Annali di Matematica (2009) 188:61–122
DOI 10.1007/s10231-008-0067-4

Partial regularity and singular sets of solutions of higher order parabolic systems

Verena Bögelein

Received: 11 June 2007 / Revised: 9 October 2007 / Published online: 4 April 2008

Abstract In the present paper we provide a broad survey of the regularity theory for non-differentiable higher order parabolic systems of the type

$$\int_{\Omega_T} u \cdot \varphi_t - A(z, u, Du, \ldots, D^m u) \cdot D^m \varphi \, \mathrm{d}z = \int_{\Omega_T} \sum_{k=0}^{m-1} B^k(z, u, Du, \ldots, D^m u) \cdot D^k \varphi \, \mathrm{d}z.$$

Initially, we prove a partial regularity result with the method of A-polycaloric approximation, which is a parabolic analogue of the harmonic approximation lemma of De Giorgi. Moreover, we prove better estimates for the maximal parabolic Hausdorff-dimension of the singular set of weak solutions, using fractional parabolic Sobolev spaces. Thereby, we also consider different situations, which yield a better dimension reduction result, including the low dimensional case and coefficients $A(z, D^m u)$, independent of the lower order derivatives of u.

Keywords Partial regularity · Singular set · Higher order parabolic systems

Mathematics Subject Classification (2000) 35D10 · 35G20 · 35K55

1 Introduction and statement of the results

Let Ω be a bounded domain in $\mathbb{R}^n$ and Ω_T the parabolic cylinder $\Omega \times (-T, 0)$ over Ω with $T > 0$. In the following we consider weak solutions $u \in L^2(-T, 0; W^{m,2}(\Omega; \mathbb{R}^N))$, $N, m \geq 1$ of higher order parabolic systems of the form

V. Bögelein (✉)
Department Mathematik, Universität Erlangen–Nürnberg, Bismarckstr. 1 1/2,
91054 Erlangen, Germany
e-mail: boegelein@mi.uni-erlangen.de

$$\int_{\Omega_T} \big(u \cdot \varphi_t - A(z, \delta u, D^m u) \cdot D^m \varphi\big)\, \mathrm{d}z = \int_{\Omega_T} B(z, \delta u, D^m u) \cdot \delta\varphi \,\mathrm{d}z \tag{1.1}$$

for all $\varphi \in C_0^\infty(\Omega_T; \mathbb{R}^N)$. Here and in the following we write $z = (x, t) \in \mathbb{R}^{n+1}$, $\varphi_t = \partial_t \varphi$ denotes the derivative with respect to the time-variable t, whence Du, respectively $D^k u$ denote the derivatives with respect to the space-variable x and $\delta u = (u, Du, \dots, D^{m-1}u)$ is the vector of lower order derivatives. We note that $D^k u = \{D^\alpha u_i\}_{i=1,\dots,N}^{|\alpha|=k}$ is an element of the vectorspace $\odot^k(\mathbb{R}^n, \mathbb{R}^N)$ of k-linear functions with values in $\mathbb{R}^N$, which can be identified with $\mathbb{R}^{N\binom{n+k-1}{k}}$. Throughout the whole paper we shall use the abbreviations $\mathscr{N} = N\binom{n+m-1}{m}$, $\mathscr{M} = N\binom{n+m-1}{m-1} = \sum_{k=0}^{m-1} \mathscr{M}_k$, where $\mathscr{M}_k = N\binom{n+k-1}{k}$, which allows us to write $D^m u \in \mathbb{R}^{\mathscr{N}}$, $D^k u \in \mathbb{R}^{\mathscr{M}_k}$ and $\delta u \in \mathbb{R}^{\mathscr{M}}$.

We consider coefficients $A\colon \Omega_T \times \mathbb{R}^{\mathscr{M}} \times \mathbb{R}^{\mathscr{N}} \to \mathrm{Hom}(\mathbb{R}^{\mathscr{N}}, \mathbb{R})$ such that $(z, \xi, p) \mapsto A(z, \xi, p)$ and $(z, \xi, p) \mapsto \partial_p A(z, \xi, p)$ are continuous on $\Omega_T \times \mathbb{R}^{\mathscr{M}} \times \mathbb{R}^{\mathscr{N}}$ and $B \equiv (B^0, \dots, B^{m-1})$ with $B^k\colon \Omega_T \times \mathbb{R}^{\mathscr{M}} \times \mathbb{R}^{\mathscr{N}} \to \mathrm{Hom}(\mathbb{R}^{\mathscr{M}_k}, \mathbb{R})$ for $k = 0, \dots, m-1$. We assume the following ellipticity and growth conditions, with $0 < \nu \le 1$ and $1 \le L < \infty$:

$$\partial_p A(z, \xi, p)\, \widetilde{p} \cdot \widetilde{p} \ge \nu\, |\widetilde{p}|^2, \tag{1.2}$$

$$|A(z, \xi, p)| \le L\,(1 + |p|), \tag{1.3}$$

$$|B(z, \xi, p)| \le L\,(1 + |p|), \tag{1.4}$$

for all $z \in \Omega_T$, $\xi \in \mathbb{R}^{\mathscr{M}}$ and $p, \widetilde{p} \in \mathbb{R}^{\mathscr{N}}$. Moreover, considering minimizers of functionals it is usual to assume growth conditions on the second derivatives of the functional, see for instance [1]. In the case of elliptic, respectively parabolic systems this corresponds to a bound on $\partial_p A$. Here, we shall assume $\partial_p A$ to be—not necessarily uniformly—bounded. More precisely, we assume that for given $M > 0$ there exists κ_M, such that

$$|\partial_p A(z, \xi, p)| \le L\, \kappa_M, \tag{1.5}$$

for all $z \in \Omega_T$, $\xi \in \mathbb{R}^{\mathscr{M}}$ and $p \in \mathbb{R}^{\mathscr{N}}$ such that $|\xi| + |p| \le M$. With respect to the variables (z, ξ) we will put only a Hölder-continuity assumption on the coefficients. Since our parabolic system is of order $2m$, the natural parabolic metric in $\mathbb{R}^{n+1}$ is

$$\mathrm{d}_{\mathscr{P}}(z, z_0) \equiv \sqrt[2m]{|x - x_0|^{2m} + |t - t_0|}, \quad \text{where } z = (x, t), z_0 = (x_0, t_0) \in \mathbb{R}^{n+1}.$$

We assume the mapping $(z, \xi) \mapsto \frac{A(z,\xi,p)}{1+|p|}$ to be—not necessarily uniformly—Hölder continuous with respect to the parabolic metric $\mathrm{d}_{\mathscr{P}}$, with Hölder exponent $\beta \in (0, 1)$, i.e.

$$|A(z, \xi, p) - A(z_0, \xi_0, p)| \le L\, \theta\big(|\xi| + |\xi_0|, \mathrm{d}_{\mathscr{P}}(z, z_0) + |\xi - \xi_0|\big)(1 + |p|), \tag{1.6}$$

for all $z, z_0 \in \Omega_T$, $\xi, \xi_0 \in \mathbb{R}^{\mathscr{M}}$ and $p \in \mathbb{R}^{\mathscr{N}}$ with

$$\theta(y, s) \equiv \min\{1, K(y)s^\beta\},$$

where $K\colon [0, \infty) \to [1, \infty)$ is non-decreasing. This will be enough to prove our partial regularity result, i.e. to show that $D^m u$ is of class $C^{\beta, \frac{\beta}{2m}}$ outside a set of $\mathscr{L}^{n+1}$-measure zero, the so called singular set.

Theorem 1.1 *Suppose that $u \in L^2(-T, 0; W^{m,2}(\Omega; \mathbb{R}^N))$ is a weak solution of the parabolic system (1.1) in Ω_T under the assumptions (1.2)–(1.6). Then*

$$D^m u \in C^{\beta, \frac{\beta}{2m}}\big(\Omega_T \backslash \Sigma; \mathbb{R}^{\mathscr{N}}\big),$$

where $\Omega_T \backslash \Sigma$ is an open subset of Ω_T with full measure, i.e. $|\Sigma| = 0$.

To prove regularity results for higher order parabolic systems, there had to be developed new techniques, in particular to overcome the difficulties arising from the lack of regularity of the intermediate derivatives $Du, \ldots, D^{m-1}u$ with respect to the time variable t. For non-differentiable second-order parabolic systems without any a priori regularity assumptions on the solution, regularity results were initially achieved by Duzaar and Mingione [19] and Duzaar et al. [20], see also [3] for systems with non-standard growth.

With regard to regularity theory it is essential to have a suitable Cacciopoli inequality at hand. In order to come along with the non-uniform bound of $\partial_p A$ in (1.5) we will need very fine estimates in the proof. Considering higher order problems, we also have to treat integrals of intermediate derivatives $Du, \ldots, D^{m-1}u$. Since the general Poincaré inequality is not applicable for weak solutions of parabolic systems, we cannot estimate them in terms of $D^m u$. Instead, we use an Interpolation-Theorem on the annulus (see Lemma 2.4), which preserves the right scaling.

After proving that the in terms of Cacciopoli's inequality rescaled solution is approximately a solution of a linear system, we can apply the so called A-polycaloric approximation lemma, which allows us to approximate the weak solution of a non-linear system by a solution of a linear system with constant coefficients. This is a parabolic analogue of the classical harmonic approximation lemma of De Giorgi (see [13,35]). The technique has its origin in Simon's approach to Allard's regularity theorem [34,35] and was used to obtain regularity results for non-linear elliptic, respectively parabolic systems in [16–19]. This technique will allow us to approximate weak solutions of the original problem with solutions of a linear parabolic system with constant coefficients. Subsequently, we can exploint good estimates for solutions of linear systems

The application of the A-polycaloric approximation lemma then leads us—after an iteration procedure—to an excess-decay estimate in points where a certain smallness condition is fulfilled. In those points we conclude $C^{\beta, \frac{\beta}{2m}}$-regularity of $D^m u$ by an integral characterization of Hölder continuous functions due to Campanato and therefore obtain a characterization of the singular set in Theorem 3.7. But contrary to the elliptic case we cannot directly conclude that the singular set has $\mathscr{L}^{n+1}$-measure zero. This is due to the fact that we cannot apply Poinaré's inequality to weak solutions of parabolic systems, since they are a priori only L^2-functions with respect to the time variable t. Therefore, we prove a sort of Poincaré inequality for weak solutions in Lemma 3.12, which exploits the absolute continuity in time of weighted means of the weak solution.

In the following we denote by Σ the singular set of u, i.e. we have $D^m u \in C^{\beta, \frac{\beta}{2m}}(\Omega_T \setminus \Sigma; \mathbb{R}^{\mathscr{N}})$. Then, from Theorem 1.1 we know that Σ is a set of $\mathscr{L}^{n+1}$-measure zero. But how "large" can Σ be? To answer this question we firstly need a quantity to measure the size of Σ. For elliptic systems, estimates of the singular set are usually expressed in terms of its Hausdorff-dimension. To get the analogous results in the case of parabolic systems, one has to express the estimates in terms of the parabolic Hausdorff-dimension. In our situation of $2m$th order parabolic systems, it is convenient to work with the following parabolic Hausdorff-dimension

$$\dim_{\mathscr{P}}(F) \equiv \inf \left\{ s > 0 : \mathscr{P}^s(F) = 0 \right\} = \sup \left\{ s > 0 : \mathscr{P}^s(F) = \infty \right\},$$

where $F \subset \mathbb{R}^{n+1}$ and

$$\mathscr{P}^s(F) \equiv \lim_{\rho \searrow 0} \inf \left\{ \sum_{i=1}^{\infty} \rho_i^s : F \subset \bigcup_{i=1}^{\infty} Q_{\rho_i}(z_i), \; 0 \le \rho_i < \rho \right\}$$

denotes the parabolic s-dimensional Hausdorff-measure of F, with $s \in [0, n+2m]$. Here, the supremum is taken over all parabolic cylinders of the form $Q_{\rho_i}(z_i) = B_{\rho_i} \times (t_i - \rho_i^{2m}, t_i)$.

In order to prove our dimension reduction result, i.e. to show better estimates for the Hausdorff-dimension of the singular set, we have to slightly reinforce our assumptions in the sense that $\partial_p A$ is uniformly bounded:

$$|\partial_p A(z, \xi, p)| \leq L, \tag{1.7}$$

for all $z \in \Omega_T$, $\xi \in \mathbb{R}^{\mathscr{M}}$ and $p, \widetilde{p} \in \mathbb{R}^{\mathscr{N}}$ and

$$|A(z, \xi, p) - A(z_0, \xi_0, p)| \leq L\, \widetilde{\theta}\big(\mathrm{d}_{\mathscr{P}}(z, z_0) + |\xi - \xi_0|\big)(1 + |p|), \tag{1.8}$$

for all $z, z_0 \in \Omega_T$, $\xi, \xi_0 \in \mathbb{R}^{\mathscr{M}}$ and $p \in \mathbb{R}^{\mathscr{N}}$, where $\widetilde{\theta}\colon [0, \infty) \mapsto [0, 1]$ is a bounded continuous concave function, such that

$$\widetilde{\theta}(s) \leq s^{\beta}, \quad s > 0. \tag{1.9}$$

Under these slightly stronger assumptions, the result of Theorem 1.1 can be improved to the following

Theorem 1.2 *Let $u \in L^2(-T, 0; W^{m,2}(\Omega; \mathbb{R}^N))$ be a weak solution of the parabolic system (1.1) in Ω_T under the assumptions (1.2)–(1.4), (1.7) and (1.8) and let Σ denote the singular set of u. Then there exists $\delta = \delta(n, m, N, \beta, L/\nu) > 0$, such that*

$$\dim_{\mathscr{P}}(\Sigma) \leq n + 2m - \delta.$$

Remark 1.3 We can quantify δ in the last theorem by $\delta = \frac{\beta\sigma}{1+\sigma}$, where σ is the exponent gained from the higher integrability of $|D^m u|$ (see Theorem 4.1). Therefore we see that

$$\lim_{\beta \searrow 0} \delta = 0 \quad \text{and} \quad \lim_{L/\nu \to \infty} \delta = 0.$$

The estimate for the singular set from Theorem 1.2 can still be improved for simpler systems, where the coefficients A do not depend on the intermediate derivatives δu, of the following type

$$\int_{\Omega_T} \big(u \cdot \varphi_t - A(z, D^m u) \cdot D^m \varphi\big)\, \mathrm{d}z = 0 \qquad \text{for all } \varphi \in C_0^{\infty}(\Omega_T; \mathbb{R}^N). \tag{1.10}$$

Then, we have

Theorem 1.4 *Let $u \in L^2(-T, 0; W^{m,2}(\Omega; \mathbb{R}^N))$ be a weak solution of the simpler system (1.10) in Ω_T under the assumptions (1.2), (1.3), (1.7) and (1.8) and denote by Σ the singular set of u. Then there exists $\delta = \delta(n, m, N, \beta, L/\nu) > 0$, such that*

$$\dim_{\mathscr{P}}(\Sigma) \leq n + 2m - 2\beta - \delta.$$

The main idea in proving this kind of dimension reduction results for non-differentiable systems is to show that $D^m u$ lies in a certain fractional Sobolev-space. The result then follows from Giusti's lemma. This method was introduced by Mingione in [30] and [31] for elliptic systems. Moreover, for elliptic systems, the dimension reduction result could be improved also in the case that the coefficients depend on u, under the additional assumption that u is Hölder-continuous with some arbitrary Hölder exponent.

Regarding our parabolic problem, this suggests that we can improve the estimate for the Hausdorff dimension of the singular set from Theorem 1.2 under the assumption that the solution u and its derivatives up to the order $m-1$ are Hölder continuous. Then, it turns out, that in fact the stronger estimate from Theorem 1.4 holds, although the coefficients A depend on δu.

Theorem 1.5 *Let $u \in L^2(-T,0;W^{m,2}(\Omega;\mathbb{R}^N)) \cap C^{m-1,\lambda,\frac{\lambda}{2m}}(\Omega_T;\mathbb{R}^N)$ with $\lambda \in (0,1)$ be a weak solution to the parabolic system (1.1) in Ω_T under the assumptions (1.2)–(1.4), (1.7) and (1.8) and denote by Σ the singular set of u. Then*

$$\dim_{\mathscr{P}}(\Sigma) \le n + 2m - 2\beta.$$

In the case of homogeneous systems, i.e. $B \equiv 0$ this inequality is strict, i.e. there exists $\delta = \delta(n,m,N,\beta,L/\nu) > 0$ such that

$$\dim_{\mathscr{P}}(\Sigma) \le n + 2m - 2\beta - \delta.$$

A main ingredient in the proof is an Interpolation-Theorem which ensures better integrability properties of a function by interpolation between fractional Sobolev spaces and Hölder spaces. In the elliptic framework, this interpolation result goes back to Campanato [8, Teorema 3.III]. Therefore, we will first show a parabolic version of this Interpolation-Theorem (see Theorem 5.7). It will be exploited to improve the fractional differentiability of $D^m u$ in each step of a finite iteration process in the proof of Lemma 6.9.

Finally, we want to point out the main difficulties appearing in the parabolic case. Since we only know that weak solutions are L^2-functions with respect to t, we cannot estimate finite differences of $D^k u$, $k = 0, \ldots, m-1$ in terms of $\partial_t D^k u$. Therefore, we have to exploit the parabolic system in Lemma 6.1 to infer a suitable similar estimate for the L^2-norm of finite differences of $D^k u$. Since this estimate is not good enough for the purpose of Theorem 1.5, we will consider second finite differences in Lemma 6.6, for which we can show better estimates. These estimates can then be carried over to first finite differences, applying a result of Domokos [14], used for the treatment of sub-elliptic equations in the Heisenberg-group. Moreover, additionally to the L^2-norm, we will also need estimates for L^{2+b}-norms ($b > 0$) of finite differences of $D^k u$. In order to transfer our estimate to a "larger" L^p-norm, we will use the Hardy–Littlewood maximal function and the sharp function in Lemma 6.5.

The assumption concerning the Hölder continuity of $D^{m-1}u$ in Theorem 1.5 indeed is fulfilled in particular situations. In small dimensions, i.e. in the case $n \le 2$ for second-order systems ($m = 1$), we know due to a result of Campanato [10, Theorem 8.II], that u is Hölder continuous on a set of full $\mathscr{P}^n$-measure. More precisely, there exists an open subset $\Omega_0 \subset \Omega_T$ and $\lambda \in (0,1)$ such that $u \in C^{\lambda,\frac{\lambda}{2}}\left(\Omega_0;\mathbb{R}^N\right)$, where $\mathscr{P}^n(\Omega_T \setminus \Omega_0) = 0$. This yields the following

Theorem 1.6 *Let $n \le 2$, $m = 1$ and $u \in L^2(-T,0;W^{1,2}(\Omega;\mathbb{R}^N))$ be a weak solution of the system (1.1) in Ω_T under the assumptions (1.2)–(1.4), (1.7) and (1.8). Then for the singular set Σ of u there holds*

$$\dim_{\mathscr{P}}(\Sigma) \le n + 2 - 2\beta.$$

In the case of homogeneous systems there there exists $\delta = \delta(n,N,\beta,L/\nu) > 0$ such that

$$\dim_{\mathscr{P}}(\Sigma) \le n + 2 - 2\beta - \delta.$$

The present paper is part of the PhD-Thesis of the author. Our intention is to provide a broad overview over regularity theory for this kind of nondifferentiable higher order parabolic systems. For this reason, and for sake of brevity, we shall concentrate here on those points containing the most important technical innovations of the paper, confining ourself to giving only a sketch of those proofs that are closer to the ones for second-order systems, and that can be found scattered in the literature. The full proofs can be nevertheless retrieved from [5].

2 Notation and preliminary material

As basic sets for our estimates we usually take parabolic cylinders. If not differently mentioned we denote $Q_\rho(z_0) \equiv B_\rho(x_0) \times (t_0 - \rho^{2m}, t_0)$, where $z_0 = (x_0, t_0) \in \mathbb{R}^{n+1}$, $\rho > 0$ and $B_\rho(x_0)$ is the open ball in $\mathbb{R}^n$ with center x_0 and radius ρ. If $z_0 = 0$, we abbreviate $Q_\rho = Q_\rho(0)$ and $B_\rho = B_\rho(0)$. Moreover, for an integrable function $v\colon Q_\rho(z_0) \to \mathbb{R}^k$, $k \in \mathbb{N}$ we write $(v)_{z_0,\rho} \equiv ⨍_{Q_\rho(z_0)} v \, \mathrm{d}z$ for its mean-value on the parabolic cylinder $Q_\rho(z_0)$.

We now summarize some conclusions of our assumptions on the coefficients that will be used in several points of the paper. From the ellipticity (1.2) of $\partial_p A$ we infer that A is monotone with respect to p, i.e. for all $z \in \Omega_T$, $\xi \in \mathbb{R}^{\mathscr{M}}$ and $p, \widetilde{p} \in \mathbb{R}^{\mathscr{N}}$ there holds

$$\big(A(z, \xi, p) - A(z, \xi, \widetilde{p})\big) \cdot (p - \widetilde{p}) \geq \nu\, |p - \widetilde{p}|^2. \tag{2.1}$$

Moreover, we will use that θ from (1.6) is a concave function with respect to s and

$$\theta\big(|\xi| + |\xi_0|, \mathrm{d}_{\mathscr{P}}(z, z_0) + |\xi - \xi_0|\big) \leq K(2|\xi_0| + 1)\, \big(\mathrm{d}_{\mathscr{P}}(z, z_0) + |\xi - \xi_0|\big)^\beta, \tag{2.2}$$

which can be seen by considering the cases $|\xi - \xi_0| \leq 1$ (then $|\xi| + |\xi_0| \leq 2|\xi_0| + 1$) and $|\xi - \xi_0| > 1$ (then the term on the right-hand side is > 1; the one on the left-hand side is always ≤ 1). We further set

$$H(s) \equiv K(2s + 1)\, (1 + s).$$

Then, combining (1.6) and (2.2) we have

$$|A(z, \xi, p) - A(z_0, \xi_0, p)| \leq L\, H(M)\, \left(\mathrm{d}_{\mathscr{P}}(z, z_0) + |\xi - \xi_0|\right)^\beta, \tag{2.3}$$

provided that $|\xi_0| \leq M$ and $|p| \leq M$. By virtue of the continuity of $\partial_p A$ there exists for each $M > 0$ a modulus of continuity $\omega_M\colon [0, \infty) \to [0, 1]$ with $\lim_{s \searrow 0} \omega_M(s) = 0$ for all $M > 0$, such that $M \mapsto \omega_M(s)$ is non-deceasing for fixed $s \geq 0$ and $s \mapsto \omega_M(s)$ is non-decreasing and $s \mapsto \omega_M(s)^2$ is concave for fixed $M > 0$, with the property that

$$|\partial_p A(z, \xi, p) - \partial_p A(z_0, \xi_0, p_0)| \leq 2L\, \kappa_M\, \omega_M\left(\mathrm{d}_{\mathscr{P}}(z, z_0)^2 + |\xi - \xi_0|^2 + |p - p_0|^2\right) \tag{2.4}$$

for all $z, z_0 \in \Omega_T$, $\xi, \xi_0 \in \mathbb{R}^{\mathscr{M}}$ and $p, p_0 \in \mathbb{R}^{\mathscr{N}}$ with $|\xi| + |p| \leq M$ and $|\xi_0| + |p_0| \leq M$.

2.1 Estimates for polynomials

In order to treat regularity problems for elliptic respectively parabolic systems one usually needs to control oscillation quantities of the solution to measure in a weak sense its regularity. In any case polynomials, especially the mean value polynomials and the minimizing polynomials, will play an important role. Now, we will establish the basic facts and estimates used throughout the paper. The first lemma is an immediate consequence of Taylor's expansion [5, Lemma A.2].

Lemma 2.1 *Let* $P\colon \mathbb{R}^n \to \mathbb{R}^N$ *be a polynomial of degree* $\le m$ *and* $B_\rho(x_0) \subset \mathbb{R}^n$ *with* $\rho \le 1$. *Then*

$$|\delta P(x) - \delta P(x_0)| \le c(m)\, \rho\, \big(|\delta P(x_0)| + |D^m P|\big) \qquad \text{for all} \quad x \in B_\rho(x_0).$$

For instance we can attain $c(m) = 2m\sqrt{m}$.

We can represent any polynomial by its mean values on balls, [15]. This representation allows us to estimate our polynomial in terms of these mean values [5, Lemma A.1].

Lemma 2.2 *Let* $P\colon \mathbb{R}^n \to \mathbb{R}^N$ *be a polynomial of degree* $\le m$ *and* $B_\rho(x_0) \subset \mathbb{R}^n$. *Then for all* $0 \le k \le m-1$ *there holds, with* $(D^j P)_{x_0,r} = \fint_{B_r(x_0)} D^j P \,\mathrm{d}y$:

$$|D^k P(x)| \le c(n,m) \sum_{j=k}^{m} \rho^{j-k} |(D^j P)_{x_0,r}| \qquad \text{for all } x \in B_\rho(x_0).$$

2.2 Technical Lemma

In some places, i.e. in the proof of the Caccioppoli inequality we will "absorb" certain integrals of the right-hand side. For this we will need the following lemma, which is standard and can be found for instance in [23, p. 161].

Lemma 2.3 *Let* $0 < \vartheta < 1$, $A, B \ge 0$, $\alpha > 0$ *and let* $f \ge 0$ *be a bounded function satisfying*

$$f(t) \le \vartheta f(s) + A(s-t)^\alpha + B$$

for all $0 < r \le t < s \le \rho$. *Then there exists a constant* $c_{\mathrm{tech}} = c_{\mathrm{tech}}(\alpha, \vartheta)$, *such that*

$$f(r) \le c_{\mathrm{tech}} \left(A(\rho - r)^{-\alpha} + B\right).$$

2.3 Interpolation Lemma

We now state an interpolation lemma for intermediate derivatives on the annulus, similar to [4, Theorem 4.14]. For our purpose the "right" scaling on the annulus will be essential. We refer to [5, Lemma B.1] for a detailed proof. Later, we will apply this lemma several times on the horizontal time slices.

Lemma 2.4 *Let* B_r, $B_R \subset \mathbb{R}^n$ *be two balls with the same center and radius* r, *respectively* R, *where* $0 < r < R \le 1$ *and let* $u \in W^{m,p}(B_R)$ *with* $p \ge 1$. *Then for any* $0 \le k \le m-1$ *and* $0 < \varepsilon \le 1$ *there holds*

$$\int_{B_R \setminus B_r} \frac{|D^k u|^p}{(R-r)^{p(m-k)}} \,\mathrm{d}x \le \varepsilon \int_{B_R \setminus B_r} |D^m u|^p \,\mathrm{d}x + c(n,m,p)\, \varepsilon^{-\frac{k}{m-k}} \int_{B_R \setminus B_r} \frac{|u|^p}{(R-r)^{pm}} \,\mathrm{d}x.$$

2.4 Steklov-means

Since weak solutions u of parabolic systems possess only very weak regularity properties with respect to the time variable t (i.e. they are not assumed to be weakly differentiable), they are in principle no admissible test-functions (also disregarding boundary values). In order to be nevertheless able to test properly, we smooth the function u in the time direction t, using the so-called Steklov-means. This also enables us to work on the time-slices $\mathbb{R}^n \times \{t\}$, even if u is only an L^2-function with respect to t. Given a function $f \in L^1(\Omega_T)$ and $0 < h < T$, we define its *Steklov-mean* by

$$[f]_h(x,t) \equiv \begin{cases} \dfrac{1}{h}\displaystyle\int_t^{t+h} f(x,s)\,\mathrm{d}s, & t \in (-T,-h), \\ 0, & t \in (-h,0), \end{cases} \tag{2.5}$$

respectively

$$[f]_{\bar h}(x,t) \equiv \begin{cases} \dfrac{1}{h}\displaystyle\int_{t-h}^{t} f(x,s)\,\mathrm{d}s, & t \in (-T+h,0), \\ 0, & t < -T+h. \end{cases}$$

Then $[f]_h \to f$ and $[f]_{\bar h} \to f$ as $h \searrow 0$ a.e. in Ω_T and

$$\partial_t[f]_h(x,t) = \tfrac{1}{h}\left(f(x,t+h) - f(x,t)\right), \quad \partial_t[f]_{\bar h}(x,t) = \tfrac{1}{h}\left(f(x,t) - f(x,t-h)\right).$$

Moreover, we have

$$\int_{t_1}^{t_2}\int_\Omega |[f]_h|^2\,\mathrm{d}x\,\mathrm{d}t \le \int_{t_1}^{t_2+h}\int_\Omega |f|^2\,\mathrm{d}x\,\mathrm{d}t \quad \text{and} \quad \int_{t_1}^{t_2}\int_\Omega |[f]_{\bar h}|^2\,\mathrm{d}x\,\mathrm{d}t \le \int_{t_1-h}^{t_2}\int_\Omega |f|^2\,\mathrm{d}x\,\mathrm{d}t \tag{2.6}$$

for $-T \le t_1 < t_2 \le -h$, respectively for $-T+h \le t_1 < t_2 \le 0$. Rewriting system (1.1) with Steklov-means $[u]_h$ of u, we obtain the following system on the time-slices $\Omega \times \{t\}$

$$\int_\Omega \partial_t[u]_h(\cdot,t)\cdot\varphi + \left[A(\cdot,t,\delta u(\cdot,t),D^m u(\cdot,t))\right]_h \cdot D^m\varphi dx$$
$$= -\int_\Omega \left[B(\cdot,t,\delta u(\cdot,t),D^m u(\cdot,t))\right]_h \cdot \delta\varphi dx \tag{2.7}$$

for all $\varphi \in W_0^{m,2}(\Omega;\mathbb{R}^N)$ and for a.e. t in $(-T,0)$.

3 Partial Regularity

3.1 A-polycaloric approximation

Our main tool in proving partial regularity is the lemma of A-polycaloric approximation, stating that whenever a function u is in a certain sense approximately a solution of a linear parabolic system, then there exists a solution g of this linear system which is in some sense "near" to u.

Lemma 3.1 *Given $\varepsilon > 0$ there is a constant $\delta = \delta(n,N,m,\nu,\Lambda,\varepsilon) \in (0,1]$ with the following property: Whenever A is a strongly elliptic bilinear form on $\mathbb{R}^{\mathcal{N}}$ with ellipticity constant $\nu > 0$ and upper bound Λ, i.e. $\nu|p|^2 \le A(p,p)$ and $A(p,\tilde p) \le \Lambda|p||\tilde p|$ for $p,\tilde p \in \mathbb{R}^{\mathcal{N}}$ and $u \in L^2(t_0-\rho^{2m},t_0;W^{m,2}(B_\rho(x_0);\mathbb{R}^N))$ with*

$$\sum_{k=0}^m \fint_{Q_\rho(z_0)} \left|\frac{D^k u}{\rho^{m-k}}\right|^2 \mathrm{d}z \le 1,$$

is approximately A-polycaloric in the sense that

$$\left| \fint_{Q_\rho(z_0)} \big(u \cdot \varphi_t - A(D^m u, D^m \varphi)\big)\, \mathrm{d}z \right| \le \delta \sup_{Q_\rho(z_0)} |D^m \varphi|, \qquad \textit{for all} \quad \varphi \in C_0^\infty(Q_\rho(z_0); \mathbb{R}^N),$$

then there exists an A-caloric function $g \in L^2(t_0 - \rho^{2m}, t_0;\, W^{m,2}(B_\rho(x_0); \mathbb{R}^N))$*, i.e.*

$$\int_{Q_\rho(z_0)} \big(g \cdot \varphi_t - A(D^m g, D^m \varphi)\big)\, \mathrm{d}z = 0 \quad \textit{for all } \varphi \in C_0^\infty(Q_\rho(z_0); \mathbb{R}^N),$$

such that

$$\sum_{k=0}^{m} \fint_{Q_\rho(z_0)} \left| \frac{D^k g}{\rho^{m-k}} \right|^2 \mathrm{d}z \le 1 \quad \textit{and} \quad \sum_{k=0}^{m-1} \fint_{Q_\rho(z_0)} \left| \frac{D^k (u-g)}{\rho^{m-k}} \right|^2 \mathrm{d}z \le \varepsilon.$$

Proof We will only sketch the proof and refer the reader to [5, Lemma 3.3], for a detailed proof. Without loss of generality we may assume that $z_0 = 0$ and $\rho = 1$. Otherwise we rescale u to $Q_1(0)$ via $v(x,t) \equiv \rho^{-m} u(x_0 + \rho x, t_0 + \rho^{2m} t)$. Thus it is enough to show the assertion on $Q \equiv Q_1(0) \equiv B \times (-1, 0)$.

Supposed the conclusion of the lemma were wrong, then there would exist $\varepsilon > 0$ and a sequence $(A_j)_{j\in\mathbb{N}}$ of bilinear forms on $\mathbb{R}^{\mathscr{N}}$ with uniform ellipticity constant $\nu > 0$ and upper bound Λ and a sequence of functions $(v_j)_{j\in\mathbb{N}}$ with $v_j \in L^2(-1, 0; W^{m,2}(B; \mathbb{R}^N))$, such that

$$\sum_{k=0}^{m} \fint_Q |D^k v_j|^2\, \mathrm{d}z \le 1 \qquad \text{and} \qquad \left| \fint_Q \big(v_j \cdot \varphi_t - A_j(D^m v_j, D^m \varphi)\big)\, \mathrm{d}z \right| \le \frac{1}{j} \sup_Q |D^m \varphi| \tag{3.1}$$

for all $\varphi \in C_0^\infty(Q; \mathbb{R}^N)$ and $j \in \mathbb{N}$, but

$$\fint_Q |v_j - g|^2\, \mathrm{d}z > \varepsilon \tag{3.2}$$

for all A_j-caloric functions g on Q with $\sum_{k=0}^m \fint_Q |D^k g|^2\, \mathrm{d}z \le 1$. By the uniform boundedness of $\|D^k v_j\|_{L^2}$ for $0 \le k \le m$ there exists a subsequence (also labelled with j), a function $v \in L^2(-1, 0;\, W^{m,2}(B; \mathbb{R}^N))$ and a bilinear form A, such that

$$\begin{cases} D^k v_j \rightharpoonup D^k v & \text{weakly in } L^2(Q; \mathbb{R}^{\mathscr{M}_k}) \quad \text{for all} \quad 0 \le k \le m, \\ A_j \to A & \text{as bilinear forms on } \mathbb{R}^{\mathscr{N}}. \end{cases}$$

Since $v \mapsto \sum_{k=0}^m \fint_Q |D^k v|^2\, \mathrm{d}z$ is weakly lower semicontinuous, we get from (3.1) that $\sum_{k=0}^m \fint_Q |D^k v|^2\, \mathrm{d}z \le 1$. Due to the convergence $A_j \to A$, the weak convergence of v_j and $D^m v_j$, the uniform boundedness of $|D^m v_j|$ in $L^2(Q)$ and (3.1), i.e. the fact that v_j is approximately A_j-caloric, we can show that v is an A-caloric function on Q, i.e.

$$\int_Q \big(v \cdot \varphi_t - A(D^m v, D^m \varphi)\big)\, \mathrm{d}z = 0 \qquad \text{for all} \quad \varphi \in C_0^\infty(Q; \mathbb{R}^N). \tag{3.3}$$

In order to derive the contradiction in (3.2) we have to show strong convergence of $D^k v_j$ to $D^k v$ in $L^2(Q; \mathbb{R}^{\mathscr{M}_k})$ for $0 \le k \le m-1$. The compactness argument which is applied in this situation in the elliptic case, cannot be used, since v_j is possibly not differentiable with respect to t. Instead, we use a different argument. Exploiting the fact that v_j is approximately A_j-caloric, we can show that

$$\lim_{h\searrow 0}\int_{-1}^{-h}\|D^k\left(v_j(\cdot,t+h)-v_j(\cdot,t)\right)\|^2_{L^2(B)}\,\mathrm{d}t=0$$

uniformly with respect to j for all $0\le k\le m-1$. Since we also know that $(D^k v_j)_{j\in\mathbb{N}}$ is uniformly bounded in $L^1_{loc}(-1,0;\,L^2(B))$, Theorem 3 in [33] ensures the existence of a subsequence $(v_j)_{j\in\mathbb{N}}$ (also labelled with j), converging strongly in $W^{m-1,2}$, i.e. for $0\le k\le m-1$ there holds

$$D^k v_j\to D^k v\qquad\text{strongly in}\quad L^2(Q;\mathbb{R}^{\mathscr{M}_k}).$$

From (3.3) we know that v is an A-caloric function. In the following we will derive the contradiction by constructing suitable A_j-caloric functions. For this, let $w_j\in C([-1,0];\,L^2(B,\mathbb{R}^N))\cap L^2(-1,0;\,W^{m,2}_0(B,\mathbb{R}^N))$, $\partial_t w_j\in L^2(-1,0;\,W^{-m,2}(B,\mathbb{R}^N))$ be the weak solution of the initial value problem:

$$\int_Q\big(w_j\cdot\varphi_t-A_j(D^m w_j,D^m\varphi)\big)\,\mathrm{d}z=-\int_Q A_j(D^m v,D^m\varphi)\,\mathrm{d}z,\quad\text{for all}\quad\varphi\in C_0^\infty(Q,\mathbb{R}^N),$$

and $w_j(\cdot,-1)\equiv 0$. Then, testing the system with $\varphi=w_j$, we can infer that

$$\sup_{t\in(-1,0)}\frac{1}{2}\|w_j(\cdot,t)\|^2_{L^2(B)}+\frac{\nu}{2}\int_Q|D^m w_j|^2\,\mathrm{d}z\to 0\quad\text{as}\ \ j\to\infty. \tag{3.4}$$

We now define $g_j\equiv v-w_j\in L^2(-1,0;\,W^{m,2}(B;\mathbb{R}^N))$. Then, g_j agrees with v on the parabolic boundary $(B\times\{-1\})\cup(\partial B\times(-1,0))$ of Q, since w_j vanishes there. Furthermore g_j is A_j-caloric and from (3.4) we infer that $\|g_j-v\|_{L^2(Q)}+\|D^m(g_j-v)\|_{L^2(Q)}\to 0$ as $j\to\infty$. This implies $g_j\to v$ in $L^2(-1,0;\,W^{m,2}(B;\mathbb{R}^N))$ and therefore, $\sum_{k=0}^m\int_Q|D^k g_j|^2\,\mathrm{d}z$ is bounded. Letting $\widetilde{g_j}\equiv\frac{g_j}{b_j}$, $b_j\equiv\max\{1,\ \sum_{k=0}^m\int_Q|D^k g_j|^2\,\mathrm{d}z\}$ we finally obtain the contradiction, because $\widetilde{g_j}$ is A_j-caloric, $\sum_{k=0}^m\int_Q|D^k\widetilde{g}_j|^2\,\mathrm{d}z\le 1$ for all $j\in\mathbb{N}$ and

$$\int_Q|D^k\left(v-\widetilde{g_j}\right)|^2\,\mathrm{d}z\le 2\int_Q|D^k\left(v-g_j\right)|^2\,\mathrm{d}z+2\left(1-\frac{1}{b_j}\right)\int_Q|D^k g_j|^2\,\mathrm{d}z\to 0$$

as $j\to\infty$, for $0\le k\le m-1$, since $D^k g_j\to D^k v$ in L^2 and $b_j\to 1$. This is the desired contradiction to (3.2), because we also know that $D^k v_j\to D^k v$ in L^2. □

3.2 Caccioppoli inequality

As usual, the first step in proving partial regularity is a suitable Caccioppoli inequality.

Lemma 3.2 *Suppose that $u\in L^2(-T,0;\,W^{m,2}(\Omega;\mathbb{R}^N))$ is a weak solution of (1.1) in Ω_T under the assumptions (1.2)–(1.6) and let $Q_\rho(z_0)\Subset\Omega_T$, with $0<\rho\le 1$. Then for $M>0$, and for all polynomials $P:\mathbb{R}^n\to\mathbb{R}^N$ independent of t of degree $\le m$, fulfilling*

$|\delta P(x)| + |D^m P| \le M$ for $x \in B_\rho(x_0)$ there holds

$$\sup_{t\in(t_0-(\rho/2)^{2m},t_0)} \fint_{B_{\rho/2}(x_0)} \frac{|u(\cdot,t)-P|^2}{\rho^{2m}}\,\mathrm{d}x + \fint_{Q_{\rho/2}(z_0)} |D^m(u-P)|^2\,\mathrm{d}z \le c_{Cac} \fint_{Q_\rho(z_0)} \frac{|u-P|^2}{\rho^{2m}} + \rho^{2\beta}\,\mathrm{d}z,$$

where c_{Cac} depends on $n, \nu, L, M, H(M), \kappa_{M+1}$.

Proof Without loss of generality we can assume that $z_0 = (x_0, t_0) = 0$, i.e. $Q_\rho(z_0) = Q_\rho = B_\rho \times (-\rho^{2m}, 0)$. We choose two cut-off functions $\eta \in C_0^\infty(B_\rho)$ and $\zeta \in C^1(\mathbb{R})$ with the properties

$$\begin{cases} \eta \equiv 1 \text{ in } B_{\rho/2}, \quad 0 \le \eta \le 1, \quad |\nabla \eta| \le c/\rho; \\ \zeta \equiv 1 \text{ on } \left(-(\rho/2)^{2m}, \infty\right), \quad \zeta \equiv 0 \text{ on } (-\infty, -\rho^{2m}), \quad 0 \le \zeta \le 1, \quad 0 \le \zeta' \le 2/\rho^{2m}. \end{cases}$$

Choosing the test-function $\varphi_h = \eta^2\zeta^2(u_h - \ell)$, where u_h denotes the Steklov-mean of u defined in (2.5), in the Steklov-formulation (2.7) of the system, we obtain for a.e. $\tau \in (-\rho^{2m}, 0)$ that

$$\int_{B_\rho} \big(\partial_t u_h \cdot \varphi_h + [A(\cdot,\cdot,u,Du)]_h \cdot D\varphi_h\big)(\cdot,\tau)\,\mathrm{d}x = 0. \tag{3.5}$$

Noting that $\partial_t P \equiv 0$, we infer for a.e. $t \in (-\rho^{2m}, 0)$ that

$$\begin{aligned} \int_{-\rho^{2m}}^{t}\int_{B_\rho} \partial_t[u]_h \cdot \varphi_h\,\mathrm{d}x\mathrm{d}\tau &= \int_{-\rho^{2m}}^{t}\int_{B_\rho} \left(\frac{1}{2}\,\partial_t\left(|[u]_h - P|^2\zeta^2\right)\eta - |[u]_h - P|^2\eta\zeta\zeta'\right)\mathrm{d}x\mathrm{d}\tau \\ &= \frac{1}{2}\int_{B_\rho} |[u]_h(\cdot,t) - P|^2\eta\zeta(t)^2\,\mathrm{d}x - \int_{-\rho^{2m}}^{t}\int_{B_\rho} |[u]_h - P|^2\eta\zeta\zeta'\,\mathrm{d}x\,\mathrm{d}\tau. \end{aligned}$$

Therefore, integrating (3.5) over $(-\rho^{2m}, t)$, using the previous identity and passing to the limit $h \searrow 0$ yields for a.e. $t \in (-\rho^{2m}, 0)$ that

$$\begin{aligned} &\frac{1}{2}\int_{B_\rho} |u(\cdot,t) - P|^2\eta\zeta(t)^2\,\mathrm{d}x + \int_{-\rho^{2m}}^{t}\int_{B_\rho} A(\cdot,\delta u, D^m u)\cdot D^m(u-P)\eta\zeta^2\,\mathrm{d}z \\ &\quad = \int_{-\rho^{2m}}^{t}\int_{B_\rho} \left(-A(\cdot,\delta u, D^m u)\cdot \mathrm{LOT}\,\zeta^2 - B(\cdot,\delta u, D^m u)\cdot \delta\varphi + |u-P|^2\eta\zeta\zeta'\right)\mathrm{d}z, \end{aligned}$$

where we have denoted $\mathrm{d}z = \mathrm{d}x\,\mathrm{d}\tau$ and $\varphi \equiv \eta\zeta^2(u-P)$ and

$$D^m\varphi = \zeta^2\Bigg(D^m(u-P)\eta + \underbrace{\sum_{k=0}^{m-1}\binom{m}{k} D^k(u-P)\odot D^{m-k}\eta}_{\equiv \mathrm{LOT}}\Bigg).$$

Furthermore we have

$$\int_{-\rho^{2m}}^{t}\int_{B_\rho} A(\cdot,\delta u, D^m P)\cdot D^m(u-P)\eta\zeta^2\,\mathrm{d}z$$
$$= \int_{-\rho^{2m}}^{t}\int_{B_\rho} \big(A(\cdot,\delta u, D^m P)\cdot D^m\varphi - A(\cdot,\delta u, D^m P)\cdot \mathrm{LOT}\,\zeta^2\big)\,\mathrm{d}z$$

and

$$\int_{-\rho^{2m}}^{t}\int_{B_\rho} A(0,\delta P(0), D^m P)\cdot D^m\varphi\,\mathrm{d}z = 0.$$

Combining the previous identities, using the monotonicity (2.1) of A, i.e. that $(A(\cdot,\delta u, D^m u) - A(\cdot,\delta u, D^m P))\cdot D^m(u-P) \ge \nu|D^m(u-P)|^2$ and noting that $\eta\equiv 1$ on B_{ρ_1} we get for a.e. $t\in(-\rho^{2m},0)$:

$$\frac{1}{2}\int_{B_{\rho_1}} |u(\cdot,t)-P|^2\zeta(t)^2\,\mathrm{d}x + \nu\int_{-\rho^{2m}}^{t}\int_{B_{\rho_1}} |D^m(u-P)|^2\zeta^2\,\mathrm{d}z$$
$$= -\int_{-\rho^{2m}}^{t}\int_{B_\rho}\big(A(\cdot,\delta u, D^m u) - A(\cdot,\delta u, D^m P)\big)\cdot \mathrm{LOT}\,\zeta^2\,\mathrm{d}z$$
$$-\int_{-\rho^{2m}}^{t}\int_{B_\rho}\big(A(\cdot,\delta u, D^m P) - A(\cdot,\delta P, D^m P)\big)\cdot D^m\varphi\,\mathrm{d}z$$
$$-\int_{-\rho^{2m}}^{t}\int_{B_\rho}\big(A(\cdot,\delta P, D^m P) - A(0,\delta P(0), D^m P)\big)\cdot D^m\varphi\,\mathrm{d}z$$
$$-\int_{-\rho^{2m}}^{t}\int_{B_\rho} B(\cdot,\delta u, D^m u)\cdot\delta\varphi\,\mathrm{d}z + \int_{-\rho^{2m}}^{t}\int_{B_\rho} |u-P|^2\eta\zeta\zeta'\,\mathrm{d}z$$
$$= I_1 + I_2 + I_3 + I_4 + I_5, \tag{3.6}$$

with the obvious meaning of $I_1 - I_5$. We now derive estimates for $I_1 - I_5$. Thereby we take $\varepsilon\in(0,1]$.

Estimate for I_1. We once again decompose $I_1 = I_{1,1} + I_{1,2} + I_{1,3}$, where

$$I_{1,1} \equiv -\int_{-\rho^{2m}}^{t}\int_{B_\rho}\big(A(\cdot,\delta u, D^m u) - A(\cdot,\delta P, D^m u)\big)\cdot \mathrm{LOT}\,\zeta^2\,\mathrm{d}z,$$

$$I_{1,2} \equiv -\int_{-\rho^{2m}}^{t}\int_{B_\rho}\big(A(\cdot,\delta P, D^m u) - A(\cdot,\delta P, D^m P)\big)\cdot \mathrm{LOT}\,\zeta^2\,\mathrm{d}z,$$

$$I_{1,3} \equiv -\int_{-\rho^{2m}}^{t}\int_{B_\rho} \big(A(\cdot,\delta P, D^m P) - A(\cdot,\delta u, D^m P)\big)\cdot \mathrm{LOT}\,\zeta^2\,\mathrm{d}z.$$

To estimate $I_{1,1}$ we use (1.6) and the assumption $|D^m P| \le M$. Then, we exploit our assumptions on θ, namely, for the term involving $1+M$ we use (2.2), (note that $|\delta P(x)| \le M$ for $x \in B_\rho$ by assumption and that $H(M) = K(2M+1)(1+M)$) and for the term involving $|D^m(u-P)|$ we use that $\theta \le 1$. Also applying Hölder's inequality in the last line we infer that

$$\begin{aligned}
|I_{1,1}| &\le L\int_{Q_{\rho_2}} \theta\,(|\delta u| + |\delta P|, |\delta u - \delta P|)\,\big((1+M) + |D^m(u-P)|\big)\;|\mathrm{LOT}|\zeta^2\,\mathrm{d}z\\
&\le L\int_{Q_{\rho_2}} \big((1+M)K(2M+1)|\delta u - \delta P|^\beta + |D^m(u-P)|\big)\,|\mathrm{LOT}|\zeta^2\,\mathrm{d}z\\
&\le L\Bigg(\int_{Q_{\rho_2}} \big(H(M)^2|\delta u - \delta P|^{2\beta} + |D^m(u-P)|^2\big)\,\zeta^2\,\mathrm{d}z\Bigg)^{\frac12}\Bigg(\int_{Q_{\rho_2}} |\mathrm{LOT}|^2\zeta^2\,\mathrm{d}z\Bigg)^{\frac12}.
\end{aligned}$$

To estimate $I_{1,2}$ we decompose $Q_{\rho_2} = S_1 \cup S_2$, where $S_1 \equiv \{z \in Q_{\rho_2} : |D^m(u-P)| \le 1\}$, $S_2 \equiv \{z \in Q_{\rho_2} : |D^m(u-P)| > 1\}$ and rewrite $I_{1,2}$ as follows

$$\begin{aligned}
I_{1,2} &\le \int_{Q_{\rho_2}} |A(\cdot,\delta P, D^m u) - A(\cdot,\delta P, D^m P)|\,|\mathrm{LOT}|\,\zeta^2\,\mathrm{d}z\\
&= \int_{S_1} (\ldots)\;\mathrm{d}z + \int_{S_2} (\ldots)\;\mathrm{d}z \equiv I_{1,2,1} + I_{1,2,2},
\end{aligned}$$

with the obvious labelling of $I_{1,2,1}$ and $I_{1,2,2}$. For $I_{1,2,1}$ we use (1.5) and Hölder's inequality and note that $|\delta P(x)| \le M$ for $x \in B_\rho$ and that $|D^m P + s(D^m u - D^m P)| \le M+1$ on S_1 to obtain

$$\begin{aligned}
I_{1,2,1} &\le \int_{S_1}\Bigg|\int_0^1 \frac{\partial A}{\partial p}\big(z,\delta P, D^m P + s(D^m u - D^m P)\big)\,(D^m u - D^m P)\,\mathrm{d}s\Bigg|\,|\mathrm{LOT}|\,\zeta^2\,\mathrm{d}z\\
&\le L\,\kappa_{M+1}\int_{S_1} |D^m(u-P)|\,|\mathrm{LOT}|\,\zeta^2\,\mathrm{d}z\\
&\le L\,\kappa_{M+1}\Bigg(\int_{Q_{\rho_2}} |D^m(u-P)|^2\zeta^2\,\mathrm{d}z\Bigg)^{\frac12}\Bigg(\int_{Q_{\rho_2}} |\mathrm{LOT}|^2\zeta^2\,\mathrm{d}z\Bigg)^{\frac12}.
\end{aligned}$$

For $I_{2,2}$ we use the growth condition (1.3) instead of (1.5), the assumption $|D^m P| \le M$, the fact that $|D^m(u-P)| > 1$ on S_2 and Hölder's inequality to obtain

$$I_{1,2,2} \le L \int_{S_2} (2 + M + |D^m u|)\, |\mathrm{LOT}|\, \zeta^2\, \mathrm{d}z \le 3L\,(1+M) \int_{S_2} |D^m(u-P)|\, |\mathrm{LOT}|\, \zeta^2\, \mathrm{d}z$$
$$\le 3L\,(1+M)\Bigg(\int_{Q_{\rho_2}} |D^m(u-P)|^2 \zeta^2\Bigg)^{\frac12}\Bigg(\int_{Q_{\rho_2}} |\mathrm{LOT}|^2 \zeta^2\, \mathrm{d}z\Bigg)^{\frac12}.$$

To estimate $I_{1,3}$ we use (2.3) (note that $|\delta P(x)| \le M$ and $|D^m P| \le M$) and Hölder's inequality

$$I_{1,3} \le L\, H(M)\Bigg(\int_{Q_{\rho_2}} |\delta u - \delta P|^{2\beta} \zeta^2\, \mathrm{d}z\Bigg)^{\frac12}\Bigg(\int_{Q_{\rho_2}} |\mathrm{LOT}|^2 \zeta^2\, \mathrm{d}z\Bigg)^{\frac12}.$$

Together the previous estimates result in

$$I_1 \le c\, L\Bigg(\int_{Q_{\rho_2}} \big(|D^m(u-P)|^2 + |\delta u - \delta P|^{2\beta}\big)\, \zeta^2\, \mathrm{d}z\Bigg)^{\frac12}\Bigg(\int_{Q_{\rho_2}} |\mathrm{LOT}|^2 \zeta^2\, \mathrm{d}z\Bigg)^{\frac12}, \tag{3.7}$$

where c depends on M, $H(M)$ and κ_{M+1}. With Young's inequality we infer that

$$I_1 \le \varepsilon \int_{Q_{\rho_2}} |D^m(u-P)|^2 \zeta^2\, \mathrm{d}z + c\,\frac{L^2}{\varepsilon} \int_{Q_{\rho_2}} \big(|\mathrm{LOT}|^2 + |\delta u - \delta P|^{2\beta}\big)\, \zeta^2\, \mathrm{d}z.$$

Estimate for I_2. We use the Hölder-continuity (1.6) of the mapping $\xi \mapsto \frac{A(z,\xi,p)}{1+|p|}$ in the form (2.3), (which is applicable since we know $|\delta P(x)| \le M$ for $x \in B_\rho$ and $|D^m P| \le M$). Applying Young's inequality afterwards yields for $\varepsilon > 0$

$$|I_2| \le L\, H(M) \int_{Q_{\rho_2}} |\delta(u-P)|^\beta \big(|D^m(u-P)| + |\mathrm{LOT}|\big)\, \zeta^2\, \mathrm{d}z$$
$$\le \varepsilon \int_{Q_{\rho_2}} |D^m(u-P)|^2 \zeta^2\, \mathrm{d}z + \int_{Q_{\rho_2}} \Big(|\mathrm{LOT}|^2 + \tfrac{L^2\, H(M)^2}{\varepsilon} |\delta(u-P)|^{2\beta}\Big)\, \zeta^2\, \mathrm{d}z.$$

Estimate for I_3. Exploiting once again the Hölder-continuity (2.3) of coefficients and the fact that $|\delta P(x) - \delta P(0)| \le c\,\rho\, M$ for $x \in B_\rho$ (by Lemma 2.1) we infer that

$$|I_3| \le c\, L\, H(M)\,(1+M)^\beta \rho^\beta \int_{Q_{\rho_2}} \big(|D^m(u-P)| + |\mathrm{LOT}|\big)\, \zeta^2\, \mathrm{d}z$$
$$\le \varepsilon \int_{Q_{\rho_2}} |D^m(u-P)|^2 \zeta^2\, \mathrm{d}z + \varepsilon \int_{Q_{\rho_2}} |\mathrm{LOT}|^2 \zeta^2\, \mathrm{d}z + \frac{c(M, H(M))\, L^2}{\varepsilon}\, \rho^{2\beta}\, |Q_\rho|.$$

Estimate for I_4. From the growth (1.4) of B and the fact that $|D^m P| \le M$ we find that

$$|I_4| \le L \int_{Q_{\rho_2}} \big(1 + M + |D^m(u-P)|\big)\, |\delta\varphi|\, \mathrm{d}z$$
$$\le \varepsilon \int_{Q_{\rho_2}} |D^m(u-P)|^2 \zeta^2\, \mathrm{d}z + \frac{2L^2}{\varepsilon} \int_{Q_{\rho_2}} \frac{|\delta((u-P)\eta)|^2}{\rho^2} \zeta^2\, \mathrm{d}z + (1+M)^2 \rho^2\, |Q_\rho|.$$

Estimate for I_5. For I_5, we only note that $|\zeta'| \le \frac{2}{\rho^{2m}} \le \frac{2}{(\rho_2-\rho_1)^{2m}}$.

We insert the estimates for $I_1 - I_5$ into (3.6), take the supremum over $t \in (-(\rho/2)^{2m}, 0)$ in the first term on the left-hand side and take $t = 0$ in the second-term and note that $\eta \equiv 1$ on B_{ρ_1} to infer that

$$\frac{1}{2} \sup_{t \in (-(\rho/2)^{2m}, 0)} \int_{B_{\rho_1}} |u(\cdot, t) - P|^2\, \mathrm{d}x + \nu \int_{Q_{\rho_1}} |D^m(u-P)|^2 \zeta^2\, \mathrm{d}z$$
$$\le 3\varepsilon \int_{Q_{\rho_2}} |D^m(u-P)|^2 \zeta^2\, \mathrm{d}z$$
$$+ c \int_{Q_{\rho_2}} \left(|\mathrm{LOT}|^2 \zeta^2 + \frac{|\delta((u-P)\eta)|^2}{\rho^2} \zeta^2 + |\delta(u-P)|^{2\beta} \zeta^2 + \frac{|u-P|^2}{(\rho_2-\rho_1)^{2m}} \right) \mathrm{d}z$$
$$+ c\, \rho^{2\beta}\, |Q_\rho|,$$

where $c = c(n, m, L, M, H(M), \kappa_{M+1}, 1/\varepsilon)$. Now we will observe bounds for the terms of lower order. Firstly we note that the integrand in the subsequent estimate differs from zero only on the annulus $B_{\rho_2} \setminus B_{\rho_1}$, due to the fact that $D^{m-k}\eta = 0$ on B_{ρ_1} for $0 \le k < m$. Applying the Interpolation-Lemma 2.4 "slicewise" on the annulus $B_{\rho_2} \setminus B_{\rho_1}$ for a.e. $t \in (-\rho^{2m}, 0)$ we obtain for $0 < \mu \le 1$ that

$$\int_{Q_{\rho_2}} |\mathrm{LOT}|^2 \zeta^2\, \mathrm{d}z \le c \sum_{k=0}^{m-1} \int_{-\rho^{2m}}^{0} \int_{B_{\rho_2} \setminus B_{\rho_1}} \frac{|D^k(u-P)|^2}{(\rho_2-\rho_1)^{2(m-k)}}\, \zeta^2\, \mathrm{d}z$$
$$\le \mu \int_{Q_{\rho_2}} |D^m(u-P)|^2 \zeta^2\, \mathrm{d}z + c(n, m, 1/\mu) \int_{Q_{\rho_2}} \frac{|u-P|^2}{(\rho_2-\rho_1)^{2m}}\, \mathrm{d}z.$$

Noting that $\rho_2 - \rho_1 \le \rho$ we get similarly

$$\int_{Q_{\rho_2}} \frac{|\delta((u-P)\eta)|^2}{\rho^2} \zeta^2\, \mathrm{d}z \le \mu \int_{Q_{\rho_2}} |D^m(u-P)|^2 \zeta^2\, \mathrm{d}z + c(n, m, 1/\mu) \int_{Q_{\rho_2}} \frac{|u-P|^2}{(\rho_2-\rho_1)^{2m}}\, \mathrm{d}z.$$

Furthermore we need the following estimate for the terms of lower order which result from the modulus of continuity of A. Using Young's inequality and the same arguments as before and noting that $\rho_2 \le \rho \le 1$ and $\rho^{\frac{2\beta(m-k)}{1-\beta}} \le \rho^{2\beta}$ for $k \le m-1$, we get for $0 < \mu \le 1$

$$\int_{Q_{\rho_2}} |\delta(u-P)|^{2\beta}\,\mathrm{d}z \le \sum_{k=0}^{m-1}\Bigg(\int_{Q_{\rho_2}} \frac{|D^k(u-P)|^2}{\rho^{2(m-k)}}\,\mathrm{d}z + |Q_\rho|\,\rho^{\frac{2\beta(m-k)}{1-\beta}}\Bigg)$$

$$\le \mu \int_{Q_{\rho_2}} |D^m(u-P)|^2\,\mathrm{d}z + c\int_{Q_{\rho_2}} \frac{|u-P|^2}{\rho^{2m}}\,\mathrm{d}z + c\,\rho^{2\beta}\,|Q_\rho|,$$

where $c = c(n, m, 1/\mu)$. Inserting the previous estimates above, choosing in turn μ and ε small enough and dividing by ν we infer that

$$\sup_{t\in(-(\rho/2)^{2m},0)} \int_{B_{\rho_1}} |u(\cdot,t)-P|^2\,\mathrm{d}x + \int_{-\rho^{2m}}^{0}\int_{B_{\rho_1}} |D^m(u-P)|^2\zeta^2\,\mathrm{d}z$$

$$\le \frac{1}{2}\int_{-\rho^{2m}}^{0}\int_{B_{\rho_2}} |D^m(u-P)|^2\zeta^2\,\mathrm{d}z + c\int_{Q_\rho} \frac{|u-P|^2}{(\rho_2-\rho_1)^{2m}}\,\mathrm{d}z + c\,\rho^{2\beta}|Q_\rho|,$$

where $c = c(n, m, \nu, L, M, H(M), \kappa_{M+1})$. Applying Lemma 2.3, we can "absorb" the first term of the right-hand side. Finally, noting that $\zeta \equiv 1$ on $(-(\rho/2)^{2m}, 0)$ we get the desired Caccioppoli inequality. □

Remark 3.3 In the case of simpler systems of the type (1.10), where A does not depend on δu, the weaker assumption $|D^m P| \le M$ on P suffices, since in this case (1.6) has the simpler form $|A(z,p) - A(z_0,p)| \le L\,\mathrm{d}_{\mathscr{P}}(z,z_0)^\beta(1+|p|)$ and the terms $I_{1,1}$, $I_{1,3}$ and I_2 do not appear in the above proof of the Caccioppoli inequality.

3.3 Linearization

Given a polynomial $P\colon \mathbb{R}^n \to \mathbb{R}^N$ of degree $\le m$ and $z_0 \in \Omega_T$ and a parabolic cylinder $Q_\rho(z_0) \Subset \Omega_T$ with $0 < \rho \le 1$. We define

$$\phi_u(z_0,\rho,D^mP) \equiv \fint_{Q_\rho(z_0)} |D^m u - D^m P|^2\,\mathrm{d}z \quad\text{and}\quad \psi_u(z_0,\rho,P) \equiv \fint_{Q_\rho(z_0)} \frac{|u-P|^2}{\rho^{2m}}\,\mathrm{d}z.$$

The conclusion of the following lemma is, that every weak solution of (1.1), fulfilling a suitable smallness condition, solves approximately a linear parabolic system. This property is needed to apply the A-polycaloric approximation lemma later.

Lemma 3.4 *Suppose that $u \in L^2(-T,0;W^{m,2}(\Omega;\mathbb{R}^N))$ is a weak solution of* (1.1) *in Ω_T under the assumptions* (1.2)–(1.6) *on A. Then for all $M > 0$, $Q_\rho(z_0) \Subset \Omega_T$, with $\rho \le 1$, $\varphi \in C_0^\infty(Q_\rho(z_0);\mathbb{R}^N)$ and each polynomial $P\colon \mathbb{R}^n \to \mathbb{R}^N$ of degree $\le m$ with $|\delta P(x_0)| + |D^m P| \le M$ for $x \in B_\rho(x_0)$ we have*

$$\left|\fint_{Q_\rho(z_0)} \left((u-P)\cdot\varphi_t - \frac{\partial A}{\partial p}\big(z_0,\delta P(x_0),D^mP\big)\,D^m(u-P)\cdot D^m\varphi\right)\mathrm{d}z\right|$$

$$\le c_{Eu}\left(\omega_{M+1}(\phi_u)\sqrt{\phi_u} + \phi_u + \psi_u + \rho^\beta\right)\sup_{Q_\rho(z_0)} |D^m\varphi|,$$

where $\phi_u \equiv \phi_u(z_0, \rho, D^m P)$ *and* $\psi_u \equiv \psi_u(z_0, \rho, P)$ *and* c_{Eu} *is of the form* $c_{Eu} = L\, c(n, m, H(M), \kappa_{M+1})$.

Proof Without loss of generality we may assume that $\sup_{Q_\rho(z_0)} |D^m \varphi| \le 1$ and that $z_0 = (x_0, t_0) = 0$. Noting that $⨍_{Q_\rho} P\varphi_t \,\mathrm{d}z = 0$ (since P does not depend on the time variable t) and $⨍_{Q_\rho} A(0, \delta P(0), D^m P) \cdot D^m \varphi \,\mathrm{d}z = 0$, we obtain from (1.1)

$$
\begin{aligned}
&⨍_{Q_\rho} \left((u - P) \cdot \varphi_t - \frac{\partial A}{\partial p} \big(0, \delta P(0), D^m P\big)\, D^m (u - P) \cdot D^m \varphi \right) \mathrm{d}z \\
&= ⨍_{Q_\rho} \left(A(0, \delta P(0), D^m u) - A(0, \delta P(0), D^m P) - \frac{\partial A}{\partial p} \big(0, \delta P(0), D^m P\big) D^m (u - P) \right) \cdot D^m \varphi \,\mathrm{d}z \\
&\quad + ⨍_{Q_\rho} \big(A(z, \delta u, D^m u) - A(z, \delta P, D^m u) \big) \cdot D^m \varphi \,\mathrm{d}z \\
&\quad + ⨍_{Q_\rho} \big(A(z, \delta P, D^m u) - A(0, \delta P(0), D^m u) \big) \cdot D^m \varphi \,\mathrm{d}z + ⨍_{Q_\rho} B(z, \delta u, D^m u) \cdot \delta\varphi \,\mathrm{d}z \\
&= I_1 + I_2 + I_3 + I_4,
\end{aligned}
$$

with the obvious meaning of $I_1 - I_4$. In the following we will derive estimates for $I_1 - I_4$.

Estimate for I_1**.** In order to use the modulus of continuity ω from (2.4), we decompose Q_ρ into $S_1 \equiv \{z \in Q_\rho : |D^m(u - P)| \le 1\}$, $S_2 \equiv \{z \in Q_\rho : |D^m(u - P)| > 1\}$ and write

$$
I_1 = \frac{1}{|Q_\rho|} \int_{S_1} (\cdots) \,\mathrm{d}z + \frac{1}{|Q_\rho|} \int_{S_2} (\cdots) \,\mathrm{d}z \equiv I_{1,1} + I_{1,2}.
$$

For the integrand of $I_{1,1}$ we write

$$
\begin{aligned}
&\left| A\big(0, \delta P(0), D^m u\big) - A\big(0, \delta P(0), D^m P\big) - \partial_p A\big(0, \delta P(0), D^m P\big) \cdot D^m (u - P) \right| \\
&\le \int_0^1 \left| \big(\partial_p A\big(0, \delta P(0), D^m P + s D^m (u - P)\big) - \partial_p A\big(0, \delta P(0), D^m P\big) \big) \cdot D^m (u - P) \right| \mathrm{d}s \\
&\le 2L\, \kappa_{M+1}\, \omega_{M+1} \big(|D^m (u - P)|^2 \big) \, |D^m (u - P)|,
\end{aligned}
$$

where we have used (2.4), the fact that $|\delta P(0)| + |D^m P + s(D^m u - D^m P)| \le M + 1$ on S_1 and $|\delta P(0)| + |D^m P| \le M + 1$. Thus, using Hölder's inequality and Jensen's inequality (note that ω_{M+1}^2 is concave), we get

$$
\begin{aligned}
|I_{1,1}| &\le 2L\, \kappa_{M+1} \left(⨍_{Q_\rho} \omega_{M+1}^2 \big(|D^m (u - P)|^2 \big) \,\mathrm{d}z \right)^{\frac{1}{2}} \left(⨍_{Q_\rho} |D^m (u - P)|^2 \,\mathrm{d}z \right)^{\frac{1}{2}} \\
&\le 2L\, \kappa_{M+1}\, \omega_{M+1}(\phi_u)\, \sqrt{\phi_u}.
\end{aligned}
$$

We estimate the integrand of $I_{1,2}$ by the use of (1.3) and (1.5), noting again that $|\delta P(0)| + |D^m P| \le M$, as well as $|D^m u - D^m P| > 1$ on S_2

$$\begin{aligned}\Big|A\left(0,\delta P(0),D^m u\right) &- A\left(0,\delta P(0),D^m P\right) - \partial_p A\left(0,\delta P(0),D^m P\right)\cdot D^m(u-P)\Big|\\ &\le L\,(1+|D^m u|) + L\,(1+|D^m P|) + 2L\,\kappa_{M+1}\,|D^m(u-P)|\\ &\le 2L\;(1+\kappa_{M+1}+M)\,|D^m(u-P)|^2,\end{aligned}$$

which directly implies

$$|I_{1,2}| \le 2L\;(1+\kappa_{M+1}+M)\;\phi_u.$$

Estimate for I_2**.** We use the modulus of continuity (1.6) of the mapping $\xi \mapsto \frac{A(z,\xi,p)}{1+|p|}$ and (2.2) and note that $|\delta P(x)| \le M$ on B_ρ by assumption. Subsequently we apply Young's inequality and the Interpolation-Lemma 2.4 to estimate the integral involving the intermediate derivatives $|\delta(u-P)/\rho|^2$ in terms of $|D^m(u-P)|^2$ and $|(u-P)/\rho^m|^2$ and observe that

$$\begin{aligned}|I_2| &\le L\,K(2M+1)\,\rho^\beta \fint_{Q_\rho}\left|\frac{\delta(u-P)}{\rho}\right|^\beta\left(1+M+|D^m(u-P)|\right)\mathrm{d}z\\ &\le 2L\,H(M)\fint_{Q_\rho}|D^m(u-P)|^2+\left|\frac{\delta(u-P)}{\rho}\right|^2+\rho^\beta\,\mathrm{d}z\\ &\le c(n,m)\;L\,H(M)\left(\phi_u+\psi_u+\rho^\beta\right).\end{aligned}$$

Estimate for I_3**.** We once again use the modulus of continuity (1.6) of the mapping $(z,\xi) \mapsto \frac{A(z,\xi,p)}{1+|p|}$ and (2.2), the assumption on P, Lemma 2.1 and Young's inequality

$$\begin{aligned}|I_3| &\le L\,K(2|\delta P(0)|+1)\fint_{Q_\rho}(2\rho+|\delta P-\delta P(0)|)^\beta\left(1+|D^m u|\right)\mathrm{d}z\\ &\le 2^\beta\rho^\beta\,L\,H(M)\fint_{Q_\rho}\left(1+M+|D^m(u-P)|\right)\mathrm{d}z \le 4L\,H(M)^2\;\left(\rho^\beta+\phi_u\right).\end{aligned}$$

Estimate for I_4**.** From (1.4), the fact that $|D^m P| \le M$, $|\delta\varphi| \le c\rho$ (note that $|D^i\varphi| \le c\rho^{m-i}$ for $0 \le i \le m-1$ since $|D^m\varphi| \le 1$) and Young's inequality we infer that

$$\begin{aligned}I_4 &\le L\fint_{Q_\rho}\left(1+|D^m u|\right)|\delta\varphi|\,\mathrm{d}z \le c\;\rho\fint_{Q_\rho}\left(1+M+|D^m(u-P)|\right)\mathrm{d}z\\ &\le c(n,m,L,M)\;(\phi_u+\rho).\end{aligned}$$

Combining the estimates for $I_1 - I_4$ we finally conclude the desired inequality. □

Remark 3.5 If A does not depend on δu, then the weaker assumption $|D^m P| \le M$ for P is sufficient, because then II does not appear in the above proof and in the estimates of I and III we do not need any restriction on $|\delta P(x_0)|$.

Given a function $g \in L^2(-T,0;W^{m,2}(\Omega;\mathbb{R}^N))$ and $Q_\sigma(z_0) \Subset \Omega_T$ with $0 < \sigma \le 1$. Then $P_{g;z_0,\sigma} : \mathbb{R}^n \to \mathbb{R}^N$ denotes the mean value polynomial of degree $\le m$, which is uniquely defined by $\fint_{Q_\sigma(z_0)} D^k g \,\mathrm{d}z = \fint_{B_\sigma(x_0)} D^k P_{g;z_0,\sigma}\,\mathrm{d}x$ for all $0 \le k \le m$ and we write

$$\psi_g(z_0,\sigma) \equiv \psi_g(z_0,\sigma,P_{g;z_0,\sigma}) \equiv \frac{1}{\sigma^{2m}} \fint_{Q_\sigma(z_0)} |g - P_{g;z_0,\sigma}|^2\,\mathrm{d}z,$$

when using this polynomial.

In order to get suitable excess-estimates for the weak solution u we will exploit good estimates for solutions of linear systems, which are stated in the next lemma, see [7, Chap. 5], or [5, Lemma 4.5].

Lemma 3.6 *Let $\Omega_T \subset \mathbb{R}^{n+1}$ and $g \in L^2(-T,0;W^{m,2}(\Omega;\mathbb{R}^N))$ be a weak solution of the linear parabolic system*

$$\int_{\Omega_T} \left(g\cdot\varphi_t - A(D^m g, D^m\varphi)\right)\mathrm{d}z = 0, \qquad \textit{for all } \varphi \in C_0^\infty(\Omega_T;\mathbb{R}^N),$$

where $A\colon \mathbb{R}^{\mathscr{N}} \times \mathbb{R}^{\mathscr{N}} \to \mathbb{R}$ is a bilinear form fulfilling

$$A(p,p) \ge \nu\,|p|^2, \qquad A(p,\widetilde{p}) \le \Lambda\,|p||\widetilde{p}|,$$

for all $p,\widetilde{p} \in \mathbb{R}^{\mathscr{N}}$. Then g is smooth in Ω_T and for $Q_\rho(z_0) \Subset \Omega_T$ there holds the following estimate:

$$\psi_g(z_0,\theta\rho) \le c_{pa}(n,N,m,\Lambda/\nu)\,\theta^2\,\psi_g(z_0,\rho) \qquad \textit{for all}\ \ \theta \in \left(0,\tfrac{1}{2}\right). \tag{3.8}$$

3.4 Characterization of regular points

In this Chapter we will establish a first characterization of regular points, i.e. we show that a regular point cannot lie in the set $\Sigma_0 \cup \Sigma_2$, where Σ_0 and Σ_2 are defined in the statement of the next theorem. However this characterization does not directly imply that the singular set has $\mathscr{L}^{n+1}$-measure zero. The Lebesgue theory then in fact ensures that $|\Sigma_2| = 0$, but for Σ_0 we cannot directly conclude this property - in contrary to the elliptic case. The problem in the parabolic case is that we cannot apply Poincaré's inequality, since u is a priori only an L^2-function with respect to t. Therefore we will prove a Poincaré type inequality valid for weak solutions in Chap. 3.5.

Theorem 3.7 *Suppose that $u \in L^2(-T,0;W^{m,2}(\Omega;\mathbb{R}^N))$ is a weak solution of* (1.1) *under the assumptions* (1.2)– (1.6). *Then*

$$D^m u \in C^{\beta,\frac{\beta}{2m}}(\Omega_T \setminus \Sigma;\mathbb{R}^{\mathscr{N}}).$$

Moreover $\Sigma \subset \Sigma_0 \cup \Sigma_2$, with

$$\Sigma_0 \equiv \left\{ z_0 \in \Omega_T : \liminf_{\rho\searrow 0} \rho^{-2m} \fint_{Q_\rho(z_0)} \left|u - \widehat{P}_{u;z_0,\rho}\right|^2\mathrm{d}z > 0 \right\},$$

$$\Sigma_2 \equiv \left\{ z_0 \in \Omega_T : \limsup_{\rho\searrow 0} \sum_{k=0}^{m} |(D^k u)_{z_0,\rho}| = \infty \right\},$$

where $\widehat{P}_{u;z_0,\rho}\colon \mathbb{R}^n \to \mathbb{R}^N$ is the unique polynomial of degree $\le m$, minimizing $P \mapsto \fint_{Q_\rho(z_0)} |u - P|^2\,\mathrm{d}z$.

Remark 3.8 For simpler systems of the type (1.10) with coefficients $A(z, D^m u)$, which are independent of δu we actually have $\Sigma \subset \Sigma_0 \cup \widetilde{\Sigma}_2$, where Σ_0 is defined above and

$$\widetilde{\Sigma}_2 \equiv \left\{ z_0 \in \Omega_T : \limsup_{\rho \searrow 0} |(D^m u)_{z_0;\rho}| = \infty \right\}.$$

We will prove the theorem in three steps. In the first step we will use the A-polycaloric approximation lemma to show that—under certain smallness assumptions—the excess $\widetilde{\psi}_u$ of u fulfills a suitable growth estimate when we enlarge the cylinder by a constant factor. Afterwards we iterate this excess estimate by showing that the smallness assumptions are also fulfilled on the smaller cylinder (under the condition that they are fulfilled on the larger cylinder). From this we conclude an excess-decay estimate for $D^m u$, which will finally, in the third step—using the integral characterization of Hölder continuous functions from Campanato—provide the Hölder continuity of $D^m u$, in those points z_0, where the smallness assumptions are fulfilled.

Proof of Theorem 3.7 Let $z_0 = (x_0, t_0) \in \Omega_T$ with $0 < \rho \le 1$, such that $Q_\rho(z_0) \Subset \Omega_T$ and let $P : \mathbb{R}^n \to \mathbb{R}^N$ be a polynomial of degree $\le m$. We define

$$\widetilde{\psi}_u(z_0, \rho, P) \equiv \psi_u(z_0, \rho, P) + \rho^{2\beta}.$$

Step 1: Applying the A-polycaloric approximation lemma we show

Lemma 3.9 *Given $M > 0$ and α with $\beta < \alpha < 1$, there exist $\vartheta \in (0, \frac{1}{4})$ and $\delta \in (0, 1]$, depending on $n, N, m, \nu, M, H(M), L\kappa_{M+1}$ and α, such that if*

$$\omega_{M+1}^2 \left(\widetilde{\psi}_u(z_0, \rho, \widehat{P}_{u;z_0,\rho}) \right) + \widetilde{\psi}_u(z_0, \rho, \widehat{P}_{u;z_0,\rho}) \le \tfrac{1}{2}\delta^2$$

and

$$3m\sqrt{m} \left(|\delta \widehat{P}_{u;z_0,\rho}(x_0)| + |D^m \widehat{P}_{u;z_0,\rho}| \right) \le M, \tag{3.9}$$

on $Q_\rho(z_0) \Subset \Omega_T$ with $0 < \rho \le 1$. Then there holds

$$\widetilde{\psi}_u(z_0, \vartheta\rho, \widehat{P}_{u;z_0,\vartheta\rho}) \le \vartheta^{2\alpha}\, \widetilde{\psi}_u(z_0, \rho, \widehat{P}_{u;z_0,\rho}) + c_3\, \rho^{2\beta},$$

with $c_3 \equiv 1+\delta^{-2}$. There we have denoted by $\widehat{P}_{u;z_0,\rho}$ and $\widehat{P}_{u;z_0,\vartheta\rho}$ the polynomials of degree $\le m$ minimizing the mapping $P \mapsto ⨍_{Q_\rho(z_0)} |u-P|^2\, \mathrm{d}z$, respectively $P \mapsto ⨍_{Q_{\vartheta\rho}(z_0)} |u-P|^2\, \mathrm{d}z$.

For simpler systems of the type (1.10)*, it is enough to require the weaker assumption $|D^m \widehat{P}_{u;z_0,\rho}| \le M$, instead of* (3.9).

Proof Without loss of generality we assume that $z_0 = (x_0, t_0) = 0$ and we abbreviate $\widehat{P}_\rho \equiv \widehat{P}_{u;0,\rho}$. We set $\psi_u \equiv \psi_u \left(0, \rho/2, \widehat{P}_\rho\right)$ and $\widetilde{\psi}_u \equiv \widetilde{\psi}_u \left(0, \rho, \widehat{P}_\rho\right)$ and observe the following monotonicity property

$$\psi_u = \psi_u(0, \rho/2, \widehat{P}_\rho) \le 2^{n+4}\, \psi_u(0, \rho, \widehat{P}_\rho) \le 2^{n+4}\, \widetilde{\psi}_u. \tag{3.10}$$

Note that from Lemma 2.1, the fact that $\rho \le 1$ and the assumption (3.9) on the polynomial $\widehat{P}_\rho$ we find for $x \in B_\rho$ that $|\delta \widehat{P}_\rho(x)| \le |\delta \widehat{P}_\rho(x) - \delta \widehat{P}_\rho(0)| + |\delta \widehat{P}_\rho(0)| \le 2m\sqrt{m}\, \rho\, (|\delta \widehat{P}_\rho(0)| + |D^m \widehat{P}_\rho|) + |\delta \widehat{P}_\rho(0)| \le 3m\sqrt{m} \left(|\delta \widehat{P}_\rho(0)| + |D^m \widehat{P}_\rho|\right) \le M$. This allows us to apply the Caccioppoli inequality, i.e. Lemma 3.2 with $P = \widehat{P}_\rho$ to conclude that

$$\phi_u \equiv \phi_u(0, \rho/2, \widehat{P}_\rho) \le c_{Cac} \left(⨍_{Q_\rho} \frac{|u - \widehat{P}_\rho|^2}{\rho^{2m}}\, \mathrm{d}z + \rho^{2\beta} \right) = c_{Cac}\, \widetilde{\psi}_u. \tag{3.11}$$

Applying Lemma 3.4 to the function $v \equiv u - \widehat{P}_\rho$ on the cylinder $Q_{\rho/2}$ (which again is permitted since $|\delta \widehat{P}_\rho(x)| \le M$ for $x \in B_\rho$), using (3.10), (3.11) and the fact that $\omega_{M+1}(cs) \le c\,\omega_{M+1}(s)$ for $c \ge 1$ (since $s \mapsto \omega_{M+1}(s)$ is concave and $\omega_{M+1}(0) = 0$) we obtain

$$\left| \int_{Q_{\rho/2}} \left(v \cdot \varphi_t - \frac{\partial A}{\partial p} \left(0, \delta \widehat{P}_\rho(0), D^m \widehat{P}_\rho\right) D^m v \cdot D^m \varphi \right) \mathrm{d}z \right| \le c_1 \left(\omega_{M+1}\left(\widetilde{\psi}_u\right) \sqrt{\widetilde{\psi}_u} + \widetilde{\psi}_u + \rho^\beta \right) \sup_{Q_{\rho/2}} |D^m \varphi|,$$

for all $\varphi \in C_0^\infty(Q_{\rho/2}; \mathbb{R}^N)$, where the constant is of the form $c_1 = L\,c(n, m, L/\nu, M, H(M), \kappa_{M+1})$.

Now we define the bilinear form

$$\mathscr{A}(p, \widetilde{p}\,) \equiv \frac{\partial A}{\partial p} \left(0, \delta \widehat{P}_\rho(0), D^m \widehat{P}_\rho\right) (p, \widetilde{p}), \qquad \text{with } p, \widetilde{p} \in \mathbb{R}^{\mathscr{N}},$$

and from the conditions (1.2) and (1.5) and the fact that $|\delta \widehat{P}_\rho(0)| + |D^m \widehat{P}_\rho| \le M$ by assumption we find that $\mathscr{A}(p, \widetilde{p}) \le L\,\kappa_{M+1}\,|p||\widetilde{p}|$, $\mathscr{A}(p, \widetilde{p}) \ge \nu\,|p|^2$ for all $p, \widetilde{p} \in \mathbb{R}^{\mathscr{N}}$, i.e. $\mathscr{A}$ fulfills the conditions of Lemma 3.1 about A-polycaloric approximation with ν and $\Lambda = L\,\kappa_{M+1}$.

For some given $\varepsilon > 0$, which will be chosen later, we determine $\delta = \delta(n, N, m, \nu, L\,\kappa_{M+1}, \varepsilon) \in (0, 1]$, to be the constant from Lemma 3.1. Furthermore we define

$$w \equiv \gamma^{-1} v = \gamma^{-1}(u - \widehat{P}_\rho), \quad \text{where } \gamma \equiv 4c_1 \sqrt{\psi_u(0, \rho, \widehat{P}_\rho) + \delta^{-2}\rho^{2\beta}}.$$

Then, for the function w we have

$$\left| \int_{Q_{\rho/2}} \left(w \cdot \varphi_t - \mathscr{A}(D^m w, D^m \varphi) \right) \mathrm{d}z \right| \le \frac{1}{4} \left(\omega_{M+1}\left(\widetilde{\psi}_u\right) + \sqrt{\widetilde{\psi}_u} + \delta \right) \sup_{Q_{\rho/2}} |D^m \varphi|$$

$$\le \left(\omega_{M+1}^2\left(\widetilde{\psi}_u\right) + \widetilde{\psi}_u + \frac{1}{2}\delta^2 \right)^{\frac{1}{2}} \sup_{Q_{\rho/2}} |D^m \varphi| \le \delta \sup_{Q_{\rho/2}} |D^m \varphi|$$

for all $\varphi \in C_0^\infty(Q_{\rho/2}; \mathbb{R}^N)$ provided we assume that $\omega_{M+1}^2\left(\widetilde{\psi}_u\right) + \widetilde{\psi}_u \le \frac{1}{2}\delta^2$. Further applying the Interpolation-Lemma 2.4 and the Caccioppoli inequality, i.e. Lemma 3.2 we infer that

$$\sum_{k=0}^{m} \fint_{Q_{\rho/2}} \left| \frac{D^k w}{(\rho/2)^{m-k}} \right|^2 \mathrm{d}z \le \fint_{Q_{\rho/2}} 2|D^m w|^2 + c \left| \frac{w}{\rho^m} \right|^2 \mathrm{d}z$$

$$\le (2c_{Cac} + c)\, \frac{\psi_u + \rho^{2\beta}}{\gamma^2} \le \frac{2c_{Cac} + c}{16 c_1^2} \le 1,$$

for $c_1 \gg 1$ big enough, depending on $n, m, \nu, L, M, H(M)$ and κ_{M+1}.

Therefore, we can apply the Lemma about A-polycaloric approximation (i.e. Lemma 3.1) to the function w on the cylinder $Q_{\rho/2}$, provided we assume that $\omega_{M+1}^2\left(\widetilde{\psi}_u\right) + \widetilde{\psi}_u \le \frac{1}{2}\delta^2$ and condition (3.9) is fulfilled. The application of Lemma 3.1 then ensures the existence of a function $g \in L^2(-(\rho/2)^{2m}, 0; W^{m,2}(B_{\rho/2}; \mathbb{R}^N))$, which is $\mathscr{A}$-polycaloric in $Q_{\rho/2}$ and has the following properties

$$\sum_{k=0}^{m} \fint_{Q_{\rho/2}} \left|\frac{D^k g}{(\rho/2)^{m-k}}\right|^2 \mathrm{d}z \le 1 \quad \text{and} \quad \fint_{Q_{\rho/2}} \left|\frac{w-g}{(\rho/2)^m}\right|^2 \mathrm{d}z \le \varepsilon. \tag{3.12}$$

Next we use the excess-estimate (3.8) for the A-polycaloric function g in order to get an estimate for the excess $\widetilde{\psi}_u$ of u. We abbreviate $P_{\theta\rho/2} \equiv P_{g;0,\theta\rho/2}\colon \mathbb{R}^n \to \mathbb{R}^N$ for the mean value polynomial of degree $\le m$ with the property that $(D^k P_{\theta\rho/2})_{0;\theta\rho/2} = (D^k g)_{0;\theta\rho/2}$ for $k = 0, \ldots, m$. First, using Lemma 2.2 and (3.12) we get an estimate for the L^2-norm of $P_{\rho/2}$

$$\begin{aligned}\fint_{Q_{\rho/2}} |P_{\rho/2}|^2 \,\mathrm{d}z &\le c \sum_{k=0}^{m} \left(\frac{\rho}{2}\right)^{2k} \fint_{Q_{\rho/2}} |(D^k P_{\rho/2})_{0;\rho/2}|^2 \,\mathrm{d}z \\ &\le c \sum_{k=0}^{m} \left(\frac{\rho}{2}\right)^{2k} \fint_{Q_{\rho/2}} |D^k g|^2 \,\mathrm{d}z \le c \left(\frac{\rho}{2}\right)^{2m}.\end{aligned}$$

Together with the excess-estimate (3.8) for g and the first inequality in (3.12), we get for all $0 < \theta < \frac{1}{2}$

$$\left(\frac{\theta\rho}{2}\right)^{-2m} \fint_{Q_{\theta\rho/2}} \left|g - P_{\theta\rho/2}\right|^2 \mathrm{d}z \le c_{pa}\, \theta^2 \left(\frac{\rho}{2}\right)^{-2m} \fint_{Q_{\rho/2}} \left|g - P_{\rho/2}\right|^2 \mathrm{d}z \le c_g\, \theta^2,$$

where $c_g = c_g(n, N, m, L\,\kappa_{M+1}/\nu)$. From (3.12) we conclude the following excess estimate for w

$$\begin{aligned}&\left(\frac{\theta\rho}{2}\right)^{-2m} \fint_{Q_{\theta\rho/2}} |w - P_{\theta\rho/2}|^2 \,\mathrm{d}z \\ &\quad\le 2\left(\frac{\theta\rho}{2}\right)^{-2m} \left(\theta^{-n-4m} \fint_{Q_{\rho/2}} |w-g|^2 \,\mathrm{d}z + \fint_{Q_{\theta\rho/2}} |g - P_{\theta\rho/2}|^2 \,\mathrm{d}z\right) \\ &\quad\le 2c_g \left(\theta^{-n-4m}\, \varepsilon + \theta^2\right).\end{aligned}$$

Rescaling to u via $w = \gamma^{-1}(u - \widehat{P}_\rho)$ this implies

$$\left(\frac{\theta\rho}{2}\right)^{-2m} \fint_{Q_{\theta\rho/2}} \left|u - \widehat{P}_\rho - \gamma\, P_{\theta\rho/2}\right|^2 \mathrm{d}z \le 2c_g\, \gamma^2 \left(\theta^{-n-4m}\varepsilon + \theta^2\right).$$

This estimate then also holds for $\widehat{P}_\rho + \gamma\, P_{\theta\rho/2}$ replaced by $\widehat{P}_{\theta\rho/2}$, where $\widehat{P}_{\theta\rho/2}$ denotes the polynomial of degree $\le m$, minimizing the mapping $P \mapsto \fint_{Q_{\theta\rho/2}} |u - P|^2 \,\mathrm{d}z$. Using the definition of γ we infer

$$\left(\frac{\theta\rho}{2}\right)^{-2m} \fint_{Q_{\theta\rho/2}} |u - \widehat{P}_{\theta\rho/2}|^2 \,\mathrm{d}z \le c_2 \left(\theta^{-n-4m}\varepsilon + \theta^2\right) \left(\psi_u(0, \rho, \widehat{P}_\rho) + \delta^{-2}\rho^{2\beta}\right),$$

where c_2 is of the form $L\, c(n, N, m, L\kappa_{M+1}/\nu, M, H(M))$. Now we choose $\varepsilon = \theta^{n+4m+2}$ to obtain

$$\psi_u\left(0, \theta\rho/2, \widehat{P}_{\theta\rho/2}\right) \le 2c_2\, \theta^2 \left(\psi_u(0, \rho, \widehat{P}_\rho) + \delta^{-2}\rho^{2\beta}\right). \tag{3.13}$$

For α with $\beta < \alpha < 1$ we choose $0 < \theta < \frac{1}{2}$ depending on $n, N, m, \nu, L, M, H(M), \kappa_{M+1}$ and α, such that $2^{1+2\alpha} c_2 \theta^2 \le \theta^{2\alpha}$. This also fixes $\varepsilon = \varepsilon(n, N, m, \nu, L, M, H(M), \kappa_{M+1}, \alpha)$ and $\delta = \delta(n, N, m, \nu, L, M, H(M), \kappa_{M+1}, \alpha) \in (0, 1]$. Putting $\vartheta \equiv \frac{\theta}{2}$ in (3.13), we conclude the first assertion of the lemma.

The second assertion, stating that for simpler systems of the type (1.10), the condition $|D^m \widehat{P}_{u;z_0,\rho}| \le M$ instead of (3.9) is sufficient, can be concluded from Remarks 3.3 and 3.5 after Lemma 3.2 and Lemma 3.4, as condition (3.9) is only needed in those two points of the proof. □

Step 2: We iterate Lemma 3.9 to obtain

Lemma 3.10 *Given $M > 1$ and α with $\beta < \alpha < 1$, there exist constants $\vartheta \in (0, \frac{1}{4})$, $\widetilde{\psi}_0$, ρ_0 and c_4, depending on n, N, m, ν, L, M, $H(M)$, κ_{M+1} and α, such that for all parabolic cylinders $Q_\rho(z_0) \Subset \Omega_T$ the conditions*

$$\begin{aligned} &(i) && 3m\sqrt{m}\, \left(|\delta \widehat{P}_\rho(x_0)| + |D^m \widehat{P}_\rho|\right) \le M, \\ &(ii) && \rho \le \rho_0, \\ &(iii) && \widetilde{\psi}_u(\rho) \le \widetilde{\psi}_0 \end{aligned}$$

imply

$$\begin{aligned} &(I)_j && \widetilde{\psi}_u(\vartheta^j \rho) \le \vartheta^{2\alpha j}\, \widetilde{\psi}_u(\rho) + c_4\, (\vartheta^j \rho)^{2\beta}, \\ &(II)_j && 3m\sqrt{m}\, \left(|\delta \widehat{P}_{\vartheta^j\rho}(x_0)| + |D^m \widehat{P}_{\vartheta^j\rho}|\right) \le 2M, \end{aligned}$$

for all $j \in \mathbb{N}$, where for given $0 < r \le 1$ we denote by $\widehat{P}_r \equiv \widehat{P}_{u;z_0,r}$ the polynomial of degree $\le m$, minimizing the mapping $P \mapsto ⨍_{Q_r(z_0)} |u - P|^2\, dz$ and $\widetilde{\psi}_u(r) \equiv \widetilde{\psi}_u(z_0, r, \widehat{P}_r)$. Furthermore the limit

$$\Gamma_{z_0} \equiv \lim_{j\to\infty} (D^m u)_{z_0,\vartheta^j\rho}$$

exists and for $0 < r \le \frac{\rho}{2}$ there holds the estimate

$$⨍_{Q_r(z_0)} |D^m u - \Gamma_{z_0}|^2\, dz \le c \left(\left(\frac{r}{\rho}\right)^{2\alpha} \psi_u(\rho) + r^{2\beta} \right),$$

where c depends on $n, N, m, \nu, L, M, H(M), \kappa_{M+1}, \alpha$ and β.

Remark 3.11 For simpler systems of the type (1.10) it is enough to require the weaker assumption $|D^m \widehat{P}_\rho| \le M$ instead of (i) and $|D^m \widehat{P}_{\vartheta^j\rho}| \le 2M$ instead of $(\mathrm{II})_j$.

Since the iteration procedure from the previous lemma is quite standard we only sketch the proof here, see [5, Lemma 4.9]. The lemma can be shown by induction. Initially, with the help of Lemma 3.9, the assertion is shown in the case $j = 1$. Subsequently, for fixed $j \in \mathbb{N}$ we suppose that $(I)_k$ and $(II)_k$ hold $k = 1, \ldots, j-1$. This enables us to apply Lemma 3.9, and hence deduce $(I)_j$ and $(II)_j$. Thus, having shown the first part of the lemma, we then consider the sequence $((D^m u)_{\vartheta^j\rho/2})$ and show that it is a Cauchy-sequence with limit Γ_{z_0}.

Step 3: Here we use the integral characterization of Hölder continuous functions due to Campanato and Da Prato to show the Hölder continuity of $D^m u$ on $\Omega_T \setminus \Sigma$.

Given $Q_\rho(z_0) \subset \Omega_T$, then $\widehat{P}_{u;z_0,\rho} \colon \mathbb{R}^n \to \mathbb{R}^N$ denotes the polynomial of degree $\le m$, which minimizes the mapping $P \mapsto ⨍_{Q_\rho(z_0)} |u(x,t) - P(x)|^2\, dx\, dt$. Let $z_0 = (x_0, t_0) \in$

$\Omega_T \backslash (\Sigma_0 \cup \Sigma_2)$. Then we can find $M_0 > 0$ and $0 < \rho \le \rho_0(M_0)$ with $Q_{2\rho}(z_0) \Subset \Omega_T$ such that $3m\sqrt{m}\,\big(|\delta \widehat{P}_{u;z,\rho}(x_0)| + |D^m \widehat{P}_{u;z,\rho}|\big) < M_0$ and $\widetilde{\psi}_u(z_0, \rho, \widehat{P}_{u;z_0,\rho}) < \widetilde{\psi}_0(M_0)$ holds, where $\rho_0(M_0)$ and $\widetilde{\psi}_0(M_0)$ are the constants from Lemma 3.10. As the mappings $z \mapsto (D^m u)_{z,\rho}$ and $z \mapsto \widetilde{\psi}_u(z, \rho, \widehat{P}_{u;z,\rho})$ are continuous, there is $0 < R \le \frac{1}{2}\rho$, such that $3m\sqrt{m}\,\big(|\delta \widehat{P}_{u;z,\rho}(x)| + |D^m \widehat{P}_{u;z_0,\rho}|\big) < M_0$ and $\widetilde{\psi}_u(z, \rho, \widehat{P}_{u;z,\rho}) < \widetilde{\psi}_0(M_0)$ for all $z = (x, t) \in Q_R(z_0)$. Moreover $Q_\rho(z) \subset Q_{2\rho}(z_0) \subset \Omega_T$. Therefore we may apply Lemma 3.10 for all $z \in Q_R(z_0)$ and infer that the limit $\Gamma_z = \lim_{j\to\infty} (D^m u)_{z;\vartheta^j\rho}$ exists, with $0 < \vartheta = \vartheta(2M_0) < \frac{1}{4}$ and for $0 < r \le \frac{\rho}{2}$ there holds

$$\fint_{Q_r(z)} |D^m u - \Gamma_z|^2 \,\mathrm{d}z' \le c\left(\left(\frac{r}{\rho}\right)^{2\alpha} \psi_u(z, \rho) + r^{2\beta}\right),$$

where c depends on $n, N, m, \nu, L, M_0, H(M_0), \kappa_{M_0+1}, \alpha$ and β. This implies (see e.g. [12]) that the Lebesgue representative $z \mapsto \Gamma_z$ of $D^m u$ is Hölder continuous on $Q_R(z_0)$ with respect to the parabolic metric with Hölder exponent β, which completes the proof of Theorem 3.7. □

Until now we have shown a first characterization of the singular set, which—as mentioned above—does not directly imply that the singular set is a set of Lebesgue measure zero. To conclude this assertion we need a Poincaré type inequality, which is proved in the next chapter.

3.5 Poincaré type inequality for weak solutions

Due to the fact that weak solutions are a priori only L^2 functions with respect to the time variable t, we cannot apply Poincaré's inequality. To get nevertheless some sort of Poincaré inequality for weak solutions, we consider the space- and time direction separately. We know that our weak solution is weakly differentiable with respect to the space variable x. This allows us to apply Poincaré's inequality in x-direction. For the time-direction we will exploit the fact that the weighted means $(u)_{\tilde\eta}(t)$—defined below—of a weak solution u to system (1.1) are absolutely continuous with respect to the time variable t.

Let $v\colon Q_\rho(z_0) \to \mathbb{R}^N$ be integrable on $Q_\rho(z_0) \Subset \Omega_T$. Moreover, let $\tilde\eta \in C_0^\infty(B_\rho(x_0))$ be a non-negative weight-function, with $\int_{B_\rho} \tilde\eta \,\mathrm{d}x = 1$. For a.e. $t \in (t_0 - \rho^{2m}, t_0)$ we define the weighted mean of v by

$$(v)_{\tilde\eta}(t) \equiv \int_{B_\rho(x_0)} v(\cdot, t)\, \tilde\eta \,\mathrm{d}x. \tag{3.14}$$

Lemma 3.12 *Given $M > 0$. Let $u \in L^2(-T, 0; W^{m,2}(\Omega; \mathbb{R}^N))$ be a weak solution of (1.1) in Ω_T under the assumptions (1.3)–(1.6). Let $Q_\rho(z_0) \Subset \Omega_T$ with $0 < \rho \le 1$ and $P_{u;z_0,\rho}\colon \mathbb{R}^n \to \mathbb{R}^N$ denotes the mean value polynomial of u of degree $\le m$, defined by $(D^k P_{u;z_0,\rho})_{x_0,\rho} = (D^k u)_{x_0,\rho}$ for $k = 0, \ldots, m$. Then, under the condition that $|\delta P_{u;z_0,\rho}(x)| + |D^m P_{u;z_0,\rho}| \le M$ for all $x \in B_\rho(x_0)$ there exists $c = c(n, N, m, L, M, K(2M + 1), \kappa_{M+1})$, such that*

$$\fint_{Q_\rho(z_0)} |u - P_{u;z_0,\rho}|^2 \,\mathrm{d}z \le c\, \rho^{2m}\left(\fint_{Q_\rho(z_0)} \left|D^m u - (D^m u)_{z_0,\rho}\right|^2 \mathrm{d}z + \rho^{2\beta}\right).$$

Remark 3.13 For simpler systems of the type (1.10) the conclusion also holds if we replace our condition on $P_{u;z_0,\rho}$ by the weaker condition $|D^m P_{u;z_0,\rho}| \le M$. This is due to the fact that (1.6) then has the simpler form $|A(z, p) - A(z_0, p)| \le L\, \mathrm{d}_{\mathscr{P}}(z, z_0)^\beta (1 + |p|)$.

Proof Without loss of generality we assume that $z_0 = 0$ and we abbreviate $P \equiv P_{u;z_0,\rho}$. We choose a weight-function $\tilde\eta \in C_0^\infty(B_\rho)$, such that $\int_{B_\rho} \tilde\eta \,\mathrm{d}x = 1$ and $\|D^\ell \tilde\eta\|_{L^2(B_\rho)} \le c(n,m)\rho^{-(\frac{n}{2}+m)}$ for $0 \le \ell \le 2m$. For any $i \in \{1\dots,N\}$ we take $\varphi\colon \mathbb{R}^{n+1} \to \mathbb{R}^N$ with $\varphi_i = \tilde\eta$ and $\varphi_j = 0$ for $j \ne i$ as test-function in the Steklov-formulation (2.7) of the system and obtain for the weighted means of $[u_i]_h$, defined in (3.14) (note that $[u]_h = ([u_1]_h,\dots,[u_N]_h)$) for a.e. $s,t \in (-\rho^{2m},0)$,

$$
\begin{aligned}
([u_i]_h)_{\tilde\eta}(t) - ([u_i]_h)_{\tilde\eta}(s) &= \int_s^t \frac{\partial([u_i]_h)_{\tilde\eta}}{\partial\tau}\,\mathrm{d}\tau \\
&= \int_s^t\int_{B_\rho} \big([A_i(\cdot,\delta u, D^m u)]_h \cdot D^m\tilde\eta + [B_i(\cdot,\delta u, D^m u)]_h \cdot \delta\tilde\eta\big)\,\mathrm{d}x\,\mathrm{d}\tau. \qquad (3.15)
\end{aligned}
$$

Letting $h \searrow 0$, enlarging the domain of integration if necessary and noting that $A(z_0,\delta P(x_0), D^m P)$ is constant, we find

$$
\begin{aligned}
|(u_i)_{\tilde\eta}(t) - (u_i)_{\tilde\eta}(s)| \le& \int_{Q_\rho} \big|A(\cdot,\delta u, D^m u) - A(\cdot,\delta u, D^m P)\big|\,|D^m\tilde\eta|\,\mathrm{d}z, \\
&+ \int_{Q_\rho} \big|A(\cdot,\delta u, D^m P) - A(\cdot,\delta P, D^m P)\big|\,|D^m\tilde\eta|\,\mathrm{d}z \\
&+ \int_{Q_\rho} |A(\cdot,\delta P, D^m P) - A(z_0,\delta P(x_0), D^m P)||D^m\tilde\eta|\,\mathrm{d}z \\
&+ \int_{Q_\rho} |B(\cdot,\delta u, D^m u)|\,|\delta\tilde\eta|\,\mathrm{d}z \\
=& I_1 + I_2 + I_3 + I_4,
\end{aligned}
$$

with the obvious meaning of $I_1 - I_4$.

Estimate for I_1. We can treat this term similar to the integral I_1 in the proof of the Caccioppoli inequality, i.e. Lemma 3.2. Proceeding this way (with $D^m\tilde\eta$ instead of LOT) we end up with the following analogue of estimate (3.7):

$$
I_1 \le L\,c(M, H(M), \kappa_{M+1})\,\|D^m\tilde\eta\|_{L^2(Q_\rho)} \left(\int_{Q_\rho} \big(|D^m(u-P)|^2 + |\delta u - \delta P|^{2\beta}\big)\,\mathrm{d}z\right)^{\frac12}.
$$

Estimate for I_2. Similarly to the estimate of I_2 in the proof of the Caccioppoli inequality (with $D^m\tilde\eta$ instead of $D^m\varphi$) we obtain with the help of (1.6) and Hölder's inequality that

$$
I_2 \le L \int_{Q_\rho} |\delta u - \delta P|^\beta |D^m\tilde\eta|\,\mathrm{d}z \le L\,\|D^m\tilde\eta\|_{L^2(Q_\rho)} \left(\int_{Q_\rho} |\delta u - \delta P|^{2\beta}\,\mathrm{d}z\right)^{\frac12}.
$$

Estimate for I_3. We can once again proceed similar to the estimate of I_3 in the proof of the Caccioppoli inequality (with $D^m\tilde\eta$ instead of $D^m\varphi$) to find that

$$I_3 \le c\, L\, (1+M)^\beta \rho^\beta \int_{Q_\rho} |D^m \tilde\eta|\, \mathrm{d}z \le c(M, H(M))\, L\, \|D^m \tilde\eta\|_{L^2(Q_\rho)}\, \rho^\beta |Q_\rho|.$$

Estimate for I_4**.** From the growth assumption (1.4) on B, the fact that $|D^m P| \le M$ and Hölder's inequality we infer

$$I_4 \le L\, \|\delta\tilde\eta\|_{L^2(Q_\rho)} \Bigg(\int_{Q_\rho} \big(1 + M + |D^m(u-P)|\big)^2\, \mathrm{d}z \Bigg)^{\frac{1}{2}}.$$

Combining the estimates for $I_1 - I_4$ and summing over $i = 1, \ldots, N$ we arrive at:

$$|(u)_{\tilde\eta}(t) - (u)_{\tilde\eta}(s)| \le c\, \|D^m \tilde\eta\|_{L^2(Q_\rho)} \Bigg(\int_{Q_\rho} \big(|D^m(u-P)|^2 + |\delta u - \delta P|^{2\beta} + \rho^{2\beta}\big)\, \mathrm{d}z \Bigg)^{\frac{1}{2}}$$
$$+ c\, \|\delta\tilde\eta\|_{L^2(Q_\rho)} \Bigg(\int_{Q_\rho} \big(|D^m(u-P)|^2 + 1\big)\, \mathrm{d}z \Bigg)^{\frac{1}{2}},$$

where $c = c(N, L, M, \kappa_{M+1})$.

We get an analogous estimate for the weighted means of $D^k u$, $k = 1, \ldots, m$, by taking $D^\alpha \tilde\eta$ instead of $\tilde\eta$ as test-function in (2.7), where α is a multiindex with $|\alpha| = k$. Then, by integration by parts (k-times) we find that $(D^\alpha u)_{\tilde\eta}(t) = (-1)^k (u)_{D^\alpha \tilde\eta}(t)$. Replacing $\tilde\eta$ with $D^\alpha \tilde\eta$ in the previous estimate and noting that $\|D^\ell \tilde\eta\|_{L^2(Q_\rho)} = \rho^m \|D^\ell \tilde\eta\|_{L^2(B_\rho)} \le c(n) \rho^m \rho^{-(\frac{n}{2}+\ell)}$ for $0 \le \ell \le 2m$ we obtain

$$|(D^k u)_{\tilde\eta}(t) - (D^k u)_{\tilde\eta}(s)| \le c\, \rho^{m-k} \Bigg(\fint_{Q_\rho} |D^m(u-P)|^2 + |\delta(u-P)|^{2\beta}\, \mathrm{d}z + \rho^{2\beta} \Bigg)^{\frac{1}{2}}, \tag{3.16}$$

where $c = c(n, m, N, L, M, \kappa_{M+1})$.

Moreover, starting once again from (3.15), but using only the bounds (1.3) and (1.4) to estimate the right-hand side, we find that

$$|(D^k u)_{\tilde\eta}(t) - (D^k u)_{\tilde\eta}(s)|^2 \le c(n, m, N, L)\, \rho^{2(m-k)} \fint_{Q_\rho} \big(1 + |D^m u|^2\big)\, \mathrm{d}z. \tag{3.17}$$

Our next goal is to obtain a suitable estimate for the term involving $\delta u - \delta P$ in (3.16). For $0 \le j \le m-1$ we consider the following decomposition

$$\begin{aligned} \fint_{Q_\rho} |D^j(u-P)|^2\, \mathrm{d}z &\le 3 \Bigg[\fint_{-\rho^{2m}}^{0} \fint_{B_\rho} \big|D^j(u-P) - \big(D^j(u-P)\big)_{\tilde\eta}\big|^2\, \mathrm{d}x\, \mathrm{d}t \\ &\quad + \fint_{-\rho^{2m}}^{0} \Bigg| \fint_{-\rho^{2m}}^{0} \big((D^j u)_{\tilde\eta}(t) - (D^j u)_{\tilde\eta}(s)\big)\, \mathrm{d}s \Bigg|^2 \mathrm{d}t + \Bigg| \fint_{-\rho^{2m}}^{0} (D^j u)_{\tilde\eta}(s)\, \mathrm{d}s - (D^j P)_{\tilde\eta} \Bigg|^2 \Bigg] \\ &= 3(J_1 + J_2 + J_3), \end{aligned} \tag{3.18}$$

with the obvious meaning of J_1 - J_3.

Estimate for J_1**.** Applying the Poincaré inequality "slicewise" to $D^j(u-P)(\cdot,t)$ we infer

$$J_1 \le c(n)\,\rho^2 \fint_{-\rho^{2m}}^{0} \fint_{B_\rho} |D^{j+1}(u-P)|^2 \,\mathrm{d}x\,\mathrm{d}t.$$

Estimate for J_3**.** Here we exploit the fact that $\fint_{Q_\rho} D^j(u-P)\,\mathrm{d}z = 0$ to bound J_3 in terms of J_1, yielding that

$$J_3 \le J_1 \le c(n)\,\rho^2 \fint_{-\rho^{2m}}^{0} \fint_{B_\rho} |D^{j+1}(u-P)|^2 \,\mathrm{d}x\,\mathrm{d}t.$$

Estimate for J_2**.** To estimate the second integral we use (3.17) with $k = j$ to find that

$$J_2 \le c(n,m,N,L)\,\rho^{2(m-j)} \fint_{Q_\rho} \left(1+|D^m u|^2\right) \mathrm{d}z.$$

Joining the estimates for $J_1 - J_3$ with (3.18) yields for $0 \le j \le m-1$ that

$$\fint_{Q_\rho} |D^j(u-P)|^2 \,\mathrm{d}z \le c\,\rho^2 \fint_{Q_\rho} |D^{j+1}(u-P)|^2 \,\mathrm{d}z + c\,\rho^{2(m-j)} \fint_{Q_\rho} (1+|D^m u|^2)\,\mathrm{d}z,$$

where $c = c(n,m,N,L)$. Starting with $j = k$ and then iterating this estimate for $j = k+1,\dots,m-1$, we obtain for $0 \le k \le m-1$

$$\fint_{Q_\rho} |D^k(u-P)|^2 \,\mathrm{d}z \le c\rho^{2(m-k)} \left[\fint_{Q_\rho} |D^m(u-P)|^2 \,\mathrm{d}z + \fint_{Q_\rho} (1+|D^m u|^2)\,\mathrm{d}z \right], \tag{3.19}$$

where $c = c(n,m,N,L)$. Summing over $k = 0,\dots,m-1$ and noting that $D^m P = (D^m u)_\rho$, $|D^m P| \le M$ by assumption and $\rho^{2(m-k)} \le \rho^2$ since $\rho \le 1$ we find

$$\fint_{Q_\rho} |\delta u - \delta P|^2 \,\mathrm{d}z \le c(n,m,N,L)\,\rho^2 \left[\fint_{Q_\rho} |D^m u - (D^m u)_\rho|^2 \,\mathrm{d}z + 1 \right].$$

Inserting this estimate in (3.16) (after an application of Hölder's inequality) and noting once again that $D^m P = (D^m u)_\rho$, we arrive at

$$|(D^k u)_{\tilde\eta}(t) - (D^k u)_{\tilde\eta}(s)| \le c\,\rho^{m-k} \left(\rho^{2\beta} + \fint_{Q_\rho} \left|D^m u - (D^m u)_\rho\right|^2 \mathrm{d}z \right)^{\frac{1}{2}}, \tag{3.20}$$

where c depends on $n, m, N, L, M, K(2M+1)$ and κ_{M+1}.

Now we are in a position to consider the integral we initially wanted to estimate. For $0 \le k \le m-1$ we once again use the decomposition from (3.18) and estimate the integrals J_1 and J_3 as we did before. But now, for the term J_2 we can use the better estimate (3.20) instead of (3.16) and we finally come up with the desired Poincaré inequality. □

Remark 3.14 From the proof of the previous lemma, in particular from (3.17), we infer the following estimate in time for the weighted means of $D^k u$: Suppose that $Q_\rho(z_0) \Subset \Omega_T$ and $\tilde\eta \in C_0^\infty(B_\rho(x_0))$ is a non-negative weight-function with $\int_{B_\rho(x_0)} \tilde\eta\,\mathrm{d}x = 1$ and

$\|D^\ell\widetilde{\eta}\|_{L^2(B_\rho(x_0))} \le c\rho^{-(\frac{n}{2}+\ell)}$ for $0 \le \ell \le 2m$. Then for $0 \le k \le m$ and for a.e. $s, t \in (t_0 - \rho^{2m}, t_0)$, the difference in time of the weighted mean of $D^k u$, defined in (3.14) can be estimated by

$$|(D^k u)_{\tilde{\eta}}(t) - (D^k u)_{\tilde{\eta}}(s)|^2 \le c(n, m, N, L)\, \rho^{2(m-k)} \fint_{Q_\rho(z_0)} \left(1 + |D^m u|^2\right) \mathrm{d}z.$$

3.6 Proof of the partial regularity result

The regularity result of Theorem 1.1 finally follows from the next theorem, since we can conclude from the Lebesgue differentiation theorem that $|\Sigma_1| = 0$ (with Σ_1 defined below).

Theorem 3.15 *Suppose that $u \in L^2(-T, 0; W^{m,2}(\Omega; \mathbb{R}^N))$ is a weak solution of (1.1) under the assumptions (1.2)– (1.6). Then $\Sigma \subset \Sigma_1 \cup \Sigma_2$, where Σ_2 is as in Theorem 3.7 and*

$$\Sigma_1 \equiv \left\{ z_0 \in \Omega_T : \liminf_{\rho \searrow 0} \fint_{Q_\rho(z_0)} |D^m u - (D^m u)_{z_0,\rho}|^2 \,\mathrm{d}z > 0 \right\}.$$

For simpler systems of the type (1.10) there even holds $\Sigma \subset \Sigma_1 \cup \widetilde{\Sigma}_2$, where $\widetilde{\Sigma}_2$ is as in Remark 3.8.

Proof We will show that $\Omega_T \backslash (\Sigma_1 \cup \Sigma_2) \subset \Omega_T \backslash (\Sigma_0 \cup \Sigma_2)$, where Σ_0 is as in Theorem 3.7. Then the assertion can be concluded from Theorem 3.7.

Let $z_0 \in \Omega_T \backslash (\Sigma_1 \cup \Sigma_2)$. Then $\limsup_{\rho \searrow 0}(|(\delta u)_{z_0,\rho}| + |(D^m u)_{z_0,\rho}|) \equiv M < \infty$, and therefore we find some ρ_0 such that $|(\delta u)_{z_0,\rho}| + |(D^m u)_{z_0,\rho}| \le M + 1$ for all $\rho \le \rho_0$. Hence, with Lemma 2.2, we find for the mean value polynomial $P_{u;z_0,\rho} : \mathbb{R}^n \to \mathbb{R}^N$ of degree $\le m$ of u on $Q_\rho(z_0)$ that $|\delta P_{u;z_0,\rho}(x)| + |D^m P_{u;z_0,\rho}| \le c\,(|(\delta P_{u;z_0,\rho})_{z_0,\rho}| + |(D^m P_{u;z_0,\rho})_{z_0,\rho}|) = c\,(|(\delta u)_{z_0,\rho}| + |(D^m u)_{z_0,\rho}|) \le c(n, m)(M + 1) \equiv c_M$. Therefore we may apply the Poincaré inequality, i.e. Lemma 3.12) for all $\rho \le \rho_0$ and conclude that

$$\rho^{-2m} \fint_{Q_\rho(z_0)} |u - P_{u;z_0,\rho}|^2 \,\mathrm{d}z \le c \left(\fint_{Q_\rho(z_0)} \left|D^m u - (D^m u)_{z_0,\rho}\right|^2 \,\mathrm{d}z + \rho^{2\beta} \right),$$

where $c = c(n, m, N, L, M, K(2c_M + 1), \kappa_{c_M})$. Since $z_0 \notin \Sigma_1$ the first term on the right-hand side vanishes as $\rho \searrow 0$. But this estimate certainly also holds for the minimizing polynomial $\widehat{P}_{u;z_0,\rho} : \mathbb{R}^n \to \mathbb{R}^N$ and therefore it implies $\rho^{-2m} \fint_{Q_\rho(z_0)} |u - \widehat{P}_{u;z_0,\rho}|^2 \,\mathrm{d}z \to 0$ as $\rho \searrow 0$. Hence, we have $z_0 \in \Omega_T \backslash (\Sigma_0 \cup \Sigma_2)$ and by Theorem 3.7, z_0 is a regular point.

For the second assertion about simpler systems we recall Remark 3.13, which allows us to replace the condition on $P_{u;z_0,\rho}$ by the weaker condition $|(D^m u)_{z_0,\rho}| = |D^m P_{u;z_0,\rho}| \le M+1$ in the proof above. Therefore $\Omega_T \backslash (\Sigma_1 \cup \widetilde{\Sigma}_2) \subset \Omega_T \backslash (\Sigma_0 \cup \widetilde{\Sigma}_2)$ and the assertion can be concluded by Remark 3.8. □

4 Higher integrability

Up to now, while proving partial regularity of $D^m u$, there was no higher integrability of $D^m u$ needed. This tool will become essential when we want to show better estimates for the Hausdorff-dimension of the singular set in Chap. 6. So here is just the right place to analyze higher integrability properties. For that purpose we put our focus on parabolic systems of the type

$$\int_{\Omega_T} \big(u \cdot \varphi_t - A(z, D^m u) \cdot D^m \varphi\big)\, dz = \int_{\Omega_T} \sum_{i=0}^{m} B^i(z, D^m u) \cdot D^i \varphi \, dz \tag{4.1}$$

for all $\varphi \in C_0^\infty(\Omega_T; \mathbb{R}^N)$, where $A : \Omega_T \times \mathbb{R}^{\mathcal{N}} \to \mathrm{Hom}(\mathbb{R}^{\mathcal{N}}, \mathbb{R})$ and $B^i : \Omega_T \times \mathbb{R}^{\mathcal{N}} \to \mathrm{Hom}(\mathbb{R}^{\mathcal{M}_i}, \mathbb{R})$ for $i = 0, \dots, m$ with

$$|B^i(z, p)| \le L\,|p| + b_i, \quad \text{for } i = 1, \dots, m-1 \qquad \text{and} \quad |B^m(z, p)| \le b_m, \tag{4.2}$$

for all $z \in \Omega_T$ and $p \in \mathbb{R}^{\mathcal{N}}$, where $b_i : \Omega_T \to \mathbb{R}_+$, $i = 1, \dots, m$. Contrary to the rest of the paper, here, we do not indicate the dependence of the coefficients on the intermediate derivatives δu, since there is no hypothesis put on the modulus of continuity with respect to δu, in the sense of (1.8). Moreover, we consider a sligtly more general right-hand side. This enables us to present the natural scaling with respect to the radius, coming from the fact that the inhommogeneity counts as a derivative. The main result of this Chapter is the following

Theorem 4.1 *Suppose that $u \in L^2(-T, 0; W^{m,2}(\Omega; \mathbb{R}^N))$ is a weak solution to the parabolic system (4.1) under the assumptions (1.2), (1.3), (4.2) and $b_i \in L^{2(1+\sigma_0)}(\Omega_T)$ with $\sigma_0 > 0$ for $i = 0, \dots, m$. Then there exists $\sigma = \sigma(n, m, L/\nu)$ with $0 < \sigma < \sigma_0$, such that $|D^m u| \in L^{2(1+\sigma)}_{loc}(\Omega_T)$ and for $Q_{2\rho}(z_0) \Subset \Omega_T$ there holds*

$$\fint_{Q_{\rho/2}(z_0)} |D^m u|^{2(1+\sigma)}\, dz \le c\bigg(\fint_{Q_{2\rho}(z_0)} |D^m u|^2\, dz\bigg)^{1+\sigma} + c + c \fint_{Q_{2\rho}(z_0)} \sum_{i=0}^{m} (\rho^{m-i}\, b_i)^{2(1+\sigma)}\, dz,$$

where $c = c(n, m, N, L/\nu)$ and $\lim_{L/\nu \to \infty} \sigma = 0$.

For second-order parabolic systems higher integrability was shown by Giaquinta and Struwe [25]. Note that higher integrabiliy results and Calderón-Zygmund estimates for more general p-Laplacean type systems were achieved in [2,28]. As usual, when showing such a higher integrability result, the main ingredients are a Caccioppoli and a Poincaré inequality. In the next lemma we state the Caccioppoli inequality, which is suitable for our purpose. Contrary to Lemma 3.2, it deals with polynomials of degree $\le m-1$. The proof is similar—but simpler — since now we can directly use the growth assumption (1.3) to estimate the terms involving A on the right-hand side, instead of the Hölder continuity assumption (1.6). For a detailed proof (in the case of more general parabolic p-Laplacean type systems) we refer to [5, Lemma 7.1].

Lemma 4.2 *Suppose that $u \in L^2(-T, 0; W^{m,2}(\Omega; \mathbb{R}^N))$ is a weak solution of the parabolic system (1.1) in Ω_T under the assumptions (1.2)–(1.6) and let $Q_r(z_0) \Subset \Omega_T$, with $0 < r \le 1$. Then for all polynomials $P : \mathbb{R}^n \to \mathbb{R}^N$ of degree $\le m-1$ there holds*

$$\sup_{t \in (t_0 - (r/2)^{2m}, t_0)} \fint_{B_{r/2}(x_0)} \frac{|u(\cdot, t) - P|^2}{r^{2m}}\, dx + \fint_{Q_{r/2}(z_0)} |D^m u|^2\, dz \le c_{Cac} \fint_{Q_r(z_0)} \frac{|u - P|^2}{r^{2m}} + \sum_{i=0}^{m} b_i^2 + 1\, dz,$$

where c_{Cac} depends on n, m and L/ν and $c_{Cac} \to \infty$ as $L/\nu \to \infty$.

The next lemma is a sort of Poincaré inequality. It can be proved similar to Lemma 3.12 [see in particular (3.19)], i.e. by applying the Poincaré inequality in x-direction and using estimates for differences in time of weighted means of u, respectively $D^k u$, $0 \le k \le m-1$. Therefore we only state the lemma without proof and refer the reader to [5, Lemma 6.4], for a detailed proof (in the case of more general parabolic p-Laplacean type systems).

Lemma 4.3 *Suppose that* $u \in L^2(-T, 0; W^{m,2}(\Omega; \mathbb{R}^N))$ *is a weak solution of the parabolic system* (1.1) *in* Ω_T *under the assumptions* (1.3)–(1.4) *and let* $Q_r(z_0) \Subset \Omega_T$, *with* $0 < r \le 1$. *Then for* $k = 0, \dots, m-1$ *and for all* $\gamma \in [1, 2]$ *there holds*

$$\fint_{Q_r(z_0)} |D^k(u - P_Q)|^\gamma \, dz \le c(n, m, N, L, \gamma)\, \rho^{\gamma(m-k)} \fint_{Q_r(z_0)} \left(|D^m u|^\gamma + \sum_{i=0}^m b_i^\gamma + 1\right) dz,$$

where $P_Q : \mathbb{R}^n \to \mathbb{R}^N$ *is the mean value polynomial of* u *of degree* $\le m-1$ *defined by* $(\delta P_Q)_{x_0,r} = (\delta u)_{z_0,r}$.

Now, we have collected all the preliminaries in order to prove our higher integrability result:

Proof of Theorem 4.1 Without loss of generality we can assume that $z_0 = 0$ and $\rho = 1$. Otherwise we consider the function $v(x, t) = \rho^{-m} u(x_0 + \rho x, t_0 + \rho^{2m} t)$, which is a solution of the parabolic system $v_t + (-1)^m \operatorname{div}^m \tilde{A}(z, D^m v) = \sum_{i=0}^m (-1)^i \operatorname{div}^i \tilde{B}^i(z, D^m v)$, with $\tilde{A}(x, t, p) = A(x_0 + \rho x, t_0 + \rho^{2m} t, p)$ and $\tilde{B}^i(x, t, p) = \rho^{m-i} B^i(x_0 + \rho x, t_0 + \rho^{2m} t, p)$ on Q_1 and infer the general result by rescaling to $Q_\rho(z_0)$.

Let $Q_{2r}(\tilde{z}) \Subset Q_1$ be a parabolic cylinder. We define $\gamma = \max\{\frac{2n}{n+2m}, 1\}$ and apply Gagliardo-Nirenberg's inequality [see e.g. [5,29,32], Theorem B.6]:

$$\begin{aligned}
\fint_{Q_r(\tilde{z})} |u - P|^2 \, dz &\le c\, r^{m\gamma} \fint_{\tilde{t}-r^{2m}}^{\tilde{t}} \sum_{k=0}^m \fint_{B_r(\tilde{x})} \frac{|D^k(u - P)|^\gamma}{r^{\gamma(m-k)}} \, dx \left(\fint_{B_r(\tilde{x})} |u - P| \, dx \right)^{1-\frac{\gamma}{2}} dt \\
&\le c(n, m)\, r^{m\gamma} \sum_{k=0}^m \fint_{Q_r(\tilde{z})} \frac{|D^k(u - P)|^\gamma}{r^{\gamma(m-k)}} \, dz \left(\sup_{t \in (\tilde{t}-r^{2m}, \tilde{t})} \fint_{B_r(\tilde{x})} |u - P| \, dx \right)^{1-\frac{\gamma}{2}} \\
&= I_1 \cdot (I_2)^{1-\frac{\gamma}{2}}
\end{aligned}$$

with the obvious meaning of I_1 and I_2.

Estimate of I_1**.** Applying the Poincaré inequality from Lemma 4.3 for each $k = 0, \dots, m$, we infer

$$I_1 \le c(n, m, N, L) \fint_{Q_r(\tilde{z})} \left(|D^m u|^\gamma + \sum_{i=0}^m b_i^\gamma + 1\right) dz.$$

Estimate of I_2**:** In turn, we apply the Caccioppoli inequality from Lemma 4.2 and the Poincaré's inequality from Lemma 4.3 to find that

$$\begin{aligned}
I_2 &\le c\, r^{2m} \fint_{Q_{2r}(\tilde{z})} \left(\frac{|u - P|^2}{r^{2m}} + \sum_{i=0}^m b_i^2 + 1 \right) dz \\
&\le c(n, m, N, L/\nu)\, r^{2m} \fint_{Q_{2r}(\tilde{z})} \left(|D^m u|^2 + \sum_{i=0}^m b_i^2 + 1 \right) dz.
\end{aligned}$$

Inserting the estimates for I_1 and I_2 above, dividing by r^{2m} and using Young's inequality, we obtain for $\varepsilon > 0$ that

$$\fint_{Q_r(\tilde{z})} \frac{|u-P|^2}{r^{2m}}\,\mathrm{d}z \le \varepsilon \fint_{Q_{2r}(\tilde{z})} |D^m u|^2\,\mathrm{d}z + c\left(\fint_{Q_r(\tilde{z})} |D^m u|^\gamma\,\mathrm{d}z\right)^{\frac{2}{\gamma}} + c\fint_{Q_{2r}(\tilde{z})} \left(\sum_{i=0}^m b_i^2 + 1\right)\mathrm{d}z,$$

where $c = c(1/\varepsilon, n, m, N, L/\nu)$. We estimate the L^2-norm of $D^m u$, using in turn the Caccioppoli inequality from Lemma 4.2 and the above estimate with the choice $\varepsilon = \frac{1}{2C_{Cac}}$

$$\begin{aligned}\fint_{Q_{r/2}(\tilde{z})} |D^m u|^2\,\mathrm{d}z &\le c_{Cac} \fint_{Q_r(\tilde{z})} \left(\frac{|u-P|^2}{r^{2m}} + \sum_{i=0}^m b_i^2 + 1\right)\mathrm{d}z \\ &\le \frac{1}{2}\fint_{Q_{2r}(\tilde{z})} |D^m u|^2\,\mathrm{d}z + c\left(\fint_{Q_r(\tilde{z})} |D^m u|^\gamma\,\mathrm{d}z\right)^{\frac{2}{\gamma}} + c\fint_{Q_{2r}(\tilde{z})} \left(\sum_{i=0}^m b_i^2 + 1\right)\mathrm{d}z,\end{aligned}$$

where $c = c(n, m, N, L/\nu)$. This is a Reverse-Hölder inequality, valid for any parabolic cylinder $Q_r(\tilde{z})$ with $Q_{2r}(\tilde{z}) \Subset Q_1$. Therefore we can apply Gehring's Theorem (see e.g. [23, Chap. V, Proposition 1.1]), which ensures the existence of $\sigma = \sigma(n, m, N, L/\nu)$, such that $|D^m u| \in L^{2(1+\sigma)}_{loc}(Q_1)$ and moreover, the asserted estimate holds.

Because the maximal size of σ depends on c and c increases with L/ν, we observe that $\sigma \searrow 0$ as $L/\nu \to \infty$. This behavior of σ is described e.g. in [6] and [36], providing an explicit specification of σ. □

5 Parabolic fractional Sobolev spaces

5.1 Preliminaries

In the case of differentiable or Lipschitz continuous coefficients one usually applies the difference quotient method in order to show that $D^m u$ is weakly differentiable with respect to x and t. Then bounds for the dimension of the singular set can be obtained with the help of Giusti's Lemma. Since the considered coefficients A are not assumed to be differentiable or Lipschitz continuous with respect to (z, ξ), we cannot apply the difference quotient method here. Instead we use fractional difference quotients, and hence we will show that $D^m u$ lies in a certain fractional Sobolev space. In this secton the definitions and basic properties about fractional Sobolev and Nikolskii spaces are given.

Let $f : \mathbb{R}^{n+1} \to \mathbb{R}^k$, $k \in \mathbb{N}$. We define the finite differences in space direction $\tau^s_h(f)$ by

$$(\tau^s_h f)(x,t) \equiv f(x + he_s, t) - f(x,t),$$

for $|h| > 0$, $x \in \Omega$ with $\operatorname{dist}(x, \partial\Omega) > |h|$ and $1 \le s \le n$ where $\{e_s\}_{1\le s\le n}$ is the standard basis of $\mathbb{R}^n$. Similarly the finite differences in time direction $\tau_h(f)$ are defined by

$$(\tau_h f)(x,t) \equiv f(x, t+h) - f(x,t),$$

for $|h| > 0$ and $t \in (-T + |h|, -|h|)$. Then, we have the following estimate: For $f \in L^p(\Omega \times (t_1, t_2 + h))$, where $\Omega \subset \mathbb{R}^n$, $t_1 < t_2$ and $h > 0$ there holds

$$\int_{t_1}^{t_2} \int_{\Omega} |\tau_h f|^p \, dx \, dt \le 2 \int_{t_1}^{t_2+h} \int_{\Omega} |f|^p \, dx \, dt, \tag{5.1}$$

and also the analogue for $h < 0$.

The proof of our dimension reduction result will be based on showing that $D^m u$ admits better fractional differentiability properties. We now define the class of fractional Sobolev-spaces, which are suitable for our purpose. Suppose that $v \in L^p(\Omega_T; \mathbb{R}^k)$ with $1 \le p < \infty$, $k \in \mathbb{N}$. Then, v belongs to the parabolic fractional Sobolev space $W^{\alpha,\gamma;\,p}(\Omega_T; \mathbb{R}^k)$ with $\alpha, \gamma \in (0, 1)$, if

$$\begin{aligned}[v]_{\alpha,\gamma;\,p;\,\Omega_T} \equiv & \int_{-T}^{0} \int_{\Omega} \int_{\Omega} \frac{|v(x,t) - v(y,t)|^p}{|x-y|^{n+\alpha p}} \, dx \, dy \, dt \\ & + \int_{\Omega} \int_{-T}^{0} \int_{-T}^{0} \frac{|v(x,t) - v(x,\tau)|^p}{|t-\tau|^{1+\gamma p}} \, dt \, d\tau \, dx < \infty. \end{aligned}$$

The local variant, i.e. the space $W^{\alpha,\gamma;\,p}_{loc}(\Omega_T; \mathbb{R}^k)$ is defined as usual. This means that $v \in W^{\alpha,\gamma;\,p}_{\mathrm{loc}}(\Omega_T; \mathbb{R}^k)$, if $v \in W^{\alpha,\gamma;\,p}(\widetilde{Q}; \mathbb{R}^k)$ for all $\widetilde{Q} \Subset \Omega_T$.

The well known relation between fractional Sobolev-spaces and Nicolskii-spaces can be adapted to the parabolic framework (see [5, Lemma 10.9]).

Lemma 5.1 *Let $v \in L^p(\Omega_T; \mathbb{R}^k)$, $1 \le p < \infty$, $k \in \mathbb{N}$. Moreover, let $B_R(x_0) \Subset \Omega$ and $(t_1, t_2) \Subset (-T, 0)$.*

(i) *Suppose that there exist $A_1, c_1 > 0$, such that for all $0 < |h| \le \min\{T + t_1, |t_2|, A_1\}$ there holds*

$$\int_{B_R(x_0)} \int_{t_1}^{t_2} |\tau_h v|^p \, dt \, dx \le c_1 \, |h|^{p\gamma} \qquad \textit{for some} \quad \gamma \in (0, 1).$$

Then, there exists $\widetilde{c}_1 = \widetilde{c}_1(p, \gamma, \widetilde{\gamma}, A_1, c_1, T + t_1, |t_2|, t_2 - t_1, \|v\|_{L^p(B_R(x_0)\times(t_1,t_2))})$, such that

$$\int_{B_R(x_0)} \int_{t_1}^{t_2} \int_{t_1}^{t_2} \frac{|v(x,t) - v(x,\tau)|^p}{|t-\tau|^{1+\widetilde{\gamma} p}} \, d\tau \, dt \, dx \le \widetilde{c}_1 \qquad \textit{for all} \ \ \widetilde{\gamma} \in (0, \gamma).$$

(ii) *Suppose that $B_{3R}(x_0) \Subset \Omega$ and that there exist $A_2 \ge 1$ and $c_2 \ge 0$, such that for all $0 < |h| \le \frac{R}{A_2}$ and $s \in \{1, \dots, n\}$ there holds*

$$\int_{t_1}^{t_2} \int_{B_{2R}(x_0)} |\tau_h^s v|^p \, dx \, dt \le c_2 \, |h|^{p\alpha} \qquad \textit{for some} \quad \alpha \in (0, 1).$$

Then, there exists $\widetilde{c}_2 = \widetilde{c}_2(n, p, R, \alpha, \widetilde{\alpha}, A_2, c_2, , \|v\|_{L^p(B_R(x_0)\times(t_1,t_2))})$*, such that*

$$\int_{t_1}^{t_2} \int_{B_R(x_0)} \int_{B_R(x_0)} \frac{|v(x,t) - v(y,t)|^p}{|x-y|^{n+\widetilde{\alpha}p}} \, dx \, dy \, dt \le \widetilde{c}_2 \quad \textit{for all} \quad \widetilde{\alpha} \in (0, \alpha).$$

The next lemma is a Poincaré type inequality for functions lying in a fractional Sobolev space (see [5, Lemma 10.10]).

Lemma 5.2 *Let* $Q_\rho(z_0) \subset \mathbb{R}^{n+1}$ *be a parabolic cylinder. Suppose that* $v \in W^{\gamma, \frac{\gamma}{2m}; p}(Q_\rho(z_0); \mathbb{R}^k)$*, with* $\gamma \in (0, 1)$*,* $1 \le p < \infty$ *and* $k \in \mathbb{N}$*. Then there holds:*

$$\int_{Q_\rho(z_0)} |v - (v)_{z_0,\rho}|^p \, dz \le c(n, p) \, \rho^{\gamma p} \, [v]^p_{\gamma, \frac{\gamma}{2m}; p; Q_\rho(z_0)}.$$

In Chap. 6.1 we will derive the required estimates only for second differences of $D^k u$, $0 \le k \le m-1$ (i.e. for $\tau_h(\tau_h D^k u)$). With the help of the following lemma we can conclude from bounds for second differences to similar bounds for first differences. This reasoning was introduced by Domokos in [14] for the treatment of sub-elliptic equations in the Heisenberg-group (see also [22,5, Lemma 10.11]).

Lemma 5.3 *Let* $f \in L^2(-T, 0)$*,* $(t_1, t_2) \Subset (-T, 0)$*,* $0 < \alpha < 1$ *and* $0 < h_0 < \frac{1}{2}\min\{|t_2|, T + t_1\}$*. Suppose that there exists* $M > 0$ *such that*

$$\int_{t_1}^{t_2} |\tau_h(\tau_h f)|^2 \, dt \le M^2 \, |h|^{2\alpha} \qquad \forall\, 0 < |h| < h_0.$$

Then

$$\int_{t_1}^{t_2} |\tau_h f|^2 \, dt \le c(\alpha, h_0) \, \left(M^2 + \|f\|_{L^2(-T,0)}\right) \, |h|^{2\alpha} \qquad \forall\, 0 < |h| < \frac{h_0}{2}.$$

The final estimate for our dimension reduction result will be concluded with the help of the following lemma. The proof is based on a measure theoretical argument, a parabolic version of the so called Giusti Lemma [5, Lemma 10.13].

Lemma 5.4 *Let* $v \in W^{\gamma, \frac{\gamma}{2m}; p}_{loc}(\Omega_T; \mathbb{R}^k)$ *with* $\gamma \in (0, 1)$*,* $p \ge 1$*,* $k \in \mathbb{N}$*. For the sets*

$$A \equiv \left\{ z_0 \in \Omega_T : \liminf_{\rho \searrow 0} \;\; ⨍_{Q_\rho(z_0)} |v - (v)_{z_0,\rho}|^p \, dz > 0 \right\},$$

$$B \equiv \left\{ z_0 \in \Omega_T : \limsup_{\rho \searrow 0} \; |(v)_{z_0,\rho}| = \infty \right\}$$

there holds

$$\dim_{\mathscr{P}}(A) \le n + 2m - \gamma p \quad \textit{and} \quad \dim_{\mathscr{P}}(B) \le n + 2m - \gamma p.$$

5.2 Interpolation for parabolic fractional Sobolev spaces

In this chapter we want to establish a parabolic version of an interpolation theorem of S. Campanato (see [11, Theorem 2.1], [9, Lemma 2.5]). It ensures better integrability properties of a function by interpolation between parabolic fractional Sobolev spaces and Hölder spaces. The first version of this interpolation theorem can be found in [8, Teorema 3.III], in which the interpolation between Sobolev and Hölder spaces is considered.

In this chapter we use parabolic cubes instead of parabolic cylinders. For $\rho > 0$ and $z_0 = (x_0, t_0) \in \mathbb{R}^{n+1}$ the associated parabolic cube is denoted by $\mathcal{Q}_\rho(z_0) \equiv C_\rho(x_0) \times (t_0 - \rho^{2m}, t_0)$, where $C_\rho(x_0) \equiv \{x \in \mathbb{R}^n : \|x - x_0\|_\infty < \rho\}$ is the euclidean cube of sidelength $2\rho > 0$ and with center $x_0 \in \mathbb{R}^n$.

The next lemma is an interpolation result between the function spaces L^2 and $W^{1+\theta;2}$ in $\mathbb{R}^n$. It will be applied later on the time-slices. For a proof we refer to [11, Appendix, Lemma 1], or [5, Lemma 10.14].

Lemma 5.5 *Let $C_\rho(x_0) \subset \mathbb{R}^n$, $\rho > 0$ be a cube and suppose that $u \in W^{1+\gamma;2}(C_\rho(x_0); \mathbb{R}^k)$, with $\gamma \in (0, 1)$, $k \in \mathbb{N}$. Then we have*

$$\int_{C_\rho(x_0)} |Du|^2 \, dx \le c[Du]_{\gamma;2;C_\rho(x_0)}^{\frac{2}{1+\gamma}} \left(\int_{C_\rho(x_0)} |u|^2 \, dx \right)^{\frac{\gamma}{1+\gamma}} + c\, \rho^{-2} \int_{C_\rho(x_0)} |u|^2 \, dx,$$

where $c = c(n, \gamma)$.

The next lemma is a parabolic version of Lemma 2 in [11, Appendix].

Lemma 5.6 *Let $\mathcal{Q}_\rho \equiv \mathcal{Q}_\rho(z_0) \subset \mathbb{R}^{n+1}$, $\rho > 0$ and suppose that $u \in L^2(t_0 - \rho^{2m}, t_0; W^{1,2}(C_\rho(x_0); \mathbb{R}^k)) \cap W^{\frac{1+\gamma}{2m};2}(t_0 - \rho^{2m}, t_0; L^2(C_\rho(x_0); \mathbb{R}^k))$ and $Du \in W^{\gamma,\frac{\gamma}{2m};2}(\mathcal{Q}_\rho; \mathbb{R}^{nk})$, for some $\gamma \in (0, 1)$, $k \in \mathbb{N}$. Then*

$$\int_{\mathcal{Q}_\rho} |Du - (Du)_{\mathcal{Q}_\rho}|^2 \, dz \le c(n, \gamma) \left([Du]_{\gamma,\frac{\gamma}{2m};2;\mathcal{Q}_\rho} + [u]_{0,\frac{1+\gamma}{2m};2;\mathcal{Q}_\rho} \right)^{\frac{2}{1+\gamma}} \times \left(\int_{\mathcal{Q}_\rho} |u - (u)_{\mathcal{Q}_\rho}|^2 \, dz \right)^{\frac{\gamma}{1+\gamma}}.$$

Proof Without loss of generality we can assume that $z_0 = (x_0, t_0) = 0$. Let $\ell_\rho : \mathbb{R}^n \to \mathbb{R}^k$ be the affine function minimizing $\ell \mapsto ⨍_{\mathcal{Q}_\rho} |u(x,t) - \ell(x)|^2 \, dx\, dt$. To bound the considered integral we firstly exploit the fact that $(Du)_{\mathcal{Q}_\rho}$ minimizes $a \mapsto ⨍_{\mathcal{Q}_\rho} |Du(z) - a|^2 \, dz$, for $a \in \mathbb{R}^{nk}$ and apply Lemma 5.5 "slicewise" to $D(u - \ell_\rho)$ (note also that $D\ell_\rho$ is constant)

$$\begin{aligned} \int_{\mathcal{Q}_\rho} |Du - (Du)_{\mathcal{Q}_\rho}|^2 \, dz &\le \int_{\mathcal{Q}_\rho} |Du - D\ell_\rho|^2 \, dz \\ &\le c \int_{-\rho^{2m}}^{0} [Du(\cdot,t)]_{\gamma;2;C_\rho}^{\frac{2}{1+\gamma}} \left(\int_{C_\rho} |u(\cdot,t) - \ell_\rho|^2 \, dx \right)^{\frac{\gamma}{1+\gamma}} dt + c\, \rho^{-2} \int_{\mathcal{Q}_\rho} |u - \ell_\rho|^2 \, dz \\ &\le c \left([Du]_{\gamma,0;2;\mathcal{Q}_\rho}^{\frac{2}{1+\gamma}} + \rho^{-2} \left(\int_{\mathcal{Q}_\rho} |u - \ell_\rho|^2 \, dz \right)^{\frac{1}{1+\gamma}} \right) \left(\int_{\mathcal{Q}_\rho} |u - \ell_\rho|^2 \, dz \right)^{\frac{\gamma}{1+\gamma}}, \end{aligned} \tag{5.2}$$

where we have used Hölder's inequality with respect to t in the second last line. Note that $c = c(n,\gamma)$. To bound the second integral on the in the right-hand side we exploit the minimality property of ℓ_ρ and infer that

$$\begin{aligned}\int_{\mathcal{Q}_\rho} |u-\ell_\rho|^2\,\mathrm{d}z &\le \int_{\mathcal{Q}_\rho} |u-(u)_{\mathcal{Q}_\rho}-(Du)_{\mathcal{Q}_\rho}\cdot x|^2\,\mathrm{d}z\\ &\le 2\int_{-\rho^{2m}}^{0}\int_{C_\rho}\big|u(x,t)-(u)_{C_\rho}(t)-(Du)_{\mathcal{Q}_\rho}\cdot x\big|^2\,\mathrm{d}x\,\mathrm{d}t+2|C_\rho|\int_{-\rho^{2m}}^{0}|(u)_{C_\rho}(t)-(u)_{\mathcal{Q}_\rho}|^2\,\mathrm{d}t\\ &= 2(I_1+I_2),\end{aligned} \tag{5.3}$$

with the obvious meaning of I_1 and I_2. To estimate I_1 we apply in turn Poincaré's inequality "slicewise" for a.e. $t\in(-\rho^{2m},0)$ and the Poincaré inequality for fractional Sobolev spaces from Lemma 5.2 to infer

$$I_1 \le c\,\rho^2\int_{\mathcal{Q}_\rho}|Du-(Du)_{\mathcal{Q}_\rho}|^2\,\mathrm{d}z \le c(n,\gamma)\,\rho^{2(1+\gamma)}\,[Du]^2_{\gamma,\frac{\gamma}{2m};\,2;\,\mathcal{Q}_\rho},$$

For II we use the fact that $|t-s|^{-1}\ge\rho^{-2m}$ for $t\neq s\in(-\rho^{2m},0)$ to obtain

$$\begin{aligned}I_2 &\le \int_{-\rho^{2m}}^{0}\fint_{-\rho^{2m}}^{0}\int_{C_\rho}|u(\cdot,t)-u(\cdot,\tau)|^2\,\mathrm{d}x\,\mathrm{d}\tau\,\mathrm{d}t\\ &\le \rho^{2(1+\gamma)}\int_{C_\rho}\int_{-\rho^{2m}}^{0}\int_{-\rho^{2m}}^{0}\frac{|u(\cdot,t)-u(\cdot,\tau)|^2}{|t-\tau|^{1+\frac{1+\gamma}{m}}}\,\mathrm{d}\tau\,\mathrm{d}t\,\mathrm{d}x = \rho^{2(1+\gamma)}\,[u]^2_{0,\frac{1+\gamma}{2m};\,2;\,\mathcal{Q}_\rho}.\end{aligned}$$

Combining the previous bounds for I_1 and I_2 with (5.3) we arrive at

$$\int_{\mathcal{Q}_\rho}|u-\ell_\rho|^2\,\mathrm{d}z \le c(n,\gamma)\,\rho^{2(1+\gamma)}\,\Big([Du]_{\gamma,\frac{\gamma}{2m};\,2;\,\mathcal{Q}_\rho}+[u]_{0,\frac{1+\gamma}{2m};\,2;\,\mathcal{Q}_\rho}\Big)^2,$$

Using this estimate to bound the second integral on the right-hand side of (5.2) and exploiting once again the minimality property of ℓ_ρ we obtain the desired estimate. □

Now we are in a position to prove the main result of this chapter.

Theorem 5.7 *Let $\mathcal{Q}_\rho(z_0)\subset\mathbb{R}^{n+1}$ be a parabolic cube and $u\in L^2(t_0-\rho^{2m},t_0;\,W^{1,2}(C_\rho(x_0);\mathbb{R}^k))\cap W^{\frac{1+\gamma}{2m};2}(t_0-\rho^{2m},t_0;\,L^2(C_\rho(x_0);\mathbb{R}^k))\cap C^{\lambda,\frac{\lambda}{2m}}(\mathcal{Q}_\rho(z_0);\mathbb{R}^k)$ and $Du\in W^{\gamma,\frac{\gamma}{2m};2}(\mathcal{Q}_\rho(z_0);\mathbb{R}^{nk})$ with $\gamma\in(0,1)$, $0<\lambda\le 1$, $k\in\mathbb{N}$. Then $|Du|\in L^s(\mathcal{Q}_\rho(z_0))$ for all $1\le s<q=\frac{2(n+2m)(1+\gamma)}{n+2m-2\lambda\gamma}$ with the estimate*

$$\int_{\mathcal{Q}_\rho(z_0)}|Du-(Du)_{z_0,\rho}|^s\,\mathrm{d}z \le c(n,m,\gamma,q,s)\,A^s\,|\mathcal{Q}_\rho(z_0)|^{1-\frac{s}{q}}.$$

Proof We consider a family of parabolic cubes $(\mathcal{Q}_{\rho_j}(z_j))_{j\in\mathbb{N}}$, which are pairwise disjoint and with $\mathcal{Q}_{\rho_j}(z_j)\subset\mathcal{Q}_\rho(z_0)$ for $j\in\mathbb{N}$. We abbreviate $\mathcal{Q}_\rho\equiv\mathcal{Q}_\rho(z_0)$ and $\mathcal{Q}_j\equiv\mathcal{Q}_{\rho_j}(z_j)$

for $j \in \mathbb{N}$. From Hölder's inequality and Lemma 5.6 we infer for each parabolic cube $\mathcal{Q}_j$, $j \in \mathbb{N}$ that

$$\begin{aligned}\int_{\mathcal{Q}_j} |Du - (Du)_{\mathcal{Q}_j}| \, \mathrm{d}z &\le |\mathcal{Q}_j|^{\frac{1}{2}} \Bigg(\int_{\mathcal{Q}_j} |Du - (Du)_{\mathcal{Q}_j}|^2 \, \mathrm{d}z \Bigg)^{\frac{1}{2}} \\ &\le c\, |\mathcal{Q}_j|^{\frac{1}{2}} \Big([Du]_{\gamma, \frac{\gamma}{2m}; 2; \mathcal{Q}_j} + [u]_{0, \frac{1+\gamma}{2m}; 2; \mathcal{Q}_j} \Big)^{\frac{1}{1+\gamma}} \\ &\quad \times \Bigg(\int_{\mathcal{Q}_j} |u - (u)_{\mathcal{Q}_j}|^2 \, \mathrm{d}z \Bigg)^{\frac{\gamma}{2(1+\gamma)}}, \end{aligned} \tag{5.4}$$

where $c = c(n, \gamma)$. Now we want to exploit the Hölder continuity of u to bound the last integral on the right-hand side. Let $z \in \mathcal{Q}_j$. Since $\operatorname{dist}_{\mathcal{P}}(\tilde{z}, z) \le c(n, m)\rho_j$ for $z, \bar{z} \in \mathcal{Q}_j$, we observe that

$$|u(z) - (u)_{\mathcal{Q}_j}| \le \fint_{\mathcal{Q}_j} |u(z) - u(\tilde{z})| \, d\tilde{z} \le (c\,\rho_j)^{\lambda} \fint_{\mathcal{Q}_j} \frac{|u(z) - u(\tilde{z})|}{\operatorname{dist}_{\mathcal{P}}(z, \tilde{z})^{\lambda}} \, d\tilde{z} \le c(n, m)\, \rho_j^{\lambda}\, [u]_{\lambda, \frac{\lambda}{2m}; \mathcal{Q}_\rho},$$

Integrating over $\mathcal{Q}_j$ and noting that $\rho_j = (2^{-n} |\mathcal{Q}_j|)^{1/(n+2m)}$ we obtain

$$\int_{\mathcal{Q}_j} |u - (u)_{\mathcal{Q}_j}|^2 \, \mathrm{d}z \le c\, \rho_j^{2\lambda}\, |\mathcal{Q}_j|\, [u]^2_{\lambda, \frac{\lambda}{2m}; \mathcal{Q}_j} = c(n, m)\, |\mathcal{Q}_j|^{1 + \frac{2\lambda}{n+2m}}\, [u]^2_{\lambda, \frac{\lambda}{2m}; \mathcal{Q}_\rho}.$$

We will use this estimate to bound the last integral on the right-hand side of (5.4). Therefore, the resulting exponent of $|\mathcal{Q}_j|$ is $\frac{1}{2} + \frac{\gamma}{2(1+\gamma)}(1 + \frac{2\lambda}{n+2m}) = 1 - \frac{1}{q}$, where q is defined in the statement of the theorem. Hence, we deduce from (5.4) that

$$\int_{\mathcal{Q}_j} |Du - (Du)_{\mathcal{Q}_j}| \, \mathrm{d}z \le c\, |\mathcal{Q}_j|^{1 - \frac{1}{q}}\, A_j,$$

with

$$A_j \equiv \Big([Du]_{\gamma, \frac{\gamma}{2m}; 2; \mathcal{Q}_j} + [u]_{0, \frac{1+\gamma}{2m}; 2; \mathcal{Q}_j} \Big)^{\frac{1}{1+\gamma}}\, [u]^{\frac{\gamma}{1+\gamma}}_{\lambda, \frac{\lambda}{2m}; \mathcal{Q}_\rho},$$

where $c = c(n, m, \gamma)$. Summing over $j = 1, \ldots, \infty$ we infer

$$\sum_{j=1}^{\infty} |\mathcal{Q}_j|^{1-q} \Bigg(\int_{\mathcal{Q}_j} |Du - (Du)_{\mathcal{Q}_j}| \, \mathrm{d}z \Bigg)^q \le c(n, m, \gamma)^q \sum_{j=1}^{\infty} A_j^q \le c(n, m, \gamma)^q\, A^q,$$

where we have also used the fact that $q/(1+\gamma) > 1$ and

$$A \equiv \Big([u]_{0, \frac{1+\gamma}{2m}; 2; \mathcal{Q}_\rho} + [Du]_{\gamma, \frac{\gamma}{2m}; 2; \mathcal{Q}_\rho} \Big)^{\frac{1}{1+\gamma}}\, [u]^{\frac{\gamma}{1+\gamma}}_{\lambda, \frac{\lambda}{2m}; 2; \mathcal{Q}_\rho}.$$

Taking the supremum over all families $(\mathcal{Q}_j)_{j \in \mathbb{N}}$ of disjoint parabolic cubes with $\mathcal{Q}_j \subset \mathcal{Q}_\rho$ for $j \in \mathbb{N}$ we conclude that

$$K^q(Du) \equiv \sup \left\{ \sum_{j=1}^{\infty} |\mathcal{Q}_j|^{1-q} \Bigg(\int_{\mathcal{Q}_j} |Du - (Du)_{\mathcal{Q}_j}| \, \mathrm{d}z \Bigg)^q : \mathcal{Q}_\rho \subset \bigcup_{j=1}^{\infty} \mathcal{Q}_j \right\} \le c^q\, A^q.$$

An application of John-Nirenberg's Theorem in a parabolic version ([27, Lemma 3], [26, Lemma 2.3], in the elliptic case) then yields for any $\mu > 0$ that

$$\left|\left\{z \in \mathcal{Q}_\rho : |Du(z) - (Du)_{\mathcal{Q}_\rho}| > \mu\right\}\right| \le c(n, m, \gamma, q) \left(\frac{A}{\mu}\right)^q .$$

For $M > 0$ we therefore obtain

$$\begin{aligned}\int_{\mathcal{Q}_\rho} |Du - (Du)_{\mathcal{Q}_\rho}|^s \, dz &= s \int_0^\infty \mu^{s-1} \, |\{z \in \mathcal{Q}_\rho : |Du(z) - (Du)_{\mathcal{Q}_\rho}| > \mu\}| d\mu \\ &= s \int_0^M \ldots d\mu + s \int_M^\infty \ldots d\mu \\ &\le M^s \, |\mathcal{Q}_\rho| + c\, s \int_M^\infty A^q \, \mu^{s-q-1} d\mu = M^s \, |\mathcal{Q}_\rho| + \frac{c\, s}{q-s} A^q \, M^{s-q},\end{aligned}$$

where $c = c(n, m, q, \gamma)$. Choosing $M = A|\mathcal{Q}_\rho|^{-\frac{1}{q}}$ we infer the desired estimate. □

Remark 5.8 At this stage we want to mention that the result of the previous theorem can also be applied on parabolic cylinders. But then we end up with a smaller radius. More precisely, let $Q_\rho(z_0)$ be a parabolic cylinder, such that the assumptions of Theorem 5.7 are fulfilled on $Q_\rho(z_0)$. Then we can conclude that $|Du| \in L^s(Q_{\rho/2}(z_0))$ for all $1 \le s < q$.

6 Dimension reduction

Since the coefficients A of our system are not assumed to be Lipschitz continuous with respect to $(z, \delta u)$ we cannot expect to derive estimates for difference quotients of $D^m u$. The best we can hope for is to controll the fractional difference quotients of $D^m u$ with denominator h^β in x-direction, respectively $h^{\frac{\beta}{2m}}$ in t-direction, where β is the Hölder exponent of the coefficients in (1.8). The method to consider fractional difference quotients was developed by Mingione in [31] and [30] for elliptic systems. In order to derive estimates for the finite differences in time $\tau_h D^m u$ of $D^m u$ we will have to consider finite differences $\tau_h(A)$ of the coefficients A. Thereby we will often use the following decomposition

$$\begin{aligned}&\tau_h \left[A(\cdot, \cdot, \delta u(\cdot, \cdot), D^m u(\cdot, \cdot))\right](x, t) \\ &\quad = A\left(x, t+h, \delta u(x, t+h), D^m u(x, t+h)\right) - A\left(x, t+h, \delta u(x, t+h), D^m u(x, t)\right) \\ &\qquad + A\left(x, t+h, \delta u(x, t+h), D^m u(x, t)\right) - A\left(x, t+h, \delta u(x, t), D^m u(x, t)\right) \\ &\qquad + A\left(x, t+h, \delta u(x, t), D^m u(x, t)\right) - A\left(x, t, \delta u(x, t), D^m u(x, t)\right) \\ &\quad \equiv \mathscr{A}(h) + \mathscr{B}(h) + \mathscr{C}(h).\end{aligned} \tag{6.1}$$

Furthermore, we denote

$$\begin{aligned}\mathscr{A}(h)(x, t) &= \int_0^1 \frac{\partial A}{\partial p}\left(x, t+h, \delta u(x, t+h), D^m u(x, t) + \vartheta \tau_h D^m u(x, t)\right) d\vartheta \\ &\quad \cdot \tau_h(D^m u)(x, t) \\ &\equiv \widetilde{\mathscr{A}}(h) \cdot \tau_h(D^m u)(x, t),\end{aligned} \tag{6.2}$$

with the obvious meaning of $\widetilde{\mathscr{A}}(h)$. The conditions (1.2) and (1.7) imply the following ellipticity and boundedness properties of $\widetilde{\mathscr{A}}(h)$

$$\widetilde{\mathscr{A}}(h)p \cdot p \geq \nu\,|p|^2, \qquad |\widetilde{\mathscr{A}}(h)| \leq L, \tag{6.3}$$

for $p \in \mathbb{R}^{\mathscr{N}}$ and consequently

$$|\mathscr{A}(h)| = |\widetilde{\mathscr{A}}(h)\,\tau_h(D^m u)| \leq L\,|\tau_h(D^m u)|. \tag{6.4}$$

The Hölder continuity of A in (1.8) provides the following bound for $\mathscr{B}(h)$ and $\mathscr{C}(h)$

$$|\mathscr{B}(h)| \leq L\,\widetilde{\theta}\,(|\tau_h(\delta u)|)\left(1+|D^m u|\right), \qquad |\mathscr{C}(h)| \leq L\,|h|^{\frac{\beta}{2m}}\,\left(1+|D^m u|\right). \tag{6.5}$$

Similarly, we can decompose the differences $\tau_h^s(A)$ of A in the space directions e_s, $s = 1, \ldots, n$. Since the proofs of the fractional differentiability of $D^m u$ in Chapts. 6.2 and 6.3 for the space direction are similar—but simpler—then the ones for the time direction, we will not accomplish the details there. We will only sketch the differences and refer to [5] for a detailed proof.

6.1 Estimates for finite differences

We first consider the derivatives $D^k u$, $0 \leq k \leq m-1$ of lower order of a weak solution u of system (1.1). Since $u \in L^2(-T, 0; W^{m,2}(\Omega, \mathbb{R}^N))$ we know that $D^k u$ is weakly differentiable with respect to the space-variable x. Therefore, we can estimate finite differences of $D^k u$, $0 \leq k \leq m-1$ by

$$\int_{t_1}^{t_2}\int_{B_R(x_0)} |\tau_h^s(D^k u)|^2\,\mathrm{d}x\,\mathrm{d}t \leq c(n)|h|^2 \int_{t_1}^{t_2}\int_{B_{R+|h|}(x_0)} |D^{k+1}u|^2\,\mathrm{d}x\,\mathrm{d}t, \tag{6.6}$$

whenever $B_{R+|h|}(x_0) \Subset \Omega$ and $(t_1, t_2) \Subset (-T, 0)$. But $D^k u$ is not necessarily weakly differentiable with respect to the time variable t. To obtain nevertheless a similar estimate, we will exploit the parabolic system. Roughly speaking, each space derivative corresponds to a "$\frac{1}{2m}$th time derivative". We know that u is m times weakly differentiable with respect to the space variable x. This suggests that u is "$\frac{1}{2}$ times differentiable" with respect to t, and $D^k u$ respectively "$\frac{m-k}{2m}$ times". In a certain sense this is the conclusion of the following lemma, where we derive suitable estimates for $|\tau_h u|$.

Lemma 6.1 *Suppose that* $u \in L^2(-T, 0; W^{m,2}(\Omega; \mathbb{R}^N))$ *is a weak solution of system* (1.1) *under the assumptions* (1.3) *and* (1.4) *and let* $(t_1, t_2) \Subset (-T, 0)$, $B_{2r}(x_0) \Subset \Omega$ *and* $0 < r \leq 1$. *Then for all* $0 < |h| < \frac{1}{2}\min\{|t_2|, T+t_1\}$ *the following estimate holds*

$$\int_{t_1}^{t_2}\int_{B_r(x_0)} \frac{|\tau_h u|^2}{|h|}\,\mathrm{d}z \leq c(n, m, L)\left(1+\frac{|h|}{r^{2m}}\right)\int_{t_1}^{t_2}\int_{B_{2r}(x_0)} \left(1+\left[|D^m u|\right]_h + |\tau_h(D^m u)|\right)^2\,\mathrm{d}z.$$

Proof Without loss of generality we can assume that $x_0 = 0$ and we show the assertion for $h > 0$, since the proof in the case $h < 0$ is similar, with $[u]_{\bar{h}}$ instead of $[u]_h$, where $[u]_{\bar{h}}$ and $[u]_h$ denote the Steklov-means of u defined in (2.5). Therefore let $0 < h < \frac{1}{2}\min\{|t_2|, T+t_1\}$.

We start with the Steklov-formulation (2.7) of the system. Due to the fact that $\tau_h u = h\,\partial_t[u]_h$, we can write $\partial_t[u]_h \cdot \varphi = \frac{\tau_h u}{h} \cdot \varphi$ for the integrand on the left-hand side of (2.7). Let $r \leq r_1 < r_2 \leq 2r$. We choose a cut-off $\eta \in C_0^\infty(B_{r_2})$ with $\eta \equiv 1$ on B_{r_1} and

$|D^j\eta| \le c(r_2 - r_1)^{-j}$ for $0 \le j \le m$. Taking the test-function $\varphi = \eta\,\tau_h u$, integrating over (t_1, t_2), using the growth assumptions (1.3) on A and (1.4) on B and Hölder's inequality, we get

$$
\begin{aligned}
&\int_{t_1}^{t_2}\int_{B_{r_2}} \frac{|\tau_h u|^2}{h}\,\eta\,\mathrm{d}z \\
&= \int_{t_1}^{t_2}\int_{B_{r_2}} \partial_t[u]_h \cdot \eta\,\tau_h u\,\mathrm{d}z \\
&= -\int_{t_1}^{t_2}\int_{B_{r_2}} \big[A(\cdot, \delta u, D^m u)\big]_h \cdot D^m(\eta\,\tau_h u) + \big[B(\cdot, \delta u, D^m u)\big]_h \cdot \delta(\eta\,\tau_h u)\,\mathrm{d}z \\
&\le c(m, L)\left(\int_{t_1}^{t_2}\int_{B_{r_2}} \big(1 + \big[|D^m u|\big]_h\big)^2\,\mathrm{d}z\right)^{\frac{1}{2}}\left(\sum_{k=0}^{m}\int_{t_1}^{t_2}\int_{B_{r_2}} |D^k(\eta\,\tau_h u)|^2\,\mathrm{d}z\right)^{\frac{1}{2}}. \qquad (6.7)
\end{aligned}
$$

To bound the second term on the right-hand side we use that $|D^{k-j}\eta| \le c\,(r_2 - r_1)^{-(k-j)}$ for $0 \le j \le k \le m$, $\eta \equiv 1$ on B_{r_1}, apply the Interpolation-lemma 2.4 and note that $D^m(\tau_h u) = \tau_h(D^m u)$ and $r_2 - r_1 \le 1$. Thus, we obtain for $0 \le k \le m$

$$
\begin{aligned}
\int_{t_1}^{t_2}\int_{B_{r_2}} |D^k(\eta\tau_h u)|^2\mathrm{d}z &\le c\sum_{j=0}^{k}\int_{t_1}^{t_2}\int_{\operatorname{spt} D^{k-j}\eta} \frac{|D^j(\tau_h u)|^2}{(r_2 - r_1)^{2(k-j)}}\mathrm{d}z \\
&\le c\int_{t_1}^{t_2}\int_{B_{r_2}} \frac{|\tau_h u|^2}{(r_2 - r_1)^{2m}} + |\tau_h(D^m u)|^2\mathrm{d}z
\end{aligned}
$$

where $c = c(n, m)$. Summing over $k = 0, \ldots, m$ this yields a bound for the second term on the right-hand side of (6.7). Recalling that $\eta \equiv 1$ on B_{r_1} and using $B_{r_2} \subset B_{2r}$ and Young's inequality we therefore obtain from (6.7) that

$$
\begin{aligned}
\int_{t_1}^{t_2}\int_{B_{r_1}} \frac{|\tau_h u|^2}{h}\,\mathrm{d}z &\le \frac{1}{2}\int_{t_1}^{t_2}\int_{B_{r_2}} \frac{|\tau_h u|^2}{h}\,\mathrm{d}z \\
&\quad + c\left(1 + \frac{h}{(r_2 - r_1)^{2m}}\right)\int_{t_1}^{t_2}\int_{B_{2r}} \big(1 + \big[|D^m u|\big]_h + |\tau_h D^m u|\big)^2\,\mathrm{d}z,
\end{aligned}
$$

where $c = c(n, m, L)$. Applying Lemma 2.3 we infer the desired estimate. □

Corollary 6.2 *Suppose that* $u \in L^2(-T, 0; W^{m,2}(\Omega; \mathbb{R}^N))$ *is a weak solution of system* (1.1) *under the assumptions* (1.3) *and* (1.4) *and let* $(t_1, t_2) \Subset (-T, 0)$, $B_{2r}(x_0) \Subset \Omega$ *with* $0 < r \le 1$. *Then for all* $0 < |h| < \frac{1}{2}\min\{|t_2|, T + t_1, r^{2m}\}$ *and all* $0 \le k \le m - 1$ *there holds*

$$\int_{t_1}^{t_2}\int_{B_r(x_0)} |\tau_h(D^k u)|^2 \,\mathrm{d}z \le c(n,m,L)\,|h|^{\frac{m-k}{m}} \int_{t_1}^{t_2}\int_{B_{2r}(x_0)} \big(1+\big[|D^m u|\big]_h + |\tau_h(D^m u)|\big)^2 \,\mathrm{d}z.$$

Proof Applying the Interpolation-Lemma 2.4 “slicewise" on the ball $B_r(x_0)$ with the choice $\varepsilon = |h|^{\frac{m-k}{m}}/r^{2(m-k)}$ (note that $|h|^{\frac{m-k}{m}}/r^{2(m-k)} \le 1$ by assumption) we infer that

$$\int_{t_1}^{t_2}\int_{B_r(x_0)} |D^k(\tau_h u)|^2 \,\mathrm{d}z \le |h|^{\frac{m-k}{m}} \int_{t_1}^{t_2}\int_{B_r(x_0)} |D^m(\tau_h u)|^2 \,\mathrm{d}z + c(n,m)\,|h|^{-\frac{k}{m}} \int_{t_1}^{t_2}\int_{B_r(x_0)} |\tau_h u|^2 \,\mathrm{d}z.$$

From Lemma 6.1 we obtain a bound for the second integral on the right-hand side (note that $|h|/r^{2m} \le 1$ by assumption). Finally noting that $D^k(\tau_h u) = \tau_h(D^k u)$ we infer the asserted estimate. □

Corollary 6.3 *Suppose that* $u \in L^2(-T,0;W^{m,2}(\Omega;\mathbb{R}^N))$ *is a weak solution of system* (1.1) *under the assumptions* (1.3) *and* (1.4) *and let* $(t_1,t_2) \Subset (-T,0)$, $B_{2r}(x_0) \Subset \Omega$ *with* $0 < r \le 1$ *and* $0 < |h| < \frac{1}{2}\min\{|t_2|, T+t_1, r^{2m}\}$. *Then for* $0 \le k \le m-1$ *there holds*

$$\int_{t_1}^{t_2}\int_{B_r(x_0)} |\tau_h(D^k u)|^2 \,\mathrm{d}z \le c(n,m,L)\,|h|^{\frac{m-k}{m}} \int_{t_1-|h|}^{t_2+|h|}\int_{B_{2r}(x_0)} \big(1+|D^m u|\big)^2 \,\mathrm{d}z.$$

Proof The conclusion immediately follows from Corollary 6.2, since we can further estimate the right-hand side with the help of (2.6) and (5.1) (and respectively their analogues for negative h). □

Remark 6.4 Under the assumptions of Corollary 6.3 we conclude with Lemma 5.1 that for $0 \le k \le m-1$ and for all $\gamma \in (0,1)$ we have

$$D^k u \in W_{loc}^{m-k,\frac{\gamma(m-k)}{2m};2}(\Omega_T;\mathbb{R}^{\mathscr{M}_k}).$$

We will also need a version of Corollary 6.2 where the exponent of $|\tau_h(D^k u)|$ is larger then 2. We will attain such an estimate by transferring the estimate from Corollary 6.2 to a “larger” L^p-norm with the help of the Hardy-Littlewood Maximal function and the Sharp function.

Lemma 6.5 *Suppose that* $u \in L^2(-T,0;W^{m,2}(\Omega;\mathbb{R}^N))$ *is a weak solution of system* (1.1) *under the assumptions* (1.3) *and* (1.4) *and let* $(t_1,t_2) \Subset (-T,0)$, $B_{2r}(x_0) \Subset \Omega$ *with* $0 < r \le 1$. *Furthermore suppose that* $|D^m u| \in L^{2+b}_{loc}(\Omega_T)$ *for some* $b > 0$. *Then there exists* $c = c(n,m,L,b)$ *such that for all* $0 \le k \le m-1$ *and* $0 < |h| < \frac{1}{2}\min\{|t_2|, T+t_1, 1\}$ *there holds*

$$\begin{aligned}\int_{t_1}^{t_2}\int_{B_r(x_0)} &|\tau_h(D^k u)|^{2+b} \,\mathrm{d}z \\ &\le c\,|h|^{\frac{(2+b)(m-k)}{2m}} \int_{t_1}^{t_2}\int_{B_{2r}(x_0)} \big(1+|D^m u| + \big[|D^m u|\big]_h + |\tau_h(D^m u)|\big)^{2+b} \,\mathrm{d}z.\end{aligned}$$

Proof We choose a parabolic cylinder $Q_\rho(\bar z) = B_\rho(\bar x)\times(\bar t-\rho^{2m},\bar t\,)\subset B_r(x_0)\times(t_1,t_2)$ and denote $Q_{2\rho,\rho^{2m}}(\bar z) = B_{2\rho}(\bar x)\times(\bar t-\rho^{2m},\bar t\,)$. Now we distinguish the two cases $\rho^{2m}\ge|h|$ and $\rho^{2m}<|h|$.

In the case $\rho^{2m}\ge|h|$ we can apply Corollary 6.2 to obtian for $0\le j\le m-1$, with $c=c(n,m,L)$:

$$\begin{aligned}\fint_{Q_\rho(\bar z)}|\tau_h D^j u-(\tau_h D^j u)_{\bar z,\rho}|^2\,\mathrm{d}z &\le 2\fint_{Q_\rho(\bar z)}|\tau_h D^j u|^2\,\mathrm{d}z\\ &\le c\,|h|^{\frac{m-j}{m}}\fint_{Q_{2\rho,\rho^{2m}}(\bar z)}\Big(1+\big[|D^m u|\big]_h+|\tau_h(D^m u)|\Big)^2\,\mathrm{d}z.\end{aligned}$$

In the case $\rho^{2m}<|h|$, let $\widetilde\eta\in C_0^\infty(B_{2\rho}(\bar x))$ be a non-negative weight-function with $\int_{B_{2\rho}(\bar x)}\widetilde\eta\,\mathrm{d}x=1$ and $\|D^\ell\widetilde\eta\|_{L^2(B_\rho(\bar x))}\le c\rho^{-(\frac n2+\ell)}$ for $0\le\ell\le 2m$. Exploiting the weighted means of $\tau_h D^j u$ defined in (3.14) we can decompose the integral similar to (3.18) in the proof of the Poincaré type inequality. Then, applying the Poincaré inequality on the horizontal slices $B_\rho(\bar x)\times\{t\}$ for a.e. $t\in(\bar t-\rho^{2m},\bar t)$ we infer

$$\begin{aligned}\fint_{\bar t-\rho^{2m}}^{\bar t}\fint_{B_\rho(\bar x)}|\tau_h D^j u-(\tau_h D^j u)_{\bar z,\rho}|^2\,\mathrm{d}x\,\mathrm{d}t &\le c(n)\,\rho^2\fint_{\bar t-\rho^{2m}}^{\bar t}\fint_{B_\rho(\bar x)}|\tau_h D^{j+1}u|^2\,\mathrm{d}x\,\mathrm{d}t\\ &\quad+3\fint_{\bar t-\rho^{2m}}^{\bar t}\fint_{\bar t-\rho^{2m}}^{\bar t}\big|(\tau_h D^j u)_{\widetilde\eta}(t)-(\tau_h D^j u)_{\widetilde\eta}(\tau)\big|^2\,\mathrm{d}\tau\,\mathrm{d}t.\end{aligned}$$

To estimate the integrand of the second integral we use the estimate for differences of weighted means from Remark 3.14 on the cylinders $Q_\rho(\bar z)$ and $Q_\rho(\bar x,\bar t+h)$. Noting that $\rho^{m-j}\le|h|^{\frac{m-j}{2m}}$ we obtain for a.e. $t,\tau\in(\bar t-\rho^{2m},\bar t\,)$ that

$$\begin{aligned}&|(\tau_h D^j u)_{\widetilde\eta}(t)-(\tau_h D^j u)_{\widetilde\eta}(\tau)|\\ &\quad\le|(D^j u)_{\widetilde\eta}(t+h)-(D^j u)_{\widetilde\eta}(\tau+h)|+|(D^j u)_{\widetilde\eta}(t)-(D^j u)_{\widetilde\eta}(\tau)|\\ &\quad\le c\,\frac{\rho^{2m}}{\rho^{m+j}}\fint_{Q_\rho(\bar z)}\big(1+|D^m u(x,t+h)|+|D^m u|\big)\,\mathrm{d}x\,\mathrm{d}t\\ &\quad\le c(n,m,L)\,|h|^{\frac{m-j}{2m}}\fint_{Q_\rho(\bar z)}\big(1+|\tau_h D^m u|+|D^m u|\big)\,\mathrm{d}z,\end{aligned}$$

where we have also used the fact that $|D^m u(x,t+h)|\le|\tau_h D^m u(x,t)|+|D^m u(x,t)|$. Inserting this above, using Hölder's inequality and noting that $\rho^2\le|h|^{\frac1m}$ we arrive at

$$\begin{aligned}&\fint_{Q_\rho(\bar z)}|\tau_h D^j u-(\tau_h D^j u)_{\bar z,\rho}|^2\,\mathrm{d}z\\ &\quad\le c\,|h|^{\frac1m}\fint_{Q_\rho(\bar z)}|\tau_h D^{j+1}u|^2\,\mathrm{d}z+c\,|h|^{\frac{m-j}{m}}\fint_{Q_\rho(\bar z)}\big(1+|\tau_h D^m u|+|D^m u|\big)^2\,\mathrm{d}z.\end{aligned}$$

Combining both cases we conclude that for $0 \le j \le m-1$ and for all parabolic cylinders $Q_\rho(\bar z) \subset B_r(x_0) \times (t_1, t_2)$ there holds

$$\begin{aligned} &\fint_{Q_\rho(\bar z)} |\tau_h D^j u - (\tau_h D^j u)_{\bar z,\rho}|^2 \,\mathrm{d}z \\ &\quad \le c\,|h|^{\frac{1}{m}} \fint_{Q_\rho(\bar z)} |\tau_h D^{j+1} u|^2 \,\mathrm{d}z + c(n,m,L)\,|h|^{\frac{m-j}{m}} \fint_{Q_{2\rho,\rho^{2m}}(\bar z)} |w_h|^2 \,\mathrm{d}z, \end{aligned}$$

where

$$w_h \equiv 1 + |D^m u| + \big[|D^m u|\big]_h + |\tau_h(D^m u)|.$$

We denote $Q_0 \equiv B_r(x_0) \times (t_1, t_2)$ and $\widehat{Q}_0 \equiv B_{2r}(x_0) \times (t_1, t_2)$. Our next aim is to derive a bound for the sharp function of $\tau_h D^j u$, defined below. For this we consider a point $\tilde z \in Q_0$ and a parabolic cylinder $Q_\rho(\bar z) \subset Q_0$ with $\tilde z \in Q_\rho(\bar z)$. From Hölder's inequality and the last estimate we infer that

$$\begin{aligned} \bigg(\fint_{Q_\rho(\bar z)} |\tau_h D^j u - (\tau_h D^j u)_{\bar z,\rho}| \,\mathrm{d}z\bigg)^2 &\le \fint_{Q_\rho(\bar z)} |\tau_h D^j u - (\tau_h D^j u)_{\bar z,\rho}|^2 \,\mathrm{d}z \\ &\le c\,|h|^{\frac{1}{m}}\, M\big(|\tau_h D^{j+1} u|^2 \chi_{Q_0}\big)(\tilde z) + c\,|h|^{\frac{m-j}{m}}\, \widetilde M\big(|w_h|^2 \chi_{\widehat Q_0}\big)(\tilde z), \end{aligned}$$

where $c = c(n, m, L)$ and M, $\widetilde M$ denote the maximal functions defined by

$$M(f)(\tilde z) \equiv \sup_{\tilde z \in Q_r(\hat z)} \fint_{Q_r(\hat z)} |f| \,\mathrm{d}z, \qquad \widetilde M(f)(\tilde z) \equiv \sup_{\tilde z \in Q_{2r,r^{2m}}(\hat z)} \fint_{Q_{2r,r^{2m}}(\hat z)} |f| \,\mathrm{d}z$$

for an integrable function $f\colon \mathbb{R}^{n+1} \to \mathbb{R}$ and $\tilde z \in \mathbb{R}^{n+1}$. Taking the supremum over all parabolic cylinders $Q_\rho(\bar z)$ with $\tilde z \in Q_\rho(\bar z)$ we find that for each $\tilde z \in Q_0$ there holds

$$[\tau_h D^j u]^\#_{Q_0}(\tilde z) \le c\,|h|^{\frac{1}{2m}}\, M\big(|\tau_h D^{j+1} u|^2 \chi_{Q_0}\big)^{\frac12}(\tilde z) + c\,|h|^{\frac{m-j}{2m}}\, \widetilde M\big(|w_h|^2 \chi_{\widehat Q_0}\big)^{\frac12}(\tilde z),$$

where $c = c(n, m, L)$ and $f^\#_{Q_0}$ denotes the localized Sharp function of f:

$$f^\#_{Q_0}(\tilde z) \equiv \sup_{Q \subset Q_0,\ \tilde z \in Q} \fint_Q |f - (f)_Q| \,\mathrm{d}z,$$

where the supremum is taken over all parabolic cylinders Q with $z_0 \in Q \subset Q_0$. Due to a result of C. Fefferman and E. M. Stein, (see [21], Theorem 5, [24], Theorem 4.8) we know that $f^\#_{Q_0} \in L^p(Q_0)$ implies that $f \in L^p(Q_0)$ (and vice versa) and for $p > 1$ the following estimate holds

$$\int_{Q_0} |f - (f)_{Q_0}|^p \,\mathrm{d}z \le c(n,p) \int_{Q_0} |f^\#_{Q_0}|^p \,\mathrm{d}z.$$

We mention that in [21] and [24] the previous estmate is proved for a sharp function, where the supremum is taken over cubes. But the proof can be adapted to the parabolic geometry with minor changes. Therefore, we infer for $0 \le j \le m-1$ that

$$\begin{aligned}\int_{Q_0} |\tau_h D^j u - (\tau_h D^j u)_{Q_0}|^{2+b}\, dz &\le c \int_{Q_0} \big|[\tau_h D^j u]^{\#}_{Q_0}\big|^{2+b}\, dz \\ &\le c\, |h|^{\frac{2+b}{2m}} \int_{Q_0} M\big(|\tau_h D^{j+1}u|^2 \chi_{Q_0}\big)^{\frac{2+b}{2}}\, dz + c\, |h|^{\frac{(2+b)(m-j)}{2m}} \int_{Q_0} \widetilde{M}\big(|w_h|^2 \chi_{\widehat{Q}_0}\big)^{\frac{2+b}{2}}\, dz \\ &\le c\, |h|^{\frac{2+b}{2m}} \int_{Q_0} |\tau_h D^{j+1}u|^{2+b}\, dz + c\, |h|^{\frac{(2+b)(m-j)}{2m}} \int_{\widehat{Q}_0} |w_h|^{2+b}\, dz,\end{aligned}$$

where we have also used the Hardy-Littlewood maximal theorem in the last line. Here $c = c(n, m, L, b)$. Since we can bound the mean value $|(\tau_h D^j u)_{Q_0}|$ with the help of Corollary 6.2, we find that

$$\int_{Q_0} |\tau_h D^j u|^{2+b}\, dz \le c\, |h|^{\frac{2+b}{2m}} \int_{Q_0} |\tau_h D^{j+1}u|^{2+b}\, dz + c\, |h|^{\frac{(2+b)(m-j)}{2m}} \int_{\widehat{Q}_0} |w_h|^{2+b}\, dz,$$

where $c = c(n, m, L, b)$. Starting with $j = k$ and iterating this estimate for $j = k+1, \dots, m-1$ and recalling the definitions of w_h, $Q_0 \equiv B_r(x_0) \times (t_1, t_2)$ and $\widehat{Q}_0 \equiv B_{2r}(x_0) \times (t_1, t_2)$, we finally conclude the desired estimate. □

In the case that $D^m u$ admits better differentiability properties, with respect to t, we expect that this also affects $D^k u$, $0 \le k \le m-1$ in some sense. More precisely, if $D^m u$ is "γ - times differentiable" with respect to t for some $\gamma \in (0, 1)$, then we would expect that u is "$\frac{1}{2} + \gamma$ - times differentiable" with respext to t (then $D^k u$ is "$\frac{m-k}{2m} + \gamma$ - times differentiable"). But we do not get this property for first finite differences. In fact, the best we could get in the proof of Lemma 6.1 is that u is "$\frac{1}{2} + \frac{\gamma}{2}$ - times differentiable" with respect to t [this can be seen from (6.7)]. Therefore we turn our attention to second differences $\tau_{-h}(\tau_h D^k u)$.

Lemma 6.6 *Let $u \in L^2(-T, 0; W^{m,2}(\Omega; \mathbb{R}^N))$ be a weak solution of (1.1) under the assumptions (1.4) and (1.7) and let $(t_1, t_2) \Subset (-T, 0)$, $B_{2R}(x_0) \Subset \Omega$ and $0 < r < R \le 1$ and $0 < |h| < \frac{1}{2}\min\{|t_2|, T + t_1, |t_2 - t_1|, r^{2m}\}$. Moreover, suppose that $|D^m u| \in L^{2+b}((t_1 - |h|, t_2 + |h|) \times B_{2R}(x_0))$ for some $b \in (0, 2\beta)$ and $\delta u \in C^{0,\lambda/(2m)}((t_1 - |h|, t_2 + |h|) \times B_{2R}(x_0))$ for some $\lambda \in (0, 1)$. Then for all $0 \le k \le m-1$ there holds*

$$\begin{aligned}\int_{t_1}^{t_2}\int_{B_r} |\tau_{-h}(\tau_h D^k u)|^2\, dz \le \frac{c\, |h|^{\frac{m-k}{m}}}{(R-r)^{2m}} \Bigg[&\int_{t_1}^{t_2}\int_{B_R} |\tau_h D^m u|^2 + |\tau_{-h} D^m u|^2\, dz \\ &+ |h|^{\frac{1}{2m}(\lambda\beta + b(1-\frac{\lambda}{2}))} \int_{t_1-|h|}^{t_2+|h|}\int_{B_{2R}} 1 + |D^m u|^{2+b}\, dz \Bigg],\end{aligned}$$

where $c = c(n, m, L, [\delta u]_{0,\lambda/(2m)})$.

Proof Without loss of generality we assume that $x_0 = 0$ and we show the assertion only for $h > 0$. We start with the Steklov-formulation (2.7) of the system for $[u]_{\bar{h}}$ instead of $[u]_h$. Taking the difference at the levels $t+h$ and t we obtain for a.e. $t \in (t_1, t_2)$

$$\int_\Omega \tau_h \left(\partial_t [u]_{\bar{h}}(\cdot, t)\right) \cdot \varphi \, \mathrm{d}x$$
$$= -\int_\Omega \tau_h \left[A(\cdot, t, \delta u, D^m u)\right]_{\bar{h}} \cdot D^m \varphi + \tau_h \left[B(\cdot, t, \delta u, D^m u)\right]_{\bar{h}} \cdot \delta\varphi \, \mathrm{d}x$$

for all $\varphi \in W_0^{m,2}(\Omega, \mathbb{R}^N)$. We choose $r \le r_1 < r_2 \le R$ and a cut-off function $\eta \in C_0^\infty(B_{r_2})$ with $0 \le \eta \le 1$, $\eta \equiv 1$ on B_{r_1} and $|D^j \eta| \le c(r_2 - r_1)^{-j}$ for $j = 0, \dots, m$. Choosing the test-function $\varphi = \eta(\tau_{-h}\tau_h u)$, integrating over (t_1, t_2) and noting that $\partial_t [u]_{\bar{h}} = \frac{\tau_{-h} u}{-h}$, we infer that

$$\int_{t_1}^{t_2}\int_{B_{r_2}} \frac{|\tau_{-h}\tau_h u|^2}{h} \eta \, \mathrm{d}z$$
$$= -\int_{t_1}^{t_2}\int_{B_{r_2}} \tau_h(\partial_t [u]_{\bar{h}}) \cdot \eta(\tau_{-h}\tau_h u) \, \mathrm{d}z$$
$$= \int_{t_1}^{t_2}\int_{B_{r_2}} \tau_h [A(\cdot, \delta u, D^m u)]_{\bar{h}} \cdot D^m(\eta\tau_{-h}\tau_h u) + \tau_h [B(\cdot, \delta u, D^m u)]_{\bar{h}} \cdot \delta(\eta\tau_{-h}\tau_h u) \, \mathrm{d}z.$$

With the notation from (6.1) we decompose $\tau_h [A]_{\bar{h}} = [\tau_h A]_{\bar{h}} = [\mathscr{A}(h)]_{\bar{h}} + [\mathscr{B}(h)]_{\bar{h}} + [\mathscr{C}(h)]_{\bar{h}}$. With Young's inequality we obtain from the above equation for $\varepsilon > 0$ that

$$\int_{t_1}^{t_2}\int_{B_{r_2}} \frac{|\tau_{-h}\tau_h u|^2}{h} \eta \, \mathrm{d}z$$
$$\le \int_{t_1}^{t_2}\int_{B_{r_2}} \varepsilon |D^m(\eta\, \tau_{-h}\tau_h u)|^2 + \tfrac{1}{\varepsilon}\left(|[\mathscr{A}(h)]_{\bar{h}}|^2 + |[\mathscr{B}(h)]_{\bar{h}}|^2 + |[\mathscr{C}(h)]_{\bar{h}}|^2\right) \mathrm{d}z$$
$$+ \int_{t_1}^{t_2}\int_{B_{r_2}} \left|\tau_h \left[B(\cdot, \delta u, D^m u)\right]_{\bar{h}}\right| |\delta(\eta\, \tau_{-h}\tau_h u)| \, \mathrm{d}z$$
$$= \varepsilon I_1 + \tfrac{1}{\varepsilon}(I_2 + I_3 + I_4) + I_5, \tag{6.8}$$

with the obvious meaning of $I_1 - I_5$.

Estimate for I_1. We recall that $D^{\ell - j}\eta \equiv 0$ on B_{r_1} for $j \le \ell - 1$. Applying the Interpolation-lemma 2.4 we obtain for $0 \le j \le m-1$, $j \le \ell \le m$ and for all $0 < \mu \le 1$ that

$$
\begin{aligned}
&\int_{t_1}^{t_2}\int_{\operatorname{spt}D^{\ell-j}\eta}\frac{|D^j(\tau_{-h}\tau_h u)|^2}{(r_2-r_1)^{2(\ell-j)}}\,\mathrm{d}z\\
&\le \int_{t_1}^{t_2}\int_{B_{r_2}}\frac{\mu}{2}(r_2-r_1)^{2(m-\ell)}|D^m(\tau_{-h}\tau_h u)|^2+\frac{c}{\mu^{\frac{j}{m-j}}}\frac{|\tau_{-h}\tau_h u|^2}{(r_2-r_1)^{2\ell}}\,\mathrm{d}z\\
&\le \mu\int_{t_1}^{t_2}\int_{B_R}|\tau_h D^m u|^2+|\tau_{-h}D^m u|^2\,\mathrm{d}z+\frac{c(n,m)}{\mu^{m-1}}\int_{t_1}^{t_2}\int_{B_{r_2}}\frac{|\tau_{-h}\tau_h u|^2}{(r_2-r_1)^{2m}}\,\mathrm{d}z, \qquad (6.9)
\end{aligned}
$$

where we have also used the fact that $(r_2-r_1)\le 1$, $|h|\le 1$ and $|\tau_{-h}\tau_h f(t)|=|2f(t)-f(t+h)-f(t-h)|\le|\tau_h f(t)|+|\tau_{-h}f(t)|$. Using the assumptions on η, i.e. that $|D^{m-j}\eta|\le c(r_2-r_1)^{-(m-j)}$ for $0\le j\le m$ and the last estimate in the case $\mu=1$ and $\ell=m$ we infer that

$$
\begin{aligned}
I_1 &\le c\sum_{j=0}^{m}\int_{t_1}^{t_2}\int_{\operatorname{spt}D^{m-j}\eta}\frac{|D^j(\tau_{-h}\tau_h u)|^2}{(r_2-r_1)^{2(m-j)}}\,\mathrm{d}z\\
&\le 2m\int_{t_1}^{t_2}\int_{B_R}|\tau_h D^m u|^2+|\tau_{-h}D^m u|^2\,\mathrm{d}z+c(n,m)\int_{t_1}^{t_2}\int_{B_{r_2}}\frac{|\tau_{-h}\tau_h u|^2}{(r_2-r_1)^{2m}}\,\mathrm{d}z.
\end{aligned}
$$

Estimate for I_2. From (2.6), the bound (6.4) for $\mathscr{A}(h)$ and noting that $|h|\le|t_2-t_1|$ we get

$$
\begin{aligned}
I_2 &\le \int_{t_1-h}^{t_2}\int_{B_R}|\mathscr{A}(h)|^2\,\mathrm{d}z\le L^2\int_{t_1-h}^{t_2}\int_{B_R}|\tau_h(D^m u)|^2\,\mathrm{d}z\\
&\le L^2\int_{t_1}^{t_2}\int_{B_R}|\tau_h(D^m u)|^2+|\tau_{-h}(D^m u)|^2\,\mathrm{d}z.
\end{aligned}
$$

Estimate for I_3. Similarly, using (2.6) and (6.5), (1.9), Hölder's inequality (with exponents $\frac{2+b}{2}$, $\frac{2+b}{b}$) and noting that $|D^m u|\in L^{2+b}$ by assumption, we obtain

$$
I_3\le L^2\Bigg(\int_{t_1-h}^{t_2}\int_{B_R}\big(1+|D^m u|\big)^{2+b}\,\mathrm{d}z\Bigg)^{\frac{2}{2+b}}\Bigg(\int_{t_1-h}^{t_2}\int_{B_R}|\tau_h(\delta u)|^{\frac{2\beta(2+b)}{b}}\,\mathrm{d}z\Bigg)^{\frac{b}{2+b}}.
$$

We now exploit the Hölder continuity of δu, i.e. the fact that $|\tau_h(\delta u)(z)|\le|h|^{\frac{\lambda}{2m}}[\delta u]_{0,\lambda/(2m)}$ for all $z\in B_R\times(t_1-h,t_2)$, to diminish the exponent $2\beta(2+b)/b$ in the last integral to $2+b$. Taking into account that $|D^m u|\in L^{2+b}_{loc}$ we apply Lemma 6.5 (with t_1,t_2 replaced by t_1-h,t_2) to infer that

$$
\begin{aligned}
&\bigg(\int_{t_1-h}^{t_2}\int_{B_R} |\tau_h(\delta u)|^{\frac{2\beta(2+b)}{b}}\,dz\bigg)^{\frac{b}{2+b}}\\
&\le |h|^{\frac{\lambda(2\beta-b)}{2m}}\,[\delta u]_{0,\frac{\lambda}{2m}}^{2\beta-b}\bigg(\int_{t_1-h}^{t_2}\int_{B_R} |\tau_h(\delta u)|^{2+b}\,dz\bigg)^{\frac{b}{2+b}}\\
&\le c\,|h|^{\frac{b}{2m}+\frac{\lambda(2\beta-b)}{2m}}\bigg(\int_{t_1-h}^{t_2}\int_{B_{2R}} \big(1+|D^m u|+\big[|D^m u|\big]_h+|\tau_h D^m u|\big)^{2+b}\,dz\bigg)^{\frac{b}{2+b}}\\
&\le c\big(n,m,L,[\delta u]_{0,\lambda/(2m)}\big)\,|h|^{\frac{b}{2m}+\frac{\lambda(2\beta-b)}{2m}}\bigg(\int_{t_1-h}^{t_2+h}\int_{B_{2R}} \big(1+|D^m u|\big)^{2+b}\,dz\bigg)^{\frac{b}{2+b}}.
\end{aligned}
$$

Here we have used (2.6) and (5.1) in the last line. Inserting this estimate above we find that

$$
I_3 \le c(n,m,L,[\delta u]_{0,\lambda/(2m)})\,|h|^{\frac{1}{2m}(\lambda\beta+b(1-\frac{\lambda}{2}))}\int_{t_1-h}^{t_2+h}\int_{B_{2R}} \big(1+|D^m u|\big)^{2+b}\,dz.
$$

Here we have used the fact that $b < 2\beta$.

Estimate for I_4. From (2.6), and (6.5), the facts that $|h| \le 1$ and $\frac{1}{2m}(\lambda\beta + b(1-\frac{\lambda}{2})) = \frac{1}{2m}(\lambda(\beta-\frac{b}{2})+b) \le \frac{\beta}{m}$, since $b \in (0, 2\beta)$ we obtain

$$
I_4 \le L^2\,|h|^{\frac{1}{2m}(\lambda\beta+b(1-\frac{\lambda}{2}))}\int_{t_1-h}^{t_2}\int_{B_R} \big(1+|D^m u|\big)^2\,dz.
$$

Estimate for I_5. From Young's inequality we get for $\varepsilon > 0$

$$
I_5 \le \frac{\varepsilon}{2m\,|h|^{\frac{1}{m}}}\int_{t_1}^{t_2}\int_{B_{r_2}} |\delta(\eta\tau_{-h}\tau_h u)|^2\,dz + \frac{2m\,|h|^{\frac{1}{m}}}{\varepsilon}\int_{t_1}^{t_2}\int_{B_{r_2}} \big|\tau_h\big[B(\cdot,\delta u,D^m u)\big]_{\bar h}\big|^2\,dz.
$$

Using the estimate (6.9) in the case $\mu = |h|^{\frac{1}{m}}$ we infer for $1 \le \ell \le m-1$ that

$$
\begin{aligned}
&\int_{t_1}^{t_2}\int_{B_{r_2}} |D^\ell(\eta\,\tau_{-h}\tau_h u)|^2\,dz\\
&\le c\sum_{j=0}^{\ell}\int_{t_1}^{t_2}\int_{\operatorname{spt}D^{\ell-j}\eta} \frac{|D^j(\tau_{-h}\tau_h u)|^2}{(r_2-r_1)^{2(\ell-j)}}\,dz\\
&\le 2m\,|h|^{\frac{1}{m}}\int_{t_1}^{t_2}\int_{B_R} |\tau_h D^m u|^2 + |\tau_{-h} D^m u|^2\,dz + c(n,m)\,|h|^{-\frac{m-1}{m}}\int_{t_1}^{t_2}\int_{B_{r_2}} \frac{|\tau_{-h}\tau_h u|^2}{(r_2-r_1)^{2m}}\,dz.
\end{aligned}
$$

Inserting this above and using the growth (1.4) of B, (2.6), and (5.1) we obtain, with $c = c(n,m,L)$

$$I_5 \le \varepsilon \int_{t_1}^{t_2}\int_{B_R} |\tau_h D^m u|^2 + |\tau_{-h} D^m u|^2 \, dz + \frac{c\varepsilon}{|h|}\int_{t_1}^{t_2}\int_{B_{r_2}} \frac{|\tau_{-h}\tau_h u|^2}{(r_2-r_1)^{2m}} \, dz$$
$$+ \frac{c|h|^{\frac{1}{m}}}{\varepsilon}\int_{t_1-h}^{t_2+h}\int_{B_R} \big(1+|D^m u|\big)^2 \, dz.$$

Combining the estimates for I_1 - I_5 with (6.8) and noting that $|Du|^2 \le 1 + |Du|^{2+b}$, we arrive at

$$\int_{t_1}^{t_2}\int_{B_{r_1}} \frac{|\tau_{-h}\tau_h u|^2}{|h|} \, dz \le \frac{c\,\varepsilon}{(r_2-r_1)^{2m}}\int_{t_1}^{t_2}\int_{B_{r_2}} \frac{|\tau_{-h}\tau_h u|^2}{|h|} \, dz$$
$$+ \frac{c}{\varepsilon}\int_{t_1}^{t_2}\int_{B_R} |\tau_h D^m u|^2 + |\tau_{-h} D^m u|^2 \, dz$$
$$+ \frac{c\,|h|^{\frac{1}{2m}(\lambda\beta + b(1-\frac{\lambda}{2}))}}{\varepsilon}\int_{t_1-h}^{t_2+h}\int_{B_{2R}} 1 + |D^m u|^{2+b} \, dz,$$

where $c = c(n, m, L, [\delta u]_{0,\lambda/(2m)})$. We choose $\varepsilon = \frac{1}{2c}(r_2-r_1)^{2m}$ and apply Lemma 2.3 in the case $\vartheta = \frac{1}{2}$, to infer the asserted estimate in the case $k = 0$. In the case $1 \le k \le m-1$ we once again apply the Interpolation-Lemma 2.4 with $(\tau_{-h}\tau_h u, 0, r, 2, (|h|/r^{2m})^{\frac{m-k}{m}})$ instead of $(u, r_1, r_2, p, \varepsilon)$, which is possible since $|h|/r^{2m} \le 1$ by assumption. This finally yields the desired estimate. □

The following lemma is the starting point for the considerations concerning finite differences of $D^m u$.

Lemma 6.7 *Let $u \in L^2(-T, 0; W^{m,2}(\Omega; \mathbb{R}^N))$ be a weak solution of (1.1) with (1.2)–(1.4), (1.7) and (1.8) and suppose that $(t_1, t_2) \Subset (-T, 0)$, $B_R(x_0) \Subset \Omega$ and $0 < r < R \le 1$. Moreover let $\eta \in C_0^\infty(B_R(x_0))$ and $\zeta \in C^1(\mathbb{R})$ be two cut-off functions with $0 \le \eta \le 1$, $0 \le \zeta \le 1$, $\eta \equiv 1$ on $B_r(x_0)$, $|D^k \eta| \le c(R-r)^{-k}$ for $0 \le k \le m$ and $\zeta \equiv 0$ on $(-\infty, t_1]$. Then for $0 < |h| < \frac{1}{2}\min\{T + t_1, |t_2|, 1\}$ there holds*

$$\int_{t_1}^{t_2}\int_{B_r(x_0)} |\tau_h(D^m u)|^2 \zeta^2 \, dz \le c\,|h|^{\frac{\beta}{m}}\int_{t_1}^{t_2}\int_{B_R(x_0)} \big(1+|D^m u|\big)^2 \zeta^2 \, dz$$
$$+ c\int_{t_1}^{t_2}\int_{B_R(x_0)} \sum_{k=0}^{m-1} \frac{|\tau_h(D^k u)|^2}{(R-r)^{2(m-k)}}\zeta^2 + \|\zeta'\|_\infty |\tau_h u|^2 \zeta \, dz$$
$$+ c\int_{t_1}^{t_2}\int_{B_R(x_0)} \big(1+|D^m u|\big)^2 \,\widetilde{\theta}\,(|\tau_h(\delta u)|)^2 \zeta^2 \, dz$$
$$+ c\left|\int_{t_1}^{t_2}\int_{B_R(x_0)} \tau_h\big[B(\cdot, \delta u, D^m u)\big]\cdot \delta(\eta^{2m}\tau_h u)\zeta^2 \, dz\right|,$$

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

We mention that the analogue estimate also holds for finite differences τ_h^s, $s = 1, \dots, n$ in space-direction and the proof can be completely adapted.

Proof Without loss of generality we can assume that $x_0 = 0$. We choose $0 < |\lambda| < \frac{1}{2}\min\{T + t_1, |t_2|, 1\}$. Our starting point is the Steklov-formulation (2.7) of the system, with λ instead of h. Taking the difference of (2.7) at the levels $t + h$ and t and using that $\tau_h(\partial_t u_\lambda) = \partial_t(\tau_h u_\lambda)$, we obtain for a.e. $t \in (t_1, t_2)$

$$\int_\Omega \left(\partial_t\,(\tau_h[u]_\lambda(\cdot,t))\cdot\varphi + \tau_h\left[A(\cdot,t,\delta u, D^m u)\right]_\lambda \cdot D^m\varphi\right) \mathrm{d}x = \int_\Omega \tau_h\left[B(\cdot,t,\delta u, D^m u)\right]_\lambda \cdot \delta\varphi\,\mathrm{d}x$$

for all $\varphi \in W_0^{m,2}(\Omega, \mathbb{R}^N)$. We now choose the test-function $\varphi_\lambda = (\tau_h[u]_\lambda)\eta^{2m}\zeta^2$, where η and ζ are specified in the statement of the lemma. Noting that $\zeta(t_1) = 0$ we calculate for the first-term on the left-hand side, integrated over (t_1, t_2):

$$\int_{t_1}^{t_2}\int_\Omega \partial_t\,(\tau_h[u]_\lambda)\cdot\varphi_\lambda\,\mathrm{d}z = \frac{1}{2}\int_\Omega \eta^{2m}|\tau_h[u]_\lambda(t_2)|^2\,\zeta^2(t_2)\,\mathrm{d}x - \int_{t_1}^{t_2}\int_\Omega \eta^{2m}|\tau_h[u]_\lambda|^2\,\zeta\zeta'\,\mathrm{d}z.$$

Now, we integrate the above system over (t_1, t_2) and insert the previous identity. Passing to the limit $\lambda \searrow 0$ and noting that the term involving $\zeta(t_2)$ is non-negative, we arrive at

$$\int_{t_1}^{t_2}\int_\Omega \tau_h\left[A(\cdot,\delta u, D^m u)\right]\cdot D^m\varphi\,\mathrm{d}z \le \int_{t_1}^{t_2}\int_\Omega \tau_h\left[B(\cdot,\delta u, D^m u)\right]\cdot\delta\varphi\,\mathrm{d}z + \int_{t_1}^{t_2}\int_\Omega |\tau_h u|^2\,\eta^{2m}\zeta\zeta'\,\mathrm{d}z,$$

where $\varphi = (\tau_h u)\eta^{2m}\zeta^2$. With the chain rule we compute for $0 \le j \le m$ that $D^j\eta^{2m} = \eta^{2m-j}\mathscr{F}_j(\eta)$, where $\mathscr{F}_j(\eta) \in \mathbb{R}^{\mathscr{M}_j}$ and $|\mathscr{F}_j(\eta)| \le c(n,m)\,(R-r)^{-j}$. With this notation we find that

$$D^m\varphi = \Bigg(D^m(\tau_h u)\,\eta^{2m} + \eta^m \underbrace{\sum_{k=0}^{m-1}\binom{m}{k} D^k(\tau_h u)\odot\mathscr{F}_{m-k}(\eta)\,\eta^k}_{\equiv \mathrm{LOT}}\Bigg)\zeta^2.$$

Recalling the decomposition of $\tau_h(A)$ from (6.1), i.e. $\tau_h(A)(x,t) = \mathscr{A}(h) + \mathscr{B}(h) + \mathscr{C}(h)$, we obtain:

$$\begin{aligned}
&\int_{t_1}^{t_2}\int_{B_R} \mathscr{A}(h)\cdot D^m(\tau_h u)\,\eta^{2m}\zeta^2\,\mathrm{d}z \\
&\quad\le \int_{t_1}^{t_2}\int_{B_R} \left(|\mathscr{B}(h)+\mathscr{C}(h)||D^m(\tau_h u)|\,\eta^{2m} + |\mathscr{A}(h)+\mathscr{B}(h)+\mathscr{C}(h)||\mathrm{LOT}|\,\eta^m\right)\zeta^2\,\mathrm{d}z \\
&\quad\quad + \left|\int_{t_1}^{t_2}\int_{B_R} \tau_h\left[B(\cdot,\delta u, D^m u)\right]\cdot\delta\varphi\,\mathrm{d}z\right| + \int_{t_1}^{t_2}\int_{B_R} |\tau_h u|^2\,\eta^{2m}\zeta\zeta'\,\mathrm{d}z.
\end{aligned}$$

With the notation from (6.2) we write $\mathscr{A}(h) = \widetilde{\mathscr{A}}(h)\tau_h(D^m u)$ and infer from the ellipticity (6.3) of $\widetilde{\mathscr{A}}(h)$ an estimate from below for the integrand on the left-hand side of the above inequality, i.e. $\mathscr{A}(h) \cdot \tau_h(D^m u) \geq \nu|\tau_h(D^m u)|^2$. Also using the bound (6.4) for $\mathscr{A}(h)$, i.e. $|\mathscr{A}(h)| \leq L|\tau_h D^m u|$, Young's inequality and $0 \leq \eta \leq 1$, we obtain

$$\begin{aligned}
&\nu \int_{t_1}^{t_2}\int_{B_R} |\tau_h(D^m u)|^2\, \eta^{2m}\zeta^2 \,\mathrm{d}z \\
&\leq \int_{t_1}^{t_2}\int_{B_R} \tfrac{\nu}{2}|\tau_h(D^m u)|^2\, \eta^{2m}\zeta^2 + c\left(|\mathscr{B}(h)|^2 + |\mathscr{C}(h)|^2 + |\mathrm{LOT}|^2\right)\zeta^2 \,\mathrm{d}z \\
&\quad + \left|\int_{t_1}^{t_2}\int_{B_R} \tau_h\left[B(\cdot,\delta u, D^m u)\right]\cdot \delta\varphi \,\mathrm{d}z\right| + \int_{t_1}^{t_2}\int_{B_R} |\tau_h u|^2\, \eta^{2m}\zeta\zeta' \,\mathrm{d}z,
\end{aligned}$$

where $c = c(\nu, L)$. Now we estimate the remaining integrals on the right-hand side. The first term can be absorbed on the left-hand side. For $\mathscr{B}(h)$ and $\mathscr{C}(h)$ we use the estimates in (6.5) and for the the terms of lower order we recall that $|\mathscr{F}_{m-k}(\eta)| \leq c(n,m)(R-r)^{-(m-k)}$ for $0 \leq k \leq m-1$. Finally, dividing by $\nu/2$ and noting that $\eta \equiv 1$ on B_r we obtain the asserted estimate. □

6.2 Fractional estimates

In this chapter we show that $D^m u$ admits certain fractional differentiability properties. For this we have to assume slightly stronger hypothesis for the coefficients A and a mild regularity for δu. Our main strategy can be described in the following way. We first derive estimates for fractional difference quotients in x- and t- direction. This implies that $D^m u$ lies in a certain Nicolskii space. Thus, due to the embedding from Lemma 5.1 we can conclude the fractional differentiability of $D^m u$ in the sense of the $W^{\gamma,\frac{\gamma}{2m}}$ spaces.

Lemma 6.8 *Let $u \in L^2(-T,0; W^{m,2}(\Omega;\mathbb{R}^N))$ be a weak solution of (1.1) under the assumptions (1.2)–(1.4), (1.7) and (1.8). Then for any $\widetilde{\Omega} \Subset \Omega$ and $(t_1,t_2) \Subset (-T,0)$ and for all $\gamma < \frac{\delta}{2}$ there holds*

$$\begin{aligned}
&\int_{\widetilde{\Omega}}\int_{t_1}^{t_2}\int_{t_1}^{t_2} \frac{|D^m u(x,t) - D^m u(x,\tau)|^2}{|t-\tau|^{1+\frac{\gamma}{m}}} \,\mathrm{d}t\,\mathrm{d}\tau\,\mathrm{d}x \\
&\quad + \int_{t_1}^{t_2}\int_{\widetilde{\Omega}}\int_{\widetilde{\Omega}} \frac{|D^m u(x,t) - D^m u(y,t)|^2}{|x-y|^{n+2\gamma}} \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t \leq c < \infty,
\end{aligned}$$

where $c = c(n,m,N,\nu,L,\beta,\gamma,\mathrm{dist}(\widetilde{\Omega},\partial\Omega), |t_2|, T+t_1, \|u\|_{L^2(-T,0;W^{m,2}(\Omega))})$ and $\delta = \frac{\beta\sigma}{1+\sigma}$, where $\sigma = \sigma(n,m,N,L/\nu) > 0$ is the exponent from Theorem 4.1 about the higher integrability of $|D^m u|$.

Moreover, in the case of simpler systems of the type (1.10), the assertion holds for all $\gamma < \beta$.

Proof Here we shall take symmetric parabolic cylinders, i.e. for $z_0 \in \mathbb{R}^{n+1}$, $R > 0$ we denote $Q_R(z_0) \equiv B_R(x_0) \times (t_0 - R^{2m}, t_0 + R^{2m})$. Now, we will start with the time-direction, i.e. the first integral in the statement of the lemma. We choose a parabolic cylinder $Q_R(z_0) \Subset \Omega_T$ with $R \le 1$. Without loss of generality we can assume that $z_0 = 0$; then $Q_R(z_0) = Q_R = B_R \times (-R^{2m}, R^{2m})$. Moreover, we take $0 < |h| \le (R/4)^{2m}$. Let $\eta \in C_0^\infty(B_{R/2})$ and $\zeta \in C_0^1(-(R/2)^{2m}, (R/2)^{2m})$ be two cut-off functions with $0 \le \eta \le 1$, $0 \le \zeta \le 1$, $\eta \equiv 1$ on $B_{R/4}$, $\zeta \equiv 1$ on $(-(R/4)^{2m}, (R/4)^{2m})$, $|D^k\eta| \le c\,R^{-k}$ for $0 \le k \le m$ and $|\zeta'| \le c\,R^{-2m}$. With this particular choice of cut-off functions and with $(R/4, R/2)$ instead of (r, R) we infer from Lemma 6.7 that

$$
\begin{aligned}
\int_{Q_{R/4}} |\tau_h(D^m u)|^2\,\mathrm{d}z \le c\,|h|^{\frac{\beta}{m}} \int_{Q_{R/2}} \left(1 + |D^m u|\right)^2\,\mathrm{d}z + c \int_{Q_{R/2}} \sum_{k=0}^{m-1} \frac{|\tau_h(D^k u)|^2}{R^{2(m-k)}}\,\mathrm{d}z \\
+ c \int_{Q_{R/2}} \left(1 + |D^m u|\right)^2 \widetilde{\theta}\left(|\tau_h(\delta u)|\right)^2\,\mathrm{d}z \\
+ c \int_{Q_{R/2}} \left|\tau_h\left[B(\cdot, \delta u, D^m u)\right]\right| |\delta(\eta^{2m}\tau_h u)|\,\mathrm{d}z \\
= c\,(I_1 + I_2 + I_3 + I_4),
\end{aligned}
\tag{6.10}
$$

with the obvious meaning of $I_1 - I_4$ and where $c = c(n, m, \nu, L)$. Our aim is to show that $\tau_h(D^m u)$ lies in a suitable Nicolskii-space. For this we still have to consider the terms $I_2 - I_4$.

Estimate for I_2. Applying Corollary 6.3 with $R/2$ instead of r, which is possible since $|h| \le (R/4)^{2m}$ and taking into account that $(R/2)^{2m} + |h| \le R^{2m}$ we infer for $0 \le k \le m-1$

$$
\int_{Q_{R/2}} |\tau_h(D^k u)|^2\,\mathrm{d}z \le c(n, m, L)\,|h|^{\frac{m-k}{m}} \int_{Q_R} \left(1 + |D^m u|^2\right)\,\mathrm{d}z. \tag{6.11}
$$

Since $|h| \le 1$ this bound implies that $I_2 \le c\,|h|^{\frac{1}{m}}$, where $c = c(n, m, L, R, \|u\|_{L^2(-T,0;W^{m,2}(\Omega))})$.

Estimate for I_3. From Theorem 4.1 we know that $|D^m u|$ is higher integrable, i.e. there exists $\sigma = \sigma(n, m, N, L/\nu) > 0$ such that $|D^m u| \in L^{2(1+\sigma)}_{loc}(\Omega_T)$. Moreover, using Hölder's inequality and the fact that $\widetilde{\theta} \le 1$ and therefore $\widetilde{\theta}(\cdot)^{\frac{2(1+\sigma)}{\sigma}} \le \widetilde{\theta}(\cdot)$ we obtain

$$
\begin{aligned}
I_3 &\le |Q_{R/2}| \left(\fint_{Q_{R/2}} \widetilde{\theta}\left(|\tau_h(\delta u)|\right)^{\frac{2(1+\sigma)}{\sigma}}\,\mathrm{d}z \right)^{\frac{\sigma}{1+\sigma}} \left(\fint_{Q_{R/2}} \left(1 + |D^m u|\right)^{2(1+\sigma)}\,\mathrm{d}z \right)^{\frac{1}{1+\sigma}} \\
&\le c(n, m, N, \nu, L)\,|Q_{R/2}| \left(\fint_{Q_{R/2}} \widetilde{\theta}\left(|\tau_h(\delta u)|\right)\,\mathrm{d}z \right)^{\frac{\sigma}{1+\sigma}} \fint_{Q_R} \left(1 + |D^m u|\right)^2\,\mathrm{d}z,
\end{aligned}
$$

To bound the first integral on the right-hand side we apply Jensen's inequality to the concave function $\widetilde{\theta}$, use the growth condition (1.9) on $\widetilde{\theta}$ and Hölder's inequality to infer that

$$\left(\fint_{Q_{R/2}} \widetilde{\theta}\left(|\tau_h(\delta u)|\right) \mathrm{d}z \right)^{\frac{\sigma}{1+\sigma}} \leq c\, \widetilde{\theta}\left(\fint_{Q_{R/2}} |\tau_h(\delta u)| \,\mathrm{d}z \right)^{\frac{\sigma}{1+\sigma}}$$
$$\leq c(n, m, L) \left(\fint_{Q_{R/2}} |\tau_h(\delta u)|^2 \,\mathrm{d}z \right)^{\frac{\beta\sigma}{2(1+\sigma)}}.$$

Inserting this above and using the bound in (6.11) and that $|h|^{\frac{m-k}{m}} \leq |h|^{\frac{1}{m}}$ for $0 \leq k \leq m-1$ we arrive at

$$I_3 \leq c \left(\fint_{Q_{R/2}} |\tau_h(\delta u)|^2 \,\mathrm{d}z \right)^{\frac{\beta\sigma}{2(1+\sigma)}} \fint_{Q_R} \left(1 + |D^m u|\right)^2 \,\mathrm{d}z \leq c\, |h|^{\frac{\beta\sigma}{2m(1+\sigma)}},$$

where $c = c(n, m, N, \nu, L, R, \|u\|_{L^2(-T,0;W^{m,2}(\Omega))})$.

Estimate for I_4**.** We first apply Hölder's inequality. Then we use the growth assumption (1.4) on B and (5.1) in the first resulting integral. To estimate the second integral we differentiate $\eta^{2m}\tau_h u$ with respect to x, taking into account that $|D^j\eta| \leq c\, R^{-j}$ for $0 \leq j \leq m-1$. Finally, exploiting the bound in (6.11) we obtain

$$I_4 \leq c(n, m)\, L \left(\int_{Q_R} \left(1 + |D^m u|\right)^2 \,\mathrm{d}z \right)^{\frac{1}{2}} \left(\sum_{k=0}^{m-1} \sum_{j=0}^{k} \int_{Q_{R/2}} \frac{|D^j(\tau_h u)|^2}{R^{2(k-j)}} \,\mathrm{d}z \right)^{\frac{1}{2}} \leq c\, |h|^{\frac{1}{2m}},$$

where $c = c(n, m, L, R, \|u\|_{L^2(-T,0;W^{m,2}(\Omega))})$.

Combining the estimates for $I_2 - I_4$ with (6.10) and taking into account that $|h| \leq 1$ and that $I_3 = I_4 = 0$ for simpler systems of the type (1.10) we obtain

$$\int_{Q_{R/4}} |\tau_h(D^m u)|^2 \,\mathrm{d}z \leq c \left(|h|^{\frac{\beta}{m}} + |h|^{\frac{1}{m}}\right) + c \left(|h|^{\frac{\beta\sigma}{2m(1+\sigma)}} + |h|^{\frac{1}{2m}}\right)$$
$$\leq c_1\, |h|^{\frac{\beta}{m}} + c_2\, |h|^{\frac{\beta\sigma}{2m(1+\sigma)}}, \tag{6.12}$$

where $c_i = c_i(n, m, \nu, L, \beta, R, \|u\|_{L^2(-T,0;W^{m,2}(\Omega))})$ for $i = 1, 2$ and $c_2 = 0$, for simpler systems of the type (1.10). Since the constants c_1 and c_2 are independent of h, we can apply Lemma 5.1 and conclude that

$$[D^m u]_{0, \frac{\gamma}{2m}; 2; Q_{R/4}} \leq c(n, m, \nu, L, \beta, \gamma, R, \|u\|_{L^2(-T,0;W^{m,2}(\Omega))}) \quad \text{for all } \gamma < \tfrac{\delta}{2} = \tfrac{\beta\sigma}{2(1+\sigma)}.$$

For systems of the type (1.10), the above statement even holds for all $\gamma < \beta$, since in this case we have $c_2 = 0$ in (6.12). Since this bound is independent of the particular considered cylinder, the desired local estimate on $\widetilde{\Omega} \times (t_1, t_2)$ follows with a standard covering argument.

The proof for the spacial fractional differentiability, is essentially similar. Therefore, we will only outline the differences. As starting point we use a spacial analogue of Lemma 6.7. The terms appearing on the right-hand side can then be estimated similarly, apart from one difference: Instead of applying Corollary 6.3 to infer (6.11), we now can exploit the fact that δu has weak derivatives in L^2 with respect to the space variable x, i.e. we can use the estimate (6.6). Then, proceeding as above we can also conclude the fractional differentiability with respect to the space direction. □

Under the additional assumption that δu is Hölder continuous we can show $D^m u \in W^{\gamma, \frac{\gamma}{2m}; 2}_{loc}(\Omega_T; \mathbb{R}^{\mathcal{N}})$ for all $\gamma < \beta$, although the coefficients A depend on δu. This is proved by an iteration argument in the following lemma. In each step of the iteration, the "order" of fractional differentiability of $D^m u$ can be improved. Within this procedure the Interpolation-Theorem 5.7 plays an essential role.

Lemma 6.9 *Let $u \in L^2(-T,0;W^{m,2}(\Omega;\mathbb{R}^N)) \cap C^{m-1,\lambda,\frac{\lambda}{2m}}(\Omega_T;\mathbb{R}^N)$ with $\lambda \in (0,1)$ be a weak solution of (1.1) under the assumptions (1.2)–(1.4), (1.7) and (1.8). Then for all $\gamma < \beta$ we have $D^m u \in W^{\gamma,\frac{\gamma}{2m};2}_{loc}(\Omega_T;\mathbb{R}^{\mathscr{N}})$ and $D^{m-1}u \in W^{\frac{1+\gamma}{2m};2}_{loc}(-T,0;L^2(\Omega;\mathbb{R}^{\mathscr{M}_{m-1}}))$. Moreover for any $\widetilde{\Omega} \Subset \Omega$ and $(t_1,t_2) \Subset (-T,0)$ there holds*

$$[D^m u]_{\gamma,\frac{\gamma}{2m};2;\widetilde{\Omega}\times(t_1,t_2)} \le c < \infty, \qquad \textit{for all } \gamma < \beta,$$

where $c = c(n,m,N,\nu,L,\beta,\lambda,\gamma,\operatorname{dist}(\widetilde{Q},\partial\Omega_T), \|u\|_{L^2(-T,0;W^{m,2}(\Omega))}, [\delta u]_{\lambda,\frac{\lambda}{2m}})$.

Proof In the following, with C, respectively C_ℓ, $\widehat{C}_\ell$, $\widetilde{C}_\ell$ (with $\ell \in \mathbb{N}$) we denote constants depending on $n,m,N,\nu,L,\beta,\lambda$, $\|u\|_{L^2(-T,0;W^{m,2}(\Omega))}$ and $[\delta u]_{\lambda,\frac{\lambda}{2m}}$. We will only indicate the additional dependencies of these constats. Moreover, we once again take symmetric parabolic cylinders, i.e. for $z_0 \in \mathbb{R}^{n+1}$, $R > 0$ we denote $Q_R(z_0) \equiv B_R(x_0) \times (t_0 - R^{2m}, t_0 + R^{2m})$. We choose such a parabolic cylinder $Q_R(z_0) \Subset \Omega_T$ with $R \le 1$ and fix $\gamma \in (0,\beta)$. In the following we will show that

$$\begin{cases} D^m u \in W^{\gamma,\frac{\gamma}{2m};2}\big(Q_{R/64^{\bar\ell}}(z_0);\mathbb{R}^{\mathscr{N}}\big), \\ D^{m-1}u \in W^{1,\frac{1+\gamma}{2m};2}\big(Q_{R/64^{\bar\ell}}(z_0);\mathbb{R}^{\mathscr{M}_{m-1}}\big), \end{cases} \tag{6.13}$$

where $\bar\ell = \bar\ell(\beta,\gamma,\lambda) \in \mathbb{N}$.

Without loss of generality, in the following we consider $h > 0$ as parameter for the finite differences and we assume that $z_0 = 0$. Initiallay we define the sequence $(b_\ell)_{\ell\in\mathbb{N}}$ of positive real numbers

$$b_0 = 0, \qquad b_1 = \beta\lambda, \qquad b_{\ell+1} = \beta\lambda + b_\ell\left(1 - \tfrac{\lambda}{2}\right).$$

Rewriting $b_\ell = \beta\lambda\sum_{i=0}^{\ell-1}(1-\frac{\lambda}{2})^i$, we see that $b_\ell \nearrow 2\beta$ as $\ell \to \infty$. Furthermore we define a sequence of radii $\rho_\ell \equiv R/64^\ell$ for $\ell \in \mathbb{N}$ and we set $Q_\ell \equiv Q_{\rho_\ell}$ and denote $\alpha Q_\ell \equiv Q_{\alpha\rho_\ell}$ and $\alpha B_\ell \equiv B_{\alpha\rho_\ell}$ for $\alpha > 0$. In the following we show by induction that for all $\ell \in \mathbb{N}$, and for all $0 < h \le \rho_\ell^{2m}$ there holds

$$\begin{cases} \displaystyle\int_{16Q_\ell} |\tau_h D^m u|^2 + |\tau_{-h} D^m u|^2 \, dz \le C_\ell \, |h|^{\frac{b_\ell}{2m}}, \\ \displaystyle\int_{16Q_\ell} \sum_{s=1}^n |\tau_h^s D^m u|^2 \, dz \le C_\ell \, |h|^{b_\ell}, \end{cases} \tag{A_ℓ}$$

where $C_\ell = C_\ell(R,\ell)$. Moreover, we have $D^m u \in W^{\frac{\vartheta b_\ell}{2},\frac{\vartheta b_\ell}{4m};2}(8Q_\ell;\mathbb{R}^{\mathscr{N}})$ for all $\vartheta \in (0,1)$ and

$$[D^m u]_{\frac{\vartheta b_\ell}{2},\frac{\vartheta b_\ell}{4m};2;8Q_\ell} \le \widehat{C}_\ell(R,\ell,\vartheta). \tag{B_ℓ}$$

Furthermore, we have $|D^m u| \in L^{2+b_{\ell-1}}(Q_\ell)$ and

$$\int_{Q_{\ell-1}} |D^m u|^{2+b_{\ell-1}} \, dz \le \widetilde{C}_\ell(R,\ell). \tag{C_ℓ}$$

We further note that C_ℓ, $\widehat{C}_\ell$ and $\widetilde{C}_\ell$ depend on R, but not on the special cylinder $Q_R = Q_R(z_0)$ and that possibly $C_\ell, \widehat{C}_\ell, \widetilde{C}_\ell \to \infty$ as $\ell \to \infty$. Therefore we will stop the iteration at some finite step $\bar{\ell}$.

The case $\ell = 1$**.** First we show that $(A)_1$ holds. For this we will once again use the estimate (6.10) from the proof of Lemma 6.8 with the same labelling for $I_1 - I_4$. But here we treat the term I_3 in a different way. Since $\widetilde{\theta}(s) \le s^\beta$ for $s > 0$ by (1.9) and due to the Hölder continuity of δu we find that

$$I_3 \le \int\limits_{Q_{R/2}} \big(1 + |D^m u|\big)^2 \, |\tau_h(\delta u)|^{2\beta} \,\mathrm{d}z \le |h|^{\frac{\beta\lambda}{m}} \, [\delta u]^{2\beta}_{0,\frac{\lambda}{2m}} \int\limits_{Q_{R/2}} \big(1 + |D^m u|\big)^2 \,\mathrm{d}z.$$

For the remaining terms I_2 and I_4 in (6.10) we use the same estimates as before. Taking also into account that $b_1 = \beta\lambda$ by definition and $|h| \le 1$ we therefore conclude that

$$\int\limits_{Q_{R/4}} |\tau_h(D^m u)|^2 \,\mathrm{d}z \le c \, \Big(|h|^{\frac{\beta}{m}} + |h|^{\frac{1}{m}} + |h|^{\frac{\beta\lambda}{m}} + |h|^{\frac{1}{2m}}\Big) \le C_1 \, |h|^{\frac{b_1}{2m}},$$

with the asserted dependence of the constant. The analogue estimate for $-h$ instead of h can be shown similarly. Since $Q_{R/4} = Q_{16\rho_1} = 16Q_1$, this shows the first bound in $(A)_1$. Since the same reasoning can also be applied for the space-direction, we also infer the second estimate in $(A)_1$. Together, this proves $(A)_1$. By Lemma 5.1 $(A)_1$ now implies that $D^m u \in W^{\frac{\vartheta b_1}{2}, \frac{\vartheta b_1}{4m}; 2}(8Q_1; \mathbb{R}^{\mathcal{N}})$ for all $\vartheta \in (0, 1)$, i.e. $(B)_1$ holds. For $(C)_1$ there is nothing to show, since $|D^m u| \in L^2$ by assumption.

Induction step $\ell \to \ell + 1$**.** We now show $(A)_{\ell+1}$, $(B)_{\ell+1}$ and $(C)_{\ell+1}$ provided that $(A)_\ell$, $(B)_\ell$ and $(C)_\ell$ hold for some $\ell \ge 1$.

First, we show $(C)_{\ell+1}$. Applying Lemma 6.6 with $(8B_\ell, 16B_\ell, -(8\rho_\ell)^{2m}, (8\rho_\ell)^{2m}, b_{\ell-1}, m-1)$ instead of $(B_r(x_0), B_R(x_0), t_1, t_2, b, k)$ and using $(A)_\ell$ and $(C)_\ell$ we obtain (note that $\lambda\beta + b_{\ell-1}(1 - \frac{\lambda}{2}) = b_\ell$ and $0 < h \le \rho_\ell^{2m}$)

$$\begin{aligned}
&\int\limits_{8Q_\ell} |\tau_{-h}(\tau_h D^{m-1} u)|^2 \,\mathrm{d}z \\
&\quad \le \frac{c|h|^{\frac{1}{m}}}{\rho_\ell^{2m}} \int\limits_{16Q_\ell} \Big(|\tau_h D^m u|^2 + |\tau_{-h} D^m u|^2 + |h|^{\frac{1}{2m}(\lambda\beta + b_{\ell-1}(1-\frac{\lambda}{2}))}(1 + |D^m u|^{2+b_{\ell-1}})\Big) \,\mathrm{d}z \\
&\quad \le C(1/\rho_\ell, C_\ell, \widetilde{C}_\ell) \, |h|^{\frac{1}{m} + \frac{b_\ell}{2m}}.
\end{aligned}$$

Since $\tau_{-h}\tau_h f(t) = -\tau_h\tau_h f(t-h)$ and $h \le \rho_\ell^{2m}$ this implies the same estimate for $\tau_h(\tau_h D^{m-1} u)$ instead of $\tau_{-h}(\tau_h D^{m-1} u)$ on the smaller cylinder $4Q_\ell$. Applying Lemma 5.3 yields a similar estimate for the first differences, i.e. for all $0 < h \le \rho_\ell^{2m}$ we have

$$\int\limits_{4Q_\ell} |\tau_h(D^{m-1} u)|^2 \,\mathrm{d}z \le C(1/\rho_\ell, b_\ell, C_\ell, \widetilde{C}_\ell) \, |h|^{\frac{1}{m} + \frac{b_\ell}{2m}}. \tag{6.14}$$

Therefore, we can apply the Lemma 5.1 with $(\frac{2+b_\ell}{4m}, \frac{2+\vartheta b_\ell}{4m})$ instead of $(\gamma, \widetilde{\gamma})$ to infer that $D^{m-1} u \in W^{0, \frac{2+\vartheta b_\ell}{4m}; 2}(2Q_\ell; \mathbb{R}^{\mathcal{M}_{m-1}})$ for all $\vartheta \in (0, 1)$ and there holds $[D^{m-1} u]_{0, \frac{2+\vartheta b_\ell}{4m}, 2Q_\ell} \le C(1/\rho_\ell, b_\ell, \vartheta, C_\ell, \widetilde{C}_\ell)$. Moreover, from $(B)_\ell$ we know that $D^m u \in W^{\frac{\vartheta b_\ell}{2}, \frac{\vartheta b_\ell}{4m}; 2}(2Q_\ell; \mathbb{R}^{\mathcal{N}})$

for all $\vartheta \in (0,1)$. Finally, by assumption we have that $D^{m-1}u \in C^{\lambda,\frac{\lambda}{2m}}(\Omega_T;\mathbb{R}^{\mathscr{M}_{m-1}})$. Therefore we can apply the Interpolation-Theorem 5.7 for u replaced with $D^{m-1}u$ in the case $\gamma = \frac{\vartheta b_\ell}{2}$ and infer that $|D^m u| \in L^s(Q_\ell)$ for all $s < \frac{(n+2m)(2+\vartheta b_\ell)}{n+2m-\lambda\vartheta b_\ell}$. Now, we choose $\vartheta \in (0,1)$, such that $2+b_\ell < \frac{(n+2m)(2+\vartheta b_\ell)}{n+2m-\lambda\vartheta b_\ell}$. Note that we can take any $\frac{n+2m}{n+2m+(2+b_\ell)\lambda} < \vartheta < 1$. Hence, we can choose $\vartheta = \vartheta(n,m,\lambda) = \frac{n+2m}{n+2m+2\lambda}$. Then, we have $|D^m u| \in L^{2+b_\ell}(Q_\ell)$ and the bound in $(C)_{\ell+1}$ holds, i.e.

$$\int_{Q_\ell} |D^m u|^{2+b_\ell}\, dz \le C(1/\rho_\ell, b_\ell, C_\ell, \widetilde{C}_\ell, \widehat{C}_\ell) = \widetilde{C}_{\ell+1}. \tag{6.15}$$

Due to Lemma 6.5 this implies that $|\tau_h(D^k u)| \in L^{2+b_\ell}(32Q_{\ell+1})$ for $0 \le k \le m-1$ (note that $\frac{1}{2}Q_\ell = 32Q_{\ell+1}$) with the estimate

$$\begin{aligned}\int_{32Q_{\ell+1}} |\tau_h(D^k u)|^{2+b_\ell}\, dz &\le C(b_\ell)\,|h|^{\frac{(2+b_\ell)(m-k)}{2m}} \int_{Q_\ell} (1+|D^m u|)^{2+b_\ell}\, dz \\ &\le C(b_\ell, \widetilde{C}_{\ell+1})\,|h|^{\frac{(2+b_\ell)(m-k)}{2m}},\end{aligned} \tag{6.16}$$

where have also used (2.6) and (5.1) and the fact that $|h| \le \rho_{\ell+1}^{2m}$.

The previous considerations can now be exploited to infer better estimates for the finite differences of $D^m u$. We choose $16\rho_{\ell+1} \le r_1 < r_2 < r_3 \le 32\rho_{\ell+1}$ with $r_2 = \frac{1}{2}(r_1+r_3)$. Then $16Q_{\ell+1} \subset Q_{r_1} \subset Q_{r_2} \subset Q_{r_3} \subset 32Q_{\ell+1}$ by construction. Moreover, we choose cut-off functions $\eta \in C_0^\infty(B_{r_2})$ and $\zeta \in C_0^1(-r_2^{2m}, r_2^{2m})$ with $0 \le \eta, \zeta \le 1$, $\eta \equiv 1$ on B_{r_1}, $\zeta \equiv 1$ on $(-r_1^{2m}, r_1^{2m})$, $|D^k\eta| \le c(r_2-r_1)^{-k}$ for $0 \le k \le m$ and $|\zeta'| \le c(r_2-r_1)^{-2m}$. From Lemma 6.7 with $(B_{r_1}, B_{r_2}, -r_2^{2m}, r_2^{2m})$ instead of $(B_r(x_0), B_R(x_0), t_1, t_2)$ and the fact that $\zeta \equiv 1$ on $(-r_1^{2m}, r_1^{2m})$ we obtain

$$\begin{aligned}\int_{Q_{r_1}} |\tau_h D^m u|^2\, dz &\le C\,|h|^{\frac{\beta}{m}} \int_{Q_{r_2}} \big(1+|D^m u|\big)^2\, dz + C\int_{Q_{r_2}} \sum_{k=0}^{m-1} \frac{|\tau_h(D^k u)|^2}{(r_2-r_1)^{2(m-k)}}\, dz \\ &\quad + C\int_{Q_{r_2}} \big(1+|D^m u|\big)^2\, \widetilde{\theta}\,(|\tau_h(\delta u)|)^2\, dz \\ &\quad + C\left|\int_{Q_{r_2}} B(\cdot,\delta u, D^m u)\cdot \tau_{-h}\big(\delta(\eta^{2m}\tau_h u)\zeta^2\big)\, dz\right| \\ &= C\left(I_1^{(\ell+1)} + I_2^{(\ell+1)} + I_3^{(\ell+1)} + I_4^{(\ell+1)}\right),\end{aligned} \tag{6.17}$$

with the obvious labelling of $I_1^{(\ell+1)} - I_4^{(\ell+1)}$. Note that in last term we have used "integration by parts for finite differences", which is applicable since $\operatorname{spt}\zeta \subset (-r_2^{2m}, r_2^{2m})$ and $r_2^{2m} + |h| \le (32\rho_{\ell+1})^{2m} + \rho_{\ell+1}^{2m} \le (32^{2m}+1)\rho_{\ell+1}^{2m} \le \rho_\ell^{2m}$ and $r_2 \le 32\rho_{\ell+1} \le \rho_\ell$; therefore $\operatorname{spt}(B\cdot\tau_{-h}(\delta(\eta^{2m}\tau_h u)\zeta^2)) \subset Q_\ell)$. We now establish bounds for $I_2^{(\ell+1)} - I_4^{(\ell+1)}$.

Estimate for $I_2^{(\ell+1)}$: From the fact that $Q_{r_2} \subset 32Q_{\ell+1}$, $(r_2-r_1) \le 1$ and Corollary 6.3, which is applicable since $|h| \le \rho_{\ell+1}^{2m}$, we infer that

$$I_2^{(\ell+1)} \le \frac{2}{(r_2-r_1)^{2m}} \sum_{k=0}^{m-1} \int_{32Q_{\ell+1}} |\tau_h(D^k u)|^2\, dz \le C\,\frac{|h|^{\frac{1}{m}}}{(r_2-r_1)^{2m}} \int_{Q_\ell} \big(1+|D^m u|^2\big)\, dz.$$

Estimate for $I_3^{(\ell+1)}$. Using (1.9), the fact that $Q_{r_2} \subset 32Q_{\ell+1}$ and Hölder's inequality we obtain

$$I_3^{(\ell+1)} \le \Bigg(\int\limits_{32Q_{\ell+1}} \big(1+|D^m u|\big)^{2+b_\ell}\, \mathrm{d}z \Bigg)^{\frac{2}{2+b_\ell}} \Bigg(\int\limits_{32Q_{\ell+1}} |\tau_h(\delta u)|^{\frac{2\beta(2+b_\ell)}{b_\ell}}\, \mathrm{d}z \Bigg)^{\frac{b_\ell}{2+b_\ell}}.$$

For the second integral on the right-hand side we use the Hölder continuity of δu, i.e. the fact that $|\tau_h(\delta u)| \le |h|^{\frac{\lambda}{2m}} [\delta u]_{0,\frac{\lambda}{2m}}$ and (6.16) to infer that

$$\begin{aligned}
&\Bigg(\int\limits_{32Q_{\ell+1}} |\tau_h(\delta u)|^{\frac{2\beta(2+b_\ell)}{b_\ell}}\, \mathrm{d}z \Bigg)^{\frac{b_\ell}{2+b_\ell}} \\
&\quad \le C(b_\ell)\, |h|^{\frac{b_\ell}{2m}+\frac{\lambda(2\beta-b_\ell)}{2m}}\, [\delta u]_{0,\frac{\lambda}{2m}}^{2\beta-b_\ell} \Bigg(\int\limits_{Q_\ell} (1+|D^m u|)^{2+b_\ell}\, \mathrm{d}z \Bigg)^{\frac{b_\ell}{2+b_\ell}}.
\end{aligned}$$

Inserting this above and taking into account that $b_\ell + \lambda(2\beta - b_\ell) = \beta\lambda + b_\ell - \lambda(b_\ell - \beta) \ge b_{\ell+1}$ since $\beta \ge b_\ell/2$ we get

$$I_3^{(\ell+1)} \le C(b_\ell)\, |h|^{\frac{b_{\ell+1}}{2m}} \int\limits_{Q_\ell} (1+|D^m u|)^{2+b_\ell}\, \mathrm{d}z.$$

Estimate for $I_4^{(\ell+1)}$. For the last term we obtain with Young's inequality for $\varepsilon > 0$ that

$$I_4^{(\ell+1)} \le \varepsilon \int\limits_{Q_\ell} \big|\tau_{-h}\big(\delta(\eta^{2m}\tau_h u)\zeta^2\big)\big|^2\, \mathrm{d}z + \frac{L^2}{\varepsilon} \int\limits_{Q_\ell} \big(1+|D^m u|\big)^2\, \mathrm{d}z = \varepsilon\, J_1 + J_2, \qquad (6.18)$$

with the obvious meaning of J_1 and J_2. To estimate J_1 we set $w \equiv \delta(\eta^{2m}\tau_h u)$ and compute:

$$\big|\tau_{-h}[w(t)\zeta(t)^2]\big| \le \big|w(t-h)\, \tau_{-h}[\zeta(t)^2]\big| + \big|\tau_{-h}[w(t)]\, \zeta(t)^2\big|.$$

Moreover, we have $|\tau_{-h}[\zeta(t)^2]| = |\zeta(t-h) - \zeta(t)|(\zeta(t-h)+\zeta(t)) \le |h|\|\zeta'\|_\infty(\zeta(t-h)+\zeta(t))$ (note that $\zeta \ge 0$) and therefore, we infer that

$$\begin{aligned}
J_1 &\le |h|^2\, \|\zeta'\|_\infty^2 \int\limits_{Q_\ell} \big|\delta\big(\eta^{2m}\tau_h u(\cdot,t-h)\big)\big|^2 \big(\zeta(t-h)^2+\zeta(t)^2\big)\, \mathrm{d}z \\
&\quad + \int\limits_{Q_\ell} \big|\tau_{-h}\big(\delta(\eta^{2m}\tau_h u(\cdot,t))\big)\big|^2 \zeta(t)^2\, \mathrm{d}z \\
&= J_{1,1} + J_{1,2}.
\end{aligned}$$

First, we recall that $|\zeta'| \le c(r_2-r_1)^{-2m}$, $0 \le \zeta \le 1$, $|D^k\eta| \le c(r_2-r_1)^{-k} \le c(r_2-r_1)^{-m}$ for $0 \le k \le m$ and $\operatorname{spt}\zeta \subset (-r_2^{2m}, r_2^{2m})$. Then we apply Corollary 6.3 (note that $Q_{r_2} \subset 32Q_{\ell+1} = \frac{1}{2}Q_\ell$ and $|h| \le \rho_{\ell+1}^{2m}$ and $|h| \le 1$) to obtain

$$J_{1,1} \le \frac{C\,|h|^2}{(r_2-r_1)^{6m}} \sum_{k=0}^{m-1} \int\limits_{-r_2^{2m}-|h|}^{r_2^{2m}+|h|} \int\limits_{B_{r_2}} |D^k\tau_h u|^2\, \mathrm{d}z \le C\, \frac{|h|^2}{(r_2-r_1)^{6m}} \int\limits_{Q_\ell} \big(1+|D^m u|\big)^2\, \mathrm{d}z.$$

For $J_{1,2}$ we note that $\operatorname{spt}(\eta\zeta) \subset Q_{r_2}$ and $|D^k\eta| \le c(r_2-r_1)^{-k} \le c(r_2-r_1)^{-m}$ for $0 \le k \le m$. Moreover, we apply Lemma 6.6 with $(B_{r_2}, B_{r_3}, -r_2^{2m}, r_2^{2m})$ instead of $(B_r(x_0), B_R(x_0), t_1, t_2)$ (note that $|h| \le \rho_{\ell+1}^{2m}$ and $\lambda\beta + b_\ell(1-\frac{\lambda}{2}) = b_{\ell+1}$) to infer that

$$J_{1,2} \le \frac{C}{(r_2-r_1)^{2m}} \sum_{k=0}^{m-1} \int_{Q_{r_2}} |\tau_{-h}\tau_h(D^k u)|^2 \, \mathrm{d}z$$

$$\le \sum_{k=0}^{m-1} \frac{C\,|h|^{\frac{m-k}{m}}}{(r_3-r_2)^{4m}} \bigg(\int_{Q_{r_3}} |\tau_h D^m u|^2 + |\tau_{-h} D^m u|^2 \, \mathrm{d}z + |h|^{\frac{b_{\ell+1}}{2m}} \int_{Q_\ell} 1 + |D^m u|^{2+b_\ell} \, \mathrm{d}z \bigg).$$

Now, we join the previous estimates for $J_{1,1}$ and $J_{1,2}$ with (6.18) and note that $|h| \le 1$, $|h|^2 \le |h|^{\frac{b_{\ell+1}}{2m}}$, $2(r_2-r_1) = (r_3-r_1)$ and $|D^m u|^2 \le 1 + |D^m u|^{2+b_\ell}$. Finally, choosing $\varepsilon = (r_2-r_1)^{6m}/(4c|h|^{\frac{1}{m}})$ we obtain

$$I_4^{(\ell+1)} \le \frac{1}{4} \int_{Q_{r_3}} |\tau_h D^m u|^2 + |\tau_{-h} D^m u|^2 \, \mathrm{d}z + \frac{C\,|h|^{\frac{b_{\ell+1}}{2m}}}{(r_3-r_1)^{6m}} \int_{Q_\ell} 1 + |D^m u|^{2+b_\ell} \, \mathrm{d}z,$$

where we have also used that $|h|^{\frac{1}{m}} \le |h|^{\frac{b_{\ell+1}}{2m}}$, $(r_3-r_1) \le 1$ and $|D^m u|^2 \le 1 + |D^m u|^{2+b_\ell}$.

Combining the previous estimates for $I_2^{(\ell+1)} - I_4^{(\ell+1)}$ with (6.17) and noting once again that $|h| \le 1$, $\frac{b_{\ell+1}}{2m} \le \frac{\beta}{m}$, $(r_3-r_1) \le 1$, $r_3 - r_1 = 2(r_2-r_1)$ and $|D^m u|^2 \le 1 + |D^m u|^{2+b_\ell}$ we arrive at

$$\int_{Q_{r_1}} |\tau_h D^m u|^2 \, \mathrm{d}z \le \frac{1}{4} \int_{Q_{r_3}} |\tau_h D^m u|^2 + |\tau_{-h} D^m u|^2 \, \mathrm{d}z$$

$$+ C(b_\ell) \frac{|h|^{\frac{b_{\ell+1}}{2m}}}{(r_3-r_1)^{6m}} \int_{Q_\ell} 1 + |D^m u|^{2+b_\ell} \, \mathrm{d}z.$$

We also obtain the analogous estimate for $-h$ instead of h, and combining them yields:

$$\int_{Q_{r_1}} |\tau_h D^m u|^2 + |\tau_{-h} D^m u|^2 \, \mathrm{d}z \le \frac{1}{2} \int_{Q_{r_3}} |\tau_h D^m u|^2 + |\tau_{-h} D^m u|^2 \, \mathrm{d}z$$

$$+ \frac{C|h|^{\frac{b_{\ell+1}}{2m}}}{(r_3-r_1)^{6m}} \int_{Q_\ell} 1 + |D^m u|^{2+b_\ell} \, \mathrm{d}z,$$

where $C = C(b_\ell)$. Applying Lemma 2.3 and using $(C)_{\ell+1}$, i.e. (6.15) we infer that

$$\int_{16Q_{\ell+1}} |\tau_h D^m u|^2 + |\tau_{-h} D^m u|^2 \, \mathrm{d}z \le C(b_\ell) \frac{|h|^{\frac{b_{\ell+1}}{2m}}}{\rho_{\ell+1}^{6m}} \int_{Q_\ell} 1 + |D^m u|^{2+b_\ell} \, \mathrm{d}z$$

$$\le C(1/\rho_{\ell+1}, b_\ell)\, \widetilde{C}_{\ell+1}\, |h|^{\frac{b_{\ell+1}}{2m}}.$$

This proves the first bound in $(A)_{\ell+1}$.

A similar estimate for the finite differences in x-direction can be obtained, when we start with a similar estimate for the space-direction, instead of Lemmas 6.7. Then we estimate the right-hand side similar to the time direction. Once again, we can exploit the weak differentiability of $D^k u$, $0 \le k \le m-1$, i.e. the estimate (6.9) instead of Lemma 6.5 and 6.6 to infer also the second estimate in $(A)_{\ell+1}$. This finishes the proof of $(A)_{\ell+1}$. Finally, from Lemma 5.1 we now conclude that $(B)_{\ell+1}$ holds.

Since $b_\ell \nearrow 2\beta$, we find for each $\gamma < \beta$ an integer $\bar\ell = \bar\ell(\beta,\gamma,\lambda) \in \mathbb{N}$, such that $2\gamma < b_{\bar\ell} < 2\beta$. Iterating up to the step $\bar\ell$ we infer from $(B)_{\bar\ell}$ that $D^m u \in W^{\gamma,\frac{\gamma}{2m};2}(Q_{R/64^{\bar\ell}};\mathbb{R}^{\mathscr{N}})$ (note that $8\rho_{\bar\ell} \ge \rho_{\bar\ell} = R/64^{\bar\ell}$). Moreover from (6.14) we know that $D^{m-1}u \in W^{1,\frac{1+\gamma}{2m};2}(Q_{R/64^{\bar\ell}};\mathbb{R}^{\mathscr{M}_{m-1}})$. This shows the assertion (6.13).

Now let $\widetilde{Q} = \widetilde{\Omega} \times (t_1,t_2) \Subset \Omega_T$. We choose $R = \frac{1}{2}\min\{\operatorname{dist}(\widetilde{\Omega},\partial\Omega), T+\widetilde{t}_1, |t_2|\}$ and cover $\widetilde{Q}$ by finitely many cylinders $Q_{R/64^{\bar\ell}}(z_i)$, $i=1,\dots,M$, with center $z_i \in \widetilde{Q}$. Using (6.13) on each cylinder $Q_{R/64^{\bar\ell}}(z_i)$ and summing over $i=1,\dots,M$ we finally obtain the assertion of Lemma 6.9. □

6.3 Proof of the results

Proof of Theorem 1.2 At first we will show that for $k=0,\dots,m$ there holds

$$D^k u \in W^{\gamma,\frac{\gamma}{2m};2}_{loc}(\Omega_T,\mathbb{R}^{\mathscr{M}_k}) \qquad \text{for all } \gamma < \delta/2, \tag{6.19}$$

where $\delta = \delta(n,m,\beta,N,L/\nu)$ is specified in Lemma 6.8. For $k=m$ this is exactly the conclusion of those Lemmas. Furthermore for $k=0,\dots,m-1$ we know from Remark 6.4 that $D^k u \in W^{m-k,\frac{\gamma(m-k)}{2m};2}_{loc}(\Omega_T,\mathbb{R}^{\mathscr{M}_k})$ for all $\gamma \in (0,1)$. Therefore we conclude that (6.19) is true.

Applying Proposition 5.4 yields that $\dim_{\mathscr{P}}(\Sigma_1) \le n+2m-\delta$ and $\dim_{\mathscr{P}}(\Sigma_2) \le n+2m-\delta$, where Σ_1 and Σ_2 are defined in Theorems 3.15 and 3.7. From Theorem 3.15 we know that the singular set Σ of $D^m u$ is contained in $\Sigma_1 \cup \Sigma_2$ and therefore $\dim_{\mathscr{P}}(\Sigma) \le \dim_{\mathscr{P}}(\Sigma_1 \cup \Sigma_2) \le n+2m-\delta$, which shows the assertion of Theorem 1.2. □

Remark 6.10 Under the conditions of Theorems 1.4 or 1.5 we can use the better fractional differentiability properties from Lemmas 6.8, respectively 6.9 to infer the better estimate $\dim_{\mathscr{P}}(\Sigma) \le n+2m-2\beta$.

We shall see in the following that this estimate can be improved slightly in the sense that the inequality is strict. This will be achieved by exploiting the differentiability of the coefficients A with respect to the last variable in such a way that we can rewrite the system as a linear system for the finite differences $\tau_h u$ respectively $\tau_h^s u$. From this linear system we obtain the higher integrability of the t- derivative of $D^m u$ of fractional order $\frac{\beta}{2m}$ and also of the x-derivative of $D^m u$ of fractional order β (see Lemmas 6.11 and 6.12). This leads us to better estimates for the fractional difference quotients of $D^m u$. Therefore we can conclude that $D^m u$ lies in some slightly better fractional Sobolev space, which will lead us directly to the dimension reduction statements of Theorems 1.4 and 1.5.

Lemma 6.11 *Let* $u \in L^2(-T,0;W^{m,2}(\Omega;\mathbb{R}^N)) \cap C^{m-1,\lambda,\frac{\lambda}{2m}}(\Omega_T;\mathbb{R}^N)$ *with* $\lambda \in (0,1)$ *be a weak solution of* (1.1) *in* Ω_T *under the assumptions* (1.2), (1.3), (1.7) *and* (1.8) *with* $B \equiv 0$. *Then there exists* $\xi = \xi(n,m,N,L/\nu) > 1$, *such that for all* $\widetilde{\Omega} \Subset \Omega$, $(t_1,t_2) \Subset (-T,0)$ *and* *for all $\gamma < \beta$ there holds*

$$\int_{\widetilde{\Omega}} \int_{t_1}^{t_2} \int_{t_1}^{t_2} \frac{|D^m u(x,t) - D^m u(x,\tau)|^{2\xi}}{|t-\tau|^{1+\frac{\xi\gamma}{m}}} \, dt \, d\tau \, dx + \int_{t_1}^{t_2} \int_{\widetilde{\Omega}} \int_{\widetilde{\Omega}} \frac{|D^m u(x,t) - D^m u(y,t)|^{2\xi}}{|x-y|^{n+2\xi\gamma}} \, dx \, dy \, dt < \infty.$$

Proof In the following, with C we shall denote a constant depending on $n, m, N, \nu, L, \beta, \lambda$, $\|u\|_{L^2(-T,0;W^{m,2}(\Omega))}$ and $[\delta u]_{\lambda,\frac{\lambda}{2m}}$. We will only indicate the additional dependencies of C.

We fix parameters γ and θ such that

$$0 < \gamma < \beta \qquad \text{and } 2 + 2\beta < \theta < q \equiv \frac{2(n+2m)(1+\gamma)}{n+2m-2\lambda\gamma}.$$

From Lemma 6.9 we already know that $D^m u \in W^{\gamma,\frac{\gamma}{2m};2}_{loc}(\Omega_T;\mathbb{R}^{\mathscr{N}})$ and $D^{m-1}u \in W^{\frac{1+\gamma}{2m};2}_{loc}(-T,0;L^2(\Omega;\mathbb{R}^{\mathscr{M}_{m-1}}))$. Moreover, by assumption we have that $D^{m-1}u \in C^{\lambda,\frac{\lambda}{2m}}(\Omega_T;\mathbb{R}^{\mathscr{M}_{m-1}})$. Therefore we can apply Lemma 5.7 to the function $D^{m-1}u$ and find that $|D^m u| \in L^s_{loc}(\Omega_T)$ for all $1 \le s < q$. In particular we have $|D^m u| \in L^\theta_{loc}(\Omega_T)$. Since the result we are going to prove is of local nature, we can suppose without loss of generality that $|D^m u| \in L^\theta(\Omega_T)$ and

$$\int_{\Omega_T} |D^m u|^\theta \, dz \le C(\gamma,\theta) < \infty. \tag{6.20}$$

We will start with the proof for the time-direction. Let $\widetilde{\Omega} \Subset \Omega$, $(t_1,t_2) \Subset (-T,0)$ and $0 < |h| < \frac{1}{2}\min\{|t_2|, T+t_1, 1\}$. We set $\widetilde{Q} \equiv \widetilde{\Omega} \times (t_1,t_2)$. Replacing φ with $\tau_{-h}\varphi$ in (1.1) we obtain after "integration by parts" for finite differences

$$\int_{\Omega_T} \big(\tau_h(u)\cdot\varphi_t - \tau_h[A(\cdot,\delta u, D^m u)]\cdot D^m\varphi\big)\, dz = 0 \qquad \text{for all } \varphi \in C_0^\infty(\widetilde{Q};\mathbb{R}^N). \tag{6.21}$$

With the notation introduced in (6.1) and (6.2) for $\widetilde{\mathscr{A}}$, $\mathscr{B}$ and $\mathscr{C}$, we rewrite $\tau_h(A) = \widetilde{\mathscr{A}}(h)\tau_h(D^m u) + \mathscr{B}(h) + \mathscr{C}(h)$ and define

$$v_h \equiv \frac{\tau_h(u)}{|h|^{\frac{\beta}{2m}}}, \qquad \widetilde{\mathscr{B}}(h) \equiv -\frac{\mathscr{B}(h)}{|h|^{\frac{\beta}{2m}}}, \qquad \widetilde{\mathscr{C}}(h) \equiv -\frac{\mathscr{C}(h)}{|h|^{\frac{\beta}{2m}}}.$$

Dividing (6.21) by $|h|^{\frac{\beta}{2m}}$, we see that v_h is a weak solution of the linear system

$$\int_{\Omega_T} \big(v_h\cdot\varphi_t - \widetilde{\mathscr{A}}(h)D^m v_h\cdot D^m\varphi\big)\, dz = \int_{\Omega_T} \big(\widetilde{\mathscr{B}}(h) + \widetilde{\mathscr{C}}(h)\big)\cdot D^m\varphi \, dz \tag{6.22}$$

for all $\varphi \in C_0^\infty(\widetilde{Q};\mathbb{R}^N)$. Due to (6.3), the coefficients satisfy $\widetilde{\mathscr{A}}(h)\, p\cdot p \ge \nu\,|p|^2$, $|\widetilde{\mathscr{A}}(h)| \le L$ and

$$|\widetilde{\mathscr{B}}(h)| \le L\,|h|^{-\frac{\beta}{2m}}\,|\tau_h\delta u|^\beta\,\big(1+|D^m u|\big), \qquad |\widetilde{\mathscr{C}}(h)| \le L\,\big(1+|D^m u|\big). \tag{6.23}$$

In the following we will apply Theorem 4.1 in order to show that $|D^m v_h|$ is higher integrable. For that purpose we firstly have to ensure that $\widetilde{\mathscr{B}}$ and $\widetilde{\mathscr{C}}$ fulfill the required integrability assumptions, i.e. that they are integrable with an exponent larger then 2. Indeed, by (6.20) we know that $\widetilde{\mathscr{C}} \in L^{\theta}(\Omega_T)$. Moreover, we set $1+\delta_1 \equiv \frac{\theta}{2+2\beta} > 1$ and show that $\widetilde{\mathscr{B}} \in L^{2\xi}$ for all $1 \le \xi \le 1+\delta_1$. Let $Q_\rho \equiv Q_\rho(z_0)$ be a parabolic cylinder with $Q_{2\rho} \Subset \Omega_T$. To estimate the following integral we first use (6.23) and Hölder's inequality. Since $\theta \ge 2\xi(1+\beta)$ we conclude from (6.20) that $|D^m u| \in L^{2\xi(1+\beta)}(\Omega_T)$. Therefore we can apply Lemma 6.5 with $b = 2\xi\beta$ to estimate the first integral appearing on the right-hand side. Hence, we obtain for $1 \le \xi \le 1+\delta_1$ that

$$\begin{aligned} \fint_{Q_\rho} |\widetilde{\mathscr{B}}(h)|^{2\xi}\, dz &\le L^{2\xi}\, |h|^{-\frac{\beta\xi}{m}} \left(\fint_{Q_\rho} |\tau_h(\delta u)|^{2\xi(1+\beta)}\, dz \right)^{\frac{\beta}{1+\beta}} \\ &\qquad \times \left(\fint_{Q_\rho} (1+|D^m u|)^{2\xi(1+\beta)}\, dz \right)^{\frac{1}{1+\beta}} \\ &\le C \fint_{Q_{2\rho}} \left(1+|D^m u| + \left[|D^m u|\right]_h + |\tau_h D^m u|\right)^{2\xi(1+\beta)}\, dz. \qquad (6.24) \end{aligned}$$

We note that, here and in the following it is important that all appearing constants are independent of h.

Higher integrability of $D^m v_h$. From the previous discussion we know that $\widetilde{\mathscr{B}}, \widetilde{\mathscr{C}} \in L^{2(1+\delta_1)}$. Therefore we can apply Theorem 4.1 to infer that there exists $\xi = \xi(n, m, N, L/\nu)$ with $1 < \xi < 1+\delta_1$, such that $|D^m v_h| \in L^{2\xi}_{loc}(\widetilde{Q})$ and for all parabolic cylinders $Q_{4\rho} \equiv Q_{4\rho}(z_0) \Subset \widetilde{Q}$ and $0 < \rho \le 1$ there holds

$$\fint_{Q_{\rho/4}} |D^m v_h|^{2\xi}\, dz \le C \left(\fint_{Q_\rho} |D^m v_h|^2\, dz \right)^{\xi} + C \fint_{Q_\rho} \left(|\widetilde{\mathscr{B}}|^{2\xi} + |\widetilde{\mathscr{C}}|^{2\xi} + 1\right) dz, \qquad (6.25)$$

Note that we have used (6.24) in the last line.

In order to estimate the first integral on the right-hand side we choose two cut-off functions $\eta \in C_0^\infty(B_{2\rho})$ and $\zeta \in C_0^1(\mathbb{R})$ with $0 \le \eta \le 1$, $0 \le \zeta \le 1$, $\eta \equiv 1$ on B_ρ, $|D^k \eta| \le c\,\rho^{-k}$ for $0 \le k \le m$ and $\zeta \equiv 0$ on $(-\infty, t_0 - (2\rho)^{2m})$, $\zeta \equiv 1$ on $(t_0 - \rho^{2m}, t_0)$ and $0 \le \zeta' \le c\,\rho^{-2m}$. We use the estimate from Lemma 6.7 with this particular choice of cut-off functions, recall the definition $v_h = \tau_h u/|h|^{\beta/(2m)}$ and the fact that $\widetilde{\theta}(s) \le s^\beta$ to infer that

$$\begin{aligned} \fint_{Q_\rho} |D^m v_h|^2\, dz &\le C \fint_{Q_{2\rho}} \left(1+|D^m u|\right)^2 dz + C\, |h|^{-\frac{\beta}{m}} \sum_{k=0}^{m-1} \fint_{Q_{2\rho}} \frac{|\tau_h(D^k u)|^2}{\rho^{2(m-k)}}\, dz \\ &\quad + C\, |h|^{-\frac{\beta}{m}} \fint_{Q_{2\rho}} \left(1+|D^m u|\right)^2 |\tau_h(\delta u)|^{2\beta}\, dz \\ &= C\,(I_1 + I_2 + I_3), \qquad (6.26) \end{aligned}$$

with the obvious meaning of $I_1 - I_3$. In the following we will derive bounds for I_2 and I_3.

Estimate for I_2. Applying Corollary 6.2 (note that $0 < |h| < \frac{1}{2}\min\{|\tau_2|, T+\tau_1, 1\}$) and taking into account that $|h| \le 1$, we obtain

$$I_2 \le C\,\rho^{-2(m-k)} \fint_{Q_{4\rho}} \left(1 + \left[|D^m u|\right]_h + |\tau_h D^m u|\right)^2 dz.$$

Estimate for I_3**.** Here, we proceed completely similar to our estimate for $\widetilde{\mathscr{B}}$, i.e. (6.24) for ξ replaced with 1. This yields that

$$I_3 \le C \int_{Q_{4\rho}} \left(1 + |D^m u| + \left[|D^m u|\right]_h + |\tau_h D^m u|\right)^{2(1+\beta)} \, dz.$$

Combining the estimates for I_2 and I_3 with (6.26), we infer an estimate for the first integral in (6.25). For the second integral we use (6.24) and (6.23). Applying also Hölder's inequality, we finally arrive at

$$\begin{aligned} \int_{Q_{\rho/4}} |D^m v_h|^{2\xi} \, dz &\le C \int_{Q_{4\rho}} \left(1 + |D^m u| + \left[|D^m u|\right]_h + |\tau_h D^m u|\right)^{2\xi(1+\beta)} \, dz \\ &\le |\Omega_T| + C(\gamma, \theta, \rho), \end{aligned}$$

where we have used (2.6), (5.1), the fact that $0 < |h| < \frac{1}{2}\min\{|t_2|, T + t_1, 1\}$ and (6.20) in the last line.

Since the choice of the cylinder $Q_{4\rho} \Subset \widetilde{Q}$ was arbitrary, we can cover any open subset $\mathscr{O} \Subset \widetilde{Q}$ by parabolic cylinders $(Q_{\rho/4}(z_i))$, $i = 1, \dots, M$ with $Q_{4\rho}(z_i) \Subset \widetilde{Q}$ and $\rho = \frac{1}{8}\min\{\mathrm{dist}(\mathscr{O}, \widetilde{Q}), 1\}$. Summing over $i = 1, \dots, M$ yields a bound of the considered integral over the whole set $\mathscr{O}$, i.e. for $\int_{\mathscr{O}} |D^m v_h|^{2\xi} \, dz$. Recalling the definition $v_h = \tau_h u/|h|^{\beta/(2m)}$ we therefore find that

$$\int_{\widetilde{Q}} |\tau_h(D^m u)|^{2\xi} \, dz = |h|^{\frac{\xi\beta}{m}} \int_{\widetilde{Q}} |D^m v_h|^{2\xi} \, dz \le C(\gamma, \theta, \mathrm{dist}(\mathscr{O}, \widetilde{Q})) \, |h|^{\frac{\xi\beta}{m}},$$

for all h with $0 < |h| < \frac{1}{2}\min\{|t_2|, T + t_1, 1\}$. Applying Lemma 5.1 and noting that $\widetilde{Q} \Subset \Omega_T$ was arbitrary, we finally conclude the asserted fractional differentiability in time direction.

The proof for the space-direction is very much similar to the one for the time-direction from above and we will not accomplish the details. We can also write our parabolic system as a linear system for $v_h = \tau_h^s(u)/|h|^\beta$. Then, applying once again Theorem 4.1 we can show higher integrability of $|D^m v_h|$. The resulting terms are estimated essentially similar as before. But now, we can use the weak differentiability of $D^k u$ for $0 \le k \le m-1$ [see (6.6)] to estimate integrals involving $\tau_h^s(D^k u)$ [instead of Lemma 6.5 in (6.24)]. Proceeding this way, we also infer the second assertion of the lemma. □

Lemma 6.12 *The conclusion of Lemma 6.11 also holds for weak solutions* $u \in L^2(-T, 0; W^{m,2}(\Omega; \mathbb{R}^N))$ *of the simpler system* (1.10) *under the assumptions* (1.2)*,* (1.7) *and* (1.8).

Proof We will only outline this proof, since it is similar to the one of Lemma 6.11. Following the proof of Lemma 6.11, the arguments leading us to (6.22) now yield

$$\int_{\Omega_T} \left(v_h \cdot \varphi_t - \widetilde{\mathscr{A}}(h) D^m v_h \cdot D^m \varphi\right) \, dz = -\int_{\Omega_T} \widetilde{\mathscr{C}}(h) \cdot D^m \varphi \, dz \qquad \text{for all } \varphi \in C_0^\infty(\widetilde{Q}; \mathbb{R}^N),$$

where $v_h \equiv \tau_h(u)/|h|^{\frac{\beta}{2m}}$ and $\widetilde{\mathscr{A}}(h)$ respectively $\widetilde{\mathscr{C}}(h) = \mathscr{C}(h)/|h|^{\frac{\beta}{2m}}$ are defined in (6.2) respectively (6.1). Due to the higher integrability of $D^m u$ from Theorem 4.1 there exists $\sigma = \sigma(n, m, N, L/\nu) > 0$, such that $|D^m u| \in L^{2(1+\sigma)}_{loc}(\Omega_T)$. Since $|\widetilde{\mathscr{C}}(h)| \le L(1 + |D^m u|)$ we infer that $|\widetilde{\mathscr{C}}(h)| \in L^{2(1+\sigma)}_{loc}(\Omega_T)$. Therefore Theorem 4.1 ensures the existence of $\xi = \xi(n,$

$m, N, L/\nu)$ with $0 < \xi < 1 + \sigma$, such that for all parabolic cylinders $Q_{4\rho}(z_0) \Subset \Omega_T$ there holds

$$\fint_{Q_{\rho/4}(z_0)} |D^m v_h|^{2\xi} \,\mathrm{d}z \le c \left(\fint_{Q_\rho(z_0)} \left|D^m v_h\right|^2 \,\mathrm{d}z \right)^{\xi} + c \fint_{Q_\rho(z_0)} \left(1 + |D^m u|^{2\xi}\right) \mathrm{d}z.$$

Using the estimate (6.12) from the proof of Lemma 6.8, (with $c_2 = 0$) we can bound the first integral on the right-hand side by a constant independent of h. Whence, the second integral can be bounded, using the higher integrability of $|D^m u|$, we mentioned above. In conclusion, we can bound the left-hand side independently of h and recalling the definition $v_h = \tau_h u/|h|^\beta$ we therefore obtain

$$\fint_{Q_{\rho/4}(z_0)} |\tau_h(D^m u)|^{2\xi} \,\mathrm{d}z = |h|^{\frac{2\xi}{m}} \fint_{Q_{\rho/4}(z_0)} \left|D^m v_h\right|^{2\xi} \,\mathrm{d}z \le |h|^{\frac{2\xi}{m}}\, c,$$

where $c = c(n, m, N, \nu, L, \beta, R, T + \widetilde{t}_0, \|u\|_{L^2(-T,0;W^{m,2}(B_{4R}(x_0))})$. Since the choice of the cylinder $Q_{4\rho}(z_0) \Subset \Omega_T$ was arbitrary, we obtain the desired result with a covering argument. □

Proof of Theorems 1.4 and 1.5 Due to Remark 6.10 we already know that the first assertion of Theorem 1.4, respectively 1.5 holds, i.e. that $\dim_{\mathscr{P}}(\Sigma) \le n + 2m - 2\beta$. We can apply Lemmas 6.12, respectively 6.11 to ensure that there exists $\xi = \xi(n, m, N, L/\nu) > 1$ such that

$$D^m u \in W^{\gamma, \frac{\gamma}{2m}; 2\xi}_{loc}(\Omega_T, \mathbb{R}^{\mathscr{N}}) \qquad \text{for all } \gamma < \beta.$$

Hence, we infer from Lemma 5.4 that $\dim_{\mathscr{P}}(\Sigma_1) \le n + 2m - 2\beta - \delta$ and $\dim_{\mathscr{P}}(\Sigma_2) \le n + 2m - 2\beta - \delta$, where $\delta = 2\beta\sigma$ in the case of Theorem 1.4, respectively $\delta = 2\beta(\xi - 1)$ in the case of Theorem 1.5. From Theorem 3.15 we know that the singular set Σ of $D^m u$ is contained in $\Sigma_1 \cup \Sigma_2$ and hence $\dim_{\mathscr{P}}(\Sigma) \le \dim_{\mathscr{P}}(\Sigma_1 \cup \Sigma_2) \le n + 2m - 2\beta - \delta$, which shows the assertion. □

References

1. Acerbi, E., Mingione, G.: Regularity results for a class of functionals with non-standard growth. Arch. Rat. Mech. Anal. **156**, 121–140 (2001)
2. Acerbi, E., Mingione, G.: Gradient estimates for a class of parabolic systems. Duke Math. J. **136**, 285–320 (2007)
3. Acerbi, E., Mingione, G., Seregin, G.A.: Regularity results for parabolic systems related to a class of non-newtonian fluids. Ann. Inst. Henri Poincaré Anal. Non Linéaire **21**(1), 25–60 (2004)
4. Adams, R.A.: Sobolev Spaces. Academic Press, New York (1978)
5. Bögelein, V.: Regularity results for weak and very weak solutions of higher order parabolic systems. Ph.D.Thesis (2007)
6. Bojarski, B., Iwaniec, T.: Analytical foundations of the theory of quasiconformal mappings in $\mathbb{R}^n$. Ann. Acad. Sci. Fenn. Ser. A I **8**, 257–324 (1983)
7. Campanato, S.: Equazioni paraboliche del secondo ordine e spazi $\mathscr{L}^{2,\theta}(\Omega, \delta)$. Ann. Mat. Pura Appl. IV Ser. **73**, 55–102 (1966)
8. Campanato, S.: Maggiorazioni interpolatorie negli spazi $H^{m,p}_\lambda(\Omega)$. Ann. Mat. Pura Appl. **75**, 261–276 (1967)
9. Campanato, S.: Differentiability of the solutions of nonlinear elliptic systems with natural growth. Ann. Mat. Pura Appl. **131**(4), 75–106 (1982)
10. Campanato, S.: On the nonlinear parabolic systems in divergence form. Hölder Continuity and Partial Hölder Continuity of the Solutions. Ann. Mat. Pura Appl. **137**(4), 83–122 (1984)

11. Campanato, S., Cannarsa, P.: Differentiability and partial Hölder continuity of the solutions of non-linear elliptic systems of order $2m$ with quadratic growth. Ann. Scuola Norm. Sup. Pisa **8**, 285–309 (1980)
12. Da Prato, G.: Spazi $\mathscr{L}^{(p,\vartheta)}(\Omega, \delta)$ e loro proprieta. Ann. Mat. Pura Appl. IV. Ser. **69**, 383–392 (1965)
13. De Giorgi, E.: Frontiere orientate di misura minima. Sem. Scuola Normale Superiore Pisa (1960–1961)
14. Domokos, A.: Differentiability of solutions for the non-degenerate p-laplacian in the heisenberg group. J. Differ. Equ. **204**, 439–470 (2004)
15. Duzaar, F., Gastel, A., Grotowski, J.F.: Optimal partial regularity for nonlinear elliptic systems of higher order. J. Math. Sci., Tokyo **8**(3), 463–499 (2001)
16. Duzaar, F., Grotowski, J.F.: Optimal interior partial regularity for nonlinear elliptic systems: the method of a-harmonic approximation. Manuscr. Math. **103**, 267–298 (2000)
17. Duzaar, F., Mingione, G.: The p-harmonic approximation and the regularity of p-harmonic maps. Calc. Var. Part. Differ. Equ. **20**, 235–256 (2004)
18. Duzaar, F., Mingione, G.: Regularity for degenerate elliptic problems via p-harmonic approximation. Ann. Inst. Henri Poincaré Anal. Non Linéaire **21**, 735–766 (2004)
19. Duzaar, F., Mingione, G.: Second order parabolic systems, optimal regularity, and singular sets of solutions. Ann. Inst. Henri Poincaré Anal. Non Linéaire **22**, 705–751 (2005)
20. Duzaar, F., Mingione, G., Steffen, K.: Second order parabolic systems with p-growth and regularity (to appear, 2008)
21. Fefferman, C., Stein, E.M.: H^p spaces of several variables. Acta Math. **129**, 137–193 (1972)
22. Föglein, A.: Regularität von Lösungen gewisser Systeme elliptischer partieller Differentialgleichungen in der Heisenbarg-Gruppe. Diplomarbeit (2005)
23. Giaquinta, M.: Multiple Integrals in the Calculus of Variations and Nonlinear Elliptic Systems. Princeton University Press, Princeton (1983)
24. Giaquinta, M.: Introduction to regularity theory for nonlinear elliptic systems (1993)
25. Giaquinta, M., Struwe, M.: On the partial regularity of weak solutions of nonlinear parabolic systems. Math. Z. **179**, 437–451 (1982)
26. Giusti, E.: Direct Methods in the Calculus f Variations. World Scientific Publishing Company, Tuck Link, Singapore (2003)
27. John, J., Nirenberg, L.: On functions of bounded mean oscillation. Commun. Pure Appl. Math. **14**, 415–426 (1961)
28. Kinnunen, J., Lewis, J.L.: Higher integrability for parabolic systems of p-laplacian type. Duke Math. J. **102**, 253–271 (2000)
29. Ladyženskaja, O.A., Solonnikov, V.A., Ural'ceva, N.N.: Linear and Quasilinear Equations of Parabolic Type, vol. 23. American Mathematical Society, Providence (1968)
30. Mingione, G.: Bounds for the singular set of solutions to non linear elliptic systems. Calc. Var. Partial Differ. Equ. **18**(4), 373–400 (2003)
31. Mingione, G.: The singular set of solutions to non-differentiable elliptic systems. Arch. Ration. Mech. Anal. **166**, 287–301 (2003)
32. Nirenberg, L.: On elliptic partial differential equations. Ann. Sc. Norm. Super., Pisa III. Ser. **123**, 115–162 (1960)
33. Simon, J.: Compact sets in the space $L^p(0, T; B)$. Ann. Mat. Pura Appl. IV. Ser. **146**, 65–96 (1987)
34. Simon, L.: Lectures on Geometric Measure Theory. Proc. Centre Math. Anal., Austr. Nat. Univ., Canberra (1983)
35. Simon, L.: Theorems on Regularity and Singularity of Energy Minimizing Maps. Lectures in Math., ETH Zrich, Birkhuser, Basel (1996)
36. Stredulinsky, E.W.: Higher integrability from reverse Hölder inequalities. Indiana Univ. Math. J. **29**, 407–413 (1980)