Advances in Differential Equations Volume 14, Numbers 1-2 (2009), 121–200

VERY WEAK SOLUTIONS OF HIGHER-ORDER DEGENERATE PARABOLIC SYSTEMS

Verena Bögelein
Department Mathematik
Friedrich-Alexander-Universität Erlangen–Nürnberg
Bismarckstr. 1 1/2, 91054 Erlangen, Germany

(Submitted by: Emmanuele DiBenedetto)

Abstract. We consider non-linear higher-order parabolic systems whose simplest model is the parabolic p-Laplacean system

$$\int_{\Omega_T} u\cdot\varphi_t - \langle |D^m u|^{p-2} D^m u, D^m\varphi\rangle\, dz = 0.$$

It turns out that the usual regularity assumptions on solutions can be weakened in the sense that going slightly below the natural integrability exponent still yields a classical weak solution. Namely, we prove the existence of some $\beta > 0$ such that $D^m u \in L^{p-\beta} \Rightarrow D^m u \in L^{p+\beta}$.

1. Introduction

The aim of this paper is to provide regularity results for solutions to possibly degenerate parabolic systems of partial differential equations below those ranges which are traditionally considered to be "natural." In regularity theory for elliptic, respectively parabolic, partial differential equations one usually considers weak solutions, lying in a certain "natural" function space associated to the problem, and dictated by the coercivity properties of the operators involved.

When starting with such a weak solution, one is interested in proving better regularity properties. But on the other hand there naturally arises the question whether this initial condition can be weakened. This leads us to the notion of very weak solutions, i.e., solutions lying slightly below the natural function space. To be more precise, let us consider the parabolic p-Laplacian system of order $2m$, $m \in \mathbb{N}$, in its weak formulation, which is the model problem for our results:

$$\int_{\Omega_T} u\cdot\varphi_t - \langle |D^m u|^{p-2} D^m u, D^m\varphi\rangle\, dz = 0 \qquad \forall\, \varphi \in C_0^\infty(\Omega_T;\mathbb{R}^N). \quad (1.1)$$

Accepted for publication: September 2008.
AMS Subject Classifications: 35D10, 35G20, 35K65.

Here, the associated natural function space for weak solutions is the parabolic Sobolev space $L^p(-T,0;W^{m,p}(\Omega;\mathbb{R}^N)) \cap L^2(\Omega_T;\mathbb{R}^N)$. But the weak formulation (1.1) of the problem already makes sense for functions $u \in L^{p-\beta}(-T,0;W^{m,p-\beta}(\Omega;\mathbb{R}^N)) \cap L^1(\Omega_T;\mathbb{R}^N)$, where $0 < \beta \le \min\{1, p-1\}$. This observation gives reason for the following.

Definition 1.1. *Let $0 < \beta \le \min\{1, p-1\}$. Then a function $u \in L^{p-\beta}(-T,0;W^{m,p-\beta}(\Omega;\mathbb{R}^N)) \cap L^2(\Omega_T;\mathbb{R}^N)$ satisfying* (1.1), *respectively* (2.1), *below is called a very weak solution of the associated parabolic system.*

The main result of this paper, presented in Theorem 2.1, is to ensure the existence of an exponent $\beta > 0$ depending only on the data such that any very weak solution already is a weak solution. At a first glance this solely is a question of higher integrability. But, accordingly it can also be interpreted as a question of uniqueness of solutions in the function space $L^{p-\beta}(-T,0;W^{m,p-\beta}(\Omega;\mathbb{R}^N)) \cap L^2(\Omega_T;\mathbb{R}^N)$. Hence, our result can be reformulated as follows: going slightly below the natural integrability exponent preserves uniqueness of solutions. But it is not known what the largest possible value of β is - not even in the elliptic case.

Although higher integrability properties [5, 16, 18, 22, 26], Calderón & Zygmund estimates [1, 2, 24, 25] and also higher regularity properties [14] for weak solutions are widely studied, the literature for solutions lying below the natural exponent is rare. For linear elliptic systems higher integrability of very weak solutions was achieved by Elcrat and Meyers [15] by a duality argument, exploiting higher integrability of the adjoint problem. To treat non-linear elliptic systems there are mainly two different approaches. The one given by Iwaniec and Sbordone [20, 21] is based on a non-linear Hodge decomposition, while the method developed by Lewis [27] makes use of the Hardy-Littlewood maximal function and the Whitney extension theorem. Note that the latter paper also treats higher-order elliptic systems. The main difficulty when considering very weak solutions is the fact that the solution itself is a priori not an admissible test-function. Therefore, one first has to construct a suitable test-function, namely via the Hodge decomposition in [21], respectively with Whitney's extension theorem in [27]. But none of these techniques applies in the parabolic case, due to the difficulties arising from the additional time direction. Nevertheless, Kinnunen and Lewis [23] succeeded to treat very weak solutions for non-linear second order parabolic systems. The main idea is to construct a suitable test-function in a very delicate way consisting of mean values of the solution multiplied by a partition of unity subordinate to a Whitney decomposition. In the present paper we extend this result to parabolic systems of higher order $2m \geq 2$. Again, the crucial point is to construct an admissible test-function. In order to approximate derivatives of order m, we will construct the test-function by mean value polynomials of order $m-1$ of the solution on Whitney-type cylinders multiplied by a partition of unity. Then we will show very fine estimates, in particular for the involved polynomials, which are consistent with the geometry coming from the Whitney cylinders.

We now briefly sketch the main points of the proof. A basic observation is that in the degenerate, respectively singular case, i.e., when $p \neq 2$, we will work on intrinsic parabolic cylinders, invented by DiBenedetto [9, 10, 11, 12, 13]. By this we mean that the scaling of the cylinders depends on the solution itself; i.e., when $|D^m u(z_0)| \approx \lambda$, then we take cylinders around z_0 with radius $\varrho > 0$ in space-direction and $\lambda^{2-p}\varrho^{2m}$ in time-direction. The scaling factor λ^{2-p} then helps to compensate the non-homogeneous behavior of the parabolic system. On the other hand, working within this intrinsic geometry complicates the proof in several respects. Since there is no uniform system of cylinders available, on which we could define the natural maximal function, we shall use the so called strong maximal function, defined in (5.1) below, which possesses only weaker properties than the natural one. For instance, it is not known that $M(|D^m u|)^{-\beta}$ is a Muckenhoupt-weight. But this is an essential property used in the proof for the elliptic case in [27].

Another basic difference compared to elliptic problems is that we cannot apply the Poincaré inequality to solutions of parabolic systems. This is due to the lack of differentiability with respect to the time direction. Nevertheless, exploiting the parabolic system we can show a Poincaré type inequality valid for solutions. Such Poincaré type inequalities are provided in Chapter 4, since they will be frequently needed throughout the whole paper.

Chapter 5 is devoted to the proof of a Caccioppoli type inequality. As already mentioned above this is the crucial step when considering very weak solutions, since the solution itself is not an admissible test-function. To construct our test-function, we will change the solution u, multiplied by a cut-off function, on the "bad set" where the strong maximal function of $|D^m u|$ is large, i.e., larger then some fixed parameter λ. The construction is as follows: We cover this bad set by intrinsic Whitney-type cylinders with scaling factor λ^{2-p}. Then we define the test-function by mean value polynomials of order $m-1$ of the solution on Whitney-type cylinders multiplied by a partition of unity. Here, we emphasize that two different systems of intrinsic cylinders are involved, namely the one on which we will prove the Caccioppoli inequality and the one coming from the Whitney cylinders. In the following we will show that the constructed function on the one hand is regular enough to play as a test-function and on the other hand it is in a certain sense similar to the original solution. This will be achieved by very fine estimates, namely estimates for the mean value polynomials and Poincaré type estimates on the Whitney cylinders. The crucial point here is that we always have to respect both systems of cylinders. Having ensured these properties we are in a position to test the parabolic system with our test-function. Next, we decompose the domain of integration into the "good set," where the test-function equals the original solution and the "bad set." Using the portion on the good set and exploiting the estimates proved before for the terms on the bad set, we obtain suitable estimates on the lower level sets of the maximal function of $|D^m u|$ at level λ. Integrating with respect to the parameter λ we finally deduce the desired Caccioppoly type inequality.

In Chapter 6 we finally prove our main result. Here we first show a reverse Hölder type inequality. This will be achieved with the help of our Caccioppoli inequality and the Poincaré type inequality we have proved before. It is worth mentioning that the reverse Hölder inequality will only hold on particular intrinsic cylinders satisfying hypotheses (6.3) and (6.4). Therefore, in the final proof of the higher integrability we carefully have to choose such intrinsic cylinders. Hence, we know that on these particular cylinders there holds a reverse Hölder inequality. This will be enough to apply a modified version of Gehring's theorem, which finally yields the desired higher integrability result.

2. Statement of the result

Let $\Omega \subset \mathbb{R}^n$ be a bounded open set and $\Omega_T \equiv \Omega \times (-T, 0)$ $(T > 0)$ the parabolic cylinder over Ω. We consider non-linear parabolic systems of the following form

$$\int_{\Omega_T} u \cdot \varphi_t - \langle A(z, D^m u), D^m \varphi \rangle \, dz = \int_{\Omega_T} \langle B(z, D^m u), \delta\varphi \rangle \, dz \tag{2.1}$$

for all $\varphi \in C_0^\infty(\Omega_T; \mathbb{R}^N)$, where $u \colon \mathbb{R}^{n+1} \to \mathbb{R}^N$ with $N, m \geq 1$ will always be a very weak solution, see Definition 1.1. Here and in the following we write $z = (x, t) \in \mathbb{R}^{n+1}$ and $\varphi_t = \partial_t \varphi$ denotes the derivative with respect to the time-variable t, whence Du, respectively $D^k u$, denotes the derivatives with respect to the space-variable x and $\delta u = (u, Du, \ldots, D^{m-1}u)$ is the vector of lower order derivatives. We note that $D^k u = \{D^\alpha u_i\}_{i=1,\ldots,N}^{|\alpha|=k}$ is an element of the vector space $\odot^k(\mathbb{R}^n, \mathbb{R}^N)$ of k-linear functions with values in $\mathbb{R}^N$. Moreover, we consider coefficients $A \colon \Omega_T \times \sum_{k=0}^{m-1} \odot^k(\mathbb{R}^n, \mathbb{R}^N) \times \odot^m(\mathbb{R}^n, \mathbb{R}^N) \to$

$\mathrm{Hom}(\odot^m(\mathbb{R}^n,\mathbb{R}^N),\mathbb{R})$ and an inhomogeneity $B \equiv (B^0,\dots,B^{m-1})$ such that $B^k\colon \Omega_T \times \sum_{k=0}^{m-1}\odot^k(\mathbb{R}^n,\mathbb{R}^N) \times \odot^m(\mathbb{R}^n,\mathbb{R}^N) \to \mathrm{Hom}(\odot^k(\mathbb{R}^n,\mathbb{R}^N),\mathbb{R})$ for $k = 0,\dots,m-1$, fulfilling p-growth conditions, which are allowed to be degenerate. Precisely, we assume that A and B are Carathéodory functions and

$$\langle A(z,q),q\rangle \ge \nu\,|q|^p - b^p, \tag{2.2}$$

$$|A(z,q)| \le L\,|q|^{p-1} + b^{p-1}, \tag{2.3}$$

$$|B(z,q)| \le L\,|q|^{p-1} + b^{p-1}, \tag{2.4}$$

for all $z \in \Omega_T$, $q \in \odot^m(\mathbb{R}^n,\mathbb{R}^N)$ and some constants $0 < \nu \le 1 \le L < \infty$ and $p > \max\{1, \frac{2n}{n+2m}\}$. The function $b\colon \Omega_T \to \mathbb{R}$ is supposed to be measurable with bounded norm $\|b\|_{L^{\hat p}(\Omega_T)} < \infty$ for some $\hat p \ge p$. Let us mention that the restriction $p > \max\{1, \frac{2n}{n+2m}\}$ is necessary in the parabolic framework, because of the embedding $W^{m,\frac{2n}{n+2m}} \hookrightarrow L^2$ (we always have to deal with the L^2-norm of u, coming from the time derivative in (2.1)).

As basic sets for our estimates we shall use parabolic cylinders where the radii in space and time are not coupled. This is due to the fact that in the degenerate case, i.e., when $p \ne 2$, no uniform system of cylinders is available and therefore one has to consider parabolic cylinders whose size depends on the solution itself. We now introduce the basic notation for parabolic cylinders and several useful sets which shall be frequently used in the following. By $Q_{z_0}(\varrho,s) \equiv B_{x_0}(\varrho)\times\mathcal{T}_{t_0}(s) \subset \mathbb{R}^{n+1}$ we denote the parabolic cylinder around $z_0 = (x_0,t_0) \in \mathbb{R}^{n+1}$, where $B_{x_0}(\varrho)$ is the open ball in $\mathbb{R}^n$ with center x_0 and radius ϱ and $\mathcal{T}_{t_0}(s) \equiv (t_0-s,t_0+s)$ the interval of length $2s$ around t_0. In the particular case $s = \varrho^{2m}$ we write $Q_{z_0}(\varrho) = Q_{z_0}(\varrho,\varrho^{2m})$. Moreover, we define $\mathcal{H}_{t_0}(s) \equiv \mathbb{R}^n\times\mathcal{T}_{t_0}(s)$ and $\mathcal{C}_{x_0}(\varrho) \equiv B_{x_0}(\varrho)\times\mathbb{R}$. If $z_0 = 0$, we omit the center; i.e., we abbreviate $Q(\varrho,s) = Q_0(\varrho,s)$ and $B(\varrho) = B_0(\varrho)$. Furthermore, we write $\alpha B(\varrho) = B(\alpha\varrho)$, $\alpha\mathcal{T}(s) = \mathcal{T}(\alpha^{2m}s)$, $\alpha Q(\varrho,s) = Q(\alpha\varrho,\alpha^{2m}s)$, ... for a ball, interval, cylinder, ... enlarged by the factor $\alpha > 0$ with respect to the parabolic metric. Finally, if $v\colon Q_{z_0}(\varrho,s) \to \mathbb{R}^k$, $k \in \mathbb{N}$, is integrable we write

$$(v)_{z_0;\varrho,s} \equiv (v)_{Q_{z_0}(\varrho,s)} \equiv \fint_{Q_{z_0}(\varrho,s)} v\,dz$$

for its mean value on $Q_{z_0}(\varrho,s)$, respectively for $w\colon B_{x_0}(\varrho) \to \mathbb{R}^k$ we write $(w)_{x_0;\varrho} \equiv (w)_{B_{x_0}(\varrho)} \equiv \fint_{B_{x_0}(\varrho)} w\,dz$. We now are in a position to state the main result of the paper:

Theorem 2.1. *Let $p > \max\{1, \frac{2n}{n+2m}\}$. Then there exists $\beta_0 = \beta_0(n, N, m, p, L/\nu) > 0$ such that the following holds: Whenever $u \in L^2(\Omega_T; \mathbb{R}^N) \cap L^{p-\beta}(-T, 0; W^{m,p-\beta}(\Omega; \mathbb{R}^N))$, with some $\beta \in (0, \beta_0]$, is a very weak solution of the parabolic system* (2.1) *under the assumptions* (2.2)-(2.4), *then we have $u \in L^p_{\mathrm{loc}}(-T, 0; W^{m,p}_{\mathrm{loc}}(\Omega; \mathbb{R}^N))$, and for any parabolic cylinder $Q_{z_0}(2\varrho) \Subset \Omega_T$ there holds*

$$\fint_{Q_{z_0}(\varrho)} |D^m u|^p \, dz \\ \leq c \Big(\fint_{Q_{z_0}(2\varrho)} (|D^m u| + b)^{p-\beta} \, dz \Big)^{1+\frac{\beta}{d}} + c \fint_{Q_{z_0}(2\varrho)} \big(1 + b^p\big) \, dz,$$

where $c = c(n, N, m, p, L/\nu)$ and

$$d = \begin{cases} 2 - \beta & \text{if } p \geq 2 \\ p - \beta - \frac{(2-p)n}{2m} & \text{if } p < 2. \end{cases}$$

Remark 2.2. Combining this result with the higher integrability result for weak solutions from [5], Theorem 1, we even achieve an integrability exponent of $|D^m u|$ larger then p. To be precise, if $b \in L^{\hat{p}}_{\mathrm{loc}}(\Omega_T)$ for some $\hat{p} > p$, then under the assumptions of Theorem 2.1 there exists $\varepsilon \in (0, \hat{p}]$ such that $u \in L^{p+\varepsilon}_{\mathrm{loc}}(-T, 0; W^{m,p+\varepsilon}_{\mathrm{loc}}(\Omega; \mathbb{R}^N))$.

Note that we will not explicitly calculate the exponent β_0 in Theorem 2.1 coming from an application of Gehring's theorem. But it is worth mentioning that $\beta_0 \searrow 0$ when $L/\nu \to \infty$, since the involved constants blow up when the fraction L/ν becomes large (see for instance [31, 6] for the dependence of the constants in Gehring's theorem). Indeed, [29] provides a counterexample of a linear elliptic equation featuring exactly this qualitative behavior.

3. Preliminary material

3.1. **Interpolation and iteration lemmata.** We now state an interpolation lemma for intermediate derivatives which can be found for instance in [3], Theorem 4.14. For the precise dependence of the constant we also refer to [4], Lemma B.1. Later, we will apply this lemma several times on the horizontal time slices.

Lemma 3.1. *Let $B_{x_0}(\varrho) \subset \mathbb{R}^n$ be a ball with radius $0 < \varrho \leq 1$ and let $u \in W^{m,p}(B_{x_0}(\varrho))$ with $p \geq 1$. Then, there exists a constant $c = c(n, m, p)$*

such that for any $0 \le k \le m-1$ and $0 < \varepsilon \le 1$ there holds

$$\int_{B_{x_0}(\varrho)} \Big|\frac{D^k u}{\varrho^{m-k}}\Big|^p dx \le \varepsilon \int_{B_{x_0}(\varrho)} |D^m u|^p \, dx + c\, \varepsilon^{-\frac{k}{m-k}} \int_{B_{x_0}(\varrho)} \Big|\frac{u}{\varrho^m}\Big|^p dx.$$

We now state Gagliardo-Nirenberg's inequality (see [28]) in a form which will be convenient for our purpose:

Lemma 3.2. *Let $B_{x_0}(\varrho) \subset \mathbb{R}^n$ with $\varrho \le 1$ and $u \in W^{m,\vartheta}(B_{x_0}(\varrho))$, $m \in \mathbb{N}$ and $1 \le \sigma, \vartheta, r \le \infty$ and $\theta \in (0,1)$ and $0 \le k \le m-1$ with $k - \frac{n}{\sigma} \le \theta(m - \frac{n}{\vartheta}) - (1-\theta)\frac{n}{r}$. Then, there exists a constant $c = c(n,m,\sigma)$ such that there holds*

$$⨍_{B_{x_0}(\varrho)} \Big|\frac{D^k u}{\varrho^{m-k}}\Big|^\sigma dx \le c\Big(\sum_{j=0}^{m} ⨍_{B_{x_0}(\varrho)} \Big|\frac{D^j u}{\varrho^{m-j}}\Big|^\vartheta dx\Big)^{\frac{\theta\sigma}{\vartheta}} \Big(⨍_{B_{x_0}(\varrho)} \Big|\frac{u}{\varrho^m}\Big|^r dx\Big)^{\frac{(1-\theta)\sigma}{r}}.$$

The next lemma is a standard iteration lemma and can be found for instance in [17], Lemma 5.1.

Lemma 3.3. *Let $0 < \theta < 1$, $B \ge 0$, $A \ge 0$ $\alpha > 0$ and $0 < r < \varrho < \infty$ and let $f \ge 0$ be a bounded function satisfying*

$$f(t) \le \theta f(s) + A(s-t)^{-\alpha} + B \qquad \forall\ r \le t < s \le \varrho.$$

Then there exists a constant $c = c(\alpha, \vartheta)$, such that

$$f(r) \le c\ \big(A(\varrho - r)^{-\alpha} + B\big).$$

3.2. **Mean value polynomials.** In order to treat regularity problems for elliptic respectively parabolic systems one usually needs to control oscillation quantities of the solutions to measure in a weak sense its regularity. In any case polynomials, especially the mean value polynomials, will play an important role. Here, we will establish the basic estimates used throughout the paper. The following lemma provides an estimate of a polynomial in terms of its mean values (see [4], Lemma A.1).

Lemma 3.4. *Let $P\colon \mathbb{R}^n \to \mathbb{R}^N$ be a polynomial of degree $\le m-1$ and $\mathcal{O}_1, \mathcal{O}_2$ bounded open sets in $\mathbb{R}^n$ with $\mathcal{O}_1 \subset \mathcal{O}_2$. Then for all $0 \le k \le m-1$ there holds (with $(D^\ell P)_{\mathcal{O}_1} = ⨍_{\mathcal{O}_1} D^\ell P\, dx$) :*

$$|D^k P(x)| \le c(n,m) \sum_{\ell=k}^{m-1} \operatorname{diam}(\mathcal{O}_2)^{\ell-k}\, |(D^\ell P)_{\mathcal{O}_1}| \qquad \forall\, x \in \mathcal{O}_2.$$

The next lemma provides an estimate for the derivative of a polynomial in terms of its L^1-norm on a ball and is a slight generalization of a result of S. Campanato (see [7], Lemma 2.I)

Lemma 3.5. *Let $P\colon \mathbb{R}^n \to \mathbb{R}^N$ be a polynomial of degree $\le m-1$ and $B_{x_0}(\varrho_1), B_{x_0}(\varrho_2)$ balls in $\mathbb{R}^n$ with radii $0<\varrho_1\le\varrho_2\le 1$. Then for $0\le k\le m-1$ there holds*

$$|D^kP(x)| \le c(n,m)\left(\tfrac{\varrho_2}{\varrho_1}\right)^{m-1}\varrho_2^{-k} \fint_{B_{x_0}(\varrho_1)} |P(y)|\,dy \qquad \forall\, x\in B_{x_0}(\varrho_2).$$

Proof. Applying in turn Lemma 3.4 with $(B_{x_0}(\varrho_1), B_{x_0}(\varrho_2))$ instead of $(\mathcal{O}_1, \mathcal{O}_2)$ and the interpolation Lemma 3.1 (note that $D^mP\equiv 0$) we infer for $x\in B_{x_0}(\varrho_2)$

$$\begin{aligned}|D^kP(x)| &\le c\sum_{\ell=k}^{m-1}\varrho_2^{\ell-k}\,|(D^\ell P)_{x_0;\varrho_1}|\\ &\le c\sum_{\ell=k}^{m-1}\varrho_2^{\ell-k}\fint_{B_{x_0}(\varrho_1)}|D^\ell P|\,dy \le c\left(\tfrac{\varrho_2}{\varrho_1}\right)^{m-1}\varrho_2^{-k}\fint_{B_{x_0}(\varrho_1)}|P|\,dy,\end{aligned}$$

where $c=c(n,m)$. This proves the desired estimate. □

An immediate consequence of Lemma 3.4 is the following (see [4], Lemma A.6).

Lemma 3.6. *Let $\mathcal{O}_1, \mathcal{O}_2$ be bounded open sets in $\mathbb{R}^n$ with $\mathcal{O}_1\subset\mathcal{O}_2$ and $\mathcal{T}\subset\mathbb{R}$ an interval. Furthermore, suppose that $u\in L^1(\mathcal{T}; W^{m,1}(\mathcal{O}_1;\mathbb{R}^N))$ and $P_{\mathcal{O}_1\times\mathcal{T}}\colon\mathbb{R}^n\to\mathbb{R}^N$ is the mean value polynomial of degree $\le m-1$ defined by $(\delta P_{\mathcal{O}_1\times\mathcal{T}})_{\mathcal{O}_1} = (\delta u)_{\mathcal{O}_1\times\mathcal{T}}$. Then for all polynomials $P\colon\mathbb{R}^n\to\mathbb{R}^N$ of degree $\le m-1$ and for all $0\le k\le m-1$ and $x\in\mathcal{O}_2$ there holds*

$$|D^k(P_{\mathcal{O}_1\times\mathcal{T}}-P)(x)| \le c(n,m)\sum_{\ell=k}^{m-1}\operatorname{diam}(\mathcal{O}_2)^{\ell-k}\fint_{\mathcal{O}_1\times\mathcal{T}}|D^\ell(u-P)|\,dz.$$

3.3. **Definition and properties of the mollification.** In Chapter 5.3 we will construct a test-function by changing u (multiplied with a cut-off function) on the "bad set" where the strong maximal function of D^mu is larger then a certain value λ. To be precise we will use a kind of mollification $[u]_h$ of u instead of u itself for this construction. The definition and basic properties of this mollification are established in this chapter.

Let $\phi\in C_0^\infty(B_1)$ be a rotationally symmetric mollifying kernel, satisfying $\int_{\mathbb{R}^n}\phi\,dx=1$, $0\le\phi\le c(n)$ and $\phi(x)=\phi(|x|)$ for all $x\in\mathbb{R}^n$. For $h>0$ we define the associated scaled kernel ϕ_h by $\phi_h(x)\equiv h^{-n}\,\phi(\frac{x}{h})$, for $x\in\mathbb{R}^n$. Now, let $\mathcal{O}\Subset\Omega$, $0<h<\operatorname{dist}(\mathcal{O},\partial\Omega)$ and $\gamma>0$ and suppose that $f\colon\Omega_T\to\mathbb{R}^k$,

$k \geq 1$ is a locally integrable function. Then, by $[f]_h$ we mean the following function

$$[f]_h(x,t) \equiv \fint_{Q_z(h,\gamma h^{2m})} \phi(\tfrac{y-x}{h}) f(y,\tau)\, d(y,\tau) = \fint_{\mathcal{T}_t(\gamma h^{2m})} (\phi_h * f)(x,\tau)\, d\tau \tag{3.1}$$

for $x \in \mathcal{O}$ and $t \in (-T+\gamma h^{2m}, -\gamma h^{2m})$. Here, we recall that $\mathcal{T}_t(\gamma h^{2m}) = (t-\gamma h^{2m}, t+\gamma h^{2m})$. Moreover, we note that for the sake of clearness we do not indicate the parameter γ in our notation, although $[f]_h$ depends on γ. We summarize some basic properties of $[f]_h$ in the next lemma. For a proof we refer to [4], Chapter 8.2.

Lemma 3.7. *Let ϕ be a mollifying kernel on $\mathbb{R}^n$, $\mathcal{O} \Subset \Omega$, $0 < h < \operatorname{dist}(\mathcal{O}, \partial\Omega)$ and $\gamma > 0$ and set $\mathcal{T} \equiv (-T+\gamma h^{2m}, -\gamma h^{2m})$. Moreover, suppose that $f\colon \Omega_T \to [0,\infty)$ is an integrable function. Then*

(i) $[f]_h \to f$ a.e. in $\mathcal{O} \times \mathcal{T}$ as $h \searrow 0$, $[f]_h$ is continuous and bounded.

(ii) For any cylinder $Q_{\mathfrak{z}}(r,s) \subset \mathcal{O} \times \mathcal{T}$ we have

$$\fint_{Q_{\mathfrak{z}}(r,s)} [f]_h\, dz \leq c(n) \fint_{Q_{\mathfrak{z}}(r+h, s+\gamma h^{2m})} f\, dz.$$

(iii) The function $[f]_h$ is differentiable with respect to t on $\mathcal{T}$ and moreover $[f]_h(\cdot,t) \in C^\infty(\mathcal{O})$ for all $t \in \mathcal{T}$.

We can reformulate our parabolic system (2.1) for the mollified function $[u]_h$: For $\mathcal{O} \Subset \Omega$, $0 < h < \operatorname{dist}(\mathcal{O}, \partial\Omega)$ and fixed $\gamma > 0$, the function $[u]_h$ satisfies

$$\int_\Omega \partial_t [u]_h(\cdot,t) \cdot \varphi + \big\langle [A(\cdot, D^m u)]_h(\cdot,t), D^m\varphi \big\rangle\, dx = -\int_\Omega \big\langle [B(\cdot, D^m u)]_h(\cdot,t), \delta\varphi \big\rangle\, dx \tag{3.2}$$

for all $\varphi \in W_0^{m,\frac{p-\beta}{1-\beta}}(\mathcal{O};\mathbb{R}^N) \cap L^2(\mathcal{O};\mathbb{R}^N)$ and almost every $t \in (-T+\gamma h^{2m}, -\gamma h^{2m})$. Moreover, (2.3), (2.4) and the fact that $L \geq 1$ imply

$$\begin{aligned} |[A(\cdot, D^m u)]_h(z)| &\leq L\, [|D^m u|^{p-1} + b^{p-1}]_h(z), \\ |[B(\cdot, D^m u)]_h(z)| &\leq L\, [|D^m u|^{p-1} + b^{p-1}]_h(z), \end{aligned} \tag{3.3}$$

for all $z \in \mathcal{O} \times (-T+\gamma h^{2m}, -\gamma h^{2m})$. For more details we once again refer to [4], Chapter 8.2.

4. Poincaré type estimates

Since we did not impose any a priori differentiability assumption with respect to time on our solutions, we cannot apply the usual Poincaré inequality. In order to have nevertheless a certain Poincaré type inequality at hand, we will exploit our parabolic system. To be more precise, the parabolic system will help us to control differences in time of the weighted means of the solution. Therefore, we first introduce the notion of a weight-function and weighted mean that will be suitable for our purpose.

Suppose that $\mathcal{O}$ is an open set in $\mathbb{R}^n$ and $\eta \in C_0^\infty(\mathcal{O})$ is a nonnegative function on $\mathcal{O}$. Moreover, let $f \in L^1(\mathcal{O} \times (t_1, t_2); \mathbb{R}^k)$ with $t_1 < t_2$, $k \in \mathbb{N}$. Then, we define the **weighted mean** of $f(\cdot, t)$ on $\mathcal{O}$ for almost every $t \in (t_1, t_2)$ by

$$(f)_\eta(t) \equiv \int_{\mathcal{O}} f(\cdot, t)\, \eta\, dx. \tag{4.1}$$

The next lemma is a Poincaré type inequality involving differences in time of weighted means rather than time derivatives of the considered function. It will be applied later in Corollary 4.3 for the solution u of our parabolic system and in Lemma 5.11 to the function v_h defined in (5.13). Since we will apply it on various parabolic cylinders or intersections of parabolic cylinders, we formulate it for very general types of sets.

Lemma 4.1. *Let $u \in L^\vartheta(-T, 0; W^{m,\vartheta}(\Omega; \mathbb{R}^N))$ with $\vartheta \geq 1$. Moreover suppose that $\mathcal{O} \Subset \Omega$ is a convex open set such that $B_y(\varrho) \subset \mathcal{O} \subset B_y(\alpha\varrho)$ for some $y \in \mathbb{R}^n$, $0 < \varrho \leq 1$ and $\alpha > 1$ and $\mathcal{T}_1, \mathcal{T}_2 \subset (-T, 0)$ are two intervals. Finally let $\eta \in C_0^\infty(\mathcal{O})$ be a nonnegative function with $\int_{\mathcal{O}} \eta\, dx = 1$ and $0 \leq \eta \leq c_\eta/\varrho^n$. Then for all $0 \leq k \leq m-1$ there holds*

$$\begin{aligned} &\fint_{\mathcal{O}\times\mathcal{T}_1} |D^k(u - P_{\mathcal{O}\times\mathcal{T}_2})|^\vartheta\, dz \\ &\qquad \leq c\ \varrho^{\vartheta(m-k)} \Big(\fint_{\mathcal{O}\times\mathcal{T}_1} |D^m u|^\vartheta\, dz + \fint_{\mathcal{O}\times\mathcal{T}_2} |D^m u|^\vartheta\, dz \Big) \\ &\qquad\quad + c \sum_{k\leq|\ell|\leq m-1} \varrho^{\vartheta(|\ell|-k)} \sup_{t_1,t_2\in\mathcal{T}_1\cup\mathcal{T}_2} |(u)_{D^\ell\eta}(t_2) - (u)_{D^\ell\eta}(t_1)|^\vartheta, \end{aligned}$$

where $c = c(n, m, \vartheta, \alpha, c_\eta)$ and $P_{\mathcal{O}\times\mathcal{T}_2} \colon \mathbb{R}^n \to \mathbb{R}^N$ denotes the mean value polynomial of u of degree $\leq m-1$ on $\mathcal{O} \times \mathcal{T}_2$, defined by $(\delta P_{\mathcal{O}\times\mathcal{T}_2})_{\mathcal{O}} = (\delta u)_{\mathcal{O}\times\mathcal{T}_2}$, and the sum is taken over all multiindices ℓ of order $k \leq |\ell| \leq m-1$.

Proof. In the following we shall abbreviate $P = P_{\mathcal{O}\times\mathcal{T}_2}$. In order to apply Poincaré's inequality slicewise with respect to x, we use the weighted means of $D^j(u-P)$ with weight-function η and consider for $k \le j \le m-1$ and almost every $t \in (-T,0)$ the following decomposition

$$\begin{aligned}
&-\!\!\!\!\!\!\int_{\mathcal{O}} |D^j(u(\cdot,t)-P)|^\vartheta\,dx \le 3^\vartheta \Big[-\!\!\!\!\!\!\int_{\mathcal{O}} \big|D^j(u(\cdot,t)-P) - \big(D^j(u(\cdot,t)-P)\big)_\eta\big|^\vartheta\,dx \\
&+ \Big|(D^j u)_\eta(t) - -\!\!\!\!\!\!\int_{\mathcal{T}_2} (D^j u)_\eta(\tau)\,d\tau\Big|^\vartheta + \Big| -\!\!\!\!\!\!\int_{\mathcal{T}_2} (D^j u)_\eta(\tau)\,d\tau - (D^j P)_\eta\Big|^\vartheta\ \Big] \\
&= 3^\vartheta \big(I(t) + II(t) + III\big),
\end{aligned} \tag{4.2}$$

with the obvious meaning of $I(t)$, $II(t)$ and III. In turn we will estimate these terms. We start with the estimate for $I(t)$. Applying Poincaré's inequality for convex domains (see e.g. [4], Lemma B.5) with respect to x to $D^j(u-P)(\cdot,t)$ we find

$$I(t) \le c(n,\vartheta,\alpha,c_\eta)\ \varrho^\vartheta -\!\!\!\!\!\!\int_{\mathcal{O}} |D^{j+1}(u(\cdot,t)-P)|^\vartheta\,dx.$$

The estimate for III is similar. Here, we exploit the mean value property of P, i.e., the fact that $-\!\!\!\!\!\int_{\mathcal{T}_2} -\!\!\!\!\!\int_{\mathcal{O}} D^j(u-P)\,dx\,dt = 0$ and then use the estimate for $I(\tau)$ from above to infer

$$\begin{aligned}
III &= \Big| -\!\!\!\!\!\!\int_{\mathcal{T}_2} -\!\!\!\!\!\!\int_{\mathcal{O}} D^j(u-P) - \big(D^j(u-P)\big)_\eta\,dx\,d\tau\Big|^\vartheta \\
&\le -\!\!\!\!\!\!\int_{\mathcal{T}_2} I(\tau)\,d\tau \le c\ \varrho^\vartheta -\!\!\!\!\!\!\int_{\mathcal{T}_2} -\!\!\!\!\!\!\int_{\mathcal{O}} |D^{j+1}(u-P)|^\vartheta\,dx\,d\tau.
\end{aligned}$$

To get an estimate for $II(t)$ we first note that from integration by parts we have

$$(D^\ell u)_\eta(t) = \int_{\mathcal{O}} D^\ell u(\cdot,t)\eta\,dx = (-1)^j \int_{\mathcal{O}} u(\cdot,t) D^\ell \eta\,dx = (-1)^j (u)_{D^\ell\eta}(t)$$

for any multiindex ℓ of order $|\ell| = j$. Therefore, taking the supremum over $\tau \in \mathcal{T}_2$ yields

$$\begin{aligned}
II(t) &\le -\!\!\!\!\!\!\int_{\mathcal{T}_2} |(D^j u)_\eta(t) - (D^j u)_\eta(\tau)|^\vartheta\,d\tau \le \sup_{\tau\in\mathcal{T}_2} |(D^j u)_\eta(t) - (D^j u)_\eta(\tau)|^\vartheta \\
&\le c \sum_{|\ell|=j} \sup_{\tau\in\mathcal{T}_2} |(u)_{D^\ell\eta}(t) - (u)_{D^\ell\eta}(\tau)|^\vartheta,
\end{aligned}$$

where $c = c(m,\vartheta)$. Combining the previous estimates for $I(t)$, $II(t)$ and III with (4.2) and integrating with respect to t over $\mathcal{T}_i$ with $i = 1, 2$ we infer for

$k \le j \le m-1$ that

$$\begin{aligned}\fint_{\mathcal{O}\times\mathcal{T}_i} |D^j(u-P)|^\vartheta dz \le c\,\varrho^\vartheta \fint_{\mathcal{O}\times\mathcal{T}_1} |D^{j+1}(u-P)|^\vartheta\, dz \\ + c\,\varrho^\vartheta \fint_{\mathcal{O}\times\mathcal{T}_2} |D^{j+1}(u-P)|^\vartheta\, dz \\ + c \sum_{|\ell|=j} \sup_{t_1,t_2\in\mathcal{T}_1\cup\mathcal{T}_2} |(u)_{D^\ell\eta}(t_2) - (u)_{D^\ell\eta}(t_1)|^\vartheta, \end{aligned} \tag{4.3}$$

where $c = c(n, m, \vartheta, \alpha, c_\eta)$. We now start with (4.3) for $j = k$ and $i = 1$ and then again apply (4.3) iteratively for $j = k+1, \dots, m-1$ and $i = 1, 2$ to infer

$$\begin{aligned}&\fint_{\mathcal{O}\times\mathcal{T}_1} |D^k(u-P)|^\vartheta\, dz \\ &\le c\ \varrho^\vartheta \Big(\fint_{\mathcal{O}\times\mathcal{T}_1} |D^{k+1}(u-P)|^\vartheta\, dz + \fint_{\mathcal{O}\times\mathcal{T}_2} |D^{k+1}(u-P)|^\vartheta\, dz \Big) \\ &\quad + c \sum_{|\ell|=k} \sup_{t_1,t_2\in\mathcal{T}_1\cup\mathcal{T}_2} |(u)_{D^\ell\eta}(t_2) - (u)_{D^\ell\eta}(t_1)|^\vartheta \\ &\le c\ \varrho^{2\vartheta} \Big(\fint_{\mathcal{O}\times\mathcal{T}_1} |D^{k+2}(u-P)|^\vartheta\, dz + \fint_{\mathcal{O}\times\mathcal{T}_2} |D^{k+2}(u-P)|^\vartheta\, dz \Big) \\ &\quad + c \sum_{k\le|\ell|\le k+1} \varrho^{\vartheta(|\ell|-k)} \sup_{t_1,t_2\in\mathcal{T}_1\cup\mathcal{T}_2} |(u)_{D^\ell\eta}(t_2) - (u)_{D^\ell\eta}(t_1)|^\vartheta \\ &\ \vdots \\ &\le c\ \varrho^{\vartheta(m-k)} \Big(\fint_{\mathcal{O}\times\mathcal{T}_1} |D^m u|^\vartheta\, dz + \fint_{\mathcal{O}\times\mathcal{T}_2} |D^m u|^\vartheta\, dz \Big) \\ &\quad + c \sum_{k\le|\ell|\le m-1} \varrho^{\vartheta(|\ell|-k)} \sup_{t_1,t_2\in\mathcal{T}_1\cup\mathcal{T}_2} |(u)_{D^\ell\eta}(t_2) - (u)_{D^\ell\eta}(t_1)|^\vartheta, \end{aligned}$$

where $c = c(n, m, \vartheta, \alpha, c_\eta)$. This proves the desired inequality. □

As already mentioned before, we now exploit the parabolic system in order to control differences in time of the weighted means of our solution.

Lemma 4.2. *Suppose that* $u \in L^{p-\beta}(-T, 0; W^{m,p-\beta}(\Omega;\mathbb{R}^N))$, $0 \le \beta \le \min\{1, p-1\}$, *fulfills system* (2.1) *with* (2.3) *and* (2.4). *Let* $\mathcal{O} \Subset \Omega$ *be an open convex set and* $\eta \in C_0^\infty(\mathcal{O})$ *a nonnegative function and* $[u]_h$ *defined according to* (3.1) *with* $0 < h < \operatorname{dist}(\mathcal{O}, \partial\Omega)$, $\gamma > 0$. *Then, there exists*

$c = c(N)$ such that for the weighted means of $[u]_h$ defined in (4.1) we have for almost every $t_1, t_2 \in (-T + \gamma h^{2m}, -\gamma h^{2m})$ the following estimate

$$|([u]_h)_\eta(t_2) - ([u]_h)_\eta(t_1)| \le c\, L \sum_{\ell=0}^{m} \|D^\ell \eta\|_{L^\infty} \int_{\mathcal{O}\times(t_1,t_2)} \big[(|D^m u| + b)^{p-1}\big]_h \, dz.$$

Proof. Let $i \in \{1, \dots N\}$. Let us recall that $[u]_h(\cdot, t)$ satisfies (3.2) for $t \in (-T + \gamma h^{2m}, -\gamma h^{2m})$. Therefore we can test (3.2) for $t \in (t_1, t_2)$ with the test-function $\varphi \colon \mathbb{R}^n \to \mathbb{R}^N$, $\varphi_i \equiv \eta$, $\varphi_j \equiv 0$ for $j \neq i$ and integrate the resulting equation with respect to t over (t_1, t_2). This gives:

$$\begin{aligned}([u_i]_h)_\eta(t_2) - ([u_i]_h)_\eta(t_1) &= \int_{t_1}^{t_2} \partial_t([u_i]_h)_\eta \, dt \\ &= -\int_{t_1}^{t_2} \int_{\mathcal{O}} \big\langle [A_i(\cdot, D^m u)]_h, D^m \eta \big\rangle + \big\langle [B_i(\cdot, D^m u)]_h, \delta\eta \big\rangle \, dx\, dt.\end{aligned}$$

Using the growth of $[A]_h$ and $[B]_h$ from (3.3) and noting that $L \ge 1$ we infer

$$\begin{aligned}|([u_i]_h)_\eta(t_2) - ([u_i]_h)_\eta(t_1)| &\le L \int_{\mathcal{O}\times(t_1,t_2)} \big[|D^m u|^{p-1} + b^{p-1}\big]_h \big(|D^m \eta| + |\delta\eta|\big) \, dz \\ &\le 2L \sum_{\ell=0}^{m} \|D^\ell \eta\|_{L^\infty} \int_{\mathcal{O}\times(t_1,t_2)} \big[(|D^m u| + b)^{p-1}\big]_h \, dz.\end{aligned}$$

Summing over $i = 1, \dots N$ we obtain the asserted estimate. □

Combining Lemma 4.1 and Lemma 4.2 we now conclude the following Poincaré type inequality for very weak solutions.

Corollary 4.3. *Suppose that $u \in L^{p-\beta}(-T, 0; W^{m,p-\beta}(\Omega; \mathbb{R}^N))$, $0 \le \beta \le \min\{1, p-1\}$, satisfies system (2.1) with (2.3) and (2.4) and let $B = B_{x_0}(\varrho) \Subset \Omega$ with $0 < \varrho \le 1$ and $\mathcal{T}_1, \mathcal{T}_2 \subset (-T, 0)$ be two intervals with $\mathcal{T}_1 \cap \mathcal{T}_2 \neq \emptyset$. Then for all $1 \le \vartheta \le p - \beta$ and all $0 \le k \le m-1$ there holds*

$$\begin{aligned}&\fint_{B\times\mathcal{T}_1} |D^k(u - P_{B\times\mathcal{T}_2})|^\vartheta \, dz \\ &\quad\le c\, \varrho^{\vartheta(m-k)} \Big[\fint_{B\times\mathcal{T}_1} |D^m u|^\vartheta \, dz + \fint_{B\times\mathcal{T}_2} |D^m u|^\vartheta \, dz \\ &\qquad\qquad + \Big(\frac{|\mathcal{T}_1 \cup \mathcal{T}_2|}{\varrho^{2m}} \fint_{B\times(\mathcal{T}_1\cup\mathcal{T}_2)} (|D^m u| + b)^{p-1} \, dz \Big)^\vartheta \Big],\end{aligned}$$

where $c = c(n, N, m, L, \vartheta)$ and $P_{B\times\mathcal{T}_2} \colon \mathbb{R}^n \to \mathbb{R}^N$ denotes the mean value polynomial of u of degree $\le m-1$ on $B \times \mathcal{T}_2$, defined by $(\delta P_{B\times\mathcal{T}_2})_B = (\delta u)_{B\times\mathcal{T}_2}$.

Proof. Let $\eta \in C_0^\infty(B)$ be a nonnegative function with $\int_B \eta\, dx = 1$ and $|D^\ell \eta| \le c/\varrho^{n+\ell}$ for $0 \le \ell \le m+k$. Applying Lemma 4.1 with this particular weight-function, and using Lemma 4.2 (after passing to the limit $h \searrow 0$) to estimate the differences in times of weighted means $(u)_{D^\ell \eta}$ appearing on the right-hand side, yields the desired estimate. □

In the previous Poincaré type inequality we have the "wrong exponent" of $|D^m u|$ on the right-hand side, namely $(⨍|D^m u|^{p-1}\, dz)^\vartheta$. Roughly speaking, in the following corollary, we introduce a special scaling of the parabolic cylinders depending on the solution itself in order to compensate this wrong exponent.

Corollary 4.4. *Suppose that* $u \in L^{p-\beta}(-T,0;W^{m,p-\beta}(\Omega;\mathbb{R}^N))$, $0 \le \beta \le \min\{1, p-1\}$, *satisfies the parabolic system* (2.1) *with* (2.3) *and* (2.4) *and let* $Q_{z_0}(\varrho,s) \Subset \Omega_T$ *with* $0 < \varrho \le 1$ *and* $0 < s \le \lambda^{2-p}\varrho^{2m}$ *and* $\lambda > 0$. *Suppose that there is a constant* $\kappa \ge 1$, *such that*

$$⨍_{Q_{z_0}(\varrho,s)} (|D^m u| + b)^\xi \, dz \le \kappa\, \lambda^\xi, \quad \textit{where } \max\{1, p-1\} \le \xi \le p - \beta. \tag{4.4}$$

Then for $1 \le \vartheta \le \xi$ *and* $0 \le k \le m$ *there holds*

$$⨍_{Q_{z_0}(\varrho,s)} |D^k(u - P_{z_0;\varrho,s})|^\vartheta \, dz \le c\ \varrho^{\vartheta(m-k)}\, \lambda^\vartheta,$$

where $c = c(n, N, m, L, \vartheta, \kappa)$ *and* $P_{z_0,\varrho,s}\colon \mathbb{R}^n \to \mathbb{R}^N$ *denotes the mean value polynomial of degree* $\le m-1$ *defined by* $(\delta P_{z_0,\varrho,s})_{x_0;\varrho} = (\delta u)_{z_0;\varrho,s}$.

Proof. Without loss of generality we assume $z_0 = 0$ and we recall that $Q(\varrho,s) = B(\varrho) \times \mathcal{T}(s)$. Applying the Poincaré type inequality from Corollary 4.3 with $\mathcal{T}_1 = \mathcal{T}_2 = \mathcal{T}(s)$ and noting that $|\mathcal{T}(s)|/\varrho^{2m} = 2s/\varrho^{2m} \le 2\lambda^{2-p}$, we obtain

$$\begin{aligned}
&⨍_{Q(\varrho,s)} |D^k(u - P_{\varrho,s})|^\vartheta \, dz \\
&\quad\le c\, \varrho^{\vartheta(m-k)} \Big[⨍_{Q(\varrho,s)} |D^m u|^\vartheta \, dz + \Big(\lambda^{2-p} ⨍_{Q(\varrho,s)} (|D^m u| + b)^{p-1} dz \Big)^\vartheta \Big] \\
&\quad\le c\, \varrho^{\vartheta(m-k)} \Big[\Big(⨍_{Q(\varrho,s)} |D^m u|^\xi dz \Big)^{\frac{\vartheta}{\xi}} + \Big(\lambda^{2-p} \Big(⨍_{Q(\varrho,s)} (|D^m u| + b)^\xi dz \Big)^{\frac{p-1}{\xi}} \Big)^\vartheta \Big] \\
&\quad\le c\, \varrho^{\vartheta(m-k)} \big(\lambda^\vartheta + (\lambda^{2-p}\, \lambda^{p-1})^\vartheta \big) = c\ \varrho^{\vartheta(m-k)}\, \lambda^\vartheta.
\end{aligned}$$

Here we have also used Hölder's inequality and hypothesis (4.4). Note that c depends on n, N, m, L, ϑ and κ. □

5. Caccioppoli type inequality

This chapter is devoted to the proof of the Caccioppoli type inequality stated below. This indeed is the crucial step to attain our higher integrability result. Thereby the main problem will be that we cannot test the parabolic system by the solution itself. Therefore, in Section 5.1 we will construct an admissible test-function in a very delicate way. Then, in Sections 5.2-5.4 we will show that this test-function satisfies the desired properties; i.e., on the one hand it fulfills the required integrability properties and on the other hand it is in a certain sense similar to the original solution. Finally, in Section 5.5 we will prove the Caccioppoli type inequality.

In order to state our Caccioppoli type inequality we will introduce the strong maximal function. Let $f\colon \mathbb{R}^{n+1} \to [-\infty, \infty]$ be a locally integrable function. Then we define the **strong maximal function** of f by

$$Mf(z) \equiv \sup_{\mathcal{Q}\ni z} \fint_{\mathcal{Q}} |f|\, d\tilde{z}, \qquad \text{for } z \in \mathbb{R}^{n+1}, \tag{5.1}$$

where the supremum is taken over all axially parallel cylinders $\mathcal{Q}$ containing z, i.e., cylinders of the form $\mathcal{Q} = B \times \mathcal{T}$, where $B \subset \mathbb{R}^n$ is a ball and $\mathcal{T} \subset \mathbb{R}$ is an interval. This is a sort of non-centered maximal function, which allows any scaling of the cylinders. An application of the Hardy-Littlewood maximal theorem in x- and t- direction shows that the Hardy-Littlewood maximal theorem still holds for this type of maximal function; i.e., M is a bounded operator of L^ϑ to L^ϑ for $1 < \vartheta < \infty$ and there holds (see [4], Lemma D.5 or [30], I.5.3(c))

$$\|Mf\|_{L^\vartheta(\mathbb{R}^{n+1})} \le c(n, \vartheta)\ \|f\|_{L^\vartheta(\mathbb{R}^{n+1})}. \tag{5.2}$$

In the sequel we shall use the following abbreviation:

$$M_{\mathcal{Q}}(z)^\xi \equiv M\big((|D^m u| + b)^\xi \chi_{\mathcal{Q}}\big)(z), \quad \text{with } \ \xi = \max\left\{\tfrac{1}{2}(1+p), p - \tfrac{1}{2}\right\}, \tag{5.3}$$

for a cylinder $\mathcal{Q} \subset \mathbb{R}^{n+1}$. Note that $\max\{1, p-1\} < \xi < p$. Now, we can state our Caccioppoli type inequality:

Lemma 5.1. *Let $\kappa \ge 1$. Then there exist $\beta_1 \in (0, \frac{1}{4}\min\{p-1, 1\}]$ and $c_{Cac} \ge 1$, depending on $n, N, m, p, L/\nu$ and κ such that the following holds: Whenever $u \in L^{p-\beta}(-T, 0; W^{m,p-\beta}(\Omega; \mathbb{R}^N)) \cap L^2(\Omega_T; \mathbb{R}^N)$, for some $0 < \beta \le \beta_1$, is a very weak solution of the parabolic system* (2.1) *satisfying* (2.2)–(2.4) *and $Q \equiv B \times \mathcal{T} \equiv Q_{z_0}(\varrho, s)$ is a parabolic cylinder such that $16Q \equiv Q_{z_0}(16\varrho, 16^{2m}s) \Subset \Omega_T$ and $s = \lambda_1^{2-p}\varrho^{2m}$ with $\lambda_1 > 0$, which satisfies*

$$\kappa^{-1}\lambda_1^{p-\beta} \le \fint_{Q} (|D^m u| + b)^{p-\beta}\, dz \tag{5.4}$$

and

$$\fint_{16Q} (|D^m u| + b)^{p-\beta}\, dz \le \kappa\ \lambda_1^{p-\beta}, \tag{5.5}$$

then we have

$$\begin{aligned}\lambda_1^{p-\beta} + \lambda_1^{p-2} \sup_{t\in\mathcal{T}} \fint_B \Big|\frac{u(\cdot,t)-P_Q}{\varrho^m}\Big|^2 \frac{dx}{\max\{M_{16Q}(\cdot,t),\lambda_1\}^\beta} \\ \le c_{Cac} \fint_{16Q} \lambda_1^{p-2-\beta}\Big|\frac{u-P_{16Q}}{\varrho^m}\Big|^2 + \Big|\frac{u-P_{16Q}}{\varrho^m}\Big|^{p-\beta} + b^{p-\beta}\, dz.\end{aligned}$$

Here $P_Q, P_{16Q} \colon \mathbb{R}^n \to \mathbb{R}^N$ *denote the mean value polynomials of degree* $\le m-1$ *associated to* Q *respectively* $16Q$ *defined by* $(\delta P_Q)_B = (\delta u)_Q$*, respectively* $(\delta P_Q)_{16B} = (\delta u)_{16Q}$.

Before we can start with the proof of the lemma, we introduce some further notation which shall be valid throughout the whole chapter. Without loss of generality we assume that $z_0 = 0$ and we abbreviate $B = B_{x_0}(\varrho)$, $\mathcal{T} = \mathcal{T}_{t_0}(s)$, $Q \equiv Q_{z_0}(\varrho, s)$, $\mathcal{H} = \mathcal{H}_{t_0}(s)$, $\mathcal{C} = \mathcal{C}_{x_0}(\varrho)$ and $\alpha B = B_{x_0}(\alpha\varrho)$, respectively $\alpha\mathcal{T} \equiv \mathcal{T}_{t_0}(\alpha^{2m}s)$, $\alpha Q \equiv Q_{z_0}(\alpha\varrho, \alpha^{2m}s)$, $\alpha\mathcal{H} \equiv \mathcal{H}_{t_0}(\alpha^{2m}s)$ and $\alpha\mathcal{C} \equiv \mathcal{C}_{x_0}(\alpha\varrho)$ for $\alpha > 0$. Then we can write $\alpha Q = \alpha B \times \alpha\mathcal{T}$. Let us note that due to the definition of the exponent ξ in (5.3) and the fact that $\beta \le \beta_1 \le \frac{1}{4}\min\{p-1, 1\}$ we have $p - \beta > \xi > 1$. Indeed, there even holds $\frac{p-\beta}{\xi}\max\big\{\frac{3p+1}{2p+2}, \frac{4p-1}{4p-2}\big\} = c(p) > 1$. Therefore, recalling the definition of M_{16Q} in (5.3), the Hardy-Littlewood maximal theorem, i.e., the estimate (5.2) with $(p-\beta)/\xi$ instead of ϑ, and the hypothesis (5.5) (note that $|16Q| = 16^{n+2m}|Q|$) we find

$$\begin{aligned}\int_{16Q} M_{16Q}^{p-\beta}\, dz &= \int_{16Q} M\big((|D^m u|+b)^\xi \chi_{16Q}\big)^{\frac{p-\beta}{\xi}}\, dz \\ &\le c \int_{16Q} (|D^m u|+b)^{p-\beta}\, dz \le c\ \kappa\lambda_1^{p-\beta}\, |Q|,\end{aligned} \tag{5.6}$$

with $c = c(n, m, p)$. Here, we note that the constant in the Hardy-Littlewood maximal theorem depends continuously on the exponent $(p-\beta)/\xi \in (c(p), p)$ with $c(p) > 1$ as noted above and therefore we can bound it by a constant depending only on n and p. Now, for $\lambda > 0$ we define

$$E(\lambda) \equiv \big\{z \in \overline{16Q} : M_{16Q}(z) \le \lambda\big\}$$

the lower level set of M_{16Q}, which we call the "good set." For $t \in \mathbb{R}$ we denote by

$$E_t(\lambda) \equiv \{x \in 16B : (x,t) \in E(\lambda)\} \tag{5.7}$$

the associated time-slice in the "good set" contained in $16B$. For a sufficiently large constant $c_E = c_E(n, m, p, \kappa) \geq 1$ we have

$$E(\lambda) \neq \emptyset \qquad \text{for all } \lambda \geq c_E \lambda_1.$$

Indeed, otherwise there would hold $M(z) > \lambda$ for all $z \in 16Q$ and we conclude from (5.6) that

$$\lambda^{p-\beta}\,|16Q| < \int_{16Q} M_{16Q}(z)^{p-\beta}\,dz \leq c(n,m,p)\;\kappa\,\lambda_1^{p-\beta}\,|16Q|,$$

which contradicts the choice $\lambda \geq c_E\lambda_1$ with a constant $c_E = c_E(n,m,p,\kappa) = (c(n,m,p)\kappa)^{1/(p-\beta)}$. Hence, in the following we shall always consider $\lambda \geq c_E\lambda_1$ and we put $\gamma = \lambda^{2-p}$.

5.1. **Construction of the test-function.** To construct our test-function we will modify the function $[u]_h$ on the "bad set," where the maximal function of $|D^m u|^\xi$ is large. To define the test-function on this set we will use mean value polynomials associated to certain Whitney cylinders having the intrinsic scaling $Q(r, \gamma r^{2m})$, for some $r > 0$. Therefore we first shall explain the Whitney-decomposition leading to these Whitney cylinders together with their basic geometric properties. We define the intrinsic parabolic metric d_λ by

$$d_\lambda(z_1, z_2) \equiv \max\left\{|x_2 - x_1|,\ \sqrt[2m]{\lambda^{p-2}|t_2 - t_1|}\,\right\},$$

for $z_1 = (x_1, t_1)$, $z_2 = (x_2, t_2) \in \mathbb{R}^{n+1}$ and we set

$$d_\lambda(\mathcal{Q}_1, \mathcal{Q}_2) \equiv \inf\{d_\lambda(z_1, z_2) : z_1 \in \mathcal{Q}_1, z_2 \in \mathcal{Q}_2\},$$

for $\mathcal{Q}_1, \mathcal{Q}_2 \subset \mathbb{R}^{n+1}$. Then the balls associated with the metric d_λ are intrinsic parabolic cylinders of the form $Q_{\mathfrak{z}}(r, \gamma r^{2m}) = \{z \in \mathbb{R}^{n+1} : d_\lambda(z, \mathfrak{z}) < r\}$ for $\mathfrak{z} \in \mathbb{R}^{n+1}$ and $r > 0$ (recall $\gamma = \lambda^{2-p}$). The next lemma is a parabolic version of the Whitney decomposition theorem (see [8] or [4], Theorem C.2).

Lemma 5.2. *There exist Whitney-type parabolic cylinders* $(Q_i)_{i=1}^\infty$, *with* $Q_i = Q_{z_i}(r_i, \gamma r_i^{2m})$ *(recall that* $\gamma = \lambda^{2-p}$ *and* $\lambda \geq c_E\lambda_1$ *is a free parameter and* $\alpha Q_i \equiv Q_{z_i}(\alpha r_i, \gamma(\alpha r_i)^{2m})$ *for* $\alpha > 0$*) having the following properties:*

- *(i)* $\mathbb{R}^{n+1}\setminus E(\lambda) = \bigcup_{i=1}^\infty Q_i$.
- *(ii) In each point of* $\mathbb{R}^{n+1}\setminus E(\lambda)$ *intersect at most* $c(n,m)$ *of the bigger cylinders* $16Q_i$.
- *(iii) For all* $i \in \mathbb{N}$ *there holds* $d_\lambda(z_i, E(\lambda)) = 4r_i$ *and hence*

$$2Q_i \subset \mathbb{R}^{n+1} \setminus E(\lambda) \quad \text{and} \qquad 4Q_i \cap E(\lambda) \neq \emptyset.$$

We can rewrite the Whitney-cylinders in the form $Q_i = B_i \times \mathcal{T}_i$, where $B_i = B_{x_i}(r_i)$ is the horizontal slice of Q_i and $\mathcal{T}_i = \mathcal{T}_{t_i}(\gamma r_i^{2m}) = (t_i - \gamma r_i^{2m}, t_i + \gamma r_i^{2m})$ the interval in the t-direction. To construct the test-function, we will multiply the mean value polynomials of $[u - P_Q]_h$ on the Whitney-cylinders with a partition of unity on $\mathbb{R}^{n+1} \setminus E(\lambda)$ subordinate to the cylinders $2Q_i$. This partition of unity will be introduced in the following lemma. A detailed construction is given in [4], Chapter 8.3.

Lemma 5.3. *There exists a partition of unity $(\omega_i)_{i=1}^{\infty}$ on $\mathbb{R}^{n+1} \setminus E(\lambda)$, i.e., $\sum_{i=1}^{\infty} \omega_i \equiv 1$ on $\mathbb{R}^{n+1} \setminus E(\lambda)$, having the following properties, where c denotes a constant depending only on n and m:*

$$\begin{cases} \omega_i \in C_0^{\infty}(2Q_i), 0 \le \omega_i \le 1 \text{ and } \omega_i \ge c \text{ on } Q_i \text{ for all } i \in \mathbb{N}, \\ |\partial_t \omega_i| \le c/(\gamma r_i^{2m}), |D^k \omega_i| \le c/r_i^k \text{ for } 0 \le k \le m \text{ and for all } i \in \mathbb{N}. \end{cases}$$

For $i \in \mathbb{N}$ we define

$$I(i) \equiv \{j \in \mathbb{N} : \operatorname{spt}\omega_j \cap \operatorname{spt}\omega_i \neq \emptyset\}.$$

From Lemma 5.2, (ii) we see that $\operatorname{card} I(i) \le c(n, m)$. Moreover if Q_i is one of the Whitney-cylinders, then for each $j \in I(i)$ the radius r_j of Q_j is comparable to the radius r_i of Q_i; i.e.,

$$r_j/3 \le r_i \le 3r_j \quad \text{and} \quad Q_j \subset 16Q_i \qquad \forall\, j \in I(i). \tag{5.8}$$

To see this we suppose that $j \in I(i)$, which means that $\operatorname{spt}\omega_j \cap \operatorname{spt}\omega_i \neq \emptyset$. Since $\operatorname{spt}\omega_j \subset 2Q_j$ and $\operatorname{spt}\omega_i \subset 2Q_i$ this certainly implies $2Q_j \cap 2Q_i \neq \emptyset$. Hence, there exists $\mathfrak{z} \in 2Q_j \cap 2Q_i$ and we conclude from Lemma 5.2, (iii) (applied for i and j)

$$4r_i = d_\lambda(z_i, E(\lambda)) \le d_\lambda(z_i, \mathfrak{z}) + d_\lambda(\mathfrak{z}, z_j) + d_\lambda(z_j, E(\lambda)) \le 2r_i + 2r_j + 4r_j.$$

Absorbing $2r_i$ on the left-hand side we see that $r_i \le 3r_j$. On the other hand, since also $i \in I(j)$ we conclude with the same argument with i and j exchanged that $r_j \le 3r_i$. Together this proves the first assertion in (5.8). For the second one, i.e., $Q_j \subset 16Q_i$, we consider a point $z \in Q_j$. Then with the help of $\mathfrak{z} \in 2Q_j \cap 2Q_i$ as introduced above we find

$$d_\lambda(z, z_i) \le d_\lambda(z, \mathfrak{z}) + d_\lambda(\mathfrak{z}, z_i) \le 3r_j + 2r_i \le 9r_i + 2r_i < 16r_i.$$

Hence, $z \in Q_j$ implies $z \in 16Q_i$. Since $j \in I(i)$ was arbitrary we conclude the second assertion in (5.8). Furthermore, we denote by Θ those $\ell \in \mathbb{N}$ where $\operatorname{spt}\omega_\ell \cap 2\mathcal{H} \neq \emptyset$ (note $2\mathcal{H} = \mathcal{H}(2^{2m}s)$):

$$\Theta \equiv \{\ell \in \mathbb{N} : \operatorname{spt}\omega_\ell \cap 2\mathcal{H} \neq \emptyset\} \tag{5.9}$$

and we decompose Θ as follows

$$\Theta_1 \equiv \{\ell \in \Theta : 16Q_\ell \subset 4\mathcal{H}\} \quad \text{and} \quad \Theta_2 \equiv \Theta \setminus \Theta_1. \tag{5.10}$$

Thus, Θ_1 are those indices where $16Q_\ell$ is completely contained in $4\mathcal{H} = \mathcal{H}(4^{2m}s)$ and Θ_2 are the remaining ones where $16Q_\ell \setminus 4\mathcal{H} \neq \emptyset$. In the sequel we will several times distinguish the cases $i \in \Theta_1$, respectively $i \in \Theta_2$. In particular we will exploit the following facts:

$$\gamma r_i^{2m} \geq s/8^{2m} \qquad \forall\ i \in \Theta_2. \tag{5.11}$$

Indeed, for $i \in \Theta_2$ we have $16Q_i \setminus 4\mathcal{H} \neq \emptyset$ and since $i \in \Theta$ we also know that $\mathrm{spt}\omega_i \cap 2\mathcal{H} \neq \emptyset$ and therefore $16Q_i \cap 2\mathcal{H} \neq \emptyset$. From this we infer $2\gamma(16r_i)^{2m} \geq (4^{2m} - 2^{2m})s \geq 2 \cdot 2^{2m}s$ and therefore $\gamma r_i^{2m} \geq s/8^{2m}$. On the other hand, for Θ_1 we shall exploit the following:

$$16Q_i \subset 16Q \qquad \forall\ i \in \Theta_1 \text{ with } 8Q \cap 16Q_i \neq \emptyset \text{ and } r_i < \varrho/4. \tag{5.12}$$

This can be seen as follows: Since $i \in \Theta_1$ we know that $16Q_i \subset 4\mathcal{H}$ (see (5.10)). Therefore we already have the inclusion for the vertical t-direction. Moreover the horizontal slice $16B_i$ of $16Q_i$ has diameter $32r_i < 8\varrho$. Then, since $8Q \cap 16Q_i \neq \emptyset$ we have that the horizontal slice is contained in a ball of radius $\leq 8\varrho + 32r_i < 16\varrho$; i.e., $16B_i \subset 16B$. Together, we deduce the assertion $16Q_i \subset 16Q$.

Now, we are in a position to construct our test-function. When proving a Caccioppoli inequality one would like to choose

$$v_h(z) \equiv v_h(x,t) \equiv [u(x,t) - P_Q(x)]_h\, \eta(x)\zeta(t) \tag{5.13}$$

as test-function. Here $P_Q \colon \mathbb{R}^n \to \mathbb{R}^N$ is the mean value polynomial of degree $\leq m-1$ defined by $(\delta P_Q)_B = (\delta u)_Q$ and $\eta \in C_0^\infty(8B)$, $\zeta \in C_0^\infty(8\mathcal{T})$ are cut-off functions with

$$\begin{cases} \eta \equiv 1 \text{ in } 4B, & 0 \leq \eta \leq 1, \quad |D^k\eta| \leq c(m)/\varrho^k \ \text{ for } 0 \leq k \leq m \\ \zeta \equiv 1 \text{ on } 4\mathcal{T}, & 0 \leq \zeta \leq 1, \quad |\zeta'| \leq 1/s. \end{cases}$$

(Recall that $\alpha B = B(\alpha\varrho)$ and $\alpha\mathcal{T} = \mathcal{T}(\alpha^{2m}s) = (-\alpha^{2m}s, \alpha^{2m}s)$ for $\alpha > 0$.) Finally, $[u - P_Q]_h$ denotes the mollification of $u - P_Q$, in the sense of (3.1) with $0 < h < \min\{\varrho, \sqrt[2m]{s/\gamma}\}$. The latter condition guaranties that $Q_z(h, \gamma h^{2m}) \subset 16Q$ for $z \in \mathrm{spt}(\eta\zeta) \subset 8Q$. Thus v_h is continuous, $\mathrm{spt}v_h \subset \mathrm{spt}(\eta\zeta) \subset 8Q$ and $v_h \to v \equiv (u - P_Q)\eta\zeta$ almost everywhere as $h \searrow 0$.

Since v is only a function of the class $L^{p-\beta}(-T, 0; W^{m,p-\beta}(\Omega))$, it is a priori no admissible test-function. In order to correct this lack of integrability, we modify v_h on the "bad set" $\mathbb{R}^{n+1} \setminus E(\lambda)$, where the maximal function of

$|D^m u|^\xi$ is large. For that purpose we use the covering of $\mathbb{R}^{n+1}\setminus E(\lambda)$ from Lemma 5.2 by the Whitney-type cylinders $(Q_i)_{i=1}^\infty$. Then, our test-function is supposed to equal v_h on the "good set" $E(\lambda)$, while on $\mathbb{R}^{n+1}\setminus E(\lambda)$ we take the mean value polynomials of degree $\leq m-1$ of v_h on $16\mathcal{C}\cap Q_i$, regularized by the partition of unity; i.e., we set

$$w_h(z) \equiv w_h(x,t) \equiv \begin{cases} v_h(z), & \text{for } z \in E(\lambda), \\ \displaystyle\sum_{i=1}^\infty \omega_i(z) P_{v_h,i}(x), & \text{for } z \in \mathbb{R}^{n+1}\setminus E(\lambda), \end{cases} \tag{5.14}$$

where $P_{v_h,i}\colon \mathbb{R}^n \to \mathbb{R}^N$ denotes the mean value polynomial of v_h on $16\mathcal{C}\cap Q_i$ of degree $\leq m-1$ defined by $(\delta P_{v_h,i})_{16B\cap B_i} = (\delta v_h)_{16\mathcal{C}\cap Q_i}$ (recall that $16\mathcal{C} = 16B \times \mathbb{R}$). We note that the function w_h is well defined since the sum in (5.14) is finite for each $z \in \mathbb{R}^{n+1}\setminus E(\lambda)$ (since $\omega_i(z) \neq 0$ for at most $c(n,m)$ indices i) and we can write

$$w_h(z) = \sum_{j\in I(i)} \omega_j(z) P_{v_h,j}(x), \qquad \text{for } z = (x,t) \in Q_i. \tag{5.15}$$

Letting $h \searrow 0$ we see that $v_h \to v = (u - P_Q)\eta\zeta$ almost everywhere and $w_h \to w$ with

$$w(z) \equiv w(x,t) \equiv \begin{cases} v(z), & \text{for } z \in E(\lambda), \\ \displaystyle\sum_{i=1}^\infty \omega_i(z) P_{v,i}(x), & \text{for } z \in \mathbb{R}^{n+1}\setminus E(\lambda), \end{cases}$$

where $P_{v,i}\colon \mathbb{R}^n \to \mathbb{R}^N$ denotes the mean value polynomial of v on $16\mathcal{C}\cap Q_i$ of degree $\leq m-1$ defined by $(\delta P_{v,i})_{16B\cap B_i} = (\delta v)_{16\mathcal{C}\cap Q_i}$. Let us mention that it is enough to consider mean value polynomials on Whitney-cylinders Q_i with $8Q\cap Q_i \neq \emptyset$, since otherwise we have $P_{v_h,i} \equiv 0$ because $\operatorname{spt} v_h \subset 8Q$ (and similarly for v).

Although not always explicitly stated, throughout all of Chapter 5 we work under the permanent assumptions of Lemma 5.1. Moreover, v_h and w_h always denote the functions defined in (5.13), respectively (5.14).

5.2. **Poincaré type estimates on the Whitney-cylinders.** From the geometrical setting it now turns out that we will have to deal with two different scales, namely the one given in the statement of Lemma 5.1, i.e., intrinsic cylinders of the type $Q(\varrho, \lambda_1^{2-p}\varrho^{2m})$, and the one of the Whitney cylinders, i.e., $Q_{z_i}(r_i, \lambda^{2-p} r_i^{2m})$. The crucial point is that both intrinsic scaling factors are independent (to be precise, we only know $\lambda \geq c_E\lambda_1$). Therefore, we will now establish certain Poincaré type estimates of integrals on the Whitney cylinders, which will be adapted to our particular geometric situation. We emphasize, that all these estimates will be based either on hypothesis (5.5) to bound the integral $\fint_{16Q}(|D^m u| + b)^{p-\beta}\,dz$ in terms of $\lambda_1^{p-\beta}$ or Lemma 5.6 (which exploits the geometric properties of the Whitney-cylinders) to bound $\fint_{16Q\cap Q_i}(|D^m u|+b)^{p-\beta}\,dz$ in terms of $\lambda^{p-\beta}$. But before we can start with these estimates we still need a rather geometrical statement.

Lemma 5.4. *Let $Q_{\mathfrak{z}}(r,\gamma r^{2m}) \subset \mathbb{R}^{n+1}$ be a parabolic cylinder with $8Q \cap Q_{\mathfrak{z}}(r,\gamma r^{2m}) \neq \emptyset$. Then for $\alpha > 1$ there holds*

$$|16Q \cap \alpha Q_{\mathfrak{z}}(r,\gamma r^{2m})| \le c(n)\ \alpha^{n+2m}|16Q \cap Q_{\mathfrak{z}}(r,\gamma r^{2m})|.$$

Proof. We recall that $\mathfrak{z} = (\mathfrak{x},\mathfrak{t})$, $\alpha Q_{\mathfrak{z}}(r,\gamma r^{2m}) = Q_{\mathfrak{z}}(\alpha r,\gamma(\alpha r)^{2m})$ and $Q = B \times \mathcal{T}$ with $B = B(\varrho)$, $\mathcal{T} = \mathcal{T}(s) = (-s,s)$. Firstly, we consider the space direction. In the case $B_{\mathfrak{x}}(r) \subset B(16\varrho)$ we have $|B(16\varrho)\cap B_{\mathfrak{x}}(\alpha r)| \le |B_{\mathfrak{x}}(\alpha r)| = \alpha^n|B_{\mathfrak{x}}(r)| = \alpha^n|B(16\varrho)\cap B_{\mathfrak{x}}(r)|$. On the other hand, if $B_{\mathfrak{x}}(r) \setminus B(16\varrho) \neq \emptyset$ we know that $B(8\varrho)\cap B_{\mathfrak{x}}(r) \neq \emptyset$, since $8Q \cap Q_{\mathfrak{z}}(r,\gamma r^{2m}) \neq \emptyset$ by assumption. Therefore there exists a ball $B_y(4\varrho)$ of radius 4ϱ such that $B_y(4\varrho) \subset B(16\varrho)\cap B_{\mathfrak{x}}(r)$ and we have $|B(16\varrho)\cap B_{\mathfrak{x}}(\alpha r)| \le |B(16\varrho)| = 4^n|B_y(4\varrho)| \le 4^n|B(16\varrho)\cap B_{\mathfrak{x}}(r)|$. Together we infer

$$|B(16\varrho)\cap B_{\mathfrak{x}}(\alpha r)| \le \max\{\alpha^n, 4^n\}\ |B(16\varrho)\cap B_{\mathfrak{x}}(r)|.$$

We can use the same argument in the time-direction. In the case that $\mathcal{T}_{\mathfrak{t}}(\gamma r^{2m}) \subset \mathcal{T}(16^{2m}s)$ we have $|\mathcal{T}(16^{2m}s)\cap \mathcal{T}_{\mathfrak{t}}(\gamma(\alpha r)^{2m})| \le |\mathcal{T}_{\mathfrak{t}}(\gamma(\alpha r)^{2m})| = \alpha^{2m}|\mathcal{T}_{\mathfrak{t}}(\gamma r^{2m})| = \alpha^{2m}|\mathcal{T}(16^{2m}s)\cap\mathcal{T}_{\mathfrak{t}}(\gamma r^{2m})|$. On the other hand, if $\mathcal{T}_{\mathfrak{t}}(\gamma r^{2m})\setminus \mathcal{T}(16^{2m}s) \neq \emptyset$ we recall that $\mathcal{T}_{\mathfrak{t}}(\gamma r^{2m}) \cap \mathcal{T}(8^{2m}s) \neq \emptyset$ by assumption and therefore one interval of $\mathcal{T}(16^{2m}s)\setminus\mathcal{T}(8^{2m}s)$ is contained in $\mathcal{T}_{\mathfrak{t}}(\gamma r^{2m})$; i.e., $|\mathcal{T}(16^{2m}s)\cap\mathcal{T}_{\mathfrak{t}}(\gamma r^{2m})| \ge 16^{2m}s - 8^{2m}s \ge (2\cdot 16^{2m}s)/4$. Thus, we have $|\mathcal{T}(16^{2m}s)\cap\mathcal{T}_{\mathfrak{t}}(\gamma(\alpha r)^{2m})| \le |\mathcal{T}(16^{2m}s)| = 2\cdot 16^{2m}s \le 4|\mathcal{T}(16^{2m}s)\cap\mathcal{T}_{\mathfrak{t}}(\gamma r^{2m})|$. Combining both cases, we find

$$|\mathcal{T}(16^{2m}s)\cap\mathcal{T}_{\mathfrak{t}}(\gamma(\alpha r)^{2m})| \le \max\{\alpha^{2m},4\}\ |\mathcal{T}(16^{2m}s)\cap\mathcal{T}_{\mathfrak{t}}(\gamma r^{2m})|.$$

Joining the previous estimates for the space and time direction we conclude the desired estimate. □

Combining the previous lemma with Lemma 3.7 yields the following.

Corollary 5.5. *Let $f\colon \mathbb{R}^{n+1} \to [0,\infty)$ be an integrable function and $[f]_h$ defined according to* (3.1), *with $0 < h < \min\{\varrho, \sqrt[2m]{s/\gamma}\}$. Then, for any parabolic cylinder $Q_{\mathfrak{z}}(r,\gamma r^{2m}) \subset \mathbb{R}^{n+1}$ there holds*

$$\fint_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})}[f]_h\,\chi_{8Q}\,dz \le c(n)\ \fint_{16Q\cap Q_{\mathfrak{z}}(r+h,\gamma(r+h)^{2m})} f\,dz.$$

Proof. We can assume without loss of generality that $8Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})\neq\emptyset$, since otherwise the integral on the left side is identically zero. Noting that $8\varrho+h\le 8\varrho+\varrho<16\varrho$ and $8^{2m}s+\gamma h^{2m}\le 8^{2m}s+s<16^{2m}s$ we find

$$\begin{aligned}\fint_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})}[f]_h\chi_{8Q}\,dz &\le \fint_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})}[f\chi_{16Q}]_h\,dz\\ &\le \tfrac{|Q_{\mathfrak{z}}(r,\gamma r^{2m})|}{|16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})|}\fint_{Q_{\mathfrak{z}}(r,\gamma r^{2m})}[f\chi_{16Q}]_h\,dz.\end{aligned}$$

From Lemma 3.7, (ii) applied with $s=\gamma r^{2m}$ we infer for the integral on the right-hand side (note that $\gamma r^{2m}+\gamma h^{2m}\le\gamma(r+h)^{2m}$)

$$\begin{aligned}\fint_{Q_{\mathfrak{z}}(r,\gamma r^{2m})}[f\chi_{16Q}]_h\,dz &\le c\fint_{Q_{\mathfrak{z}}(r+h,\gamma(r+h)^{2m})}f\chi_{16Q}\,dz\\ &= c\,\tfrac{|16Q\cap Q_{\mathfrak{z}}(r+h,\gamma(r+h)^{2m})|}{|Q_{\mathfrak{z}}(r+h,\gamma(r+h)^{2m})|}\fint_{16Q\cap Q_{\mathfrak{z}}(r+h,\gamma(r+h)^{2m})}f\,dz,\end{aligned}$$

where $c=c(n,m)$. Inserting this above and noting that $|Q_{\mathfrak{z}}(r,\gamma r^{2m})|/|Q_{\mathfrak{z}}(r+h,\gamma(r+h)^{2m})|=(\frac{r}{r+h})^{n+2m}$, while $|16Q\cap Q_{\mathfrak{z}}(r+h,\gamma(r+h)^{2m})|/|16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})|\le c(n)(\frac{r+h}{r})^{n+2m}$ by Lemma 5.4 applied for $\alpha=\frac{r+h}{r}$ (note that we have assumed $8Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})\neq\emptyset$) we deduce the desired estimate. □

The reason for using the Whitney-decomposition is to get uniform estimates for the mean values of $|D^m u|$ on the Whitney-cylinders. This - in a more refined version - is the content of the following.

Lemma 5.6. *Let $\lambda\ge c_E\lambda_1$ and $Q_{\mathfrak{z}}(r,\gamma r^{2m})\subset\mathbb{R}^{n+1}$ be a parabolic cylinder such that $8Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})\neq\emptyset$ and $4Q_{\mathfrak{z}}(r,\gamma r^{2m})\cap E(\lambda)\neq\emptyset$. Then for $1\le\vartheta\le\xi$, where ξ is from (5.3), we have*

$$\fint_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})}(|D^m u|+b)^{\vartheta}\,dz\le c(n,m)\ \lambda^{\vartheta}.$$

Proof. As usual we denote $\mathfrak{z}=(\mathfrak{x},\mathfrak{t})$. The idea of the proof is to enlarge the cylinder $Q_{\mathfrak{z}}(r,\gamma r^{2m})$ by the factor 4 and then exploit the fact that the bigger cylinder has nonempty intersection with $E(\lambda)$, the set where the maximal function of $|D^m u|^{\xi}$ is $\le\lambda$. Let us mention that $16Q\cap 4Q_{\mathfrak{z}}(r,\gamma r^{2m})$ is no cylinder in the case that $16B\cap 4B_{\mathfrak{x}}(r)$ is no ball. However, since $8Q\cap 4Q_{\mathfrak{z}}(r,\gamma r^{2m})\neq\emptyset$ and therefore $8B\cap 4B_{\mathfrak{x}}(r)\neq\emptyset$ we can always find a ball B' with the properties $16B\cap 4B_{\mathfrak{x}}(r)\subset B'$ and $|B'|\le c(n)|16B\cap 4B_{\mathfrak{x}}(r)|$. Indeed, in the case $4B_{\mathfrak{x}}(r)\subset 16B$ we take $B'=4B_{\mathfrak{x}}(r)$, while in the case $4B_{\mathfrak{x}}(r)\setminus 16B\neq\emptyset$ there exists a ball $B_y(\varrho)\subset 16B\cap 4B_{\mathfrak{x}}(r)$ of radius ϱ, since $8B\cap 4B_{\mathfrak{x}}(r)\neq\emptyset$. Taking $B'=16B$ we have $|B'|=16^n|B_y(\varrho)|\le 16^n|16B\cap$

$4B_{\mathfrak{x}}(r)|$. Therefore we can find a cylinder Q', such that $16Q \cap 4Q_{\mathfrak{z}}(r,\gamma r^{2m}) \subset Q'$ and $|Q'| \le c(n)|16Q \cap 4Q_{\mathfrak{z}}(r,\gamma r^{2m})|$. Since $4Q_{\mathfrak{z}}(r,\gamma r^{2m}) \cap E(\lambda) \ne \emptyset$ by assumption and $E(\lambda) \subset \overline{16Q}$, we then also have $\overline{Q'} \cap E(\lambda) \ne \emptyset$. Hence, there exists $\hat{z} \in \overline{Q'} \cap E(\lambda)$ and with Hölder's inequality (note that $\vartheta \le \xi$) and the definition of M_{16Q} in (5.3) we conclude

$$\begin{aligned}\fint_{16Q\cap 4Q_{\mathfrak{z}}(r,\gamma r^{2m})} (|D^m u|+b)^\vartheta\, dz &\le c\left(\fint_{Q'} (|D^m u|+b)^\xi \chi_{16Q}\, dz\right)^{\vartheta/\xi}\\ &\le c\, M_{16Q}(\hat{z})^\vartheta \le c\,\lambda^\vartheta,\end{aligned}$$

where $c = c(n,m)$. We now want to replace $16Q \cap 4Q_{\mathfrak{z}}(r,\gamma r^{2m})$ by $16Q \cap Q_{\mathfrak{z}}(r,\gamma r^{2m})$ in the left-hand side of the last inequality. Applying Lemma 5.4 in the case $\alpha = 4$, which is possible since $Q_{\mathfrak{z}}(r,\gamma r^{2m}) \cap 8Q \ne \emptyset$ by assumption, we find that $|16Q \cap 4Q_{\mathfrak{z}}(r,\gamma r^{2m})| \le c(n,m)\,|16Q \cap Q_{\mathfrak{z}}(r,\gamma r^{2m})|$. Hence, the left-hand side of the previous inequality can be bounded from below by $c^{-1} \fint_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})} (|D^m u| + b)^\vartheta\, dz$. This finally yields the desired estimate. □

Corollary 5.7. *Let* $\lambda \ge c_E\lambda_1$ *and suppose that* $Q_{\mathfrak{z}}(r,\gamma r^{2m}) \subset \mathbb{R}^{n+1}$ *is a parabolic cylinder with* $4Q_{\mathfrak{z}}(r,\gamma r^{2m}) \cap E(\lambda) \ne \emptyset$. *Then for* $0 < h < \min\{\varrho, \sqrt[2m]{s/\gamma}\}$ *and* $1 \le \vartheta \le \xi$ *there holds*

$$\fint_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})} \big[(|D^m u| + b)^\vartheta\big]_h \chi_{8Q}\, dz \le c(n,m)\ \lambda^\vartheta.$$

Proof. We can assume without loss of generality that $8Q \cap Q_{\mathfrak{z}}(r,\gamma r^{2m}) \ne \emptyset$, since otherwise the considered integral is equal to zero. From Corollary 5.5 we infer

$$\begin{aligned}&\fint_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})} \big[(|D^m u| + b)^\vartheta\big]_h \chi_{8Q}\, dz\\ &\qquad\le c(n)\ \fint_{16Q\cap Q_{\mathfrak{z}}(r+h,\gamma(r+h)^{2m})} (|D^m u| + b)^\vartheta\, dz.\end{aligned}$$

Since $8Q \cap Q_{\mathfrak{z}}(r,\gamma r^{2m}) \ne \emptyset$ and $4Q_{\mathfrak{z}}(r,\gamma r^{2m}) \cap E(\lambda) \ne \emptyset$ we know that also the enlarged cylinder $Q_{\mathfrak{z}}(r+h,\gamma(r+h)^{2m})$ has nonempty intersection with $8Q$ as well as $4Q_{\mathfrak{z}}(r+h,\gamma(r+h)^{2m}) \cap E(\lambda) \ne \emptyset$. Therefore we can apply Lemma 5.6 with $16Q\cap\, Q_{\mathfrak{z}}(r+h,\gamma(r+h)^{2m})$ instead of $16Q \cap Q_{\mathfrak{z}}(r,\gamma r^{2m})$ to conclude the desired estimate. □

Since we distinguish the lower and upper level sets $E(\lambda)$ and $\mathbb{R}^{n+1} \setminus E(\lambda)$, where the maximal function of $|D^m u|^\xi$ is smaller, respectively larger then λ

and later we integrate with respect to λ, we heuristically interpret $|D^m u| \approx \lambda$. With this in mind we can think of the subsequent lemma as a sort of Poincaré inequality. This Poincarè type inequality can in particular be applied on the Whitney-cylinders Q_i, or more precisely on $16Q \cap Q_i$.

Lemma 5.8. *Let $\lambda \ge c_E \lambda_1$ and $Q_{\mathfrak{z}}(r, \gamma r^{2m}) \subset \mathbb{R}^{n+1}$ be a parabolic cylinder such that $B_{\mathfrak{z}}(r) \subset 16B$, $8Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m}) \neq \emptyset$ and $4Q_{\mathfrak{z}}(r, \gamma r^{2m}) \cap E(\lambda) \neq \emptyset$. Then for $1 \le \vartheta \le \xi$ and $0 \le k \le m$ there holds*

$$\fint_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})} |D^k(u - P_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})})|^\vartheta \, dz \le c \; r^{\vartheta(m-k)} \lambda^\vartheta,$$

where $c = c(n, N, m, L, \vartheta)$ and $P_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})} \colon \mathbb{R}^n \to \mathbb{R}^N$ denotes the mean value polynomial of u of degree $\le m-1$ on $16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m})$ defined by $(\delta P_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})})_{16B\cap B_{\mathfrak{x}}(r)} = (\delta u)_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})}$.

Proof. Since $B_{\mathfrak{x}}(r) \subset 16B$ we find that $16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m})$ is a parabolic cylinder of the form $16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m}) = B_{\mathfrak{x}}(r) \times \mathcal{I}$, where $\mathcal{I} \subset \mathbb{R}$ is an interval with $|\mathcal{I}| \le 2\gamma r^{2m} = 2\lambda^{2-p} r^{2m}$. Moreover, from the assumptions, we know that $8Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m}) \neq \emptyset$ and $4Q_{\mathfrak{z}}(r, \gamma r^{2m}) \cap E(\lambda) \neq \emptyset$. Therefore we can apply Lemma 5.6 to infer that the assumption (4.4) of Corollary 4.4 is in force with $(16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m}), c(n, m))$ instead of $(Q_{z_0}(\varrho, s), \kappa)$. The application of the corollary then yields the desired estimate with a constant depending only on n, N, m, L and ϑ. □

The next lemma again is a Poincaré type inequality for very weak solutions, similar to the one from Lemma 5.8. But now we subtract the mean value polynomial on the cylinder Q, rather than on $16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m})$.

Lemma 5.9. *Let $\lambda \ge c_E \lambda_1$ and suppose that $Q_{\mathfrak{z}}(r, \gamma r^{2m})$ is a parabolic cylinder such that $8Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m}) \neq \emptyset$ and $4Q_{\mathfrak{z}}(r, \gamma r^{2m}) \cap E(\lambda) \neq \emptyset$. Then for $1 \le \vartheta \le \xi$ and $0 \le k \le m$ there holds*

$$\fint_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})} |D^k(u - P_Q)|^\vartheta \, dz \le c \; \varrho^{\vartheta(m-k)} \, \lambda^\vartheta,$$

where c depends on n, N, m, L, κ. Here $P_Q \colon \mathbb{R}^n \to \mathbb{R}^N$ denotes the mean value polynomial of degree $\le m-1$ defined by $(\delta P_Q)_B = (\delta u)_Q$.

Proof. First we show that it is enough to prove the asserted estimate with P_{16Q} instead of P_Q. For this we in turn apply Lemma 3.6, enlarge the domain of integration from Q to $16Q$ and then apply Corollary 4.4 (note that the hypothesis (4.4) is fulfilled due to (5.5)) to infer for $x \in 16B$

$$
\begin{aligned}
|D^k(P_Q - P_{16Q})(x)| &\le c \sum_{\ell=k}^{m-1} (16\varrho)^{\ell-k} \fint_Q |D^\ell(u - P_{16Q})|\, dz \\
&\le c\ \varrho^{m-k}\, \lambda_1 \le c\ \varrho^{m-k}\, \lambda,
\end{aligned}
$$

where $c = c(n, N, m, L, \kappa)$. Therefore it remains to estimate the integral $\fint_{16Q \cap Q_{\mathfrak{z}}(r,\gamma r^{2m})} |u - P_{16Q}|^\vartheta\, dz$. For that purpose we distinguish two cases. In the following we write $\mathfrak{z} \equiv (\mathfrak{x}, \mathfrak{t})$, as usual.

In **the case** $r \ge \varrho$ we know from the assumption $8Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m}) \neq \emptyset$ that also $8B \cap B_{\mathfrak{x}}(r) \neq \emptyset$. This together with $r \ge \varrho$ yields that $16B \subset 32B_{\mathfrak{x}}(r)$. Then we have $16Q \cap 32Q_{\mathfrak{z}}(r, \gamma r^{2m}) = 16B \times (16\mathcal{T} \cap 32\mathcal{T}_{\mathfrak{t}}(\gamma r^{2m}))$. Moreover, from Lemma 5.4 we know that $|16Q \cap 32Q_{\mathfrak{z}}(r, \gamma r^{2m})|/|16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m})| \le c(n, m)$. Therefore we can enlarge the domain of integration from $16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m})$ to $16Q \cap 32Q_{\mathfrak{z}}(r, \gamma r^{2m})$ and then apply the Poincaré inequality from Corollary 4.3 with $(16B, 16\mathcal{T} \cap 32\mathcal{T}_{\mathfrak{t}}(\gamma r^{2m}), 16\mathcal{T})$ instead of $(B, \mathcal{T}_1, \mathcal{T}_2)$ to obtain

$$
\begin{aligned}
\fint_{16Q \cap Q_{\mathfrak{z}}(r,\gamma r^{2m})} |D^k(u - P_{16Q})|^\vartheta\, dz &\le c \fint_{16Q \cap 32Q_{\mathfrak{z}}(r,\gamma r^{2m})} |D^k(u - P_{16Q})|^\vartheta\, dz \\
&\le c\, \varrho^{\vartheta(m-k)} \Big[\fint_{16Q \cap 32Q_{\mathfrak{z}}(r,\gamma r^{2m})} |D^m u|^\vartheta\, dz + \fint_{16Q} |D^m u|^\vartheta\, dz \\
&\qquad + \Big(\frac{s}{\varrho^{2m}} \fint_{16Q} (|D^m u| + b)^{p-1} dz \Big)^\vartheta \Big] \\
&= c(n, N, m, L, \vartheta)\ \varrho^{\vartheta(m-k)}\, (I + II + III),
\end{aligned}
$$

with the obvious meaning of I - III. To estimate I we notice that from our assumptions on the cylinder $Q_{\mathfrak{z}}(r, \gamma r^{2m})$ we also have $8Q \cap 32Q_{\mathfrak{z}}(r, \gamma r^{2m}) \neq \emptyset$ and $4 \cdot 32Q_{\mathfrak{z}}(r, \gamma r^{2m}) \cap E(\lambda) \neq \emptyset$, as well as $\vartheta \le \xi$. Therefore we can apply Lemma 5.6 for the cylinder $32Q_{\mathfrak{z}}(r, \gamma r^{2m})$ to infer $I \le c(n, m)\ \lambda^\vartheta$.

The estimates for II and III are similar. From Hölder's inequality, the fact that $s/\varrho^{2m} = \lambda_1^{2-p}$ and $\vartheta \le \xi \le p - \beta$, as well as $p - 1 \le p - \beta$, hypothesis (5.5) and $\lambda_1 \le \lambda$ we get

$$
\begin{aligned}
II + III &\le \Big(\fint_{16Q} |D^m u|^{p-\beta} dz \Big)^{\frac{\vartheta}{p-\beta}} + \Big[\lambda_1^{2-p} \Big(\fint_{16Q} (|D^m u| + b)^{p-\beta} dz \Big)^{\frac{p-1}{p-\beta}} \Big]^\vartheta \\
&\le c\ \lambda_1^\vartheta \le c\ \lambda^\vartheta,
\end{aligned}
$$

with $c = c(\kappa, \vartheta)$. Inserting the estimates for $I - III$ above, we deduce the assertion in the case $r \ge \varrho$.

In **the case** $r < \varrho$, we enlarge $Q_{\mathfrak{z}}(r, \gamma r^{2m})$ iteratively by the factor 2, so that we can proceed on the largest cylinder as in the first case. On the intermediate cylinders we exploit the assumption that $4Q_{\mathfrak{z}}(r, \gamma r^{2m})$ has nonempty intersection with $E(\lambda)$. Now, let $\ell \in \mathbb{N}$, such that $2^{\ell-1} r < \varrho \le 2^\ell r$. Then $\operatorname{diam}(B_{\mathfrak{r}}(2^\ell r)) = 2^{\ell+1} r < 4\varrho$, which together with the assumption $8B \cap B_{\mathfrak{r}}(r) \neq \emptyset$ implies $B_{\mathfrak{r}}(2^\ell r) \subset 16B$. For $0 \le j \le \ell$ we set $\widetilde{Q}_j \equiv 16Q \cap 2^j Q_{\mathfrak{z}}(r, \gamma r^{2m})$. Then by construction we have $16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m}) = \widetilde{Q}_0 \subset \widetilde{Q}_1 \subset \cdots \subset \widetilde{Q}_\ell$ and there holds

$$\begin{aligned}
&\Big(-\!\!\!\!\!\!\int_{\widetilde{Q}_0} |D^k(u - P_{16Q})|^\vartheta dz \Big)^{\frac{1}{\vartheta}} \le \Big(-\!\!\!\!\!\!\int_{\widetilde{Q}_0} |D^k(u - P_{\widetilde{Q}_0})|^\vartheta \, dz \Big)^{\frac{1}{\vartheta}} \\
&+ \sum_{j=1}^{\ell} \Big(-\!\!\!\!\!\!\int_{\widetilde{Q}_0} |D^k(P_{\widetilde{Q}_{j-1}} - P_{\widetilde{Q}_j})|^\vartheta \, dz \Big)^{\frac{1}{\vartheta}} + \Big(-\!\!\!\!\!\!\int_{\widetilde{Q}_0} |D^k(P_{\widetilde{Q}_\ell} - P_{16Q})|^\vartheta \, dz \Big)^{\frac{1}{\vartheta}} \\
&= I + II + III,
\end{aligned} \tag{5.16}$$

where $P_{\widetilde{Q}_j} : \mathbb{R}^n \to \mathbb{R}^N$ denote the mean value polynomials of u of degree $\le m-1$ defined by $(\delta P_{\widetilde{Q}_j})_{B_{\mathfrak{r}}(2^j r)} = (\delta u)_{\widetilde{Q}_j}$. Before we estimate I - III we mention that we are allowed to apply Lemma 5.8 on the cylinders $\widetilde{Q}_j$ due to the fact that $B_{\mathfrak{r}}(2^\ell r) \subset 16B$. An estimate for I therefore directly follows from Lemma 5.8 applied to the cylinder $\widetilde{Q}_0 = 16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m})$; i.e., $I \le c\, r^{m-k} \lambda \le c\, \varrho^{m-k} \lambda$.

For the estimate of II we first note that $B_{\mathfrak{r}}(2^\ell r) \subset 16B$ and from the assumptions we know that $8Q \cap \widetilde{Q}_j \neq \emptyset$ and $4\widetilde{Q}_j \cap E(\lambda) \neq \emptyset$. Therefore, the assumptions of Lemma 5.8 are satisfied for $(16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m}), \vartheta)$ replaced by $(\widetilde{Q}_j, 1)$. Applying in turn Lemma 3.6, Lemma 5.4 to find $|\widetilde{Q}_j|/|\widetilde{Q}_{j-1}| \le c(n,m)$ and Lemma 5.8 as indicated then yields for $x \in B_{\mathfrak{r}}(2^j r)$

$$\begin{aligned}
|D^k(P_{\widetilde{Q}_{j-1}} - P_{\widetilde{Q}_j})(x)| &\le c \sum_{i=k}^{m-1} (2^j r)^{i-k} -\!\!\!\!\!\!\int_{\widetilde{Q}_{j-1}} |D^i(u - P_{\widetilde{Q}_j})| \, dz \\
&\le c \sum_{i=k}^{m-1} (2^j r)^{\ell-k} -\!\!\!\!\!\!\int_{\widetilde{Q}_j} |D^i(u - P_{\widetilde{Q}_j})| \, dz \le c\, (2^j r)^{m-k} \lambda,
\end{aligned}$$

where $c = c(n, N, m, L)$. From this estimate and the fact that $r < 2^{1-\ell}\varrho$ we infer

$$II \le c \sum_{j=1}^{\ell} (2^j r)^{m-k} \lambda = c\, \varrho^{m-k} \lambda \sum_{j=1}^{\ell} \big(\tfrac{2^j r}{\varrho}\big)^{m-k} \le c\, \varrho^{m-k} \lambda \sum_{j=1}^{\ell} (2^{j+1-\ell})^{m-k}$$

$$\leq c\,\varrho^{m-k}\lambda,$$

where $c = c(n, N, m, L, \vartheta)$.

Finally we come to the estimate for III. To bound the difference of the polynomials in III we use Lemma 3.6 with $(B_{\mathfrak{x}}(2^\ell r), B_{\mathfrak{x}}(2^\ell r), \mathcal{I}_j, P_{\widetilde{Q}_\ell}, P_{16Q})$ instead of $(\mathcal{O}_1, \mathcal{O}_2, \mathcal{T}, P_{\mathcal{O}_1\times\mathcal{T}}, P)$. Since we have $2^\ell r \geq \varrho$ and $\widetilde{Q}_\ell = 16Q \cap 2^\ell Q_{\mathfrak{z}}(r, \gamma r^{2m})$ we then are in position to use the results from the first case (with $\vartheta = 1$) in order to estimate the remaining integrals by $c\,\varrho^{m-i}\lambda$, and due to our choice of ℓ we also have $2^\ell r \leq 2\varrho$. Proceeding this way we deduce

$$III \leq c \sum_{i=k}^{m-1} (2^\ell r)^{i-k} \fint_{\widetilde{Q}_\ell} |D^i(u - P_{16Q})|\,dz \leq c \sum_{i=k}^{m-1} \varrho^{i-k} \varrho^{m-i}\lambda \leq c\,\varrho^{m-k}\,\lambda,$$

where $c = c(n, N, m, L, \kappa)$. Joining the previous estimates for I - III with (5.16), taking this to the power ϑ and noting again that $\widetilde{Q}_0 = 16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m})$ we infer the assertion of the lemma also in the remaining case $r < \varrho$. □

From the last lemma we immediately conclude a first Poincaré type estimate for the function v_h defined in (5.13).

Corollary 5.10. *Let $\lambda \geq c_E \lambda_1$ and $Q_{\mathfrak{z}}(r, \gamma r^{2m}) \subset \mathbb{R}^{n+1}$ be a parabolic cylinder with $4Q_{\mathfrak{z}}(r, \gamma r^{2m}) \cap E(\lambda) \neq \emptyset$. Then, for $0 \leq k \leq m$ and $1 \leq \vartheta \leq \xi$ there holds*

$$\fint_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})} |D^k v_h|^\vartheta\,dz \leq c(n, N, m, L, \kappa, \vartheta)\ \varrho^{\vartheta(m-k)}\,\lambda^\vartheta.$$

Proof. Recalling the definition $v_h = [u - P_Q]_h \eta\zeta$ we first compute $D^k v_h = \sum_{\ell=0}^{k} \binom{k}{\ell} D^{k-\ell}\eta \odot D^\ell [u - P_Q]_h \zeta$. From the fact that $\operatorname{spt}(\eta\zeta) \subset 8Q$, $0 \leq \zeta \leq 1$ and $|D^{k-\ell}\eta| \leq c/\varrho^{k-\ell}$ we therefore obtain

$$\begin{aligned}\fint_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})} |D^k v_h|^\vartheta dz &\leq c \sum_{\ell=0}^{k} \varrho^{-\vartheta(k-\ell)} \fint_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})} |D^\ell [u - P_Q]_h|^\vartheta \chi_{8Q}\,dz \\ &\leq c \sum_{\ell=0}^{k} \varrho^{-\vartheta(k-\ell)}\,\varrho^{\vartheta(m-\ell)}\,\lambda^\vartheta \leq c\,\varrho^{\vartheta(m-k)}\,\lambda^\vartheta.\end{aligned}$$

Here we have also applied Corollary 5.5 and Lemma 5.9 in the second line (note that $h \leq \min\{\varrho, \sqrt[2m]{s/\gamma}\}$ and we can assume without loss of generality that $8Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m}) \neq \emptyset$, since otherwise the considered integral is identically zero). Note that the constant c depends on n, N, m, L, κ and ϑ. This proves the desired estimate. □

The next lemma is a sort of improved Poincaré inequality for v_h, which will be applied on the Whitney cylinders Q_i with $i \in \Theta_1$ later. Since two different configurations are involved which correspond to the choice of the cut-off function and the Whitney decomposition, it will be important to have the smaller radius in the right-hand side of the estimate.

Lemma 5.11. *Let $\lambda \geq c_E\lambda_1$ and $Q_{\mathfrak{z}}(r, \gamma r^{2m}) \subset 4\mathcal{H}$ be a parabolic cylinder satisfying $4Q_{\mathfrak{z}}(r, \gamma r^{2m}) \cap E(\lambda) \neq \emptyset$. Then, for $0 \leq k \leq m$ and $\vartheta \geq 1$ there holds*

$$\begin{aligned}&\fint_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})} |D^k(v_h - P)|^\vartheta dz\\ &\qquad\qquad \leq c\,\min\{r,\varrho\}^{\vartheta(m-k)}\Big(\lambda^\vartheta + \delta_{\vartheta>\xi}\fint_{16Q\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})}|D^m v_h|^\vartheta dz\Big),\end{aligned}$$

where $P\colon \mathbb{R}^n \to \mathbb{R}^N$ denotes the mean value polynomial of v_h of degree $\leq m-1$ on $16\mathcal{C} \cap Q_{\mathfrak{z}}(r, \gamma r^{2m})$ defined by $(\delta P)_{16B\cap B_{\mathfrak{x}}(r)} = (\delta v_h)_{16\mathcal{C}\cap Q_{\mathfrak{z}}(r,\gamma r^{2m})}$ and $\delta_{\vartheta>\xi} = 1$ if $\vartheta > \xi$ and $\delta_{\vartheta>\xi} = 0$ if $1 \leq \vartheta \leq \xi$. The constant c depends only on n, N, m, L, κ and ϑ.

Proof. Without loss of generality we can assume $8Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m}) \neq \emptyset$. Otherwise we have $P \equiv 0 \equiv v_h$ on $Q_{\mathfrak{z}}(r, \gamma r^{2m})$, since $\mathrm{spt} v_h \subset 8Q$ by construction. Then the asserted estimates hold trivially. We start proving the first assertion of the lemma; i.e., we first consider $\vartheta \geq 1$.

Initially, we observe that from our assumption $Q_{\mathfrak{z}}(r, \gamma r^{2m}) \subset 4\mathcal{H}$ it follows that $\mathcal{T}_{\mathfrak{t}}(\gamma r^{2m}) \subset 4\mathcal{T}$. Therefore, we find that our domain of integration equals $16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m}) = (16B \cap B_{\mathfrak{x}}(r)) \times \mathcal{T}_{\mathfrak{t}}(\gamma r^{2m})$. In order to apply the Poincaré type inequality from Lemma 4.1 for $\mathcal{O} = 16B \cap B_{\mathfrak{x}}(r)$, we first have to construct an appropriate weight-function $\widetilde{\eta} \in C_0^\infty(16B \cap B_{\mathfrak{x}}(r))$. From the geometric situation it is clear that we have to distinguish two cases. In the case $B_{\mathfrak{x}}(r) \subset 16B$ we have $16B \cap B_{\mathfrak{x}}(r) = B_{\mathfrak{x}}(r)$. Then we choose $\widetilde{\eta} \in C_0^\infty(B_{\mathfrak{x}}(r))$ such that $\widetilde{\eta} \geq 0$, $\int_{B_{\mathfrak{x}}(r)} \widetilde{\eta}\,dx = 1$ and $|D^\ell\widetilde{\eta}| \leq c/r^{n+\ell}$ for $0 \leq \ell \leq 2m$. On the other hand, in the case $B_{\mathfrak{x}}(r) \setminus 16B \neq \emptyset$ we know that $r \geq 4\varrho$ since $8B \cap B_{\mathfrak{x}}(r) \neq \emptyset$. This allows us to choose a ball $B_y(\varrho)$ contained in $16B \cap B_{\mathfrak{x}}(r)$. Moreover, since $\mathrm{diam}(16B) = 32\varrho$ we have $16B \cap B_{\mathfrak{x}}(r) \subset B_y(32\varrho)$. Then we choose $\widetilde{\eta} \in C_0^\infty(B_y(\varrho))$ such that $\widetilde{\eta} \geq 0$, $\int_{B_y(\varrho)} \widetilde{\eta}\,dx = 1$ and $|D^\ell\widetilde{\eta}| \leq c/\varrho^{n+\ell}$ for $0 \leq \ell \leq 2m$. Together both cases yield a weight-function $\widetilde{\eta} \in C_0^\infty(16B \cap B_{\mathfrak{x}}(r))$ satisfying $\widetilde{\eta} \geq 0$, $\int_{16B\cap B_{\mathfrak{x}}(r)} \widetilde{\eta}\,dx = 1$ and $|D^\ell\widetilde{\eta}| \leq c\,\max\{r^{-(n+\ell)}, \varrho^{-(n+\ell)}\}$ for $0 \leq \ell \leq 2m$, where the smallest possible value of c depends on n and m. Because $16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m}) = 16B \cap$

$B_{\mathfrak{x}}(r) \times \mathcal{T}_{\mathfrak{t}}(\gamma r^{2m})$ satisfies the set of geometric assumptions from Lemma 4.1 we now use the weight-function $\widetilde{\eta}$ to construct the weighted means of v_h, i.e., $(v_h)_{D^\alpha \widetilde{\eta}}$, and apply the lemma with $(v_h, \widetilde{\eta}, 16B \cap B_{\mathfrak{x}}(r), \mathcal{T}_{\mathfrak{t}}(\gamma r^{2m}), \mathcal{T}_{\mathfrak{t}}(\gamma r^{2m}))$ instead of $(u, \eta, \mathcal{O}, \mathcal{T}_1, \mathcal{T}_2)$ to infer

$$\begin{aligned} &⨍_{16Q \cap Q_{\mathfrak{z}}(r,\gamma r^{2m})} |D^k(v_h - P)|^\vartheta \, dz \\ &\quad \le c \min\{\varrho, r\}^{\vartheta(m-k)} ⨍_{16Q \cap Q_{\mathfrak{z}}(r,\gamma r^{2m})} |D^m v_h|^\vartheta \, dz \\ &\qquad + c \sum_{k \le |\alpha| \le m-1} \min\{\varrho, r\}^{\vartheta(|\alpha|-k)} \sup_{t_1, t_2 \in \mathcal{T}_{\mathfrak{t}}(\gamma r^{2m})} |(v_h)_{D^\alpha \widetilde{\eta}}(t_2) - (v_h)_{D^\alpha \widetilde{\eta}}(t_1)|^\vartheta. \end{aligned} \tag{5.17}$$

We still have to estimate the second term on the right-hand side. As we mentioned above, we have $\mathcal{T}_{\mathfrak{t}}(\gamma r^{2m}) \subset 4\mathcal{T}$. This implies, by the choice of ζ and the cut-off function in time in the definition of v_h, that $\zeta \equiv 1$ on $\mathcal{T}_{\mathfrak{t}}(\gamma r^{2m})$. So $v_h(x,t) = [u(x,t) - P_Q(x)]_h \eta(x) = [u(x,t)]_h \eta(x) - [P_Q(x)]_h \eta(x)$ for $(x,t) \in Q_{\mathfrak{z}}(r, \gamma r^{2m})$. Since $[P_Q(x)]_h \eta(x)$ is independent of t, this term cancels out in the following calculation, with α being a multiindex of order $k \le |\alpha| \le m-1$:

$$\begin{aligned} |(v_h)_{D^\alpha \widetilde{\eta}}(t_2) - (v_h)_{D^\alpha \widetilde{\eta}}(t_1)| &= |([u]_h \eta)_{D^\alpha \widetilde{\eta}}(t_2) - ([u]_h \eta)_{D^\alpha \widetilde{\eta}}(t_1)| \\ &= |([u]_h)_{\eta D^\alpha \widetilde{\eta}}(t_2) - ([u]_h)_{\eta D^\alpha \widetilde{\eta}}(t_1)|. \end{aligned}$$

Applying Lemma 4.2 with $\eta D^\alpha \widetilde{\eta}$ instead of η and recalling that $\operatorname{spt}(\eta D^\alpha \widetilde{\eta}) \subset 8B \cap B_{\mathfrak{x}}(r)$ we infer

$$\begin{aligned} &|(v_h)_{D^\alpha \widetilde{\eta}}(t_2) - (v_h)_{D^\alpha \widetilde{\eta}}(t_1)| \\ &\qquad \le c \sum_{\ell=0}^m \|D^\ell(\eta D^\alpha \widetilde{\eta})\|_{L^\infty} \int_{8Q \cap Q_{\mathfrak{z}}(r,\gamma r^{2m})} \big[(|D^m u| + b)^{p-1}\big]_h \, dz, \end{aligned} \tag{5.18}$$

where $c = c(N, L)$. In order to estimate the right-hand side of (5.18) we distinguish between the cases $r > \varrho$ and $r \le \varrho$.

In **the case** $r > \varrho$ we have

$$\begin{aligned} |D^\ell(\eta D^\alpha \widetilde{\eta})| &\le c \sum_{j=0}^{\ell} |D^{\ell-j}\eta| |D^j D^\alpha \widetilde{\eta}| \\ &\le c \sum_{j=0}^{\ell} \varrho^{j-\ell} \max\{r^{-n-j-|\alpha|}, \varrho^{-n-j-|\alpha|}\} \le c\, \varrho^{-(n+m+|\alpha|)}. \end{aligned}$$

Now using in turn Lemma 3.7 (ii), the fact that $h \le \min\{\varrho, \sqrt[2m]{s/\gamma}\}$ (see the construction in (3.1)) and therefore $8\varrho + h \le 16\varrho$ and $8^{2m}s + \gamma h^{2m} \le 16^{2m}s$, Hölder's inequality (note that $p-1 \le p-\beta$), the hypothesis (5.5) and finally $s = \lambda_1^{2-p}\varrho^{2m}$ we can continue the estimate in (5.18) as follows

$$\begin{aligned}|(v_h)_{D^\alpha\widetilde{\eta}}(t_2) - (v_h)_{D^\alpha\widetilde{\eta}}(t_1)| &\le c\,\varrho^{-(n+m+|\alpha|)}\,|16Q| \fint_{16Q} (|D^m u| + b)^{p-1}\,dz \\ &\le c\,\varrho^{-(n+m+|\alpha|)}\,|16Q| \Big(\fint_{16Q} (|D^m u| + b)^{p-\beta} dz\Big)^{\frac{p-1}{p-\beta}} \\ &\le c\,\varrho^{-(m+|\alpha|)}\,s\,\lambda_1^{p-1} = c\,\varrho^{m-|\alpha|}\,\lambda_1 \le c\,\varrho^{m-|\alpha|}\lambda,\end{aligned}$$

where c depends on n, N, m, L and κ. Joining this inequality with (5.17), we conclude the desired estimate in the case $r > \varrho$.

In **the case** $r \le \varrho$ we have

$$\begin{aligned}|D^\ell(\eta D^\alpha\widetilde{\eta})| &\le c \sum_{j=0}^{\ell} |D^{\ell-j}\eta||D^j D^\alpha \widetilde{\eta}| \\ &\le c \sum_{j=0}^{\ell} \varrho^{j-\ell} \max\{r^{-n-j-|\alpha|}, \varrho^{-n-j-|\alpha|}\} \le c\, r^{-(n+m+|\alpha|)}.\end{aligned}$$

Due to the assumptions we know that $4Q_{\mathfrak{z}}(r, \gamma r^{2m}) \cap E(\lambda) \neq \emptyset$. Therefore we can apply Corollary 5.7 in the case $\vartheta = p-1$ to estimate the integral on the right-hand side of (5.18). Proceeding this way we infer

$$\begin{aligned}|(v_h)_{D^\alpha\widetilde{\eta}}(t_2) - (v_h)_{D^\alpha\widetilde{\eta}}(t_1)| &\le c\; r^{-(n+m+|\alpha|)}|16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m})|\,\lambda^{p-1} \\ &\le c\; r^{m-|\alpha|}\gamma\,\lambda^{p-1} = c\; r^{m-|\alpha|}\lambda,\end{aligned}$$

where we have used $\gamma = \lambda^{2-p}$ in the last line and $c = c(n, m, L)$. Combining this inequality with (5.17) we deduce the desired estimate also in the remaining case $r \le \varrho$. Note that in any case the constant depends only on n, N, m, L, κ and ϑ. This proves the assertion of the lemma when $\vartheta > \xi$. Finally, in the case $1 \le \vartheta \le \xi$ we can apply Corollary 5.10 to further estimate the integral $\fint_{16Q \cap Q_{\mathfrak{z}}(r,\gamma r^{2m})} |D^m v_h|^\vartheta\,dz$ in terms of λ^ϑ. This finishes the proof of Lemma 5.11. □

In the subsequent lemma we derive estimates for the mean value polynomials of v_h which are a consequence of the previous Poincaré type estimates. In Chapter 5.3 this will immediately lead to bounds of the test-function w_h on the "bad set."

Lemma 5.12. *Let $\lambda \ge c_E\lambda_1$ and $Q_i = B_i \times \mathcal{T}_i = B_{x_i}(r_i) \times \mathcal{T}_{t_i}(\gamma r_i^{2m})$ be one of the Whitney-cylinders and $P_{v_h,i}\colon \mathbb{R}^n \to \mathbb{R}^N$ the mean value polynomial of v_h of degree $\le m-1$ defined in (5.14) by $(\delta P_{v_h,i})_{16B\cap B_i} = (\delta v_h)_{16\mathcal{C}\cap Q_i}$. Then for any $0 \le k \le m-1$ and $\mathfrak{x} \in 16B \cap 16B_i$ we have*

$$|D^k P_{v_h,i}(\mathfrak{x})| \le c\,\varrho^{m-k}\,\lambda \tag{5.19}$$

and

$$|D^k P_{v_h,i}(\mathfrak{x})| \le c\,r_i^{-k}\Big(\min\big\{r_i^m\lambda, \sqrt{s\lambda^p}\big\} + \fint_{16Q\cap Q_i} |v_h|\,dz\Big). \tag{5.20}$$

If $i \in \Theta_1$, then for any $j \in I(i)$, $0 \le k \le m-1$ and $\mathfrak{x} \in 16B \cap 16B_i$ there holds

$$|D^k P_{v_h,j}(\mathfrak{x}) - D^k P_{v_h,i}(\mathfrak{x})| \le c\,\min\{r_i, \varrho\}^{m-k}\,\lambda. \tag{5.21}$$

In any case the constant c depends only on n, N, m, L and κ.

Proof. From the definition of v_h in (5.13) we recall that $\operatorname{spt} v_h \subset 8Q$. Therefore we can assume without loss of generality that $8Q \cap Q_i \ne \emptyset$; otherwise we have $P_{v_h,i} \equiv 0$.

We start with the proof of (5.19). Here, we first estimate the mean value polynomial $|D^k P_{v_h,i}(\mathfrak{x})|$ for $\mathfrak{x} \in 16B \cap 16B_i$ by an application of Lemma 3.4 with $(\mathcal{O}_1, \mathcal{O}_2, P)$ replaced by $(16B \cap B_i, 16B \cap 16B_i, P_{v_h,i})$. This yields

$$|D^k P_{v_h,i}(\mathfrak{x})| \le c(n,m) \sum_{\ell=k}^{m-1} \min\{r_i, \varrho\}^{\ell-k} |(D^\ell v_h)_{16\mathcal{C}\cap Q_i}|. \tag{5.22}$$

Here we have also used $\operatorname{diam}(16B \cap 16B_i) \le 32\min\{r_i, \varrho\}$ and the defining property of the mean value polynomial $P_{v_h,i}$; i.e., $(D^\ell P_{v_h,i})_{16B\cap B_i} = (D^\ell v_h)_{16\mathcal{C}\cap Q_i}$. Since $\operatorname{spt} v_h \subset 8Q \subset 16Q$ and $|16Q \cap Q_i|/|16\mathcal{C} \cap Q_i| \le 1$ we can now apply Corollary 5.10 to further estimate

$$|D^k P_{v_h,i}(\mathfrak{x})| \le c \sum_{\ell=k}^{m-1} \varrho^{\ell-k} \fint_{16Q\cap Q_i} |D^\ell v_h|\,dz \le c \sum_{\ell=k}^{m-1} \varrho^{\ell-k}\,\varrho^{m-\ell}\,\lambda = c\,\varrho^{m-k}\,\lambda,$$

yielding the bound in (5.19).

Next we will prove the estimate (5.20). Applying the interpolation lemma, i.e., Lemma 3.1, on the horizontal slices of Q_i, i.e., on B_i, we obtain

$$\begin{aligned} |(D^\ell v_h)_{16\mathcal{C}\cap Q_i}| &\le \tfrac{1}{|16\mathcal{C}\cap Q_i|} \int_{Q_i} |D^\ell v_h|\,dz \\ &\le \tfrac{c}{|16\mathcal{C}\cap Q_i|} \Big(r_i^{m-\ell} \int_{Q_i} |D^m v_h|\,dz + r_i^{-\ell} \int_{Q_i} |v_h|\,dz \Big) \end{aligned}$$

$$\le c \,\tfrac{|16Q\cap Q_i|}{|16\mathcal{C}\cap Q_i|}\, r_i^{m-\ell}\lambda + c\, r_i^{-\ell} \fint_{16Q\cap Q_i} |v_h|\, dz.$$

Here, we have applied Corollary 5.10 in the last line, which is allowed since $4Q_i \cap E(\lambda) \neq \emptyset$ by Lemma 5.2, (iii) and $\int_{Q_i} |D^m v_h|\, dz = \int_{16Q\cap Q_i} |D^m v_h|\, dz$ because $\mathrm{spt} v_h \subset 16Q$ and we have used the fact that $|16Q\cap Q_i|/|16\mathcal{C}\cap Q_i| \le 1$. Moreover, we even have (note that $\gamma = \lambda^{2-p}$)

$$\tfrac{|16Q\cap Q_i|}{|16C\cap Q_i|} \le \min\big\{1, \tfrac{s}{\gamma r_i^{2m}}\big\} \le \min\Big\{1, \sqrt{\tfrac{s}{\gamma r_i^{2m}}}\Big\} = \tfrac{1}{r_i^m \lambda}\, \min\big\{r_i^m \lambda, \sqrt{s\lambda^p}\big\}.$$

Inserting this above and then using the resulting estimate for $|(D^\ell v_h)_{16\mathcal{C}\cap Q_i}|$ in (5.22) yields the bound in (5.20).

Finally, we come to the proof of (5.21). We fix $i \in \Theta_1$ (see (5.10) for the definition of Θ_1). Then $16Q_i \subset 4\mathcal{H}$ and for $j \in I(i)$ we know that $r_i \le 3r_j$ and $Q_j \subset 16Q_i$ by (5.8). We first note that (5.19) applied to $|D^k P_{v_h,i}|$ and $|D^k P_{v_h,j}|$ yields

$$|D^k P_{v_h,j}(\mathfrak{x}) - D^k P_{v_h,i}(\mathfrak{x})| \le |D^k P_{v_h,j}(\mathfrak{x})| + |D^k P_{v_h,i}(\mathfrak{x})| \le c\, \varrho^{m-k}\lambda. \quad (5.23)$$

Hence, it remains to consider the case where r_i is small compared to ϱ, i.e., when $r_i \le \varrho/4$. In this case we know from (5.12) that $16Q_i \subset 16Q$; i.e., $16Q \cap 16Q_i = 16Q_i$, as well as $16Q\cap Q_j = Q_j$ for $j \in I(i)$. Denoting by $\widehat{P}_{v_h,i}$ the mean value polynomial of degree $\le m-1$, defined by $(\delta \widehat{P}_{v_h,i})_{16B_i} = (\delta v_h)_{16Q_i}$ we can apply Lemma 3.6 with $(B_j, 16B_i, \mathcal{T}_{t_j}(\gamma r_j^{2m}), P_{v_h,j}, \widehat{P}_{v_h,i})$ instead of $(\mathcal{O}_1, \mathcal{O}_2, \mathcal{T}, P_{\mathcal{O}_1\times\mathcal{T}}, P)$. The application of the lemma, together with the fact that $|16Q_i|/|Q_j| = (16r_i/r_j)^{n+2m} \le c(n,m)$ then yields, for $\mathfrak{x} \in 16B_i$,

$$\begin{aligned}|D^k P_{v_h,j}(\mathfrak{x}) - D^k \widehat{P}_{v_h,i}(\mathfrak{x})| &\le c \sum_{\ell=k}^{m-1} r_i^{\ell-k} \fint_{Q_j} |D^\ell (v_h - \widehat{P}_{v_h,i})|\, dz \\ &\le c \sum_{\ell=k}^{m-1} r_i^{\ell-k} \fint_{16Q_i} |D^\ell (v_h - \widehat{P}_{v_h,i})|\, dz,\end{aligned}$$

where $c = c(n,m)$. We now apply Lemma 5.11 with $\vartheta = 1 \le \xi$ to estimate the integrals appearing on the right-hand side of the previous inequality by $c r_i^{m-\ell}\lambda$. This is possible because $16Q_i \cap E(\lambda) \neq \emptyset$ and $16Q_i \subset 4\mathcal{H}$ by assumption. Proceeding this way we arrive at

$$|D^k P_{v_h,j}(\mathfrak{x}) - D^k \widehat{P}_{v_h,i}(\mathfrak{x})| \le c\, \sum_{\ell=k}^{m} r_i^{\ell-k} r_i^{m-\ell}\, \lambda = c\, r_i^{m-k}\, \lambda.$$

Recalling that by construction $i \in I(i)$ the previous estimate with i instead of j yields immediately:

$$
\begin{aligned}
&|D^k P_{v_h,j}(\mathfrak{x}) - D^k P_{v_h,i}(\mathfrak{x})| \\
&\le |D^k P_{v_h,j}(\mathfrak{x}) - D^k \widehat{P}_{v_h,i}(\mathfrak{x})| + |D^k P_{v_h,i}(\mathfrak{x}) - D^k \widehat{P}_{v_h,i}(\mathfrak{x})| \le c\, r_i^{m-k}\, \lambda.
\end{aligned}
$$

Together with (5.23) this implies the desired estimate with a constant $c = c(n, N, m, L, \kappa)$. □

5.3. **Properties of the test-function.** In the following we will establish bounds for w_h and its derivatives on the "bad set" $(16B \times 2\mathcal{T}) \setminus E(\lambda)$. As before we work on the permanent assumptions of Lemma 5.1 and we recall that v_h and w_h denote the functions defined in (5.13), respectively (5.14).

Lemma 5.13. *Suppose that $\lambda \ge c_E \lambda_1$. Then w_h is bounded on $16Q \setminus E(\lambda)$ with*

$$
\|w_h\|_{L^\infty(16Q\setminus E(\lambda);\mathbb{R}^N)} \le c(n, N, m, L, \kappa)\ \varrho^m\, \lambda.
$$

Proof. Let $z = (x, t) \in 16Q \setminus E(\lambda)$. Then there exists a Whitney cylinder Q_i with $z \in Q_i$. From the definition of w_h, i.e., from (5.15), we infer that $w_h(z) = \sum_{j\in I(i)} \omega_j(z) P_{v_h,j}(x)$. Using the bound (5.19) from Lemma 5.12 for the mean value polynomials $P_{v_h,j}$ we see that $|P_{v_h,j}(x)| \le c\, \varrho^m \lambda$. This yields the desired estimate, since $(\omega_i)_{i\in\mathbb{N}}$ forms a partition of unity. Note that $c = c(n, N, m, L, \kappa)$. □

Lemma 5.14. *Suppose that $\lambda \ge c_E \lambda_1$ and Q_i is one of the Whitney-cylinders. Then, for any $0 \le k \le m$ and $\mathfrak{z} \in 16Q \cap \operatorname{spt}\omega_i$ we can bound $D^k w_h$ as follows:*

$$
|D^k w_h(\mathfrak{z})| \le c\, r_i^{-k} \Big(\min\big\{ r_i^m \lambda, \sqrt{s\lambda^p} \big\} + \fint_{16Q\cap 16Q_i} |v_h|\, dz \Big). \tag{5.24}
$$

In the case $i \in \Theta_1$, or $i \in \Theta_2$ and $\varrho \le r_i$ there holds

$$
|D^k w_h(\mathfrak{z})| \le c\ \varrho^{m-k}\, \lambda. \tag{5.25}
$$

In the case $i \in \Theta_1$ we have for any $0 < \varepsilon \le 1$

$$
|D^k w_h(\mathfrak{z})| \le c\ r_i^{m-k} \Big(\tfrac{\lambda}{\varepsilon} + \tfrac{\varepsilon}{\lambda\, r_i^{2m}} \fint_{B_i} |P_{v_h,i}|^2\, dx \Big). \tag{5.26}
$$

In the case $i \in \Theta_2$ there holds

$$
|D^k w_h(\mathfrak{z})| \le c\ r_i^{m-k} \Big(\lambda + \tfrac{\lambda^{1-p}}{s} \fint_{16Q\cap 16Q_i} |v_h|^2\, dz \Big). \tag{5.27}
$$

In any case the constant c depends only on n, N, m, L and κ.

Proof. Let $\mathfrak{z} = (\mathfrak{x}, \mathfrak{t}) \in 16Q \cap \operatorname{spt}\omega_i$ for some $i \in \mathbb{N}$. With the definition of w_h, i.e., (5.15), we compute

$$D^k w_h(\mathfrak{z}) = \sum_{\ell=0}^{k} \binom{k}{\ell} \sum_{j \in I(i)} D^{k-\ell}\omega_j(\mathfrak{z}) \odot D^\ell P_{v_h,j}(\mathfrak{x}) = \sum_{\ell=0}^{k} I^{(\ell)}, \tag{5.28}$$

with the obvious meaning of $I^{(\ell)}$.

We first prove (5.24). From the bound (5.20) from Lemma 5.12 for $|D^\ell P_{v_h,j}(\mathfrak{x})|$, the fact that $|D^{k-\ell}\omega_j| \le c/r_j^{k-\ell}$ and $1/r_j \le 3/r_i$ by (5.8) we find

$$|I^{(\ell)}| \le c\, r_i^{-k} \Big(\min\big\{r_i^m \lambda, \sqrt{s\lambda^p}\big\} + \sum_{j \in I(i)} ⨍_{16Q \cap Q_j} |v_h|\, dz \Big).$$

We now want to estimate the remaining sum of mean value integrals of $|v_h|$ on $16Q \cap Q_j$ in terms of the mean value integral on $16Q \cap 16Q_i$. We recall from (5.8) that $Q_j \subset 16Q_i$ for any $j \in I(i)$. On the other hand we also have $Q_i \subset 16Q_j$, since $i \in I(j)$, which implies $16Q_i \subset 4 \cdot 16Q_j$ (note $r_i \le 3r_j$). In the case that $8Q \cap Q_j \neq \emptyset$ (and we only have to consider such $j \in I(i)$ with $8Q \cap Q_j \neq \emptyset$, since $\operatorname{spt}v_h \subset 8Q$) we now can apply Lemma 5.4 to infer $|16Q \cap 16Q_i| \le c|16Q \cap 4 \cdot 16Q_j| \le c(n,m)|16Q \cap Q_j|$. This allows us to conclude

$$\sum_{j \in I(i)} ⨍_{16Q \cap Q_j} |v_h|\, dz \le c(n,m)\ ⨍_{16Q \cap 16Q_i} |v_h|\, dz, \tag{5.29}$$

which combined with the previous estimate for $|I^{(\ell)}|$ and (5.28) proves the bound in (5.24).

Now, we consider the the case when $i \in \Theta_2$. From Young's inequality we infer

$$⨍_{16Q \cap 16Q_i} |v_h|\, dz \le r_i^m \lambda + r_i^{-m}\lambda^{-1} ⨍_{16Q \cap 16Q_i} |v_h|^2\, dz.$$

Noting that $r_i^{-2m} \le 8^{2m}\lambda^{2-p}/s$ for $i \in \Theta_2$ by (5.11) and inserting this in the estimate (5.24), we deduce the bound in (5.27). Finally, if $\varrho \le r_i$, we have $|D^{k-\ell}\omega_j(\mathfrak{z})| \le c/r_j^{k-\ell} \le c/r_j^{k-\ell} \le c/\varrho^{k-\ell}$ (recall (5.8)). This together with the estimate (5.19) for $|D^\ell P_{v_h,j}(\mathfrak{x})|$ yields $|I^{(\ell)}| \le c\,\varrho^{m-k}\lambda$. Joining this with (5.28) proves (5.25) in the case $i \in \Theta_2$ and $\varrho \le r_i$.

We now turn our attention to the case when $i \in \Theta_1$. Here we use the fact that $\sum_{j \in I(i)} \omega_j \equiv 1$ on $\operatorname{spt}\omega_i$, so that $\sum_{j \in I(i)} D^{k-\ell}\omega_j \equiv 0$ for $0 \le \ell \le k-1$

on $\mathrm{spt}\omega_i$. Therefore we obtain for $0 \le \ell \le k-1$

$$\begin{aligned}|I^{(\ell)}| &= \Big|\sum_{j\in I(i)} D^{k-\ell}\omega_j(\mathfrak{z}) \odot \big(D^\ell P_{v_h,j}(\mathfrak{x}) - D^\ell P_{v_h,i}(\mathfrak{x})\big)\Big| \\ &\le \sum_{j\in I(i)} |D^{k-\ell}\omega_j(\mathfrak{z})|\,|D^\ell P_{v_h,j}(\mathfrak{x}) - D^\ell P_{v_h,i}(\mathfrak{x})| \\ &\le c\sum_{j\in I(i)} r_j^{-(k-\ell)} \min\{r_i^{m-\ell}, \varrho^{m-\ell}\}\,\lambda \le c\ \min\big\{r_i^{m-k}, \varrho^{m-k}\big\}\,\lambda.\end{aligned}$$

Here we have used in the third line the fact that $|D^{k-\ell}\omega_j| \le c/r_j^{k-\ell}$ and (5.21) from Lemma 5.12 (note that $\mathfrak{z} \in 16Q\cap\mathrm{spt}\omega_i \subset 16Q\cap 16Q_i$), while in the last line we used $r_j^{-1} \le 3r_i^{-1}$ from (5.8) and the fact that $r_i^{-(k-\ell)} \le \varrho^{-(k-\ell)}$ if the minimum is achieved for ϱ^{m-k}. To estimate the remaining term $I^{(k)}$ we recall that $0 \le \omega_j \le 1$ and we use again (5.19) to bound $|D^k P_{v_h,j}(\mathfrak{x})|$ by $c\,\varrho^{m-k}\lambda$, respectively the fact that $D^m P_{v_h,i} \equiv 0$, to find that $|I^{(k)}| \le c\,\varrho^{m-k}\lambda$. Joining the previous estimates for $I^{(\ell)}$, $0 \le \ell \le k$ with (5.28) we conclude the bound in (5.25) also for the remaining case $i \in \Theta_1$. This finishes the proof of (5.25). It now remains to show (5.26). For this we estimate $I^{(k)}$ from (5.28) in a different way. Using Lemma 5.12, i.e., estimate (5.21), Lemma 3.5 with $(B_i, 2B_i, P_{v_h,i})$ instead of $(\mathcal{O}_1, \mathcal{O}_2, P)$ (note that $\mathfrak{z} \in \mathrm{spt}\omega_i \subset 2Q_i$) and finally Young's inequality we obtain for $\varepsilon \ge 0$

$$\begin{aligned}|I^{(k)}| &\le \sum_{j\in I(i)} \omega_j(\mathfrak{z})\,|D^k P_{v_h,j}(\mathfrak{x}) - D^k P_{v_h,i}(\mathfrak{x})| + |D^k P_{v_h,i}(\mathfrak{x})| \\ &\le c\ r_i^{m-k}\,\lambda + c\ r_i^{-k} ⨍_{B_i} |P_{v_h,i}|\,dx \\ &\le c\ r_i^{m-k}\,\lambda + c\ r_i^{m-k}\Big(\frac{\lambda}{\varepsilon}\ + \frac{\varepsilon}{\lambda\, r_i^{2m}} ⨍_{B_i} |P_{v_h,i}|^2\,dx\Big).\end{aligned}$$

Joining this estimate for $I^{(k)}$ and the previous estimates for $I^{(\ell)}$, $0 \le \ell \le k-1$ with (5.28) we deduce the inequality in (5.26). Finally, we mention that the asserted dependencies of the constants from the indicated parameters follow the ones in Lemma 5.12. □

Lemma 5.15. *Suppose that $\lambda \ge c_E\lambda_1$ and Q_i is one of the Whitney-cylinders. Then, for all $\mathfrak{z} \in 16Q \cap \mathrm{spt}\omega_i$ we can bound the time-derivative $\partial_t w_h$ as follows:*

$$|\partial_t w_h(\mathfrak{z})| \le \tfrac{c}{\gamma r_i^{2m}}\ \Big(\min\{r_i^m\lambda, \sqrt{s\lambda^p}\} + ⨍_{16Q\cap 16Q_i} |v_h|\,dz\Big). \tag{5.30}$$

In the case $i \in \Theta_1$ there holds

$$|\partial_t w_h(\mathfrak{z})| \le \tfrac{c}{\gamma r_i^{2m}} \min\{r_i^m, \varrho^m\}\ \lambda. \tag{5.31}$$

In the case $i \in \Theta_2$ we have

$$|\partial_t w_h(\mathfrak{z})| \le \tfrac{c}{s}\ \varrho^m \lambda. \tag{5.32}$$

In any case the constant c depends only on n, N, m, L and κ.

Proof. To derive the estimate for the time derivatives $\partial_t w_h$ we argue similarly to the estimation of $D^k w_h$ from Lemma 5.14. Let $\mathfrak{z} = (\mathfrak{x}, \mathfrak{t}) \in 16Q \cap \mathrm{spt}\omega_i$ for some $i \in \mathbb{N}$. Recalling the definition of w_h from (5.15) and the fact that $\partial_t P_{v_h,j} = 0$ we compute

$$\partial_t w_h(\mathfrak{z}) = \sum_{j \in I(i)} \partial_t \omega_j(\mathfrak{z}) P_{v_h,j}(\mathfrak{x}). \tag{5.33}$$

We first prove (5.30). Using (5.20) from Lemma 5.12 to estimate $|P_{v_h,j}(\mathfrak{x})|$ and the fact that $|\partial_t \omega_j| \le c/(\gamma r_j^{2m}) \le c/(\gamma r_i^{2m})$ for $j \in I(i)$ (recall (5.8)) we obtain

$$\begin{aligned} |\partial_t w_h(\mathfrak{z})| &\le \sum_{j \in I(i)} |\partial_t \omega_j(\mathfrak{z})|\, |P_{v_h,j}(\mathfrak{x})| \\ &\le \tfrac{c}{\gamma r_i^{2m}} \Big(\min\{r_i^m \lambda, \sqrt{s\lambda^p}\} + \sum_{j \in I(i)} \fint_{16Q \cap Q_j} |v_h|\, dz \Big). \end{aligned}$$

Now, using the estimate (5.29) from the proof of Lemma 5.14 we can bound the mean value integrals of $|v_h|$ on $16Q \cap Q_j$ in terms of the mean value integral on $16Q \cap 16Q_i$. This yields the bound in (5.30).

Now we come to the proof of (5.32), where we consider $i \in \Theta_2$. Here we use (5.19) from Lemma 5.12 to bound $|P_{v_h,j}(\mathfrak{x})|$ by $c\,\varrho^m \lambda$ and the fact that $|\partial_t \omega_j| \le c/(\gamma r_i^{2m}) \le c/s$ for $i \in \Theta_2$ (see (5.11)). Inserting this in (5.33) we infer the bound in (5.32).

It now remains to prove (5.31), where we consider $i \in \Theta_1$. Noting that $\sum_{j \in I(i)} \partial_t \omega_j \equiv 0$ on $\mathrm{spt}\omega_i$ and $|\partial_t \omega_j| \le c/(\gamma r_i^{2m})$ for $j \in I(i)$ (see (5.8)) and using estimate (5.21) from Lemma 5.12 we conclude

$$|\partial_t w_h(\mathfrak{z})| = \Big| \sum_{j \in I(i)} \partial_t \omega_j(\mathfrak{z}) \left(P_{v_h,j}(\mathfrak{x}) - P_{v_h,i}(\mathfrak{x}) \right) \Big| \le \tfrac{c}{\gamma r_i^{2m}} \min\{r_i^m, \varrho^m\}\, \lambda.$$

This finally proves (5.31). As before, the asserted dependencies of the constants from the indicated parameters follow from those in Lemma 5.12. □

Remark 5.16. Since the bounds in Lemmas 5.13 - 5.15 are independent of h and $D^k w_h \to D^k w$ almost everywhere as $h \searrow 0$ for $0 \le k \le m$ the assertions also hold for w instead of w_h. Similarly, by the dominated convergence theorem we can pass to the limit $h \searrow 0$ in Lemma 5.12. This leads to similar bounds for $P_{v,i}$ in terms of $\varrho^{m-k}\lambda$ and $r_i^{m-k}\lambda + r_i^{-k} \fint_{16Q\cap Q_i} |v|\,dz$.

Corollary 5.17. *Let $\lambda \ge c_E\lambda_1$. Then, $\partial_t w_h$ and $D^k w_h$, $0 \le k \le m$ are locally bounded uniformly with respect to h on $(16B \times 2\mathcal{T})\backslash E(\lambda)$ and there exists a constant $c = c(n, N, m, L, \kappa)$ such that*

$$|D^k w_h(\mathfrak{z})| \le c\ \varrho^{m-k}\,\lambda\left(1 + (\lambda/\lambda_1)^{\frac{2-p}{2}}\right),$$

for all $\mathfrak{z} \in (16B \times 2\mathcal{T})\backslash E(\lambda)$ and $0 \le k \le m$ and

$$|\partial_t w_h(\mathfrak{z})| \le c\ \varrho^m\,\lambda\left(\tfrac{1}{\gamma r_i^{2m}} + \tfrac{1}{s}\right),$$

for all $\mathfrak{z} \in (16B \times 2\mathcal{T})\backslash E(\lambda)$ and $i \in \mathbb{N}$ such that $\mathfrak{z} \in Q_i$.

Proof. Let $\mathfrak{z} \in (16B\times 2\mathcal{T})\backslash E(\lambda)$. Then we find a Whitney cylinder Q_i with $\mathfrak{z} \in Q_i$. This means that $Q_i \cap 2\mathcal{H} \ne \emptyset$ and therefore $\mathrm{spt}\omega_i \cap 2\mathcal{H} \ne \emptyset$, so that $i \in \Theta = \Theta_1 \cup \Theta_2$ (see the definition of Θ in (5.9)). In the case $i \in \Theta_1$ or $i \in \Theta_2$ and $\varrho \le r_i$ the estimate for $|D^k w_h|$ directly follows from Lemma 5.14, i.e., from (5.25), while in the remaining case when $i \in \Theta_2$ and $r_i < \varrho$ we use (5.24) to infer

$$|D^k w_h(\mathfrak{z})| \le c\,r_i^{m-k}\left(\lambda + \frac{1}{r_i^m}\fint_{16Q\cap 16Q_i}|v_h|\,dz\right) \le c\,\varrho^{m-k}\left(\lambda + (\lambda/\lambda_1)^{\frac{2-p}{2}}\,\lambda\right),$$

where we have also used $r_i^{-m} \le 8^m(\lambda^{2-p}/s)^{\frac{1}{2}} = 8^m\varrho^{-m}(\lambda/\lambda_1)^{\frac{2-p}{2}}$ for $i \in \Theta_2$ from (5.11) and Corollary 5.10 to bound the integral $\fint_{16Q\cap 16Q_i}|v_h|\,dz$ from above by $c\,\varrho^m\lambda$ (note that $16Q_i \cap E(\lambda) \ne \emptyset$). This proves the assertion of the desired estimate for $|D^k w_h|$, $k = 1, \ldots, m$. The estimate for $|\partial_t w_h|$ directly follows from Lemma 5.15, i.e., estimate (5.31) and (5.32). $\Box$

Lemma 5.18. *Suppose that $\lambda \ge c_E\lambda_1$. Then there exists a constant $c = c(n, N, m, L, \kappa)$ such that*

$$\int_{16Q\setminus E(\lambda)} |w_h|^2\,dz \le c\ s\,\lambda^p\,|16Q \setminus E(\lambda)| + c\int_{8Q\setminus E(\lambda)}|v_h|^2\,dz.$$

Proof. Since $16Q \setminus E(\lambda)$ is covered by Whitney-cylinders Q_i, we can decompose the domain of integration as follows: $16Q \setminus E(\lambda) = \sum_i 16Q \cap Q_i$. Then on each of the sets $16Q \cap Q_i$ we use estimate (5.24) from Lemma 5.14 to bound $|w_h|$ in terms of $\sqrt{s\lambda^p} + \fint_{16Q\cap Q_i}|v_h|\,dz$. Finally, we in turn apply

Hölder's inequality, sum over $i \in \mathbb{N}$ and recall that $\operatorname{spt} v_h \subset 8Q$ to conclude the desired estimate. □

Lemma 5.19. *Suppose that $\lambda \geq c_E \lambda_1$ and $i \in \Theta_1$. Then, for $\vartheta \geq 1$ we have*

$$\fint_{16Q \cap Q_i} |v_h - w_h|^\vartheta \, dz \leq c\, r_i^{\vartheta m} \Big(\lambda^\vartheta + \delta_{\vartheta > \xi} \fint_{16Q \cap Q_i} |D^m v_h|^\vartheta \, dx \Big),$$

where c=$c(n, N, m, L, \kappa, \vartheta)$ and $\delta_{\vartheta > \xi} = 1$ if $\vartheta > \xi$ and $\delta_{\vartheta > \xi} = 0$ if $1 \leq \vartheta \leq \xi$.

Proof. For $z = (x, t) \in Q_i$ we infer from the definition of w_h and the fact that $\sum_{j \in I(i)} \omega_j(z) = 1$ that

$$\begin{aligned} |v_h(z) - w_h(z)| &= \Big| v_h(z) - \sum_{j \in I(i)} \omega_j(z) P_{v_h, j}(x) \Big| \\ &\leq |v_h(z) - P_{v_h, i}(x)| + \sum_{j \in I(i)} \omega_j(z) |P_{v_h, i}(x) - P_{v_h, j}(x)|. \end{aligned}$$

Now, we integrate with respect to z over $16Q \cap Q_i$. The first term on the right-hand side of the resulting inequality is estimated by an application of Lemma 5.11 with Q_i instead of $Q_{\mathfrak{z}}(r, \gamma r^{2m})$ (note that $Q_i \subset 4\mathcal{H}$ since $i \in \Theta_1$ and $4Q_i \cap E(\lambda) \neq \emptyset$), whereas for the remaining term we use the pointwise estimate (5.21) with $k = 0$ from Lemma 5.12. Proceeding this way we conclude the desired estimate. The asserted dependence of the constant c from the indicated parameters follows from those in Lemma 5.11 and 5.12. □

Lemma 5.20. *Suppose that $\lambda \geq c_E \lambda_1$. Then, for $1 \leq \vartheta \leq \xi$ we have*

$$\int_{(16B \times 2\mathcal{T}) \setminus E(\lambda)} |\partial_t w_h \cdot (v_h - w_h)|^\vartheta \, dz \leq c\, \lambda^{\vartheta p} \, |16Q \setminus E(\lambda)| + \frac{c}{s^\vartheta} \int_{8Q} |v_h|^{2\vartheta} \, dz,$$

where $c = c(n, N, m, L, \kappa, \vartheta)$.

Proof. The strategy of the proof is to decompose $(16B \times 2\mathcal{T}) \setminus E(\lambda)$ in Whitney-cylinders Q_i and to establish the estimate on each cylinder separately. Since the domain of integration is $(16B \times 2\mathcal{T}) \setminus E(\lambda)$ we only have to consider those Whitney-cylinders Q_i with $2\mathcal{H} \cap Q_i \neq \emptyset$, which means that $i \in \Theta = \Theta_1 \cup \Theta_2$ (recall the definition of Θ from (5.9)). We distinguish between $i \in \Theta_1$ and $i \in \Theta_2$.

In the case $i \in \Theta_1$ we first use (5.31), i.e., the sup-estimate for $\partial_t w_h$ on Q_i with $i \in \Theta_1$. Then we apply Lemma 5.19 to bound the remaining integral (note that we have assumed $\vartheta \leq \xi$). Recalling also the definition $\gamma = \lambda^{2-p}$

we obtain

$$\begin{aligned}\int_{16Q\cap Q_i} |\partial_t w_h \cdot (v_h - w_h)|^\vartheta \, dz &\le c \left(\frac{\lambda}{\gamma r_i^m}\right)^\vartheta \left(r_i^m \lambda\right)^\vartheta |16Q \cap Q_i| \\ &= c\, \lambda^{\vartheta p} |16Q \cap Q_i|.\end{aligned}$$

In the case $i \in \Theta_2$ we first exploit the bound (5.30) for $|\partial_t w_h|$ and Young's inequality to infer

$$\begin{aligned}&\int_{16Q\cap Q_i} |\partial_t w_h \cdot (v_h - w_h)|^\vartheta \, dz \\ &\qquad \le \frac{c}{(\gamma r_i^{2m})^\vartheta} \left(r_i^m \lambda + \fint_{16Q\cap 16Q_i} |v_h| \, dz\right)^\vartheta \int_{16Q\cap Q_i} |v_h|^\vartheta + |w_h|^\vartheta \, dz.\end{aligned}$$

To further estimate the right side we use the bound (5.24) from Lemma 5.14 for $|w_h|$ and Young's and Hölder's inequality, yielding that

$$\begin{aligned}&\int_{16Q\cap Q_i} |\partial_t w_h \cdot (v_h - w_h)|^\vartheta \, dz \\ &\qquad \le \frac{c}{(\gamma r_i^{2m})^\vartheta} \left(\left(r_i^m \lambda\right)^{2\vartheta} + \fint_{16Q\cap 16Q_i} |v_h|^{2\vartheta} \, dz\right) |16Q \cap Q_i| \\ &\qquad \le c\, \lambda^{\vartheta p} |16Q \cap Q_i| + \frac{c}{s^\vartheta} \int_{16Q\cap 16Q_i} |v_h|^{2\vartheta} \, dz.\end{aligned}$$

Here, in the last line we have also used $\gamma = \lambda^{2-p}$ and the fact that $(\gamma r_i^{2m})^{-1} \le c/s$ for $i \in \Theta_2$. We now sum the previous inequalities with respect to $i \in \Theta = \Theta_1 \cup \Theta_2$. Noting that $(16B \times 2\mathcal{T}) \setminus E(\lambda) \subset \bigcup_{i\in\Theta} Q_i$ (by construction), that locally in each point at most $c(n, m)$ of the Whitney-cylinders intersect, and finally that $\operatorname{spt} v_h \subset 8Q$, leads us to the desired estimate. □

The next lemma is a geometric statement that will be needed later in order to ensure that certain Whitney cylinders are from Θ_1.

Lemma 5.21. *Let $z \in 2\mathcal{H} \setminus E(\lambda)$ such that $d_\lambda(z, E(\lambda)) \le \frac{3}{8} \sqrt[2m]{s/\gamma}$. Then for all Whitney-cylinders $Q_i = Q_{z_i}(r_i, \gamma r_i^{2m})$ with $z \in Q_i$ we have $r_i \le \frac{1}{3} d_\lambda(z, E(\lambda))$ and $i \in \Theta_1$.*

Proof. Since $z \in Q_i \cap 2\mathcal{H} \subset \operatorname{spt}\omega_i \cap 2\mathcal{H}$ we have $i \in \Theta$. Moreover, because $z \in Q_i$ we infer from Lemma 5.2, (iii) that $r_i \le \frac{1}{3} d_\lambda(Q_i, E(\lambda)) \le \frac{1}{3} d_\lambda(z, E(\lambda))$, proving the first assertion of the lemma. Furthermore, together with our assumption on $d_\lambda(z, E(\lambda))$ we find $r_i \le \frac{1}{8} \sqrt[2m]{s/\gamma}$. To show that $i \in \Theta_1$ we have to ensure $16Q_i \subset 4\mathcal{H}$ (see (5.10)). The height of $16Q_i$ is estimated by

$$2\,\gamma(16 r_i)^{2m} = 2 \cdot 16^{2m}\, \gamma r_i^{2m} \le 2 \cdot 2^{2m}\, s \le 2^{2m}(2^{2m} - 1)\, s = 4^{2m} s - 2^{2m} s.$$

Together with the fact that $16Q_i \cap 2\mathcal{H} \neq \emptyset$ this implies $Q_i \subset 4\mathcal{H}$, so that $i \in \Theta_1$ (recall that $2\mathcal{H} = \mathcal{H}(2^{2m}s)$ and $4\mathcal{H} = \mathcal{H}(4^{2m}s)$). $\square$

Lemma 5.22. *Let $\lambda \geq c_E\lambda_1$. Then $D^k w_h$ is continuous on $16B \times 2\mathcal{T}$ for $0 \leq k \leq m-1$. Moreover, the functions $D^k w_h(\cdot, t)$ are locally Lipschitz-continuous on $16B$ for $0 \leq k \leq m-1$ and for all $t \in 2\mathcal{T}$. On compact subsets of $16B \times \{t\}$ the Lipschitz constant is independent of h and t.*

Proof. We first recall that the restriction of w_h to the set $E(\lambda)$ equals v_h, while the restriction to $\mathbb{R}^{n+1} \setminus E(\lambda)$ is constructed from a Whitney type decomposition by gluing together the local mean value polynomials of v_h on the Whitney cylinders by a partition of unity. On this bad set, i.e., on $(16B \times 2\mathcal{T}) \setminus E(\lambda)$, we have already established in Corollary 5.17 local L^∞-bounds for the space and time derivatives $D^k w_h$, respectively $\partial_t w_h$.

The main purpose of the lemma is to show that the two functions, i.e., $w_h|_{E(\lambda)}$ and $w_h|_{\mathbb{R}^{n+1}\setminus E(\lambda)}$, match along the boundary of $E(\lambda)$. To this end we first show that $D^k w_h$ is continuous on $(16B \times 2\mathcal{T})$ for $0 \leq k \leq m-1$. Having established the continuity we will prove that on each horizontal slice, i.e., for $t \in 2\mathcal{T}$, the function $D^k w_h(\cdot, t)$ is locally Lipschitz continuous.

Claim 1. For $0 \leq k \leq m-1$ the function

$$V_h^{(k)} \equiv \begin{cases} D^k v_h|_{E(\lambda)} & \text{on } \; E(\lambda) \\ D^k w_h|_{\mathbb{R}^{n+1}\setminus E(\lambda)} & \text{on } \; \mathbb{R}^{n+1} \setminus E(\lambda) \end{cases}$$

is continuous on $16B \times 2\mathcal{T}$.

Proof of Claim 1. We first note that $D^k w_h$ is of class C^∞ on the open set $\mathbb{R}^{n+1} \setminus E(\lambda)$ by construction. This follows directly from the definition of $w_h = \sum_{i=1}^\infty \omega_i P_{v_h,i}$ since ω_i is of class C^∞, $P_{v_h,i}$ are polynomials of degree $\leq m-1$ and the sum is locally finite (see (5.15)).

Now we consider $z_1 \in (16B \times 2\mathcal{T}) \cap E(\lambda)$. Then from the definition of v_h in (5.13) we see that $D^k v_h$ is continuous at z_1. To show the continuity of $V_h^{(k)}$ in z_1 we consider $V_h^{(k)}(z_1) - V_h^{(k)}(z_2)$ in the limit $16B \times 2\mathcal{T} \ni z_2 \to z_1$, which is equivalent to $d_\lambda(z_1, z_2) \to 0$ for fixed $\lambda > 0$.

In the case $z_2 \in E(\lambda)$ the definition of $V_h^{(k)}$ implies

$$|V_h^{(k)}(z_1) - V_h^{(k)}(z_2)| = |D^k v_h(z_1) - D^k v_h(z_2)|.$$

In the case $z_2 = (x_2, t_2) \in 16B \cap 2\mathcal{T} \setminus E(\lambda)$ we have $z_2 \in Q_i = Q_{z_i}(r_i, \gamma r_i)$ for some Whitney cylinder $Q_i \subset \mathbb{R}^{n+1} \setminus E(\lambda)$. Moreover, we assume that $d_\lambda(z_1, z_2) \leq \frac{3}{8} \min\{\varrho, \sqrt[2m]{s/\gamma}\}$, which is a possible assumption, since we consider $d_\lambda(z_1, z_2) \to 0$. This implies $d_\lambda(z_2, E(\lambda)) \leq d_\lambda(z_1, z_2) \leq \frac{3}{8} \sqrt[2m]{s/\gamma}$.

Therefore, we can apply Lemma 5.21 in the case $z = z_2$ to infer that $i \in \Theta_1$ and

$$r_i \le \tfrac{1}{3}\, d_\lambda(z_2, E(\lambda)) \le \tfrac{1}{3}\, d_\lambda(z_1, z_2) \le \tfrac{1}{8}\, \varrho. \tag{5.34}$$

At this stage we can assume $8Q \cap 16Q_i \neq \emptyset$, since otherwise for any $j \in I(i)$ we would have $P_{v_h,j} \equiv 0$ and $V_h^{(k)}(z_2) = 0 = D^k v_h(z_2)$. From (5.12) we then conclude $16Q_i \subset 16Q$ and hence we have $Q_j \subset 16Q_i \subset 16Q$ for all $j \in I(i)$ (recall (5.8)). From the local representation of w_h from (5.15) and the fact that $\sum_{j\in I(i)} \omega_j(z_2) = 1$ (note that $z_2 \in Q_i$) we infer:

$$\begin{aligned} V_h^{(k)}(z_1) - V_h^{(k)}(z_2) &= D^k v_h(z_1) - D^k\Big(\sum_{j\in I(i)} \omega_j P_{v_h,j}\Big)(z_2) \\ &= \sum_{j\in I(i)} \omega_j(z_2)\big(D^k v_h(z_1) - D^k P_{v_h,j}(x_2)\big) \\ &\quad - \sum_{\ell=0}^{k-1} \binom{k}{\ell} \sum_{j\in I(i)} D^{k-\ell}\omega_j(z_2) \odot D^\ell P_{v_h,j}(x_2) \\ &= I + II, \end{aligned} \tag{5.35}$$

with the obvious meaning of I and II.

To estimate I we use $0 \le \omega_j \le 1$ and then exploit the defining property of $D^k P_{v_h,j}$, i.e., $(D^k v_h)_{Q_j} = (D^k P_{v_h,j})_{B_j}$, to decompose the remaining term as follows:

$$\begin{aligned} |I| &\le c \sum_{j\in I(i)} |D^k v_h(z_1) - (D^k v_h)_{Q_j}| + c \sum_{j\in I(i)} |(D^k P_{v_h,j})_{B_j} - D^k P_{v_h,j}(x_2)| \\ &\le c \sup_{z\in 16Q_i} |D^k v_h(z_1) - D^k v_h(z)| + c \sum_{j\in I(i)} \sup_{x\in 16B_i} |D^k P_{v_h,j}(x) - D^k P_{v_h,j}(x_2)|. \end{aligned}$$

Here, in the second line we have taken into account that $Q_j \subset 16Q_i$, respectively $B_j \subset 16B_i$, for $j \in I(i)$. To estimate the difference of the polynomials we use Lemma 5.12, (5.19) and the fact that $\varrho \le 1$ to infer for $x \in 16B_i$, $x_2 \in B_i$ and $j \in I(i)$,

$$\begin{aligned} |D^k P_{v_h,j}(x) - D^k P_{v_h,j}(x_2)| &\le \sup_{y\in 16B_i} |D^{k+1} P_{v_h,j}(y)||x - x_2| \\ &\le c\, \varrho^{m-k-1} r_i\, \lambda \le c\, r_i\, \lambda, \end{aligned}$$

yielding that the second term on the right side in the estimate for $|I|$ is bounded by $c\, r_i \lambda$, with $c = c(n, N, m, L, \kappa)$.

Now we come to the estimate for II. Using the fact that $D^{k-\ell}(\sum_{j\in I(i)}\omega_j) \equiv 0$ for $0\le \ell\le k-1$ on Q_i and $z_2\in Q_i$ we can rewrite II as follows:

$$\begin{aligned}|II| &= \Big|\sum_{\ell=0}^{k-1}\binom{k}{\ell}\sum_{j\in I(i)} D^{k-\ell}\omega_j(z_2)\odot\big(D^\ell P_{v_h,j}(x_2)-D^\ell P_{v_h,i}(x_2)\big)\Big|\\ &\le c\sum_{\ell=0}^{k-1}\sum_{j\in I(i)} r_j^{-(k-\ell)}|D^\ell P_{v_h,j}(x_2)-D^\ell P_{v_h,i}(x_2)|\\ &\le c\sum_{\ell=0}^{k-1} r_i^{-(k-\ell)}\, r_i^{m-\ell}\,\lambda \le c\, r_i^{m-k}\,\lambda\le c\, r_i\,\lambda.\end{aligned}$$

Here we have also used $r_i\le 3r_j$ by (5.8) and (5.21) from Lemma 5.12 to estimate the difference $|D^\ell P_{v_h,j}(x_2)-D^\ell P_{v_h,i}(x_2)|$ from above by $c\,r_i^{m-\ell}\lambda$ (which is justified since $i\in\Theta_1$). Joining the previous estimates for I and II with (5.35) we finally arrive at

$$|V_h^{(k)}(z_1)-V_h^{(k)}(z_2)|\le c\sup_{z\in 16Q_i}|D^k v_h(z_1)-D^k v_h(z)| + c\,r_i\,\lambda.$$

To show the continuity of $V_h^{(k)}$ in z_1 we still have to ensure that the right-hand side of the previous inequality tends to zero as $z_2\to z_1$. From (5.34) we know that $r_i\searrow 0$ as $z_2\to z_1$. This also implies that $d_\lambda(z_1,z)$, with $z\in 16Q_i, z_1\in Q_i$ tends to zero. Therefore we can exploit the continuity of $D^k v_h$ to deduce that also the first term on the right-hand side tends to zero as $(16B\times 2\mathcal{T})\setminus E(\lambda)\ni z_2\to z_1$. This together with the estimate from the case $z_2\in(16B\times 2\mathcal{T})\cap E(\lambda)$ proves the continuity of $V_h^{(k)}$ at $z_1\in(16B\times 2\mathcal{T})\cap E(\lambda)$, and since z_1 was an arbitrary point in $(16B\times 2\mathcal{T})\cap E(\lambda)$ we deduce the continuity of $V_h^{(k)}$ on this set. Since $V_h^{(k)}$ is by construction C^∞ on $\mathbb{R}^{n+1}\setminus E(\lambda)$ the claim follows.

Claim 2. For all $t\in 2\mathcal{T}$ and for any $0\le k\le m-1$ the function $x\mapsto V_h^{(k)}(x,t)$ is locally Lipschitz-continuous on $16B$.

Proof of Claim 2. As mentioned above, on the open set $(16B\times 2\mathcal{T})\setminus E(\lambda)$ the function $V_h^{(k)}$ is of class C^∞ and from Corollary 5.17 we infer for all $x\in 16B\setminus E_t(\lambda)$ (recall the definition (5.7) of $E_t(\lambda)$)

$$|DV_h^{(k)}(x,t)| = |D^{k+1}w_h(x,t)|\le c\,\varrho^{m-k-1}\lambda\left(1+(\lambda/\lambda_1)^{\frac{2-p}{2}}\right),$$

so that $V_h^{(k)}$ is locally Lipschitz continuous on $16B\setminus E_t(\lambda)$.

Now, we consider $16B \cap E_t(\lambda)$, which we assume to be nonempty. Otherwise we have $16B \subset 16B \setminus E_t(\lambda)$ and therefore the argument from above shows that $V_h^{(k)}(\cdot,t)$ is Lipschitz continuous on $16B$ with Lipschitz constant $c\,\varrho^{m-k-1}\,\lambda\,(1+(\lambda/\lambda_1)^{(2-p)/2})$. We choose $x_1 \in 16B \cap E_t(\lambda)$ and consider $x_2 \in 16B$ with

$$|x_1 - x_2| \le \min\big\{\varrho,\ \sqrt[2m]{s/\gamma}\big\}. \tag{5.36}$$

We now distinguish the cases whether or not the point x_2 is contained in $E_t(\lambda)$.

In the case $x_2 \in E_t(\lambda)$ we set $r = 2|x_1 - x_2|$ and define $\widetilde{B}_0 \equiv B_{\frac{x_1+x_2}{2}}(r)$ and $\widetilde{Q}_0 \equiv Q_{(\frac{x_1+x_2}{2},t)}(r,\gamma r^{2m})$. We note that by the choice of x_2 in (5.36) we have $2\gamma r^{2m} = 2\gamma\, 2^{2m}|x_1-x_2|^{2m} \le 2\cdot 2^{2m}s \le 4^{2m}s - 2^{2m}s$. Together with $t \in 2\mathcal{T}$ this shows that $\widetilde{Q}_0 \subset 4\mathcal{H}$. Moreover, from (5.36) we infer $2r = 4|x_1 - x_2| \le 4\varrho$. Therefore, we can assume $\widetilde{Q}_0 \subset 16Q$; otherwise we would have $8Q \cap \widetilde{Q}_0 = \emptyset$ which means that $v_h \equiv 0$ on $\widetilde{Q}_0$, and since $(x_1,t),(x_2,t) \in \widetilde{Q}_0$ the Lipschitz estimate holds trivially. Since (x_1,t), $(x_2,t) \in E(\lambda)$ we have $V_h^{(k)}(x_1,t) = D^k v_h(x_1,t)$ and $V_h^{(k)}(x_2,t) = D^k v_h(x_2,t)$ and therefore

$$\begin{aligned}
|V_h^{(k)}(x_1,t) - V_h^{(k)}(x_2,t)| &\le |D^k v_h(x_1,t) - D^k P_{\widetilde{Q}_0}(x_1)| \\
&\quad + |D^k P_{\widetilde{Q}_0}(x_1) - D^k P_{\widetilde{Q}_0}(x_2)| + |D^k v_h(x_2,t) - D^k P_{\widetilde{Q}_0}(x_2)| \\
&= I + II + III,
\end{aligned} \tag{5.37}$$

where $P_{\widetilde{Q}_0}\colon \mathbb{R}^n \to \mathbb{R}^N$ denotes the mean value polynomial of v_h on $\widetilde{Q}_0$ of degree $\le m-1$ defined by $(\delta P_{\widetilde{Q}_0})_{\widetilde{B}_0} = (\delta v_h)_{\widetilde{Q}_0}$. To infer an estimate for I we consider the nested sequence of concentric cylinders $\widetilde{Q}_j \equiv \frac{1}{2^j}Q_{(x_1,t)}(r,\gamma r^{2m})$ with center (x_1,t). Moreover, we denote $\widetilde{B}_j \equiv \frac{1}{2^j}B_{x_1}(r)$. Hence, for any $i \in \mathbb{N}$ we can decompose I as follows:

$$\begin{aligned}
I &\le |D^k v_h(x_1,t) - (D^k v_h)_{\widetilde{Q}_i}| + |(D^k P_{\widetilde{Q}_i})_{\widetilde{B}_i} - D^k P_{\widetilde{Q}_i}(x_1)| \\
&\qquad + \sum_{j=0}^{i-1} |D^k P_{\widetilde{Q}_{j+1}}(x_1) - D^k P_{\widetilde{Q}_j}(x_1)| \\
&= I_1^{(i)} + I_2^{(i)} + I_3^{(i)},
\end{aligned} \tag{5.38}$$

where $P_{\widetilde{Q}_j}$ is defined by $(\delta P_{\widetilde{Q}_j})_{\widetilde{B}_j} = (\delta v_h)_{\widetilde{Q}_j}$ as usual and we have used $(D^k v_h)_{\widetilde{Q}_i} = (D^k P_{\widetilde{Q}_i})_{\widetilde{B}_i}$. Obviously, (note that $D^k v_h$ is continuous) we have $I_1^{(i)} \to 0$ as $i \to \infty$. Before we come to the estimate for the second term, i.e., $I_2^{(i)}$, we will bound the oscillation of the polynomials $D^k P_{\widetilde{Q}_i}$. To be precise, we use the Taylor expansion of $D^k P_{\widetilde{Q}_i}$ in x_1 and Lemma 3.6 in the case $P \equiv 0$ (note that $\operatorname{diam}(\widetilde{B}_i) = 2 \cdot 2^{-i} r$), to infer for $x \in \widetilde{B}_i$, $i \in \mathbb{N}_0$ that

$$\begin{aligned} |D^k P_{\widetilde{Q}_i}(x) - D^k P_{\widetilde{Q}_i}(x_1)| &\le \sum_{j=k+1}^{m-1} \frac{1}{(j-k)!} |x - x_1|^{j-k} \, |D^j P_{\widetilde{Q}_i}(x_1)| \\ &\le c \sum_{\ell=k+1}^{m-1} (2^{-i} r)^{\ell-k} \fint_{\widetilde{Q}_i} |D^\ell v_h| \, dz \le c \; 2^{-i} r \, \lambda. \end{aligned} \tag{5.39}$$

Here, in the last line we have used $2^{-i} r \le 1$ and Lemma 5.10 (recall that $\widetilde{Q}_i \subset 16Q$ and $\widetilde{Q}_i \cap E(\lambda) \neq \emptyset$ and $\varrho \le 1$) to bound the integrals $\fint_{\widetilde{Q}_i} |D^\ell v_h| \, dz$ by $c \, \varrho^{m-\ell} \lambda \le c \, \lambda$. Note that $c = c(n, N, m, L, \kappa)$. This immediately leads us to the following estimate for $I_2^{(i)}$ for all $i \in \mathbb{N}$:

$$I_2^{(i)} \le \fint_{\widetilde{B}_i} |D^k P_{\widetilde{Q}_i}(x) - D^k P_{\widetilde{Q}_i}(x_1)| \, dx \le c \; 2^{-i} r \, \lambda.$$

Now, we come to the estimate of $I_3^{(i)}$ for $i \in \mathbb{N}_0$. Applying Lemma 3.6 with $(\mathcal{O}_1, \mathcal{O}_2, \mathcal{T}, \; P_{\mathcal{O}_1 \times \mathcal{T}}, P)$ replaced by $(\widetilde{B}_{j+1}, \widetilde{B}_{j+1}, \widetilde{Q}_{j+1}, P_{\widetilde{Q}_{j+1}}, P_{\widetilde{Q}_j})$ we can estimate the difference of two subsequent polynomials as follows (note that $\operatorname{diam}(\widetilde{B}_{j+1}) = 2 \cdot 2^{-(j+1)} r$):

$$|D^k P_{\widetilde{Q}_{j+1}}(x_1) - D^k P_{\widetilde{Q}_j}(x_1)| \le c \sum_{\ell=k}^{m-1} \big(2^{-(j+1)} r\big)^{\ell-k} \fint_{\widetilde{Q}_j} |D^\ell (v_h - P_{\widetilde{Q}_j})| \, dz,$$

where $c = c(n, m)$. Note that we have $\widetilde{Q}_j \subset \widetilde{Q}_0 \subset 16Q \cap 4\mathcal{H}$ and $\widetilde{Q}_j \cap E(\lambda) \neq \emptyset$ for all $j \in \mathbb{N}_0$. Therefore, we can apply Lemma 5.11 with $(\widetilde{Q}_j, \ell, 1)$ instead of $(16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m}), k, \vartheta)$ for $j \in \mathbb{N}_0$, $0 \le \ell \le m-1$ to estimate the integrals on the right side from above in terms of $(2^{-j} r)^{m-\ell} \, \lambda$. Summing over $j = 0, \dots, i-1$ we therefore obtain

$$I_3^{(i)} \le c \sum_{j=0}^{i-1} \sum_{\ell=k}^{m-1} \big(2^{-(j+1)} r\big)^{\ell-k} \, (2^{-j} r)^{m-\ell} \, \lambda \le c \sum_{j=0}^{i-1} (2^{-j} r)^{m-k} \, \lambda \le c \; r^{m-k} \, \lambda,$$

where $c = c(n, N, m, L, \kappa)$ and $i \in \mathbb{N}$ is arbitrary. Since the previous bound is uniform with respect to $i \in \mathbb{N}$ we can pass to the limit $i \to \infty$ in (5.38) and obtain $I \le c \, \lambda |x_1 - x_2|^{m-k}$ (recall $r = 2|x_1 - x_2|$). The same argument for x_2 instead of x_1 yields a similar estimate for III, i.e., $III \le c \, \lambda |x_1 - x_2|^{m-k}$.

Finally we come to the estimate for II. Here, we use (5.39) with $i = 0$ and $x = x_2 \in \widetilde{B}_0$ to infer

$$II \le c\,\lambda\, r = c\,\lambda\,|x_1 - x_2|,$$

where $c = c(n, N, m, L, \kappa)$. Combining the previous estimates for I - III with (5.37) gives the desired Lipschitz bound

$$|V_h^{(k)}(x_1, t) - V_h^{(k)}(x_2, t)| \le c\,\lambda\,|x_1 - x_2| \tag{5.40}$$

for $x_1, x_2 \in E_t(\lambda)$ and $|x_1 - x_2|$ satisfying (5.36).

Finally, we consider the case $x_2 \in 16B \setminus E_t(\lambda)$. On the line segment connecting x_2 and x_1 (starting at x_2) we choose the "first point" x_E in $E_t(\lambda)$; this means that $x_E \in (x_2, x_1]$, $x_E \in E_t(\lambda)$ and $[x_2, x_E) \subset 16B \setminus E_t(\lambda)$. This is possible since $E(\lambda)$ is closed. Due to the continuity of $V_h^{(k)}(\cdot, t)$ on $16B$ there exists $\delta > 0$ such that $|V_h^{(k)}(x_E, t) - V_h^{(k)}(x, t)| < \lambda|x_1 - x_2|$ for any $x \in \mathbb{R}^n$ with $|x - x_E| < \delta$. We fix $\hat{x}$ in the open segment (x_2, x_E) with $|x_E - \hat{x}| < \delta$. By the choice of x_E we know $\hat{x} \in 16B \setminus E_t(\lambda)$ and

$$|V_h^{(k)}(x_E, t) - V_h^{(k)}(\hat{x}, t)| < \lambda\,|x_1 - x_2|.$$

The line connecting $\hat{x}$ with x_2 is completely contained in $16B \setminus E_t(\lambda)$. Therefore we can use the pointwise estimate for $D^{k+1} w_h$ from Corollary 5.17 as follows (note $V_h^{(k)} = D^k w_h$ on $16B \setminus E_t(\lambda)$):

$$\begin{aligned} |V_h^{(k)}(\hat{x}, t) - V_h^{(k)}(x_2, t)| &= \Big| \int_0^1 \frac{d}{d\sigma} D^k w_h(\sigma\hat{x} + (1-\sigma)x_2, t)\, d\sigma \Big| \\ &\le \int_0^1 |\langle D^{k+1} w_h(\sigma\hat{x} + (1-\sigma)x_2, t), \hat{x} - x_2\rangle|\, d\sigma \\ &\le c\ \varrho^{m-k-1} \left(1 + (\lambda/\lambda_1)^{\frac{2-p}{2}}\right) \lambda\,|\hat{x} - x_2|. \end{aligned}$$

At this stage it only remains to derive an estimate for the line connecting x_1 and x_E. But this can easily be inferred from the first case because now x_1, x_E are both contained in $16B \cap E(\lambda)$. Hence, we can use (5.40) for x_E instead of x_2 to infer

$$|V_h^{(k)}(x_1, t) - V_h^{(k)}(x_E, t)| \le c\,\lambda\,|x_1 - x_E|.$$

Joining the previous estimates we finally arrive at

$$|V_h^{(k)}(x_1, t) - V_h^{(k)}(x_2, t)| \le c(n, N, m, L)\ \left(1 + (\lambda/\lambda_1)^{\frac{2-p}{2}}\right) \lambda\,|x_1 - x_2|.$$

Together, both cases show that $V_h^{(k)}$ is locally Lipschitz continuous for $0 \le k \le m - 1$ on horizontal slices of $16B \times \{t\}$. Therefore, we know that

$V_h^{(k)}(\cdot,t)$ is weakly differentiable for all $t \in 2\mathcal{T}$. It now remains to show that $D^k w_h \equiv V_h^{(k)}$ almost everywhere on $16B \times 2\mathcal{T}$ for $0 \le k \le m$. We know that $V_h^{(0)} \equiv w_h$ from the definition of $V_h^{(0)}$. Now, let $\ell \in \{0,\dots,m-1\}$. Since $\mathbb{R}^{n+1} \setminus E(\lambda)$ is open, we have $DV_h^{(\ell)} = V_h^{(\ell+1)}$ on $\mathbb{R}^{n+1} \setminus E(\lambda)$. Moreover, we know that $V_h^{(\ell)}(\cdot,t)$ and $D^\ell v_h(\cdot,t)$ are weakly differentiable with respect to x for all $t \in 2\mathcal{T}$. Therefore $DV_h^{(\ell)}(x,t) = D^{\ell+1}v_h(x,t)$ for almost every $x \in \{y \in 16B : V_h^{(\ell)}(y,t) = D^\ell v_h(y,t)\}$. Since $V_h^{(\ell)} = D^\ell v_h$ on $E(\lambda)$ by definition, we therefore have $DV_h^{(\ell)} = D^{\ell+1}v_h$ almost everywhere on $(16B \times 2\mathcal{T}) \cap E(\lambda)$. Taking into account that $D^{\ell+1}v_h = V_h^{(\ell+1)}$ on $E(\lambda)$ by the definition of $V_h^{(\ell+1)}$ we conclude $DV_h^{(\ell)} = V_h^{(\ell+1)}$ almost everywhere on $(16B \times 2\mathcal{T}) \cap E(\lambda)$. Altogether, we deduce that $DV_h^{(\ell)} = V_h^{(\ell+1)}$ almost everywhere on $16B \times 2\mathcal{T}$ for all $0 \le \ell \le m-1$. Hence, for $0 \le k \le m$ we have $D^k w_h = D^k V_h^{(0)} = D^{k-1}V_h^{(1)} = \dots = V_h^{(k)}$ almost everywhere on $16B \times 2\mathcal{T}$. Thus, we conclude the claimed Lipschitz continuity of $D^k w_h$ from the Lipschitz continuity of $V_h^{(k)}$. □

Lemma 5.23. *Let $\lambda \ge c_E\lambda_1$. Then $|v_h - w_h|^2$ admits a weak derivative with respect to t on $16B \times 2\mathcal{T}$.*

Proof. For $f\colon \mathbb{R}^{n+1} \to \mathbb{R}$ and $\delta > 0$ we define the finite difference with respect to time by

$$\tau_\delta f(z) = \tau_\delta f(x,t) = f(x,t+\delta) - f(x,t).$$

In the following we will show for $1 < \vartheta < \min\{2,\xi\}$ (for the definition of ξ see (5.3)) that

$$\mathcal{J} \equiv \int_{16B\times\mathcal{T}(2^{2m}s-\delta)} \big|\tau_\delta|v_h - w_h|^2\big|^\vartheta \, dz \le K\ \delta^\vartheta \qquad \forall\, \delta \in (0, s/8^{2m}]. \quad (5.41)$$

The constant K will be independent of δ, but it will depend on h and λ. Therefore we are not allowed to pass to the limit $h \searrow 0$. Once (5.41) is established we can conclude that $\partial_t|v_h - w_h|^2$ exists (in the weak sense) on $16B \times 2\mathcal{T}$. This follows for instance from Lemma 7.24 in [19].

We now fix $1 < \vartheta < \min\{2,\xi\}$ and $\delta \in (0, s/8^{2m}]$ and for $k > 0$ we define

$$E_k \equiv \Big\{z \in 16B \times 2\mathcal{T} : d_\lambda(z, E(\lambda)) \le k \sqrt[2m]{\delta/\gamma}\Big\}$$

and decompose the integral in (5.41) as follows:

$$\mathcal{J} = \int_{(16B\times\mathcal{T}(2^{2m}s-\delta))\setminus E_2} |\dots|^\vartheta \, dz + \int_{(16B\times\mathcal{T}(2^{2m}s-\delta))\cap E_2} |\dots|^\vartheta \, dz$$

$$\equiv I+II, \tag{5.42}$$

with the obvious meaning of I and II.

We start with the estimate of I. First we remark that if $z=(x,t)\in(16B\times\mathcal{T}(2^{2m}s-\delta))\setminus E_2$, then $d_\lambda((x,t),E(\lambda))\geq 2\sqrt[2m]{\delta/\gamma}$, so that

$$\begin{aligned} d_\lambda\big((x,t+\delta),E(\lambda)\big) &\geq d_\lambda\big((x,t),E(\lambda)\big)-d_\lambda\big((x,t+\delta),(x,t)\big)\\ &\geq 2\sqrt[2m]{\delta/\gamma}-\sqrt[2m]{\delta/\gamma}=\sqrt[2m]{\delta/\gamma}, \end{aligned}$$

so that (x,t) and $(x,t+\delta)$ are both contained in $(16B\times 2\mathcal{T})\setminus E_1\subset(16B\times 2\mathcal{T})\setminus E(\lambda)$ (and also the segment $[(x,t),(x,t+\delta)]$). Since v_h and w_h are differentiable with respect to t on $(16B\times 2\mathcal{T})\setminus E(\lambda)$ the integral I can be estimated as follows:

$$\begin{aligned} I &\leq \delta^\vartheta\int_{(16B\times 2\mathcal{T})\setminus E(\lambda)}\big|\partial_t|v_h-w_h|^2\big|^\vartheta\,dz\\ &\leq 2\,\delta^\vartheta\int_{(16B\times 2\mathcal{T})\setminus E(\lambda)}|\partial_t v_h\cdot(v_h-w_h)|^\vartheta+|\partial_t w_h\cdot(v_h-w_h)|^\vartheta\,dz\\ &=2\,\delta^\vartheta\,(I_1+I_2), \end{aligned}$$

with the obvious meaning of I_1 and I_2. First, we observe that I_2 is bounded due to Lemma 5.20. To see that also I_1 is bounded we use in turn Young's inequality, the L^∞- bound for w_h from Lemma 5.13, and the fact that v_h is bounded (see Lemma 3.7, (i)) and differentiable with respect to t with derivative $\partial_t v_h=\eta\zeta_t[u-P_Q]_h+\eta\zeta\partial_t[u-P_Q]_h\in L^2(16Q)$ (see Lemma 3.7, (iii)). We obtain

$$\begin{aligned} I_1 &\leq \int_{(16B\times 2\mathcal{T})\setminus E(\lambda)}|\partial_t v_h|^2+|v_h-w_h|^{\frac{2}{2-\vartheta}}\,dz\\ &\leq c\int_{16Q}|\partial_t v_h|^2+|v_h|^{\frac{2}{2-\vartheta}}+(\varrho^m\lambda)^{\frac{2}{2-\vartheta}}\,dz, \end{aligned}$$

where $c=c(n,N,m,L,\kappa)$.

We now come to the estimate for II. To justify the following computations we first observe that $|D^k v_h|$ is bounded for $0\leq k\leq m$. This is due to the definition of v_h, i.e., $v_h=\eta\zeta[u-P_Q]_h$, the fact that $D^\ell[u-P_Q]_h=[D^\ell(u-P_Q)]_h$, $|D^\ell(u-P_Q)|\in L^1(16Q)$ for $1\leq\ell\leq m$ and Lemma 3.7, (i). Now, let $z=(x,t)\in(16B\times\mathcal{T}(2^{2m}s-\delta))\cap E_2$. Then $d_\lambda((x,t),E(\lambda))\leq 2\sqrt[2m]{\delta/\gamma}$ and therefore $d_\lambda((x,t+\delta),E(\lambda))\leq d_\lambda((x,t),E(\lambda))+d_\lambda((x,t+\delta),(x,t))\leq$

$3\sqrt[2m]{\delta/\gamma}$, so that $(x, t+\delta) \in E_3$. This implies

$$
\begin{aligned}
II &\le 2 \int_{(16B\times\mathcal{T}(2^{2m}s-\delta))\cap E_2} |(v_h - w_h)(x,t+\delta)|^{2\vartheta} + |(v_h - w_h)(x,t)|^{2\vartheta}\, dz \\
&\le 4 \int_{E_3\setminus E(\lambda)} |v_h - w_h|^{2\vartheta}\, dz.
\end{aligned}
$$

Here we have also used that we only have to take into account those $z \in E_3 \setminus E(\lambda)$, since $w_h \equiv v_h$ on $E(\lambda)$. For such z we have $z \in 16B \times 2\mathcal{T}$ and $d_\lambda(z, E(\lambda)) < 3\sqrt[2m]{\delta/\gamma}$ and $z \notin E(\lambda)$. Therefore $z \in Q_i = Q_{z_i}(r_i, \gamma r_i^{2m})$ for some $i \in \Theta$. More precisely we have $i \in \Theta_1$. This can be seen from Lemma 5.21. Indeed, we have $z \in 2\mathcal{H} \setminus E(\lambda)$ and since $\delta \le s/8^{2m}$ we also have $d_\lambda(z, E(\lambda)) < 3\sqrt[2m]{\delta/\gamma} \le \frac{3}{8}\sqrt[2m]{s/\gamma}$. Therefore we can apply Lemma 5.21 to infer that $i \in \Theta_1$ and $r_i \le \frac{1}{3} d_\lambda(z, E(\lambda)) < \sqrt[2m]{\delta/\gamma}$. This proves the claim and hence $E_3 \setminus E(\lambda)$ is covered by $\bigcup_{i\in\Theta_1,\, E_3\cap Q_i\neq\emptyset} Q_i$. This allows us to decompose the domain of integration in our estimate of II into Whitney-cylinders $Q_i \in \Theta_1$, fulfilling the assumptions of Lemma 5.19. We therefore obtain

$$
\begin{aligned}
II &\le 4 \sum_{i\in\Theta_1,\, E_3\cap Q_i\neq\emptyset} \int_{16Q\cap Q_i} |v_h - w_h|^{2\vartheta}\, dz \\
&\le c \sum_{i\in\Theta_1,\, E_3\cap Q_i\neq\emptyset} r_i^{2m\vartheta} \int_{16Q\cap Q_i} |D^m v_h|^{2\vartheta} + \lambda^{2\vartheta}\, dz.
\end{aligned}
$$

Recalling that $r_i < \sqrt[2m]{\delta/\gamma}$, the integral II can be further estimated by (recall that $\gamma = \lambda^{2-p}$)

$$
\begin{aligned}
II &\le c\, \delta^{\vartheta}\gamma^{-\vartheta} \sum_{i\in\mathbb{N},\, Q_i\cap E_3\neq\emptyset} \int_{16Q\cap Q_i} |D^m v_h|^{2\vartheta} + \lambda^{2\vartheta}\, dz \\
&\le c\, \delta^{\vartheta} \int_{16Q} \lambda^{(p-2)\vartheta} |D^m v_h|^{2\vartheta} + \lambda^{p\vartheta}\, dz,
\end{aligned}
$$

where $c = c(n, N, m, L, p, \kappa)$ and the integral is bounded since $|D^m v_h|$ is bounded as already mentioned. Finally, joining the estimates for I_1, I_2 and II with (5.42) we infer the desired estimate, i.e., (5.41) with a constant $K < \infty$ depending on $n, N, m, L, \vartheta, p, \varrho, s, \lambda, h$ and u, but independent of δ. This proves the lemma, i.e., $\partial_t |v_h - w_h|^2$ exists in the weak sense on $16B \times 2\mathcal{T}$. □

5.4. **Estimates on the "bad set".** In the proof of the Caccioppoli inequality we will decompose the domain of integration into the "good" and "bad set," where the maximal function of $|D^m u|^\xi$ is small, respectively large. Therefore we will need to control in particular certain integrals on this bad set, i.e., on $(16B \times 2\mathcal{T}) \setminus E(\lambda)$ respectively on horizontal slices of this set. Our aim in this chapter is to establish these estimates (see Lemma 5.25 and Lemma 5.27).

Lemma 5.24. *Let $\lambda \geq c_E \lambda_1$. Then there exists $c = c(n, N, m, L, \kappa)$ such that for any $0 \leq k \leq m$ and $i \in \Theta$ there hold the following estimates:*

$$\int_{8Q \cap \mathrm{spt}\omega_i} \big[(|D^m u| + b)^{p-1}\big]_h \, |D^k w_h| \, dz \leq c\, \varrho^{m-k} \Big[\lambda^p |16Q \cap 2Q_i| + \tfrac{\delta_2(i)}{s} \int_{8Q \cap 16Q_i} |v_h|^2 dz\Big] \tag{5.43}$$

and in the case $r_i < \varrho/4$ we have for any $0 < \varepsilon \leq 1$

$$\int_{8Q \cap \mathrm{spt}\omega_i} \big[(|D^m u| + b)^{p-1}\big]_h \, |D^k w_h| \, dz \tag{5.44}$$
$$\leq c\, r_i^{m-k} \Big[\tfrac{\lambda^p}{\varepsilon} \, |16Q \cap 2Q_i| + \tfrac{\delta_2(i)}{s} \int_{8Q \cap 16Q_i} |v_h|^2 \, dz + \varepsilon \, \delta_1(i) \int_{B_i} |P_{v_h, i}|^2 \, dx\Big].$$

Here $\delta_1(i) \equiv 1$ if $i \in \Theta_1$ and $\delta_1(i) \equiv 0$ otherwise and $\delta_2(i) \equiv 1$ if $i \in \Theta_2$ and $\delta_2(i) \equiv 0$ otherwise.

Proof. Let $Q_i = Q_{z_i}(r_i, \gamma r_i^{2m})$ be a Whitney-cylinder with $i \in \Theta$. From Lemma 5.2, (iii) we know that $4Q_i$ has non-empty intersection with $E(\lambda)$. Then the same is certainly also true for $8Q_i$, i.e., $8Q_i \cap E(\lambda) \neq \emptyset$. This allows us to apply Corollary 5.7 with $16Q \cap 2Q_i$ instead of $16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m})$. Noting that $8Q \cap \mathrm{spt}\omega_i \subset 16Q \cap 2Q_i$ and $p - 1 \leq \xi$ (see 5.3) we therefore obtain

$$\int_{8Q \cap \mathrm{spt}\omega_i} \big[(|D^m u| + b)^{p-1}\big]_h \, |D^k w_h| \, dz \leq c(n, m) \; |16Q \cap 2Q_i| \, \lambda^{p-1} \sup_{8Q \cap \mathrm{spt}\omega_i} |D^k w_h|. \tag{5.45}$$

We now distinguish several cases in order to find suitable bounds for $|D^k w_h|$ on $8Q \cap \mathrm{spt}\omega_i$.

In the cases $i \in \Theta_1$ or $i \in \Theta_2$ with $\varrho \leq r_i$ we infer from Lemma 5.14, i.e., from (5.25), that

$$\sup_{8Q \cap \mathrm{spt}\omega_i} |D^k w_h| \leq c \; \varrho^{m-k} \, \lambda,$$

yielding (5.43) in the cases $i \in \Theta_1$ or $i \in \Theta_2$ with $\varrho \le r_i$. In the case $i \in \Theta_2$ we obtain from (5.27)

$$\sup_{8Q\cap \mathrm{spt}\omega_i} |D^k w_h| \le c\, r_i^{m-k} \Big(\lambda + \tfrac{\lambda^{1-p}}{s} \fint_{16Q\cap 16Q_i} |v_h|^2\, dz\Big).$$

Inserting this in (5.45) we infer (5.43) in the remaining case $i \in \Theta_2$ and $r_i < \varrho$ (note $\mathrm{spt} v_h \subset 8Q$ and $|16Q\cap 16Q_i| \le c(n,m)|16Q\cap 2Q_i|$ by Lemma 5.4). This finishes the proof of (5.43). Moreover, we also deduce (5.44) in the case $i \in \Theta_2$. Therefore we are now left with the proof of (5.44) in the case $i \in \Theta_1$ (with $r_i < \varrho/4$). Since we can assume $8Q \cap 16Q_i \ne \emptyset$ (otherwise the right side in (5.45) is zero) we can use (5.12) to conclude that $16Q_i \subset 16Q$. From Lemma 5.14, i.e., from (5.26), we infer for $0 < \varepsilon \le 1$

$$\sup_{8Q\cap \mathrm{spt}\omega_i} |D^k w_h| \le c\, r_i^{m-k} \Big(\tfrac{\lambda}{\varepsilon} + \tfrac{\varepsilon}{\lambda r_i^{2m}} \fint_{B_i} |P_{v_h,i}|^2\, dx\Big).$$

Joining this with (5.45) (note $|16Q\cap 2Q_i| = |2Q_i| = c\,\gamma r_i^{2m}|B_i|$ in the present situation and $\gamma = \lambda^{2-p}$) we infer (5.44) in the remaining case $i \in \Theta_1$ and $r_i < \varrho/4$. This finishes the proof of the lemma. □

Lemma 5.25. *Let $\lambda \ge c_E\lambda_1$. Then there exists $c = c(n,N,m,L,\kappa)$ such that for any $0 \le k \le m$ there holds*

$$\int_{(8B\times 2\mathcal{T})\setminus E(\lambda)} \big[(|D^m u| + b)^{p-1}\big]_h |D^k w_h|\, dz \le c\,\varrho^{m-k}\Big(\lambda^p\,|16Q\setminus E(\lambda)| + \frac{1}{s}\int_{8Q} |v_h|^2\, dz\Big).$$

Proof. The strategy of the proof is to decompose $(8B\times 2\mathcal{T})\setminus E(\lambda)$ in Whitney-cylinders Q_i and to establish the estimate on each cylinder separately. Since the domain of integration is $(8B\times 2\mathcal{T})\setminus E(\lambda)$ we only have to consider those Whitney-cylinders Q_i with $2\mathcal{H}\cap Q_i \ne \emptyset$, which means (recall the definition of Θ from (5.9)) that $i \in \Theta$ and therefore Lemma 5.24 is applicable. At this stage we recall that $(8B\times 2\mathcal{T})\setminus E(\lambda) \subset \bigcup_{i\in\Theta} Q_i$ (by construction) and that locally in each point at most $c(n,m)$ of the Whitney-cylinders intersect. Moreover, from Lemma 5.2, (iii) we know that $2Q_i \subset 16Q\setminus E(\lambda)$. Therefore, using Lemma 5.24, i.e., estimate (5.43), for each Whitney-cylinder Q_i with $i \in \Theta$ and then summing with respect to i yields the desired estimate with a constant depending only on n, N, m, L, κ. □

As already mentioned, in the proof of the Caccioppoli type inequality we will have to control certain integrals on the "bad set" $(16B\times 2\mathcal{T})\setminus E(\lambda)$.

For this we will use the estimates from the last lemma. They are a direct consequence of the properties of the test-function w_h, respectively w, that we obtained so far in Lemma 5.14. But we also have to control certain integrals on horizontal slices $16B \setminus E_t(\lambda)$, $t \in 2\mathcal{T}$. For this we will need some additional information, which can be achieved by testing the parabolic system with the test-function w on a single Whitney-cylinder Q_i, $i \in \Theta$. To be precise, we use $\eta\zeta\omega_i$ as a cut-off function, and test the system with $(\eta\zeta\omega_i)w$. The global test-function w can later be recovered by summation over all Whitney-cylinders.

Lemma 5.26. *Let $\lambda \geq c_E\lambda_1$. Then for any $i \in \Theta_1$, $0 < \varepsilon \leq 1$ and almost every $t \in 2\mathcal{T}$ we have*

$$\Big|\int_{16B}\big((v-P_{v,i})\cdot w\,\omega_i\big)(\cdot,t)\,dx\Big| \leq \tfrac{c}{\varepsilon}\,\lambda^p\,|16Q\cap 2Q_i|+\varepsilon\,\delta_1(i)\int_{B_i}|P_{v,i}|^2dx, \quad (5.46)$$

where $\delta_1(i) \equiv 1$ if $r_i < \varrho/4$ and $\delta_1(i) \equiv 0$ otherwise. When $i \in \Theta_2$ then for almost every $t \in 2\mathcal{T}$ there holds

$$\Big|\int_{16B}\big(v\cdot w\,\omega_i\big)(\cdot,t)\,dx\Big| \leq c\,\lambda^p\,|16Q\cap 2Q_i| + \frac{c}{s}\int_{8Q\cap 16Q_i}|u-P_Q|^2\,dz. \quad (5.47)$$

In any case the constant c depends only on n, N, m, L and κ.

Proof. Let $t \in 2\mathcal{T}$ and $i \in \Theta$. In the weak formulation of our parabolic system in the form (3.2) we choose as test-function $\varphi = \eta\zeta\omega_i w$, which is admissible, since $w = \sum_{j\in I(i)}\omega_j P_{v,j}$ is a finite sum of C^∞-functions on $\mathrm{spt}\omega_i$. Integrating the resulting equation with respect to $\tau \in (t_i-\gamma(2r_i)^{2m}, t)$ we get

$$\int_{t_i-\gamma(2r_i)^{2m}}^{t}\int_{\mathbb{R}^n}\Big[\partial_t[u]_h\cdot\eta\zeta\omega_i w + \big\langle[A(\cdot,D^m u)]_h, D^m(\eta\omega_i w)\zeta\big\rangle + \big\langle[B(\cdot,D^m u)]_h, \delta(\eta\omega_i w)\zeta\big\rangle\Big]\,dxd\tau = 0.$$

Now, let $P\colon \mathbb{R}^n \to \mathbb{R}^N$ be a polynomial of degree $\leq m-1$. We do not further specify P at this stage, since it will be chosen later in two different ways. Noting that $\partial_t[P]_h \equiv 0$, and $\omega_i(\cdot,t_i-\gamma(2r_i)^{2m}) \equiv 0$ since $\mathrm{spt}\omega_i \subset 2Q_i$ by construction we infer with integration by parts for almost every $t \in 2\mathcal{T}$ that

$$\begin{aligned}&\int_{t_i-\gamma(2r_i)^{2m}}^{t}\int_{\mathbb{R}^n}\partial_t[u]_h\cdot\eta\zeta\omega_i w\,dx\,d\tau\\ &= \int_{t_i-\gamma(2r_i)^{2m}}^{t}\int_{\mathbb{R}^n}\partial_t[u-P]_h\cdot\eta\zeta\omega_i w\,dx\,d\tau\end{aligned}$$

$$= \int_{\mathbb{R}^n} \big([u-P]_h \cdot \eta\zeta\omega_i w\big)(\cdot,t)\,dx - \int_{t_i-\gamma(2r_i)^{2m}}^{t} \int_{\mathbb{R}^n} [u-P]_h \cdot \partial_t(\zeta\omega_i w)\eta\,dxd\tau$$
$$\to \int_{\mathbb{R}^n} \big((u-P) \cdot \eta\zeta\omega_i w\big)(\cdot,t)\,dx - \int_{t_i-\gamma(2r_i)^{2m}}^{t} \int_{\mathbb{R}^n} (u-P) \cdot \partial_t(\zeta\omega_i w)\eta\,dxd\tau$$

in the limit $h \searrow 0$. We now insert this in the previous equation and pass to the limit $h \searrow 0$ also in the remaining terms. Then, we use the growth conditions (2.3) and (2.4) imposed on A and B and recall that $\mathrm{spt}(\eta\zeta) \subset 8Q$ by construction to arrive at

$$\begin{aligned}
&\Big|\int_{16B} \big((u-P) \cdot \eta\zeta\omega_i w\big)(\cdot,t)\,dx\Big| \\
&\le c\,L \int_{8Q} \big(|D^m u|+b\big)^{p-1}\big(|D^m(\eta\omega_i w)| + |\delta(\eta\omega_i w)|\big)\,\zeta\,dz \\
&\quad + c\int_{8Q} |u-P|\,|\partial_t(\zeta\omega_i w)|\eta\,dz = c(p,L)\ (I+II),
\end{aligned} \tag{5.48}$$

with the obvious meaning of I and II. Since $\mathrm{spt}(\eta\zeta\omega_i) \subset 8Q \cap 2Q_i$ we have to distinguish different cases depending on the smaller scale in space respectively time direction, i.e., if r_i is small respectively large compared to ϱ and γr_i^{2m} is small respectively large compared to s.

We start with an estimate for I. Recalling that $|D^\ell \eta| \le c\varrho^{-\ell}$ and $|D^\ell \omega_i| \le c r_i^{-\ell}$ for $0 \le \ell \le m$ we find for $0 \le k \le m$

$$\begin{aligned}
|D^k(\eta\omega_i w)(z)| &= \Big|\sum_{\ell=0}^{k} \binom{k}{\ell} D^{k-\ell}(\eta\omega_i)(z) \odot D^\ell w(z)\Big| \\
&\le c\sum_{\ell=0}^{k} \max\{\varrho^{-(k-\ell)}, r_i^{-(k-\ell)}\}|D^\ell w(z)|.
\end{aligned}$$

Since $\varrho \le 1$ we have $\max\{\varrho^{-(k-\ell)}, r_i^{-(k-\ell)}\} \le \max\{\varrho^{-(m-\ell)}, r_i^{-(m-\ell)}\}$ yielding for I that

$$I \le c\sum_{k=0}^{m} \max\big\{\varrho^{-(m-k)}, r_i^{-(m-k)}\big\} \int_{8Q\cap \mathrm{spt}\omega_i} (|D^m u| + b)^{p-1}|D^k w|\,dz. \tag{5.49}$$

We now come to the **proof of (5.47)**, where we consider $i \in \Theta_2$. Here, we choose in (5.48) the polynomial $P \equiv P_Q$. To estimate the remaining integrals in (5.49) we apply Lemma 5.24. To be precise, in the case $r_i \ge \varrho/4$ we use (5.43), while in the case $r_i < \varrho/4$ we use (5.44) (note that $\delta_1(i) = 0$

and we can choose $\varepsilon = 1$ and pass to the limit $h \searrow 0$) to infer

$$I \le c\,\lambda^p\,|16Q \cap 2Q_i| + \frac{c}{s}\int_{8Q\cap 16Q_i} |u - P_Q|^2\,dz.$$

Here we have also used that $|v| = \eta\zeta|u - P_Q| \le |u - P_Q|$. Therefore it remains to estimate II from (5.48). In order to obtain such an estimate we first consider $\partial_t(\zeta\omega_i w)$. Using the fact that $|\partial_t\zeta| \le c/s$ and $|\partial_t\omega_j| \le c/(\gamma r_j^{2m}) \le c/(\gamma r_i^{2m}) \le c/s$ for $j \in I(i)$ with $i \in \Theta_2$ (this follows from (5.8) and (5.11)) and (5.24) from Lemma 5.14 to bound $|w|$ and (5.30) from Lemma 5.15 to bound $|\partial_t w|$ and Remark 5.16 which ensures that the lemma can be applied with $h = 0$ we find on $16Q \cap \mathrm{spt}\omega_i$ that

$$|\partial_t(\zeta\omega_i w)| = |\partial_t(\zeta\omega_i)w + (\zeta\omega_i)\partial_t w| \le \frac{c}{s}\,\Big(\sqrt{s\lambda^p} + \fint_{16Q\cap 16Q_i} |v|\,dz\Big).$$

Next, we recall that $\mathrm{spt}\omega_i \subset 2Q_i$, $\mathrm{spt}v \subset 8Q$ and $|v| \le |u - P_Q|$ and use Young's inequality to infer that

$$\begin{aligned} II &\le \frac{c}{s}\int_{8Q\cap 2Q_i} |u - P_Q|\,dz \cdot \Big(\sqrt{s\lambda^p} + \fint_{16Q\cap 16Q_i} |v|\,dz\Big) \\ &\le c\,\lambda^p\,|16Q \cap 2Q_i| + \frac{c}{s}\int_{8Q\cap 16Q_i} |u - P_Q|^2\,dz. \end{aligned}$$

Joining the previous estimates for I and II with (5.48) and recalling the definition $v = (u - P_Q)\eta\zeta$ we conclude the asserted estimate (5.47).

We now come to the **proof of (5.46)**, where we consider $i \in \Theta_1$. From (5.49) and Lemma 5.24, i.e., from (5.43) in the case $\varrho \le 4r_i$ and (5.44) in the case $r_i < \varrho/4$, we deduce

$$I \le c\,\Big(\tfrac{\lambda^p}{\varepsilon}\,|16Q \cap 2Q_i| + \varepsilon\,\delta_1(i)\int_{B_i} |P_{v,i}|^2\,dx\Big),$$

where $\varepsilon \in (0, 1]$ and $\delta(i) = 1$ in the case $r_i < \varrho/4$ and $\delta(i) = 0$ otherwise. We now turn our attention to the estimation of II and start with the case $r_i \ge \varrho/4$. Here we choose once again $P \equiv P_Q$ in (5.48). We first observe that $\zeta \equiv 1$ on $\mathrm{spt}\omega_i$ since $i \in \Theta_1$. Then we use $0 \le \omega_i \le 1$, $|\partial_t\omega_i| \le c/(\gamma r_i^{2m})$ and (5.13) (i.e., the L^∞- bound for w) and (5.31) (the L^∞- bound for $\partial_t w$ on $16Q \cap \mathrm{spt}\omega_i$) to bound $|\partial_t(\omega_i w)|$ by $c\,\lambda^{p-1}/\varrho^m$ (here we again note that $\gamma = \lambda^{2-p}$). This gives

$$II \le \int_{8Q} |u - P_Q||\partial_t(\omega_i w)|\,dz \le \tfrac{c\,\lambda^{p-1}}{\varrho^m}\int_{8Q\cap 2Q_i} |u - P_Q|\,dz \le c\,\lambda^p|16Q \cap 2Q_i|.$$

Here, for the estimate of the remaining integral we have applied Lemma 5.9 with $(16Q \cap 2Q_i, 0, 1)$ instead of $(16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m}), k, \vartheta)$. This is possible since we can assume $8Q \cap 2Q_i \neq \emptyset$ (otherwise the integral is zero) and $4Q_i \cap E(\lambda) \neq \emptyset$ by construction of the Whitney-cylinders. At this stage we can argue similarly as in the case $i \in \Theta_2$. We first note that $v = \eta(u - P_Q)$ in $2\mathcal{H}$ and then use the previous estimates for I and II to obtain from (5.48) for the case $i \in \Theta_1$ and $r_i \geq \varrho/4$:

$$\Big| \int_{16B} \big(v \cdot \omega_i w\big)(\cdot, t)\, dx \Big| = \Big| \int_{16B} \big((u - P_Q) \cdot \eta\omega_i w\big)(\cdot, t)\, dx \Big| \leq c\, \lambda^p\, |16Q \cap 2Q_i|.$$

Now, using $0 \leq \omega_i \leq 1$, the L^∞- bound for w from Lemma 5.13, (5.19) from Lemma 5.12, the fact that $Q_i \subset 4\mathcal{H}$ (since $i \in \Theta_1$) and therefore $|16B \cap 2B_i| = |16Q \cap 2Q_i|/|2\mathcal{T}_i| = |16Q \cap 2Q_i|/(2\gamma(2r_i)^{2m}) \leq 2^{2m}\varrho^{-2m}\lambda^{p-2}|16Q \cap 2Q_i|$ we infer

$$\int_{16B} \big(|P_{v,i}|\, \omega_i\, |w|\big)(\cdot, t)|\, dx \leq c\ \varrho^{2m}\, \lambda^2\, |16B \cap 2B_i| \leq c\ \lambda^p\, |16Q \cap 2Q_i|.$$

Joining this with the previous estimate we conclude (5.46) in the case $i \in \Theta_1$, $r_i \geq \varrho/4$.

We now come to the remaining case $i \in \Theta_1$, $r_i < \varrho/4$. Since we can assume $8Q \cap 2Q_i \neq \emptyset$ (otherwise this would imply $8Q \cap 2\mathrm{spt}\omega_i = \emptyset$ and we would have $v \equiv 0$ on $\mathrm{spt}\omega_i \subset 2Q_i$ and $P_{v,i} = 0$ and the asserted estimate holds trivially), we infer from (5.12) that $16Q_i \subset 16Q$. In the present case we choose in (5.48) the polynomial $P = P_{16Q_i}$, where P_{16Q_i} denotes the mean value polynomial of order $m-1$ defined by $(\delta P_{16Q_i})_{16B_i} = (\delta u)_{16Q_i}$. This choice will enable us to avoid a factor $(\varrho/r_i)^{2m}$ in the estimate for II. The precise argument is as follows: We first use $0 \leq \omega_i \leq 1$, $|\partial_t \omega_i| \leq c/(\gamma r_i^{2m})$, (5.31) and (5.26) for the case $k = 0$ to obtain for $0 < \varepsilon \leq 1$ the following pointwise bound on $\mathrm{spt}\omega_i$:

$$\begin{aligned} |\partial_t(\omega_i w)| &\leq \frac{c}{\gamma r_i^m} \Big(\frac{\lambda}{\varepsilon} + \frac{\varepsilon}{\lambda\, r_i^{2m}} \ -\!\!\!\!\!\!\int_{B_i} |P_{v,i}|^2\, dx \Big) \\ &= \frac{c}{r_i^m} \Big(\frac{\lambda^{p-1}}{\varepsilon\, r_i^m} + \frac{\varepsilon}{\lambda\, |Q_i|} \int_{B_i} |P_{v,i}|^2\, dx \Big). \end{aligned} \tag{5.50}$$

Since $16Q_i \subset 16Q$ and $16Q_i \cap E(\lambda) \neq \emptyset$ and $8Q \cap 16Q_i \neq \emptyset$, we can apply Lemma 5.8 with $(16Q_i, 1)$ instead of $(16Q \cap Q_{\mathfrak{z}}(r, \gamma r^{2m}), \vartheta)$ to obtain for $0 \leq k \leq m-1$

$$\int_{16Q_i} |D^k(u - P_{16Q_i})|\, dz \leq c\ (16r_i)^{m-k}\, \lambda\, |16Q_i| \leq c\ r_i^{m-k}\, \lambda\, |Q_i|. \tag{5.51}$$

To estimate II in (5.48) we note that $\zeta \equiv 1$ on $2Q_i \subset 4\mathcal{H}$ (by $i \in \Theta_1$). Moreover $\mathrm{spt}\omega_i \subset 2Q_i$. Taking this into account and using (5.50) and (5.51) with $k = 0$ we infer for any $0 < \varepsilon \le 1$ that

$$II = \int_{2Q_i} |u - P_{16Q_i}|\,|\partial_t(\omega_i w)|\eta\,dz \le \frac{c}{\varepsilon}\,\lambda^p\,|Q_i| + c\,\varepsilon\int_{B_i} |P_{v,i}|^2\,dx.$$

Inserting this and the estimate for I in (5.48) (note that we have chosen $P \equiv P_{16Q_i}$ and $\zeta \equiv 1$ on $\mathrm{spt}\omega_i$) we arrive at

$$\Big|\int_{16B} \big((u - P_{16Q_i}) \cdot \eta\omega_i w\big)(\cdot, t)\,dx\Big| \le \frac{c}{\varepsilon}\,\lambda^p\,|Q_i| + \varepsilon\int_{B_i} |P_{v,i}|^2\,dx \qquad (5.52)$$

for all $\varepsilon \in (0, 1]$. To be correct, we have to replace ε by ε/c in the argument above. However, the left-hand side of the previous estimate is not exactly the left-hand side of the assertion of the lemma. In order to control the difference we note that

$$\begin{aligned}(v - P_{v,i}) \cdot \omega_i w &= \big((u - P_Q)\eta - P_{v,i}\big) \cdot \omega_i w\\ &= (u - P_{16Q_i}) \cdot \eta\omega_i w + \big((P_{16Q_i} - P_Q)\eta - P_{v,i}\big) \cdot \omega_i w,\end{aligned}$$

where we have used $\zeta \equiv 1$ on $\mathrm{spt}\omega_i$. Therefore, we need to estimate $((P_{16Q_i} - P_Q)\eta - P_{v,i})(x)$ for x with $z = (x, t) \in \mathrm{spt}\omega_i \subset 2Q_i$. This can be achieved by using the mean value polynomial as follows: With the abbreviation $\mathcal{P}(x) = ((P_{16Q_i} - P_Q)\eta - P_{v,i})(x)$ we find for $x \in 2B_i = B_{x_i}(2r_i)$ that

$$\begin{aligned}|\mathcal{P}(x)| &\le |\mathcal{P}(x) - (\mathcal{P})_{B_i}| + |(\mathcal{P})_{B_i}| \le 4r_i \sup_{2B_i} |D\mathcal{P}| + |(\mathcal{P})_{B_i}|\\ &\le (4r_i)^2 \sup_{2B_i} |D^2\mathcal{P}| + |(\mathcal{P})_{B_i}| + 4r_i\,|(D\mathcal{P})_{B_i}|\\ &\ \ \vdots\\ &\le (4r_i)^m \sup_{2B_i} |D^m\mathcal{P}| + c\sum_{k=0}^{m-1} r_i^k\,|(D^k\mathcal{P})_{B_i}| = J_1 + J_2,\end{aligned}$$

with the obvious meaning of J_1 and J_2. To estimate J_1 we first observe for $x \in 2B_i$ that

$$|D^m\big((P_{16Q_i} - P_Q)\eta\big)(x)| \le c\sum_{k=0}^{m-1} \varrho^{-(m-k)}|D^k(P_{16Q_i} - P_Q)(x)|.$$

In order to control $|D^k(P_{16Q_i} - P_Q)|$ we apply Lemma 3.6 with $(16B_i, 16B_i, 16\mathcal{T}_i, P_{16Q_i}, P_Q)$ instead of $(\mathcal{O}_1, \mathcal{O}_2, \mathcal{T}, P_{\mathcal{O}_1\times\mathcal{T}}, P)$ to infer for $0 \le k \le m - 1$

and $x \in 2B_i$

$$|D^k(P_{16Q_i} - P_Q)(x)| \le c \sum_{j=k}^{m-1} r_i^{j-k} \fint_{16Q_i} |D^j(u - P_Q)|\,dz \le c \sum_{j=k}^{m-1} r_i^{j-k} \varrho^{m-j}\lambda$$
$$\le \varrho^{m-k}\lambda.$$

Here we have also applied Lemma 5.9, which is justified since $16Q_i \cap E(\lambda) \neq \emptyset$ by construction, $16Q_i \subset 16Q$ and we can assume $8Q \cap 16Q_i \neq \emptyset$ (otherwise we have $P_{v,i} \equiv 0$ and $v \equiv 0$ on $\mathrm{spt}\omega_i$). Inserting this above and noting that $D^m\mathcal{P} = D^m((P_{16Q_i} - P_Q)\eta)$ since $D^m P_{v,i} = 0$ we find

$$J_1 = (4r_i)^m \sup_{2B_i} |D^m\big((P_{16Q_i} - P_Q)\eta\big)| \le c\, r_i^m\, \lambda.$$

Now we start considering J_2. Recalling that $(D^k P_{v,i})_{B_i} = (D^k v)_{Q_i}$ and $v = (u - P_Q)\eta$ on $16Q_i$ we see that $(D^k\mathcal{P})_{B_i} = (D^k((P_{16Q_i} - P_Q)\eta - v))_{B_i} = (D^k((P_{16Q_i} - u)\eta))_{B_i}$ Using this information, the fact that $|D^{k-\ell}\eta| \le c/\varrho^{k-\ell} \le c/r_i^{k-\ell}$ and (5.51) we obtain

$$\begin{aligned} J_2 &\le c \sum_{k=0}^{m-1} r_i^k \fint_{Q_i} |D^k\big((u - P_{16Q_i})\eta\big)|\,dz \\ &\le c \sum_{k=0}^{m-1} r_i^k \sum_{\ell=0}^{k} \fint_{Q_i} |D^{k-\ell}\eta||D^\ell(u - P_{16Q_i})|\,dz \\ &\le c \sum_{k=0}^{m-1} \sum_{\ell=0}^{k} r_i^\ell \fint_{16Q_i} |D^\ell(u - P_{16Q_i})|\,dz \le c\, r_i^m\, \lambda. \end{aligned}$$

Together the estimates for J_1 and J_2 yield $|\mathcal{P}(x)| \le c\, r_i^m\, \lambda$ for $x \in 2B_i$, which combined with the facts that $\mathrm{spt}\omega_i(\cdot, t) \subset 2Q_i$, $0 \le \omega_i \le c$ and the bound for $|w|$ from Lemma 5.14, (5.26) leads us to

$$\int_{16B} \big|\big((P_{16Q_i} - P_Q)\eta - P_{v,i}\big) \cdot (\omega_i w)(\cdot, t)\big|\,dx \le c\, r_i^m \lambda \int_{2B_i} |(\omega_i w)(\cdot, t)|\,dx$$
$$\le \tfrac{c}{\varepsilon}\, \lambda^p |Q_i| + \varepsilon \int_{B_i} |P_{v,i}|^2\,dx.$$

Here we have also used the definition of Q_i, i.e., $|Q_i| = |B_i| \cdot 2\gamma r_i^{2m}$ and $\gamma = \lambda^{2-p}$. The assertion of the lemma in the case $i \in \Theta_1$ with $r_i < \varrho/4$ (note that $\delta_1(i) = 1$ in this case) now follows from the last estimate and (5.52). Finally, we mention that the asserted dependencies of the constants from the indicated parameters follow those in Lemmas 5.14, 5.15 and 5.24. □

As we have mentioned above, in the proof of the Caccioppoli type inequality, we will have to control a certain integral, i.e., the integral of $|v|^2-|v-w|^2$ on the horizontal slices $16B \setminus E_t(\lambda)$, $t \in 2\mathcal{T}$. The required estimate will be achieved in the following lemma by bringing together the information we have obtained so far, in particular we will use the estimates from Lemma 5.26.

Lemma 5.27. *Let $\lambda \geq c_E\lambda_1$. There exists $c = c(n, N, m, L, \kappa)$, such that for almost every $t \in 2\mathcal{T}$ there holds*

$$\int_{16B\setminus E_t(\lambda)} \big(|v|^2 - |v{-}w|^2\big)(\cdot,t)\,dx \geq -c\,\lambda^p\,|16Q\setminus E(\lambda)| - \frac{c}{s}\int_{16Q} |u - P_Q|^2\,dz.$$

Proof. We denote by

$$\Lambda \equiv \big\{i \in \Theta : \mathrm{spt}\omega_i \cap 16B \times \{t\} \neq \emptyset \text{ and } |v|+|w| \not\equiv 0 \text{ on } \mathrm{spt}\omega_i \cap (16B\times\{t\})\big\}.$$

Note that $\mathrm{spt}\omega_i \cap 16B \times \{t\} \neq \emptyset$ already implies $i \in \Theta$ and if $v \equiv w \equiv 0$ on $\mathrm{spt}\omega_i \cap 16B \times \{t\}$, then

$$\int_{\mathrm{spt}\omega_i\cap 16B\times\{t\}} |v|^2 - |v-w|^2\,dx = 0.$$

Since $\Lambda \subset \Theta = \Theta_1 \cup \Theta_2$ we can decompose $\Lambda = \Lambda_1 \cup \Lambda_2$, where $\Lambda_1 = \Lambda \cap \Theta_1$ and $\Lambda_2 = \Lambda \cap \Theta_2$. Noting that $\sum_{i\in\Lambda}\omega_i(\cdot,t) \equiv 1$ on $16B \setminus E_t(\lambda)$ we can rewrite the left-hand side of the asserted inequality as follows:

$$\begin{aligned}&\int_{16B\setminus E_t(\lambda)} \big(|v|^2 - |v{-}w|^2\big)(\cdot,t)\,dx \qquad (5.53)\\ &= \sum_{i\in\Lambda_1}\int_{16B}\big(\omega_i\,(\ldots)\big)(\cdot,t)\,dx + \sum_{i\in\Lambda_2}\int_{16B}\big(\omega_i\,(\ldots)\big)(\cdot,t)\,dx = I + II.\end{aligned}$$

In turn we estimate I and II.

We start with the estimate of II. Here we rewrite the integrand as follows: $|v|^2-|v{-}w|^2 = 2v\cdot w - |w|^2$ and estimate the first term by the use of Lemma 5.26, i.e., the estimate (5.47), to obtain

$$\begin{aligned}\sum_{i\in\Lambda_2}\Big|\int_{16B}(\omega_i v\cdot w)(\cdot,t)\,dx\Big| &\leq c\sum_{i\in\Lambda_2}\Big[\lambda^p\,|16Q\cap 2Q_i| + \frac{1}{s}\int_{8Q\cap 16Q_i}|u{-}P_Q|^2dz\Big]\\ &\leq c\,\lambda^p\,|16Q\setminus E(\lambda)| + \frac{c}{s}\int_{16Q}|u - P_Q|^2\,dz,\end{aligned}$$

where $c = c(n, N, m, L, \kappa)$. For the second term in II involving $|w|^2$ we apply Lemma 5.14, i.e., the estimate (5.24), and Hölder's inequality to obtain

$$\begin{aligned}\sum_{i\in\Lambda_2}\int_{16B}(\omega_i|w|^2)(\cdot,t)\,dx &\le c\sum_{i\in\Lambda_2}|16B\cap 2B_i|\Big(\sqrt{s\lambda^p}+\fint_{16Q\cap 16Q_i}|v|\,dz\Big)^2\\ &\le c\sum_{i\in\Lambda_2}|16B\cap 2B_i|\Big(s\,\lambda^p+\fint_{16Q\cap 16Q_i}|v|^2\,dz\Big)\\ &\le c\sum_{i\in\Lambda_2}\Big(\lambda^p\,|16Q\cap 2Q_i|+\frac{1}{s}\int_{16Q\cap 16Q_i}|v|^2\,dz\Big)\\ &\le c\,\lambda^p\,|16Q\setminus E(\lambda)|+\frac{c}{s}\int_{16Q\setminus E(\lambda)}|u-P_Q|^2\,dz,\end{aligned}$$

where $c = c(n, N, m, L, \kappa)$. Here we have also used the fact that $|16B \cap 2B_i|/|16Q \cap 2Q_i| = |16\mathcal{T} \cap 2\mathcal{T}_i|^{-1} \le c \max\{1/s, 1/(\gamma r_i^{2m})\} \le c/s$ (note for $i \in \Lambda$ we have $2\mathcal{H} \cap \mathrm{spt}\omega_i \ne \emptyset$ by definition and therefore $2\mathcal{T} \cap 2\mathcal{T}_i \ne \emptyset$. Moreover, from the fact that $i \in \Lambda_2 \subset \Theta_2$ we infer $1/(\gamma r_i^{2m}) \le c/s$ (see (5.11))) and $2Q_i \subset \mathbb{R}^{n+1} \setminus E(\lambda)$ from Lemma 5.2, (iii) and $|v| \le |u - P_Q|$. Combining the last two estimates we therefore arrive at

$$\begin{aligned}|II| &= \sum_{i\in\Lambda_2}\int_{16B}\big(\omega_i\,(2v\cdot w-|w|^2)\big)(\cdot,t)\,dx\\ &\le c\,\lambda^p\,|16Q\setminus E(\lambda)|+\frac{c}{s}\int_{16Q}|u-P_Q|^2\,dz.\end{aligned}$$

Now we start with the estimate for I. We can rewrite I as follows:

$$I=\sum_{i\in\Lambda_1}\int_{16B}\Big(\omega_i\big(|P_{v,i}|^2+2(v-P_{v,i})\cdot w\big)-\omega_i|w-P_{v,i}|^2\Big)(\cdot,t)\,dx = I_1+I_2,$$

with the obvious meaning of I_1 and I_2. To estimate I_2 we use the definition of w, the fact that $\sum_{j\in I(i)}\omega_j \equiv 1$ on $\mathrm{spt}\omega_i$ and (5.21) to obtain for $z = (x, t) \in 16Q \cap \mathrm{spt}\omega_i$:

$$\begin{aligned}|w(z)-P_{v,i}(x)| &= \Big|\sum_{j\in I(i)}\omega_j(z)(P_{v,j}-P_{v,i})(x)\Big|\\ &\le \sum_{j\in I(i)}\omega_j(z)|P_{v,j}(x)-P_{v,i}(x)| \le c\,r_i^m\,\lambda.\end{aligned}$$

Using the previous pointwise bound and the fact that $|16B \cap 2B_i|/|16Q \cap 2Q_i| \le 1/(2\gamma r_i^{2m})$ for $i \in \Lambda_1$ (note that $i \in \Lambda_1$ implies $i \in \Theta_1$ and therefore

$Q_i \subset 2\mathcal{H}$) we get

$$|I_2| \le c\,\lambda^2 \sum_{i\in\Lambda_1} r_i^{2m}\,|16B\cap 2B_i| \le c\,\lambda^p \sum_{i\in\Lambda_1} |16Q\cap 2Q_i| \le c\,\lambda^p\,|16Q\setminus E(\lambda)|,$$

where $c = c(n,N,m,L,\kappa)$. To get an estimate for I_1 we use Lemma 5.26 with $i\in\Theta_1$ and obtain for any fixed $0<\varepsilon\le 1$ that

$$\begin{aligned} I_1 &\ge \sum_{i\in\Lambda_1} \Big(\int_{16B} \omega_i(\cdot,t)|P_{v,i}|^2\,dx - \varepsilon\,\delta_1(i)\int_{B_i} |P_{v,i}|^2\,dx - \tfrac{c}{\varepsilon}\,\lambda^p\,|16Q\cap 2Q_i|\Big) \\ &= I_{1,1} - I_{1,2} - I_{1,3}, \end{aligned}$$

where c depends on n,N,m,L,κ and δ_1 was defined in Lemma 5.26, i.e., $\delta_1(i)=1$ if $r_i<\varrho/4$ and $\delta_1(i)=0$ otherwise. $\varepsilon>0$ will be specified later. In the sequel we will estimate $I_{1,i}$ for $i=1,2,3$. Performing the summation in $I_{1,3}$ we find $I_{1,3}\le \frac{c}{\varepsilon}\,\lambda^p|16Q\setminus E(\lambda)|$.

To estimate $I_{1,1}$ and $I_{1,2}$ we want to use the positive terms $\int_{16B}\omega_i(\cdot,t)$ $|P_{v,i}|^2\,dx$ from $I_{1,1}$ in order to compensate some of the negative terms of $I_{1,2}$. For this we proceed as follows: First, we choose $j(i)$ such that $Q_{j(i)}\cap \mathbb{R}^n\times\{t\}\ne\emptyset$. Then, there are two cases: In the case that $j(i)\in\Lambda_1$, we can compensate $|B_i|\sum_{k=0}^{m-1} r_i^{2k}\,|(D^kP_{v,j(i)})_{B_{j(i)}}|^2$ by one of the "good" positive terms from $I_{1,1}$ (here we have to choose ε in an appropriate way). Otherwise, i.e., in the case that $j(i)\in\Lambda_2\subset\Theta_2$, we can use the estimate (5.20) from Lemma 5.12.

We only have to consider those $i\in\Lambda_1$ with $\delta_1(i)\ne 0$, i.e., $r_i<\varrho/4$. Then $Q_j\subset 16Q_i\subset 16Q$ for all $j\in I(i)$ by (5.12) (note that $i\in\Lambda_1\subset\Theta_1$ and we only have to consider those i with $8B\cap 16B_i\ne\emptyset$; otherwise $P_{v,i}\equiv 0$ and we can choose $j(i)=i$). Since $\operatorname{spt}\omega_i\cap 16B\times\{t\}\ne\emptyset$ there exists $j(i)\in I(i)$ such that $t\in\mathcal{T}_{j(i)}$; in particular $j(i)\in\Lambda$. By construction we have $\omega_{j(i)}(\cdot,t)\ge c(n,m)>0$ on $B_{j(i)}$ (see Lemma 5.3). With this in mind we now apply in turn Lemma 5.12, i.e., (5.21), Lemma 3.5 with $(B_{j(i)},16B_{j(i)},P_{v,j(i)},0)$ instead of $(B_{x_0}(\varrho_1),B_{x_0}(\varrho_2),P,k)$ (note that $B_i\subset 16B_{j(i)}$ and $16r_{j(i)}/r_i\le 48$) and Hölder's inequality to deduce that

$$\begin{aligned} \int_{B_i}|P_{v,i}|^2\,dx &\le 2\int_{B_i}|P_{v,i}-P_{v,j(i)}|^2 + |P_{v,j(i)}|^2\,dx \\ &\le c\,r_i^{2m}\lambda^2|B_i| + c\int_{B_{j(i)}}|P_{v,j(i)}|^2\,dx \\ &\le c\,r_i^{2m}\lambda^2|B_i| + c_1\int_{16B}\omega_{j(i)}(\cdot,t)|P_{v,j(i)}|^2\,dx, \end{aligned}$$

with $c = c(n, N, m, L, \kappa)$ and $c_1 = c_1(n, m)$. We now consider those $i \in \Theta_1$ for which the chosen $j(i)$ is also in Λ_1, i.e., we let $\Lambda_1' \equiv \{i \in \Lambda_1 : j(i) \in \Lambda_1\}$. We decompose the sum $I_{1,2} = \sum_{i\in\Lambda_1}(\dots) = \sum_{i\in\Lambda_1'}(\dots) + \sum_{i\in\Lambda_1\setminus\Lambda_1'}(\dots)$ and set $c_2 = c_2(n, m) = \max_{\ell\in\mathbb{N}} \operatorname{card} I(\ell)$. For fixed $\ell \in \Lambda_1$ we consider $i \in \Lambda_1'$ with $j(i) = \ell$. Then $i \in I(\ell)$ and therefore the number of those i is bounded by c_2. We now choose $\varepsilon \equiv \varepsilon(n, m) = (c_1c_2)^{-1} < 1$. Then

$$
\begin{aligned}
\sum_{i\in\Lambda_1'}(\dots) &= \varepsilon \sum_{i\in\Lambda_1'} \delta_1(i) \int_{B_i} |P_{v,i}|^2\,dx \\
&\le \varepsilon \sum_{i\in\Lambda_1'} \delta_1(i) \left(c\, r_i^{2m}\, \lambda^2\, |B_i| + c_1 \int_{16B} \omega_{j(i)}(\cdot, t)|P_{v,j(i)}|^2\,dx \right) \\
&\le \varepsilon\, c \sum_{i\in\Lambda_1'} \lambda^p\,|Q_i| + \varepsilon\, c_1c_2 \sum_{\ell\in\Lambda_1} \int_{16B} \omega_\ell(\cdot, t)|P_{v,\ell}|^2\,dx \\
&= c\,\lambda^p\,|16Q \setminus E(\lambda)| + I_{1,1},
\end{aligned}
$$

which means that $-\sum_{i\in\Lambda_1'}(\dots) + I_{1,1} \ge -c\,\lambda^p\,|16Q \setminus E(\lambda)|$. Next, we consider the sum $\sum_{i\in\Lambda_1\setminus\Lambda_1'}(\dots)$. For $i \in \Lambda_1\setminus\Lambda_1'$ we know that $j(i) \in \Lambda$ (by the choice of $j(i)$) and $j(i) \notin \Lambda_1$. Therefore $j(i) \in I(i) \cap \Lambda_2$. Hence, $j(i) \in \Theta_2$. From (5.8) and (5.11) we therefore infer $1/(\gamma r_i^{2m}) \le 3^{2m}/(\gamma r_{j(i)}^{2m}) \le c/s$. From the beginning of the proof we already know that $Q_i \subset 16Q_i \subset 16Q$ for $i \in \Lambda_1$ with $\delta_1(i) \ne 0$. Using also (5.20) and Hölder's inequality we obtain

$$
\begin{aligned}
\sum_{i\in\Lambda_1\setminus\Lambda_1'}(\dots) &= \frac{1}{c_1c_2} \sum_{i\in\Lambda_1\setminus\Lambda_1'} \delta_1(i) \int_{B_i} |P_{v,i}|^2\,dx \\
&\le c \sum_{i\in\Lambda_1\setminus\Lambda_1'} \delta_1(i) \left(r_i^{2m}\lambda^2|B_i| + \frac{1}{\gamma r_i^{2m}} \int_{Q_i} |v|^2\,dz \right) \\
&\le c \sum_{i\in\Lambda_1\setminus\Lambda_1'} \delta_1(i) \left(\lambda^p\,|Q_i| + \frac{1}{s} \int_{Q_i} |u - P_Q|^2\,dz \right) \\
&\le c\,\lambda^p\,|16Q \setminus E(\lambda)| + \frac{c}{s} \int_{16Q} |u - P_Q|^2\,dz,
\end{aligned}
$$

with $c = c(n, N, m, L, \kappa)$. Joining the previous estimates with (5.53) we get the desired estimate. □

5.5. **Proof of the Caccioppoli type inequality.** In order to prove our Caccioppoli type inequality we will test the parabolic system with the test-function w_h. Then we decompose the domain of integration into the "good set" $E(\lambda)$ and its complement the "bad set," where the maximal function of $|D^m u|$ is smaller, respectively larger then a fixed parameter $\lambda \geq c_E\lambda_1$. We will use the integrals on the good set, whereas we will use the estimates from Lemma 5.25 and 5.27 to bound the integrals on the "bad set." Next, we will integrate the resulting inequality with respect to λ. Finally, we will have to bound $\int_Q \frac{|D^m u|^p}{M(z)^\beta}\,dz$ from below in terms of $\int_Q |D^m u|^{p-\beta}\,dz$. Here, we cannot apply the Muckenhoupt theorem, since we do not know that $M(|D^m u|)^{-\beta}$ is a Muckenhoupt weight in the parabolic case since we are using the strong maximal function. Therefore we will follow a different method (see [23]). This finally will lead us to the Caccioppli type inequality.

Proof. [Proof of Lemma 5.1] In the first step of the proof we test the parabolic system with w_h. Let $\lambda \geq c_E\lambda_1$, $\tau \in 2\mathcal{T}$ and $0 < h < \min\{\varrho, \sqrt[2m]{s/\gamma}\}$ where $\gamma = \lambda^{2-p}$. The function $\varphi = \eta w_h(\cdot,\tau)$ is an admissible test-function in (3.2), i.e., in the averaged form, since $D^k w_h(\cdot,\tau)$ is locally Lipschitz continuous on $16B$ for $0 \leq k \leq m-1$ (as shown in Lemma 5.22) and η has support in $16B$. Hence, choosing $\varphi = \eta w_h(\cdot,\tau)$ as a test-function in (3.2) and integrating with respect to τ over $(t_1,t) \subset 2\mathcal{T}$ we obtain

$$\int_{t_1}^t \int_{16B} \Big[\partial_\tau[u]_h\eta \cdot w_h + \langle [A(\cdot,D^m u)]_h, D^m(\eta w_h)\rangle + \langle [B(\cdot,D^m u)]_h, \delta(\eta w_h)\rangle\Big]\,dx\,d\tau = 0. \tag{5.54}$$

For the first term on the right side we recall the definition of $E_\tau(\lambda)$ from (5.7) and $v_h = [u-P_Q]_h\eta$ on $16B \times (t_1,t) \subset 16B \times 2\mathcal{T}$ since $\zeta \equiv 1$ on $2\mathcal{T}$ and observe that $\partial_\tau[P_Q]_h \equiv 0$. Then we exploit the fact that $v_h \equiv w_h$ on $E(\lambda)$, and finally that $|w_h - v_h|^2$ is weakly differentiable with respect to τ on $16B \times 2\mathcal{T}$ by Lemma 5.23. Using these ingredients the following chain of equalities holds:

$$\begin{aligned}
&\int_{t_1}^t \int_{16B} \partial_\tau[u]_h\eta \cdot w_h\,dx\,d\tau = \int_{t_1}^t \int_{16B} \partial_\tau v_h \cdot w_h\,dx\,d\tau \\
&= \tfrac12 \int_{t_1}^t \int_{16B} \partial_\tau |v_h|^2\,dx\,d\tau + \int_{t_1}^t \int_{16B\setminus E_\tau(\lambda)} \partial_\tau v_h \cdot (w_h - v_h)\,dx\,d\tau \\
&= \tfrac12 \int_{t_1}^t \int_{16B} \partial_\tau\big(|v_h|^2 - |w_h - v_h|^2\big)\,dxd\tau + \int_{t_1}^t \int_{16B\setminus E_\tau(\lambda)} \partial_\tau w_h \cdot (w_h - v_h)\,dxd\tau \\
&= I_1(t) - I_1(t_1) + I_2,
\end{aligned}$$

where we have abbreviated

$$I_1(\tau) = \tfrac{1}{2}\int_{16B} \big(|v_h|^2 - |w_h{-}v_h|^2\big)(\cdot,\tau)\,dx$$

and

$$I_2 = \int_{t_1}^{t}\int_{16B\setminus E_\tau(\lambda)} \partial_\tau w_h\cdot(w_h{-}v_h)\,dx\,d\tau.$$

Later, we will use $I_1(t)$ and estimate $I_1(t_1)$ and I_2. We start with the estimation of I_2. For this we in turn apply Lemma 5.20 in the case $\vartheta = 1$ (note that $|v_h| \le |[u-P_Q]_h|$) to infer

$$\begin{aligned}|I_2| &\le \int_{(16B\times 2\mathcal{T})\setminus E(\lambda)} |\partial_\tau w_h\cdot(w_h - v_h)|\,dz\\ &\le c\ \lambda^p\,|16Q\setminus E(\lambda)| + \frac{c}{s}\int_{8Q} |[u-P_Q]_h|^2\,dz,\end{aligned}$$

with $c = c(n,N,m,L,\kappa)$. Now, we use the fact that $t_1 \in 2\mathcal{T} = (-2^{2m}s, 2^{2m}s)$ is at our disposal. Therefore we choose $t_1 \in (-2^{2m}s, -s)$ (which we can do) such that

$$\begin{aligned}|I_1(t_1)| &= \Big|\int_{16B}\big(|v_h|^2{-}|w_h{-}v_h|^2\big)(\cdot,t_1)\,dx\Big|\\ &\le \frac{1}{s}\int_{-2^{2m}s}^{-s}\Big|\int_{16B}|v_h|^2{-}|w_h{-}v_h|^2\,dx\Big|\,d\tau.\end{aligned}$$

Recalling that $w_h = v_h$ on $E(\lambda)$, $\mathrm{spt} v_h \subset 8Q$, $|v_h| \le |[u-P_Q]_h|$ and using Lemma 5.18, the right-hand side of the previous estimate can be bounded as follows:

$$\begin{aligned}|I_1(t_1)| &\le \frac{1}{s}\int_{8Q}|v_h|^2\,dz + \frac{1}{s}\int_{16Q\setminus E(\lambda)}|w_h|^2\,dz\\ &\le c\ \lambda^p\,|16Q\setminus E(\lambda)| + \frac{c}{s}\int_{8Q}|[u-P_Q]_h|^2\,dz,\end{aligned}$$

where $c = c(n,N,m,L,\kappa)$. Now, we consider the remaining term in (5.54) (i.e., the term involving A and B). Here, we decompose the integral as follows:

$$\int_{t_1}^{t}\int_{E_\tau(\lambda)}\big(\dots\big)\,dx\,d\tau + \int_{t_1}^{t}\int_{16B\setminus E_\tau(\lambda)}\big(\dots\big)\,dx\,d\tau = II_1 + II_2.$$

Later, we will use the term II_1 and estimate II_2. To bound II_2 we use the growth estimates (3.3), $\mathrm{spt}\eta \subset 8B$, $D^\ell\eta \le c/\varrho^\ell$, $\varrho \le 1$, Lemma 5.25 and

$|v_h| \le |[u-P_Q]_h|$ to infer

$$\begin{aligned}|II_2| &\le c\ \sum_{k=0}^{m}\int_{t_1}^{t}\int_{16B\setminus E_\tau(\lambda)}\big[(|D^m u|+b)^{p-1}\big]_h\,|D^k(\eta w_h)|\,dx\,d\tau\\ &\le c\ \sum_{k=0}^{m}\varrho^{-(m-k)}\int_{(8B\times 2\mathcal{T})\setminus E(\lambda)}\big[(|D^m u|+b)^{p-1}\big]_h\,|D^k w_h|\,dz\\ &\le c\ \lambda^p\,|16Q\setminus E(\lambda)|+\frac{c}{s}\int_{8Q}|[u-P_Q]_h|^2\,dz,\end{aligned}$$

where $c=c(n,N,m,L,\kappa)$. Connecting the previous estimates with (5.54) (which means $I_1(t)+II_2\le |I_1(t_1)|+|I_2|+|II_2|$), noting that $w_h=v_h$ on $E(\lambda)$ and passing to the limit $h\searrow 0$ we arrive at

$$\begin{aligned}&\tfrac12\int_{16B}\big(|v|^2-|w-v|^2\big)(\cdot,t)\,dx\\ &\qquad+\int_{t_1}^{t}\int_{E_\tau(\lambda)}\langle A(\cdot,D^m u),D^m(\eta v)\rangle+\langle B(\cdot,D^m u),\delta(\eta v)\rangle\,dx\,d\tau\\ &\qquad\qquad\le c\,\lambda^p\,|16Q\setminus E(\lambda)|+\frac{c}{s}\int_{16Q}|u-P_Q|^2\,dz,\end{aligned}$$

where $c=c(n,N,m,L,\kappa)$. We note that t_1 is now a fixed time in $(-2^{2m}s,-s)$ and $t\in 2\mathcal{T}$ with $t>t_1$ is arbitrary. Hence, we can take any $t\in(-s,2^{2m}s)$ (and hence any $t\in\mathcal{T}=(-s,s)$). In the first integral we decompose $16B$ into $E_t(\lambda)$ and $16B\setminus E_t(\lambda)$. Then, on $E_t(\lambda)$ we have $v=w$, so that $\int_{E_t(\lambda)}|v-w|^2(\cdot,t)dx=0$ and on $16B\setminus E_t(\lambda)$ we use Lemma 5.27 to obtain

$$\begin{aligned}&\int_{16B}\big(|v|^2-|w-v|^2\big)(\cdot,t)\,dx\\ &\qquad=\int_{E_t(\lambda)}|v(\cdot,t)|^2\,dx+\int_{16B\setminus E_t(\lambda)}\big(|v|^2-|w-v|^2\big)(\cdot,t)\,dx\\ &\qquad\ge\int_{E_t(\lambda)}|v(\cdot,t)|^2\,dx-c\ \lambda^p\,|16Q\setminus E(\lambda)|-\frac{c}{s}\int_{16Q}|u-P_Q|^2\,dz,\end{aligned}$$

with $c=c(n,N,m,L,\kappa)$. Together with the second-to-last estimate (where we choose $t=s$ in the second term on the left side) this yields for almost every $t\in\mathcal{T}$ (with the obvious abbreviation for $(\dots)$)

$$\tfrac12\int_{E_t(\lambda)}|v(\cdot,t)|^2\,dx+\int_{t_1}^{s}\int_{E_\tau(\lambda)}\big(\dots\big)\,dx\,d\tau$$

$$\leq c\,\lambda^p\,|16Q\setminus E(\lambda)| + \frac{c}{s}\int_{16Q}|u-P_Q|^2\,dz.$$

Now, let $\beta \in (0, \frac{1}{4}\min\{p-1,1\}]$. In the next step we multiply the last inequality by $\lambda^{-(1+\beta)}$ and integrate with respect to λ over the interval $(c_E\lambda_1, \infty)$. This enables us to reduce the integrability exponent in our estimates. For almost every $t \in (-s,s)$ we get

$$\begin{aligned}\int_{c_E\lambda_1}^{\infty} \lambda^{-(1+\beta)} \Big(\tfrac{1}{2}\int_{E_t(\lambda)}|v(\cdot,t)|^2\,dx + \int_{t_1}^{s}\int_{E_\tau(\lambda)}(\dots)\,dx\,d\tau\Big)\,d\lambda \\ \leq c\int_{c_E\lambda_1}^{\infty}\lambda^{-(1+\beta)}\Big(\lambda^p\,|16Q\setminus E(\lambda)| + \frac{1}{s}\int_{16Q}|u-P_Q|^2\,dz\Big)\,d\lambda. \quad (5.55)\end{aligned}$$

We now define the truncated maximal function $\mathfrak{M}(z) = \max\{M_{16Q}(z), \lambda_1\}$ and apply Fubini's theorem, yielding for the first integral of the left-hand side

$$\begin{aligned}\int_{c_E\lambda_1}^{\infty}\lambda^{-(1+\beta)}\int_{E_t(\lambda)}|v(\cdot,t)|^2\,dx\,d\lambda &\geq \int_{16B}|v(\cdot,t)|^2\int_{c_E\mathfrak{M}(x,t)}^{\infty}\lambda^{-(1+\beta)}\,d\lambda\,dx \\ &= \frac{1}{\beta\,c_E^{\beta}}\int_{16B}\frac{|v(\cdot,t)|^2}{\mathfrak{M}(\cdot,t)^{\beta}}\,dx.\end{aligned}$$

Similarly we obtain for the second integral appearing on the left-hand side of (5.55):

$$\int_{c_E\lambda_1}^{\infty}\lambda^{-(1+\beta)}\int_{t_1}^{s}\int_{E_\tau(\lambda)}(\dots)\,dx\,d\tau\,d\lambda = \frac{1}{\beta\,c_E^{\beta}}\int_{t_1}^{s}\int_{16B}(\dots)\,\frac{dx\,d\tau}{\mathfrak{M}(x,\tau)^{\beta}}.$$

Finally, we treat the right-hand side of (5.55). Taking into account the definition of $E(\lambda)$ and using (5.6) we find for the first part

$$\begin{aligned}\int_{c_E\lambda_1}^{\infty}\lambda^{p-1-\beta}\,|16Q\setminus E(\lambda)|\,d\lambda &= \int_{c_E\lambda_1}^{\infty}\lambda^{p-1-\beta}\,|\{z\in 16Q : M_{16Q}(z) > \lambda\}|\,d\lambda \\ &\leq \tfrac{1}{p-\beta}\int_{16Q}M_{16Q}(z)^{p-\beta}\,dz \leq c\,\lambda_1^{p-\beta}\,|Q|,\end{aligned}$$

where $c = c(n,m,p,\kappa)$. The second part can easily be estimated since it does not depend on λ:

$$\frac{1}{s}\int_{c_E\lambda_1}^{\infty}\lambda^{-(1+\beta)}\int_{16Q}|u-P_Q|^2\,dz\,d\lambda = \frac{1}{\beta\,c_E^{\beta}}\int_{16Q}\lambda_1^{-\beta}\,\frac{|u-P_Q|^2}{s}\,dz.$$

Joining the previous observations with (5.55) and multiplying by βc_E^β we deduce for almost every $t \in (-s,s)$

$$\tfrac{1}{2}\int_{16B} \frac{|v(\cdot,t)|^2}{\mathfrak{M}(\cdot,t)^\beta}\,dx + \mathcal{A} + \mathcal{B} \le c\,\beta\,\lambda_1^{p-\beta}\,|Q| + c\int_{16Q} \lambda_1^{-\beta}\,\frac{|u-P_Q|^2}{s}\,dz, \quad (5.56)$$

where the $c = c(n,N,m,L,p,\kappa)$ and we have abbreviated

$$\mathcal{A} \equiv \int_{t_1}^{s}\int_{16B} \langle A(\cdot,D^m u), D^m(\eta v)\rangle \frac{dx\,d\tau}{\mathfrak{M}(x,\tau)^\beta}$$

and

$$\mathcal{B} \equiv \int_{t_1}^{s}\int_{16B} \langle B(\cdot,D^m u), \delta(\eta v)\rangle\,\frac{dx\,d\tau}{\mathfrak{M}(x,\tau)^\beta}.$$

The third and final step is devoted to estimating the second integral of the left-hand side of (5.56) from below by the use of the ellipticity of the vector field A. Recalling that $v = \eta(u-P_Q)$ on $16B \times 2\mathcal{T}$, $\operatorname{spt}\eta \subset 8B$, $D^m P_Q \equiv 0$, the ellipticity bound (2.2) of A, the growth condition (2.3) imposed on A and the fact that $\eta \equiv 1$ on $4B$ and $Q \subset 4B \times (t_1,s)$ we infer

$$\begin{aligned}
\mathcal{A} &= \int_{t_1}^{s}\int_{8B} \langle A(\cdot,D^m u), \eta^2 D^m u\rangle \\
&\quad + \sum_{k=0}^{m-1} \binom{m}{k} \langle A(\cdot,D^m u), D^{m-k}\eta^2 \odot D^k(u-P_Q)\rangle\ \frac{dx\,d\tau}{\mathfrak{M}(x,\tau)^\beta} \\
&\ge \int_{t_1}^{s}\int_{8B} \big(\nu|D^m u|^p - b^p\big)\eta^2 \frac{dx\,d\tau}{\mathfrak{M}(x,\tau)^\beta} \\
&\quad - c\sum_{k=0}^{m-1}\int_{16Q} (|D^m u|{+}b)^{p-1}\left|\frac{D^k(u-P_Q)}{\varrho^{m-k}}\right|\ \frac{dz}{\mathfrak{M}(z)^\beta} \\
&\ge \nu\int_Q \frac{|D^m u|^p}{\mathfrak{M}^\beta}\,dz - \int_{16Q}\frac{b^p}{\mathfrak{M}^\beta}\,dz \\
&\quad - c\sum_{k=0}^{m-1}\int_{16Q} \frac{\big(|D^m u|{+}b\big)^{p-1}}{\mathfrak{M}^\beta}\left|\frac{D^k(u-P_Q)}{\varrho^{m-k}}\right| dz \\
&\equiv \nu\,\mathcal{A}_1 - \mathcal{A}_2 - \mathcal{A}_3,
\end{aligned}$$

with the obvious meaning of $\mathcal{A}_1$ - $\mathcal{A}_3$ and $c = c(n,m,L)$. We start considering $\mathcal{A}_1$. Here we want to eliminate in $\int_Q \frac{|D^m u|^p}{\mathfrak{M}^\beta}\,dz$ the weight $\mathfrak{M}^{-\beta}$. In the elliptic case this can be done by Muckenhoupt's theorem (see [27], proof of Theorem 1). However, in the parabolic setting such an argument is not available since we are using the strong maximal function (defined in (5.1)).

This resembles the fact that no uniform system of cylinders is available due to the scaling of the system. At this point we exploit the properties of the truncated strong maximal function $\mathfrak{M}(z) = \max\{M_{16Q}(z), \lambda_1\}$. We start by decomposing Q into $E \equiv \{z \in Q : |D^m u(z)| \geq \beta\, \mathfrak{M}(z)\}$ and its complement $Q \setminus E$. Then, for $z \in Q \setminus E$ we have either $|D^m u(z)| < \beta M_{16Q}(z)$ or $|D^m u(z)| < \beta\lambda_1$. Hence, we find

$$\begin{aligned}\int_Q |D^m u|^{p-\beta}\, dz &= \int_E |D^m u|^{p-\beta}\, dz + \int_{Q\setminus E} |D^m u|^{p-\beta}\, dz\\ \leq \beta^{-\beta} \int_Q \frac{|D^m u|^p}{\mathfrak{M}^\beta}\, dz &+ \beta^{p-\beta} \int_Q M_{16Q}^{p-\beta} + \lambda_1^{p-\beta}\, dz \leq 2\, \mathcal{A}_1 + c\, \beta\, \lambda_1^{p-\beta} |Q|.\end{aligned}$$

Here we have exploited the fact that $\beta^{-\beta} \leq 2$ for $\beta > 0$ and $\beta^{p-\beta} \leq \beta$ since $p - \beta \geq 1$, the definition of $\mathcal{A}_1$ and (5.6). Note that $c = c(n, m, p, \kappa)$, which follows from the dependencies of the constant in (5.6) and those of $c_E = c_E(n, m, p, \kappa)$. For the estimate of $\mathcal{A}_2$ we note that $b \leq \mathfrak{M}$ almost everywhere to get

$$\mathcal{A}_2 = \int_{16Q} \frac{b^p}{\mathfrak{M}^\beta}\, dz \leq \int_{16Q} b^{p-\beta}\, dz.$$

To estimate the sum in $\mathcal{A}_3$ we use $|D^m u| + b \leq \mathfrak{M}$, Young's inequality, the interpolation Lemma 3.1 and hypothesis (5.5). This gives for any $0 \leq k \leq m-1$ and $0 < \tilde\varepsilon \leq 1$:

$$\begin{aligned}&\int_{16Q} \frac{(|D^m u| + b)^{p-1}}{\mathfrak{M}^\beta} \Big| \frac{D^k(u - P_Q)}{\varrho^{m-k}} \Big|\, dz\\ &\quad\leq \int_{16Q} (|D^m u| + b)^{p-1-\beta} \Big| \frac{D^k(u - P_Q)}{\varrho^{m-k}} \Big|\, dz\\ &\quad\leq \tilde\varepsilon \int_{16Q} (|D^m u| + b)^{p-\beta}\, dz + c_{\tilde\varepsilon} \int_{16Q} \Big| \frac{D^k(u - P_Q)}{\varrho^{m-k}} \Big|^{p-\beta}\, dz\\ &\quad\leq \tilde\varepsilon \int_{16Q} (|D^m u| + b)^{p-\beta} dz + \tilde\varepsilon \int_{16Q} |D^m u|^{p-\beta} dz + c_{\tilde\varepsilon} \int_{16Q} \Big| \frac{u - P_Q}{\varrho^m} \Big|^{p-\beta} dz\\ &\quad\leq 2\tilde\varepsilon\, \kappa\, \lambda_1^{p-\beta}\, |16Q| + c_{\tilde\varepsilon} \int_{16Q} \Big| \frac{u - P_Q}{\varrho^m} \Big|^{p-\beta}\, dz, \qquad (5.57)\end{aligned}$$

where $c_{\tilde\varepsilon} = c(n, N, m, p, 1/\tilde\varepsilon)$. Noting that $|16Q| = 16^{n+2m}|Q|$ and choosing $\tilde\varepsilon$ suitably (i.e., of the form $\tilde\varepsilon \equiv \varepsilon/c$ with some large c) we find for $\varepsilon > 0$ that

$$\mathcal{A}_3 \leq \varepsilon\, \lambda_1^{p-\beta}\, |Q| + c_\varepsilon \int_{16Q} \Big| \frac{u - P_Q}{\varrho^m} \Big|^{p-\beta}\, dz,$$

where $c_\varepsilon = c(n,m,L,\kappa,1/\varepsilon)$. Similarly, we estimate $\mathcal{B}$ from (5.56). Using the growth assumption (2.4) for B, the fact that $|D^\ell \eta| \le c/\varrho^\ell$ and $\varrho \le 1$ and (5.57) (for sufficiently small $\tilde\varepsilon$, i.e., for $\tilde\varepsilon \equiv \varepsilon/c$ with a suitably large c) we obtain

$$\begin{aligned}\mathcal{B} &\le c\,L \sum_{\ell=0}^{m-1}\sum_{k=0}^{\ell}\int_{16Q} \frac{(|D^m u|+b)^{p-1}}{\mathfrak{M}^\beta}\Big|\frac{D^k(u-P_Q)}{\varrho^{\ell-k}}\Big|\,dz\\ &\le \varepsilon\,\lambda_1^{p-\beta}\,|Q| + c_\varepsilon \int_{16Q}\Big|\frac{u-P_Q}{\varrho^m}\Big|^{p-\beta}dz,\end{aligned}$$

where $c_\varepsilon = c(n,N,m,L,p,\kappa,1/\varepsilon)$. Joining the previous estimates for $\mathcal{A}_1$ - $\mathcal{A}_3$ and $\mathcal{B}$ with (5.56), recalling that $s = \varrho^{2m}\lambda_1^{2-p}$ and rearranging terms we arrive at

$$\begin{aligned}&\tfrac12 \int_B \frac{|v(\cdot,t)|^2}{\mathfrak{M}(\cdot,t)^\beta}\,dx + \tfrac{\nu}{2}\int_Q |D^m u|^{p-\beta}\,dz \\ &\le c_\varepsilon \int_{16Q} \lambda_1^{p-2-\beta}\Big|\frac{u-P_Q}{\varrho^m}\Big|^2 + \Big|\frac{u-P_Q}{\varrho^m}\Big|^{p-\beta} + b^{p-\beta}\,dz + \tilde c\,(\beta+\varepsilon)\,\lambda_1^{p-\beta}|Q|,\end{aligned}\tag{5.58}$$

where $c_\varepsilon = c_\varepsilon(n,N,m,L,p,\kappa,1/\varepsilon)$ and $\tilde c = \tilde c(n,m,p,\kappa)$. From the hypothesis (5.4) we get

$$\lambda_1^{p-\beta}\,|Q| \le \kappa \int_Q (|D^m u|+b)^{p-\beta}\,dz \le 2^{p-1}\,\kappa \int_Q |D^m u|^{p-\beta} + b^{p-\beta}\,dz, \tag{5.59}$$

yielding a bound for the second term on the left-hand side from below. At this stage the proof is almost finished. We only have to replace in the estimate (5.58) the polynomial P_Q with P_{16Q}. From Lemma 3.6, the interpolation Lemma 3.1, Hölder's inequality and hypothesis (5.5) we find for $x \in 16B$ and $\tilde\varepsilon \in (0,1]$

$$\begin{aligned}|P_Q(x) - P_{16Q}(x)| &\le c \sum_{\ell=0}^{m-1} \varrho^\ell \fint_Q |D^\ell(u-P_{16Q})|\,dz\\ &\le \tilde\varepsilon\,\varrho^m \fint_Q |D^m u|\,dz + c_{\tilde\varepsilon}\fint_Q |u-P_{16Q}|\,dz\\ &\le c\,\tilde\varepsilon\,\varrho^m\lambda_1 + c_{\tilde\varepsilon}\fint_{16Q}|u-P_{16Q}|\,dz,\end{aligned}$$

where $c = c(n,m,\kappa)$ and $c_{\tilde\varepsilon} = c_{\tilde\varepsilon}(n,m,1/\tilde\varepsilon)$. Now we use this estimate with a suitable choice of $\tilde\varepsilon \ll 1$ to replace P_Q by P_{16Q} on the right side of (5.58), while we employ (5.59) to bound the second term on the left side from below. Thus, we infer for almost every $t \in (-s,s)$ that

$$\tfrac{1}{2}\int_B \frac{|v(\cdot,t)|^2}{\mathfrak{M}(\cdot,t)^\beta}\,dx + \big(\tfrac{\nu}{2^p\kappa} - \tilde c\,(\varepsilon+\beta)\big)\lambda_1^{p-\beta}|Q| \\ \le c_\varepsilon \int_{16Q} \lambda_1^{p-2-\beta}\Big|\frac{u-P_{16Q}}{\varrho^m}\Big|^2 + \Big|\frac{u-P_{16Q}}{\varrho^m}\Big|^{p-\beta} + b^{p-\beta} dz,$$

where $c_\varepsilon = c_\varepsilon(n,N,m,L,p,\kappa,1/\varepsilon)$ and $\tilde c = \tilde c(n,N,m,L,p,\kappa)$. We first fix $\varepsilon = \nu/(6\cdot 2^p\kappa\tilde c)$. This also fixes $c = c(n,m,N,p,L/\nu,\kappa)$. Then we choose $\beta \in (0, \frac{1}{4}\min\{p-1,1\})$ such that $\beta \le \nu/(6\cdot 2^p\kappa\tilde c)$. With these choices and recalling that $v(\cdot,t) = u(\cdot,t) - P_Q$ on B the previous estimate (divided by $\nu/(2\cdot 2^p\kappa)$) turns into

$$\int_B \frac{|u(\cdot,t)-P_Q|^2}{\mathfrak{M}(\cdot,t)^\beta}\,dx + \lambda_1^{p-\beta}|Q| \\ \le c\int_{16Q} \lambda_1^{p-2-\beta}\Big|\frac{u-P_{16Q}}{\varrho^m}\Big|^2 + \Big|\frac{u-P_{16Q}}{\varrho^m}\Big|^{p-\beta} + b^{p-\beta}\,dz,$$

with $c = c(n,N,m,p,L/\nu,\kappa)$. Taking mean values (note $|16Q| = 16^{n+2m}|Q|$ and $|Q| = 2\varrho^{2m}\lambda_1^{2-p}|B|$) and recalling that $\mathfrak{M}(z) = \max\{M_{16Q}(z),\lambda_1\}$ the desired Caccioppoli inequality follows . □

6. Higher integrability of very weak solutions

The following lemma is a consequence of Gagliardo-Nirenberg's theorem and the Hardy-Littlewood maximal theorem. It will be applied later for the particular cases $\sigma = 2$ and $\sigma = p-\beta$.

Lemma 6.1. *Let $\kappa \ge 1$, $p > 1$ and $0 \le \beta \le \frac{1}{4}\min\{p-1,1\}$ and $u \in L^{p-\beta}(-T,0;\, W^{m,p-\beta}(\Omega;\mathbb{R}^N)) \cap L^2(\Omega_T;\mathbb{R}^N)$ and let $Q \equiv B\times\mathcal{T} \equiv Q_{z_0}(\varrho,s)$ be a parabolic cylinder with $0<\varrho\le 1$, $s = \lambda^{2-p}\varrho^{2m}$ and $\lambda>0$, such that $\alpha Q \Subset \Omega_T$ for some $1\le\alpha\le 16$. Supposing that*

$$\fint_{\alpha Q} (|D^m u| + b)^{p-\beta}\,dz \le \kappa\ \lambda^{p-\beta}, \tag{6.1}$$

then for $1 \le \sigma \le \max\{2, p-\beta\}$ there exists $c = c(n,m,p,\kappa)$ such that

$$\fint_Q \Big|\frac{u-P_Q}{\varrho^m}\Big|^\sigma dz \le c\,\big(\lambda^\beta J\big)^{\frac{\sigma-\vartheta}{2}}\Big(\sum_{k=0}^m \fint_Q \Big|\frac{D^k(u-P_Q)}{\varrho^{m-k}}\Big|^{\vartheta\frac{rp}{rp-\beta(\sigma-\vartheta)}} dz\Big)^{\frac{rp-\beta(\sigma-\vartheta)}{rp}},$$

with $r = \frac{2(p-\beta)}{p}$ *and* $\vartheta \equiv \max\{1, \frac{n\sigma}{n+rm}\}$ *and* $P_Q\colon \mathbb{R}^n \to \mathbb{R}^N$ *denoting the mean value polynomial of* u *of degree* $\le m-1$*, defined by* $(\delta P_\varrho)_B = (\delta u)_Q$ *and (with* $M_{\alpha Q}$ *from* (5.3))

$$J \equiv \sup_{t\in\mathcal{T}} \fint_B \Big|\frac{u(\cdot,t)-P_Q}{\varrho^m}\Big|^2 \frac{dx}{\max\{M_{\alpha Q}(\cdot,t),\lambda\}^\beta}.$$

Proof. Without loss of generality we assume $z_0 = 0$ and we set $\theta \equiv \frac{\vartheta}{\sigma} \in (0,1)$. In the following we want to apply the Gagliardo-Nirenberg inequality, i.e., Lemma 3.2 in the case $k=0$. We note that our choice of exponents is admissible, since

$$\theta\big(m - \tfrac{n}{\vartheta}\big) - (1-\theta)\tfrac{n}{r} = \tfrac{\vartheta m}{\sigma} - \tfrac{n}{\sigma} - \tfrac{n}{r} + \tfrac{\vartheta n}{r\sigma} = -\tfrac{n}{\sigma} + \tfrac{\vartheta(mr+n)-n\sigma}{r\sigma} \ge -\tfrac{n}{\sigma}.$$

Applying Gagliardo-Nirenberg's inequality slicewise with respect to $t \in \mathcal{T}$ we get (recall $\theta\sigma = \vartheta$)

$$\fint_Q \Big|\frac{u-P_Q}{\varrho^m}\Big|^\sigma dz \le c \fint_{\mathcal{T}} \Big(\fint_B \sum_{k=0}^m \Big|\frac{D^k(u-P_Q)}{\varrho^{m-k}}\Big|^\vartheta dx\Big)\Big(\fint_B \Big|\frac{u-P_Q}{\varrho^m}\Big|^r dx\Big)^{\frac{\sigma-\vartheta}{r}} dt.$$

We now abbreviate $\mathfrak{M}(z) \equiv \max\{M_{\alpha Q}(z),\lambda\}$. Using Hölder's inequality (note that $r<2$) and then taking the supremum over $t\in\mathcal{T}$ we can estimate the second integral on the right-hand side and we obtain for almost every $t \in \mathcal{T}$

$$\begin{aligned}\fint_B \Big|\frac{u(\cdot,t)-P_Q}{\varrho^m}\Big|^r dx &= \fint_B \Big(\Big|\frac{u(\cdot,t)-P_Q}{\varrho^m}\Big|^2 \mathfrak{M}(\cdot,t)^{-\beta}\Big)^{\frac{r}{2}} \mathfrak{M}(\cdot,t)^{\frac{\beta r}{2}}\, dx \\ &\le J^{\frac{r}{2}} \Big(\fint_B \mathfrak{M}(\cdot,t)^{p-\beta}\, dx\Big)^{\frac{2-r}{2}}.\end{aligned}$$

Inserting this above and applying again Hölder's inequality (note $\frac{rp}{\beta(\sigma-\vartheta)} > 1$) leads us to

$$\begin{aligned}&\fint_Q \Big|\frac{u-P_Q}{\varrho^m}\Big|^\sigma dz \\ &\le c\, J^{\frac{\sigma-\vartheta}{2}} \fint_{\mathcal{T}} \Big(\fint_B \sum_{k=0}^m \Big|\frac{D^k(u-P_Q)}{\varrho^{m-k}}\Big|^\vartheta dx\Big)\Big(\fint_B \mathfrak{M}^{p-\beta}\, dx\Big)^{\frac{\beta(\sigma-\vartheta)}{rp}} dt \\ &\le c\, J^{\frac{\sigma-\vartheta}{2}} \Big(\fint_Q \sum_{k=0}^m \Big|\frac{D^k(u-P_Q)}{\varrho^{m-k}}\Big|^{\vartheta\frac{rp}{rp-\beta(\sigma-\vartheta)}} dz\Big)^{\frac{rp-\beta(\sigma-\vartheta)}{rp}} \Big(\fint_Q \mathfrak{M}^{p-\beta} dz\Big)^{\frac{\beta(\sigma-\vartheta)}{rp}}.\end{aligned}$$

The second integral appearing on the right-hand side of the last inequality is bounded by $c\,\lambda^{p-\beta}$, which follows from the Hardy-Littlewood maximal theorem, i.e., from (5.2) and the hypothesis (6.1) (completely similar to (5.6)):

$$\fint_Q \mathfrak{M}^{p-\beta}\,dz \le \fint_Q M_{\alpha Q}^{p-\beta}\,dz + \lambda^{p-\beta} \le c \fint_{\alpha Q} (|D^m u| + b)^{p-\beta}\,dz + \lambda^{p-\beta} \le c\,\lambda^{p-\beta},$$

where $c = c(n,m,p,\kappa)$. Inserting this above and noting that the resulting exponent of λ equals $\frac{\beta(\sigma-\vartheta)(p-\beta)}{rp} = \frac{\beta(\sigma-\vartheta)}{2}$, this proves the desired estimate. □

Lemma 6.2. *Let* $\max\{1, \frac{2n}{n+2m}\} < p < 2+\beta$. *Then, under the assumptions of Lemma 5.1 there holds*

$$\fint_Q |u - P_Q|^2\,dz \le c(n,N,m,p,L/\nu,\kappa)\ \varrho^{2m}\,\lambda^2.$$

Proof. Without loss of generality we assume that $z_0 = 0$ and we recall from Lemma 5.1 that $Q = Q(\varrho, s)$, where $s = \lambda^{2-p}\varrho^{2m}$. We choose $1 \le \alpha_1 < \alpha_2 \le 16$ and denote for $i = 1,2$ by $\alpha_i Q = Q(\alpha_i\varrho, \alpha_i^{2m}s)$ the intrinsic parabolic cylinder of "radius" $\alpha_i\varrho$ and by $P_{\alpha_i}\colon \mathbb{R}^n \to \mathbb{R}^N$ the mean value polynomials of u of degree $\le m-1$, defined by $(\delta P_{\alpha_i})_{\alpha_i B} = (\delta u)_{\alpha_i Q}$. Then, from (5.5) we infer by enlarging the domain of integration from $\alpha_2 Q$ to $16Q$ that the hypothesis (6.1) of Lemma 6.1 with $(\alpha_1\varrho, \alpha_2\varrho)$ instead of $(\varrho, \alpha\varrho)$ is fulfilled, since $\fint_{\alpha_2 Q}(\dots)\,dz \le 16^{n+2m} \fint_{16Q}(\dots)\,dz \le 16^{n+2m}\kappa\,\lambda^{p-\beta}$. The application of the lemma in the particular case $\sigma = 2$ then yields (recall that $1 \le \alpha_2/\alpha_1 \le 16$)

$$\fint_{\alpha_1 Q} \Big|\frac{u - P_{\alpha_1}}{\varrho^m}\Big|^2 dz \le c\,\big(\lambda^\beta J\big)^{\frac{2-\vartheta}{2}} \Big(\sum_{k=0}^m \fint_{\alpha_1 Q} \Big|\frac{D^k(u - P_{\alpha_1})}{\varrho^{m-k}}\Big|^{\vartheta\frac{rp}{rp-\beta(2-\vartheta)}}\,dz\Big)^{\frac{rp-\beta(2-\vartheta)}{rp}}, \quad (6.2)$$

where $r = \frac{2(p-\beta)}{p}$ and $\vartheta = \max\{1, \frac{2n}{n+rm}\}$ and $c = c(n,m,p,\kappa)$ and

$$J = \sup_{t\in\alpha_1\mathcal{T}} \fint_{\alpha_1 B} \Big|\frac{u(\cdot,t) - P_{\alpha_1}}{\varrho^m}\Big|^2 \frac{dx}{\max\{M_{\alpha_2 Q}(\cdot,t), \lambda\}^\beta}.$$

We again note that hypothesis (5.5) is also valid for $\alpha_1\varrho$ instead of 16ϱ, with a possibly larger constant, since $|16Q|/|\alpha_1 Q| \le 16^{n+2m}$. Therefore, we can apply Corollary 4.4 on the cylinder $\alpha_1 Q$ to bound the integrals on $\alpha_1 Q$ in (6.2) by $c\,\lambda^\vartheta$. To estimate J we will use the first term in the Caccioppoli inequality from Lemma 5.1 applied on the cylinders $\alpha_1 Q$ and $\alpha_2 Q$. To be precise, we have proved the Cacciopoli inequality only on the cylinders $Q, 16Q$. Nevertheless, by a different choice of cut off functions subordinate to $\alpha_1 Q, \alpha_2 Q$ instead of $Q, 16Q$ (i.e., such that $\operatorname{spt}(\eta\zeta) \subset \alpha_2 Q$ and $\eta\zeta \equiv 1$ on $\alpha_1 Q$), we can attain the Cacciopoli inequality also on the cylinders $\alpha_1 Q, \alpha_2 Q$. But for the sake of legibility we did not carry this out. Finally, let us mention that hypotheses (5.4) and (5.5) are also satisfied for $(\alpha_1 Q, \alpha_2 Q)$ instead of $(Q, 16Q)$ with a possibly larger constant, since $|\alpha_1 Q|/|Q| \le 16^{n+2m}$ and $|16Q|/|\alpha_2 Q| \le 16^{n+2m}$. Therefore, from Lemma 5.1 we infer

$$\begin{aligned}\lambda^\beta J &\le c\Big[\fint_{\alpha_2 Q} \Big| \frac{u - P_{\alpha_2}}{(\alpha_2\varrho - \alpha_1\varrho)^m} \Big|^2 dz + \lambda^{2-p+\beta} \fint_{\alpha_2 Q} \Big| \frac{u - P_{\alpha_2}}{(\alpha_2\varrho - \alpha_1\varrho)^m} \Big|^{p-\beta} + b^{p-\beta} dz \Big] \\ &\le c\,\varrho^{-2m} \fint_{\alpha_2 Q} \Big| \frac{u - P_{\alpha_2}}{(\alpha_2 - \alpha_1)^m} \Big|^2 dz + c\,\lambda^2,\end{aligned}$$

where $c = c(n, N, m, p, L/\nu, \kappa)$. Here, in the last line we have applied Young's inequality (note that $p - \beta < 2$) to estimate the second term on the right side and hypothesis (5.5) to estimate the term involving $b^{p-\beta}$. Joining this estimate with (6.2) (recall that we have bounded the integrals on the right side by $c\,\lambda^\vartheta$) we arrive at

$$\begin{aligned}\fint_{\alpha_1 Q} |u - P_{\alpha_1}|^2\, dz &\le c\ \varrho^{\vartheta m}\, \lambda^\vartheta \Big(\fint_{\alpha_2 Q} \Big| \frac{u - P_{\alpha_2}}{(\alpha_2 - \alpha_1)^m} \Big|^2 dz + \lambda^2 \Big)^{\frac{2-\vartheta}{2}} \\ &\le \tfrac{1}{2} \fint_{\alpha_2 Q} |u - P_{\alpha_2}|^2\, dz + c\ \big(\alpha_2 - \alpha_1\big)^{-2m(\frac{2}{\vartheta} - 1)} \varrho^{2m}\, \lambda^2.\end{aligned}$$

Here we have again used Young's inequality in the last line and $c = c(n, N, m, p, \nu, L, \kappa)$. Now we can apply the iteration Lemma 3.3 to "absorb" the integral of the right side on the left-hand side. This finally yields the desired estimate. □

6.1. **Reverse-Hölder type inequality.**

Lemma 6.3. *Let $\kappa \ge 1$. Then there exists $\beta_1 \in (0, \frac{1}{4}\min\{p-1, 1\})$ and $c \ge 1$ depending on $n, m, N, p, L/\nu, \kappa$ such that the following holds: Whenever $u \in L^{p-\beta}(-T, 0; W^{m,p-\beta}(\Omega; \mathbb{R}^N)) \cap L^2(\Omega_T; \mathbb{R}^N)$ for some $0 < \beta \le \beta_1$ is a very weak solution of the parabolic system (2.1) under the assumptions (2.2)–(2.4) and $Q \equiv Q_{z_0}(\varrho, s)$ is a parabolic cylinder such that $16^2 Q \Subset \Omega_T$, and $s = \lambda^{2-p}\varrho^{2m}$ with $\lambda > 0$, which satisfies*

$$\kappa^{-1}\lambda^{p-\beta} \le \fint_Q (|D^m u| + b)^{p-\beta}\, dz \tag{6.3}$$

and

$$\fint_{16^2Q} (|D^m u| + b)^{p-\beta}\, dz \le \kappa\ \lambda^{p-\beta}, \tag{6.4}$$

then there holds

$$\lambda^{p-\beta} \le c\, \Big(\fint_{16Q} |D^m u|^q\, dz \Big)^{\frac{p-\beta}{q}} + c \fint_{16Q} b^{p-\beta}\, dz, \tag{6.5}$$

where

$$q \equiv \begin{cases} \max\{p-\frac{1}{2}, \overline{q}\}, & \overline{q} = \frac{np(p-\beta)}{p(n+2m)-\beta m(2+p-\beta)} \quad \text{if } p-\beta \ge 2 \\ \max\{\frac{1}{2}(1+p), \overline{q}\}, & \overline{q} = \frac{2np}{p(n+2m)-\beta 4m} \quad \text{if } \frac{2n}{n+2m} < p-\beta < 2. \end{cases}$$

Proof. Without loss of generality we can assume that $z_0 = 0$. Again, for $\alpha \ge 1$ we write $\alpha Q = \alpha B \times \alpha \mathcal{T}$, i.e., $\alpha B = B(\alpha\varrho)$, $\alpha\mathcal{T} = \mathcal{T}(\alpha^{2m}s)$. We choose $\beta_1 \in (0,1)$ according to Lemma 5.1 and assume that $\beta \in (0, \beta_1]$. Now, we want to apply Lemma 5.1 on the cylinders Q and $16Q$. Therefore we first have to ensure that the hypotheses (5.4) and (5.5) are satisfied. Indeed, (5.4) directly corresponds with (6.3) and (6.4) yields (5.5) by enlarging the domain of integration from $16Q$ to 16^2Q, i.e., $\fint_{16Q}(\dots)\, dz \le 16^{n+2m} \fint_{16^2Q}(\dots)\, dz \le 16^{n+2m}\kappa\lambda^{p-\beta}$. Hence, hypotheses (5.4) and (5.5) of the Caccioppoli inequality from Lemma 5.1 are fulfilled with $16^{n+2m}\kappa$ instead of κ and the application yields

$$\begin{aligned} \lambda^{p-\beta} &\le c_{Cac} \fint_{16Q} \lambda^{p-2-\beta} \Big| \frac{u - P_{16Q}}{\varrho^m} \Big|^2 + \Big| \frac{u - P_{16Q}}{\varrho^m} \Big|^{p-\beta} + b^{p-\beta}\, dz \\ &= c_{Cac} \Big(I_2 + I_{p-\beta} + \fint_{16Q} b^{p-\beta}\, dz \Big), \end{aligned} \tag{6.6}$$

with the obvious meaning of I_2, $I_{p-\beta}$ and $c_{Cac} = c_{Cac}(n, N, m, p, L/\nu, \kappa)$. To estimate I_σ for $\sigma = 2$ or $\sigma = p - \beta$ we first apply Lemma 6.1 to obtain

$$\begin{aligned} I_\sigma &= \lambda^{p-\beta-\sigma} \fint_{16Q} \Big| \frac{u - P_{16Q}}{\varrho^m} \Big|^\sigma dz \\ &\le c\, \lambda^{p-\beta-\sigma} \left(\lambda^\beta J\right)^{\frac{\sigma-\vartheta}{2}} \Big(\sum_{k=0}^m \fint_{16Q} \Big| \frac{D^k(u - P_{16Q})}{\varrho^{m-k}} \Big|^{\vartheta \frac{rp}{rp-\beta(\sigma-\vartheta)}} dz \Big)^{\frac{rp-\beta(\sigma-\vartheta)}{rp}}, \end{aligned} \tag{6.7}$$

where $r = \frac{2(p-\beta)}{p}$ and $\vartheta = \max\{1, \frac{n\sigma}{n+rm}\}$ and

$$J = \sup_{t \in 16\mathcal{T}} \fint_{16B} \Big| \frac{u(\cdot,t) - P_{16Q}}{\varrho^m} \Big|^2 \frac{dx}{\max\{M_{16^2Q}(\cdot,t), \lambda\}^\beta}.$$

The idea to estimate J is now to use the sup-term of the Caccioppoli inequality. With the same arguments as before we find that the hypotheses (5.4) and (5.5) are satisfied with $(16Q, 16^2Q, 16^{n+2m}\kappa)$ instead of $(Q, 16Q, \kappa)$. So the application of the Caccioppoli inequality, i.e., Lemma 5.1, is justified and yields

$$\lambda^\beta J \le c_{Cac}\Big[\fint_{16^2Q} \Big|\frac{u-P_{16^2Q}}{\varrho^m}\Big|^2 dz + \lambda^{2-p+\beta} \fint_{16^2Q} \Big|\frac{u-P_{16^2Q}}{\varrho^m}\Big|^{p-\beta} + b^{p-\beta}\, dz\Big],$$

where $c = c(n, N, m, p, L/\nu, \kappa)$. Our next goal is to bound the right-hand side in the previous inequality by $c\,\lambda^2$. To estimate the first integral we apply Corollary 4.4 with $(16^2Q, 2, 0)$ instead of $(Q_{z_0}(\varrho, s), \vartheta, k)$ when $p - \beta \ge 2$, whereas in the case $p - \beta < 2$ we apply Lemma 6.2 instead. For the second term we again use Corollary 4.4 with $(16^2Q, p - \beta, 0)$ instead of $(Q_{z_0}(\varrho, s), \vartheta, k)$ and for the third term we recall that due to hypothesis (6.4) we have $\fint_{16^2Q} b^{p-\beta}\, dz \le \kappa\,\lambda^{p-\beta}$. Therefore, proceeding this way, we deduce that $\lambda^\beta J \le c\,\lambda^2$ with a constant c depending on n, N, m, p, ν, L and κ. Having arrived at this stage we want to ensure that the exponent $\vartheta\frac{rp}{rp-\beta(\sigma-\vartheta)}$ of $|D^k(u - P_{16Q})|$ in (6.7) is in any case bounded from above by q. First, we note that $\overline{q} = \max\{\frac{2np}{p(n+2m)-4\beta m}, \frac{np(p-\beta)}{p(n+2m)-m\beta(2+p-\beta)}\}$. Therefore, in the case that ϑ attends the value $\frac{n\sigma}{n+rm}$ we have for $\sigma = 2$ or $\sigma = p - \beta$:

$$\begin{aligned}\vartheta\, \frac{rp}{rp-\beta(\sigma-\vartheta)} &= \frac{n\sigma}{n+rm}\, \frac{rp}{rp-\beta\sigma(1-\frac{n}{n+rm})} = \frac{n\sigma rp}{rp(n+rm)-\beta\sigma rm} \\ &= \frac{n\sigma p}{p(n+rm)-\beta\sigma m} = \frac{n\sigma p}{np+2(p-\beta)m-\beta\sigma m} \\ &= \frac{n\sigma p}{p(n+2m)-\beta m(2+\sigma)} \le \overline{q} \le q.\end{aligned}$$

Moreover, in the case that $\vartheta = 1$ we see that for $\beta \ll 1$ we have

$$\vartheta\, \frac{rp}{rp-\beta(\sigma-\vartheta)} = \frac{rp}{rp-\beta(\sigma-1)} = 1 + \frac{\beta p(\sigma-1)}{r-\beta(\sigma-1)} \le q.$$

Therefore, recalling in (6.7) that $\lambda^\beta J \le c\,\lambda^2$ and applying in turn Hölder's and Young's inequality, we obtain for $\varepsilon > 0$

$$\begin{aligned}I_\sigma &\le c\,\lambda^{p-\beta-\vartheta}\Big(\sum_{k=0}^m \fint_{16Q}\Big|\frac{D^k(u-P_{16Q})}{\varrho^{m-k}}\Big|^q dz\Big)^{\frac{\vartheta}{q}} \\ &\le \varepsilon\,\lambda^{p-\beta} + c_\varepsilon\Big(\sum_{k=0}^m \fint_{16Q}\Big|\frac{D^k(u-P_{16Q})}{\varrho^{m-k}}\Big|^q dz\Big)^{\frac{p-\beta}{q}},\end{aligned}$$

where $c_\varepsilon = c_\varepsilon(n, N, m, p, L/\nu, \kappa, 1/\varepsilon)$. Now, we consider the remaining integrals in the previous estimate for I_σ. Here, we first recall the notation $16Q = 16B \times 16\mathcal{T}$, where $16\mathcal{T} = \mathcal{T}(16^{2m}s)$. Applying the Poincaré type inequality from Corollary 4.3 with $(16B, 16\mathcal{T}, 16\mathcal{T})$ instead of $(B, \mathcal{T}_1, \mathcal{T}_2)$ and noting that $|16\mathcal{T}|/(16\varrho)^{2m} = 2 \cdot 16^{2m}s/(16\varrho)^{2m} = 2\lambda^{2-p}$, we obtain for $0 \le k \le m-1$

$$\fint_{16Q} \Big|\frac{D^k(u - P_{16Q})}{\varrho^{m-k}}\Big|^q dz \le c\fint_{16Q} |D^m u|^q\, dz + c\Big(\lambda^{2-p}\fint_{16Q} (|D^m u| + b)^{p-1}\, dz\Big)^q,$$

where $c = c(n, N, m, L, p)$. To estimate the second term on the right-hand side we use Hölder's inequality (note that $q \ge p-1$) and the hypothesis (6.3) when $p < 2$, respectively (6.4) when $p > 2$), yielding that

$$\begin{aligned}\fint_{16Q} (|D^m u| + b)^{p-1}\, dz &= \Big(\dots\Big)^{1-\frac{1}{p-1}}\Big(\dots\Big)^{\frac{1}{p-1}} \\ &\le \Big(\fint_{16Q} (|D^m u| + b)^{p-\beta}\, dz\Big)^{\frac{p-2}{p-\beta}}\Big(\fint_{16Q} (|D^m u| + b)^q\, dz\Big)^{\frac{1}{q}} \\ &\le \kappa^{\frac{p-2}{p-\beta}}\, \lambda^{p-2}\Big(\fint_{16Q} (|D^m u| + b)^q\, dz\Big)^{\frac{1}{q}}.\end{aligned}$$

Inserting the last two estimates above and applying again Hölder's inequality, we deduce in our estimate of I_σ with $\sigma = 2$ or $\sigma = p - \beta$ that

$$I_\sigma \le \varepsilon\, \lambda^{p-\beta} + c_\varepsilon\, \Big(\fint_{16Q} |D^m u|^q\, dz\Big)^{\frac{p-\beta}{q}} + c_\varepsilon \fint_{16Q} b^{p-\beta}\, dz,$$

where $c = c(n, N, m, p, L/\nu, \kappa, 1/\varepsilon)$. Joining this estimate with (6.6) we finally arrive at

$$\lambda^{p-\beta} \le 2c_{Cac}\, \varepsilon\, \lambda^{p-\beta} + c_\varepsilon\, \Big(\fint_{16Q} |D^m u|^q\, dz\Big)^{\frac{p-\beta}{q}} + c_\varepsilon \fint_{16Q} b^{p-\beta}\, dz,$$

where $c_{Cac} = c_{Cac}(n, N, m, p, L/\nu, \kappa)$ and $c_\varepsilon = c_\varepsilon(n, N, m, p, L/\nu, \kappa, 1/\varepsilon)$. Choosing $\varepsilon = 1/(4c_{Cac})$ we can absorb $\frac{1}{2}\lambda^{p-\beta}$ on the left side, proving the desired estimate (6.5). □

6.2. **Proof of the main result.** We will conclude the higher integrability of $|D^m u|$ from the following version of Gehring's theorem, where a reverse Hölder type inequality valid only on particular cylinders is enough. It can be retrieved from [5], Lemma 15 by a different choice of the involved cylinders and exponents.

Lemma 6.4. *Let $\lambda_0 \geq 1$, $\kappa \geq 1$, $1 \leq q < p-\beta$, $\hat{p} \geq p$ and $f \in L^{p-\beta}_{loc}(Q_2)$, $k \in L^{\hat{p}}_{loc}(Q_2)$ with $Q_2 \equiv Q(2, 2^{2m})$. Suppose that for each $\lambda \geq \lambda_0$ and almost every $\mathfrak{z} \in Q_2$ with $f(\mathfrak{z}) > \lambda$ there exists a parabolic cylinder $Q \equiv Q_{\mathfrak{z}}(\varrho, s) \subset Q_2$ around $\mathfrak{z}$ such that*

$$\kappa^{-1}\lambda^{p-\beta} \leq \fint_Q f^{p-\beta}\,dz \leq \kappa\Big(\fint_Q f^q\,dz\Big)^{\frac{p-\beta}{q}} + \kappa\fint_Q k^{p-\beta}\,dz \leq \kappa^2\lambda^{p-\beta}.$$

Then there exists $\varepsilon_0 = \varepsilon_0(\kappa, p-\beta-q) > 0$ and $c = c(\kappa, p-\beta-q)$ such that $f \in L^{p-\beta+\varepsilon_1}_{\mathrm{loc}}(Q_2)$, with $\varepsilon_1 = \min\{\varepsilon_0, \hat{p}-p+\beta\}$ and there holds

$$\int_{Q_2} f^{p-\beta+\varepsilon}\,dz \leq c\,\lambda_0^{\varepsilon}\int_{Q_2} f^{p-\beta}\,dz + c\int_{Q_2} k^{p-\beta+\varepsilon}\,dz \qquad \forall\ \varepsilon \in (0, \varepsilon_1].$$

Proof of Theorem 2.1. Without loss of generality we can assume that $\varrho = 1$ and $z_0 = 0$. Otherwise we consider $v(x,t) = \varrho^{-m}u(x_0+\varrho x, t_0+\varrho^{2m}t)$ on $Q(1,1)$ and get the general result by rescaling to $Q_{z_0}(\varrho, \varrho^{2m})$. We set $Q_1 \equiv Q(1,1)$ and $Q_2 \equiv Q(2, 2^{2m})$ and define the parabolic distance of $z \in Q_2$ to the boundary of Q_2 by

$$\mathrm{d}_{\mathcal{P}}(z) = \inf_{\bar{z} \in \mathbb{R}^{n+1}\setminus Q_2} \min\big\{|x-\bar{x}|,\ \sqrt[2m]{|t-\bar{t}|}\big\}.$$

Furthermore, let $\beta_1 = \beta_1(n, N, m, p, L/\nu)$ be the constant from Lemma 6.3 and $\beta \in (0, \beta_1]$ such that $p - \beta > \frac{2n}{n+2m}$. On Q_2 we now define the functions

$$g \equiv |D^m u| + b \quad \text{and} \qquad f \equiv d_{\mathcal{P}}^{\alpha}\, g, \quad \text{with } \alpha = \tfrac{n+2m}{d}.$$

Note that d was introduced in the statement of the theorem. Finally, we set

$$\lambda_0^d \equiv \fint_{Q_2} g^{p-\beta}\,dz + 1 \tag{6.8}$$

and consider

$$\lambda \geq B^{\frac{1}{d}}\lambda_0, \qquad \text{where} \ \ B \equiv 2^{10(n+2m)}. \tag{6.9}$$

Now, suppose that $\mathfrak{z}$ is a point in Q_2 with $f(\mathfrak{z}) > \lambda$. We then denote by $r_{\mathfrak{z}}$ its parabolic distance to ∂Q_2, i.e., $r_{\mathfrak{z}} \equiv d_{\mathcal{P}}(\mathfrak{z})$, and choose

$$\gamma \equiv \gamma(\mathfrak{z}) \equiv (r_{\mathfrak{z}}^{-\alpha}\lambda)^{2-p} \tag{6.10}$$

as scaling factor of our parabolic cylinders.

We first consider **the case** $p \geq 2$. Let us note that $r_{\mathfrak{z}}^{\alpha} \leq 2^{\alpha} \leq B^{\frac{1}{d}}\lambda_0 \leq \lambda$ and $p \geq 2$ imply $\gamma = (r_{\mathfrak{z}}^{-\alpha}\lambda)^{2-p} \leq 1$. Hence, working on intrinsic cylinders of the type $Q_{\mathfrak{z}}(R, \gamma R^{2m})$ with $0 < R \leq r_{\mathfrak{z}}$ we have the inclusion $Q_{\mathfrak{z}}(R, \gamma R^{2m}) \subset Q_2$. The next step of the proof consists in finding such an appropriate intrinsic parabolic cylinder around $\mathfrak{z}$ on which the hypothesis (6.3) and (6.4) of Lemma 6.3 are fulfilled. For this aim we first note that for R satisfying $r_{\mathfrak{z}}/2^9 \le R < r_{\mathfrak{z}}$ there holds

$$\fint_{Q_{\mathfrak{z}}(R,\gamma R^{2m})} g^{p-\beta}\,dz \le \frac{|Q_2|}{|Q_{\mathfrak{z}}(R,\gamma R^{2m})|}\fint_{Q_2} g^{p-\beta}\,dz = \frac{2^{n+2m}}{R^{n+2m}\,\gamma}\;\lambda_0{}^d$$
$$\le \frac{2^{10(n+2m)}}{r_{\mathfrak{z}}^{n+2m}\,\gamma}\;\frac{\lambda^d}{B} = (r_{\mathfrak{z}}^{-\alpha}\lambda)^{p-2}\,(r_{\mathfrak{z}}^{-\alpha}\lambda)^d = (r_{\mathfrak{z}}^{-\alpha}\lambda)^{p-\beta}.$$

Here we have used in turn the definitions of λ_0, γ, α, d, B and $\lambda > B^{\frac{1}{d}}\lambda_0$. Furthermore, the Lebesgue differentiation theorem ensures that for almost every $\mathfrak{z} \in Q_2$ with $f(\mathfrak{z}) > \lambda$ there holds

$$\lim_{r\searrow 0}\fint_{Q_{\mathfrak{z}}(r,\gamma r^{2m})} g^{p-\beta}\,dz = g(\mathfrak{z})^{p-\beta} = \big(d_{\mathcal{P}}(\mathfrak{z})^{-\alpha} f(\mathfrak{z})\big)^{p-\beta} > \big(r_{\mathfrak{z}}^{-\alpha}\lambda\big)^{p-\beta}.$$

The last two estimates yield on the one hand a cylinder, namely $Q_{\mathfrak{z}}(R,\gamma R^{2m})$, which is too large, and on the other hand a cylinder which is too small. By the continuous dependence of the integral on the domain of integration there must be at least one cylinder in between, for which equality holds; i.e., there exists a radius $\varrho = \varrho(\mathfrak{z})$ with $0 < \varrho \le r_{\mathfrak{z}}/2^9$ such that

$$\fint_{Q_{\mathfrak{z}}(\varrho,\gamma\varrho^{2m})} g^{p-\beta}\,dz = \big(r_{\mathfrak{z}}^{-\alpha}\lambda\big)^{p-\beta}, \quad \fint_{Q_{\mathfrak{z}}(R,\gamma R^{2m})} g^{p-\beta}\,dz \le \big(r_{\mathfrak{z}}^{-\alpha}\lambda\big)^{p-\beta} \tag{6.11}$$

for all R with $\varrho \le R \le r_{\mathfrak{z}}$. We now set $Q \equiv Q_{\mathfrak{z}}(\varrho,\gamma\varrho^{2m})$ and $\alpha Q \equiv Q_{\mathfrak{z}}(\alpha\varrho,\gamma(\alpha\varrho)^{2m})$ for $\alpha > 0$. Then $2^9 Q \Subset Q_2$. From (6.11) we conclude that the hypothesis (6.3) and (6.4) of Lemma 6.3 are fulfilled with $(r_{\mathfrak{z}}^{-\alpha}\lambda, 1)$ instead of (λ,κ); i.e.,

$$\big(r_{\mathfrak{z}}^{-\alpha}\lambda\big)^{p-\beta} = \fint_{Q} g^{p-\beta}\,dz \quad\text{and}\qquad \fint_{2^8 Q} g^{p-\beta}\,dz \le \big(r_{\mathfrak{z}}^{-\alpha}\lambda\big)^{p-\beta}. \tag{6.12}$$

The application of the lemma then yields the following reverse-Hölder inequality

$$\big(r_{\mathfrak{z}}^{-\alpha}\lambda\big)^{p-\beta} \le c\,\bigg(\fint_{2^8 Q}|D^m u|^q\,dz\bigg)^{\frac{p-\beta}{q}} + c\,\fint_{2^8 Q} b^{p-\beta}\,dz, \tag{6.13}$$

where $q < p-\beta$ is the exponent defined in Lemma 6.3 and $c = c(n,m,N,p,\nu,L)$. Since $\varrho \le r_{\mathfrak{z}}/2^9$ and $\gamma \le 1$ we have for all $z \in 2^8 Q$ that $\mathrm{d}_{\mathcal{P}}(z) \le \min\{r_{\mathfrak{z}} + 2^8\varrho, (r_{\mathfrak{z}}^{2m} + \gamma(2^8\varrho)^{2m})^{\frac{1}{2m}}\} \le \frac{3}{2} r_{\mathfrak{z}}$ and $d_{\mathcal{P}}(z) \ge \min\{r_{\mathfrak{z}} - 2^8\varrho, (r_{\mathfrak{z}}^{2m} - \gamma(2^8\varrho)^{2m})^{\frac{1}{2m}}\} \ge \frac{1}{2} r_{\mathfrak{z}}$. Recalling the definition $f(z) = \mathrm{d}_{\mathcal{P}}^{\alpha}(z)\,g(z)$ this yields

$$c^{-1}\,f(z) \le r_{\mathfrak{z}}^{\alpha}\,g(z) \le c\,f(z) \qquad \text{for all } z \in 2^8 Q, \tag{6.14}$$

with a constant $c = c(n, m) \geq 1$. We now define $k \equiv \mathrm{d}^{\alpha}_{\mathcal{P}}\, b$ and show that there exists a constant $c = c(n, m, N, p, L/\nu)$, such that

$$\begin{aligned} c^{-1}\ \lambda^{p-\beta} &\overset{(a)}{\leq} \fint_{2^8 Q} f^{p-\beta}\, dz \\ &\overset{(b)}{\leq} c\Big(\fint_{2^8 Q} f^{q}\, dz\Big)^{\frac{p-\beta}{q}} + c \fint_{2^8 Q} k^{p-\beta}\, dz \overset{(c)}{\leq} c^2\ \lambda^{p-\beta}. \end{aligned} \tag{6.15}$$

The bound (a) follows from (6.12) and (6.14) and by enlarging the domain of integration from Q to $2^8 Q$ (note that $|Q|/|2^8 Q| = 2^{8(n+2m)}$)

$$\lambda^{p-\beta} \overset{(6.12)}{=} r_{\mathfrak{z}}^{\alpha(p-\beta)} \fint_{Q} g^{p-\beta}\, dz \overset{(6.14)}{\leq} c \fint_{Q} f^{p-\beta}\, dz \leq c(n, m) \fint_{2^8 Q} f^{p-\beta}\, dz.$$

The bound (b) can be derived by the use of (6.14), (6.12), (6.13) and (6.14), since

$$\begin{aligned} \fint_{2^8 Q} f^{p-\beta}\, dz &\overset{(6.14)}{\leq} c\, r_{\mathfrak{z}}^{\alpha(p-\beta)} \fint_{2^8 Q} g^{p-\beta}\, dz \overset{(6.12)}{\leq} c\ \lambda^{p-\beta} \\ &\overset{(6.13)}{\leq} c\, r_{\mathfrak{z}}^{\alpha(p-\beta)}\Big(\fint_{2^8 Q} |D^m u|^{q}\, dz\Big)^{\frac{p-\beta}{q}} + c\, r_{\mathfrak{z}}^{\alpha(p-\beta)} \fint_{2^8 Q} b^{p-\beta}\, dz \\ &\leq c\Big(\fint_{2^8 Q} (r_{\mathfrak{z}}^{\alpha} g)^{q}\, dz\Big)^{\frac{p-\beta}{q}} + c \fint_{2^8 Q} (r_{\mathfrak{z}}^{\alpha} b)^{p-\beta}\, dz \\ &\overset{(6.14)}{\leq} c\Big(\fint_{2^8 Q} f^{q}\, dz\Big)^{\frac{p-\beta}{q}} + c \fint_{2Q} k^{p-\beta}\, dz. \end{aligned}$$

Finally, the bound (c) follows from Hölder's inequality, the definitions of f and k, (6.14) and (6.12):

$$\begin{aligned} \Big(\fint_{2^8 Q} f^{q}\, dz\Big)^{\frac{p-\beta}{q}} + \fint_{2^8 Q} k^{p-\beta}\, dz &\leq \fint_{2^8 Q} f^{p-\beta}\, dz \\ &\overset{(6.14)}{\leq} c\, r_{\mathfrak{z}}^{\alpha(p-\beta)} \fint_{2^8 Q} g^{p-\beta}\, dz \overset{(6.12)}{\leq} c\ \lambda^{p-\beta}. \end{aligned}$$

Hence, for almost every $\mathfrak{z} \in Q_2$ with $f(\mathfrak{z}) > \lambda$ there exists a parabolic cylinder Q with center $\mathfrak{z}$ such that (6.15) holds. Therefore we can apply Lemma 6.4 with $(c, 2^8 Q)$ instead of (κ, Q), yielding that for any $0 < \beta \leq \beta_1$ there exists $\varepsilon_0 = \varepsilon_0(n, N, m, p, p-\beta, L/\nu) > 0$ such that $f \in L^{p-\beta+\varepsilon_1}_{\mathrm{loc}}(Q_2)$, with $\varepsilon_1 = \min\{\varepsilon_0, \hat{p} - p + \beta\}$ and there holds

$$\int_{Q_2} f^{p-\beta+\varepsilon}\, dz \le c\, \lambda_0^{\varepsilon} \int_{Q_2} f^{p-\beta}\, dz + c \int_{Q_2} k^{p-\beta+\varepsilon}\, dz \qquad \text{for all } \varepsilon \in (0, \varepsilon_1],$$

where $c = c(n, N, m, p, p-\beta, L/\nu)$. For $0 \le \beta \le \beta_1$ we denote by $\varepsilon_{p-\beta} > 0$ the gain in the integrability exponent coming from the application of Gehring's lemma. Since $\varepsilon_{p-\beta}$ continuously depends on $p-\beta$ and $\varepsilon_p > 0$, there exists $\beta_0 \in (0, \beta_1]$ such that $\varepsilon_{p-\beta} \ge \beta$ for all $\beta \in (0, \beta_0]$. For such $\beta \in (0, \beta_0]$ we infer from the last inequality with the choice $\varepsilon = \beta$ (this is possible since $\beta \in (0, \varepsilon_{p-\beta}]$)

$$\int_{Q_2} f^{p}\, dz \le c\, \lambda_0^{\beta} \int_{Q_2} f^{p-\beta}\, dz + c \int_{Q_2} k^{p}\, dz. \tag{6.16}$$

Now, using in turn that $|D^m u| \le g \le f$ on Q_1 (since for $z \in Q_1$ we have $\mathrm{d}_{\mathcal{P}}(z) \ge \min\{1, \sqrt[2m]{2^{2m}-1}\} \ge 1$), $|Q_2|/|Q_1| = 2^{n+2m}$ (and hence $⨍_{Q_1} |D^m u|^p\, dz \le 2^{2+2m} ⨍_{Q_2} f^p\, dz$), the estimate (6.16), $f \le 2^{\alpha} g$, $k \le 2^{\alpha} b$ on Q_2 (since $\mathrm{d}_{\mathcal{P}}(z) \le 2$ for $z \in Q_2$), the definitions $\lambda_0 = (⨍_{Q_2} g^{p-\beta}\, dz + 1)^{\frac{1}{d}}$ and $g = |D^m u| + b$ and Young's inequality we obtain

$$\begin{aligned} ⨍_{Q_1} |D^m u|^p\, dz &\le c\, \lambda_0^{\beta} ⨍_{Q_2} g^{p-\beta}\, dz + c ⨍_{Q_2} b^p\, dz \\ &\le c \left(⨍_{Q_2} (|D^m u| + b)^{p-\beta}\, dz \right)^{1+\frac{\beta}{d}} + c ⨍_{Q_2} \big(1 + b^p\big)\, dz, \end{aligned}$$

where $c = c(n, N, m, p, L/\nu)$. This proves the asserted estimate in the case $p \ge 2$.

We now deal with **the case** $\max\{1, \frac{2n}{n+2m}\} < p < 2$. The basic change with respect to the case $p \ge 2$ is that we now switch to the sub-quadratic scaling, i.e., we consider intrinsic cylinders of the type $Q_{\mathfrak{z}}(\gamma^{-\frac{1}{2m}} R, R^{2m})$. The parameter λ_0 is still defined as in (6.8) and λ is again chosen according to (6.9). We recall that $\mathfrak{z}$ is a point in Q_2 with $f(\mathfrak{z}) > \lambda$ and γ was defined in (6.10). But in contrast to the case $p \ge 2$ we now have $\gamma = (r_{\mathfrak{z}}^{-\alpha}\lambda)^{2-p} \ge 1$ by virtue of $r_{\mathfrak{z}}^{-\alpha}\lambda \ge 1$ and $2 - p > 0$. Hence, for $0 < R < r_{\mathfrak{z}}$ we have the inclusion $Q_{\mathfrak{z}}(\gamma^{-\frac{1}{2m}} R, R^{2m}) \subset Q_2$. Now, once again we have to find a suitable intrinsic parabolic cylinder around $\mathfrak{z}$ for which the conditions (6.3) and (6.4) of Lemma 6.3 is fulfilled and which is contained in Q_2. Initially, we show that (6.4) is satisfied for radii $r_{\mathfrak{z}}/2^9 \le R \le r_{\mathfrak{z}}$. From the definitions of $\lambda_0, \gamma, \alpha, d, B$ (particularly that $n + 2m = \alpha d$ and $(2-p)\frac{n}{2m} + d = p - \beta$)

we infer

$$\fint_{Q_{\mathfrak{z}}(\gamma^{-\frac{1}{2m}}R, R^{2m})} g^{p-\beta}\,dz \le \frac{|Q_2|}{|Q(\gamma^{-\frac{1}{2m}}R, R^{2m})|}\fint_{Q_2} g^{p-\beta}\,dz = \frac{2^{n+2m}}{R^{n+2m}\,\gamma^{-\frac{n}{2m}}}\;\lambda_0^d$$

$$\le \frac{2^{10(n+2m)}}{r_{\mathfrak{z}}^{n+2m}\,\gamma^{-\frac{n}{2m}}}\;\frac{\lambda^d}{B} = (r_{\mathfrak{z}}^{-\alpha}\lambda)^{(2-p)\frac{n}{2m}}\,(r_{\mathfrak{z}}^{-\alpha}\lambda)^d = (r_{\mathfrak{z}}^{-\alpha}\lambda)^{p-\beta}.$$

Now we can continue as in the case $p \ge 2$, but using cubes of the type $Q = Q_{\mathfrak{z}}(\gamma^{-\frac{1}{2m}}\varrho, \varrho^{2m})$ instead of those of the type $Q_{\mathfrak{z}}(\varrho, \gamma\varrho^{2m})$. Proceeding this way we first conclude the estimate (6.16), which immediately yields the assertion for the case $p < 2$ with completely the same arguments as before. This finishes the proof of Theorem 2.1. □

References

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

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

[3] R.A. Adams, "Sobolev Spaces," Academic Press, New York, 1978.

[4] V. Bögelein, *Regularity results for weak and very weak solutions of higher order parabolic systems*, PhD thesis, Friedrich-Alexnder-Universität Erlangen-Nürnberg, 2007.

[5] V. Bögelein, *Higher integrability for weak solutions of higher order degenerate parabolic systems,* Ann. Acad. Sci. Fenn., Math., 33 (2008), 387–412.

[6] B. Bojarski and T. Iwaniec, *Analytical foundations of the theory of quasiconformal mappings in $\mathbb{R}^n$,* Ann. Acad. Sci. Fenn., Math., 8 (1983), 257–324.

[7] S. Campanato, *Proprietà di una famiglia di spazi funzionali,* Ann. Scuola Norm. Sup. Pisa, 18 (1964), 137–160.

[8] R.R. Coifman and G. Weiss, "Analyse Harmonique Non-Commutative sur Certains Espaces Homogenes," Springer-Verlag, Berlin, Heidelberg, New York, 1971.

[9] E. DiBenedetto, *On the local behaviour of solutions of degenerate parabolic equations with measurable coefficients,* Ann. Scuola Norm. Sup. Pisa Cl. Sci., 13 (1986), 487–535.

[10] E. DiBenedetto, *Intrinsic harnack type inequalities for solutions of certain degenerate parabolic equations,* Arch. Rational Mech. Anal., 100 (1988), 129–147.

[11] E. DiBenedetto, "Degenerate Parabolic Equations," Springer-Verlag, Universitytext xv, 387, New York, NY, 1993.

[12] E. DiBenedetto and A. Friedman, *Regularity of solutions of nonlinear degenerate parabolic systems,* J. Reine Angew. Math., 349 (1984), 83–128.

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

[14] F. Duzaar and G. Mingione, *Second order parabolic systems, optimal regularity, and singular sets of solutions,* Ann. Inst. Henri Poincaré Anal. Non Linéaire, 22 (2005), 705–751.

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

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

[17] M. Giaquinta, "Introduction to Regularity Theory for Nonlinear Elliptic Systems," Birkhäuser Verlag, ETH Zürich. Basel, 1993.

[18] M. Giaquinta and M. Struwe, *On the partial regularity of weak solutions of nonlinear parabolic systems,* Math. Z., 179 (1982), 437–451.

[19] D. Gilbarg and N.S. Trudinger, "Elliptic Partial Differential Equations of Second Order," Springer-Verlag, Berlin, Heidelberg, New York, 1998.

[20] T. Iwaniec, *p-harmonic tensors and quasiregular mappings,* Ann. Math., 136 (1992), 589–624.

[21] T. Iwaniec and C. Sbordone, *Weak minima of variational integrals,* J. Reine Angew. Math., 454 (1994), 143–161.

[22] J. Kinnunen and J.L. Lewis, *Higher integrability for parabolic systems of p-Laplacian type,* Duke Math. j., 102 (2000), 253–271.

[23] J. Kinnunen and J.L. Lewis, *Very weak solutions of parabolic systems of p-Laplacian type,* Ark. Mat., 40 (2002), 105–132.

[24] J. Kinnunen and S. Zhou, *A local estimate for nonlinear equations with discontinuous coefficients,* Commun. Partial Differ. Equations, 24 (1999), 2043–2068.

[25] J. Kinnunen and S. Zhou, *A boundary estimate for nonlinear equations with discontinuous coefficients,* Differ. Integral Equ., 14 (2001), 475–492.

[26] J. Kristensen and G. Mingione, *The singular set of minima of integral functionals,* Arch. Ration. Mech. Anal., 180 (2006), 331–398.

[27] J.L. Lewis, *On very weak solutions to certain elliptic systems,* Comm. Partial Differential Equations, 18 (1993), 1515–1537.

[28] L. Nirenberg, *On elliptic partial differential equations,* Ann. Sc. Norm. Super. Pisa, III., Ser. 123 (1960), 115–162.

[29] J. Serrin, Pathological solutions of elliptic differential equations, Ann. Sc. Norm. Super. Pisa, Cl. Sci., 3, 18 (1964), 385–387.

[30] E.M. Stein, "Singular Integrals and Differentiability Properties of Functions," Princeton University Press, Princeton, New Jersey, 1970.

[31] E.W. Stredulinsky, *Higher integrability from reverse Hölder inequalities,* Indiana Univ. Math. J., 29 (1980), 407–413.