Rev. Mat. Iberoam. **30** (2014), no. 4, 1355–1386
DOI 10.4171/RMI/817
© European Mathematical Society

Calderón–Zygmund estimates for parabolic $p(x,t)$-Laplacian systems

Paolo Baroni and Verena Bögelein

Abstract. We prove local Calderón–Zygmund estimates for weak solutions of the evolutionary $p(x,t)$-Laplacian system

$$\partial_t u - \operatorname{div}\big(a(x,t)|Du|^{p(x,t)-2}Du\big) = \operatorname{div}\big(|F|^{p(x,t)-2}F\big)$$

under the classical hypothesis of logarithmic continuity for the variable exponent $p(x,t)$. More precisely, we show that the spatial gradient Du of the solution is as integrable as the right-hand side F, i.e.,

$$|F|^{p(\cdot)} \in L^q_{\mathrm{loc}} \implies |Du|^{p(\cdot)} \in L^q_{\mathrm{loc}} \quad \text{for any } q>1,$$

together with quantitative estimates. Thereby we allow the presence of eventually discontinuous coefficients $a(x,t)$, requiring only a VMO condition with respect to the spatial variable x.

1. Introduction

The aim in this paper is to provide a Calderón–Zygmund theory for weak solutions of the parabolic $p(x,t)$-Laplacian system

$$\partial_t u - \operatorname{div}\big(a(z)|Du|^{p(z)-2}Du\big) = \operatorname{div}\big(|F|^{p(z)-2}F\big) \quad \text{in } \Omega_T := \Omega\times(0,T). \tag{1.1}$$

Here, Ω is an open set in $\mathbb{R}^n$ with $n\geq 2$ and Ω_T denotes the parabolic cylinder over Ω. Since we consider the case of systems, the solution is a possibly vector valued function $u\colon \Omega_T\to\mathbb{R}^N$ with $N\geq 1$. With respect to the variable exponent $p(x,t)$ we assume logarithmic continuity which is classical in the theory of variable exponent problems, while the regularity assumption on $a(x,t)$ encompasses a large class of not necessarily continuous coefficients, including the ones of the splitting form $a(x,t)=b(x)c(t)$, where $b(\cdot)$ belongs to the class of VMO functions and $c(\cdot)$ is merely a measurable bounded function. For the precise assumptions we refer

Mathematics Subject Classification (2010): Primary 35D30; Secondary 35K40, 35K65, 35K67.
Keywords: Gradient estimates, degenerate parabolic systems, nonstandard growth condition, parabolic p-Laplacian.

to Section 2 below. As usual we consider weak solutions u of (1.1), meaning that they belong to a certain parabolic Sobolev space, in particular $|Du|^{p(\cdot)} \in L^1$; see Definition 2.1 below. The existence of such weak solutions is ensured by a result of Antontsev and Shmarev [7], [8]. Our intention in this paper is to establish a local Calderón–Zygmund theory for weak solutions to (1.1) and can be summarized as follows: For any $q > 1$ we prove the inclusion

$$|F|^{p(\cdot)} \in L^q_{\mathrm{loc}}(\Omega_T) \implies |Du|^{p(\cdot)} \in L^q_{\mathrm{loc}}(\Omega_T)$$

together with quantitative gradient estimates. At this stage we mention that when the exponent $p(\cdot)$ is constant the result has been obtained in [4]; the stationary case of our result was treated in [3]. We stress that, apart from the parabolic scaling deficit, our quantitative gradient estimate is homogeneous in the sense that the constant is independent of the $p(\cdot)$-energy of Du. To our knowledge this fact is new even for the stationary elliptic case.

Parabolic systems of the type considered in (1.1) are simplified versions of systems arising in the mathematical modeling of certain phenomena in fluid dynamics, such as models for non-Newtonian fluids, especially electro-rheological fluids. The peculiarity of these fluids is that their viscosity depends strongly on the external electromagnetic field and therefore varies in space and time. A mathematical model for electro-rheological fluids has been developed by Růžička in [32] and exhibits a $p(x,t)$-growth structure in the nonlinear diffusion term. For simplified versions of this model partial regularity results can be found in [5], [21] and for the stationary case in [2], [13]. Other applications (in the case of equations) are models for flows in porous media [6], [25].

Compared to the stationary case there are only a few regularity results for parabolic problems with nonstandard growth. The first we mention, which in turn is the starting point for almost any other regularity result in this area, is the self-improving property of higher integrability, i.e., the existence of some $\varepsilon > 0$, depending only on the structural constants, such that

$$|Du|^{p(\cdot)} \in L^{1+\varepsilon}_{\mathrm{loc}}(\Omega_T).$$

This result was first established in the case of the $p(x,t)$-Laplacian equation by Antontsev and Zhikov [9], and later for a quite general class of parabolic systems with $p(x,t)$-growth independently by Zhikov and Pastukhova [34] and Bögelein and Duzaar [10]. With regard to Hölder regularity, in the scalar case Chen and Xu [15] proved that weak solutions of the parabolic $p(x,t)$-Laplacian equation are locally bounded and Hölder continuous. The local Hölder continuity of the spatial gradient Du for the parabolic $p(x,t)$-Laplacian system has recently been established by Bögelein and Duzaar [11]. As already mentioned, for more general parabolic systems with nonstandard growth partial regularity results can be found in [5], [21].

The history of Calderón–Zygmund estimates for nonlinear problems begins in the elliptic setting. The result for the p-Laplacian equation, i.e., the scalar case $N = 1$, has been obtained by Iwaniec [26], while the vectorial case $N > 1$ has been treated by DiBenedetto and Manfredi [20]. The extension to elliptic equations with VMO coefficients has been achieved by Kinnunen and Zhou [29]. General elliptic equations, also involving nonstandard growth conditions, have been treated by Acerbi and Mingione [3], who built on previous ideas of Caffarelli and Peral [14] valid for homogeneous equations with highly oscillating coefficients. For the case of higher order systems with nonstandard growth conditions we refer to Habermann [24]. The anisotropic character of the evolutionary p-Laplacian system makes it impossible to use elliptic techniques in the parabolic setting. Indeed it was initially not clear how to transfer such results to the parabolic setting. The result has finally been achieved by Acerbi and Mingione [4] who introduced the necessary new tools for developing a local Calderón–Zygmund theory for the time dependent, parabolic case (see also Misawa [31] for the special case $F \in \mathrm{BMO}$). Later, extensions to general parabolic systems were obtained by Duzaar, Mingione, and Steffen [22] while a Calderón–Zygmund theory for evolutionary obstacle problems can be found in [12], [33].

The main difficulty when considering the time dependent parabolic case comes from the nonhomogeneous scaling behavior of the system, in the sense that the solution multiplied by a constant is in general no longer a solution. Note that this problem appears already in the standard growth case when $p(\cdot) \equiv p$ with $p \neq 2$. As a consequence, all local estimates for the solution (such as energy estimates or reverse Hölder inequalities) become inhomogeneous and the use of maximal operators, which are typically used in the proof of Calderón–Zygmund estimates, becomes delicate. The technique how for overcoming this problem goes back to the pioneering work of DiBenedetto and Friedman [19]. The idea is to choose a system of parabolic cylinders different from the standard one, whose space-time scaling depends on the local behavior of the solution itself. In a certain sense this allows rebalancing the nonhomogeneous scaling of the parabolic p-Laplacian system. This technique is by now classical and is the core of the proof of almost any regularity result for degenerate parabolic problems. The strategy is to find parabolic cylinders of the form

$$Q_\varrho^{(\lambda)}(z_o) := B_\varrho(x_o) \times \big(t_o - \lambda^{2-p}\varrho^2, t_o + \lambda^{2-p}\varrho^2\big), \quad z_o = (x_o, t_o) \tag{1.2}$$

such that the scaling parameter $\lambda > 0$ and the average of $|Du|^p$ over $Q_\varrho^{(\lambda)}(z_o)$ are coupled in the following way:

$$\fint_{Q_\varrho^{(\lambda)}(z_o)} |Du|^p \, dz \approx \lambda^p.$$

Such cylinders are called intrinsic cylinders or cylinders with intrinsic coupling. The delicate aspect of this coupling clearly lies in the fact that the value of the integral average must be comparable to the scaling factor λ, which itself is involved in the construction of the support of the integral. On such intrinsic cylinders, i.e., when $|Du|$ is comparable to λ in the above sense, the parabolic p-Laplacian system $\partial_t u = \operatorname{div}(|Du|^{p-2}Du)$ behaves in a certain sense like $\partial_t u = \lambda^{p-2}\Delta u$. Therefore, using intrinsic cylinders of the type $Q_\varrho^{(\lambda)}(z_o)$ we can rebalance the multiplicative factor λ^{p-2} for instance by rescaling u in time by a factor λ^{2-p}.

When dealing with the case of nonstandard growth the construction of such a uniform system of intrinsic cylinders is not possible anymore, since the exponent p appears in the scaling parameter λ^{2-p} and therefore the scaling will depend on the particular point z_o. This means that the scaling of the intrinsic cylinder will in fact depend on space and time, so that we have to deal with a *nonuniform intrinsic geometry*. More precisely, we consider cylinders of the form

$$Q_\varrho^{(\lambda)}(z_o) := B_\varrho(x_o) \times \big(t_o - \lambda^{(2-p_o)/p_o}\varrho^2, t_o + \lambda^{(2-p_o)/p_o}\varrho^2\big), \quad p_o := p(z_o), \tag{1.3}$$

with an intrinsic coupling of the form (where for simplicity we omit the role of the right-hand side F):

$$⨍_{Q_\varrho^{(\lambda)}(z_o)} |Du|^{p(z)}\, dz \approx \lambda.$$

Note that, differently from (1.2), we made a change of parameter $\lambda^{p_o} \leftrightarrow \lambda$, so that the right-hand side is independent of p_o and hence independent of z_o. The main difficulty now comes from the fact that the heuristics we described above for the standard growth case do not apply for the case of nonstandard growth, e.g. on $Q_\varrho^{(\lambda)}(z_o)$ the parabolic $p(x,t)$-Laplacian system behaves like $\partial_t u = \lambda^{(p(z)-2)/p(z)}\Delta u$ such that the multiplicative factor $\lambda^{(p(z)-2)/p(z)}$ does not cancel the scaling factor $\lambda^{(2-p_o)/p_o}$. This problem will be solved by a parabolic localization argument which has its origin in [10]; see Section 4.

Now, we briefly describe the strategy of the proof of our main result. Since we have to work on a system of nonuniform intrinsic cylinders of the type (1.3) there is no uniform maximal function available. For this reason we avoid the use of any maximal operator in the proof by a technique which goes back to [4]. Instead of maximal operators we construct a covering of the super-level sets

$$\big\{|Du(z)|^{p(z)} > \lambda\big\}, \quad \lambda \gg 1$$

by exit cylinders $Q_{\varrho_i}^{(\lambda)}(z_i)$, $i = 1, \dots, \infty$, defined according to (1.3) on which we have

$$⨍_{Q_{\varrho_i}^{(\lambda)}(z_i)} |Du|^{p(\cdot)} + M(|F|+1)^{p(\cdot)}\, dz \approx \lambda.$$

Thereby, $M \gg 1$ is a suitably chosen parameter depending on the structural constants of the problem. Then, we know that

$$⨍_{Q_{\varrho_i}^{(\lambda)}(z_i)} |Du|^{p(\cdot)}\, dz \lesssim \lambda \quad \text{and} \quad ⨍_{Q_{\varrho_i}^{(\lambda)}(z_i)} (|F|+1)^{p(\cdot)}\, dz \lesssim \frac{\lambda}{M}. \tag{1.4}$$

Therefore, if M is large u solves (here we suppose $a \equiv 1$ for simplicity)

$$\partial_t u - \operatorname{div}\big(|Du|^{p(z)-2}Du\big) \approx 0 \quad \text{on } Q_{\varrho_i}^{(\lambda)}(z_i)$$

approximately. This heuristic suggests comparing u to the solution w of

$$\begin{cases} \partial_t w - \operatorname{div}\big(|Dw|^{p(z_i)-2}Dw\big) = 0 & \text{in } Q_{\varrho_i}^{(\lambda)}(z_i), \\ w = u & \text{on } \partial_{\mathcal{P}} Q_{\varrho_i}^{(\lambda)}(z_i). \end{cases}$$

To be precise, this will be done in a two step comparison argument. Here, we stress that the comparison argument relies strongly on the localization technique introduced in [1], [3]; we replace the variable exponent $p(z)-2$ by the constant exponent $p(z_i)-2$ and we control the error, via estimates in $L\log L$ spaces, using the logarithmic continuity assumption (2.4). The advantage now is that the theory of DiBenedetto and Friedman [18] ensures that Dw satisfies an a priori L^∞-estimate. Via the comparison argument this L^∞-estimate can be transferred into estimates for Du on the super-level sets. At this stage the final result follows by a standard argument using Fubini's theorem.

2. Statement of the result

As already mentioned, Ω will be a bounded open set in $\mathbb{R}^n$, $n\geq 2$, and $\Omega_T := \Omega\times(0,T)$, $T>0$, will denote the space-time cylinder over Ω. Moreover, we shall consider an exponent function $p:\Omega_T\to(2n/(n+2),\infty)$. For $k\in\mathbb{N}$ we define $L^{p(\cdot)}(\Omega_T,\mathbb{R}^k)$ to be the set of those measurable functions $v\colon\Omega_T\to\mathbb{R}^k$ such that $|v|^{p(\cdot)}\in L^1(\Omega_T)$, i.e.,

$$L^{p(\cdot)}(\Omega_T,\mathbb{R}^k) := \left\{v\colon\Omega_T\to\mathbb{R}^k : \int_{\Omega_T}|v|^{p(\cdot)}\,dz<\infty\right\}.$$

As usual we shall deal with weak solutions to (1.1) which are specified in the following:

Definition 2.1. A map $u\in L^2(\Omega_T,\mathbb{R}^N)\cap L^1\big(0,T;W^{1,1}(\Omega,\mathbb{R}^N)\big)$ is a weak solution of the parabolic system (1.1) if and only if $Du\in L^{p(\cdot)}(\Omega_T,\mathbb{R}^{Nn})$ and

$$\int_{\Omega_T} u\cdot\varphi_t - \langle a(\cdot)|Du|^{p(\cdot)-2}Du, D\varphi\rangle\,dz = \int_{\Omega_T}\langle |F|^{p(\cdot)-2}F, D\varphi\rangle\,dz \tag{2.1}$$

holds for every test function $\varphi\in C_0^\infty(\Omega_T,\mathbb{R}^N)$.

Since our problem is local in nature, it is not restrictive to assume the existence of γ_1 and γ_2 such that

$$\frac{2n}{n+2}<\gamma_1\leq p(z)\leq\gamma_2<\infty\quad\text{for all } z\in\Omega_T. \tag{2.2}$$

The lower bound $\gamma_1>2n/(n+2)$ is unavoidable even in the constant exponent case $p(\cdot)\equiv p$; see Chapters 5 and 8 of [17]. With respect to the regularity of $p(\cdot)$ we will assume the strong logarithmic continuity condition

$$|p(z)-p(\tilde z)|\leq\omega\big(d_{\mathcal{P}}(z,\tilde z)\big)\quad\text{for any } z,\tilde z\in\Omega_T, \tag{2.3}$$

where the parabolic metric $d_{\mathcal{P}}$ is given by

$$d_{\mathcal{P}}(z,\tilde z) := \max\left\{|x-\tilde x|, \sqrt{|t-\tilde t\,|}\right\}\quad\text{for } z=(x,t),\ \tilde z=(\tilde x,\tilde t)\in\mathbb{R}^{n+1}$$

and $\omega : [0, +\infty) \to [0, +\infty)$ is a nondecreasing modulus of continuity satisfying

$$\limsup_{\varrho \downarrow 0} \omega(\varrho) \log\left(\frac{1}{\varrho}\right) = 0, \tag{2.4}$$

an assumption which was first considered in [1] and then used in [3] and [13]. By virtue of (2.4) we may assume that there exists $R_1 \in (0, 1]$ depending on $\omega(\cdot)$ such that

$$\omega(\varrho) \log\left(\frac{1}{\varrho}\right) \le 1 \quad \text{for all } \varrho \in (0, R_1]. \tag{2.5}$$

For the coefficient function $a : \Omega_T \to \mathbb{R}$ we will assume its measurability and that

$$\nu \le a(z) \le L \quad \text{for any } z \in \Omega_T \tag{2.6}$$

holds with some constants $0 < \nu \le 1 \le L$. With regard to its regularity, we will only assume that it satisfies a VMO condition with respect to the spatial variable. More precisely, writing

$$(a)_{x_o,\varrho}(t) := \fint_{B_\varrho(x_o)} a(x,t)\, dx \quad \text{for } B_\varrho(x_o) \subset \Omega,$$

we assume that there exists $\tilde{\omega} \colon [0, \infty) \to [0, 1]$ such that

$$\sup_{B_\varrho(x_o) \subset \Omega, 0 < \varrho \le r} \fint_{B_\varrho(x_o)} \big| a(x,t) - (a)_{x_o,\varrho}(t) \big|\, dx \le \tilde{\omega}(r) \tag{2.7}$$

for a.e. $t \in (0, T)$ any $r > 0$ and

$$\lim_{r \downarrow 0} \tilde{\omega}(r) = 0. \tag{2.8}$$

Here, we stress that with respect to time we assume nothing more than measurability. Moreover, our assumptions on a allow product coefficients of the type $a(x,t) = b(x)\, c(t)$, with $b \in \mathrm{VMO}(\Omega) \cap L^\infty(\Omega)$ and $c \in L^\infty(0, T)$. Now we are ready to state our main result.

Theorem 2.2. *Let u be a weak solution of the parabolic system (1.1) where $p \colon \Omega_T \to [\gamma_1, \gamma_2]$ and $a \colon \Omega_T \to [\nu, L]$ satisfy assumptions (2.2)–(2.4) and (2.6)–(2.8). Moreover, assume that $|F|^{p(\cdot)} \in L^q_{\mathrm{loc}}(\Omega_T)$ for some $q > 1$. Then we have*

$$|Du|^{p(\cdot)} \in L^q_{\mathrm{loc}}(\Omega_T).$$

Moreover, for $K \ge 1$ there exist a radius

$$R_o = R_o(n, N, \nu, L, \gamma_1, \gamma_2, K, \omega(\cdot), \tilde{\omega}(\cdot), q) > 0$$

and a constant $c \equiv c(n, N, \nu, L, \gamma_1, \gamma_2, q)$ such that the following holds: if

$$\int_{\Omega_T} |Du|^{p(\cdot)} + (|F| + 1)^{p(\cdot)}\, dz \le K, \tag{2.9}$$

then for every parabolic cylinder $Q_{2R} \equiv Q_{2R}(\mathfrak{z}_o) \Subset \Omega_T$ with $R \in (0, R_o]$, there holds

$$\fint_{Q_R} |Du|^{p(\cdot)q}\, dz \le c\Big[\fint_{Q_{2R}} |Du|^{p(\cdot)}\, dz + \Big(\fint_{Q_{2R}} (|F|+1)^{p(\cdot)q}\, dz\Big)^{1/q}\Big]^{1+d(p_o)(q-1)}, \tag{2.10}$$

where

$$d(p_o) := \begin{cases} \dfrac{p_o}{2} & \text{if } p_o \ge 2, \\[2mm] \dfrac{2p_o}{p_o(n+2)-2n} & \text{if } p_o < 2, \end{cases} \qquad \text{with } p_o := p(\mathfrak{z}_o). \tag{2.11}$$

We note that the constant c in Theorem 2.2 remains stable when $q \downarrow 1$ and it blows up, i.e., $c \to \infty$ when $q \to \infty$. Moreover, due to an improved localization technique we are able to prove the gradient estimate (2.10) with a constant c independent of the parameter K, which was not known even in the elliptic case; see [3], [24].

Remark 2.3. The same result holds true if we assume, instead of the VMO condition (2.8), that the BMO seminorm of a with respect to x is small, i.e., that

$$[a]_{\mathrm{BMO}} := \sup_{r>0} \tilde{\omega}(r) \le \varepsilon_{\mathrm{BMO}}$$

with some constant $\varepsilon_{\mathrm{BMO}} > 0$ depending on $n, N, \nu, L, \gamma_1, \gamma_2$, and q.

3. Preliminaries and notation

3.1. Notation

For a point $z_o \in \mathbb{R}^{n+1}$ we shall always write $z_o = (x_o, t_o)$ with $x_o \in \mathbb{R}^n$ and $t_o \in \mathbb{R}$ and we shall consider, as we did for instance in the statement of Theorem 2.2, symmetric parabolic cylinders around z_o of the form $Q_\varrho(z_o) := B_\varrho(x_o) \times (t_o - \varrho^2, t_o + \varrho^2)$. Moreover, in the course of the proof of our main result, in order to rebalance the nonhomogeneity of the parabolic system, we shall also deal with scaled cylinders of the form

$$Q_\varrho^{(\lambda)}(z_o) := B_\varrho(x_o) \times \Lambda_\varrho^{(\lambda)}(z_o), \tag{3.1}$$

where $\lambda > 0$ and

$$\Lambda_\varrho^{(\lambda)}(z_o) := \big(t_o - \lambda^{(2-p_o)/p_o}\varrho^2, t_o + \lambda^{(2-p_o)/p_o}\varrho^2\big).$$

In any case, when considering a particular cylinder $Q_\varrho^{(\lambda)}(z_o)$ with center z_o, we denote by p_o the value of $p(\cdot)$ at the center of the cylinder, i.e., $p_o \equiv p(z_o)$. Note that such a system of scaled cylinders is nonuniform in the sense that the scaling $\lambda^{(2-p_o)/p_o}$ depends on the particular point z_o via $p_o \equiv p(z_o)$. In the particular case $\lambda = 1$ the cylinders $Q_\varrho^{(1)}(z_o)$ reduce to the standard parabolic ones, i.e., $Q_\varrho^{(1)}(z_o) \equiv Q_\varrho(z_o)$. By $\chi Q_\varrho^{(\lambda)}(z_o)$, for a constant $\chi > 1$, we denote the χ-times enlarged cylinder, i.e., $\chi Q_\varrho^{(\lambda)}(z_o) := Q_{\chi\varrho}^{(\lambda)}(z_o)$. For $g \in L^1_{\mathrm{loc}}(C, \mathbb{R}^k)$, with $C \subset \Omega_T$ of strictly positive measure, we shall write

$$(g)_C := \fint_C g(z)\,dz := \frac{1}{|C|}\int_C g(z)\,dz$$

for the mean value of g on C. Finally, α_n denotes the Lebesgue measure of the unit ball $B_1(0)$ in $\mathbb{R}^n$. We will denote by c a generic constant greater than one, possibly varying from line to line. We will indicate the dependencies of the constants in parentheses. For compactness of notation we will denote by the word *data* exactly the set of parameters n, N, ν, L, γ_1, and γ_2, so that by writing $c(data, M)$ we will mean that the constant c depends on n, N, ν, L, γ_1, γ_2 and moreover upon M. Constants we need to recall will be denoted with special symbols, such as c_{DiB}, $\tilde{c}$, c_*, and c_ℓ.

3.2. Preliminaries

The following lemma is a standard iteration lemma and can be found for instance in [23], Lemma 6.1.

Lemma 3.1. *Let $\phi : [R, 2R] \to [0, \infty)$ be a function such that*

$$\phi(r_1) \le \frac{1}{2}\,\phi(r_2) + \mathcal{A} + \frac{\mathcal{B}}{(r_2 - r_1)^\beta} \quad \text{for every } R \le r_1 < r_2 \le 2R,$$

where $\mathcal{A}, \mathcal{B} \ge 0$ and $\beta > 0$. Then

$$\phi(R) \le c(\beta)\left[\mathcal{A} + \frac{\mathcal{B}}{R^\beta}\right].$$

The next lemma is a useful tool when dealing with p-growth problems. The continuous dependence of the constant on p allows considering instead a constant depending on γ_1 and γ_2 when $p \in [\gamma_1, \gamma_2]$.

Lemma 3.2. *Let $p \in [\gamma_1, \gamma_2]$. Then there exists a constant $c \equiv c(n, N, \gamma_1, \gamma_2)$ such that for any A, $B \in \mathbb{R}^{Nn}$, not both zero, there holds*

$$\left(|A|^2 + |B|^2\right)^{(p-2)/2} |B - A|^2 \le c\left\langle |B|^{p-2}B - |A|^{p-2}A, B - A\right\rangle.$$

The following lemma can be deduced from Lemma 2.2 of [16]. Note that the dependence of the constant on γ_2 instead of p can be deduced from the proof of the lemma.

Lemma 3.3. *Let $p \in [\gamma_1, \gamma_2]$. Then there exists a constant $c_\ell \equiv c_\ell(\gamma_2)$ such that for any A, $B \in \mathbb{R}^{Nn}$ there holds*

$$|A|^p \le c_\ell\,|B|^p + c_\ell\left(|A|^2 + |B|^2\right)^{(p-2)/2} |B - A|^2.$$

Finally, we state a useful estimate which is a consequence of Iwaniec's inequality for Orlicz spaces [27]; see also inequality (28) of [3]. Fix $\beta > 0$, $Q \subset \mathbb{R}^{n+1}$, and $g \in L^\sigma(Q)$ for some $\sigma > 1$. Then, there holds

$$\fint_Q |g| \log^\beta\Big(e + \frac{|g|}{(g)_Q}\Big)\,dz \le c(\sigma,\beta)\Big(\fint_Q |g|^\sigma\,dz\Big)^{1/\sigma} \quad \text{for all } \sigma > 1. \tag{3.2}$$

The constant $c(\sigma,\beta)$ blows up when $\sigma \downarrow 1$. Moreover, $c(\sigma,\beta)$ depends continuously on β and therefore it can be replaced by a constant $c(\sigma,\gamma_1,\gamma_2)$ if $\beta \in [\gamma_2', \gamma_1']$.

4. Nonuniform intrinsic geometry

Lemma 4.1 yields a parabolic localization technique. Obviously the difficulty stems from the necessity of coupling the technique of intrinsic geometry with the localization needed to treat the variable exponent growth conditions. As we already pointed out in the introduction, this will be achieved using a *nonuniform intrinsic geometry*, i.e., a system of cylinders as defined in (3.1) whose scaling depends on the particular point considered. The following lemma is due to [10].

Lemma 4.1. *Let $\kappa, K, M \ge 1$ and $p\colon \Omega_T \to [\gamma_1,\gamma_2]$ satisfy (2.3) and (2.5). Then there exists a radius $\varrho_o \equiv \varrho_o(n,\gamma_1,\kappa,K,M,\omega(\cdot)) \in (0,R_1]$ such that the following holds: whenever $Du, F \in L^{p(\cdot)}(\Omega_T,\mathbb{R}^{Nn})$ satisfy (2.9) and $Q_\varrho^{(\lambda)}(z_o) \subset \Omega_T$ is a parabolic cylinder with $\varrho \in (0,\varrho_o]$ and $\lambda \ge 1$ such that*

$$\lambda \le \kappa \fint_{Q_\varrho^{(\lambda)}(z_o)} |Du|^{p(\cdot)} + M(|F|+1)^{p(\cdot)}\,dz, \tag{4.1}$$

then we have

$$\lambda \le \Big(\frac{\Gamma}{4\varrho^{n+2}}\Big)^{p_o/2}, \quad p_2 - p_1 \le \omega(\Gamma\varrho^\alpha) \quad \textit{and} \quad \lambda^{\omega(\Gamma\varrho^\alpha)} \le e^{3np_o/\alpha}, \tag{4.2}$$

where

$$p_o := p(z_o), \quad p_1 := \inf_{Q_\varrho^{(\lambda)}(z_o)} p(\cdot), \quad p_2 := \sup_{Q_\varrho^{(\lambda)}(z_o)} p(\cdot)$$

and

$$\Gamma := 4\beta_n\kappa K M, \quad \beta_n := \max\{1,(2\alpha_n)^{-1}\}, \quad \alpha := \min\big\{1, \gamma_1\tfrac{n+2}{4} - \tfrac{n}{2}\big\}. \tag{4.3}$$

Proof. We first deduce from (4.1), (2.9) (recall that $Q_\varrho^{(\lambda)}(z_o) \subset \Omega_T$), and the definitions of Γ and β_n in (4.3) the following bound for λ:

$$\lambda \le \frac{\kappa K M}{|Q_\varrho^{(\lambda)}(z_o)|} = \frac{\kappa K M}{2\,\alpha_n \varrho^{n+2}}\,\lambda^{(p_o-2)/p_o} \le \frac{\beta_n \kappa K M}{\varrho^{n+2}}\,\lambda^{(p_o-2)/p_o} = \frac{\Gamma}{4\varrho^{n+2}}\,\lambda^{(p_o-2)/p_o}.$$

Rewriting this inequality we obtain $(4.2)_1$. Now, we come to the proof of $(4.2)_2$. We define

$$\varrho_o := R_1^{1/\alpha}\,\Gamma^{-2/\alpha} \le R_1 \le 1 \tag{4.4}$$

and assume that $\varrho \le \varrho_o$. Keeping in mind the definitions of α and Γ this determines ϱ_o as a constant depending on n, γ_1, K, M, κ and $\omega(\cdot)$. From (2.3) and the fact that $\lambda \ge 1$ we obtain the following preliminary bound for the oscillation of $p(\cdot)$ on $Q_\varrho^{(\lambda)}(z_o)$:

$$p_2 - p_1 \le \omega\big(2\varrho + \sqrt{2}\,\lambda^{(2-p_o)/(2p_o)}\,\varrho\big) \le \omega\big(2\varrho + \sqrt{2}\,\lambda^{(2-\gamma_1)/(2p_o)}\,\varrho\big)\,.$$

In the case $\gamma_1 \ge 2$ this leads us to

$$p_2 - p_1 \le \omega(4\varrho)\,,$$

while in the case $2n/(n+2) < \gamma_1 < 2$ we infer from $(4.2)_1$ that

$$\begin{aligned} p_2 - p_1 &\le \omega\big(4\lambda^{(2-\gamma_1)/(2p_o)}\varrho\big) \le \omega\big(4\big(\tfrac{\Gamma}{4}\big)^{(2-\gamma_1)/4}\varrho^{1-(2-\gamma_1)(n+2)/4}\big)\\ &\le \omega\big(\Gamma\varrho^{\gamma_1(n+2)/4-n/2}\big). \end{aligned}$$

Note that the restriction $\gamma_1 > 2n/(n+2)$ ensures that the exponent of ϱ is positive, i.e., $\gamma_1(n+2)/4 - n/2 > 0$. Combining the estimates from the cases $\gamma_1 \ge 2$ and $\gamma_1 < 2$ and recalling that $\varrho \le 1$ we arrive at

$$p_2 - p_1 \le \omega(\Gamma\varrho^\alpha)\,,$$

which proves $(4.2)_2$. Finally, we come to the proof of $(4.2)_3$. Using the definition of ϱ_o in (4.4) and the logarithmic bound (2.5) (which is applicable since $R_1/\Gamma \le R_1$) we obtain

$$\Gamma^{\omega(\Gamma\varrho^\alpha)} \le \Gamma^{\omega(\Gamma\varrho_o^\alpha)} \le \Gamma^{\omega(R_1/\Gamma)} \le \big(\tfrac{\Gamma}{R_1}\big)^{\omega(R_1/\Gamma)} = \exp\big[\omega\big(\tfrac{R_1}{\Gamma}\big)\log\big(\tfrac{\Gamma}{R_1}\big)\big] \le \mathrm{e}\,.$$

Moreover, by similar reasoning and using the last inequality we get

$$\begin{aligned} \varrho^{-\omega(\Gamma\varrho^\alpha)} &= \Gamma^{\omega(\Gamma\varrho^\alpha)/\alpha}\,(\Gamma\varrho^\alpha)^{-\omega(\Gamma\varrho^\alpha)/\alpha} \le \mathrm{e}^{\,1/\alpha}\,(\Gamma\varrho^\alpha)^{-\omega(\Gamma\varrho^\alpha)/\alpha}\\ &= \mathrm{e}^{\,1/\alpha}\,\exp\big[\tfrac{\omega(\Gamma\varrho^\alpha)}{\alpha}\log\tfrac{1}{\Gamma\varrho^\alpha}\big] \le \mathrm{e}^{\,2/\alpha}. \end{aligned}$$

At this stage $(4.2)_3$ follows from $(4.2)_1$ and the previous two inequalities since

$$\lambda^{\omega(\Gamma\varrho^\alpha)} \le \big(\Gamma\varrho^{-(n+2)}\big)^{p_o\omega(\Gamma\varrho^\alpha)/2} \le \mathrm{e}^{\,p_o/2+p_o(n+2)/\alpha} \le \mathrm{e}^{\,3np_o/\alpha}\,.$$

This completes the proof of the lemma. □

Since the family of intrinsic cylinders is nonuniform, in the sense that the scaling depends on the center of the cylinder, we need the following nonuniform version of the Vitali covering theorem, which can be found as Lemma 7.1 of [10]. Note that we can choose $L_1 = 1$ due to assumption (2.5) and that we replaced M by KM which is more suitable in our setting.

Lemma 4.2. *Let $K, M, \lambda \geq 1$ and let $p : \Omega_T \to [\gamma_1, \gamma_2]$ fulfill assumptions (2.3) and (2.5). Then there exists $\chi \equiv \chi(n, \gamma_1) \geq 5$ and $\varrho_1 = \varrho_1(n, \gamma_1, K, M) \in (0, 1]$ such that the following is true: let $\mathcal{F} = \{Q_i\}_{i\in\mathcal{I}}$ be a family of axially parallel parabolic cylinders of the form*

$$Q_i \equiv Q^{(\lambda)}_{\varrho_i}(z_i) \equiv B_{\varrho_i}(x_i) \times \big(t_i - \lambda^{(2-p(z_i))/p(z_i)}\varrho_i^2, t_i + \lambda^{(2-p(z_i))/p(z_i)}\varrho_i^2\big)$$

with uniformly bounded radii, in the sense that there holds

$$\varrho_i \leq \min\Big\{\varrho_1, \big[\beta_n K M \lambda^{-2/p(z_i)}\big]^{1/(n+2)}\Big\} \quad \forall\, i \in \mathcal{I} \tag{4.5}$$

with β_n defined in (4.3). Then there exists a countable subcollection $\mathcal{G} \subset \mathcal{F}$ of disjoint parabolic cylinders such that

$$\bigcup_{Q\in\mathcal{F}} Q \subset \bigcup_{Q\in\mathcal{G}} \chi Q.$$

5. Higher integrability

In this section we provide a higher integrability result for solutions to homogeneous parabolic $p(x,t)$-Laplacian systems that will be crucial later in the proof of certain comparison estimates. We consider the parabolic system

$$\partial_t v - \operatorname{div}\big(a(z)|Dv|^{p(z)-2}Dv\big) = 0 \quad \text{on } A \times (t_1, t_2) =: \mathfrak{A}, \tag{5.1}$$

where $A \subset \mathbb{R}^n$ is an open set and $t_1 < t_2$. Then, we have the following higher integrability result from [10] (compare with Theorem 2.2 there), see also [34].

Theorem 5.1. *Suppose that $p\colon \mathfrak{A} \to [\gamma_1, \gamma_2]$ satisfies (2.3) and (2.5) and that $a\colon \mathfrak{A} \to \mathbb{R}$ satisfies (2.6). Then there exists $\varepsilon_o \equiv \varepsilon_o(data) > 0$ such that the following holds: whenever a function $v \in L^2(\mathfrak{A}, \mathbb{R}^N) \cap L^1(t_1, t_2; W^{1,1}(A, \mathbb{R}^N))$ with $Dv \in L^{p(\cdot)}(\mathfrak{A}, \mathbb{R}^{Nn})$ is a weak solution to the parabolic system (5.1) on $\mathfrak{A}$, we have that*

$$Dv \in L^{p(\cdot)(1+\varepsilon_o)}_{\mathrm{loc}}\big(\mathfrak{A}, \mathbb{R}^{Nn}\big). \tag{5.2}$$

Moreover, for any $K \geq 1$ there exists a radius $\varrho_2 \equiv \varrho_2(n, \gamma_1, \gamma_2, K, \omega(\cdot)) \in (0, R_1]$ such that there holds: if

$$\int_{\mathfrak{A}} (|Dv| + 1)^{p(\cdot)}\, dz \leq K \tag{5.3}$$

and $\varepsilon \in (0, \varepsilon_o]$, then for any parabolic cylinder $Q_{2\varrho}(z_o) \subset \mathfrak{A}$ with $\varrho \in (0, \varrho_2]$ we have

$$\fint_{Q_\varrho(z_o)} |Dv|^{p(\cdot)(1+\varepsilon)}\, dz \leq c\Big(\fint_{Q_{2\varrho}(z_o)} |Dv|^{p(\cdot)}\, dz\Big)^{1+\varepsilon d(p(z_o))} + c \tag{5.4}$$

for a constant $c \equiv c(data)$ and with $d(\cdot)$ defined in (2.11).

Note that the quantitative higher integrability estimate (5.4) is nonhomogeneous, in the sense that the exponents of $|Dv|$ on both sides of the inequality are different. In Corollary 5.2 we deduce a homogeneous version of this estimate valid on intrinsic cylinders of the type (5.5). In order to understand that inequality (5.7) is homogeneous, one has to interpret $\lambda \approx \fint |Dv|^{p(\cdot)}\,dz$ in a heuristic sense which will become clear later.

Corollary 5.2. *Let $K, c_*, \hat{c} \geq 1$ and suppose that $p\colon \mathfrak{A} \to [\gamma_1, \gamma_2]$ satisfies (2.3) and (2.5) and that $a\colon \mathfrak{A} \to \mathbb{R}$ fulfills (2.6). Then there exist $\varepsilon_o \equiv \varepsilon_o(data) > 0$, $c \equiv c(data, c_*, \hat{c}) \geq 1$, and $\varrho_2 \equiv \varrho_2(n, \gamma_1, \gamma_2, K, \omega(\cdot)) \in (0, R_1]$ such that the following holds: whenever $v \in L^2(\mathfrak{A}, \mathbb{R}^N) \cap L^1(t_1, t_2; W^{1,1}(A, \mathbb{R}^N))$ with $Dv \in L^{p(\cdot)}(\mathfrak{A}, \mathbb{R}^{Nn})$ is a weak solution to the parabolic system (5.1) satisfying (5.3) and*

$$\fint_{Q_{2\varrho}^{(\lambda)}(z_o)} |Dv|^{p(\cdot)}\,dz \leq c_* \lambda \tag{5.5}$$

for some cylinder $Q_{2\varrho}^{(\lambda)}(z_o) \subset \mathfrak{A}$ with $\varrho \in (0, \varrho_2]$ and $\lambda \geq 1$ satisfying

$$\lambda^{p_2 - p_1} \leq \hat{c}, \quad \textit{where} \quad p_1 := \inf_{Q_{2\varrho}^{(\lambda)}(z_o)} p(\cdot), \quad p_2 := \sup_{Q_{2\varrho}^{(\lambda)}(z_o)} p(\cdot), \tag{5.6}$$

then we have (5.2) and

$$\fint_{Q_{\varrho}^{(\lambda)}(z_o)} |Dv|^{p(\cdot)(1+\varepsilon_o)}\,dz \leq c\,\lambda^{1+\varepsilon_o}. \tag{5.7}$$

Proof. Without loss of generality we assume that $z_o = 0$. We let ε_o and ϱ_2 be the constants appearing in Theorem 5.1. The strategy now is to rescale the problem from $Q_\varrho^{(\lambda)}$ and $Q_{2\varrho}^{(\lambda)}$ to the standard parabolic cylinders Q_ϱ and $Q_{2\varrho}$ via a transformation in time and then apply Theorem 5.1. We start with the case $p_o := p(0) \geq 2$ and define for $(x,t) \in Q_{2\varrho}$ the rescaled exponent

$$\tilde{p}(x,t) := p\big(x, \lambda^{(2-p_o)/p_o}\, t\big),$$

the rescaled function

$$\tilde{v}(x,t) := \lambda^{-1/p_o}\, v\big(x, \lambda^{(2-p_o)/p_o}\, t\big)$$

and the rescaled coefficient

$$\tilde{a}(x,t) := \lambda^{(\tilde{p}(x,t)-p_o)/p_o}\, a\big(x, \lambda^{(2-p_o)/p_o}\, t\big).$$

Then, $\tilde{v}$ is a weak solution of the parabolic system

$$\partial_t \tilde{v} - \operatorname{div}\big(\tilde{a}(\cdot)|D\tilde{v}|^{\tilde{p}(\cdot)-2} D\tilde{v}\big) = 0 \quad \text{in } Q_{2\varrho}. \tag{5.8}$$

In order to apply the higher integrability Theorem 5.1 to $\tilde{v}$ we have to ensure that the hypotheses on $\tilde{p}$ and $\tilde{a}$ are satisfied. Since $p_o \geq 2$ and $\lambda \geq 1$ we have

$$\begin{aligned} |\tilde{p}(x_1,t_1) - \tilde{p}(x_2,t_2)| &= \big|p\big(x_1, \lambda^{(2-p_o)/p_o}\, t_1\big) - p\big(x_2, \lambda^{(2-p_o)/p_o}\, t_2\big)\big| \\ &\leq \omega\big(\max\big\{|x_1 - x_2|, \lambda^{(2-p_o)/(2p_o)}\sqrt{|t_1 - t_2|}\big\}\big) \\ &\leq \omega\big(\max\big\{|x_1 - x_2|, \sqrt{|t_1 - t_2|}\big\}\big) = \omega\big(d_{\mathcal{P}}\big((x_1,t_1),(x_2,t_2)\big)\big). \end{aligned} \tag{5.9}$$

Moreover by (2.6) and (5.6) there holds

$$\nu/\hat{c} \le \nu\,\lambda^{-(p_2-p_1)/p_o} \le \tilde{a}(x,t) \le L\,\lambda^{(p_2-p_1)/p_o} \le \hat{c}\,L. \tag{5.10}$$

Therefore, we are allowed to apply Theorem 5.1 with $(\nu/\hat{c}, \hat{c}\,L)$ instead of (ν, L) to the function $\tilde{v}$ on $Q_\varrho, Q_{2\varrho}$ to infer that $D\tilde{v} \in L^{p(\cdot)(1+\varepsilon_o)}_{\mathrm{loc}}(Q_{2\varrho}, \mathbb{R}^{Nn})$ and, moreover, the quantitative estimate

$$\fint_{Q_\varrho} |D\tilde{v}|^{\tilde{p}(\cdot)(1+\varepsilon_o)}\,dz \le c\Big(\fint_{Q_{2\varrho}} |D\tilde{v}|^{\tilde{p}(\cdot)}\,dz\Big)^{1+\varepsilon_o d(p_o)} + c$$

holds for a constant $c \equiv c(data)$. Note that $p_o = p(0) = \tilde{p}(0)$. Scaling from v to $\tilde{v}$ and back and using the preceding estimate, (5.5), and (5.6) several times we find that

$$\begin{aligned}
\fint_{Q_\varrho^{(\lambda)}} |Dv|^{p(\cdot)(1+\varepsilon_o)}\,dz &= \fint_{Q_\varrho} \lambda^{\frac{\tilde{p}(\cdot)}{p_o}(1+\varepsilon_o)}|D\tilde{v}|^{\tilde{p}(\cdot)(1+\varepsilon_o)}\,dz \\
&\le c\,\lambda^{1+\varepsilon_o} \fint_{Q_\varrho} |D\tilde{v}|^{\tilde{p}(\cdot)(1+\varepsilon_o)}\,dz \\
&\le c\,\lambda^{1+\varepsilon_o}\Big(\fint_{Q_{2\varrho}} |D\tilde{v}|^{\tilde{p}(\cdot)}\,dz\Big)^{1+\varepsilon_o d(p_o)} + c\,\lambda^{1+\varepsilon_o} \\
&= c\,\lambda^{1+\varepsilon_o}\Big(\fint_{Q_{2\varrho}^{(\lambda)}} \lambda^{-p(\cdot)/p_o}\,|Dv|^{p(\cdot)}\,dz\Big)^{1+\varepsilon_o d(p_o)} + c\,\lambda^{1+\varepsilon_o} \\
&\le c\,\lambda^{\varepsilon_o(1-d(p_o))}\Big(\fint_{Q_{2\varrho}^{(\lambda)}} |Dv|^{p(\cdot)}\,dz\Big)^{1+\varepsilon_o d(p_o)} + c\,\lambda^{1+\varepsilon_o} \\
&\le c\,(data, c_*, \hat{c})\,\lambda^{1+\varepsilon_o}.
\end{aligned} \tag{5.11}$$

This proves the lemma in the case $p_o \ge 2$. In the case $p_o < 2$ we define similarly to the preceding

$$\tilde{p}(x,t) := p\big(\lambda^{(p_o-2)/(2p_o)}\,x, t\big), \quad \tilde{v}(x,t) := \lambda^{-1/2}\,v\big(\lambda^{(p_o-2)/(2p_o)}\,x, t\big)$$

and

$$\tilde{a}(x,t) := \lambda^{(\tilde{p}(x,t)-p_o)/p_o}\,a\big(\lambda^{(p_o-2)/(2p_o)}\,x, t\big) \tag{5.12}$$

for $(x,t) \in Q_{2\tilde{\varrho}}$, where $\tilde{\varrho} := \lambda^{(2-p_o)/(2p_o)}\varrho$. A straightforward computation shows that

$$D\tilde{v}(x,t) = \lambda^{-1/p_o}\,Dv\big(\lambda^{(p_o-2)/(2p_o)}\,x, t\big) \quad \text{in } Q_{2\tilde{\varrho}},$$

and that $\tilde{v}$ is a weak solution of the system (5.8) in $Q_{2\tilde{\varrho}}$, where $\tilde{a}$, $\tilde{v}$ and $\tilde{p}$ are this time the quantities defined just above. Notice that the estimate (5.10) holds also for the vector field defined in (5.12), while the verification of (5.9) in this case is analogous to the verification in the previous case. Applying again Theorem 5.1 and repeating the computations in (5.11) we obtain the assertion of the lemma also in the case $p_o < 2$. □

6. A priori estimates

In Theorem 6.1 we state the gradient bound of DiBenedetto and Friedman ([17] and [18]) for parabolic standard growth problems. Later we will transfer these a priori estimates via a comparison argument to our nonstandard growth problem. Therefore, in this section we consider parabolic systems with constant p-growth of the type

$$w_t - \operatorname{div}\big(\tilde{a}(t)|Dw|^{p-2}Dw\big) = 0 \quad \text{on } A \times (t_1, t_2) =: \mathfrak{A}, \tag{6.1}$$

with $p > 2n/(n+2)$ and $\tilde{a}\colon (t_1, t_2) \to \mathbb{R}$. Here A is an open set in $\mathbb{R}^n$ and $t_1 < t_2$. Moreover, we define

$$C_\varrho^{(\lambda)}(z_o) := B_\varrho(x_o) \times \big(t_o - \lambda^{(2-p)/p}\,\varrho^2, t_o + \lambda^{(2-p)/p}\,\varrho^2\big).$$

Note that the scaling of this system of cylinders does not depend on the center z_o. Later we will apply Theorem 6.1 with the choice $p = p_o \equiv p(z_o)$, and hence the cylinders $C_\varrho^{(\lambda)}(z_o)$ will coincide with the ones defined in (3.1). As mentioned above, the next theorem is a consequence of the gradient bounds proved in [17] and [18]. The precise statement for the case $p \geq 2$ can be found in Lemma 1 of [4] (replacing λ by $\lambda^{1/p}$), and for the case $2n/(n+2) < p < 2$ in Lemma 2 of [4] (replacing ϱ by $\lambda^{(p-2)/2}\varrho$ and subsequently λ by $\lambda^{1/p}$).

Theorem 6.1. *Let $w \in C^0(t_1, t_2; L^2(A, \mathbb{R}^N)) \cap L^p(t_1, t_2; W^{1,p}(A, \mathbb{R}^N))$ be a weak solution to (6.1) in $\mathfrak{A}$ with $\tilde{a}\colon (t_1, t_2) \to \mathbb{R}$ satisfying $\nu \leq \tilde{a} \leq L$ for some constants $0 < \nu \leq 1 \leq L$. Moreover suppose that*

$$\fint_{C_{2\varrho}^{(\lambda)}(z_o)} |Dw|^p \, dz \leq c_* \, \lambda$$

holds for some cylinder $C_{2\varrho}^{(\lambda)}(z_o) \Subset \mathfrak{A}$, where c_* is a given positive constant. Then there exists a constant $c_{\mathrm{DiB}} \geq 1$, depending on n, N, p, ν, L, and c_* such that

$$\sup_{C_\varrho^{(\lambda)}(z_o)} |Dw| \leq c_{\mathrm{DiB}} \, \lambda^{1/p}.$$

7. Comparison estimates

In this section we prove two different comparison estimates. The first compares the weak solution u of the original inhomogeneous parabolic system (1.1) to the solution v of the associated homogeneous parabolic system (7.2) below. The second compares v to the solution w of the frozen parabolic system (7.14). Both the parabolic localization Lemma 4.1 and the homogeneous version of the higher integrability estimate from Corollary 5.2 will be crucial in order to achieve homogeneous comparison estimates.

Now we let $K \geq 1$ and suppose that (2.9) is satisfied. Next, we fix κ and $M \geq 1$ to be specified later. In the following we consider a cylinder $Q := Q_\varrho^{(\lambda)}(z_o)$ with

$z_o = (x_o, t_o) \in \Omega_T$, $\varrho \in (0,1]$, and $\lambda \geq 1$ defined according to (3.1) and which satisfies $2Q := Q^{(\lambda)}_{2\varrho}(z_o) \Subset \Omega_T$ and

$$\frac{\lambda}{\kappa} \leq \fint_{2Q} |Du|^{p(\cdot)}\,dz + \fint_{2Q} M(|F|+1)^{p(\cdot)}\,dz \leq \lambda. \tag{7.1}$$

Moreover, we abbreviate $B := B_\varrho(x_o)$ and $\Lambda := \Lambda^{(\lambda)}_\varrho(t_o)$ so that $Q \equiv B \times \Lambda$ and define

$$p_o := p(z_o), \quad p_1 := \inf_{2Q} p(\cdot) \quad \text{and} \quad p_2 := \sup_{2Q} p(\cdot).$$

By $v \in L^2(2Q, \mathbb{R}^N) \cap L^1(2\Lambda; W^{1,1}(2B, \mathbb{R}^N))$ with $Dv \in L^{p(\cdot)}(2Q, \mathbb{R}^{Nn})$ we denote the unique solution of the homogeneous initial-boundary value problem

$$\begin{cases} \partial_t v - \operatorname{div}\big(a(z)|Dv|^{p(z)-2}Dv\big) = 0 & \text{in } 2Q, \\ v = u & \text{on } \partial_{\mathcal{P}} 2Q. \end{cases} \tag{7.2}$$

Here the parabolic boundary $\partial_{\mathcal{P}} 2Q$ is given by

$$\partial_{\mathcal{P}} 2Q := (\partial 2B \times 2\Lambda) \cup (\overline{2B} \times \{t_o - \lambda^{2-p_o}(2\varrho)^2\}).$$

Note that the existence of v can be inferred from small modifications of results in [8]. Our first aim is to prove suitable energy and comparison estimates for the comparison function v. Hence we subtract the weak formulation of the parabolic system (7.2) from the one of (1.1) given in (2.1). This yields

$$\int_{2Q} (u-v)\cdot \partial_t \varphi\,dz - \int_{2Q} a(\cdot)\big\langle |Du|^{p(\cdot)-2}Du - |Dv|^{p(\cdot)-2}Dv, D\varphi\big\rangle\,dz = \int_{2Q} \big\langle |F|^{p(\cdot)-2}F, D\varphi\big\rangle\,dz$$

for any $\varphi \in C^\infty_0(2Q, \mathbb{R}^N)$. For $\theta > 0$ and $\tau := t_o + \lambda^{(2-p_o)/p_o}\,(2\varrho)^2$ we define

$$\chi_\theta(t) := \begin{cases} 1 & \text{on } (-\infty, \tau-\theta], \\ -\frac{1}{\theta}(t-\tau) & \text{on } (\tau-\theta, \tau), \\ 0 & \text{on } [\tau, \infty). \end{cases} \tag{7.3}$$

Since $Du - Dv \in L^{p(\cdot)}(2Q, \mathbb{R}^{Nn})$ and $u = v$ on $\partial_{\mathcal{P}} 2Q$ in the sense of traces, we are (formally) allowed to choose $\varphi = (u-v)\chi_\theta$ in the preceding identity. We note that the argument can be made rigorous via the use of Steklov averages and an approximation argument; since this is standard we omit the details. This choice of φ together with the observation that

$$\begin{aligned} \int_{2Q} (u-v)\cdot \partial_t[(u-v)\chi_\theta]\,dz &= -\int_{2Q} \partial_t (u-v)\cdot (u-v)\chi_\theta\,dz \\ &= -\frac{1}{2}\int_{2Q} \partial_t |u-v|^2 \chi_\theta\,dz = \frac{1}{2}\int_{2Q} |u-v|^2 \partial_t \chi_\theta\,dz \\ &= -\frac{1}{2\theta}\int_{\tau-\theta}^{\tau}\int_{2B} |u-v|^2\,dz \overset{\theta\downarrow 0}{\to} -\frac{1}{2}\int_{2B} |u-v|^2(\cdot,\tau)\,dx \leq 0 \end{aligned} \tag{7.4}$$

leads, after letting $\theta \downarrow 0$, to

$$\int_{2Q} a(\cdot)\big\langle |Du|^{p(\cdot)-2}Du - |Dv|^{p(\cdot)-2}Dv, D(u-v)\big\rangle \, dz \qquad (7.5)$$
$$\leq - \int_{2Q} \big\langle |F|^{p(\cdot)-2}F, D(u-v)\big\rangle \, dz.$$

This inequality will be used in the following in two different directions. The first one will lead to an energy inequality for Dv. Rearranging terms and taking into account that $\nu \leq a(\cdot) \leq L$ we find

$$\begin{aligned}\nu \int_{2Q} |Dv|^{p(\cdot)} \, dz &\leq L \int_{2Q} \big(|Du|^{p(\cdot)-1}|Dv| + |Dv|^{p(\cdot)-1}|Du|\big) \, dz \\ &\quad + \int_{2Q} |F|^{p(\cdot)-1}\big(|Du| + |Dv|\big) \, dz \\ &\leq \frac{\nu}{2} \int_{2Q} |Dv|^{p(\cdot)} \, dz + c \int_{2Q} \big(|Du|^{p(\cdot)} + |F|^{p(\cdot)}\big) \, dz,\end{aligned}$$

where in the last line we applied Young's inequality and $c \equiv c(\nu, L, \gamma_1, \gamma_2)$. Absorbing the first integral of the right-hand side into the left and subsequently using (7.1) we get the following *energy estimate for* Dv:

$$\int_{2Q} |Dv|^{p(\cdot)} \, dz \leq c \int_{2Q} \big(|Du|^{p(\cdot)} + |F|^{p(\cdot)}\big) \, dz \leq c(data) \, \lambda |Q|. \qquad (7.6)$$

We now come to the proof of the comparison estimate. Starting again from (7.5) we use Lemma 3.2 and Young's inequality to infer

$$\begin{aligned}\nu \int_{2Q} \big(|Du|^2 + |Dv|^2\big)^{(p(\cdot)-2)/2} |Du - Dv|^2 \, dz &\leq c \int_{2Q} |F|^{p(\cdot)-1}\big(|Du| + |Dv|\big) \, dz \\ &\leq c \, M^{-(\gamma_1-1)/\gamma_1} \int_{2Q} \big(|Du|^{p(\cdot)} + |Dv|^{p(\cdot)} + M|F|^{p(\cdot)}\big) \, dz,\end{aligned}$$

where the constant c depends on n, N, L, γ_1 and γ_2. Finally, using (7.1) and the energy estimate (7.6) this leads us to the *first comparison estimate* we were looking for:

$$\int_{2Q} \big(|Du|^2 + |Dv|^2\big)^{(p(\cdot)-2)/2} |Du - Dv|^2 \, dz \leq c(data) \, M^{-(\gamma_1-1)/\gamma_1} \, \lambda \, |Q|. \qquad (7.7)$$

Now we let $\varepsilon_o = \varepsilon_o(data) > 0$ be the higher integrability exponent from Corollary 5.2 and set

$$\varrho_3 := \min\{\varrho_o/2, \varrho_2\} \in (0, 1],$$

where ϱ_o is the radius appearing in the localization Lemma 4.1 and ϱ_2 the radius for the higher integrability from Corollary 5.2. Note that ϱ_3 depends on $data, \kappa, K, M, \omega(\cdot)$. In the course of the proof we shall further decrease the value of ϱ_3 when necessary, but without changing its dependencies. In the following we assume that

$$\varrho \le \varrho_3.$$

Thanks to assumption (7.1) we can apply Lemma 4.1 on $2Q$ which yields that

$$p_2 - p_1 \le \omega\big(\Gamma(2\varrho)^\alpha\big) \quad \text{and} \quad \lambda^{p_2-p_1} \le \lambda^{\omega(\Gamma(2\varrho)^\alpha)} \le \mathrm{e}^{3np_o/\alpha} \le \mathrm{e}^{3n\gamma_2/\alpha}, \tag{7.8}$$

where Γ and α are defined according to (4.3). Note that for the second estimate we also used that $\lambda \ge 1$. Therefore, assumption (5.6) of Corollary 5.2 is satisfied with $\hat{c} \equiv \hat{c}(n,\gamma_1,\gamma_2) := \mathrm{e}^{3n\gamma_2/\alpha}$. Due to the energy estimate (7.6) we know that also assumption (5.5) is satisfied with c_* replaced by the constant $c(data)$ from (7.6). The application of Corollary 5.2 then ensures that $Dv \in L^{p(\cdot)(1+\varepsilon_o)}(Q,\mathbb{R}^{Nn})$ and moreover

$$⨍_Q |Dv|^{p(\cdot)(1+\varepsilon_o)}\,dz \le c(data)\,\lambda^{1+\varepsilon_o}. \tag{7.9}$$

Next, we decrease the value of ϱ_3 in such a way that

$$\omega\big(\Gamma(2\varrho_3)^\alpha\big) \le \frac{\varepsilon_1}{\gamma_1'}, \quad \text{where } \varepsilon_1 := \sqrt{1+\varepsilon_o} - 1 \le \varepsilon_o \tag{7.10}$$

is satisfied. Then, by (7.8), for any $z \in 2Q$ there holds

$$\begin{aligned} p_o(1+\varepsilon_1) &\le p(z)\big(1+\omega\big(\Gamma(2\varrho)^\alpha\big)\big)(1+\varepsilon_1) \le p(z)\big(1+\omega\big(\Gamma(2\varrho_3)^\alpha\big)\big)(1+\varepsilon_1) \\ &< p(z)(1+\varepsilon_1)^2 = p(z)(1+\varepsilon_o), \end{aligned}$$

and therefore we have $Dv \in L^{p_o(1+\varepsilon_1)}(Q,\mathbb{R}^{Nn})$ together with the estimate

$$\begin{aligned} ⨍_Q |Dv|^{p_o(1+\varepsilon_1)}\,dz &\le ⨍_Q |Dv|^{p(\cdot)(1+\omega(\Gamma(2\varrho)^\alpha))(1+\varepsilon_1)}\,dz + 1 \\ &\le \Big(⨍_Q |Dv|^{p(\cdot)(1+\varepsilon_o)}\,dz\Big)^{(1+\omega(\Gamma(2\varrho)^\alpha))(1+\varepsilon_1)/(1+\varepsilon_o)} + 1 \\ &\le c\,\lambda^{(1+\omega(\Gamma(2\varrho)^\alpha))(1+\varepsilon_1)} + 1 = c\,\lambda^{1+\varepsilon_1}\lambda^{\omega(\Gamma(2\varrho)^\alpha)(1+\varepsilon_1)} + 1 \\ &\le c(data)\,\lambda^{1+\varepsilon_1}, \end{aligned} \tag{7.11}$$

where we used Hölder's inequality, (7.9), (7.8), and the fact that $\lambda \ge 1$. For later reference we also provide the following estimate using (7.8) and (7.10):

$$\begin{aligned} p_o'(p_2-1) &= p_o\Big(1+\frac{p_2-p_o}{p_o-1}\Big) \le p_o\Big(1+\frac{\omega(\Gamma(2\varrho_3)^\alpha)}{\gamma_1-1}\Big) \\ &\le p_o\big(1+\tfrac{\varepsilon_1}{\gamma_1}\big) < p_o(1+\varepsilon_1). \end{aligned} \tag{7.12}$$

Together with (7.11), Hölder's inequality, and (7.8) this implies

$$\begin{aligned} ⨍_Q |Dv|^{p_o'(p_2-1)}\,dz &\le \Big(⨍_Q |Dv|^{p_o(1+\varepsilon_1)}\,dz\Big)^{\frac{p_2-1}{(p_o-1)(1+\varepsilon_1)}} \\ &\le c\,\lambda^{(p_2-1)/(p_o-1)} = c\,\lambda^{1+(p_2-p_o)/(p_o-1)} \le c(data)\,\lambda. \end{aligned} \tag{7.13}$$

We now define

$$\tilde{a}(t) := (a)_{x_o,\varrho}(t) := \fint_{B_\varrho(x_o)} a(\cdot,t)\,dx \quad \text{for any } t \in (0,T).$$

Note that $\nu \le \tilde{a}(t) \le L$ for any $t \in (0,T)$ as a consequence of (2.6). By

$$w \in C^0\big(\Lambda, L^2(B;\mathbb{R}^N)\big) \cap L^{p_o}\big(\Lambda, W^{1,p_o}(B;\mathbb{R}^N)\big)$$

we denote the unique solution to the initial-boundary value problem

$$\begin{cases} \partial_t w - \operatorname{div}\big(\tilde{a}(t)|Dw|^{p_o-2}Dw\big) = 0 & \text{in } Q, \\ w = v & \text{on } \partial_{\mathcal{P}}Q. \end{cases} \tag{7.14}$$

Now we start deriving energy and comparison estimates for w. As before, we subtract the weak formulations of (7.2) and (7.14) and test the result with $\varphi := (v-w)\chi_\theta$, where χ_θ is defined in (7.3). Here we recall that $Dv \in L^{p_o}(Q,\mathbb{R}^{Nn})$ by (7.11) and therefore φ is (formally) admissible as a test function. Proceeding as before, i.e., treating the terms involving the time derivatives with the argument performed in (7.4) and passing to the limit $\theta \downarrow 0$ we obtain

$$\int_Q \big\langle a(\cdot)|Dv|^{p(\cdot)-2}Dv - \tilde{a}(t)|Dw|^{p_o-2}Dw, D(v-w)\big\rangle\,dz \le 0. \tag{7.15}$$

First, we shall use this inequality to get an energy estimate for Dw. Rearranging terms, taking into account that $\nu \le a(\cdot) \le L$ and $\nu \le \tilde{a}(\cdot) \le L$, and also applying Young's inequality we find

$$\begin{aligned} \nu \fint_Q |Dw|^{p_o}\,dz &\le L \fint_Q \big(|Dw|^{p_o-1}|Dv| + |Dv|^{p(\cdot)-1}|Dw|\big)\,dz \\ &\le \frac{\nu}{2}\fint_Q |Dw|^{p_o}\,dz + c\fint_Q \big(|Dv|^{p_o} + |Dv|^{p_o'(p(\cdot)-1)}\big)\,dz \end{aligned}$$

with a constant $c \equiv c(\nu, L, \gamma_1, \gamma_2)$. Absorbing the first integral of the right-hand side into the left and using Hölder's inequality, (7.11), (7.13), and the fact that $\lambda \ge 1$ we get the following *energy estimate for Dw*:

$$\fint_Q |Dw|^{p_o}\,dz \le c\Big[\fint_Q |Dv|^{p_o}\,dz + \fint_Q |Dv|^{p_o'(p_2-1)}\,dz + 1\Big] \le c(data)\,\lambda. \tag{7.16}$$

In order to obtain a comparison estimate we once again start from (7.15) which can be rewritten as follows:

$$\begin{aligned} &\fint_Q \tilde{a}(t)\big\langle |Dv|^{p_o-2}Dv - |Dw|^{p_o-2}Dw, D(v-w)\big\rangle\,dz \\ &\qquad \le \fint_Q \big(\tilde{a}(t) - a(\cdot)\big)\big\langle |Dv|^{p_o-2}Dv, D(v-w)\big\rangle\,dz \\ &\qquad\quad + \fint_Q a(\cdot)\big\langle |Dv|^{p_o-2}Dv - |Dv|^{p(\cdot)-2}Dv, D(v-w)\big\rangle\,dz. \end{aligned}$$

Using Lemma 3.2 and the fact that $\nu \le \tilde{a}(\cdot) \le L$ we obtain

$$\begin{aligned}
&\fint_Q \big(|Dv|^2 + |Dw|^2\big)^{(p_o-2)/2}|Dv - Dw|^2\,dz \\
&\qquad\le c\Big[\fint_Q |\tilde{a}(t) - a(\cdot)||Dv|^{p_o-1}|Dv - Dw|\,dz \\
&\qquad\quad + \fint_Q \big||Dv|^{p_o-1} - |Dv|^{p(\cdot)-1}\big||Dv - Dw|\,dz\Big] =: c\,[\mathrm{I} + \mathrm{II}],
\end{aligned} \tag{7.17}$$

where $c \equiv c(\nu, L, \gamma_1, \gamma_2)$. Now we estimate separately the two terms. For the first we use Hölder's inequality several times, (7.11), (7.16), the fact that $a(\cdot) \le L$, $\tilde{a}(t) \le L$, (2.7), and $\tilde{\omega} \le 1$ to infer that

$$\begin{aligned}
\mathrm{I} &\le c\Big(\fint_Q |a(t) - a(\cdot)|^{p_o'}|Dv|^{p_o}\,dz\Big)^{1/p_o'}\Big(\fint_Q \big(|Dv|^{p_o} + |Dw|^{p_o}\big)\,dz\Big)^{1/p_o} \\
&\le c\Big(\fint_Q |a(t) - a(\cdot)|^{p_o'(1+\varepsilon_1)/\varepsilon_1}\,dz\Big)^{\frac{\varepsilon_1}{p_o'(1+\varepsilon_1)}}\Big(\fint_Q |Dv|^{p_o(1+\varepsilon_1)}\,dz\Big)^{\frac{1}{p_o'(1+\varepsilon_1)}}\lambda^{1/p_o} \\
&\le c\,\lambda^{1/p_o'+1/p_o}\big[\tilde{\omega}(\varrho)\big]^{\frac{\varepsilon_1}{\gamma_1'(1+\varepsilon_1)}} \le c(data)\big[\tilde{\omega}(\varrho)\big]^{\varepsilon_1/(2\gamma_1')}\,\lambda.
\end{aligned}$$

In order to estimate II we first use (7.8) to find that for any $z \in Q$ and $b \ge 0$ there holds

$$\begin{aligned}
|b^{p_o-1} - b^{p(z)-1}| &\le |p_o - p(z)| \sup_{\sigma\in[p_1-1,p_2-1]} b^\sigma|\log b| \\
&\le \omega\big(\Gamma(2\varrho)^\alpha\big)\Big[b^{p_2-1}\log\big(\mathrm{e} + b^{p_o'(p_2-1)}\big) + \tfrac{1}{e(\gamma_1-1)}\Big],
\end{aligned}$$

where in the last line we used $b^\sigma|\log b| \le \frac{1}{\mathrm{e}\,(\gamma_1-1)}$ for $b \in [0,1]$ and $\sigma \in [p_1-1, p_2-1]$ and $b^\sigma|\log b| \le b^{p_2-1}\log(\mathrm{e} + b^{p_o'(p_2-1)})$ for $b > 1$ and $\sigma \in [p_1-1, p_2-1]$. This together with Hölder's inequality, (7.11) and (7.16) yields

$$\begin{aligned}
\mathrm{II} &\le c\,\omega\big(\Gamma(2\varrho)^\alpha\big)\fint_Q \Big[|Dv|^{p_2-1}\log\big(\mathrm{e} + |Dv|^{p_o'(p_2-1)}\big) + 1\Big]|Dv - Dw|\,dz \\
&\le c\,\omega\big(\Gamma(2\varrho)^\alpha\big)\Big(\fint_Q \Big[|Dv|^{p_2-1}\log\big(\mathrm{e} + |Dv|^{p_o'(p_2-1)}\big) + 1\Big]^{p_o'}dz\Big)^{1/p_o'} \\
&\quad\cdot\Big(\fint_Q |Dv - Dw|^{p_o}\,dz\Big)^{1/p_o} \\
&\le c\,\omega\big(\Gamma(2\varrho)^\alpha\big)\Big(\fint_Q \Big[|Dv|^{p_2-1}\log\big(\mathrm{e} + |Dv|^{p_o'(p_2-1)}\big) + 1\Big]^{p_o'}dz\Big)^{1/p_o'}\lambda^{1/p_o},
\end{aligned}$$

where $c \equiv c(data)$. Next, we note that

$$\log(\mathrm{e} + ab) \le \log(\mathrm{e} + a) + \log(\mathrm{e} + b) \qquad \forall\, a, b \ge 0,$$

which together with Young's inequality allows to further estimate II as

$$\begin{aligned} \mathrm{II} &\le c\,\omega\big(\Gamma(2\varrho)^\alpha\big)\,\lambda^{1/p_o}\Big[\fint_Q |Dv|^{p_o'(p_2-1)}\log^{p_o'}\Big(\mathrm{e}+\frac{|Dv|^{p_o'(p_2-1)}}{\big(|Dv|^{p_o'(p_2-1)}\big)_Q}\Big)\,dz \\ &\qquad + \log^{p_o'}\Big(\mathrm{e}+\big(|Dv|^{p_o'(p_2-1)}\big)_Q\Big)\fint_Q |Dv|^{p_o'(p_2-1)}\,dz+1\Big]^{1/p_o'} \\ &= c(data)\,\omega\big(\Gamma(2\varrho)^\alpha\big)\,\lambda^{1/p_o}\big[\mathrm{II}_1+\mathrm{II}_2+1\big]^{1/p_o'}, \end{aligned} \tag{7.18}$$

with the obvious meanings for II_1 and II_2. In order to estimate II_1 we apply inequality (3.2) with the choices $g=|Dv|^{p_o'(p_2-1)}$ and

$$\sigma := \frac{1+\varepsilon_1}{1+\varepsilon_1/\gamma_1} = c(data) > 1$$

to infer that

$$\mathrm{II}_1 \le c(data)\Big(\fint_Q |Dv|^{p_o'(p_2-1)\sigma}\,dz\Big)^{1/\sigma}.$$

To the integral on the right-hand side we apply Hölder's inequality (this is justified by (7.12)). Subsequently using (7.11) and (7.8) we obtain

$$\begin{aligned} \mathrm{II}_1 &\le c\Big(\fint_Q |Dv|^{p_o(1+\varepsilon_1/\gamma_1)\sigma}\,dz\Big)^{\frac{1}{\sigma}\cdot\frac{p_2-1}{(p_o-1)(1+\varepsilon_1/\gamma_1)}} = c\Big(\fint_Q |Dv|^{p_o(1+\varepsilon_1)}\,dz\Big)^{\frac{p_2-1}{(p_o-1)(1+\varepsilon_1)}} \\ &\le c\,\lambda^{(p_2-1)/(p_o-1)} = c\,\lambda^{1+(p_2-p_o)/(p_o-1)} \le c(data)\,\lambda. \end{aligned}$$

Now, we come to the estimate for II_2 in (7.18). From (7.13) and $(4.2)_1$ we get

$$\big(|Dv|^{p_o'(p_2-1)}\big)_Q = \fint_Q |Dv|^{p_o'(p_2-1)}\,dz \le c\,\lambda \le c(data,\kappa)\Big(\frac{KM}{\varrho^{n+2}}\Big)^{p_o/2}.$$

Using this estimate, again (7.13), the fact that $\log(cx)\le c\log(x)$ for $c\ge 1$, and that we can always assume $c(KM/\varrho^{n+2})^{p_o/2}\ge \mathrm{e}$ by possibly reducing the value of ϱ_3, we find

$$\mathrm{II}_2 \le \log^{p_o'}\Big(\mathrm{e}+c\Big(\frac{KM}{\varrho^{n+2}}\Big)^{p_o/2}\Big)\fint_Q |Dv|^{p_o'(p_2-1)}\,dz \le c\,M^{p_o'}\log^{p_o'}\Big(\frac{K}{\varrho}\Big)\,\lambda$$

with $c\equiv c(data,\kappa)$. Combining the estimates for II_1 and II_2 with (7.18) we obtain

$$\mathrm{II} \le c(data,\kappa)\,\omega\big(\Gamma(2\varrho)^\alpha\big)\,M\log\Big(\frac{K}{\varrho}\Big)\,\lambda.$$

Substituting the preceding estimates for I and II into (7.17) we get

$$\begin{aligned} &\fint_Q \big(|Dv|^2+|Dw|^2\big)^{(p_o-2)/2}|Dv-Dw|^2 dz \\ &\qquad\le c(data,\kappa)\Big[\omega\big(\Gamma(2\varrho)^\alpha\big)M\log\Big(\frac{K}{\varrho}\Big)+\big[\tilde\omega(\varrho)\big]^{\varepsilon_1/(2\gamma_1')}\Big]\lambda. \end{aligned}$$

Here, we still want to replace the exponent p_o in the integral on the left-hand side by $p(\cdot)$. This is achieved with the help of Hölder's inequality as follows:

$$\begin{aligned}
&\fint_{\frac12 Q} \big(|Dv|^2+|Dw|^2\big)^{(p(\cdot)-2)/2}|Dv-Dw|^2\,dz \\
&\qquad\le \Big(\fint_{\frac12 Q} \big(|Dv|^2+|Dw|^2\big)^{(p_o-2)/2}|Dv-Dw|^2\,dz\Big)^{1/2} \\
&\qquad\quad\cdot \Big(\fint_{\frac12 Q} \big(|Dv|^2+|Dw|^2\big)^{(2p(\cdot)-p_o-2)/2}|Dv-Dw|^2\,dz\Big)^{1/2} \\
&\qquad\le c\Big[\omega\big(\Gamma(2\varrho)^\alpha\big)M\log\Big(\frac{K}{\varrho}\Big)+\big[\tilde\omega(\varrho)\big]^{\varepsilon_1/(2\gamma_1')}\Big]^{1/2}\lambda^{1/2} \\
&\qquad\quad\cdot \Big(\fint_{\frac12 Q} |Dv|^{2p(\cdot)-p_o}+|Dw|^{2p(\cdot)-p_o}\,dz\Big)^{1/2}.
\end{aligned} \tag{7.19}$$

In order to further estimate the integral on the right-hand side we use that fact that $2p(\cdot)-p_o\le p(\cdot)(1+\omega(\Gamma(2\varrho)^\alpha))\le p(\cdot)(1+\omega(\Gamma\varrho_3^\alpha))\le p(\cdot)(1+\varepsilon_o)$ which is a consequence of (7.10), Hölder's inequality, (7.9), (7.8), and $\lambda\ge 1$ to infer that

$$\begin{aligned}
\fint_{\frac12 Q} |Dv|^{2p(\cdot)-p_o}\,dz &\le 2^{n+2}\fint_Q |Dv|^{p(\cdot)(1+\omega(\Gamma(2\varrho)^\alpha))}\,dz+1 \\
&\le 2^{n+2}\Big(\fint_Q |Dv|^{p(\cdot)(1+\varepsilon_o)}\,dz\Big)^{(1+\omega(\Gamma(2\varrho)^\alpha))/(1+\varepsilon_o)}+1 \\
&\le c\,\lambda^{1+\omega(\Gamma(2\varrho)^\alpha)}+1\le c(data,\kappa)\,\lambda.
\end{aligned}$$

Moreover, since the parabolic system $(7.14)_1$ is of the same type as (6.1), by (7.16) we can apply Theorem 6.1 which yields that

$$\sup_{\frac12 Q}|Dw|\le c_{\mathrm{DiB}}\,\lambda^{1/p_o}. \tag{7.20}$$

Note that c_{DiB} initially depends on n,N,ν,L and p_o. Since the dependence upon p_o is continuous it can be replaced by a larger constant depending on γ_1 and γ_2 instead of p_o, i.e., $c_{\mathrm{DiB}}=c_{\mathrm{DiB}}(data)$. Therefore, using (7.20) and (7.8) we can bound also the integral involving Dw in terms of λ. Inserting this in (7.19) we deduce the *second comparison estimate* sought:

$$\begin{aligned}
&\int_{\frac12 Q} \big(|Dv|^2+|Dw|^2\big)^{(p(\cdot)-2)/2}|Dv-Dw|^2\,dz \\
&\qquad\le c(data,\kappa)\Big[\omega\big(\Gamma(2\varrho)^\alpha\big)M\log\Big(\frac{K}{\varrho}\Big)+\big[\tilde\omega(\varrho)\big]^{\varepsilon_1/(2\gamma_1')}\Big]^{1/2}\lambda|Q|.
\end{aligned} \tag{7.21}$$

Note that this estimate holds for any cylinder $\frac12 Q\equiv Q^{(\lambda)}_{\varrho/2}(z_o)$ with $\lambda\ge 1$ and $\varrho\in(0,\varrho_3]$ such that $2Q$ satisfies the intrinsic relation (7.1) and $2Q\Subset\Omega_T$. We recall that $\varrho_3\in(0,1]$ depends on $data,\kappa,K,M$ and $\omega(\cdot)$.

8. Proof of the Calderón–Zygmund estimate

This section is devoted to the proof of Theorem 2.2. We shall proceed in several steps.

8.1. A stopping time argument

Here we construct a covering of the upper level set of $|Du|^{p(\cdot)}$ with respect to some parameter λ by intrinsic cylinders. The argument uses a certain *stopping time argument*, which has its origin in [28] and subsequently has been refined in [4], with the introduction of the "weight" M (see (1.4)), together with the Vitali covering argument from Lemma 4.2.

We let $K \geq 1$ and suppose that (2.9) is satisfied and consider a standard parabolic cylinder $Q_R \equiv Q_R(\mathfrak{z}_o)$ such that $Q_{2R} \Subset \Omega_T$. Then we fix $M \geq 1$ to be specified later and define

$$\lambda_o := \Big[\fint_{Q_{2R}} |Du|^{p(\cdot)} + M(|F|+1)^{p(\cdot)}\, dz \Big]^d \geq 1, \quad \text{where } d := \sup_{Q_{2R}} d(p(\cdot)) \tag{8.1}$$

and $d(\cdot)$ is defined as in (2.11). Next, following Section 9 in [30] (see also Section 4 of [4]), we fix two numbers $R \leq r_1 < r_2 \leq 2R$ such that $Q_R \subset Q_{r_1} \subset Q_{r_2} \subset Q_{2R}$, all the cylinders sharing the same center $\mathfrak{z}_o$. In the following we shall consider λ such that

$$\lambda > B\lambda_o, \quad \text{where } B := \Big(\frac{8\,\chi\, R}{r_2 - r_1} \Big)^{(n+2)d} \tag{8.2}$$

and for $z_o \in Q_{r_1}$ we consider radii ϱ satisfying

$$\min\big\{1, \lambda^{(p_o-2)/(2p_o)}\big\} \frac{r_2 - r_1}{4\chi} \leq \varrho \leq \min\big\{1, \lambda^{(p_o-2)/(2p_o)}\big\} \frac{r_2 - r_1}{2}, \tag{8.3}$$

where $p_o := p(z_o)$ and $\chi \equiv \chi(n,\gamma_1) \geq 5$ denotes the constant appearing in Lemma 4.2. Note that these choices of λ and ϱ ensure that $Q_\varrho^{(\lambda)}(z_o) \subset Q_{r_2}$ for any $z_o \in Q_{r_1}$. Next, we want to prove that for any $z_o \in Q_{r_1}$ there holds

$$\fint_{Q_\varrho^{(\lambda)}(z_o)} |Du|^{p(\cdot)} + M(|F|+1)^{p(\cdot)}\, dz < \lambda. \tag{8.4}$$

Indeed, enlarging the domain of integration from $Q_\varrho^{(\lambda)}(z_o)$ to Q_{2R} and recalling the definition of λ_o from (8.1) we infer that

$$\begin{aligned} &\fint_{Q_\varrho^{(\lambda)}(z_o)} |Du|^{p(\cdot)} + M(|F|+1)^{p(\cdot)}\, dz \\ &\quad \leq \frac{|Q_{2R}|}{|Q_\varrho^{(\lambda)}(z_o)|} \fint_{Q_{2R}} |Du|^{p(\cdot)} + M(|F|+1)^{p(\cdot)}\, dz = \Big(\frac{2R}{\varrho}\Big)^{n+2} \lambda^{(p_o-2)/p_o}\, \lambda_o^{1/d}. \end{aligned}$$

Now we distinguish the cases $p_o \ge 2$ and $p_o < 2$. If $p_o \ge 2$, then $1/d \le 1/d(p_o) = 2/p_o$ and $\min\{1, \lambda^{(p_o-2)/(2p_o)}\} = 1$, so that, using also the choice of ϱ from (8.3) we obtain

$$\fint_{Q_\varrho^{(\lambda)}(z_o)} |Du|^{p(\cdot)} + M(|F|+1)^{p(\cdot)}\,dz \le \Big(\frac{8\,\chi\,R}{r_2 - r_1}\Big)^{n+2} \lambda^{(p_o-2)/p_o}\,\lambda_o^{1/d}$$
$$< B^{1/d}\,\lambda^{(p_o-2)/p_o}\,B^{-1/d}\,\lambda^{1/d} = \lambda^{(p_o-2)/p_o}\,\lambda^{1/d} \le \lambda.$$

If $\gamma_1 \le p_o < 2$, we have

$$1/d \le 1/d(p_o) = 1 - n(2-p_o)/(2p_o) \quad \text{and} \quad \min\{1, \lambda^{(p_o-2)/(2p_o)}\} = \lambda^{(p_o-2)/(2p_o)},$$

and therefore we get

$$\fint_{Q_\varrho^{(\lambda)}(z_o)} |Du|^{p(\cdot)} + M(|F|+1)^{p(\cdot)}\,dz \le \Big(\frac{8\,\chi\,R\,\lambda^{(2-p_o)/(2p_o)}}{r_2 - r_1}\Big)^{n+2} \lambda^{(p_o-2)/p_o}\,\lambda_o^{1/d}$$
$$= B^{1/d}\,\lambda^{n(2-p_o)/(2p_o)}\,\lambda_o^{1/d} < B^{1/d}\,\lambda^{n(2-p_o)/(2p_o)}\,B^{-1/d}\,\lambda^{1/d}$$
$$= \lambda^{n(2-p_o)/(2p_o)}\,\lambda^{1/d} \le \lambda.$$

Hence, in any case we have proved that (8.4) holds.

For λ as in (8.2) we consider the upper level set

$$E(\lambda, r_1) := \big\{z \in Q_{r_1} : z \text{ is a Lebesgue point of } |Du| \text{ and } |Du(z)|^{p(z)} > \lambda\big\}.$$

In the following we show that also a converse inequality holds true for small radii and for points $z_o \in E(\lambda, r_1)$. Indeed, by the Lebesgue differentiation theorem (see [10], equation (7.9)) we infer for any $z_o \in E(\lambda, r_1)$ that

$$\lim_{\varrho \downarrow 0} \fint_{Q_\varrho^{(\lambda)}(z_o)} |Du|^{p(\cdot)} + M(|F|+1)^{p(\cdot)}\,dz \ge |Du(z_o)|^{p(z_o)} > \lambda.$$

From the preceding reasoning we conclude that the last inequality yields a radius for which the considered integral takes a value larger than λ, while (8.4) states that the integral is smaller than λ for any radius satisfying (8.3). Therefore, the continuity of the integral yields the existence of a maximal radius ϱ_{z_o} satisfying

$$0 < \varrho_{z_o} < \min\big\{1, \lambda^{(p_o-2)/(2p_o)}\big\}\frac{r_2 - r_1}{4\chi} \tag{8.5}$$

such that

$$\fint_{Q_{\varrho_{z_o}}^{(\lambda)}(z_o)} |Du|^{p(\cdot)} + M(|F|+1)^{p(\cdot)}\,dz = \lambda. \tag{8.6}$$

By saying that ϱ_{z_o} is maximal we mean that, for every

$$\varrho \in (\varrho_{z_o}, \min\{1, \lambda^{(p_o-2)/(2p_o)}\}(r_2 - r_1)/2],$$

inequality (8.4) holds. With this choice of ϱ_{z_o} we define concentric parabolic cylinders centered at $z_o \in E(\lambda, r_1)$ as follows:

$$Q^0_{z_o} := Q^{(\lambda)}_{\varrho_{z_o}}(z_o),\; Q^1_{z_o} := Q^{(\lambda)}_{\chi\varrho_{z_o}}(z_o),\; Q^2_{z_o} := Q^{(\lambda)}_{2\chi\varrho_{z_o}}(z_o),\; Q^3_{z_o} := Q^{(\lambda)}_{4\chi\varrho_{z_o}}(z_o). \tag{8.7}$$

Then, we have $Q^0_{z_o} \subset Q^1_{z_o} \subset Q^2_{z_o} \subset Q^3_{z_o} \subset Q_{r_2}$ and for $j \in \{0, \dots, 3\}$ there holds

$$\frac{\lambda}{(4\chi)^{n+2}} \le \fint_{Q^j_{z_o}} |Du|^{p(\cdot)} + M(|F|+1)^{p(\cdot)}\, dz \le \lambda. \tag{8.8}$$

Note that the upper bound follows from (8.6) and the maximal choice of the stopping radius ϱ_{z_o}, while the lower bound follows from (8.6) by enlarging the domain of integration from $Q^0_{z_o}$ to $Q^j_{z_o}$ and taking into account that $|Q^j_{z_o}|/|Q^0_{z_o}| \le (4\chi)^{n+2}$.

8.2. Estimates on intrinsic cylinders

Now we fix a particular cylinder $Q^0_{z_o}$ and define the comparison functions v and w as the unique solutions to the initial-boundary value problems (7.2) and (7.14) with $Q^3_{z_o}$ and $Q^2_{z_o}$ instead of $2Q$ and Q. Thanks to (8.8) we know that (7.1) is satisfied with $\kappa = \kappa(n, \gamma_1) = (4\chi)^{n+2}$. Moreover, we assume that

$$R \le R_o \le \varrho_3,$$

where $\varrho_3 = \varrho_3(data, K, M, \omega(\cdot)) \in (0, 1]$ denotes the radius introduced after (7.21) for the choice $\kappa = (4\chi)^{n+2}$. This ensures that we can apply (7.7), (7.20), and (7.21) with $\kappa = (4\chi)^{n+2}$ for any radius smaller than ϱ_3. Therefore, from (7.20) applied with $\kappa = (4\chi)^{n+2}$ we infer that

$$\sup_{Q^1_{z_o}} |Dw| \le c_{\mathrm{DiB}}\, \lambda^{1/p_o}, \tag{8.9}$$

where $c_{\mathrm{DiB}} = c_{\mathrm{DiB}}(data) \ge 1$. In the following we denote the constant from Lemma 3.3 by $c_\ell = c_\ell(\gamma_2) \ge 1$. For A chosen to depend on $data$ according to

$$A \ge 2\, c_\ell^2\, c_{\mathrm{DiB}}^{\gamma_2}\, \mathrm{e}^{3n/\alpha} \ge 1$$

we consider $z \in Q^1_{z_o} \cap E(A\lambda, r_1)$. Our aim now is to deduce a suitable estimate for $|Du(z)|^{p(z)}$. Applying Lemma 3.3 twice yields

$$\begin{aligned} |Du(z)|^{p(z)} &\le c_\ell^2\, |Dw(z)|^{p(z)} + c_\ell^2 \left(|Dv(z)|^2 + |Dw(z)|^2\right)^{(p(z)-2)/2} |Dv(z) - Dw(z)|^2 \\ &\quad + c_\ell \left(|Du(z)|^2 + |Dv(z)|^2\right)^{(p(z)-2)/2} |Du(z) - Dv(z)|^2. \end{aligned} \tag{8.10}$$

Next, we prove that

$$\begin{aligned} |Dw(z)|^{p(z)} &\le \left(|Du(z)|^2 + |Dv(z)|^2\right)^{(p(z)-2)/2} |Du(z) - Dv(z)|^2 \\ &\quad + \left(|Dv(z)|^2 + |Dw(z)|^2\right)^{(p(z)-2)/2} |Dv(z) - Dw(z)|^2 \end{aligned} \tag{8.11}$$

holds. Indeed, if (8.11) fails to hold we obtain from (8.9), $(4.2)_{2,3}$ of Lemma 4.1 (which is applicable due to (8.8)), the fact that $z \in E(A\lambda, r_1)$, and (8.10), that

$$|Dw(z)|^{p(z)} \le c_{\mathrm{DiB}}^{p(z)}\, \lambda^{p(z)/p_o} \le c_{\mathrm{DiB}}^{\gamma_2}\, \mathrm{e}^{\,3n/\alpha}\, \lambda$$
$$< \frac{1}{A}\, c_{\mathrm{DiB}}^{\gamma_2}\, \mathrm{e}^{\,3n/\alpha}\, |Du(z)|^{p(z)} \le \frac{1}{A}\, 2\, c_\ell^2\, c_{\mathrm{DiB}}^{\gamma_2}\, \mathrm{e}^{\,3n/\alpha}\, |Dw(z)|^{p(z)}.$$

However this contradicts the choice of A and hence (8.11) is proved. Therefore, combining (8.10) and (8.11) we get

$$|Du(z)|^{p(z)} \le 2\, c_\ell^2 \left(|Du(z)|^2 + |Dv(z)|^2\right)^{(p(z)-2)/2} |Du(z) - Dv(z)|^2$$
$$+ 2\, c_\ell^2 \left(|Dv(z)|^2 + |Dw(z)|^2\right)^{(p(z)-2)/2} |Dv(z) - Dw(z)|^2.$$

Integrating over $Q^1_{z_o} \cap E(A\lambda, r_1)$ and using the comparison estimates (7.7) and (7.21) with $\kappa = (4\chi)^{n+2}$ we obtain

$$\int_{Q^1_{z_o}\cap E(A\lambda, r_1)} |Du|^{p(\cdot)}\, dz \le 2\, c_\ell^2 \int_{Q^1_{z_o}} \left(|Du|^2 + |Dv|^2\right)^{(p(\cdot)-2)/2} |Du - Dv|^2\, dz$$
$$+ 2\, c_\ell^2 \int_{Q^1_{z_o}} \left(|Dv|^2 + |Dw|^2\right)^{(p(\cdot)-2)/2} |Dv - Dw|^2\, dz$$
$$\le c(data)\, G(M,R)\, \lambda\, |Q^0_{z_o}|\,, \tag{8.12}$$

where

$$G(M,R) := \sup_{\varrho\in(0,R]} \left[\frac{1}{M^{2-2/\gamma_1}} + \omega\big(\Gamma(2\varrho)^\alpha\big) M \log\left(\frac{K}{\varrho}\right) + \left[\tilde\omega(\varrho)\right]^{\frac{\varepsilon_1}{2\gamma_2'}}\right]^{1/2}. \tag{8.13}$$

Note that $M \ge 1$ is yet to be chosen and α and Γ are defined according to (4.3). Moreover, we recall that this estimate holds for any $\lambda > B\lambda_o$ and $z_o \in E(\lambda, r_1)$.

Next, we will infer a bound for the measure of the cylinder $Q^0_{z_o}$. From (8.6) we have

$$|Q^0_{z_o}| = \frac{1}{\lambda} \int_{Q^0_{z_o}} |Du|^{p(\cdot)}\, dz + \frac{1}{\lambda} \int_{Q^0_{z_o}} M(|F|+1)^{p(\cdot)}\, dz. \tag{8.14}$$

We split the first integral of (8.14) as

$$\int_{Q^0_{z_o}} |Du|^{p(\cdot)}\, dz = \int_{Q^0_{z_o}\cap\{|Du|^{p(\cdot)}\le\lambda/4\}} |Du|^{p(\cdot)}\, dz + \int_{Q^0_{z_o}\cap E(\lambda/4, r_2)} |Du|^{p(\cdot)}\, dz$$
$$\le \frac{\lambda}{4}\, |Q^0_{z_o}| + \int_{Q^0_{z_o}\cap E(\lambda/4, r_2)} |Du|^{p(\cdot)}\, dz,$$

and similarly the second integral as

$$\int_{Q^0_{z_o}} M(|F|+1)^{p(\cdot)}\, dz \le \frac{\lambda}{4}\, |Q^0_{z_o}| + \int_{Q^0_{z_o}\cap\{M(|F|+1)^{p(\cdot)}>\lambda/4\}} M(|F|+1)^{p(\cdot)}\, dz.$$

Inserting the last two estimates into (8.14) we can absorb the term $|Q^0_{z_o}|/2$ from the right-hand side into the left, yielding the estimate

$$|Q^0_{z_o}| \leq \frac{2}{\lambda}\int_{Q^0_{z_o}\cap E(\lambda/4,r_2)} |Du|^{p(\cdot)}\,dz + \frac{2}{\lambda}\int_{Q^0_{z_o}\cap\{M(|F|+1)^{p(\cdot)}>\lambda/4\}} M(|F|+1)^{p(\cdot)}\,dz.$$

Using this estimate in (8.12) we obtain for a constant $c \equiv c(data)$ that

$$\begin{aligned}\int_{Q^1_{z_o}\cap E(A\lambda,r_1)} |Du|^{p(\cdot)}\,dz &\leq c\,G(M,R)\int_{Q^0_{z_o}\cap E(\lambda/4,r_2)} |Du|^{p(\cdot)}\,dz\\ &\quad + c\,G(M,R)\int_{Q^0_{z_o}\cap\{M(|F|+1)^{p(\cdot)}>\lambda/4\}} M(|F|+1)^{p(\cdot)}\,dz. \end{aligned} \tag{8.15}$$

8.3. Estimates on level sets

Here we extend the estimate (8.15) to the super level set $E(A\lambda, r_1)$. To this end we first construct a suitable covering of $E(\lambda, r_1)$ by intrinsic cylinders of the type considered in the preceding steps. Here, we recall from the preceding two steps that for every $z_o \in E(\lambda, r_1)$ there exists a radius ϱ_{z_o} satisfying (8.5) such that on the cylinders $Q^j_{z_o}$, $j \in \{0,\ldots,3\}$ the estimates (8.8) and (8.15) hold. Next we want to apply the Vitali-type covering argument from Lemma 4.2. To this end we note that (8.6) and $(4.2)_1$ (with $\kappa = 1$) imply that

$$\lambda \leq \Big(\frac{\beta_n M K}{\varrho_{z_o}^{n+2}}\Big)^{p(z_o)/2}.$$

This ensures that assumption (4.5) of Lemma 4.2 is satisfied for the family $\mathcal{F} := \{Q^0_{z_o}\}$ of parabolic cylinders with center $z_o \in E(\lambda, r_1)$ (note that by possibly reducing the value of R_o we can ensure that $\varrho_{z_o} \leq R \leq R_o \leq \varrho_1$, where ϱ_1 is the radius from Lemma 4.2). Applying the lemma then yields the existence of a countable subfamily $\{Q^0_{z_i}\}_{i=1}^\infty \subset \mathcal{F}$ of pairwise disjoint parabolic cylinders, such that the χ-times enlarged cylinders $Q^1_{z_i}$ cover the set $E(A\lambda, r_1)$, i.e.,

$$E(A\lambda, r_1) \subset E(\lambda, r_1) \subset \bigcup_{i\in\mathbb{N}} Q^1_{z_i}.$$

Moreover, for the 4χ-times enlarged cylinders $Q^3_{z_i}$ we know that $Q^3_{z_i} \subset Q_{r_2}$. Here, we have used the notation from (8.7) with z_o replaced by z_i. Since we know that on any of the cylinders $Q^1_{z_i}$, $i \in \mathbb{N}$, the estimate (8.15) holds, we obtain, after summing over $i \in \mathbb{N}$, that

$$\begin{aligned}\int_{E(A\lambda,r_1)} |Du|^{p(\cdot)}\,dz &\leq c\,G(M,R)\int_{E(\lambda/4,r_2)} |Du|^{p(\cdot)}\,dz\\ &\quad + c\,G(M,R)\int_{Q_{r_2}\cap\{M(|F|+1)^{p(\cdot)}>\lambda/4\}} M(|F|+1)^{p(\cdot)}\,dz, \end{aligned} \tag{8.16}$$

where $c \equiv c(data)$. We recall that this estimate holds for every $\lambda > B\lambda_o$.

8.4. Raising the integrability exponent

At this stage we would like to multiply both sides of (8.16) by λ^{q-2} and subsequently integrate with respect to λ over $(B\lambda_o,\infty)$. This, formally would lead to an $L^{p(\cdot)q}$ estimate of Du after absorbing $\int |Du|^{p(\cdot)q}\,dz$ on the left-hand side. However, this step is not allowed since the integral might be infinite. This problem will be overcome in the following by a truncation argument. For $k \ge B\lambda_o$ we define the truncation operator

$$T_k : [0,+\infty) \to [0,k], \quad T_k(\sigma) := \min\{\sigma, k\}$$

and

$$E_k(A\lambda, r_1) := \big\{z \in Q_{r_1} : T_k\big(|Du(z)|^{p(z)}\big) > A\lambda\big\}.$$

Then, from inequality (8.16) we deduce that

$$\begin{aligned}\int_{E_k(A\lambda,r_1)} |Du|^{p(\cdot)}\,dz &\le c\,G(M,R)\int_{E_k(\lambda/4,r_2)} |Du|^{p(\cdot)}\,dz\\ &\quad + c\,G(M,R)\int_{Q_{r_2}\cap\{M(|F|+1)^{p(\cdot)}>\lambda/4\}} M(|F|+1)^{p(\cdot)}\,dz.\end{aligned} \tag{8.17}$$

This can be seen as follows. In the case $k \le A\lambda$ we have $E_k(A\lambda, r_1) = \emptyset$ and therefore (8.17) holds trivially. In the case $k > A\lambda$ the inequality (8.17) follows since $E_k(A\lambda, r_1) = E(A\lambda, r_1)$ and $E_k(\lambda/4, r_2) = E(\lambda/4, r_2)$. Therefore, multiplying both sides of (8.17) by λ^{q-2} and integrating with respect to λ over $(B\lambda_o,+\infty)$, we obtain

$$\begin{aligned}&\int_{B\lambda_o}^{\infty} \lambda^{q-2}\int_{E_k(A\lambda,r_1)} |Du|^{p(\cdot)}\,dz\,d\lambda\\ &\qquad\le c\,G(M,R)\int_{B\lambda_o}^{\infty} \lambda^{q-2}\int_{E_k(\lambda/4,r_2)} |Du|^{p(\cdot)}\,dz\,d\lambda\\ &\qquad\quad + c\,G(M,R)\int_{B\lambda_o}^{\infty} \lambda^{q-2}\int_{Q_{r_2}\cap\{M(|F|+1)^{p(\cdot)}>\lambda/4\}} M(|F|+1)^{p(\cdot)}\,dz\,d\lambda.\end{aligned} \tag{8.18}$$

Using Fubini's theorem we get, for the integral on the left-hand side of (8.18), that

$$\begin{aligned}&\int_{B\lambda_o}^{\infty} \lambda^{q-2}\int_{E_k(A\lambda,r_1)} |Du|^{p(\cdot)}\,dz\,d\lambda\\ &\quad= \int_{E_k(AB\lambda_o,r_1)} |Du|^{p(\cdot)}\int_{B\lambda_o}^{T_k(|Du(z)|^{p(z)})/A} \lambda^{q-2}\,d\lambda\,dz\\ &\quad= \frac{1}{q-1}\Big[\frac{1}{A^{q-1}}\int_{E_k(AB\lambda_o,r_1)} |Du|^{p(\cdot)}T_k\big(|Du|^{p(\cdot)}\big)^{q-1}\,dz\\ &\qquad - (B\lambda_o)^{q-1}\int_{E_k(AB\lambda_o,r_1)} |Du|^{p(\cdot)}\,dz\Big]\\ &\quad\ge \frac{1}{q-1}\Big[\frac{1}{A^{q-1}}\int_{Q_{r_1}} |Du|^{p(\cdot)}T_k\big(|Du|^{p(\cdot)}\big)^{q-1}\,dz - (B\lambda_o)^{q-1}\int_{Q_{r_1}} |Du|^{p(\cdot)}\,dz\Big],\end{aligned}$$

where in the last line we used the decomposition

$$Q_{r_1} = E_k(AB\lambda_o, r_1) \cup (Q_{r_1} \setminus E_k(AB\lambda_o, r_1))$$

and the fact that $T_k(|Du|^{p(\cdot)}) \le AB\lambda_o$ on $Q_{r_1} \setminus E_k(AB\lambda_o, r_1)$. Again by Fubini's theorem we obtain, for the first integral on the right-hand side of (8.18),

$$\begin{aligned}\int_{B\lambda_o}^{\infty} \lambda^{q-2} \int_{E_k(\lambda/4,r_2)} |Du|^{p(\cdot)}\, dz\, d\lambda &= \int_{E_k(B\lambda_o/4,r_2)} |Du|^{p(\cdot)} \int_{B\lambda_o}^{4T_k(|Du|^{p(\cdot)})} \lambda^{q-2}\, d\lambda\, dz \\ &\le \frac{4^{q-1}}{q-1} \int_{Q_{r_2}} |Du|^{p(\cdot)} T_k\big(|Du|^{p(\cdot)}\big)^{q-1}\, dz\end{aligned}$$

and, analogously, for the integral involving the right-hand side F,

$$\begin{aligned}&\int_{B\lambda_o}^{\infty} \lambda^{q-2} \int_{Q_{r_2}\cap\{M(|F|+1)^{p(\cdot)}>\lambda/4\}} M(|F|+1)^{p(\cdot)}\, dz\, d\lambda \\ &\quad= \int_{Q_{r_2}\cap\{M(|F|+1)^{p(\cdot)}>B\lambda_o/4\}} M(|F|+1)^{p(\cdot)} \int_{B\lambda_o}^{4M(|F|+1)^{p(\cdot)}} \lambda^{q-2}\, d\lambda\, dz \\ &\quad\le \frac{4^{q-1}M^q}{q-1} \int_{Q_{r_2}} (|F|+1)^{p(\cdot)q}\, dz.\end{aligned}$$

Hence, combining the preceding estimates with (8.18) we get

$$\begin{aligned}\int_{Q_{r_1}} |Du|^{p(\cdot)} T_k\big(|Du|^{p(\cdot)}\big)^{q-1}\, dz &\le (AB\lambda_o)^{q-1} \int_{Q_{r_1}} |Du|^{p(\cdot)}\, dz \\ &\quad + \bar{c}\, A^{q-1} G(M,R) \int_{Q_{r_2}} |Du|^{p(\cdot)} T_k\big(|Du|^{p(\cdot)}\big)^{q-1}\, dz \\ &\quad + \bar{c}\, A^{q-1} M^q G(M,R) \int_{Q_{r_2}} (|F|+1)^{p(\cdot)q}\, dz, \end{aligned} \tag{8.19}$$

where $\bar{c} = \bar{c}(data)$. Note that the estimate stays stable as $q \downarrow 1$.

8.5. Choice of the parameters

We now make the choices of the parameters M and R_o so that $\bar{c}A^{q-1}G(M,R) \le 1/2$ whenever $R \le R_o$. First, we choose $M = M(data, q) \ge 1$ large enough so that

$$\frac{\bar{c}\, A^{q-1}}{M^{1-1/\gamma_1}} \le \frac{1}{4}.$$

Next, we decrease the value of R_o, now depending on $data, K, \omega(\cdot), \tilde{\omega}(\cdot)$ and q, in such a way that for any $\varrho \le R_o$ we have

$$\bar{c}\, A^{q-1} \Big[\omega\big(\Gamma(2\varrho)^{\alpha}\big) M \log\Big(\frac{K}{\varrho}\Big) + \big[\tilde{\omega}(\varrho)\big]^{\varepsilon_1/(2\gamma_1')} \Big]^{1/2} \le \frac{1}{4}. \tag{8.20}$$

Note that this is possible due to the assumptions (2.4) and (2.8). Recalling the definition of G in (8.13) we therefore have $\bar{c}A^{q-1}G(M,R) \le 1/2$ for any $R \le R_o$. Using this in (8.19) we get

$$\begin{aligned}\int_{Q_{r_1}} &|Du|^{p(\cdot)} T_k\big(|Du|^{p(\cdot)}\big)^{q-1}\,dz \\ &\le \frac{1}{2}\int_{Q_{r_2}} |Du|^{p(\cdot)} T_k\big(|Du|^{p(\cdot)}\big)^{q-1}\,dz \\ &\quad + c\Big(\frac{R}{r_2-r_1}\Big)^{\beta}\lambda_o^{q-1}\int_{Q_{2R}}|Du|^{p(\cdot)}\,dz + c\int_{Q_{2R}}(|F|+1)^{p(\cdot)q}\,dz,\end{aligned}$$

where $\beta \equiv (n+2)(q-1)d$ and $c \equiv c(data, q)$. At this point we apply Lemma 3.1 with

$$\phi(r) \equiv \int_{Q_r} |Du|^{p(\cdot)}\, T_k\big(|Du|^{p(\cdot)}\big)^{q-1}\,dz,$$

and

$$\mathcal{A} \equiv c\int_{Q_{2R}}(|F|+1)^{p(\cdot)q}\,dz \quad \text{and} \quad \mathcal{B} \equiv c\,R^{\beta}\lambda_o^{q-1}\int_{Q_{2R}}|Du|^{p(\cdot)}\,dz,$$

yielding

$$\int_{Q_R} |Du|^{p(\cdot)}\, T_k\big(|Du|^{p(\cdot)}\big)^{q-1}\,dz \le c(\beta)\Big[\mathcal{A} + \frac{\mathcal{B}}{R^{\beta}}\Big].$$

Passing to the limit $k \to \infty$, which is possible by Fatou's lemma, and taking averages we find that

$$\fint_{Q_R}|Du|^{p(\cdot)q}\,dz \le c\Big[\lambda_o^{q-1}\fint_{Q_{2R}}|Du|^{p(\cdot)}\,dz + \fint_{Q_{2R}}(|F|+1)^{p(\cdot)q}\,dz\Big]. \tag{8.21}$$

Note that $c \equiv c(data, q)$, since β depends continuously on $p(\cdot)$, i.e., the dependence on $p(\cdot)$ via the parameter d can be replaced by a dependence on γ_1 and γ_2. Since $Q_{2R} \Subset \Omega_T$ was arbitrary, we have thus proved the first assertion in Theorem 2.2, i.e., that $|Du|^{p(\cdot)} \in L^q_{\mathrm{loc}}(\Omega_T)$. There remains to show the estimate (2.10).

8.6. Adjusting the exponent

Here we first observe that (8.21), together with the definition of λ_o in (8.1), leads to the estimate (2.10) in Theorem 2.2, but with d instead of $d(p_o)$, where $p_o := p(\mathfrak{z}_o)$ and $\mathfrak{z}_o$ is the center of the cylinder $Q_{2R} \equiv Q_{2R}(\mathfrak{z}_o)$. We recall that d was defined in (2.11) and $d \ge d(p_o)$. In order to decrease the exponent from d to $d(p_o)$ we need to show a bound of the form

$$\mathfrak{E}^{d-d(p_o)} \le c(n,\gamma_1), \quad \text{where} \quad \mathfrak{E} := \fint_{Q_{2R}} |Du|^{p(\cdot)} + (|F|+1)^{p(\cdot)q}\,dz. \tag{8.22}$$

To this end we first deduce an upper bound for $d - d(p_o)$ in terms of $\omega(R)$. Since $d(p(\cdot))$ is continuous there exists $\hat{z} \in \overline{Q_R}$ such that $d = d(p(\hat{z}))$. From the definition of $d(\cdot)$ in (2.11) we observe that

$$d(p_o) \ge \max\Big\{\frac{p_o}{2}, \frac{2p_o}{p_o(n+2)-2n}\Big\}.$$

In the following we distinguish the cases where $p(\hat z)$ is larger or smaller than two. In the case $p(\hat z) \ge 2$ we get from (2.5) that

$$d - d(p_o) = \frac{p(\hat z)}{2} - d(p_o) \le \frac{p(\hat z)}{2} - \frac{p_o}{2} \le \frac{1}{2}\,\omega(R),$$

while in the case $p(\hat z) < 2$ we have $p(\hat z) \le p_o$ and therefore we find in a similar way that

$$\begin{aligned} d - d(p_o) &\le \frac{2p(\hat z)}{p(\hat z)(n+2)-2n} - \frac{2p_o}{p_o(n+2)-2n} \\ &= \frac{4n(p_o - p(\hat z))}{[p(\hat z)(n+2)-2n][p_o(n+2)-2n]} \le \frac{4n}{[\gamma_1(n+2)-2n]^2}\,\omega(R). \end{aligned}$$

Hence, in either case we have proved that $d - d(p_o) \le c(n,\gamma_1)\omega(R)$. Recalling the definition of $\mathfrak{E}$ from (8.22) and using (2.9) we thus obtain

$$\mathfrak{E}^{d-d(p_o)} \le c(n,\gamma_1)\big[R^{-(n+2)}K\big]^{c(n,\gamma_1)\omega(R)} \le c(n,\gamma_1).$$

We note that the last inequality is a consequence of the logarithmic continuity of ω from (2.5), since $R^{-\omega(R)} \le e$ and

$$K^{\omega(R)} = \exp\big[\omega(R)\log K\big] \le \exp\big[\omega(R)\log\big(\tfrac{1}{R}\big)\big] \le \mathrm{e}$$

provided $R \le R_o \le \min\{R_1, 1/K\}$, where R_1 is the radius from (2.5). This finishes the proof of (8.22) and by the reasoning above we therefore obtain the asserted estimate (2.10). Thus we have completed the proof of Theorem 2.2. □

8.7. Proof of Remark 2.3

Here it is enough to ensure that we can choose $R_o > 0$ and $\varepsilon_{\mathrm{BMO}}$ in such a way that (8.20) is satisfied. Assuming for instance

$$[a]_{\mathrm{BMO}} \le \varepsilon_{\mathrm{BMO}} := \Big(\frac{1}{8\,\bar c A^{q-1}}\Big)^{4\gamma_1'/\varepsilon_1} \quad \text{and} \quad \omega\big(\Gamma(2\varrho)^\alpha\big)M\log\Big(\frac{K}{\varrho}\Big) \le \Big(\frac{1}{8\,\bar c A^{q-1}}\Big)^2$$

for any $\varrho \le R_o$ we conclude that (8.20) holds, since $\tilde\omega(\varrho) \le [a]_{\mathrm{BMO}}$. The rest of the proof is completely the same as the proof of Theorem 2.2.

References

[1] Acerbi, E. and Mingione, G.: Regularity results for a class of functionals with non-standard growth. *Arch. Ration. Mech. Anal.* **156** (2001), no. 2, 121–140.

[2] Acerbi, E. and Mingione, G.: Regularity results for stationary electro-rheological fluids. *Arch. Ration. Mech. Anal.* **164** (2002), no. 3, 213–259.

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

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

[5] Acerbi, E., Mingione, G. and Seregin, G. A.: Regularity results for parabolic systems related to a class of non-Newtonian fluids. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **21** (2004), no. 1, 25–60.

[6] Antontsev, S. and Shmarev, S.: A model porous medium equation with variable exponent of nonlinearity: existence, uniqueness and localization properties of solutions. *Nonlinear Anal.* **60** (2005), no. 3, 515–545.

[7] Antontsev, S. and Shmarev, S.: Parabolic equations with anisotropic nonstandard growth conditions. In *Free boundary problems*, 33–44. International Series of Numerical Mathematics 154, Birkhäuser, Basel, 2007.

[8] Antontsev, S. and Shmarev, S.: Anisotropic parabolic equations with variable nonlinearity. *Publ. Mat.* **53** (2009), no. 2, 355–399.

[9] Antontsev, S. and Zhikov, V. V.: Higher integrability for parabolic equations of $p(x,t)$-Laplacian type. *Adv. Differential Equations* **10** (2005), no. 9, 1053–1080.

[10] Bögelein, V. and Duzaar, F.: Higher integrability for parabolic systems with nonstandard growth and degenerate diffusions. *Publ. Mat.* **55** (2011), no. 1, 201–250.

[11] Bögelein, V. and Duzaar, F.: Hölder estimates for parabolic $p(x,t)$-Laplacian systems. *Math. Ann.* **354** (2012), no. 3, 907–938.

[12] Bögelein, V., Duzaar, F. and Mingione, G.: Degenerate problems with irregular obstacles. *J. Reine Angew. Math.* **650** (2011), 107–160.

[13] Bögelein, V., Duzaar, F., Habermann, J. and Scheven, C.: Stationary electrorheological fluids: low order regularity for systems with discontinuous coefficients. *Adv. Calc. Var.* **5** (2012), no. 1, 1–57.

[14] Caffarelli, L. A. and Peral, I.: On $W^{1,p}$ estimates for elliptic equations in divergence form. *Comm. Pure Appl. Math.* **51** (1998), no. 1, 1–21.

[15] Chen, Y. and Xu, M.: Hölder continuity of weak solutions for parabolic equations with nonstandard growth conditions. *Acta Math. Sin. (Engl. Ser.)* **22** (2006), no. 3, 793–806.

[16] Cupini, G., Fusco, N. and Petti, R.: Hölder continuity of local minimizers. *J. Math. Anal. Appl.* **235** (1999), no. 2, 578–597.

[17] DiBenedetto, E.: *Degenerate parabolic equations.* Universitext, Springer-Verlag, New York, 1993.

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

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

[20] DiBenedetto, E. and Manfredi, J.: On the higher integrability of the gradient of weak solutions of certain degenerate elliptic systems. *Amer. J. Math.* **115** (1993), no. 5, 1107–1134.

[21] Duzaar, F. and Habermann, J.: Partial regularity for parabolic systems with non standard growth. *J. Evol. Equ.* **12** (2012), no. 1, 203–244.

[22] Duzaar, F., Mingione, G. and Steffen, K.: Parabolic systems with polynomial growth and regularity. *Mem. Amer. Math. Soc.* **214** (2011), no. 1005.

[23] Giusti, E.: *Direct methods in the calculus of variations.* World Scientific, Tuck Link, Singapore, 2003.

[24] Habermann, J.: Calderón–Zygmund estimates for higher order systems with $p(x)$ growth. *Math. Z.* **258** (2008), no. 2, 427–462.

[25] Henriques, E. and Urbano, J. M.: Intrinsic scaling for PDEs with an exponential nonlinearity. *Indiana Univ. Math. J.* **55** (2006), no. 5, 1701–1721.

[26] Iwaniec, T.: Projections onto gradient fields and L^p-estimates for degenerated elliptic operators. *Studia Math.* **75** (1983), no. 3, 293–312.

[27] Iwaniec, T. and Verde, A.: On the operator $\mathcal{L}(f) = f \log |f|$. *J. Funct. Anal.* **169** (1999), no. 2, 391–420.

[28] Kinnunen, J. and Lewis, J. L.: Higher integrability for parabolic systems of p-Laplacian type. *Duke Math. J.* **102** (2000), no. 2, 253–271.

[29] Kinnunen, J. and Zhou, S.: A local estimate for nonlinear equations with discontinuous coefficients. *Commun. Partial Differential Equations* **24** (1999), no. 11-12, 2043–2068.

[30] Mingione, G.: The Calderón–Zygmund theory for elliptic problems with measure data. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **6** (2007), no. 2, 195–261.

[31] Misawa, M.: L^q estimates of gradients for evolutional p-Laplacian systems. *J. Differential Equations* **219** (2005), no. 2, 390–420.

[32] Růžička, M. M.: *Electrorheological fluids: modeling and mathematical theory.* Lecture Notes in Mathematics 1748, Springer-Verlag, Berlin, 2000.

[33] Scheven, C.: Existence of localizable solutions to nonlinear parabolic problems with irregular obstacles. To appear in *Manuscripta Math.* DOI: 10.1007/s00229-014-0684-8.

[34] Zhikov, V. V. and Pastukhova, S. E.: On the property of higher integrability for parabolic systems of variable order of nonlinearity. *Mat. Zametki* **87** (2010), no. 2, 179–200. Translation in *Math. Notes* **87** (2010), no. 1-2, 169–188.

Received October 11, 2012.

Paolo Baroni: Department of Mathematics, Uppsala Universitet, Lägerhyddsvägen 1, 75106 Uppsala, Sweden.
E-mail: paolo.baroni@math.uu.se

Verena Bögelein: Department Mathematik, Universität Erlangen–Nürnberg, Cauerstrasse 11, 91058 Erlangen, Germany.
E-mail: boegelein@mi.uni-erlangen.de

This work was supported by ERC grant 207573 “Vectorial Problems”.