = \gamma^m R^2 f(Ry,\gamma^{m-1}R^2s) \), and \( \tilde{A} \) which satisfies structure conditions analogous to (1.2).
By Lemma 3.2, with the previous scaling, we find
\[ \gamma^{m+1} R^2 \int_{Q_{\gamma^{-1}R^2,2R}} |D(\frac{u^{m+1}}{\gamma^m})|^2 \, dx \, dt = \int_{Q_{1,2}} |D(\frac{\tilde{u}^{m+1}}{\Gamma})|^2 \, dy \, ds \leq c \int_{Q_{2,2}} [\tilde{u}^{m+1} + f^{m+1}] \, dy \, ds \leq C(f). \]

We introduce the notation
\[ F = \left| D(\frac{\tilde{u}^{m+1}}{\Gamma}) \right|^2 \chi_{Q_{2,2}}. \]

We fix
\[ \hat{b} = \frac{4}{m+1}. \]

**Remark 5.6.** The value of \( \hat{b} \) is a direct consequence of the inhomogeneity which characterizes the energy estimates of Lemma 3.1. Before proceeding with the main argument, let us give a purely heuristic justification of (5.11), using some elementary dimensional analysis.

Denoting with the symbol \([\varpi]\) the dimension of the quantity \(\varpi\) (e.g., \([\tau]\) is a time, \([\rho]\) is a length), we notice that the integrands of both sides of (3.1) have dimension
\[ \frac{[u]^2}{[\tau]} + \frac{[u]^{m+1}}{[\rho]^2}; \]
therefore, the homogeneity is restored once we choose a time-length \(\tau\) such that
\[ [\tau] = [u]^{1-m}[\rho]^2. \]

When studying boundedness or pointwise properties of the solution, such as Harnack inequalities (see [DGV], and the references therein), one scales the time variable by a factor \(\theta\) which is not dimensionless: one defines
\[ \tau = \theta \rho^2 \]
and relies on *intrinsic cylinders* reflecting the degeneracy of the equation. The scaling parameter is
\[ \theta = \left( \frac{u(x_o,t_o)}{c} \right)^{1-m}, \]
and the corresponding family of cylinders is
\[ Q_{\theta \rho^2, \rho}(z_o) = (t_o - \theta \rho^2, t_o + \theta \rho^2) \times B_\rho(x_o). \]
However, we are here interested in integrability properties of \( |Du^{m+1}| \) and heuristically, the scaling parameter should now be a quantity related to the gradient \( Du^{m+1} \), and not to the solution \( u \).

From the relation
\[
\left[ Du^{m+1} \right] = \left[ u \right]^{m+1} [\rho]^{-1},
\]
substituting in (5.12), we have
\[
[\tau] = \left[ u \right]^{1-m} [\rho]^2 = \left[ u^{m+1} \right] \frac{2(1-m)}{m+1} [\rho]^2 = \left[ Du^{m+1} \right] \frac{2(1-m)}{m+1} [\rho]^2 + \frac{2(1-m)}{m+1}.
\]
This relation shows that, when dealing with the higher integrability of the gradient, cumbersome how it may look, the time scaling one ends up working with is
\[
\tau \approx \theta \rho^{\frac{4}{m+1}} \quad \text{with} \quad [\theta] = \left[ Du^{m+1} \right] \frac{2(1-m)}{m+1},
\]
and the exponent of \( \rho \) is precisely the value of \( \hat{b} \) given in (5.11).

We will apply Lemma 5.3 with respect to \( \tilde{u} \) and this choice of \( \hat{b} \) on the sub-intrinsic initial cylinder \( Q_{2,2} \). To simplify the notation, for \( y \in Q_{1,1} \) and \( r \in (0,1) \) we denote the sub-intrinsic cylinder \( Q_s(r,y),r \) (\( y \)) defined in Lemma 5.3 with \( Q(r,y) \).

**LEMMA 5.7.** Fix \( c_2 \in [1,\infty) \). For \( \frac{1}{2} \leq a < b \leq 1 \), there exists a parameter \( \lambda \), defined by
\[
\lambda_{a,b} = \frac{C(f)}{|b-a|^\tau},
\]
with \( \tau \) depending only on the data, such that, if for \( z \in Q_{a,a} \) and the corresponding cylinder \( Q(\rho,z) \), the intersection \( Q(c_2 \rho,z) \cap Q_{b,b} \) is not empty, then
\[
\int_{Q(\rho,z)} F \, dx \, dt \leq \lambda_{a,b}.
\]
Here \( C(f) = c \int_{Q_{2,2}} [f^{m+1} + \tilde{u}^{m+1}] \, dx \, dt \), with \( c > 0 \) depending only on the data.

**Proof.** If \( Q(c_2 \rho,z) \cap Q_{b,b} \neq \emptyset \), then either \( c_2 \rho > (b-a) \), or \( s(c_2 \rho,z) > (b-a) \). In the latter case, by (2) of Lemma 5.1, we have that \( (b-a) < c \rho^{\hat{b}} \). This implies that
for $\tilde{b} = \min \{1, \tilde{b}\}$ in any case $c \rho \tilde{b} > (b - a)$. Therefore, (6) of Lemma 5.1 implies

$$\int_{Q(\rho, z)} F \, dx \, dt \leq \frac{C(f)}{\theta_{\rho, z}|b - a|^{(n+2)/\tilde{b}}} \leq \frac{C(f)}{\theta_{1, z}|b - a|^{(n+2 + n^2)/\tilde{b}}} \leq \frac{C(f)}{|b - a|^\tau},$$

where $\theta_{1, z}$ is defined via Lemma 5.3. This defines $\tau$. \hfill \square

Now, we introduce the related intrinsic maximal function

$$\mathcal{M}(g)(z) = \sup_{Q(r, y) \ni z, y \in Q_{1, 1}} \int_{Q(r, y)} |g| \, dx \, dt,$$

and we define

$$M^*(g)(t, x) := \sup_{t \in I \subset (-2, 2), x \in B \subset B_2} \iint_{I \times B} |g| \, dx \, dt.$$ 

Notice that we have

$$|g| \leq \mathcal{M}(g) \leq M^*(g) \quad \text{a.e.},$$

therefore, $\mathcal{M}$ is continuous from $L^q \to L^q$, whenever $g \in L^q$.

We define the level sets of $F$ by

$$O_\lambda := \{\mathcal{M}(F) > \lambda\}.$$ 

The next proposition is the core of the proof of the higher integrability. It constructs a covering, which allows to exploit the reverse Hölder estimates of the previous section in a suitable way. It is a covering of Calderón-Zygmund type for $F$, built using cylinders scaled with respect to $\tilde{u}$.

The fact that the scaling is done with respect to $\tilde{u}$, and not with respect to the function whose level sets are covered, i.e., $F$, makes things quite delicate. In this context, this seems the right way to proceed, instead of relying on the by-now standard approach of parabolic intrinsic Calderón-Zygmund covering, originally introduced by Kinnunen & Lewis [KL00].

**Proposition 5.8.** Let $\frac{1}{2} \leq a < b \leq 1$ and $\lambda_{a, b}$ the quantity defined in Lemma 5.7 for a proper choice of the parameter $c_2$. Let $q \in (0, 1)$ be the exponent defined by the reverse Hölder estimates of Propositions 4.3 and 4.9.

For every $\lambda > \lambda_{a, b}$, and every $z \in O_\lambda \cap Q_{a, a}$, there exist $Q_z \subset Q_z^* \subset Q_z^{**} \subset Q_{b, b}$, which satisfy the following properties:

1. $|Q_z|, |Q_z^*|, |Q_z^{**}|$ are of comparable size.
(2) For any \( y \in Q^*_z \)

\[
\lambda \leq c \int_{Q_z} F \, dx \, dt \leq c \left( \int_{Q^*_z} F^q \, dx \, dt \right)^{\frac{1}{q}} + cM^* (\frac{r^{m+1}}{m}) (y).
\]

(3) \( \int_{Q^*_{z}} F \, dx \, dt \leq 2\lambda \).

Moreover, the set \( O_\lambda \cap Q_{a,a} \) can be covered by a family of cylinders \( Q_i^{**} := Q_{z,i}^{**} \subset Q_{b,b} \), such that the cylinders of the family \( Q_i^* \) are pairwise disjoint.

**Proof.** Let \( c_2 \in [1, \infty) \) the parameter introduced in Lemma 5.7: choose \( c_2 = 4c_1 \), where \( c_1 \) is the constant determined in Lemma 5.3, and let \( \lambda_{a,b} \) be the corresponding quantity defined in Lemma 5.7. We fix \( \lambda > \lambda_{a,b} \), and for \( z \in O_\lambda \cap Q_{a,a} \) we choose \( y_z \) and \( r_z \), such that \( Q(r_z, y_z) \ni z \) and

\[
\lambda < \int_{Q(r_z, y_z)} F \, dx \, dt \quad \text{and} \quad \int_{Q(\rho, \xi)} F \, dx \, dt \leq 2\lambda
\]

for all \( Q(\rho, \xi) \supset Q(r_z, y_z) \); such a cylinder certainly exists by the very definition of the set \( O_\lambda \). In the following notation, for \( c \in \mathbb{R}^+ \), we let \( \tilde{c} = \max \{c^{1/\tilde{b}}, c\} \). According to the following table we will carefully choose cylinders, for \( z \in O_\lambda \cap Q_{a,a} \).

**Case 1.** There exists \( \rho \in [4r_z, 8r_z] \) such that \( Q(\rho, y_z) \) is intrinsic and the degenerate alternative holds in \( Q(\rho, y_z) \). Then we let

- \( Q_z = Q(\rho, y_z), \quad Q_z^* = 2Q(\rho, y_z), \quad Q_z^{**} = Q(4c_1 \rho, y_z) \).

**Case 2.** There exists \( \rho \in [4r_z, 8r_z] \) such that \( Q(\rho, y_z) \) is intrinsic and the non-degenerate alternative holds in \( Q(\rho, y_z) \). Then we let

- \( Q_z = \frac{1}{4}Q(\rho, y_z), \quad Q_z^* = \frac{1}{4}Q(\rho, y_z), \quad Q_z^{**} = Q(2c_1 \rho, y_z) \).

**Case 3.** There exists no \( \rho \in [4r_z, 8r_z] \) such that \( Q(\rho, y_z) \) is intrinsic. Then we let

- \( Q_z = Q(4r_z, y_z), \quad Q_z^* = 2Q(4r_z, y_z), \quad Q_z^{**} = Q(16c_1 r_z, y_z) \).

The choice of \( c_2 \) implies that \( Q_z^{**} \subset Q_{b,b} \) in all the above cases, as otherwise there is a contradiction to \( \lambda > \lambda_{a,b} \) by (5.17).

On the one hand, (8) of Lemma 5.1 implies that \( Q_z^{**} \supset Q_z \ni Q(r_z, y_z) \), and therefore we conclude that

\[
\int_{Q^*_z} F \, dx \, dt \leq 2\lambda.
\]
On the other hand, since (7) of Lemma 5.1 implies that $|Q_z| \approx |Q_z^{**}| \approx |Q(r_z, y_z)|$, by (5.17) we find that

\begin{equation}
\lambda \leq c \int_{Q_z} F \, dx dt.
\end{equation}

The proof of the reverse Hölder inequality has to be split in several sub-cases.

**Case 1.** In this case $Q_z$ is intrinsic and the degenerate alternative holds. In order to apply Proposition 4.3, we have to check that $2Q_z$ is sub-intrinsic. To prove this, firstly observe $2Q_z \subset Q(2\rho, y_z)$, which is sub-intrinsic by construction. Secondly, since $2Q$ and $Q(2\rho, y_z)$ have comparable measure by (7) of Lemma 5.1, we obtain that $2Q_z$ is sub-intrinsic. Consequently, (5.18) and Proposition 4.3 imply

\[ \lambda \leq c \int_{Q_z} F \, dx dt \leq c \left( \int_{Q_z^{**}} F^q \, dx dt \right)^{\frac{1}{q}} + c M^*(\tilde{f}^{m+1}_m)(y), \]

for any $y \in Q_z^{**}$.

**Case 2.** In this case $4Q_z$ is intrinsic and the non-degenerate alternative holds. Proposition 4.9 and (5.18) directly imply

\[ \lambda \leq c \int_{Q_z} F \, dx dt \leq c \left( \int_{Q_z^{**}} F^q \, dx dt \right)^{\frac{1}{q}} + M^*(\tilde{f}^{m+1}_m)(y), \]

for any $y \in Q_z^{**}$.

**Case 3.** This is the most delicate part. We begin by applying Proposition 3.4 on the cylinder $Q_z$, with $\delta = \tilde{\delta} \in (0, 1)$ to be chosen later. Together with (5.18), we find

\[ \lambda \leq c \int_{Q_z} F \, dx dt \leq c \left( \int_{Q_z^{**}} F^q \, dx dt \right)^{\frac{1}{q}} + M^*(\tilde{f}^{m+1}_m)(y) + \frac{2\delta}{(4r_z)^2} \left( u^{m+1} \right)_{Q_z^{**}} + \frac{1}{2\theta^{m+1}_{4r_z, y_z}}. \]

By (3), (6), and (8) of Lemma 5.1 we find that

\[ (\tilde{u}^{m+1})_{Q_z^{**}} \leq c (\tilde{u}^{m+1})_{Q(24r_z, y_z)} \leq \frac{c}{\theta^{m+1}_{24r_z, y_z}} \leq \frac{c}{\theta^{m+1}_{4r_z, y_z}}. \]
This implies that

$$\tag{5.19} \lambda \leq c \int_{Q_z} F \, dx \, dt \leq c \left( \int_{Q_z} F^q \, dx \, dt \right)^{\frac{1}{q}} + M^*(\tilde{f}^{m+1}) (y) + \frac{c \delta}{(4r_z)^2 \theta_{\frac{m+1}{2}y_z}}. $$

In the following we will show that

$$\tag{5.20} \frac{1}{(4r_z)^2 \theta_{\frac{m+1}{2}y_z}} \leq c \lambda + M^*(\tilde{f}^{m+1}) (y). $$

Once this is proven, then for a proper choice of $\delta$ by absorption the result follows from (5.19).

We begin by defining $\rho_z := \inf \{ r \in [4r_z, 1] : Q(r, y_z) \}$ is intrinsic. Since we are in Case 3, we find that $\rho_z \in (8r_z, 1];$ (5) of Lemma 5.1, the choice of $\hat{b}$ in (5.11), and the previous fact together imply that

$$\tag{5.21} \frac{1}{(4r_z)^2 \theta_{\frac{m+1}{2}y_z}} \leq \frac{1}{\rho_z^2 \theta_{\frac{m+1}{2}y_z}}. $$

In the simple case $\rho_z \in (\frac{1}{2}, 1]$, we further estimate by (6) of Lemma 5.1

$$ \frac{1}{(4r_z)^2 \theta_{\frac{m+1}{2}y_z}} \leq c \leq \lambda, $$

by the choice of $\lambda_{a,b}$.

In the difficult case $\rho_z \in (8r_z, \frac{1}{2}]$, since $Q(r, y_z)$ is intrinsic, by (5.21) we find that

$$ \frac{c}{(4r_z)^2 \theta_{\frac{m+1}{2}y_z}} \leq c \int_{Q(\rho_z, y_z)} \frac{\tilde{u}^{m+1}}{\theta_{\frac{m+1}{2}y_z}} \, dx \, dt. $$

In order to apply Proposition 4.3 on the cylinder $Q(\rho_z, y_z)$, which is intrinsic, we have to show that $2Q(\rho_z, y_z)$ is sub-intrinsic and that in $Q(\rho_z, y_z)$ the degenerate alternative holds. We first show that in $Q(\rho_z, y_z)$ the degenerate condition is satisfied: this is due to the fact, that by the choice of $\rho$ we find that $Q(r, y_z)$ is strictly sub-intrinsic for all $r \in [\frac{1}{2} \rho_z, \rho_z]$. Now (6), (5), and (4) of Lemma 5.1 imply

$$ (\tilde{u}^{m+1})^\frac{m+1}{2} \leq \frac{1}{\theta_{\frac{m+1}{2}y_z}} \leq \left( \frac{1}{2} \right)^{2-\hat{b}} \frac{1}{\theta_{\frac{m+1}{2}y_z}} \leq \left( \frac{1}{2} \right)^{2-\hat{b}} (\tilde{u}^{m+1})^\frac{m+1}{2}.$$  

Lemma 5.2 implies that in $Q(\rho_z, y_z)$ the degenerate alternative holds.
That $2Q(\rho_z, y_z)$ is sub-intrinsic follows again by the fact that $2Q(\rho_z, y_z)$ has size comparable to $Q(2\rho_z, y_z)$, which is sub-intrinsic and a superset of $2Q(\rho_z, y_z)$ by (7), (3), and (8) of Lemma 5.1.

Finally Proposition 4.3, Jensen's inequality and (5.17) imply that

$$
\frac{c}{(4r_z)^2} \int_{Q(\rho_z, y_z)} \frac{\tilde{u}^{m+1}}{\rho_z^2} \, dx \, dt \leq c \int_{Q(\rho_z, y_z)} F \, dx \, dt + cM^*(\frac{\tilde{f}^{m+1}}{m})(y)
$$

$$
\leq c\lambda + cM^*(\frac{\tilde{f}^{m+1}}{m})(y).
$$

This concludes the construction of $Q_z, Q^*_z$ and $Q^{**}_z$.

The covering is gained by applying Corollary 5.5. Indeed, since

$$
O_\lambda \cap Q_{a,a} \subset \bigcup_{z \in O_\lambda \cap Q_{a,a}} Q^*_z,
$$

due to Lemma 5.3 and (7) of Lemma 5.1, Corollary 5.5 implies that we can find an at most countable subfamily of disjoint cylinders $Q^*_i = Q^*_z$, such that

$$
O_\lambda \cap Q_{a,a} \subset \bigcup_{i} Q^{**}_i,
$$

which concludes the proof.

We can finally conclude and prove the higher integrability result.

**Theorem 5.9.** (Intrinsic) Let $u \geq 0$ be a local, weak solution to (1.1)-(1.2) in the space-time cylinder $E_T$ for $m > 1$. There exist an exponent $p > 1$ and a constant $c$ that depend only on the data, such that for any sub-intrinsic parabolic cylinder

$$
\frac{C}{\theta_o} \geq \left( \int_{Q_{2\theta_o R^2, 2R}} u^{m+1} \, dx \, dt \right)^{\frac{m-1}{m+1}},
$$

we have

$$
\left( \int_{Q_{\frac{1}{2}\theta_o R^2, \frac{1}{2}R}} |Du|^{2p} \, dx \, dt \right)^{\frac{m-1}{m(m+1)}} \leq c \left( \int_{Q_{2\theta_o R^2, 2R}} \frac{p(m+1)}{m} \, dx \, dt \right)^{\frac{m-1}{m+1}}
$$

$$
+ \frac{c'}{R \left( \frac{2(m-1)}{m} \right)^{\frac{1}{m+1}} \theta_o},
$$

where $c'$ additionally depends on $C$. 

Proof. We define the so-called bad set

\[ U_\lambda := O_\lambda \cap \left\{ M^* \left( \frac{f}{m} \right)^\chi_{Q_{2,2}} \leq \varepsilon \lambda \right\}, \]

for some \( \varepsilon \), which will be chosen later. We proceed by providing a re-distributional estimate. We take \( \frac{1}{2} \leq a < b \leq 1 \) and the corresponding covering constructed in Proposition 5.8. We start by

\[
|Q_i| = |Q_i \cap U_\lambda| + |Q_i \cap (U_\lambda)^c|.
\]

Let us first consider the case

\[
\frac{|Q_i \cap U_\lambda|}{|Q_i|} \geq \frac{1}{2}.
\]

This implies that there exists \( y \in Q_i \), such that \( M^* \left( \frac{f}{m} \right)^\chi_{Q_{2,2}}(y) \leq \varepsilon \lambda \). We can apply the reverse Hölder estimate of Proposition 5.8, and obtain for some \( \gamma \in (0,1) \) that

\[
\lambda^q \leq c \left( \int_{Q_i} F \, dx \, dt \right)^q \leq \frac{c}{|Q_i|} \int_{Q_i} F^q \chi_{\{F > \gamma \lambda\}} \, dx \, dt + c(\gamma \lambda)^q + c\varepsilon \lambda^q.
\]

We now choose \( \gamma \), and \( \varepsilon \) conveniently small, such that \( c(\gamma \lambda)^q + c\varepsilon \lambda^q = \frac{1}{2} \lambda^q \) and find

\[
\lambda |Q_i| \leq c \lambda^{1-q} \int_{Q_i} F^q \chi_{\{F > \gamma \lambda\}} \, dx \, dt.
\]

On the other hand,

\[
|Q_i| \leq 2 |Q_i \cap (U_\lambda)^c| \implies \lambda |Q_i| \leq 2 \lambda |Q_i \cap (U_\lambda)^c|.
\]

Therefore, in any case,

\[
\lambda |Q_i| \leq c \lambda^{1-q} \int_{Q_i} F^q \chi_{\{F > \gamma \lambda\}} \, dx \, dt + 2 \lambda |Q_i \cap (U_\lambda)^c|.
\]
We proceed by using the last estimates as well as the fact that \( \{Q_i^{**}\}_i \) covers the set \( O_\lambda \cap Q_{a,a} \):

\[
\int_{Q_{a,a} \cap O_\lambda} F \, dx \, dt \\
\leq \sum_i \int_{Q_i^{**}} F \, dx \, dt \\
\leq 2\lambda \sum_i |Q_i^{**}| \\
\leq c\lambda \sum_i |Q_i| \\
\leq c\lambda^{1-q} \sum_i \int_{Q_i^{**}} F^q \chi_{\{F > \gamma \lambda\}} \, dx \, dt + 2\lambda |Q_i \cap (U_\lambda)^c| \\
\leq c\lambda^{1-q} \int_{Q_{b,b}} F^q \chi_{\{F > \gamma \lambda\}} \, dx \, dt + 2\lambda |Q_{b,b} \cap \{M^*(\tilde{f}_{m+1} \chi_{Q_{2,2}}) > \bar{\varepsilon} \lambda\}|.
\]

We pick \( \alpha \in (0, 1) \) to be chosen later, \( k \in \mathbb{N} \), multiply the above estimate by \( \lambda^{-\alpha} \), and integrate from \( \lambda_{a,b} \) to \( k \) with respect to \( \lambda \). This implies

\[
(I) := \int_{\lambda_{a,b}}^k \lambda^{-\alpha} \int_{Q_{a,a} \cap O_\lambda} F \, dx \, dt \, d\lambda \leq c \int_{\lambda_{a,b}}^k \lambda^{1-q-\alpha} \int_{Q_{b,b}} F^q \chi_{\{F > \gamma \lambda\}} \, dx \, dt \, d\lambda \\
+ 2 \int_{\lambda_{a,b}}^k \lambda^{1-\alpha} \left| Q_{b,b} \cap \{M^*(\tilde{f}_{m+1} \chi_{Q_{2,2}}) > \bar{\varepsilon} \lambda\} \right| \, d\lambda =: (II) + (III).
\]

We estimate from below

\[
(I) \geq \int_0^k \lambda^{-\alpha} \int_{Q_{a,a}} F \chi_{\{F > \lambda\}} \, dx \, dt \, d\lambda - \frac{\lambda_{a,b}^{2-\alpha}}{1-\alpha} \\
= \int_{Q_{a,a}} F \left( \min\{F(x,t), k\} \right) \lambda^{-\alpha} \, d\lambda \, dx \, dt - \frac{\lambda_{a,b}^{2-\alpha}}{1-\alpha} \\
= \frac{1}{1-\alpha} \int_{Q_{a,a}} F \min\{F, k\}^{1-\alpha} \, dx \, dt - \frac{\lambda_{a,b}^{2-\alpha}}{1-\alpha}.
\]

The bound from above is analogous

\[
(II) \leq c \int_{\gamma \lambda_{a,b}}^{\gamma k} \lambda_{a,b}^{1-q-\alpha} \int_{Q_{b,b}} F^q \chi_{\{F > \lambda\}} \, dx \, dt \, d\lambda \\
\leq \frac{c}{2-q-\alpha} \int_{Q_{b,b}} F^q \min\{F, k\}^{2-q-\alpha} \, dx \, dt.
\]
Finally, \((III)\) is estimated by the continuity of the maximal function and the classical integral representation via level sets. We calculate and estimate

\[
(III) = \int_{\lambda_{a,b}}^{k} \lambda^{1-\alpha} |Q_{b,b} \cap \{ M^*(\tilde{f}^{m+1}_{m}) > \tilde{\varepsilon} \lambda \} | d\lambda
\]

\[
\leq c \int_{Q_{b,b}} \left| M^*(\tilde{f}^{m+1}_{m} \chi_{Q_{2,2}}) \right|^{2-\alpha} dx dt
\]

\[
\leq c \int_{Q_{2,2}} \tilde{f}^{(m+1)(2-\alpha)}_{m} dx dt.
\]

All together, using the definition of \(\lambda_{a,b}\) from (5.13), we find that

\[
\frac{1}{1-\alpha} \int_{Q_{a,a}} F \min \{F, k\}^{1-\alpha} dx dt \leq \frac{c}{2-q-\alpha} \int_{Q_{b,b}} F^q \min \{F, k\}^{2-q-\alpha} dx dt
\]

\[
+ \frac{c|b-a|^{-(2-\alpha)}}{1-\alpha} + c \int_{Q_{2,2}} \tilde{f}^{(m+1)(2-\alpha)}_{m} dx dt.
\]

Now, we fix \(\alpha \in (0,1)\) in such a way, that

\[
\frac{c(1-\alpha)}{2-q-\alpha} \leq \frac{1}{2}.
\]

This implies

\[
\int_{Q_{a,a}} F \min \{F, k\}^{1-\alpha} dx dt \leq \frac{1}{2} \int_{Q_{b,b}} F \min \{F, k\}^{1-\alpha} dx dt
\]

\[
+ \frac{c|b-a|^{-(2-\alpha)}}{1-\alpha} + c \int_{Q_{2,2}} \tilde{f}^{(m+1)(2-\alpha)}_{m} dx dt.
\]

Finally, the interpolation Lemma 6.1 of [Giu03] implies that for every \(k \in \mathbb{N}\)

\[
\int_{Q_{2,2}^{1/2}} F \min \{F, k\}^{1-\alpha} dx dt \leq c + c \int_{Q_{2,2}} \tilde{f}^{(m+1)(2-\alpha)}_{m} dx dt.
\]

Letting \(k \to \infty\) for \(p = 2 - \alpha > 1\) yields that

\[
\int_{Q_{2,2}^{1/2}} \left| D\tilde{u}_{m+1}^{\frac{m+1}{2}} \right|^{2p} dx dt \leq c + c \int_{Q_{2,2}} \tilde{f}^{(m+1)(2-\alpha)}_{m} dx dt.
\]

This implies the desired result by scaling back to \(u\). \(\square\)

**Theorem 5.10.** (parabolic) Let \(u \geq 0\) be a local, weak solution to (1.1)-(1.2) in the space-time cylinder \(E_T^m\) for \(m > 1\). There exist an exponent \(p > 1\)
and a constant $c$, depending only on the data, such that for any parabolic cylinder $Q_{R^2,R} \subset E_T$ with

$$\left( \int_{Q_{R^2,R}} u^{m+1} \, dx \, dt \right)^{m-1} = K,$$

we have

$$\left( \int_{\frac{1}{2} Q_{R^2,R}} |Du|^{\frac{m+1}{2p}} \, dx \, dt \right)^{\frac{m-1}{p(m+1)}} \leq c\sqrt{K} \left( \int_{Q_{R^2,R}} \frac{p(m+1)}{m} f \, dx \, dt \right)^{\frac{m-1}{p(m+1)}} + cK^{\frac{3}{2}} + c. \quad (5.22)$$

**Proof.** Estimate (5.22) is proved by covering $Q_{R^2,R}$ with proper sub-intrinsic cylinders. As (1.1)-(1.2) is essentially invariant under the classical parabolic scaling, without loss of generality, we may assume $R = 1$.

We define $K$ as the number for which

$$\left( \int_{Q_{1,1}} u^{m+1} \, dx \, dt \right)^{m-1} = K.$$

Now let $r \in (0, 1]$ and $s \in (0, 1]$; any $Q_{s,r} \subset Q_{1,1}$ satisfies

$$\left( \int_{Q_{s,r}} u^{m+1} \, dx \, dt \right)^{m-1} \leq K.$$

If $\frac{2(m+1)|B_r|^{m-1}}{s^2} \geq K$, then the cylinder is sub-intrinsic, since

$$\left( \int_{Q_{s,r}} u^{m+1} \, dx \, dt \right)^{m-1} \leq \frac{r^{2(m+1)}|B_r|^{m-1}}{s^2}.$$

If $K \leq 1$, we can pick $s,r = 1$ and the result follows by Theorem 5.9. If $K > 1$, we choose $r = 1$ and $s = \frac{1}{\sqrt{K}}$. We can then cover $Q_{1,1}$ by $N$ sub-cylinders of the above type, where

$$\left\lfloor \sqrt{K} \right\rfloor \leq N \leq \left\lceil \sqrt{K} \right\rceil,$$

and $\lfloor \cdot \rfloor$, $\lceil \cdot \rceil$ are the floor and ceiling functions, respectively. This concludes the proof by Theorem 5.9. \qed
REFERENCES

[AM07] E. Acerbi and G. Mingione, Gradient estimates for a class of parabolic systems, *Duke Math. J.* **136** (2007), no. 2, 285–320.

[Bare52] G. I. Barenblatt, On some unsteady motions of a liquid and gas in a porous medium, *Akad. Nauk SSSR. Prikl. Mat. Meh.* **16** (1952), 67–78.

[Ben83] P. Bénilan, A strong regularity $L^p$ for solution of the porous media equation, *Contributions to Nonlinear Partial Differential Equations (Madrid, 1981)*, *Res. Notes in Math.*, vol. 89, Pitman, Boston, MA, 1983, pp. 39–58.

[Bög08] V. Bögelein, Higher integrability for weak solutions of higher order degenerate parabolic systems, *Ann. Acad. Sci. Fenn. Math.* **33** (2008), no. 2, 387–412.

[BZW13] V. Bögelein, F. Duzaar, and G. Mingione, The regularity of general parabolic systems with degenerate diffusion, *Memo. Amer. Math. Soc.* **221** (2013), no. 1041, vi+143.

[BP10] V. Bögelein and M. Parviainen, Self-improving property of nonlinear higher order parabolic systems near the boundary, *NoDEA Nonlinear Differential Equations Appl.* **17** (2010), no. 1, 21–54.

[BCDKS15] D. Breit, A. Cianchi, L. Diening, T. Kuusi, and S. Schwarzacher, Pointwise Calderón-Zygmund gradient estimates for the $p$-Laplace system, *J. Math. Pures Appl.* **98** (2017), 146–190.

[DiB93] E. DiBenedetto, *Degenerate Parabolic Equations*, Universitext, Springer-Verlag, New York, 1993.

[DBF85] E. DiBenedetto and A. Friedman, Hölder estimates for nonlinear degenerate parabolic systems, *J. Reine Angew. Math.* **357** (1985), 1–22.

[DGV] E. DiBenedetto, U. Gianazza, and V. Vespri, *Harnack’s Inequality for Degenerate and Singular Parabolic Equations*, Springer Monogr. Math., Springer-Verlag, New York, 2012.

[DKS11] L. Diening, P. Kaplický, and S. Schwarzacher, BMO estimates for the $p$-Laplacian, *Nonlinear Anal.* **75** (2012), no. 2, 637–650.

[DLSV11] L. Diening, D. Lengeler, B. Stroffolini, and A. Verde, Partial regularity for minimizers of quasiconvex functionals with general growth, *SIAM J. Math. Anal.* **44** (2012), no. 5, 3594–3616.

[EV92] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, revised ed., *Texts Math.*, CRC Press, Boca Raton, FL, 2015.

[FSch15] J. Frehse and S. Schwarzacher, On regularity of the time derivative for degenerate parabolic systems, *SIAM J. Math. Anal.* **47** (2015), no. 5, 3917–3943.

[Geh73] F. W. Gehring, The $L^p$-integrability of the partial derivatives of a quasiconformal mapping, *Acta Math.* **130** (1973), 265–277.

[GM79] M. Giaquinta and G. Modica, Regularity results for some classes of higher order nonlinear elliptic systems, *J. Reine Angew. Math.* **311/312** (1979), 145–169.

[GS82] M. Giaquinta and M. Struwe, On the partial regularity of weak solutions of nonlinear parabolic systems, *Math. Z.* **179** (1982), no. 4, 437–451.

[Giusti03] E. Giusti, *Direct Methods in the Calculus of Variations*, World Scientific Publishing, River Edge, NJ, 2003.
[KL00] J. Kinnunen and J. L. Lewis, Higher integrability for parabolic systems of \( p \)-Laplacian type, *Duke Math. J.* 102 (2000), no. 2, 253–271.

[KM14/1] T. Kuusi and G. Mingione, A nonlinear Stein theorem, *Calc. Var. Partial Differential Equations* 51 (2014), no. 1-2, 45–86.

[KM14/2] T. Kuusi and G. Mingione, Riesz potentials and nonlinear parabolic equations, *Arch. Ration. Mech. Anal.* 212 (2014), no. 3, 727–780.

[ME75] N. G. Meyers and A. Elcrat, Some results on regularity for solutions of non-linear elliptic systems and quasi-regular functions, *Duke Math. J.* 42 (1975), 121–136.

[Par09a] M. Parviainen, Global gradient estimates for degenerate parabolic equations in nonsmooth domains, *Ann. Mat. Pura Appl. (4)* 188 (2009), no. 2, 333–358.

[Par09b] M. Parviainen, Reverse Hölder inequalities for singular parabolic equations near the boundary, *J. Differential Equations* 246 (2009), no. 2, 512–540.

[Sch13] S. Schwarzacher, Hölder-Zygmund estimates for degenerate parabolic systems, *J. Differential Equations* 256 (2014), no. 7, 2423–2448.

[Ste93] E. M. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton Math. Ser., vol. 43, Princeton University Press, Princeton, NJ, 1993.

[Str80] E. W. Stredulinsky, Higher integrability from reverse Hölder inequalities, *Indiana Univ. Math. J.* 29 (1980), no. 3, 407–413.