Rev. Mat. Iberoamericana **27** (2011), no. 3, 751–801

# Partial regularity for subquadratic parabolic systems by $\mathcal{A}$-caloric approximation

**Christoph Scheven**

**Abstract**

We establish a partial regularity result for weak solutions of nonsingular parabolic systems with subquadratic growth of the type

$$\partial_t u - \operatorname{div}\, a(x,t,u,Du) = B(x,t,u,Du),$$

where the structure function $a$ satisfies ellipticity and growth conditions with growth rate $\frac{2n}{n+2} < p < 2$. We prove Hölder continuity of the spatial gradient of solutions away from a negligible set. The proof is based on a variant of a harmonic type approximation lemma adapted to parabolic systems with subquadratic growth.

## 1. Introduction and statement of the result

Throughout this paper, we assume that $\Omega \subset \mathbb{R}^n$ is a bounded smooth domain of dimension $n \geq 2$ and we fix $T > 0$ and $N \geq 1$. We consider weak solutions $u : \Omega \times (-T, 0) \to \mathbb{R}^N$ of nonsingular parabolic systems of the type

$$\partial_t u(z) - \operatorname{div}\, a(z,u,Du) = B(z,u,Du) \qquad \text{for } z \in \Omega \times (-T,0), \tag{1.1}$$

where the structure function $a$ satisfies some standard ellipticity and growth conditions with polynomial growth rate $p \in (\frac{2n}{n+2}, 2)$. For the inhomogeneity $B$, we consider either a controllable growth condition or the natural growth condition under an additional smallness assumption on the solution. For the precise statement of the assumptions, we refer to Section 2.

*2000 Mathematics Subject Classification*: 35K40, 35B65.
*Keywords*: Parabolic systems, partial regularity, harmonic approximation, singular set, subquadratic growth.

In general, solutions of parabolic systems (1.1) can not be expected to be regular everywhere on the domain. Even in the elliptic case, various examples of solutions with singularities are known [15, 27, 43], see also [42, 40, 39] for some examples in the parabolic case. Everywhere regularity can only be expected for systems with special structure such as the evolutionary $p$-Laplacian system

$$\partial_t u - \operatorname{div}(|Du|^{p-2}Du) = 0,$$

for which the regularity problem was settled by the fundamental contributions of DiBenedetto and Friedman [12, 13, 11], see [32] for some further generalizations. Other situations in which an everywhere regularity result is available are the case of a low-dimensional domain [35, 28, 34] or the case of bounded solutions [41, 25]. In the general case, however, one can only expect partial regularity results, that is regularity away from a singular set that is in some sense small. Results of this type have been established for quasilinear systems with $a(z, u, Du) = \tilde{a}(z, u)Du$, cf. [26], or for systems of $p$-Laplacian type structure [33]. The partial regularity for general parabolic systems of the type (1.1) was a longstanding open problem until it was solved by Duzaar and Mingione [20] for parabolic systems with quadratic growth and by Duzaar, Mingione and Steffen [22] for systems with polynomial growth with growth rate $p > 2$, cf. also [17, 4, 5] for results on boundary regularity. Their proofs are based on an adaptation of the harmonic approximation method to the parabolic setting. The harmonic approximation lemma in its basic form can be found in the book of Simon [38] and goes back to ideas of de Giorgi developed for the regularity theory of minimal surfaces [14]. Until now, the technique of harmonic approximation has been developed further and adapted to various settings in the regularity theory, cf. [21] for a survey on the numerous applications of harmonic type approximation lemmas. One of the advantages of the harmonic approximation method is that it avoids the application of reverse Hölder inequalities, which are not available in the nonquadratic parabolic case due to the anisotropic scaling of the system (1.1) in the case $p \neq 2$, see [30] for related problems.

All results on partial regularity stated above are concerned with nondegenerate parabolic systems. The degenerate case, as treated in the elliptic setting in [18, 19, 21], contains serious difficulties that are due to the appearance of degenerate parabolic cylinders. First results in this direction were achieved in [6]. In the subquadratic case $\frac{2n}{n+2} < p < 2$ however, the question of partial regularity remained open so far even for nonsingular parabolic systems. This case has been treated only in the elliptic setting, cf. [10, 8, 9, 16, 2]. In the present paper, we fill this gap in the theory and prove partial regularity for nonsingular parabolic systems with subquadratic growth.

Our main result is the following.

**Theorem 1.1.** *Let* $u \in C^0(-T,0;L^2(\Omega,\mathbb{R}^N)) \cap L^p(-T,0;W^{1,p}(\Omega,\mathbb{R}^N))$ *be a weak solution of the parabolic system* (1.1)*, where the structure function satisfies the conditions* (A1) *to* (A5) *stated in Section* 2*, and furthermore one of the following conditions is satisfied.*

(*i*) (Controllable growth) *The inhomogeneity* $B$ *satisfies condition* (B1)*, or*

(*ii*) (Natural growth) *the inhomogeneity satisfies condition* (B2)*, and the solution is bounded by* $\sup_{\Omega_T}|u| \le M_u$*, where* $M_u$ *is small enough to ensure* $2M_u\,\Lambda_1(M_u) < \nu$.

*Then there is an open subset* $\Omega^u \subset \Omega_T$ *with*

$$Du \in C^{\beta,\beta/2}_{\mathrm{loc}}(\Omega^u,\mathbb{R}^{Nn}) \qquad \text{and} \qquad |\Omega_T \setminus \Omega^u| = 0.$$

*Moreover, the solution satisfies* $u \in C^{\alpha,\alpha/2}_{\mathrm{loc}}(\Omega^u,\mathbb{R}^N)$ *for every* $\alpha \in (0,1)$.

We point out that our proof requires only minimal assumptions on the derivative $D_\xi a$ of the structure function, see (A4) and (A5). Moreover, the Hölder exponent $\beta$ of the gradient of the solution is the optimal one that is prescribed by the Hölder exponent of the structure function. In general, no higher regularity of the solution can be expected, as demonstrated by a counterexample in [2] in the elliptic setting.

Moreover, we have the following characterization of the singular set. We use the notation $Q_\rho(z_0)$ for parabolic cylinders as introduced in Section 3 below.

**Proposition 1.2.** *In the situation of the preceding theorem, the singular set satisfies moreover*

$$\Omega_T \setminus \Omega^u \subset \Sigma_1^u \cup \Sigma_2^u,$$

*where*

$$\Sigma_1^u := \left\{ z_0 \in \Omega_T \,:\, \liminf_{\rho\searrow 0} \fint_{Q_\rho(z_0)} \left|Du - (Du)_{z_0,\rho}\right|^p dz > 0 \right\}$$

*and*

$$\Sigma_2^u := \left\{ z_0 \in \Omega_T \,:\, \limsup_{\rho\searrow 0} \left(|u_{z_0,\rho}| + |(Du)_{z_0,\rho}|\right) = \infty \right\}.$$

*In the case of a structure function* $a(z,u,\xi) \equiv a(z,\xi)$ *that does not depend on* $u$*, the above statement remains true if we replace the set* $\Sigma_2^u$ *by*

$$\widehat{\Sigma}_2^u := \left\{ z_0 \in \Omega_T \,:\, \limsup_{\rho\searrow 0} |(Du)_{z_0,\rho}| = \infty \right\}.$$

Finally, we can characterize the singular set also in terms of the function

$$V(\xi) = \xi(1 + |\xi|^2)^{\frac{p-2}{4}}.$$

This characterization will be more useful for the dimension reduction of the singular set that we will address in a forthcoming paper [36].

**Proposition 1.3.** *Under the assumptions of* Theorem 1.1*, there holds*

$$\Omega_T \setminus \Omega^u \subset S_1^u \cup S_2^u,$$

*where*

$$S_1^u := \left\{ z_0 \in \Omega_T \,:\, \liminf_{\rho \searrow 0} ⨍_{Q_\rho(z_0)} \left|V(Du) - [V(Du)]_{z_0,\rho}\right|^2 dz > 0 \right\}$$

*and*

$$S_2^u := \left\{ z_0 \in \Omega_T \,:\, \limsup_{\rho \searrow 0} \big( |u_{z_0,\rho}| + |[V(Du)]_{z_0,\rho}| \big) = \infty \right\}.$$

*Again, if the structure function $a$ is independent from $u$, the set $S_2^u$ can be replaced by*

$$\widehat{S}_2^u := \left\{ z_0 \in \Omega_T \,:\, \limsup_{\rho \searrow 0} \left|[V(Du)]_{z_0,\rho}\right| = \infty \right\}.$$

Next, we shortly describe the strategy of the proof and the organization of the paper. We start by stating our general assumptions in Section 2 and by gathering some necessary basic material in Section 3. The first step in the proof is the derivation of a Caccioppoli type inequality, which will be established in Section 4. Because of the time derivative appearing in the system (1.1), the Caccioppoli inequality holds only with the $L^2$-norm of $u$ on the right-hand side, see (4.2). This is the reason why throughout the paper, we will use the $L^2$-excess for the solution $u$ itself, while the gradient $Du$ will be measured in terms of the function $V$, cf. (6.1). In particular, this principle is expressed in the specific form of the $\mathcal{A}$-caloric approximation lemma that we state in Section 5. In Section 6 we use a linearization argument in order to prove that weak solutions of (1.1) are approximately $\mathcal{A}$-caloric provided their excess is small. Consequently, we may use the $\mathcal{A}$-caloric approximation lemma to find an $\mathcal{A}$-caloric map that is close to the solution. Using an elementary estimate for $\mathcal{A}$-caloric maps from Section 7, we thus derive the crucial decay estimate in Section 8. This yields the first Regularity Theorem 8.3, which characterizes the singular set in terms of the $L^2$-excess of the solution. In order to conclude that the singular set is negligible, we need a Poincaré-Sobolev-type inequality for solutions which bounds the $L^2$-excess of $u$ in terms of only the spatial gradient $Du$. This will be established in Section 9. The final characterization of the singular set, which implies in particular Theorem 1.1, is given in Section 10.

## 2. General assumptions

On a domain $\Omega_T := \Omega\times(-T,0)$, where $\Omega \subset \mathbb{R}^n$ is a bounded smooth domain and $T > 0$, we consider weak solutions $u : \Omega_T \to \mathbb{R}^N$ of parabolic systems of the form

$$\partial_t u(z) - \operatorname{div}\, a(z,u(z),Du(z)) = B(z,u(z),Du(z)) \quad \text{for } z=(x,t)\in\Omega_T \tag{2.1}$$

in the distributional sense, and where

$$a : \Omega_T \times \mathbb{R}^N \times \mathbb{R}^{Nn} \to \mathbb{R}^{Nn},$$
$$B : \Omega_T \times \mathbb{R}^N \times \mathbb{R}^{Nn} \to \mathbb{R}^{N}.$$

Here and in the sequel, we identify $\mathbb{R}^{Nn}$ with the space of linear functions $\mathbb{R}^n \to \mathbb{R}^N$. The scalar product on $\mathbb{R}^{Nn}$ will be denoted by $\langle \cdot,\cdot\rangle$, while the scalar product on $\mathbb{R}^N$ will written by a single dot.

***Assumptions on the structure function a.*** We assume that the total derivative $D_\xi a(z,u,\xi)$ with respect to the $\xi$-variable exists and impose the following ellipticity and growth conditions on $a$, for some $\boldsymbol{\frac{2n}{n+2} < p < 2}$.

$$\big\langle D_\xi a(z,u,\xi)\zeta,\zeta\big\rangle \ge \nu(1+|\xi|^2)^{\frac{p-2}{2}}|\zeta|^2 \tag{A1}$$

$$\big|a(z,u,\xi)\big| \le \Lambda(1+|\xi|^2)^{\frac{p-1}{2}} \tag{A2}$$

for all $z\in\Omega_T$, $u\in\mathbb{R}^N$ and $\xi,\zeta\in\mathbb{R}^{Nn}$, where $0<\nu\le\Lambda$ are given constants. With respect to the variables $(z,u)$, we assume the following continuity property. We write $\mathrm{d}_{\mathrm{par}}$ for the standard parabolic distance, cf. (3.1).

$$\frac{|a(z,u,\xi)-a(z_0,u_0,\xi)|}{(1+|\xi|^2)^{\frac{p-1}{2}}} \le 2\Lambda\,\theta\big(|u|+|u_0|, \mathrm{d}_{\mathrm{par}}(z,z_0)+|u-u_0|\big) \tag{A3}$$

for all $z,z_0\in\Omega_T$, $u,u_0\in\mathbb{R}^N$ and $\xi\in\mathbb{R}^{Nn}$, with a modulus of continuity

$$\theta(y,r) = \min\big\{1, K_\theta(y)r^\beta\big\}, \qquad \text{for all } y,r\in[0,\infty)$$

for some non-decreasing function $K_\theta : [0,\infty)\to[1,\infty)$ and $\beta\in(0,1)$. In the following, we will frequently use the fact

$$\theta\big(|u|+|u_0|, \mathrm{d}_{\mathrm{par}}(z,z_0)+|u-u_0|\big) \le K_\theta(2|u_0|+1)(\mathrm{d}_{\mathrm{par}}(z,z_0)+|u-u_0|)^\beta, \tag{2.2}$$

which can easily be checked by distinguishing the cases $|u-u_0|\le 1$ and $|u-u_0|>1$. Concerning the growth of the derivative $D_\xi a$, we assume only that there is a nondecreasing function $K_D : [0,\infty)\to(0,\infty)$ with

$$|D_\xi a(z,u,\xi)| \le \Lambda K_D(|u|+|\xi|) \tag{A4}$$

for all $z\in\Omega_T$, $u\in\mathbb{R}^N$ and $\xi\in\mathbb{R}^{Nn}$.

Moreover we assume that the map $(z,u,\xi)\mapsto D_\xi a(z,u,\xi)$ is continuous with local moduli of continuity $\omega_M:[0,\infty)\to[0,1]$ with $\lim_{s\searrow 0}\omega_M(s)=0$ for all $M>0$ such that

$$\begin{aligned}|D_\xi a(z,u,\xi)-D_\xi a(z_0,u_0,\xi_0)|&\le\\ &\le 2\Lambda K_D(M)\omega_M\big(\mathrm{d}^2_{\mathrm{par}}(z,z_0)+|u-u_0|^2+|\xi-\xi_0|^2\big)\end{aligned}\tag{A5}$$

for all $z,z_0\in\Omega_T$, $u,u_0\in\mathbb{R}^N$ and $\xi,\xi_0\in\mathbb{R}^{Nn}$ with $|u|+|\xi|\le M$ and $|u_0|+|\xi_0|\le M$. We can assume without loss of generality that $M\mapsto\omega_M(s)$ is non-decreasing for every $s>0$ and that $s\mapsto\omega_M^2(s)$ is a concave and non-decreasing function for every fixed $M>0$.

***Assumptions on the inhomogeneity B.*** For the inhomogeneity, we consider one of the following two alternative conditions. We will assume that $B$ satisfies either the *controllable growth condition*

$$|B(z,u,\xi)|\le\Lambda(1+|\xi|^2)^{\frac{p}{4}}\tag{B1}$$

for all $z\in\Omega_T$, $u\in\mathbb{R}^N$ and $\xi\in\mathbb{R}^{Nn}$, or the *natural growth condition*

$$|B(z,u,\xi)|\le\Lambda_1(M)|\xi|^p+\Lambda_2(M)\tag{B2}$$

for any $M>0$ and all $(z,u,\xi)\in\Omega_T\times\mathbb{R}^N\times\mathbb{R}^{Nn}$ with $|u|\le M$, where $\Lambda_1(M),\Lambda_2(M)>0$ depend only on $M$. In the case of the latter condition, we will need to restrict ourselves to bounded solutions of (2.1), where the bound $M_u:=\sup_{\Omega_T}|u|$ satisfies the *smallness assumption*

$$2\Lambda_1(M_u)M_u<\nu.\tag{2.3}$$

Such a smallness condition is necessary for a partial regularity result even in the elliptic case, see [24, Chapter VI]. In the parabolic case, a similar assumption has been used e.g. in [25, 26].

By classical results, we may assume that weak solutions exist in the function space $C^0(-T,0;L^2(\Omega,\mathbb{R}^N))\cap L^p(-T,0;W^{1,p}(\Omega,\mathbb{R}^N))$ for every given Cauchy-Dirichlet boundary data, see e.g. [31].

## 3. Preliminary material

**Notation.** Throughout this paper, we will denote the space variable with $x$ and the time variable with $t$; moreover, we will use notations like $z=(x,t)$ or $z_0=(x_0,t_0)$ for points in space-time.

For $x_0\in\mathbb{R}^n$ and $\rho>0$, we will write $B_\rho(x_0)$ for the open ball with radius $\rho$ and center $x_0$ in $\mathbb{R}^n$. If the center is zero, we will often omit it for the sake of brevity. Furthermore, we will work with *parabolic cylinders*

$$Q_\rho(z_0):=B_\rho(x_0)\times(t_0-\rho^2,t_0),$$

where $z_0 = (x_0, t_0) \in \mathbb{R}^n \times \mathbb{R}$. These cylinders can be interpreted as the half-balls with respect to the *parabolic metric*

$$\mathrm{d}_{\mathrm{par}}(z_1, z_2) := \max\left\{|x_1 - x_2|, \sqrt{|t_1 - t_2|}\,\right\} \tag{3.1}$$

for $z_i = (x_i, t_i), i = 1, 2$. For a general cylinder $Q := B \times (s, t)$, where $B \subset \mathbb{R}^n$ and $s < t$, we write

$$\partial_{\mathrm{par}} Q = (B \times \{s\}) \cup (\partial B \times [s, t])$$

for the *parabolic boundary* of $Q$, while the set $\overline{Q} \setminus \partial_{\mathrm{par}} Q$ is called the *parabolic interior* of $Q$.

For a function $f \in L^1(Q_\rho(z_0), \mathbb{R}^k)$ we abbreviate as usually $f_{z_0,\rho} := ⨍_{Q_\rho(z_0)} f\,dz$. In the case $z_0 = 0$ we will frequently omit the parameter $z_0$ and simply write $f_\rho$.

***The functions V and W.*** We define functions

$$V(A) = \frac{A}{(1 + |A|^2)^{\frac{2-p}{4}}} \qquad \text{and} \qquad W(A) = \frac{A}{(1 + |A|)^{\frac{2-p}{2}}} \tag{3.2}$$

for all $A \in \mathbb{R}^k$, where $k \in \mathbb{N}$. First of all we note that both functions are equivalent in the sense

$$2^{\frac{p-2}{4}} |V(A)| \le |W(A)| \le |V(A)| \qquad \text{for all } A \in \mathbb{R}^k. \tag{3.3}$$

The reason why we will sometimes use the function $W$ instead of $V$ is the fact that

$$\mathbb{R}^k \ni A \to |W(A)|^2 \qquad \text{is a convex function,} \tag{3.4}$$

cf. for example [16, Sect. 3]. We recall the following standard inequalities for later reference.

**Lemma 3.1.** ([8]). *Let $1 < p < 2$ and $V$ as defined above. Then we have for any $A, B \in \mathbb{R}^k$ and $r > 0$*

(*i*) $2^{\frac{p-2}{4}} \min\{|A|, |A|^{p/2}\} \le |V(A)| \le \min\{|A|, |A|^{p/2}\}$

(*ii*) $|V(rA)| \le \max\{r, r^{p/2}\}|V(A)|$

(*iii*) $\dfrac{p}{2}\dfrac{|B - A|}{(1 + |A|^2 + |B|^2)^{\frac{2-p}{4}}} \le |V(B) - V(A)| \le c(k,p)\dfrac{|B - A|}{(1 + |A|^2 + |B|^2)^{\frac{2-p}{4}}}$

(*iv*) $|V(A) - V(B)| \le c(k, p)|V(A - B)|$

(*v*) $|V(A - B)| \le c(p, M)|V(A) - V(B)|$ *if* $|A| \le M$, $B \in \mathbb{R}^k$.

The inequalities $(i)$ and $(ii)$ readily follow from the definition of $V$, the estimates $(iii)$ are proved in [1, Lemma 2.2] and the claim $(iv)$ is a consequence of $(iii)$. Finally, the proof of $(v)$ can be found in [8, Lemma 2.1].

Furthermore, we will need the following

**Lemma 3.2.** ([1, Lemma 2.1]). *For every* $1 < p < 2$ *there holds*

$$1 \le \frac{\int_0^1 (1 + |A + s(B-A)|^2)^{\frac{p-2}{2}}\, ds}{(1 + |A|^2 + |B|^2)^{\frac{p-2}{2}}} \le \frac{8}{p-1} \qquad \text{for any } A, B \in \mathbb{R}^k.$$

***Affine Functions.*** For a given function $u \in L^2(Q_\rho(z_0), \mathbb{R}^N)$ we write $\ell_{z_0,\rho}(z) = \ell_{z_0,\rho}(x)$ for the unique affine function $\mathbb{R}^n \to \mathbb{R}^N$ minimizing the functional

$$\ell \mapsto \fint_{Q_\rho(z_0)} |u - \ell|^2\, dz.$$

We recall some properties of this function. A straightforward calculation shows that $\ell_{z_0,\rho}(x) = u_{z_0,\rho} + A_{z_0,\rho}(x - x_0)$, where

$$D\ell_{z_0,\rho} = A_{z_0,\rho} = \frac{n+2}{\rho^2} \fint_{Q_\rho(z_0)} u \otimes (x - x_0)\, dz.$$

The following lemma can be proven analogously to [29, Lemma 2].

**Lemma 3.3.** *For* $u \in L^2(Q_\rho(z_0), \mathbb{R}^N)$ *and* $\tau \in (0,1)$, *let* $\ell_{z_0,\rho}$ *and* $\ell_{z_0,\tau\rho}$ *be the affine functions* $\mathbb{R}^n \to \mathbb{R}^N$ *defined as above for the radii* $\rho$ *and* $\tau\rho$, *respectively. Then we have*

$$\left| D\ell_{z_0,\rho} - D\ell_{z_0,\tau\rho} \right|^2 \le \frac{n(n+2)}{(\tau\rho)^2} \fint_{Q_{\tau\rho}(z_0)} |u - \ell_{z_0,\rho}|^2\, dz \tag{3.5}$$

*and furthermore*

$$\left| D\ell_{z_0,\rho} - (Du)_{z_0,\rho} \right|^2 \le \frac{n(n+2)}{\rho^2} \fint_{Q_\rho(z_0)} \left| u(z) - u_{z_0,\rho} - (Du)_{z_0,\rho}(x - x_0) \right|^2 dz. \tag{3.6}$$

Analogously to the estimate (3.5), we get

$$|u_{z_0,\rho} - u_{z_0,\tau\rho}|^2 = \left| \fint_{Q_{\tau\rho}(z_0)} (u - \ell_{z_0,\rho})\, dz \right|^2 \le \fint_{Q_{\tau\rho}(z_0)} |u - \ell_{z_0,\rho}|^2\, dz, \tag{3.7}$$

where we used the fact $\int_{Q_\rho(z_0)} A(x - x_0)\, dz = 0$ for all $A \in \mathbb{R}^{Nn}$ and Jensen's inequality. For later reference, we state the following consequence of (3.5) and (3.7).

$$(\tau\rho)^{-2} |u_{z_0,\rho} - u_{z_0,\tau\rho}|^2 + \left| D\ell_{z_0,\rho} - D\ell_{z_0,\tau\rho} \right|^2 \le \Big( \frac{n+1}{\tau\rho} \Big)^2 \fint_{Q_{\tau\rho}(z_0)} |u - \ell_{z_0,\rho}|^2\, dz \tag{3.8}$$

for all $u \in L^2(Q_\rho(z_0), \mathbb{R}^N)$ and $\tau \in (0,1)$.

Finally, we state a well-known elementary lemma which will be useful in several situations. Its proof can be found in [24, Lemma V.3.1].

**Lemma 3.4.** *For $R_0 < R_1$, let $f : [R_0, R_1] \to [0, \infty)$ be a bounded function and assume that for all $R_0 < \sigma < \rho < R_1$, there holds*

$$f(\sigma) \le \vartheta f(\rho) + \frac{A}{(\rho - \sigma)^\alpha} + B$$

*for nonnegative constants $A, B$, $\alpha$ and a parameter $\vartheta \in (0, 1)$. Then we have the estimate*

$$f(\sigma_0) \le c\Big(\frac{A}{(\rho_0 - \sigma_0)^\alpha} + B\Big)$$

*for all $R_0 \le \sigma_0 < \rho_0 \le R_1$ and some constant $c = c(\alpha, \vartheta)$.*

## 4. A Caccioppoli type inequality

The first step in the proof of partial regularity is to establish a Caccioppoli type inequality. The same proof also yields an estimate for the $L^2$-norms of $u$ on the time slices. The precise statement is as follows.

**Theorem 4.1.** *Let $M > 0$ and consider a map $u \in C^0(-T, 0; L^2(\Omega, \mathbb{R}^N)) \cap L^p(-T, 0; W^{1,p}(\Omega, \mathbb{R}^N))$ that weakly solves the system* (2.1)*, under the assumptions* (A1) *to* (A4)*. Furthermore, we require that one of the following alternatives hold.*

- *(i)* Controllable growth. *The inhomogeneity satisfies* (B1)*, or*
- *(ii)* Natural growth. *The inhomogeneity satisfies* (B2)*, and the solution $u$ and the map $\ell$ are bounded by $|u|, |\ell| \le M_u$ on $\Omega_T$, where the bound satisfies the smallness condition*

$$2M_u \Lambda_1(M_u) < \nu.$$

*We choose an arbitrary parabolic cylinder $Q_\rho(z_0) \subset \Omega_T$ with $\rho \in (0, 1)$, a radius $\sigma \in [\frac{\rho}{2}, \rho)$ and an affine function $\ell : \mathbb{R}^n \to \mathbb{R}^N$ with $|\ell(x_0)| + |D\ell| \le M$. Then there holds*

$$\sup_{t \in (t_0 - \sigma^2, t_0)} \fint_{B_\sigma(x_0)} \Big|\frac{u(x,t) - \ell(x)}{\sigma}\Big|^2 dx + \fint_{Q_\sigma(z_0)} |V(Du - D\ell)|^2 \, dz$$
$$\le \frac{c_0 \rho^2}{(\rho - \sigma)^2} \fint_{Q_\rho(z_0)} \Big|\frac{u - \ell}{\rho}\Big|^2 dz + c_0 \rho^{2\beta}. \tag{4.1}$$

*Here, the constant $c_0$ depends only on $M, K_\theta(\cdot), K_D(\cdot), n, N, p, \Lambda$ and $\nu$, and in the case of (ii) additionally on $M_u, \Lambda_1(M_u)$ and $\Lambda_2(M_u)$.*

*If in the case of (i), the structure function $a(z,u,\xi) \equiv a(z,\xi)$ does not depend on $u$, the above estimate holds for all affine maps $\ell : \mathbb{R}^n \to \mathbb{R}^N$ satisfying the weaker assumption $|D\ell| \le M$.*

**Remark 4.2.** In particular, the estimate from the above lemma implies the following Caccioppoli type inequality on any parabolic cylinder $Q_\rho(z_0) \subset \Omega_T$.

$$\fint_{Q_{\rho/2}(z_0)} |V(Du - D\ell)|^2 \, dz \le 4c_0 \fint_{Q_\rho(z_0)} \Big|\frac{u-\ell}{\rho}\Big|^2 dz + c_0 \rho^{2\beta}, \tag{4.2}$$

where the constant $c_0$ may depend on $\ell$. This inequality is weaker than the corresponding statement in the elliptic case, where instead of $|\frac{u-\ell}{\rho}|^2$, the term $|V(\frac{u-\ell}{\rho})|^2$ appears on the right-hand side, see [8, 16, 2]. However, the evolutionary term in the parabolic system (1.1) enforces a term which is quadratic in $u$, so that (4.2) can not be expected to be improved. At this point the assumption $p > \frac{2n}{n+2}$ is crucial in order to guarantee that the right-hand side in (4.2) is finite.

The peculiar form (4.2) of the Caccioppoli inequality makes it natural to work with the standard $L^2$-excess solution $u$, while for the gradient $Du$, we will define the excess by means of the function $V$ as introduced in (6.1) below.

**Proof.** Unless stated differently, we write $c$ for constants that depend at most on the structure constants listed in the theorem. In order to facilitate the notation, we may assume $z_0 = 0$. We fix an affine function $\ell : \mathbb{R}^n \to \mathbb{R}^N$ with $|\ell| + |D\ell| \le M$.

The idea is to test the system (2.1) with functions of the form

$$\varphi(x,t) := \zeta_\varepsilon(t)\psi^2(x)\big(u(x,t) - \ell(x)\big),$$

where $\psi \in C^\infty_{\mathrm{cpt}}(B_\rho, [0,1])$ is a cut-off function in space with $\psi \equiv 1$ on $B_\sigma$ and $|D\psi| \le \frac{2}{\rho-\sigma}$ on $B_\rho$, while $\zeta_\varepsilon \in C^\infty(\mathbb{R})$, for any $s \in (-\sigma^2, 0)$ and $\varepsilon \in (0, \sigma^2 + s)$, is a Lipschitz-continuous cut-off function in time with the following properties.

$$\begin{cases} \zeta_\varepsilon \equiv 0 & \text{on } (-\infty, -\rho^2] \\ |\zeta_\varepsilon'| \le \frac{2}{(\rho-\sigma)^2} & \text{on } (-\rho^2, -\sigma^2] \\ \zeta_\varepsilon \equiv 1 & \text{on } (-\sigma^2, s-\varepsilon] \\ \zeta_\varepsilon(t) = -\frac{1}{\varepsilon}(t-s) & \text{on } (s-\varepsilon, s] \\ \zeta_\varepsilon \equiv 0 & \text{on } (s, \infty) \end{cases}$$

These test functions are formally not admissible since they are not smooth, and not even weakly differentiable in the time direction. Nevertheless, their use as test function can be justified by standard approximation techniques or by the use of Steklov averages. Here we will shortly sketch the first method. For any function $f$ and $\delta > 0$, we write $f_\delta$ for the mollification of $f$ by a standard symmetric mollifier with smoothing radius $\delta$. For sufficiently small values of $\delta$, we may test (2.1) with

$$\varphi_\delta := \big[\zeta_\varepsilon \psi^2 \big(u-\ell\big)_\delta\big]_\delta.$$

Keeping in mind that $\ell$ does not depend on $t$, we deduce the equation

$$\begin{aligned}\int_{Q_\rho} \langle a(z,u,Du), D\varphi_\delta\rangle \, dz &= \int_{Q_\rho} (u-\ell)\cdot \partial_t \varphi_\delta \, dz + \int_{Q_\rho} B(z,u,Du)\cdot \varphi_\delta \, dz \\ &= \int_{Q_\rho} (u-\ell)_\delta \cdot \partial_t(\zeta_\varepsilon \psi^2 (u-\ell)_\delta)\, dz \\ &\quad + \int_{Q_\rho} \big[B(z,u,Du)\big]_\delta \cdot \zeta_\varepsilon \psi^2 (u-\ell)_\delta \, dz.\end{aligned} \tag{4.3}$$

Here we compute for the first integral on the right-hand side

$$\begin{aligned}\int_{Q_\rho} (u-\ell)_\delta \cdot \partial_t(\zeta_\varepsilon \psi^2 (u-\ell)_\delta)\, dz &= \\ &= \frac{1}{2}\int_{Q_\rho} \zeta_\varepsilon \psi^2 \frac{\partial}{\partial t}|(u-\ell)_\delta|^2 \, dz + \int_{Q_\rho} \zeta_\varepsilon' \psi^2 |(u-\ell)_\delta|^2 \, dz \\ &= \frac{1}{2}\int_{Q_\rho} \zeta_\varepsilon' \psi^2 |(u-\ell)_\delta|^2 \, dz,\end{aligned} \tag{4.4}$$

by an integration by parts in the last step. Next we note that we have the following convergences as $\delta \searrow 0$.

$$\begin{aligned} \varphi_\delta &\to \zeta_\varepsilon \psi^2 (u-\ell) && \text{in } W^{1,p} \\ (u-\ell)_\delta &\to u-\ell && \text{in } L^2 \text{ and almost everywhere} \\ [B(\cdot,u,Du)]_\delta &\to B(\cdot,u,Du) && \text{almost everywhere.}\end{aligned}$$

Consequently, by plugging (4.4) into (4.3) and letting $\delta \searrow 0$, we deduce

$$\begin{aligned}\int_{Q_\rho} \zeta_\varepsilon \psi^2 \big\langle a(z,u,Du), Du - D\ell\big\rangle \, dz &= \\ &= -2\int_{Q_\rho} \zeta_\varepsilon \psi \big\langle a(z,u,Du), D\psi \otimes (u-\ell)\big\rangle \, dz \\ &\quad + \frac{1}{2}\int_{Q_\rho} \zeta_\varepsilon' \psi^2 |u-\ell|^2 \, dz + \int_{Q_\rho} \zeta_\varepsilon \psi^2 B(z,u,Du)\cdot (u-\ell)\, dz.\end{aligned} \tag{4.5}$$

Next we note that we have the following pointwise estimate, which is a consequence of the ellipticity condition (A1) on the structure function $a$.

$$
\begin{aligned}
(4.6)\quad & \big\langle a(z,u,Du)-a(z,u,D\ell), Du-D\ell\big\rangle = \\
&= \int_0^1 \Big\langle D_\xi a(z,u,D\ell+s(Du-D\ell))(Du-D\ell), Du-D\ell\Big\rangle\, ds \\
&\geq \nu \int_0^1 (1+|D\ell+s(Du-D\ell)|^2)^{\frac{p-2}{2}}|Du-D\ell|^2\, ds \\
&\geq \nu(1+|D\ell|^2+|Du|^2)^{\frac{p-2}{2}}|Du-D\ell|^2,
\end{aligned}
$$

where we used Lemma 3.2 in the last step. Writing $\varphi := \zeta_\varepsilon\psi^2(u-\ell)$ and putting together (4.6) and (4.5), we infer

$$
\begin{aligned}
&\nu\int_{Q_\rho} \zeta_\varepsilon\psi^2(1+|D\ell|^2+|Du|^2)^{\frac{p-2}{2}}|Du-D\ell|^2\, dz \leq \\
&\quad\leq \int_{Q_\rho} \zeta_\varepsilon\psi^2\big\langle a(z,u,Du)-a(z,u,D\ell), Du-D\ell\big\rangle\, dz \\
&\quad= -2\int_{Q_\rho} \zeta_\varepsilon\psi\big\langle a(z,u,Du)-a(z,u,D\ell), D\psi\otimes(u-\ell)\big\rangle\, dz \\
&\qquad - \int_{Q_\rho} \big\langle a(z,u,D\ell), D\varphi\big\rangle\, dz \\
&\qquad + \frac{1}{2}\int_{Q_\rho} \zeta_\varepsilon'\psi^2|u-\ell|^2\, dz \\
&\qquad + \int_{Q_\rho} \zeta_\varepsilon\psi^2 B(z,u,Du)\cdot(u-\ell)\, dz \\
(4.7)\qquad &\quad =: I+II+III+IV.
\end{aligned}
$$

The rest of the proof consists of the estimation of the terms $I$ to $IV$.

**Estimate for *I*.** We decompose $I = I_1+I_2+I_3$, where

$$
\begin{aligned}
I_1 &:= -2\int_{Q_\rho} \zeta_\varepsilon\psi\big\langle a(z,u,Du)-a(z,\ell,Du), D\psi\otimes(u-\ell)\big\rangle\, dz \\
I_2 &:= -2\int_{Q_\rho} \zeta_\varepsilon\psi\big\langle a(z,\ell,Du)-a(z,\ell,D\ell), D\psi\otimes(u-\ell)\big\rangle\, dz \\
I_3 &:= -2\int_{Q_\rho} \zeta_\varepsilon\psi\big\langle a(z,\ell,D\ell)-a(z,u,D\ell), D\psi\otimes(u-\ell)\big\rangle\, dz.
\end{aligned}
$$

For the first and the last term, we employ the continuity assumption (A3) on $a$, the properties of $\psi$ and the assumption $|D\ell| \le M$ in order to estimate

$$
\begin{aligned}
I_1+I_3 &\le c(p)\,\Lambda\int_{Q_\rho} \zeta_\varepsilon\psi\,\theta(|u|+|\ell|,|u-\ell|)\,(1+|Du|^{p-1}+|D\ell|^{p-1})\,\Big|\frac{u-\ell}{\rho-\sigma}\Big|\,dz\\
&\le c(p)\Lambda\int_{Q_\rho} \zeta_\varepsilon\psi\,\theta(|u|+|\ell|,|u-\ell|)(1+M^{p-1})\Big|\frac{u-\ell}{\rho-\sigma}\Big|\,dz\\
&\quad + c(p)\Lambda\int_{Q_\rho\cap\{|Du-D\ell|>1\}} \zeta_\varepsilon\psi\,\theta(|u|+|\ell|,|u-\ell|)|Du-D\ell|^{p-1}\Big|\frac{u-\ell}{\rho-\sigma}\Big|\,dz\\
&=: I'+I''.
\end{aligned}
$$

For the first integral, we use the estimate (2.2) for $\theta$ and $|\ell| \le M$, which yields

$$
\begin{aligned}
I' &\le c(p)\Lambda\,K_\theta(2M+1)(1+M^{p-1})\int_{Q_\rho} \rho^\beta\Big|\frac{u-\ell}{\rho-\sigma}\Big|^{1+\beta}\,dz\\
&\le c\int_{Q_\rho}\Big(\rho^{\frac{2\beta}{1-\beta}}+\Big|\frac{u-\ell}{\rho-\sigma}\Big|^2\Big)\,dz \le c\Big(\rho^{2\beta}|Q_\rho| + \int_{Q_\rho}\Big|\frac{u-\ell}{\rho-\sigma}\Big|^2\,dz\Big), \qquad (4.8)
\end{aligned}
$$

where we used Young' inequality with exponents $\frac{2}{1-\beta}$ and $\frac{2}{1+\beta}$. For the term $I''$, we simply estimate $\theta \le 1$ and get

$$
\begin{aligned}
I'' &\le c(p)\Lambda\int_{Q_\rho\cap\{|Du-D\ell|>1\}} \zeta_\varepsilon\psi\,|Du-D\ell|^{p-1}\,\Big|\frac{u-\ell}{\rho-\sigma}\Big|\,dz\\
&\le c(p)\Lambda\int_{Q_\rho\cap\{|Du-D\ell|>1\}} \zeta_\varepsilon\psi\,|Du-D\ell|^{\frac{p}{2}}\,\Big|\frac{u-\ell}{\rho-\sigma}\Big|\,dz\\
&\le c(p)\Lambda\int_{Q_\rho} \zeta_\varepsilon\psi\,|V(Du-D\ell)|\,\Big|\frac{u-\ell}{\rho-\sigma}\Big|\,dz\\
&\le \frac{\kappa}{2}\int_{Q_\rho}\zeta_\varepsilon\psi^2|V(Du-D\ell)|^2\,dz + c(p,\kappa)\Lambda\int_{Q_\rho}\Big|\frac{u-\ell}{\rho-\sigma}\Big|^2\,dz \qquad (4.9)
\end{aligned}
$$

for every $\kappa\in(0,1)$, where we used first $p-1\le\frac{p}{2}$, then applied Lemma 3.1 $(i)$ and finally Young's inequality. For the estimate of $I_2$, we distinguish the cases $|Du-D\ell|\le 1$ and $|Du-D\ell|>1$. In the first case, we employ the growth property (A4) of $D_\xi a$ and the fact $|\ell|+|D\ell|+|Du-D\ell|\le 2M+1$, with the result

$$
\begin{aligned}
&-2\big\langle a(z,\ell,Du)-a(z,\ell,D\ell), D\psi\otimes(u-\ell)\big\rangle =\\
&\quad = -2\int_0^1 \big\langle D_\xi a(z,\ell,D\ell+s(Du-D\ell))(Du-D\ell), D\psi\otimes(u-\ell)\big\rangle\,ds\\
&\quad \le 2\Lambda K_D(2M+1)\,|Du-D\ell|\,\Big|\frac{u-\ell}{\rho-\sigma}\Big|\\
&\quad \le c(p)\Lambda K_D(2M+1)\,|V(Du-D\ell)|\,\Big|\frac{u-\ell}{\rho-\sigma}\Big|
\end{aligned}
$$

by Lemma 3.1($i$). In the case $|Du - D\ell| > 1$, we simply make use of the growth assumption (A2) on $a$ in order to estimate

$$
\begin{aligned}
-2\big\langle a(z,\ell,Du) - a(z,\ell,D\ell), D\psi \otimes (u-\ell)\big\rangle \le \\
&\le c(p)\Lambda(2 + |Du|^{p-1} + M^{p-1})\Big|\frac{u-\ell}{\rho-\sigma}\Big| \\
&\le c(p)\Lambda(1 + M^{p-1})\,|Du - D\ell|^{p-1}\Big|\frac{u-\ell}{\rho-\sigma}\Big| \\
&\le c(p)\Lambda(1 + M^{p-1})\,|V(Du - D\ell)|\,\Big|\frac{u-\ell}{\rho-\sigma}\Big|
\end{aligned}
$$

since

$$|Du - D\ell|^{p-1} \le |Du - D\ell|^{\frac{p}{2}} \le c(p)|V(Du - D\ell)|,$$

provided $|Du - D\ell| > 1$. Putting together the last two estimates, we arrive at

$$
\begin{aligned}
I_2 &\le c\int_{Q_\rho} \zeta_\varepsilon\psi|V(Du - D\ell)|\,\Big|\frac{u-\ell}{\rho-\sigma}\Big|\,dz \\
&\le \frac{\kappa}{2}\int_{Q_\rho} \zeta_\varepsilon\psi^2|V(Du - D\ell)|^2\,dz + c(\kappa)\int_{Q_\rho}\Big|\frac{u-\ell}{\rho-\sigma}\Big|^2\,dz
\end{aligned}
\tag{4.10}
$$

by Young's inequality, for every $\kappa \in (0,1)$, where $c(\kappa)$ depends only on $\kappa$ and the data listed in the theorem. Collecting the estimates (4.8) to (4.10), we arrive at

$$
\begin{aligned}
I &= I_2 + I' + I'' \\
&\le \kappa\int_{Q_\rho} \zeta_\varepsilon\psi^2|V(Du - D\ell)|^2\,dz + c(\kappa)\Big(\int_{Q_\rho}\Big|\frac{u-\ell}{\rho-\sigma}\Big|^2\,dz + \rho^{2\beta}|Q_\rho|\Big).
\end{aligned}
\tag{4.11}
$$

**Estimate for *II*.** Since the test function $\varphi$ has compact support in $Q_\rho$, integrating by parts on the time slices yields

$$\int_{Q_\rho} \big\langle a(0,t,\ell(0),D\ell), D\varphi\big\rangle\,dx\,dt = 0.$$

Consequently, we can write

$$
\begin{aligned}
II &= -\int_{Q_\rho} \big\langle a(z,u,D\ell) - a(z,\ell,D\ell), D\varphi\big\rangle\,dz \\
&\quad - \int_{Q_\rho} \big\langle a(x,t,\ell,D\ell) - a(0,t,\ell(0),D\ell), D\varphi\big\rangle\,dx\,dt \\
&=: II_1 + II_2.
\end{aligned}
$$

For the estimate of the first term, we employ the continuity (A3) of $a$ with respect to the second variable, together with the assumption $|D\ell| \le M$, which yields

$$\begin{aligned} II_1 &\le c(p)\Lambda(1+M^{p-1}) \int_{Q_\rho} \theta(|u|+|\ell|, |u-\ell|)\, |D\varphi|\, dz \\ &\le c \int_{Q_\rho\cap\{|u-\ell|\le 1\}} \theta(|u|+|\ell|, |u-\ell|)\, |D\varphi|\, dz + c\int_{Q_\rho\cap\{|u-\ell|>1\}} |D\varphi|\, dz \\ &=: II_{1,1} + II_{1,2}. \end{aligned}$$

Here, we used the fact $\theta \le 1$ in the estimate of the last integral. For the estimate of $II_{1,1}$, we employ the property (2.2) of $\theta$ and infer

$$\begin{aligned} II_{1,1} &\le cK_\theta(2M+1) \int_{Q_\rho\cap\{|u-\ell|\le 1\}} |u-\ell|^\beta \Big(\zeta_\varepsilon\psi^2|Du-D\ell| + \Big|\frac{u-\ell}{\rho-\sigma}\Big|\Big)\, dz \\ &\le c \int_{Q_\rho\cap\{|u-\ell|\le 1\}} \zeta_\varepsilon\psi^2 |u-\ell|^\beta \big(|V(Du-D\ell)| + |V(Du-D\ell)|^{2/p}\big)\, dz \\ &\quad + \int_{Q_\rho} \rho^\beta \Big|\frac{u-\ell}{\rho-\sigma}\Big|^{1+\beta} dz. \end{aligned}$$

Here we applied Lemma 3.1($i$), distinguishing the cases $|Du-D\ell| \le 1$ and $|Du-D\ell| > 1$. Next, we apply Young's inequality to each of the three summands. Keeping in mind that $\frac{p}{p-1} > 2$ and $|u-\ell| \le 1$ on the domain of integration, we deduce

$$\begin{aligned} II_{1,1} \le \frac{\kappa}{3} \int_{Q_\rho} \zeta_\varepsilon\psi^2 |V(Du-D\ell)|^2\, dz + c(\kappa)\int_{Q_\rho} |u-\ell|^{2\beta}\, dz \\ + c\rho^{\frac{2\beta}{1-\beta}}|Q_\rho| + c\int_{Q_\rho} \Big|\frac{u-\ell}{\rho-\sigma}\Big|^2 dz \end{aligned}$$

for any $\kappa \in (0,1)$. Applying Young's inequality once more, we arrive at

$$\begin{aligned} II_{1,1} \;&\le \frac{\kappa}{3} \int_{Q_\rho} \zeta_\varepsilon\psi^2 |V(Du-D\ell)|^2\, dz \\ &\quad + c(\kappa)\Big(\rho^{\frac{2\beta}{1-\beta}}|Q_\rho| + \int_{Q_\rho} \Big|\frac{u-\ell}{\rho-\sigma}\Big|^2 dz\Big). \end{aligned} \tag{4.12}$$

The term $II_{1,2}$ can be estimated by means of Lemma 3.1($i$) and Young's inequality as follows.

$$
\begin{aligned}
II_{1,2} &\le \int_{Q_\rho\cap\{|u-\ell|>1\}} \Big(\zeta_\varepsilon\psi^2|Du-D\ell| + \Big|\frac{u-\ell}{\rho-\sigma}\Big|\Big)\,dz \\
&\le \frac{\kappa}{3}\int_{Q_\rho}\zeta_\varepsilon\psi^2|V(Du-D\ell)|^2\,dz + c(\kappa)\int_{Q_\rho\cap\{|u-\ell|>1\}}\Big(1+\Big|\frac{u-\ell}{\rho-\sigma}\Big|^2\Big)\,dz \\
&\le \frac{\kappa}{3}\int_{Q_\rho}\zeta_\varepsilon\psi^2|V(Du-D\ell)|^2\,dz + c(\kappa)\int_{Q_\rho}\Big|\frac{u-\ell}{\rho-\sigma}\Big|^2\,dz. \qquad (4.13)
\end{aligned}
$$

Here we used $|u-\ell|>1$ and $\rho-\sigma\le 1$ in the last step. For the term $II_2$, we employ again the continuity assumption (A3) and the property (2.2) with the result

$$
\begin{aligned}
II_2 &\le c(p)\Lambda(1+M^{p-1})\int_{Q_\rho}\theta(|\ell|+|\ell(0)|,|x|+|\ell-\ell(0)|)\,|D\varphi|\,dz \\
&\le c\int_{Q_\rho}K_\theta(2M+1)(1+M)^\beta\rho^\beta\Big(\zeta_\varepsilon\psi^2|Du-D\ell| + \Big|\frac{u-\ell}{\rho-\sigma}\Big|\Big)\,dz \\
&= c\Big(\int_{Q_\rho}\zeta_\varepsilon\psi^2\rho^\beta|Du-D\ell|\,dz + \int_{Q_\rho}\rho^\beta\Big|\frac{u-\ell}{\rho-\sigma}\Big|\,dz\Big) \\
&=: II_{2,1}+II_{2,2}.
\end{aligned}
$$

Lemma 3.1($i$) yields, similarly as above,

$$
II_{2,1} \le c\int_{Q_\rho}\zeta_\varepsilon\psi^2\Big(\rho^\beta|V(Du-D\ell)| + \rho^\beta|V(Du-D\ell)|^{2/p}\Big)\,dz,
$$

from which we infer by Young's inequality, keeping in mind that $\rho\le 1$ and $p\le 2$,

$$
II_{2,1} \le c(\kappa)\rho^{2\beta}|Q_\rho| + \frac{\kappa}{3}\int_{Q_\rho}\zeta_\varepsilon\psi^2|V(Du-D\ell)|^2\,dz. \qquad (4.14)
$$

Finally, we apply Young's inequality to the term $II_{2,2}$ with the result

$$
II_{2,2} \le c\Big(\rho^{2\beta}\,|Q_\rho| + \int_{Q_\rho}\Big|\frac{u-\ell}{\rho-\sigma}\Big|^2\,dz\Big). \qquad (4.15)
$$

Combining the estimates (4.12) to (4.15), we infer

$$
II \le c(\kappa)\Big(\rho^{2\beta}|Q_\rho| + \int_{Q_\rho}\Big|\frac{u-\ell}{\rho-\sigma}\Big|^2\,dz\Big) + \kappa\int_{Q_\rho}\zeta_\varepsilon\psi^2|V(Du-D\ell)|^2\,dz. \qquad (4.16)
$$

**Estimate for *III*.** By the choice of $\zeta_\varepsilon$, in particular since $|\zeta_\varepsilon'| \le \frac{2}{(\rho-\sigma)^2}$ on $[-\rho^2, -\sigma^2]$, we have

$$III \le \int_{-\rho^2}^{-\sigma^2} \int_{B_\rho} \psi^2 \Big|\frac{u-\ell}{\rho-\sigma}\Big|^2 \,dx\,dt - \frac{1}{\varepsilon}\int_{s-\varepsilon}^{s}\int_{B_\rho} \psi^2 |u-\ell|^2 \,dx\,dt. \tag{4.17}$$

**Estimate for *IV* in the case of controllable growth.** In the case of (B1), we can estimate, using $|D\ell| \le M$,

$$\begin{aligned} IV &\le \Lambda \int_{Q_\rho} \zeta_\varepsilon \psi^2 (1+|Du|^2)^{\frac{p}{4}} \,|u-\ell|\,dz \\ &\le c\int_{Q_\rho} |u-\ell|\,dz + c\int_{Q_\rho \cap \{|Du-D\ell|>1\}} \zeta_\varepsilon \psi^2 |Du-D\ell|^{\frac{p}{2}}\,|u-\ell|\,dz. \end{aligned}$$

Keeping in mind the estimate $|Du-D\ell|^{\frac{p}{2}} \le c(p)|V(Du-D\ell)|$, which holds under the assumption $|Du-D\ell|>1$ according to Lemma 3.1($i$), we deduce after an application of Young's inequality

$$IV \le \kappa \int_{Q_\rho} \zeta_\varepsilon \psi^2 |V(Du-D\ell)|^2\,dz + c(\kappa)\Big(\int_{Q_\rho}\Big|\frac{u-\ell}{\rho-\sigma}\Big|^2 dz + \rho^2|Q_\rho|\Big), \tag{4.18}$$

for any $\kappa \in (0,1)$.

**Estimate for *IV* in the case of natural growth.** In this case we use the assumption $|u| \le M_u$ in order to estimate

$$\begin{aligned} IV &\le \int_{Q_\rho} \zeta_\varepsilon \psi^2 |B(z,u,Du)|\,|u-\ell|\,dz \\ &\le \Lambda_1(M_u)\int_{Q_\rho} \zeta_\varepsilon \psi^2 |Du|^p\,|u-\ell|\,dz + \Lambda_2(M_u)\int_{Q_\rho} |u-\ell|\,dz. \end{aligned} \tag{4.19}$$

Here we estimate by Young's inequality for an arbitrary $\delta > 0$

(4.20)
$$\begin{aligned} |Du|^p &\le \big(1+|D\ell|^2+|Du|^2\big)^{\frac{p}{2}} \\ &\le \big[1+(1+\delta)|Du-D\ell|^2 + (2+\delta^{-1})|D\ell|^2\big]\big(1+|D\ell|^2+|Du|^2\big)^{\frac{p-2}{2}} \\ &\le (1+\delta)|Du-D\ell|^2\big(1+|D\ell|^2+|Du|^2\big)^{\frac{p-2}{2}} + (2+\delta^{-1})(1+M^2)^{\frac{p}{2}}, \end{aligned}$$

where we used the assumption $|D\ell| \le M$ and $p \le 2$ in the last step. We determine $\delta > 0$ by requiring that $\tilde\nu := (1+\delta)2M_u\Lambda_1(M_u)$ is the arithmetic mean of $2M_u\Lambda_1(M_u)$ and $\nu$, that is

$$\delta := \frac{\nu - 2M_u\Lambda_1(M_u)}{4M_u\Lambda_1(M_u)}.$$

Note that $\delta > 0$ because of the smallness condition (2.3), and furthermore there holds $\tilde{\nu} < \nu$. We apply the estimate (4.20) with this choice of $\delta$ and plug it into (4.19). Keeping in mind that $|u-\ell| \le 2M_u$ by assumption, we infer

$$\begin{aligned} IV &\le (1+\delta)2M_u\,\Lambda_1(M_u)\int_{Q_\rho}\zeta_\varepsilon\psi^2\big(1+|D\ell|^2+|Du|^2\big)^{\frac{p-2}{2}}|Du-D\ell|^2\,dz \\ &\quad + c\int_{Q_\rho}|u-\ell|\,dz \\ &\le \tilde{\nu}\int_{Q_\rho}\zeta_\varepsilon\psi^2\big(1+|D\ell|^2+|Du|^2\big)^{\frac{p-2}{2}}|Du-D\ell|^2\,dz \\ &\quad + c\Big(\int_{Q_\rho}\Big|\frac{u-\ell}{\rho}\Big|^2\,dz+\rho^2|Q_\rho|\Big), \end{aligned} \tag{4.21}$$

where the constant $c$ depends on $p, M, \Lambda_1(M_u), \Lambda_2(M_u)$ and $\nu$.

**Final conclusion in the case of controllable growth.** Recalling the estimate (4.7), the estimates (4.11), (4.16), (4.17) and (4.18) yield

$$\begin{aligned} \frac{1}{\varepsilon}\int_{s-\varepsilon}^{s}\int_{B_\rho}\psi^2|u-\ell|^2\,dz + \nu\int_{Q_\rho}\zeta_\varepsilon\psi^2(1+|D\ell|^2+|Du|^2)^{\frac{p-2}{2}}|Du-D\ell|^2\,dz \le \\ \le c(\kappa)\Big(\rho^{2\beta}|Q_\rho| + \int_{Q_\rho}\Big|\frac{u-\ell}{\rho-\sigma}\Big|^2\,dz\Big) + 3\kappa\int_{Q_\rho}\zeta_\varepsilon\psi^2|V(Du-D\ell)|^2\,dz. \end{aligned} \tag{4.22}$$

Here we can estimate the left-hand side from below by means of Lemma 3.1 $(iii)$ and $(v)$. Keeping in mind the assumption $|D\ell| \le M$, we deduce the pointwise bound

$$\begin{aligned} \nu(1+|D\ell|^2+|Du|^2)^{\frac{p-2}{2}}|Du-D\ell|^2 &\ge c(Nn,p,\nu)|V(Du)-V(D\ell)|^2 \\ &\ge c(Nn,p,\nu,M)|V(Du-D\ell)|^2. \end{aligned}$$

Plugging this into (4.22) and choosing $\kappa \in (0,1)$ small enough, depending on $Nn, p, \nu$ and $M$, we can absorb the last integral appearing in (4.22). Letting $\varepsilon$ tend to zero, we thus deduce

$$\begin{aligned} \int_{B_\sigma\times\{s\}}|u-\ell|^2\,dx + \int_{-\sigma^2}^{s}\int_{B_\sigma}|V(Du-D\ell)|^2\,dz \le \\ \le \int_{B_\rho\times\{s\}}\psi^2|u-\ell|^2\,dx + \int_{-\sigma^2}^{s}\int_{B_\rho}\psi^2|V(Du-D\ell)|^2\,dz \\ \le c\Big(\rho^{2\beta}|Q_\rho| + \int_{Q_\rho}\Big|\frac{u-\ell}{\rho-\sigma}\Big|^2\,dz\Big) \end{aligned}$$

for all $s \in (-\sigma^2, 0)$. After taking means, this yields the desired estimate.

**Final conclusion in the case of natural growth.** In this case we employ (4.21) in order to estimate $IV$, and as above the estimates (4.7), (4.11), (4.16) and (4.17). This yields, similarly to (4.22),

$$\begin{aligned}\frac{1}{\varepsilon}\int_{s-\varepsilon}^{s}\int_{B_\rho}\psi^2|u-\ell|^2\,dz+\nu\int_{Q_\rho}\zeta_\varepsilon\psi^2(1+|D\ell|^2+|Du|^2)^{\frac{p-2}{2}}|Du-D\ell|^2\,dz\\ \le c(\kappa)\Big(\rho^{2\beta}|Q_\rho|+\int_{Q_\rho}\Big|\frac{u-\ell}{\rho-\sigma}\Big|^2\,dz\Big)+2\kappa\int_{Q_\rho}\zeta_\varepsilon\psi^2|V(Du-D\ell)|^2\,dz\\ +\tilde\nu\int_{Q_\rho}\zeta_\varepsilon\psi^2(1+|D\ell|^2+|Du|^2)^{\frac{p-2}{2}}|Du-D\ell|^2\,dz.\end{aligned}\tag{4.23}$$

Here, the last integral can be absorbed since $\tilde\nu<\nu$ by the choice of $\tilde\nu$. The remaining integrals can be treated exactly in the same way as in the case of controllable growth above, and we arrive at the same estimate, now with a constant that may depend additionally on $M_u,\Lambda_1(M_u)$ and $\Lambda_2(M_u)$.

In the case of a system without $u$-dependence, a careful inspection of the proof shows that we do not need a bound on $|\ell(x_0)|$. This is due to the fact that the functions $K_\theta$ and $K_D$ in the assumptions (A3) and (A4) can be chosen independently from $u$ in this case. ■

## 5. Subquadratic $\mathcal{A}$-caloric approximation

The technique of the $\mathcal{A}$-caloric approximation has been developed in [20, 22] for the proof of partial regularity for parabolic systems. We also refer to [3] for a higher-order analogue. Here we give a version of an $\mathcal{A}$-caloric approximation lemma that is tailored to the case of subquadratic growth.

We consider bilinear forms $\mathcal{A}$ on $\mathbb{R}^{Nn}$ that are positive and bounded in the sense

$$\mathcal{A}(\xi,\xi)\ge\lambda|\xi|^2\qquad\text{and}\qquad\mathcal{A}(\xi,\eta)\le L|\xi|\,|\eta|\tag{5.1}$$

for all $\xi,\eta\in\mathbb{R}^{Nn}$.

**Definition 5.1.** A map $h\in L^1(t_0-R^2,t_0;W^{1,1}(B_R(x_0),\mathbb{R}^N))$ is called $\mathcal{A}$-caloric iff it satisfies the linear parabolic system

$$\int_{Q_R(z_0)}\Big(h\cdot\partial_t\varphi-\mathcal{A}(Dh,D\varphi)\Big)\,dz=0,\tag{5.2}$$

for arbitrary $\varphi\in C_0^\infty(Q_R(z_0),\mathbb{R}^N)$.

By classical results, $\mathcal{A}$-caloric maps are smooth, cf. Section 7. The main result in this section is the following lemma, which will provide suitable $\mathcal{A}$-caloric comparison maps. Similar results have been proven in the case $p\ge 2$ in [20, 22] and in the elliptic case for $p<2$ in [16].

**Lemma 5.2.** *Let $0 < \lambda \le L$ be given. Then for any $\varepsilon > 0$ there is a $\delta > 0$, depending on $p, n, N, \lambda, L$ and $\varepsilon$, such that the following holds. Assume that $\mathcal{A}$ is a bilinear form on $\mathbb{R}^{Nn}$ with the properties* (5.1) *and assume that $w \in L^\infty(t_0 - \rho^2, t_0; L^2(B_\rho(x_0), \mathbb{R}^N)) \cap L^p(t_0 - \rho^2, t_0; W^{1,p}(B_\rho(x_0), \mathbb{R}^N))$ is approximately $\mathcal{A}$-caloric in the sense*

$$\left| \fint_{Q_\rho(z_0)} \Big( w \cdot \partial_t \varphi - \mathcal{A}(Dw, D\varphi) \Big)\, dz \right| \le \delta \sup_{Q_\rho(z_0)} |D\varphi| \tag{5.3}$$

*for all $\varphi \in C_0^\infty(Q_\rho(z_0), \mathbb{R}^N)$. If furthermore, $w$ satisfies*

$$\sup_{t \in (t_0 - \rho^2, t_0)} \fint_{B_\rho(x_0)} \left| \frac{w(x,t)}{\rho} \right|^2 dx + \fint_{Q_\rho(z_0)} |V(Dw)|^2\, dz \le 1, \tag{5.4}$$

*then there exists an $\mathcal{A}$-caloric map $h \in C^\infty(Q_{\rho/2}(z_0), \mathbb{R}^N)$ with*

$$\fint_{Q_{\rho/2}(z_0)} \Big( \Big| \frac{h}{\rho/2} \Big|^2 + |V(Dh)|^2 \Big)\, dz \le 2^{n+5}$$

*and*

$$\fint_{Q_{\rho/2}(z_0)} \Big| \frac{w - h}{\rho/2} \Big|^2 dz \le \varepsilon.$$

**Remark 5.3.** We point out that if the map $w$ solves a parabolic system and thus satisfies a Caccioppoli inequality as in Theorem 4.1, an analogous statement holds if (5.4) is replaced by the weaker assumption

$$\rho^{-2} \fint_{Q_\rho(z_0)} |w|^2\, dz \le 1.$$

In the general case however, a bound on both integrals in (5.4) is crucial in order to show strong compactness in $L^2$ by a compactness principle of Simon [37].

**Proof.** By scaling invariance, it suffices to prove the lemma in the case $\rho = 1$ and $z_0 = (0,0)$. The general case can be deduced from this by rescaling in the way

$$w_{z_0,\rho}(x,t) := \frac{w(x_0 + \rho x, t_0 + \rho^2 t)}{\rho} \qquad \text{for all } (x,t) \in Q_1,$$

applying the lemma on the cylinder $Q_1$ in order to get an $\mathcal{A}$-caloric map $h_{z_0,\rho} \in C^\infty(Q_{1/2}, \mathbb{R}^N)$ and scaling back by

$$h(x,t) := \rho\, h_{z_0,\rho}\Big( \frac{x - x_0}{\rho}, \frac{t - t_0}{\rho^2} \Big) \qquad \text{for all } (x,t) \in Q_{\rho/2}(z_0).$$

Thus, we assume for contradiction that the lemma is not valid on $Q_1$. In that case, we could find an $\varepsilon > 0$, sequences $\mathcal{A}_k$ of bilinear forms with (5.1) and maps $w_k \in L^\infty(-1,0;L^2(B_1,\mathbb{R}^N)) \cap L^p(-1,0;W^{1,p}(B_1,\mathbb{R}^N))$ such that for every $k \in \mathbb{N}$, the map $w_k$ is approximately $\mathcal{A}_k$-caloric in the sense

$$\left| \fint_{Q_1} \Big( w_k \partial_t \varphi - \mathcal{A}_k(Dw_k, D\varphi) \Big)\, dz \right| \le \frac{1}{k} \sup_{Q_1} \|D\varphi\|_{L^\infty} \tag{5.5}$$

for all $\varphi \in C_0^\infty(Q_1,\mathbb{R}^N)$, and satisfies the estimate

$$\sup_{t\in(-1,0)} \fint_{B_1} |w_k(x,t)|^2\, dx + \fint_{Q_1} |V(Dw_k)|^2\, dz \le 1, \tag{5.6}$$

but for all $\mathcal{A}_k$-caloric maps $h \in C^\infty(Q_{1/2},\mathbb{R}^N)$ with

$$\fint_{Q_{1/2}} \big(4|h|^2 + |V(Dh)|^2\big)\, dz \le 2^{n+5},$$

there holds

$$4\fint_{Q_{1/2}} |w_k - h|^2\, dz > \varepsilon. \tag{5.7}$$

By an approximation argument, we can assume that (5.5) holds for all $\varphi \in W^{1,\infty}(Q_1,\mathbb{R}^N)$ with zero boundary values. The estimate (5.6) implies in particular

$$\fint_{Q_1} |w_k|^2\, dz \le 1, \tag{5.8}$$

and, in view of Lemma 3.1($i$),

$$\begin{aligned} \int_{Q_1} |Dw_k|^p\, dz &= \int_{Q_1\cap\{|Dw_k|\le 1\}} |Dw_k|^p\, dz + \int_{Q_1\cap\{|Dw_k|>1\}} |Dw_k|^p\, dz \\ &\le c(p) \left( \int_{Q_1} |V(Dw_k)|^p\, dz + \int_{Q_1} |V(Dw_k)|^2\, dz \right) \le c(p)|Q_1|, \end{aligned} \tag{5.9}$$

where we used Hölder's inequality and (5.6) in the last step. By this estimate and (5.8), we can achieve convergence in the following sense, if necessary after passing to a subsequence. There is a map $w \in L^2(Q_1,\mathbb{R}^N) \cap L^p(-1,0;W^{1,p}(B_1,\mathbb{R}^N))$ and a bilinear form $\mathcal{A}$ satisfying (5.1), such that

$$\begin{cases} w_k \rightharpoonup w & \text{weakly in } L^2(Q_1,\mathbb{R}^N) \\ Dw_k \rightharpoonup Dw & \text{weakly in } L^p(Q_1,\mathbb{R}^{Nn}) \\ \mathcal{A}_k \to \mathcal{A} & \text{as bilinear forms} \end{cases} \tag{5.10}$$

as $k \to \infty$.

By lower semicontinuity of the $L^p$-norm with respect to weak convergence, we infer from (5.8) and (5.9) that

$$\fint_{Q_1} \big(|w|^2 + |Dw|^p\big)\, dz \le c(p).$$

First, we check that $w$ is $\mathcal{A}$-caloric. To this end, we write for an arbitrary map $\varphi \in C_0^\infty(Q_1, \mathbb{R}^N)$

$$\int_{Q_1} \big(w \cdot \partial_t\varphi - \mathcal{A}(Dw, D\varphi)\big)\, dz = \int_{Q_1} \big((w - w_k) \cdot \partial_t\varphi - \mathcal{A}(Dw - Dw_k, D\varphi)\big)\, dz$$
$$+ \int_{Q_1} (\mathcal{A}_k - \mathcal{A})(Dw_k, D\varphi)\, dz + \int_{Q_1} \big(w_k \cdot \partial_t\varphi - \mathcal{A}_k(Dw_k, D\varphi)\big)\, dz.$$

Letting $k \to \infty$, the first integral tends to zero by the weak convergence $w_k \rightharpoonup w$ in $L^2$ and $Dw_k \rightharpoonup Dw$ in $L^p$, the second integral tends to zero by the convergence $\mathcal{A}_k \to \mathcal{A}$ and the uniform bound (5.9), while the last integral vanishes in the limit because of (5.5). We arrive at

$$\int_{Q_1} \big(w \cdot \partial_t\varphi - \mathcal{A}(Dw, D\varphi)\big)\, dz = 0 \tag{5.11}$$

for all $\varphi \in C_0^\infty(Q_1, \mathbb{R}^N)$, which means that $w$ is $\mathcal{A}$-caloric. In particular, classical theory implies $w \in C^\infty(Q_1, \mathbb{R}^N)$.

Next we want to prove strong convergence $w_k \to w$ in $L^2(Q_{1/2}, \mathbb{R}^N)$. For this we employ a technique from [20] in order to show some uniform continuity property of $w_k$ in the time direction. For fixed times $-1 < r < s < 0$, we test (5.5) with functions of the form $\varphi(x,t) = \zeta_\kappa(t)\psi(x)$, where $\psi \in C_0^\infty(B_1, \mathbb{R}^N)$ is arbitrary and $\zeta_\kappa$ is defined by

$$\zeta_\kappa(t) := \begin{cases} \frac{1}{\kappa}(t-s) & \text{for } s \le t \le s+\kappa \\ 1 & \text{for } s+\kappa \le t \le r-\kappa \\ -\frac{1}{\kappa}(t-r) & \text{for } r-\kappa \le t \le r \end{cases}$$

and extended by zero elsewhere. This defines a Lipschitz-continuous map for every $0 < \kappa < \frac{1}{2}(r-s)$. Plugging this choice of $\varphi$ into (5.5) we infer, keeping in mind that $|\zeta_\kappa| \le 1$,

$$\Big|\int_{Q_1} w_k(x,t) \cdot \zeta_\kappa'(t)\psi(x)\, dx\, dt\Big| \le$$
$$\le \int_s^r \int_{B_1} |\mathcal{A}_k(Dw_k, D\psi)|\, dx\, dt + \frac{|Q_1|}{k}\, \|D\psi\|_{L^\infty(B_1)}$$
$$\le \Big(L \int_s^r \int_{B_1} |Dw_k|\, dx\, dt + \frac{|Q_1|}{k}\Big)\|D\psi\|_{L^\infty(B_1)}$$
$$\le c(n,p,L)\Big((r-s)^{\frac{p-1}{p}} + \frac{1}{k}\Big)\|D\psi\|_{L^\infty(B_1)},$$

where we applied Hölder's inequality and the bound (5.9) in the last step. On the other hand, the left-hand side of this inequality can be written as

$$\begin{aligned}\left|\int_{Q_1} w_k(x,t)\cdot\zeta_\kappa'(t)\psi(x)\,dx\,dt\right| &= \\ &= \left|\int_{B_1}\left(\frac{1}{\kappa}\int_s^{s+\kappa} w_k(x,t)\,dt - \frac{1}{\kappa}\int_{r-\kappa}^{r} w_k(x,t)\,dt\right)\psi(x)\,dx\right| \\ &\xrightarrow[\kappa\searrow 0]{} \left|\int_{B_1}(w_k(\cdot,s)-w_k(\cdot,r))\,\psi\,dx\right|\end{aligned}$$

for almost all times $-1<s<r<0$. Putting together the last two formulas, we arrive at

$$\left|\int_{B_1}(w_k(\cdot,s)-w_k(\cdot,r))\,\psi\,dx\right| \le c(n,p,L)\Big((r-s)^{\frac{p-1}{p}}+\frac{1}{k}\Big)\|D\psi\|_{L^\infty(B_1)}$$

for every $\psi\in C_0^\infty(B_1,\mathbb{R}^N)$. Now we fix some $l>\frac{n+2}{2}$ and employ the Sobolev embedding $W^{l,2}(B_1,\mathbb{R}^N)\hookrightarrow W^{1,\infty}(B_1,\mathbb{R}^N)$ and the fact that $C_0^\infty(B_1,\mathbb{R}^N)$ is dense in $W_0^{l,2}(B_1,\mathbb{R}^N)$. We deduce that the above estimate holds for every $\psi\in W_0^{l,2}(B_1,\mathbb{R}^N)$ in the form

$$\left|\int_{B_1}(w_k(\cdot,s)-w_k(\cdot,r))\,\psi\,dx\right| \le c(n,p,L,l)\Big((r-s)^{\frac{p-1}{p}}+\frac{1}{k}\Big)\|\psi\|_{W_0^{l,2}(B_1)}.$$

This implies in particular that for almost every $s\in(-1,0)$ and $h\in(0,|s|)$, we have the estimate

$$\|w_k(\cdot,s)-w_k(\cdot,s+h)\|_{W^{-l,2}(B_1)} \le c(n,p,L,l)\Big(h^{\frac{p-1}{p}}+\frac{1}{k}\Big),$$

and consequently,

$$\int_{-1}^{-h}\|w_k(\cdot,s)-w_k(\cdot,s+h)\|^p_{W^{-l,2}(B_1)}\,ds \le c(n,p,L,l)\Big(h^{p-1}+\frac{1}{k^p}\Big).$$

This, together with the uniform estimates (5.8) and (5.9), enables us to apply a compactness principle of Simon (see [37, Thm. 5]) to the triple of Banach spaces

$$W^{1,p}(B_1,\mathbb{R}^N)\hookrightarrow L^2(B_1,\mathbb{R}^N)\hookrightarrow W^{-l,2}(B_1,\mathbb{R}^N),$$

where in particular the first embedding is compact because of the assumption $p>\frac{2n}{n+2}$. From the cited theorem we infer the compactness

$$w_k\to w \qquad \text{strongly in } L^p(-1,0;L^2(B_1,\mathbb{R}^N)),\text{ as } k\to\infty.$$

Combining this with the property (5.6) of the maps $w_k$, we derive

$$\begin{aligned}\int_{Q_{1/2}} |w_k - w|^2\,dz &\le \\ &\le \Big( \sup_{t\in(-1/4,0)} \int_{B_{1/2}} |w_k - w|^2\,dx \Big)^{1-\frac{p}{2}} \int_{-1/4}^{0} \Big[ \int_{B_{1/2}} |w_k - w|^2\,dx \Big]^{\frac{p}{2}} dt \\ &\xrightarrow[k\to\infty]{} 0\end{aligned}$$

since the first factor on the right-hand side is bounded independently from $k \in \mathbb{N}$ because of (5.6) and the fact that $w$ is smooth in the parabolic interior of $Q_1$, while the second factor tends to zero because of the convergence $w_k \to w$ in the $L^p$-$L^2$-norm. We thus have shown

$$w_k \to w \qquad \text{strongly in } L^2(Q_{1/2}, \mathbb{R}^N), \text{ as } k \to \infty. \tag{5.12}$$

The next step is the construction of $\mathcal{A}_k$-caloric comparison maps. We choose the maps $v_k \in C^0(-1/4, 0; L^2(B_{1/2}, \mathbb{R}^N)) \cap L^2(-1/4, 0; W^{1,2}(B_{1/2}, \mathbb{R}^N))$ as the unique solutions of the Cauchy-Dirichlet problems

$$\begin{cases} \displaystyle\int_{Q_{1/2}} \Big( v_k \cdot \partial_t \varphi - \mathcal{A}_k(Dv_k, D\varphi) \Big)\, dz = 0 & \text{for all } \varphi \in C_0^\infty(Q_{1/2}, \mathbb{R}^N) \\ v_k = w & \text{on } \partial_{\mathrm{par}} Q_{1/2} \end{cases} \tag{5.13}$$

For the existence we refer to [31]. By classical results, the comparison maps satisfy $v_k \in C^\infty(\overline{Q_{1/2}}, \mathbb{R}^N)$, since $w$ is smooth in the parabolic interior of the cylinder $Q_1$.

We claim that $Dv_k \to Dw$ strongly in $L^2(Q_{1/2}, \mathbb{R}^N)$, as $k \to \infty$. Since $v_k$ and $w$ are smooth, we may integrate by parts with respect to the time variable in the equations (5.11) and (5.13). Testing the resulting equations with $\varphi = v_k - w \in C^\infty\big(\overline{Q_{1/2}}, \mathbb{R}^N\big)$, we infer, since $v_k - w$ vanishes on the parabolic boundary of $Q_{1/2}$,

$$\begin{aligned}\int_{Q_{1/2}} \partial_t(v_k - w) \cdot (v_k - w)\,dz + \int_{Q_{1/2}} \mathcal{A}_k(Dv_k - Dw, Dv_k - Dw)\,dz \\ = \int_{Q_{1/2}} (\mathcal{A} - \mathcal{A}_k)(Dw, Dv_k - Dw)\,dz.\end{aligned} \tag{5.14}$$

Furthermore, since

$$\begin{aligned}\int_{Q_{1/2}} \partial_t(v_k - w) \cdot (v_k - w)\,dz &= \frac{1}{2} \int_{-1/4}^{0} \frac{d}{dt} \int_{B_{1/2}} |v_k - w|^2\,dx\,dt \\ &= \frac{1}{2} \int_{B_{1/2}} |v_k(\cdot, 0) - w(\cdot, 0)|^2\,dx \ge 0,\end{aligned}$$

we infer from (5.14) and the properties (5.1) of $\mathcal{A}_k$

$$\lambda \int_{Q_{1/2}} |Dv_k - Dw|^2 \, dz \le \int_{Q_{1/2}} \mathcal{A}_k(Dv_k - Dw, Dv_k - Dw) \, dz$$
$$\le |\mathcal{A}_k - \mathcal{A}| \int_{Q_{1/2}} |Dw| \, |Dv_k - Dw| \, dz.$$

Applying Young's inequality and absorbing one of the resulting terms on the left-hand side, this yields

$$\frac{\lambda}{2} \int_{Q_{1/2}} |Dv_k - Dw|^2 \, dz \le \frac{|\mathcal{A}_k - \mathcal{A}|^2}{2\lambda} \int_{Q_{1/2}} |Dw|^2 \, dz \underset{k\to\infty}{\longrightarrow} 0.$$

Sobolev embedding on the time slices yields furthermore

$$\lim_{k\to\infty} \int_{Q_{1/2}} \big( |v_k - w|^2 + |Dv_k - Dw|^2 \big) \, dz = 0, \tag{5.15}$$

which implies in turn by Lemma 3.1(*iii*)

$$\lim_{k\to\infty} \int_{Q_{1/2}} \big( |v_k - w|^2 + |V(Dv_k) - V(Dw)|^2 \big) \, dz = 0. \tag{5.16}$$

From the equivalence of $V$ and $W$ and the convexity of $|W|^2$, stated in (3.3) respectively (3.4), we infer the following lower semicontinuity property.

$$\fint_{Q_1} \big( 4|w|^2 + |V(Dw)|^2 \big) \, dz \le 4 \fint_{Q_1} \big( |w|^2 + |W(Dw)|^2 \big) \, dz$$
$$\le 4 \liminf_{k\to\infty} \fint_{Q_1} \big( |w_k|^2 + |W(Dw_k)|^2 \big) \, dz.$$

Because of $|W| \le |V|$ by (3.3) and the uniform estimate (5.6), this implies

$$\fint_{Q_1} \big( 4|w|^2 + |V(Dw)|^2 \big) \, dz \le 4.$$

Combining this with (5.16), we deduce

$$\fint_{Q_{1/2}} \big( 4|v_k|^2 + |V(Dv_k)|^2 \big) \, dz \le 2^{n+5} \tag{5.17}$$

for all sufficiently large values of $k \in \mathbb{N}$. From the convergence established in (5.12) and (5.15), we know furthermore

$$\fint_{Q_{1/2}} |w_k - v_k|^2 \, dz \le 2 \fint_{Q_{1/2}} |w_k - w|^2 \, dz + 2 \fint_{Q_{1/2}} |w - v_k|^2 \, dz \le \frac{\varepsilon}{4}$$

for large values of $k \in \mathbb{N}$. Because of (5.17), this estimate is in contradiction to (5.7), since $v_k$ is an $\mathcal{A}_k$-caloric map. This completes the proof. ∎

## 6. Approximate $\mathcal{A}$-caloricity by linearization

**Definition 6.1.** Let $u \in L^2(Q_R(z_0),\mathbb{R}^N)\cap L^p(t_0-\rho^2,t_0;W^{1,p}(B_\rho(x_0),\mathbb{R}^N))$, $A:\mathbb{R}^n\to\mathbb{R}^N$ a linear function and $\ell:\mathbb{R}^n\to\mathbb{R}^N$ an affine function. We define excess functionals

$$\Phi_V(z_0,\rho,A) := \fint_{Q_\rho} |V(Du-A)|^2\,dz \quad\text{and}\quad \Psi_2(z_0,\rho,\ell) := \fint_{Q_\rho}\Big|\frac{u-\ell}{\rho}\Big|^2 dz \tag{6.1}$$

and abbreviate furthermore $\widetilde{\Psi}_2(z_0,\rho,\ell) := \Psi_2(z_0,\rho,\ell)+\rho^{2\beta}$. When the choice of $z_0$ and $\ell$ is clear from the context, we will frequently omit these arguments and write, for example, $\Phi_V(\rho) := \Phi_V(z_0,\rho,\ell)$.

**Lemma 6.2.** *Let $u \in C^0(-T,0;L^2(\Omega,\mathbb{R}^N))\cap L^p(-T,0;W^{1,p}(\Omega,\mathbb{R}^N))$ be a weak solution of* (2.1), *under the general assumptions stated in* Section 2. *Let $\ell:\mathbb{R}^n\to\mathbb{R}^N$ be any affine function with $|\ell(x_0)|+|D\ell|\le M$ and $Q_{2\rho}(z_0)\subset\Omega_T$ a parabolic cylinder with $\rho\in(0,1)$. We define the bilinear form $\mathcal{A}$ on $\mathbb{R}^{Nn}$ by*

$$\mathcal{A}(\xi,\eta) := \big\langle D_\xi a(z_0,\ell(x_0),D\ell)\xi,\eta\big\rangle \qquad \text{for all } \xi,\eta\in\mathbb{R}^{Nn}. \tag{6.2}$$

*Then, the function $u-\ell:(x,t)\mapsto u(x,t)-\ell(x)$ satisfies*

$$\begin{aligned}\fint_{Q_\rho(z_0)} \Big((u-\ell)\cdot\partial_t\varphi - \mathcal{A}(Du-D\ell,D\varphi)\Big)\,dz &\le \\ &\le c_1\Big[\omega_{M+1}\big(\Phi_V(\rho)\big)\Phi_V^{1/2}(\rho)+\Phi_V(\rho)+\Psi_2(\rho)+\rho^\beta\Big]\sup_{Q_\rho(z_0)}|D\varphi| \\ &\le c_2\Big[\omega_{M+1}\big(\widetilde{\Psi}_2(2\rho)\big)\widetilde{\Psi}_2^{1/2}(2\rho)+\Psi_2(2\rho)+\rho^\beta\Big]\sup_{Q_\rho(z_0)}|D\varphi|\end{aligned}$$

*for all test functions $\varphi\in C_0^\infty(Q_\rho(z_0),\mathbb{R}^N)$. The constants $c_1$ and $c_2$ depend only on $M, K_\theta(\cdot), K_D(\cdot), n, N, p, \Lambda$ and $\nu$, and in the case of the natural growth condition* (B2) *additionally on $M_u$, $\Lambda_1(M_u)$ and $\Lambda_2(M_u)$.*

*Again, in the case of controllable growth* (B1) *and systems without $u$-dependence, the estimate holds for all affine maps with $|D\ell|\le M$.*

**Proof.** We note that in both the cases of controllable and of natural growth, the right-hand side of the system (2.1) satisfies an estimate of the form

$$|B(z,u,Du)|\le c(1+|Du|^p) \tag{6.3}$$

with a constant $c$ depending only on the data stated in the lemma. In this proof, we do not need any further information on $B$. Let $\varphi\in C_0^\infty(Q_\rho(z_0),\mathbb{R}^N)$ be a test function with $\sup_{Q_\rho(z_0)}|D\varphi|\le 1$. Because of the homogeneity of the claimed estimate, it suffices to prove it under this constraint on $\varphi$. We use the fact that $u$ is a solution of (2.1) together with the identities

$$\fint_{Q_\rho(z_0)} \langle a(z_0,\ell(z_0),D\ell), D\varphi\rangle\,dz = 0 \qquad\text{and}\qquad \fint_{Q_\rho(z_0)} \ell\cdot\partial_t\varphi\,dz = 0$$

and deduce

$$\begin{aligned}(6.4)\quad \fint_{Q_\rho(z_0)} \Big((u-\ell)\cdot\partial_t\varphi - \mathcal{A}(Du-D\ell,D\varphi)\Big)\,dz =\\ = \fint_{Q_\rho(z_0)} \Big\langle a(z_0,\ell(z_0),Du) - a(z_0,\ell(z_0),D\ell)\\ - D_\xi a(z_0,\ell(z_0),D\ell)(Du-D\ell), D\varphi\Big\rangle\,dz\\ + \fint_{Q_\rho(z_0)} \langle a(z,u,Du) - a(z,\ell,Du), D\varphi\rangle\,dz\\ + \fint_{Q_\rho(z_0)} \langle a(z,\ell,Du) - a(z_0,\ell(z_0),Du), D\varphi\rangle\,dz\\ - \fint_{Q_\rho(z_0)} B(z,u,Du)\cdot\varphi\,dz\\ =: I + II + III - IV.\end{aligned}$$

We will estimate the integrand of $I$ in different ways depending on whether $|Du-D\ell|\le 1$ or $|Du-D\ell|>1$. In the first case, we use (A5) in order to estimate

$$\begin{aligned}&\big|a(z_0,\ell(z_0),Du) - a(z_0,\ell(z_0),D\ell) - D_\xi a(z_0,\ell(z_0),D\ell)(Du-D\ell)\big| =\\ &= \Big|\int_0^1 D_\xi a(z_0,\ell(z_0),D\ell+s(Du-D\ell)) - D_\xi a(z_0,\ell(z_0),D\ell)\,ds\,(Du-D\ell)\Big|\\ &\le 2\Lambda K_D(M+1)\,\omega_{M+1}(|Du-D\ell|^2)\,|Du-D\ell|\\ &\le c(p)\Lambda K_D(M+1)\,\omega_{M+1}\big(2^{\frac{2-p}{2}}|V(Du-D\ell)|^2\big)\,|V(Du-D\ell)|\end{aligned}$$

by Lemma 3.1($i$), since $|Du-D\ell|\le 1$. In the case $|Du-D\ell|>1$, we simply employ the growth assumptions (A2) and (A4) on $a$ and $D_\xi a$ in order to estimate, using $|\ell(z_0)|+|D\ell|\le M$ by assumption,

$$\begin{aligned}\big|a(z_0,\ell(z_0),Du) &- a(z_0,\ell(z_0),D\ell) - D_\xi a(z_0,\ell(z_0),D\ell)(Du-D\ell)\big| \le\\ &\le \Lambda(1+|Du|^2)^{\frac{p-1}{2}} + \Lambda(1+|D\ell|^2)^{\frac{p-1}{2}} + \Lambda K_D(M)|Du-D\ell|\\ &\le c\Lambda(1+M^{p-1}+K_D(M))|Du-D\ell|^p\\ &\le c\Lambda(1+M^{p-1}+K_D(M))|V(Du-D\ell)|^2,\end{aligned}$$

where we used $|Du-D\ell|>1$ and Lemma 3.1($i$) in the last two steps.

Combining the last two estimates and keeping in mind that $|D\varphi| \le 1$, we infer

$$\begin{aligned}|I| &\le c \fint_{Q_\rho(z_0)} \omega_{M+1}\big(2^{\frac{2-p}{2}}|V(Du-D\ell)|^2\big)\,|V(Du-D\ell)|\,dz + c\Phi_V(\rho)\\ &\le c\bigg(\fint_{Q_\rho(z_0)} \omega_{M+1}^2\big(2^{\frac{2-p}{2}}|V(Du-D\ell)|^2\big)\,dz\bigg)^{1/2}\Phi_V(\rho)^{1/2} + c\Phi_V(\rho)\\ &\le c\,\omega_{M+1}\big(2^{\frac{2-p}{2}}\Phi_V(\rho)\big)\Phi_V(\rho)^{1/2} + c\Phi_V(\rho)\end{aligned}$$

by the Cauchy-Schwarz inequality and Jensen's inequality for the concave function $s \mapsto \omega_{M+1}^2(s)$. The concavity of $\omega_{M+1}$ and $\omega_{M+1}(0) = 0$ imply furthermore $\omega_{M+1}(rs) \le r\omega_{M+1}(s)$ for every $r \ge 1$. Thus, we can simplify the above estimate to

$$|I| \le c\,\omega_{M+1}(\Phi_V(\rho))\Phi_V(\rho)^{1/2} + c\Phi_V(\rho). \tag{6.5}$$

Next we turn our attention to the estimate of $II$. Here we use the continuity assumption (A3) on $a$ and (2.2) in order to estimate

$$\begin{aligned}|II| &\le c(p)\Lambda K_\theta(2M+1)\fint_{Q_\rho(z_0)} |u-\ell|^\beta\,(1+|Du|^{p-1})\,dz\\ &\le \frac{c}{|Q_\rho|}\int_{Q_\rho(z_0)\cap\{|Du-D\ell|>1\}} \rho^\beta\Big|\frac{u-\ell}{\rho}\Big|^\beta|Du-D\ell|^{p-1}\,dz\\ &\quad + c(1+M^{p-1})\fint_{Q_\rho(z_0)} \rho^\beta\Big|\frac{u-\ell}{\rho}\Big|^\beta\,dz\\ &\le \frac{c}{|Q_\rho|}\int_{Q_\rho(z_0)\cap\{|Du-D\ell|>1\}} |Du-D\ell|^p\,dz + c\fint_{Q_\rho(z_0)} \Big|\frac{u-\ell}{\rho}\Big|^2\,dz + c\rho^\beta\\ &\le c\big(\Phi_V(\rho)+\Psi_2(\rho)+\rho^\beta\big),\end{aligned} \tag{6.6}$$

where we used Young's inequality, once with exponents $\frac{2p}{2-p\beta} > 1$, $\frac{2}{\beta}$ and $\frac{p}{p-1}$, and once with exponents $\frac{2}{2-\beta} > 1$ and $\frac{2}{\beta}$, and furthermore the facts $\rho \le 1$ and Lemma 3.1($i$). Similarly we estimate, using $\mathrm{d}_{\mathrm{par}}(z,z_0) + |\ell(z)-\ell(z_0)| \le (1+M)\rho$ for all $z \in Q_\rho(z_0)$,

$$\begin{aligned}|III| &\le c(p)\Lambda K_\theta(2M+1)\fint_{Q_\rho(z_0)} (1+M)^\beta\rho^\beta(1+|Du|^{p-1})\,dz\\ &\le \frac{c}{|Q_\rho|}\int_{Q_\rho(z_0)\cap\{|Du-D\ell|>1\}} \rho^\beta|Du-D\ell|^{p-1}\,dz + c\rho^\beta\\ &\le \frac{c}{|Q_\rho|}\int_{Q_\rho(z_0)\cap\{|Du-D\ell|>1\}} |Du-D\ell|^p\,dz + c\rho^\beta\\ &\le c\Phi_V(\rho) + c\rho^\beta.\end{aligned} \tag{6.7}$$

Here we used again Young's inequality, $\rho \leq 1$ and Lemma 3.1($i$) in the last two steps. Finally, we can bound the term $IV$ with the help of (6.3) by

$$|IV| \leq \sup_{Q_\rho(z_0)} |\varphi| \fint_{Q_\rho(z_0)} |B(z,u,Du)|\, dz \leq c\rho \fint_{Q_\rho(z_0)} (1+|Du|^p)\, dz,$$

where we used $\sup_{Q_\rho(z_0)} |\varphi| \leq c\rho \sup_{Q_\rho(z_0)} |D\varphi| \leq c\rho$ in the last step. Using Lemma 3.1($i$) and $|D\ell| \leq M$, we can estimate this term further by

$$\begin{aligned} |IV| &\leq c\rho(1+M^p) + c\rho \fint_{Q_\rho(z_0)\cap\{|Du-D\ell|>1\}} |Du-D\ell|^p\, dz \\ &\leq c\rho(1+M^p) + c\rho \fint_{Q_\rho(z_0)} |V(Du-D\ell)|^2\, dz \\ &= c\rho + c\rho\Phi_V(\rho). \end{aligned} \tag{6.8}$$

Putting together (6.5), (6.6), (6.7) and (6.8), we infer from (6.4)

$$\begin{aligned} &\fint_{Q_\rho(z_0)} \Big((u-\ell)\cdot\partial_t\varphi - \mathcal{A}(Du-D\ell, D\varphi)\Big)\, dz \leq \\ &\qquad \leq c\left[\omega_{M+1}(\Phi_V(\rho))\Phi_V(\rho)^{1/2} + \Phi_V(\rho) + \Psi_2(\rho) + \rho^\beta\right] \end{aligned}$$

for all $\varphi \in C_0^\infty(Q_\rho(z_0), \mathbb{R}^N)$ with $\sup_{Q_\rho(z_0)} |D\varphi| \leq 1$. This implies the first inequality claimed in the lemma. For the second one, we estimate the right-hand side with the help of the Caccioppoli inequality from Theorem 4.1 in the form of (4.2). This yields the bound

$$\begin{aligned} &\omega_{M+1}(\Phi_V(\rho))\Phi_V(\rho)^{1/2} + \Phi_V(\rho) + \Psi_2(\rho) + \rho^\beta \leq \\ &\qquad \leq 2c_0^{1/2}\omega_{M+1}(4c_0\widetilde{\Psi}_2(2\rho))\widetilde{\Psi}_2^{1/2}(2\rho) + (4c_0 + 2^{n+4})(\Psi_2(2\rho) + \rho^\beta). \end{aligned}$$

Using once more the concavity of $\omega_{M+1}$ in the form $\omega_{M+1}(4c_0 s) \leq 4c_0\omega_{M+1}(s)$, we establish the second estimate. ∎

## 7. Estimates for linear problems

In this section, we derive an excess estimate for $\mathcal{A}$-caloric maps, which will be used as comparison maps for the original problem. This estimate has been established in [7, 4] for smooth maps or maps in $L^2(-T,0;W^{1,2}(\Omega,\mathbb{R}^N))$. For our purposes, we need the same estimate under the weaker assumption $u \in L^p(-T,0;W^{1,p}(\Omega,\mathbb{R}^N))$. For this we employ an approximation argument similar to the one in [8]. The precise result reads as follows.

**Lemma 7.1.** *Assume that* $h \in L^1(t_0-R^2, t_0; W^{1,1}(B_R(x_0), \mathbb{R}^N))$ *is* $\mathcal{A}$*-caloric, with* $\mathcal{A}$ *satisfying* (5.1). *Then there holds* $h \in C^\infty(Q_R(z_0), \mathbb{R}^N)$ *and we have the following excess estimate for every* $\rho \in (0, R)$ *and* $\tau \in (0, \frac{1}{2})$.

$$\fint_{Q_{\tau\rho}(z_0)} \left| \frac{h - \ell^{(h)}_{z_0,\tau\rho}}{\tau\rho} \right|^2 dz \le c_3 \tau^2 \fint_{Q_\rho(z_0)} \left| \frac{h - \ell^{(h)}_{z_0,\rho}}{\rho} \right|^2 dz \tag{7.1}$$

*with the affine functions* $\ell^{(h)}_{z_0,r}(x) := (h)_{z_0,r} + (Dh)_{z_0,r}(x - x_0)$ *for* $r = \tau\rho$ *and* $r = \rho$ *and a constant* $c_3 = c_3(n, N, L/\lambda)$.

**Proof.** It only remains to prove that $u \in L^2_{\mathrm{loc}}(t_0-R^2, t_0; W^{1,2}_{\mathrm{loc}}(B_R(x_0), \mathbb{R}^N))$, since under this assumption, a proof of the assertion can be found e.g. in [7, 4]. By a scaling argument, we may assume $z_0 = 0$ and $R = 1$. We fix a radius $S \in (\frac{1}{2}, 1)$ and consider mollifications

$$h_\varepsilon(x) := \varphi_\varepsilon * h(x) \qquad \text{for all } x \in Q_S,$$

where $\varphi \in C^\infty_0(Q_1)$ denotes a standard smoothing kernel and the smoothing radius satisfies $\varepsilon \in (0, 1 - S)$. The maps $h_\varepsilon \in C^\infty(Q_S, \mathbb{R}^N)$ are again $\mathcal{A}$-caloric because of the linearity of the differential equation (5.2). We claim that for all radii $\sigma < S$, they satisfy the uniform estimates

$$\sup_{Q_\sigma} |Dh_\varepsilon| \le c_\sigma \fint_{Q_S} |Dh_\varepsilon|\, dz \le c_\sigma \fint_{Q_1} |Dh|\, dz \tag{7.2}$$

with a constant $c_\sigma$ independent from $\varepsilon$. For this we use the Caccioppoli-type estimate

$$\int_{Q_\sigma} |\partial_t^j D^k Dh_\varepsilon|^2\, dz \le \frac{c}{(\rho - \sigma)^{4j+2k}} \int_{Q_\rho} |Dh_\varepsilon|^2\, dz$$

which holds for all radii $0 < \sigma < \rho < S$ and all $j, k \in \mathbb{N}$. This estimate can be derived for the smooth $\mathcal{A}$-caloric maps $h_\varepsilon$ by standard arguments (see e.g. [7, Chapter 5]). By the Sobolev embedding $W^{n+2,2} \hookrightarrow L^\infty$, the above bound implies for all radii $\sigma < \rho$ with $\rho, \sigma \in (\frac{S}{2}, S)$

$$\begin{aligned}
\sup_{Q_\sigma} |Dh_\varepsilon|^2 &\le c(n) \sum_{j+k \le n+2} \sigma^{4j+2k} \fint_{Q_\sigma} |\partial_t^j D^k Dh_\varepsilon|^2\, dz \\
&\le \frac{c(n)}{(\rho - \sigma)^{4(n+2)}} \fint_{Q_\rho} |Dh_\varepsilon|^2\, dz \\
&\le \frac{1}{2} \sup_{Q_\rho} |Dh_\varepsilon|^2 + \frac{c(n)}{(\rho - \sigma)^{8(n+2)}} \Big( \fint_{Q_S} |Dh_\varepsilon|\, dz \Big)^2,
\end{aligned}$$

where we used the facts $\sigma \le 1$, $\rho - \sigma \le 1$ and Young's inequality. The first term in the last line can be absorbed with the help of Lemma 3.4. This yields the first estimate of the claim (7.2). The second one is a standard estimate for convolutions. Letting $\varepsilon \searrow 0$ in (7.2), we infer that $Dh \in L^\infty(Q_\sigma, \mathbb{R}^{Nn})$ for every $\sigma < S$, and since $S \in (\frac{1}{2}, 1)$ was arbitrary, we have in particular $h \in L^2_{\mathrm{loc}}(-1, 0; W^{1,2}_{\mathrm{loc}}(B_1, \mathbb{R}^N))$, as desired. ∎

## 8. A decay estimate

In the sequel, for a given map $u$ and any $z_0 \in \Omega_T$ and $r > 0$ we will abbreviate $\ell_{z_0,r} = \ell^{(u)}_{z_0,r}$ for the unique affine function $\mathbb{R}^n \to \mathbb{R}^N$ minimizing the functional

$$\ell \mapsto \fint_{Q_r(z_0)} |u - \ell|^2 \, dz.$$

**Lemma 8.1.** *We assume that the hypotheses listed in* Section 2 *are in force and let $M > 0$ and $\alpha \in (\beta, 1)$. Then there are constants $\kappa_M, \tau \in (0,1)$, depending only on $\alpha, M, K_\theta(\cdot), K_D(\cdot), p, n, N, \Lambda, \nu$ and in the case of natural growth* (B2) *additionally on $M_u, \Lambda_1(M_u)$ and $\Lambda_2(M_u)$, such that the following holds. For all $z_0 = (x_0, t_0) \in \mathbb{R}^n \times \mathbb{R}$ and $\rho \in (0,1)$ and any weak solution $u \in C^0(t_0 - \rho^2, t_0; L^2(B_\rho(x_0), \mathbb{R}^N)) \cap L^p(t_0 - \rho^2, t_0; W^{1,p}(B_\rho(x_0), \mathbb{R}^N))$ of* (2.1) *that satisfies*

$$|u_{z_0,\rho}| + |D\ell_{z_0,\rho}| \le M \tag{8.1}$$

*together with the smallness assumption*

$$\Psi_2(z_0, \rho, \ell_{z_0,\rho}) + \rho^\beta \le \kappa_M, \tag{8.2}$$

*we have the excess estimate*

$$\Psi_2(z_0, \tau\rho, \ell_{z_0,\tau\rho}) \le \tau^{2\alpha}\big[\Psi_2(z_0, \rho, \ell_{z_0,\rho}) + \rho^{2\beta}\big].$$

*In the case of systems without $u$-dependence, the assumption* (8.1) *can be replaced by $|D\ell_{z_0,\rho}| \le M$.*

**Proof.** In this proof, we will abbreviate $\Psi_2(r) := \Psi_2(z_0, r, \ell_{z_0,\rho})$, $\widetilde{\Psi}_2(r) := \Psi_2(r) + r^{2\beta}$ and $\Phi_V(r) := \Phi_V(z_0, r, D\ell_{z_0,\rho})$ for any $r \in (0, \rho)$. We let $\mathcal{A}$ be the bilinear form given in (6.2), with the choice $\ell = \ell_{z_0,\rho}$. By the growth and ellipticity conditions (A1) and (A4) it satisfies

$$\mathcal{A}(\xi,\xi) \ge \nu(1+M^2)^{\frac{p-2}{2}}|\xi|^2 \qquad \text{and} \qquad \mathcal{A}(\xi,\eta) \le \Lambda K_D(M)\,|\xi|\,|\eta|$$

for all $\xi, \eta \in \mathbb{R}^{Nn}$. For an $\varepsilon > 0$ to be chosen later, we apply Lemma 5.2 with $\lambda = \nu(1+M^2)^{\frac{p-2}{2}}$ and $L = \Lambda K_D(M)$, which specifies the constant $\delta > 0$, depending at most on $\varepsilon, p, n, N, \nu, M, \Lambda$ and $K_D(M)$. With the constants $c_0$ and $c_2$ determined by Theorem 4.1 and Lemma 6.2, respectively, we define

$$\mu := \max\Big\{\frac{3c_2}{\delta},\ 2^{\frac{n}{2}+2},\ 2\sqrt{c_0}\Big\}\,\widetilde{\Psi}_2^{1/2}(\rho)$$

and choose

$$\kappa_M := \min\Big\{\frac{\delta^2}{9c_2^2},\ \frac{1}{2^{n+4}},\ \frac{1}{4c_0}\Big\} \in (0,1).$$

Assuming (8.2) with this choice of $\kappa_M$, which implies in particular $\widetilde{\Psi}_2(\rho) \le \kappa_M$, we infer

$$\mu \le \max\Big\{\frac{3c_2}{\delta},\ 2^{\frac{n}{2}+2},\ 2\sqrt{c_0}\Big\}\,\kappa_M^{1/2} = 1.$$

By Lemma 6.2, the rescaled map $w := \mu^{-1}(u - \ell_{z_0,\rho})$ satisfies

$$\begin{aligned}
&\fint_{Q_{\rho/2}(z_0)} \Big(w \cdot \partial_t\varphi - \mathcal{A}(Dw, D\varphi)\Big)\,dz \le \\
&\qquad \le \frac{c_2}{\mu}\Big[\omega_{M+1}(\widetilde{\Psi}_2(\rho))\widetilde{\Psi}_2(\rho)^{1/2} + \Psi_2(\rho) + \rho^{\beta}\Big] \sup_{Q_{\rho/2}(z_0)} |D\varphi| \\
&\qquad \le \frac{3c_2}{\mu}\,\widetilde{\Psi}_2^{1/2}(\rho) \sup_{Q_{\rho/2}(z_0)} |D\varphi| \\
&\qquad \le \delta \sup_{Q_{\rho/2}(z_0)} |D\varphi|,
\end{aligned} \tag{8.3}$$

where we subsequently used $\omega_{M+1} \le 1$, $\Psi_2(\rho) \le \kappa_M \le 1$ and the definition of $\mu$ in the last two steps. Since $\mu \in (0,1)$, Lemma 3.1($ii$) implies $|V(Dw)| \le \mu^{-1}|V(Du - D\ell_{z_0,\rho})|$. Consequently, the Caccioppoli inequality established in Theorem 4.1 implies

$$\begin{aligned}
&\sup_{t\in(t_0-\rho^2/4,t_0)} \fint_{B_{\rho/2}(x_0)} \Big|\frac{w(x,t)}{\rho/2}\Big|^2 dx + \fint_{Q_{\rho/2}(z_0)} |V(Dw)|^2\,dz \le \\
&\qquad \le \frac{1}{\mu^2}\Big[\sup_{t\in(t_0-\rho^2/4,t_0)} \fint_{B_{\rho/2}(x_0)} \Big|\frac{u(x,t) - \ell_{z_0,\rho}(x)}{\rho/2}\Big|^2 dx \\
&\qquad\qquad + \fint_{Q_{\rho/2}(z_0)} |V(Du - D\ell_{z_0,\rho})|^2\,dz\Big] \\
&\qquad \le \frac{4c_0}{\mu^2}\widetilde{\Psi}_2(\rho) \le 1
\end{aligned}$$

by the choice of $\mu$. Because of the last two estimates, we may apply the $\mathcal{A}$-caloric approximation Lemma 5.2 on the cylinder $Q_{\rho/2}(z_0)$, which provides an $\mathcal{A}$-caloric function $h \in C^\infty(Q_{\rho/4}(z_0), \mathbb{R}^N)$ with

$$\fint_{Q_{\rho/4}(z_0)} \Big( \Big|\frac{h}{\rho/4}\Big|^2 + |V(Dh)|^2 \Big)\, dz \le 2^{n+5} \tag{8.4}$$

and

$$\fint_{Q_{\rho/4}(z_0)} \Big|\frac{w-h}{\rho/4}\Big|^2\, dz \le \varepsilon. \tag{8.5}$$

We write $\ell^{(h)}_{z_0,r}(x) := (h)_{z_0,r} + (Dh)_{z_0,r}(x-x_0)$ for any $r \in (0,\rho)$. For $\tau \in (0, \frac{1}{8})$ to be chosen small later, we estimate

$$\begin{aligned}\fint_{Q_{\tau\rho}(z_0)} \Big|\frac{w-\ell^{(h)}_{z_0,\tau\rho}}{\tau\rho}\Big|^2\, dz &\le \\ &\le 2(4\tau)^{-n-4} \fint_{Q_{\rho/4}(z_0)} \Big|\frac{w-h}{\rho/4}\Big|^2\, dz + 2 \fint_{Q_{\tau\rho}(z_0)} \Big|\frac{h-\ell^{(h)}_{z_0,\tau\rho}}{\tau\rho}\Big|^2\, dz.\end{aligned}$$

Here, we can estimate by Lemma 7.1

$$\fint_{Q_{\tau\rho}(z_0)} \Big|\frac{h-\ell^{(h)}_{z_0,\tau\rho}}{\tau\rho}\Big|^2\, dz \le c_3\tau^2 \fint_{Q_{\rho/4}(z_0)} \Big|\frac{h-\ell^{(h)}_{z_0,\rho/4}}{\rho/4}\Big|^2\, dz \le c\tau^2,$$

where we used (8.4) in the last step. Combining the last three estimates, we arrive at

$$\fint_{Q_{\tau\rho}(z_0)} \Big|\frac{w-\ell^{(h)}_{z_0,\tau\rho}}{\tau\rho}\Big|^2\, dz \le c(\tau^{-n-4}\varepsilon + \tau^2) \tag{8.6}$$

for every $\tau \in (0, \frac{1}{8})$. Scaling back, we find by the minimizing property of $\ell_{z_0,\tau\rho}$, the definition of $w$, (8.6) and the definition of $\mu$

$$\begin{aligned}\fint_{Q_{\tau\rho}(z_0)} \Big|\frac{u-\ell_{z_0,\tau\rho}}{\tau\rho}\Big|^2\, dz &\le \fint_{Q_{\tau\rho}(z_0)} \Big|\frac{u-\ell_{z_0,\rho}-\mu\ell^{(h)}_{z_0,\tau\rho}}{\tau\rho}\Big|^2\, dz \\ &= \mu^2 \fint_{Q_{\tau\rho}(z_0)} \Big|\frac{w-\ell^{(h)}_{z_0,\tau\rho}}{\tau\rho}\Big|^2\, dz \le c\mu^2(\tau^{-n-4}\varepsilon+\tau^2) \le c(\tau^{-n-4}\varepsilon+\tau^2)\widetilde{\Psi}_2(\rho).\end{aligned}$$

For a given $\alpha \in (\beta, 1)$, we now choose first $\tau \in (0, \frac{1}{8})$ and then $\varepsilon \in (0,1)$ sufficiently small in order to ensure

$$\fint_{Q_{\tau\rho}(z_0)} \Big|\frac{u-\ell_{z_0,\tau\rho}}{\tau\rho}\Big|^2\, dz \le \tau^{2\alpha}\widetilde{\Psi}_2(\rho).$$

This is the claimed estimate. We point out that the choices of $\varepsilon$ and $\tau$ depend only on $n, \delta, c_0, c_2$ and $c_3$, and consequently only on the data listed in the lemma. ∎

**Lemma 8.2.** *Under the assumptions listed in* Section 2*, we let* $M > 2$*. Then there are constants* $\varepsilon_0, \rho_0 \in (0,1)$ *and* $c_4 > 0$*, depending only on* $M, K_\theta(\cdot), K_D(\cdot)$*,* $p$*,* $n, N, \Lambda, \nu$*, and in the case* (B2) *of natural growth additionally on* $M_u, \Lambda_1(M_u)$ *and* $\Lambda_2(M_u)$*, with the following property. For any* $u \in C^0(-T,0;L^2(\Omega,\mathbb{R}^N)) \cap L^p(-T,0;W^{1,p}(\Omega,\mathbb{R}^N))$ *that is a weak solution of the system* (2.1) *and* $Q_R(z_0) \subset\subset \Omega$*, assume that we have the bound*

$$|u_{z_0,R}| + |(Du)_{z_0,R}| < M - 2 \tag{8.7}$$

*and the smallness properties* $R \le \rho_0$ *and*

$$R^{-2} \fint_{Q_R(z_0)} \big|u - u_{z_0,R} - (Du)_{z_0,R}(x - x_0)\big|^2 \, dz < \varepsilon_0^2.$$

*Then there is a neighborhood* $U$ *of* $z_0$ *such that for all* $w \in U$ *and all* $\rho \in (0, \frac{R}{2}]$*, there holds*

$$\fint_{Q_\rho(w)} |Du - (Du)_{w,\rho}|^p \, dz \le c_4 \Big(\frac{\rho}{R}\Big)^{p\beta}.$$

*In case the structure function* $a$ *in the system* (2.1) *does not depend on* $u$*, we may replace the condition* (8.7) *by* $|(Du)_{z_0,R}| < M - 2$*.*

**Proof.** We may fix an arbitrary $\alpha \in (\beta, 1)$, for convenience we choose $\alpha = \frac{1+\beta}{2}$. Furthermore we let $\kappa_M, \tau \in (0,1)$ be the constants determined by Lemma 8.1. We choose the constants $\varepsilon_0, \rho_0 \in (0,1)$ so small that

$$\varepsilon_0^2 + \frac{\rho_0^{2\beta}}{\tau^{2\beta} - \tau^{2\alpha}} + \rho_0^\beta \le \kappa_M \quad \text{and} \quad \varepsilon_0^2 + \frac{\rho_0^{2\beta}}{\tau^{2\beta} - \tau^{2\alpha}} \le \tau^{n+4} \frac{(1 - \tau^\beta)^2}{2(n+1)^2}. \tag{8.8}$$

By the absolute continuity of the integral, we can choose a neighborhood $U$ of $z_0$ such that for all $w \in U$

$$|u_{w,R}| + |(Du)_{w,R}| < M - 2 \tag{8.9}$$

and

$$R^{-2} \fint_{Q_R(w)} \big|u - u_{w,R} - (Du)_{w,R}(x - x_0)\big|^2 \, dz < \varepsilon_0^2. \tag{8.10}$$

First we note that (8.9) and (8.10) imply by (3.6)

$$\begin{aligned} |u_{w,R}| + |D\ell_{w,R}| &\le |u_{w,R}| + |(Du)_{w,R}| \\ &\quad + \bigg[\frac{n(n+2)}{R^2} \fint_{Q_R(w)} \big|u - u_{w,R} - (Du)_{w,R}(x - x_0)\big|^2 \, dz\bigg]^{1/2} \\ &\le M - 2 + \sqrt{n(n+2)}\, \varepsilon_0 \le M - 1 \end{aligned} \tag{8.11}$$

since $\varepsilon_0^2 \le \frac{1}{(n+1)^2} \le \frac{1}{n(n+2)}$ by the choice of $\varepsilon_0$.

Moreover, by the minimizing property of $\ell_{w,R}$ we have

$$\Psi_2(w,R,\ell_{w,R}) \le R^{-2} \fint_{Q_R(w)} \big|u - u_{w,R} - (Du)_{w,R}(x-x_0)\big|^2\,dz \le \varepsilon_0^2. \tag{8.12}$$

Now we apply Lemma 8.1, which is possible since $\varepsilon_0^2 + \rho_0^\beta \le \kappa_M$. We infer, with $\tau \in (0,1)$ determined by the lemma,

$$\Psi_2(w,\tau R,\ell_{w,\tau R}) \le \tau^{2\alpha}\big[\Psi_2(w,R,\ell_{w,R}) + R^{2\beta}\big]. \tag{8.13}$$

In particular,

$$\Psi_2(w,\tau R,\ell_{w,\tau R}) + (\tau R)^\beta \le (\varepsilon_0^2 + \rho_0^{2\beta}) + \rho_0^\beta \le \kappa_M. \tag{8.14}$$

By (3.8), we can estimate

$$\begin{aligned} &|u_{w,\tau R}| + |D\ell_{w,\tau R}| \le \\ &\qquad \le |u_{w,R}| + |D\ell_{w,R}| + \sqrt{2}\,\frac{n+1}{\tau R}\left[\fint_{Q_{\tau R}(w)} |u-\ell_{w,R}|^2\,dz\right]^{1/2} \\ &\qquad \le M - 1 + \sqrt{2}\,\frac{n+1}{\tau^{n/2+2}}\,\Psi_2^{1/2}(w,R,\ell_{w,R}) \le M \end{aligned} \tag{8.15}$$

by (8.11), (8.12) and the choice of $\varepsilon_0$ in (8.8). By (8.14) and (8.15), we may apply Lemma 8.1 again on the ball $Q_{\tau R}(w)$. Continuing in this fashion, we successively derive the estimates

$$\begin{aligned} \Psi_2(w,\tau^k R,\ell_{w,\tau^k R}) &\le \tau^{2\alpha k}\Psi_2(w,R,\ell_{w,R}) + \tau^{2\beta k}R^{2\beta}\sum_{j=1}^{k}\tau^{2\alpha j - 2\beta j} \\ &\le \tau^{2\beta k}\left[\Psi_2(w,R,\ell_{w,R}) + \frac{\tau^{2\alpha}}{\tau^{2\beta}-\tau^{2\alpha}}R^{2\beta}\right] \\ &\le \tau^{2\beta k}\left[\varepsilon_0^2 + \frac{\rho_0^{2\beta}}{\tau^{2\beta}-\tau^{2\alpha}}\right] \end{aligned} \tag{8.16}$$

for every $k \in \mathbb{N}$. This implies in particular, by our choice of the constants in (8.8),

$$\Psi_2(w,\tau^k R,\ell_{w,\tau^k R}) + (\tau^k R)^\beta \le \varepsilon_0^2 + \frac{\rho_0^{2\beta}}{\tau^{2\beta}-\tau^{2\alpha}} + \rho_0^\beta \le \kappa_M. \tag{8.17}$$

Similarly to (8.15) we infer furthermore, using (8.11) and (8.16),

$$\begin{aligned} |u_{w,\tau^k R}| + |D\ell_{w,\tau^k R}| &\le |u_{w,R}| + |D\ell_{w,R}| + \sqrt{2}\,\frac{n+1}{\tau^{n/2+2}}\sum_{j=0}^{k-1}\Psi_2^{1/2}(w,\tau^j R,\ell_{w,\tau^j R}) \\ &\le M - 1 + \sqrt{2}\,\frac{n+1}{\tau^{n/2+2}}\sum_{j=0}^{k-1}\tau^{\beta j}\left[\varepsilon_0^2 + \frac{\rho_0^{2\beta}}{\tau^{2\beta}-\tau^{2\alpha}}\right]^{1/2} \le M \end{aligned} \tag{8.18}$$

by the choice of $\varepsilon_0$ and $\rho_0$ according to (8.8).

During the iteration, the last two estimates (8.17) and (8.18) guarantee that Lemma 8.1 is applicable in each step. Furthermore, the inequality (8.18) ensures that we may apply the Caccioppoli inequality from Theorem 4.1 to infer, for all $k \in \mathbb{N}_0$,

$$\begin{aligned}(8.19)\qquad &\fint_{Q_{\tau^k R/2}(w)} \big|V\big(Du - D\ell_{w,\tau^k R}\big)\big|^2\,dz \le \\ &\qquad\le\ c\big(\Psi_2(w,\tau^k R,\ell_{w,\tau^k R}) + \tau^{2\beta k}R^{2\beta}\big) \le c\tau^{2\beta k}\end{aligned}$$

by the decay estimate (8.16). Combining this with Lemma 3.1(i), we infer, distinguishing the cases $|Du - D\ell_{w,\tau^k R}| \le 1$ and $|Du - D\ell_{w,\tau^k R}| > 1$

$$\begin{aligned}&\fint_{Q_{\tau^k R/2}(w)} |Du - (Du)_{w,\tau^k R/2}|^p\,dz \le \\ &\le c\fint_{Q_{\tau^k R/2}(w)} |Du - D\ell_{w,\tau^k R}|^p\,dz \\ &\le c\fint_{Q_{\tau^k R/2}(w)} \big|V\big(Du - D\ell_{w,\tau^k R}\big)\big|^p\,dz + c\fint_{Q_{\tau^k R/2}(w)} \big|V\big(Du - D\ell_{w,\tau^k R}\big)\big|^2\,dz \\ &\le c(\tau^{p\beta k} + \tau^{2\beta k}) \le c\tau^{p\beta k},\end{aligned}$$

where we applied Hölder's inequality and (8.19) in the penultimate estimate. This is the claim for radii of the form $\rho = \tau^k R/2$, $k \in \mathbb{N}_0$. For general $\rho \in (0, \frac{R}{2}]$, the claim follows by a standard argument. ■

By the characterization of Hölder continuous maps with respect to the parabolic metric by Campanato–Da Prato [23], the last lemma yields a first characterization of the singular set.

**Theorem 8.3.** *Let $u \in C^0(-T,0;L^2(\Omega,\mathbb{R}^N)) \cap L^p(-T,0;W^{1,p}(\Omega,\mathbb{R}^N))$ be a weak solution to* (2.1)*, under the assumptions stated in* Section 2*. Then there is an open set $\Omega^u \subset \Omega_T$ with $Du \in C^{\beta,\beta/2}_{\mathrm{loc}}(\Omega^u,\mathbb{R}^{Nn})$, and the singular set satisfies $\Omega_T \setminus \Omega^u \subset \widetilde{\Sigma}^u_1 \cup \Sigma^u_2$, where*

$$\widetilde{\Sigma}^u_1 := \Big\{z_0 \in \Omega_T \,:\, \liminf_{\rho\searrow 0} \rho^{-2}\fint_{Q_\rho(z_0)} \big|u - u_{z_0,\rho} - (Du)_{z_0,\rho}(x - x_0)\big|^2\,dz > 0\Big\}$$

*and*

$$\Sigma^u_2 := \Big\{z_0 \in \Omega_T \,:\, \limsup_{\rho\searrow 0} \big(|u_{z_0,\rho}| + |(Du)_{z_0,\rho}|\big) = \infty\Big\}.$$

*In the case of systems without u-dependence, the set $\Sigma^u_2$ can be replaced by*

$$\widehat{\Sigma}^u_2 := \Big\{z_0 \in \Omega_T \,:\, \limsup_{\rho\searrow 0} |(Du)_{z_0,\rho}| = \infty\Big\}.$$

Note that it is a priori not clear if $\big|\widetilde{\Sigma}^u_1\big| = 0$. This will be established in the next section.

## 9. A Poincaré type inequality for solutions

In order to deduce the regularity of $u$ outside of a negligible set, we need a Poincaré type inequality for the solution that involves only spatial derivatives. For this we will need some control on the change of the mean values over the time slices with respect to the time. This will be accomplished in the following lemma.

We begin with the observation that in both the cases of the controllable growth condition (B1) and the natural growth condition (B2), we can assume that the inhomogeneity satisfies

$$|B(z,u,\xi)| \le \Lambda_B(1+|\xi|^p) \tag{9.1}$$

for all $z \in \Omega_T$, $u \in \mathbb{R}^N$ and $\xi \in \mathbb{R}^{Nn}$, where the constant $\Lambda_B$ depends on $\Lambda$ and $p$ in the case of controllable growth and on $\Lambda_1(M_u)$ and $\Lambda_2(M_u)$ in the case of natural growth.

**Lemma 9.1.** *For every $M > 0$ there is a constant $c_5$, depending only on $p, n, N, M, K_\theta(\cdot)$ and $K_D(\cdot)$, with the following property. Assume that we have a weak solution $u \in C^0(-T,0;L^2(\Omega,\mathbb{R}^N)) \cap L^p(-T,0;W^{1,p}(\Omega,\mathbb{R}^N))$ of* (2.1)*, where the assumptions* (A2) *to* (A4) *and* (9.1) *are in force. Furthermore let $u_0 \in \mathbb{R}^N$ be a constant and $A \in \mathbb{R}^{Nn}$ a linear map with $|u_0|+|A| \le M$, and let $Q_\sigma(z_0) \subset \Omega_T$ be a parabolic cylinder with $\sigma \in (0,1]$. Then, at all times $r,s \in (t_0-\sigma^2,t_0)$, we have the following estimate for an arbitrary test function $\psi \in W^{1,p}_0 \cap L^\infty(B_\sigma(x_0),\mathbb{R}^N)$.*

$$\begin{aligned}
\Big|\int_{B_\sigma(x_0)} \big[u(\cdot,r)-u(\cdot,s)\big]\cdot\psi\,dx\Big| \le{}& \\
&\le c_5\Lambda|r-s|^{\frac1p}|Q_\sigma|^{\frac{p-1}{p}}\,\|D\psi\|_{L^p}\bigg(\Big[-\!\!\!\!\!\!\int_{Q_\sigma(z_0)}|Du-A|^p\,dz\Big]^{\frac{p-1}{p}}+\min\{1,Y\}\bigg)\\
&\quad+\Lambda_B\|\psi\|_{L^\infty}\,|Q_\sigma|-\!\!\!\!\!\!\int_{Q_\sigma(z_0)}(1+|Du|^p)\,dz,
\end{aligned}$$

*where*

$$Y := Y(\sigma,z_0;u_0) := \sigma^{\tilde\beta}\Big[-\!\!\!\!\!\!\int_{Q_\sigma(z_0)}\Big(1+\Big|\frac{u-u_0}{\sigma}\Big|^2\Big)\,dz\Big]^{\frac{\tilde\beta}{2}}$$

*with $\tilde\beta := \min\{\beta, 2(p-1)/p\}$.*

*In the case of systems without u-dependence, the term $Y$ may be replaced by $\sigma^{\tilde\beta}$.*

**Proof.** Throughout this proof, we assume $z_0 = 0$ for notational convenience and write $c$ for universal constants depending at most on $p, n, N, M, K_\theta(\cdot)$ and $K_D(\cdot)$. We test equation (2.1) with functions of the form $\varphi(x,t) := \zeta_h(t)\psi(x)$, where $\psi \in C_0^\infty(B_\sigma,\mathbb{R}^N)$ is arbitrary and $\zeta_h$ is a Lipschitz continuous cut-off function in time that is defined as follows, for arbitrary times $s < r$ in the interval $(-\sigma^2, 0)$.

$$\zeta_h(t) := \begin{cases} \frac{1}{h}(t-s) & \text{for } s < t \le s+h \\ 1 & \text{for } s+h < t \le r-h \\ -\frac{1}{h}(t-r) & \text{for } r-h < t \le r \end{cases}$$

and $\zeta_h \equiv 0$ elsewhere. Here we assumed that $0 < h < \frac{1}{2}(r-s)$. Testing the parabolic system with this choice of $\varphi$, we find

$$\int_{Q_\sigma} u(z) \cdot \zeta_h'(t) \psi(x)\, dx\, dt = \\ = \int_{Q_\sigma} \big\langle a(z,u,Du), D\psi(x) \big\rangle\, \zeta_h(t)\, dx\, dt - \int_{Q_\sigma} B(z,u,Du) \cdot \psi(x)\, \zeta_h(t)\, dx\, dt,$$

where we abbreviate as usually $z = (x,t)$. Letting $h$ tend to zero, we infer that for all times $s < r$, we have

$$\begin{aligned} \Big| \int_{B_\sigma} (u(x,r) - u(x,s)) \cdot \psi(x)\, dx \Big| \le \\ &\le \|\psi\|_{L^\infty} \int_s^r \int_{B_\sigma} |B(z,u,Du)|\, dx\, dt \\ &\quad + \Big| \int_s^r \int_{B_\sigma} \big\langle a(z,u,Du), D\psi(x) \big\rangle\, dx\, dt \Big| \\ &\le \Lambda_B \|\psi\|_{L^\infty} \int_s^r \int_{B_\sigma} (1 + |Du|^p)\, dx\, dt \\ &\quad + (r-s)^{\frac{1}{p}} \bigg[ \int_s^r \Big| \int_{B_\sigma} \big\langle a(z,u,Du), D\psi(x) \big\rangle\, dx \Big|^{\frac{p}{p-1}} dt \bigg]^{\frac{p-1}{p}}. \end{aligned} \tag{9.2}$$

Here we used (9.1) and Hölder's inequality in the last step. Next we estimate the inner integral in the last line at a fixed time $t$. For this we use the fact

$$\int_{B_\sigma} \big\langle a(0,t,u_0,A), D\psi \big\rangle\, dx = 0,$$

from which we infer

$$\begin{aligned} \Big| \int_{B_\sigma} \big\langle a(x,t,u,Du), D\psi(x) \big\rangle\, dx \Big| \le \\ &\le \int_{B_\sigma} \big| \big\langle a(x,t,u,Du) - a(0,t,u_0,Du), D\psi(x) \big\rangle \big|\, dx \\ &\quad + \int_{B_\sigma} \big| \big\langle a(0,t,u_0,Du) - a(0,t,u_0,A), D\psi(x) \big\rangle \big|\, dx \\ &=: I + II. \end{aligned}$$

The first term can be estimated by the continuity assumption (A3) as

$$
\begin{aligned}
I &\le c(p)\Lambda \int_{B_\sigma\times\{t\}} \theta(|u|+|u_0|,|x|+|u-u_0|)(1+|Du|^{p-1})|D\psi|\,dx \\
&\le c(p)\Lambda \int_{B_\sigma\times\{t\}} \theta(|u|+|u_0|,|x|+|u-u_0|)(1+M^{p-1})|D\psi|\,dx \\
&\quad + c(p)\Lambda \int_{B_\sigma\times\{t\}} \theta(|u|+|u_0|,|x|+|u-u_0|)|Du-A|^{p-1}|D\psi|\,dx \\
&=: I_1+I_2.
\end{aligned}
$$

Here we used the assumption $|A|\le M$ for the last estimate. In order to bound the term $I_1$, we first apply Hölder's inequality and then the fact that $\theta \le 1$ and $\tilde\beta \le \beta$ by definition, which yields

$$
I_1 \le c(p,M)\,\Lambda\,\|D\psi\|_{L^p}\Big(\int_{B_\sigma\times\{t\}} \big[\theta(|u|+|u_0|,|x|+|u-u_0|)\big]^{\frac{\tilde\beta}{\beta}\frac{p}{p-1}}\,dx\Big)^{\frac{p-1}{p}}. \tag{9.3}
$$

For the estimate of $I_2$, we simply use $\theta\le 1$ and Hölder's inequality, which implies

$$
I_2 \;\le\; c(p)\Lambda\|D\psi\|_{L^p}\Big(\int_{B_\sigma\times\{t\}} |Du-A|^p\,dx\Big)^{\frac{p-1}{p}}. \tag{9.4}
$$

The term $II$ will be estimated in different ways depending on whether $|Du-A|\le 1$ or $|Du-A|>1$. In the first case, we observe $|u_0|+|A+\tau(Du-A)|\le |u_0|+|A|+|Du-A|\le M+1$ for all $\tau\in(0,1)$, so that by (A4),

$$
\begin{aligned}
&\big|\big\langle a(0,t,u_0,Du)-a(0,t,u_0,A), D\psi(x)\big\rangle\big| \le \\
&\qquad \le \int_0^1 \big|\big\langle D_\xi a(0,t,u_0,A+\tau(Du-A))(Du-A), D\psi\big\rangle\big|\,d\tau \\
&\qquad \le \Lambda K_D(M+1)|Du-A|\,|D\psi| \le \Lambda K_D(M+1)|Du-A|^{p-1}\,|D\psi|
\end{aligned}
$$

since $|Du-A|\le 1$ and $p<2$. In the case $|Du-A|>1$, we use instead the growth assumption (A2) on $a$, which implies

$$
\begin{aligned}
\big|\big\langle a(0,t,u_0,Du)-a(0,t,u_0,A), D\psi(x)\big\rangle\big| &\le \\
&\le c(p)\Lambda(1+M^{p-1}+|Du|^{p-1})\,|D\psi| \\
&\le c(p)\Lambda(1+M^{p-1})|Du-A|^{p-1}\,|D\psi|.
\end{aligned}
$$

Combining the last two estimates, we arrive at

$$
\begin{aligned}
II &\le c\Lambda \int_{B_\sigma\times\{t\}} |Du-A|^{p-1}\,|D\psi|\,dx \\
&\le c\Lambda\|D\psi\|_{L^p}\Big(\int_{B_\sigma\times\{t\}} |Du-A|^p\,dx\Big)^{\frac{p-1}{p}},
\end{aligned} \tag{9.5}
$$

by Hölder's inequality. Putting together the estimates (9.3), (9.4) and (9.5), we have shown

$$\begin{aligned}
\Big| \int_{B_\sigma} \big\langle a(x,t,u,Du), D\psi(x) \big\rangle \, dx \Big|^{\frac{p}{p-1}} \le \\
&\le c\,\Lambda^{\frac{p}{p-1}} \, \|D\psi\|_{L^p}^{\frac{p}{p-1}} \int_{B_\sigma \times \{t\}} |Du - A|^p \, dx \\
&\quad + c\,\Lambda^{\frac{p}{p-1}} \, \|D\psi\|_{L^p}^{\frac{p}{p-1}} \int_{B_\sigma \times \{t\}} \big[\theta(|u| + |u_0|, |x| + |u-u_0|)\big]^{\frac{\tilde\beta}{\beta}\frac{p}{p-1}} \, dx.
\end{aligned}$$

Integrating with respect to time, we infer from (9.2)

$$\begin{aligned}
\Big| \int_{B_\sigma} (u(x,r) - u(x,s)) \cdot \psi(x) \, dx \Big| \le \\
&\le \Lambda_B \|\psi\|_{L^\infty} \int_s^r \int_{B_\sigma} (1 + |Du|^p) \, dx \, dt \\
&\quad + c\Lambda (r-s)^{\frac{1}{p}} \|D\psi\|_{L^p} \bigg( \Big[ \int_{Q_\sigma} |Du - A|^p \, dz \Big]^{\frac{p-1}{p}} \\
&\qquad + \Big[ \int_{Q_\sigma} \big[\theta(|u| + |u_0|, |x| + |u-u_0|)\big]^{\frac{\tilde\beta}{\beta}\frac{p}{p-1}} \, dz \Big]^{\frac{p-1}{p}} \bigg) \\
&\le \Lambda_B \|\psi\|_{L^\infty} \, |Q_\sigma| \fint_{Q_\sigma} (1 + |Du|^p) \, dz \\
&\quad + c\,\Lambda (r-s)^{\frac{1}{p}} \, \|D\psi\|_{L^p} \, |Q_\sigma|^{\frac{p-1}{p}} \bigg( \Big[ \fint_{Q_\sigma} |Du - A|^p \, dz \Big]^{\frac{p-1}{p}} \\
&\qquad + \Big[ \fint_{Q_\sigma} \big[\theta(|u| + |u_0|, |x| + |u-u_0|)\big]^{\frac{2}{\beta}} \, dz \Big]^{\frac{\tilde\beta}{2}} \bigg)
\end{aligned}$$

by Hölder's inequality, since $\tilde\beta \frac{p}{p-1} \le 2$ by definition. The last integral can be estimated by the property (2.2) of $\theta$, combined with the assumption $|u_0| \le M$. We get

$$\begin{aligned}
\fint_{Q_\sigma} \big[\theta(|u|+|u_0|, |x| + |u-u_0|)\big]^{\frac{2}{\beta}} \, dz \le \\
&\le K_\theta (2M+1)^{\frac{2}{\beta}} \fint_{Q_\sigma} \big(\sigma + |u - u_0|\big)^2 \, dz \\
&= K_\theta (2M+1)^{\frac{2}{\beta}} \, \sigma^2 \fint_{Q_\sigma} \Big( 1 + \Big| \frac{u-u_0}{\sigma} \Big|^2 \Big) \, dz = c Y^{2/\tilde\beta}
\end{aligned}$$

with $Y$ as in the statement of the lemma. Alternatively, we can estimate the integral using the fact $\theta \le 1$. Combining both estimates, we arrive at

$$
\begin{aligned}
&\Big| \int_{B_\sigma} (u(x,r) - u(x,s)) \cdot \psi(x)\, dx \Big| \\
&\quad \le \Lambda_B \|\psi\|_{L^\infty} |Q_\sigma| \fint_{Q_\sigma} (1 + |Du|^p)\, dz \\
&\qquad + c\Lambda (r-s)^{\frac{1}{p}} \|D\psi\|_{L^p} |Q_\sigma|^{\frac{p-1}{p}} \Big( \Big[ \fint_{Q_\sigma} |Du - A|^p\, dz \Big]^{\frac{p-1}{p}} + \min\{1, Y\} \Big)
\end{aligned}
$$

for all $\psi \in C_0^\infty(B_\sigma, \mathbb{R}^N)$ and all times $r < s$ in $(-\sigma^2, 0)$. By an approximation argument, the above estimate also holds for all $\psi \in W_0^{1,p} \cap L^\infty(B_\sigma, \mathbb{R}^N)$. This establishes the claim. ■

**Lemma 9.2** (Poincaré-Sobolev inequality for solutions)**.** *Assume that the function $u \in C^0(-T, 0; L^2(\Omega, \mathbb{R}^N)) \cap L^p(-T, 0; W^{1,p}(\Omega, \mathbb{R}^N))$ weakly solves the system* (2.1)*, under the assumptions stated in* Section 2*, and $Q_{2R}(z_0) \subset \Omega_T$ for some $R \in (0,1)$. Furthermore assume that $A : \mathbb{R}^n \to \mathbb{R}^N$ is a linear map and $|u_{z_0,R}| + |A| \le M$ for some $M > 0$. Then there holds*

$$
\begin{aligned}
&R^{-2} \fint_{Q_R(z_0)} |u - u_{z_0,R} - A(x - x_0)|^2\, dz \le \\
&\qquad \le c_6 \Big( \fint_{Q_{2R}(z_0)} |Du - A|^p\, dz \Big)^{\frac{2}{p}} + c_6 \Big( \fint_{Q_{2R}(z_0)} |Du - A|^p\, dz \Big)^{\frac{2}{p}(p-1)} \\
&\qquad\quad + c_6 R^{2\tilde\beta} \Big( \fint_{Q_{2R}(z_0)} (1 + |Du|^p)\, dz \Big)^2,
\end{aligned}
$$

*where $\tilde\beta = \min\{\beta, 2(p-1)/p\}$, with a constant $c_6$ that depends only on $M, K_\theta(\cdot), K_D(\cdot), n, N, p, \Lambda, \nu$, and in the case of natural growth additionally on $M_u, \Lambda_1(M_u)$ and $\Lambda_2(M_u)$.*

*In the case of a system without u-dependence, the estimate holds without a condition on $|u_{z_0,R}|$.*

**Proof.** We assume again that $z_0 = 0$. Let $\sigma$ and $\rho$ be two radii with $R \le \sigma < \rho \le 2R$. We choose a symmetric smoothing kernel $\psi \in C_0^\infty(B_1)$ with $\int_{B_1} \psi\, dx = 1$ and $\|\psi\|_{L^\infty} + \|D\psi\|_{L^\infty} \le 2(n+2)|B^n|^{-1}$. The rescaled functions $\psi_R(x) := R^{-n}\psi(\frac{x}{R})$ satisfy

$$
\|D\psi_R\|_{L^p} \le c(n) R^{-1-n\frac{p-1}{p}} \le c(n) \sigma^{-1-n\frac{p-1}{p}} \tag{9.6}
$$

$$
\|\psi_R\|_{L^\infty} \le c(n) R^{-n} \le c(n) \sigma^{-n}. \tag{9.7}
$$

In the sequel, we will employ several different notions of means of $u$. For means over the time slice at time $t \in (-R^2, 0)$, we shall write

$$\tilde{u}_R(t) := \fint_{B_R} u(x,t)\,dx \qquad \text{and} \qquad \tilde{u}_R^{\psi}(t) := \int_{B_R} u(x,t)\psi_R(x)\,dx,$$

and for means over the cylinder $Q_R$, we will use the notations

$$u_R := \fint_{Q_R} u(z)\,dz = \fint_{-R^2}^0 \tilde{u}_R(t)\,dt \qquad \text{and} \qquad u_R^{\psi} := \fint_{-R^2}^0 \tilde{u}_R^{\psi}(t)\,dt.$$

We will refer to the means involving $\psi$ as $\psi$-means later. At first, we want to compare the different notions of means. We apply Lemma 9.1 with $u_0 = u_R$, using $|u_R| + |A| \le M$. Since $\psi_R \in C_0^\infty(B_\sigma)$, the lemma yields for any $t \in (-\sigma^2, 0)$

$$\begin{aligned}
(9.8)\quad & \big|\tilde{u}_R^{\psi}(t) - u_R^{\psi}\big| \le \\
&\le \fint_{-R^2}^0 \big|\tilde{u}_R^{\psi}(t) - \tilde{u}_R^{\psi}(s)\big|\,ds \le 4\fint_{-\sigma^2}^0 \big|\tilde{u}_R^{\psi}(t) - \tilde{u}_R^{\psi}(s)\big|\,ds \\
&\le c\sigma^{2/p}\big(\sigma^{n+2}\big)^{\frac{p-1}{p}} \|D\psi_R\|_{L^p}\Big(\Big[\fint_{Q_\sigma} |Du - A|^p\,dz\Big]^{\frac{p-1}{p}} + \min\{1, Y\}\Big) \\
&\quad + c\sigma^{n+2}\|\psi_R\|_{L^\infty}\fint_{Q_\sigma} (1 + |Du|^p)\,dz \\
&\le c\sigma\Big(\Big[\fint_{Q_{2R}} |Du - A|^p\,dz\Big]^{\frac{p-1}{p}} + \min\{1, Y\}\Big) + c\sigma^2\fint_{Q_{2R}} (1 + |Du|^p)\,dz
\end{aligned}$$

by (9.6) and (9.7), where $Y = Y(\sigma, 0; u_R)$ with the notation introduced in Lemma 9.1. Furthermore, we note that by Jensen's and Poincaré's inequalities, there holds

$$\begin{aligned}
\big|u_R - u_R^{\psi}\big|^p &\le \fint_{-R^2}^0 |\tilde{u}_R(t) - \tilde{u}_R^{\psi}(t)|^p\,dt \\
&\le 2^{p-1}\fint_{Q_R} \big(|u(x,t) - \tilde{u}_R(t) - Ax|^p + |u(x,t) - \tilde{u}_R^{\psi}(t) - Ax|^p\big)\,dz \\
(9.9)\qquad &\le c(p,n)R^p\fint_{Q_R} |Du - A|^p\,dz \le c(p,n)\sigma^p\fint_{Q_{2R}} |Du - A|^p\,dz.
\end{aligned}$$

Here we applied the Poincaré inequality separately on the time slices, once for functions with vanishing mean value and once for those with vanishing $\psi$-mean value.

Next, we give an estimate for the behavior of the $\psi$-means under a change of the radius. For this we subsequently use the symmetry of $\psi$, Jensen's inequality and the fact $\sigma \in [R, 2R]$, which leads to the estimate

$$\begin{aligned}\fint_{-\sigma^2}^0 |\tilde u_R^\psi(t) - \tilde u_\sigma^\psi(t)|^p\,dt &= \fint_{-\sigma^2}^0 \left|\int_{B_R} (u(x,t) - \tilde u_\sigma^\psi(t) - Ax)\psi_R(x)\,dx\right|^p dt\\ &\le c\fint_{-\sigma^2}^0 \left(\fint_{B_\sigma} |u(x,t) - \tilde u_\sigma^\psi(t) - Ax|^2\,dx\right)^{\frac p2} dt.\end{aligned}$$

Combining this with the Minkowski inequality, we infer

$$\begin{aligned}&\fint_{-\sigma^2}^0 \left(\fint_{B_\sigma}\left|\frac{u(x,t) - \tilde u_R^\psi(t) - Ax}{\sigma}\right|^2 dx\right)^{\frac p2} dt \le\\ &\qquad\le c\fint_{-\sigma^2}^0 \left(\fint_{B_\sigma}\left|\frac{u(x,t) - \tilde u_\sigma^\psi(t) - Ax}{\sigma}\right|^2 dx\right)^{\frac p2} dt.\end{aligned}$$

Now we slicewise apply the Poincaré inequality for functions with vanishing $\psi$-mean value to the right-hand side of the above estimate, using $p > \frac{2n}{n+2}$. This leads to

$$\fint_{-\sigma^2}^0 \left(\fint_{B_\sigma}\left|\frac{u(x,t) - \tilde u_R^\psi(t) - Ax}{\sigma}\right|^2 dx\right)^{\frac p2} dt \le c\fint_{Q_\sigma} |Du - A|^p\,dz. \tag{9.10}$$

We apply Minkowski's inequality and employ the estimates (9.10), (9.8) and (9.9) with the result

$$\begin{aligned}&\fint_{-\sigma^2}^0 \left(\fint_{B_\sigma}\left|\frac{u(x,t) - u_R - Ax}{\sigma}\right|^2 dx\right)^{\frac p2} dt \le\\ &\qquad\le c\fint_{-\sigma^2}^0 \left(\fint_{B_\sigma}\left|\frac{u(x,t) - \tilde u_R^\psi(t) - Ax}{\sigma}\right|^2 dx\right)^{\frac p2} dt\\ &\qquad\quad + c\fint_{-\sigma^2}^0 \sigma^{-p}|\tilde u_R^\psi(t) - u_R^\psi|^p\,dt + c\sigma^{-p}|u_R^\psi - u_R|^p\\ &\qquad\le c\fint_{Q_{2R}} |Du - A|^p\,dz + c\Big[\fint_{Q_{2R}} |Du - A|^p\,dz\Big]^{p-1}\\ &\qquad\quad + c\min\{1, Y^p\} + cR^p\Big[1 + \fint_{Q_{2R}} |Du|^p\,dz\Big]^p.\end{aligned}$$

We introduce the abbreviations

$$\Psi_2^A(r) := \fint_{Q_r}\left|\frac{u(z) - u_R - Ax}{r}\right|^2 dz$$

and

$$\widehat{\Phi}_p^A(r) := \fint_{Q_r} |Du - A|^p \, dz + \Big[\fint_{Q_r} |Du - A|^p \, dz\Big]^{p-1}$$

for all $r \in [R, 2R]$. With this notation, the last inequality reads

$$\begin{aligned}(9.11)\quad \fint_{-\sigma^2}^0 \Big(\fint_{B_\sigma} \Big|\frac{u(x,t) - u_R - Ax}{\sigma}\Big|^2 dx\Big)^{\frac{p}{2}} dt \le \\ \le c\Big(\widehat{\Phi}_p^A(2R) + \min\{1, Y^p\} + R^p\Big[1 + \fint_{Q_{2R}} |Du|^p \, dz\Big]^p\Big).\end{aligned}$$

From Theorem 4.1 with $\ell(x) = u_R + Ax$ we know, since $|u_R| + |A| \le M$,

$$\begin{aligned}\sup_{-\sigma^2 < s < 0} \fint_{B_\sigma} \Big|\frac{u(x,s) - u_R - Ax}{\sigma}\Big|^2 dx \le \\ (9.12)\quad \le c_0 \frac{(2R)^2}{(\rho - \sigma)^2} \fint_{Q_\rho} \Big|\frac{u(x,t) - u_R - Ax}{\rho}\Big|^2 dz + c_0 (2R)^{2\beta}.\end{aligned}$$

Combining the estimates (9.12) and (9.11), we infer

$$\begin{aligned}\Psi_2^A(\sigma) &= \fint_{Q_\sigma} \Big|\frac{u - u_R - Ax}{\sigma}\Big|^2 dz \\ &\le c \sup_{-\sigma^2 < s < 0} \Big(\fint_{B_\sigma} \Big|\frac{u(x,s) - u_R - Ax}{\sigma}\Big|^2 dx\Big)^{1-\frac{p}{2}} \times \\ &\qquad \times \fint_{-\sigma^2}^0 \Big(\fint_{B_\sigma} \Big|\frac{u(x,t) - u_R - Ax}{\sigma}\Big|^2 dx\Big)^{\frac{p}{2}} dt \\ &\le c \Big(\frac{R^{2-p}}{(\rho - \sigma)^{2-p}} \big[\Psi_2^A(\rho)\big]^{1-\frac{p}{2}} + R^{\beta(2-p)}\Big) \times \\ &\qquad \times \Big(\widehat{\Phi}_p^A(2R) + \min\big\{1, Y^p\big\} + R^p\Big[1 + \fint_{Q_{2R}} |Du|^p \, dz\Big]^p\Big).\end{aligned}$$

Applying Young's inequality with exponents $\frac{2}{2-p}$ and $\frac{2}{p}$, we conclude

$$\begin{aligned}(9.13)\quad \Psi_2^A(\sigma) &\le \frac{1}{2}\Psi_2^A(\rho) + R^{2\beta} \\ &\quad + c\Big(\frac{R}{\rho - \sigma}\Big)^{\frac{2}{p}(2-p)} \Big(\big[\widehat{\Phi}_p^A(2R)\big]^{\frac{2}{p}} + \min\big\{1, Y^2\big\} \\ &\qquad + R^2\Big[1 + \fint_{Q_{2R}} |Du|^p \, dz\Big]^2\Big)\end{aligned}$$

for all $\sigma < \rho$ in the interval $[R, 2R]$. In order to estimate the term $Y$, we apply this estimate with the choice $A = 0$ and infer from Young's inequality, since $p - 1 \in [0, 1]$,

$$\begin{aligned}\Psi_2^0(\sigma) \le& \frac{1}{2}\Psi_2^0(\rho) + R^{2\beta} \\ &+ c\Big(\frac{R}{\rho-\sigma}\Big)^{\frac{2}{p}(2-p)}\bigg(\big[\widehat{\Phi}_p^0(2R)\big]^{\frac{2}{p}} + \Big[1 + ⨍_{Q_{2R}} |Du|^p\,dz\Big]^2\bigg) \\ \le& \frac{1}{2}\Psi_2^0(\rho) + c\Big(\frac{R}{\rho-\sigma}\Big)^{\frac{2}{p}(2-p)}\Big[1 + ⨍_{Q_{2R}} |Du|^p\,dz\Big]^2,\end{aligned}$$

where we used the facts $R \le 1$ and $\frac{R}{\rho-\sigma} \ge 1$ in the last step. Since the above estimate holds for all radii $\sigma < \rho$ in $[R, 2R]$, we may apply Lemma 3.4 with $\sigma_0 = \sigma$ and $\rho_0 = \rho$, which gives

$$⨍_{Q_\sigma} \Big|\frac{u-u_R}{\sigma}\Big|^2 dz = \Psi_2^0(\sigma) \le c\Big(\frac{R}{\rho-\sigma}\Big)^{\frac{2}{p}(2-p)}\Big[1 + ⨍_{Q_{2R}} |Du|^p\,dz\Big]^2$$

for all $\sigma < \rho$ in $[R, 2R]$. We can thus estimate, since $\sigma \le 2R$ and $\frac{R}{\rho-\sigma} \ge 1$,

$$\begin{aligned}Y^2 &= \sigma^{2\tilde\beta}\Big[⨍_{Q_\sigma}\Big(1 + \Big|\frac{u-u_R}{\sigma}\Big|^2\Big)\,dz\Big]^{\tilde\beta} \\ &\le cR^{2\tilde\beta}\Big(\frac{R}{\rho-\sigma}\Big)^{\frac{2}{p}(2-p)\tilde\beta}\Big[1 + ⨍_{Q_{2R}} |Du|^p\,dz\Big]^{2\tilde\beta}.\end{aligned}$$

Plugging this estimate into (9.13) and keeping in mind that $\frac{R}{\rho-\sigma} \ge 1$, $R \le 1$ and $\tilde\beta < 1$, we conclude

$$\begin{aligned}\Psi_2^A(\sigma) \le& \frac{1}{2}\Psi_2^A(\rho) + R^{2\beta} \\ &+ c\Big(\frac{R}{\rho-\sigma}\Big)^{\frac{2}{p}(2-p)(1+\tilde\beta)}\bigg(\big[\widehat{\Phi}_p^A(2R)\big]^{\frac{2}{p}} + R^{2\tilde\beta}\Big[1 + ⨍_{Q_{2R}} |Du|^p\,dz\Big]^2\bigg).\end{aligned}$$

We apply again Lemma 3.4, now with $\sigma_0 = R$ and $\rho_0 = 2R$, and arrive at

$$\begin{aligned}⨍_{Q_R} \Big|\frac{u-u_R-Ax}{R}\Big|^2 dz &= \Psi_2^A(R) \\ &\le c\bigg(\big[\widehat{\Phi}_p^A(2R)\big]^{\frac{2}{p}} + R^{2\tilde\beta}\Big[1 + ⨍_{Q_{2R}} |Du|^p\,dz\Big]^2\bigg),\end{aligned}$$

since $R \le 1$ and $\tilde\beta \le \beta$. Recalling the definition of $\widehat{\Phi}_p^A$, we derive the claim. ■

## 10. Characterization of the singular set

In this section, we will complete the proofs of Theorem 1.1 and the Propositions 1.2 and 1.3.

**Proof of Proposition 1.2.** For the proof of the proposition, we will show that every point $z_0 \in \Omega_T \setminus (\Sigma_1^u \cup \Sigma_2^u)$ is a regular point. First we observe that for any such point $z_0$, by the definition of $\Sigma_1^u$ and $\Sigma_2^u$ we can find a sequence $\rho_k \searrow 0$ with

$$\lim_{k\to\infty} \fint_{Q_{2\rho_k}(z_0)} |Du - (Du)_{z_0,2\rho_k}|^p \, dz = 0 \qquad \text{and}$$
$$M := \lim_{k\to\infty} \big(|u_{z_0,\rho_k}| + |(Du)_{z_0,\rho_k}|\big) < \infty.$$

Consequently, there also holds

$$\lim_{k\to\infty} \fint_{Q_{2\rho_k}(z_0)} |Du - (Du)_{z_0,\rho_k}|^p \, dz = 0, \tag{10.1}$$

and

$$\limsup_{k\to\infty} \fint_{Q_{2\rho_k}(z_0)} |Du|^p \, dz \le \tag{10.2}$$
$$\le 2^{p-1} \lim_{k\to\infty} \fint_{Q_{2\rho_k}(z_0)} |Du - (Du)_{z_0,\rho_k}|^p \, dz + 2^{p-1} \lim_{k\to\infty} |(Du)_{z_0,\rho_k}|^p < \infty.$$

Applying the Poincaré inequality from Lemma 9.2 with $R = \rho_k$ and $A = (Du)_{z_0,\rho_k}$, we thus infer for some $\tilde\beta > 0$ and the function $\varphi(t) := t^{2/p} + t^{2(p-1)/p}$

$$\rho_k^{-2} \fint_{Q_{\rho_k}(z_0)} \big|u - u_{z_0,\rho_k} - (Du)_{z_0,\rho_k}(x - x_0)\big|^2 \, dz \le$$
$$\le c\,\varphi\Big(\fint_{Q_{2\rho_k}(z_0)} |Du - (Du)_{z_0,\rho_k}|^p \, dz\Big) + c\rho_k^{2\tilde\beta} \Big(\fint_{Q_{2\rho_k}(z_0)} (1 + |Du|^p) \, dz\Big)^2$$
$$\xrightarrow[k\to\infty]{} 0,$$

where we applied (10.1) and (10.2) in the last step. Thus, Theorem 8.3 yields that $Du$ is of class $C^{\beta,\beta/2}$ in a neighborhood of $z_0$, as claimed. ■

**Proof of Theorem 1.1.** Lebesgue's differentiation theorem implies $|\Sigma_1^u| = 0 = |\Sigma_2^u|$. Consequently, Proposition 1.2 implies that the singular set is negligible, more precisely $Du \in C^{\beta,\beta/2}_{\mathrm{loc}}(\Omega^u, \mathbb{R}^{Nn})$ for an open set with $|\Omega_T \setminus \Omega^u| = 0$. In order to show that $u \in C^{\alpha,\alpha/2}_{\mathrm{loc}}(\Omega^u, \mathbb{R}^N)$ for every $\alpha \in (0,1)$, we note that for every $w_0 \in \Omega^u$ there is a neighborhood $U$ of $w_0$ and a constant $L = L(w_0) \geq 1$ with $\|Du\|_{C^{\beta,\beta/2}(U)} \leq L$ on $U$. The Poincaré inequality from Lemma 9.2 thus implies for every parabolic cylinder $Q_{2\rho}(z_0) \subset U$ with $\rho \leq 1$

$$\begin{aligned}\rho^{-2} &\fint_{Q_\rho(z_0)} \left|u - u_{z_0,\rho} - (Du)_{z_0,\rho}(x - x_0)\right|^2 dz \leq \\ &\leq c(L\rho^\beta)^2 + c(L\rho^\beta)^{2(p-1)} + c\rho^{2\tilde{\beta}}(1 + L^p)^2 \leq c(1 + L^{2p}),\end{aligned}$$

since $\rho \leq 1$ and $\tilde{\beta} > 0$. From this we infer

$$\begin{aligned}\fint_{Q_\rho(z_0)} &\left|u - u_{z_0,\rho}\right|^2 dz \leq \\ &\leq 2\fint_{Q_\rho(z_0)} \left|u - u_{z_0,\rho} - (Du)_{z_0,\rho}(x - x_0)\right|^2 dz + 2\rho^2 |(Du)_{z_0,\rho}|^2 \\ &\leq c(1 + L^{2p})\rho^2\end{aligned}$$

for all $z_0 \in U$ and all sufficiently small radii $\rho > 0$. By the characterization of Hölder continuous maps by Campanato-Da Prato [23], this implies the claim $u \in C^{\alpha,\alpha/2}_{\mathrm{loc}}(\Omega^u, \mathbb{R}^N)$ for all $\alpha \in (0, 1)$. ■

**Proof of Proposition 1.3.** We choose an arbitrary point $z_0 \in \Omega_T \setminus (S^u_1 \cup S^u_2)$ and claim that $z_0$ is a regular point. By the choice of $z_0$, we can find a sequence $\rho_k \searrow 0$ with

$$\lim_{k\to\infty} \fint_{Q_{2\rho_k}(z_0)} \left|V(Du) - [V(Du)]_{z_0,2\rho_k}\right|^2 dz = 0 \tag{10.3}$$

and

$$\limsup_{k\to\infty} \fint_{Q_{2\rho_k}(z_0)} |V(Du)|^2 dz < \infty.$$

Here we argued similarly as in the derivation of (10.2) for the last estimate. By Lemma 3.1($i$), we know furthermore

$$\limsup_{k\to\infty} \fint_{Q_{2\rho_k}(z_0)} |Du|^p dz < \infty. \tag{10.4}$$

Next we employ Lemma 3.1($iii$), which yields for any $A_k \in \mathbb{R}^{Nn}$

$$|Du - A_k|^p \quad \leq \quad c(1 + |A_k|^2 + |Du|^2)^{\frac{2-p}{2}\frac{p}{2}} |V(Du) - V(A_k)|^p$$

so that by Hölder's inequality,

$$\begin{aligned}\limsup_{k\to\infty} \fint_{Q_{2\rho_k}(z_0)} |Du-(Du)_{z_0,\rho_k}|^p\,dz \le \\ &\le c \limsup_{k\to\infty} \fint_{Q_{2\rho_k}(z_0)} |Du - A_k|^p\,dz \\ &\le c \limsup_{k\to\infty} \Big(1+|A_k|^p + \fint_{Q_{2\rho_k}(z_0)} |Du|^p\,dz\Big)^{\frac{2-p}{2}} \times \\ &\qquad \times \Big(\fint_{Q_{2\rho_k}(z_0)} |V(Du)-V(A_k)|^2\,dz\Big)^{\frac{p}{2}}.\end{aligned}$$

Here, we choose $A_k \in \mathbb{R}^{Nn}$ with $V(A_k) = [V(Du)]_{z_0,2\rho_k}$, which satisfies

$$\limsup_{k\to\infty} |A_k| < \infty \tag{10.5}$$

because of Lemma 3.1($i$) and $z_0 \notin S_2^u$. Plugging this choice of $A_k$ in the above estimate and using (10.5), (10.4) and (10.3), we deduce

$$\lim_{k\to\infty} \fint_{Q_{2\rho_k}(z_0)} |Du-(Du)_{z_0,\rho_k}|^p\,dz = 0. \tag{10.6}$$

In the same fashion as in the proof of Proposition 1.2, we deduce from (10.6) and (10.4) that

$$\lim_{k\to\infty} \rho_k^{-2} \fint_{Q_{\rho_k}(z_0)} \big|u - u_{z_0,\rho_k} - (Du)_{z_0,\rho_k}(x-x_0)\big|^2\,dz = 0.$$

Thus, Lemma 8.2 and the characterization of Hölder continuous maps by Campanato-Da Prato [23] imply that $z_0$ is a regular point. ■

## References


[1] Acerbi, E. and Fusco, N.: Regularity for minimizers of nonquadratic functionals: the case $1 < p < 2$. *J. Math. Anal. Appl.* **140** (1989), no. 1, 115–135.

[2] Beck, L.: Partial regularity for weak solutions of nonlinear elliptic systems: the subquadratic case. *Manuscripta Math.* **123** (2007), no. 4, 453–491.

[3] Bögelein, V.: Partial regularity and singular sets of solutions of higher order parabolic systems. *Ann. Mat. Pura Appl.* **188** (2009), no. 1, 61–122.

[4] Bögelein, V., Duzaar, F. and Mingione, G.: The boundary regularity of non-linear parabolic systems I. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **27** (2010), no. 1, 201–255.

[5] Bögelein, V., Duzaar, F. and Mingione, G.: The boundary regularity of non-linear parabolic systems II. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **27** (2010), no. 1, 145–200.

[6] Bögelein, V., Duzaar, F. and Mingione, G.: The regularity of general parabolic systems with degenerate diffusions. Preprint, 2010.

[7] Campanato, S.: Equazioni paraboliche del secondo ordine e spazi $\mathcal{L}^{2,\theta}(\Omega, \delta)$. *Ann. Mat. Pura Appl.* **73** (1966), 55–102.

[8] Carozza, M., Fusco, N. and Mingione, G.: Partial regularity of minimizers of quasiconvex integrals with subquadratic growth. *Ann. Mat. Pura Appl.* **175** (1998), 141–164.

[9] Carozza, M. and Mingione, G.: Partial regularity of minimizers of quasiconvex integrals with subquadratic growth: the general case. *Ann. Polon. Math.* **77** (2001), no. 3, 219–243.

[10] Carozza, M. and Passarelli, A.: A regularity theorem for minimisers of quasiconvex integrals: The case $1 < p < 2$. *Proc. Roy. Soc. Edinburgh Sect. A* **126** (1996), no. 6, 1181–1199.

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

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

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

[14] De Giorgi, E.: *Frontiere orientate di misura minima.* Seminario di Matematica della Scuola Normale Superiore di Pisa, 1960-61. Editrice Tecnico Scientifica, Pisa, 1961.

[15] De Giorgi, E.: Un esempio di estremali discontinue per un problema variazionale di tipo ellittico. *Boll. Un. Mat. Ital.* **1** (1968), 135–137.

[16] Duzaar, F., Grotowski, J.and Kronz, M.: Regularity of almost minimizers of quasi-convex variational integrals with subquadratic growth. *Ann. Mat. Pura Appl.* **184** (2005), no. 4, 421–448.

[17] Duzaar, F., Kristensen, J. and Mingione, G.: The existence of regular boundary points for non-linear elliptic systems. *J. Reine Angew. Math.* **602** (2007), 17–58.

[18] Duzaar, F. and Mingione, G.: The $p$-harmonic approximation and the regularity of $p$-harmonic maps. *Calc. Var. Partial Differential Equations* **20** (2004), no. 3, 235–256.

[19] Duzaar, F. and Mingione, G.: Regularity for degenerate elliptic problems via $p$-harmonic approximation. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **21** (2004), no. 5, 735–766.

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

[21] Duzaar, F. and Mingione, G.: Harmonic type approximation lemmas. *J. Math. Anal. Appl.* **352** (2009), no. 1, 301–335.

[22] Duzaar, F., Mingione, G. and Steffen, K.: Parabolic systems with polynomial growth and regularity. To appear in *Mem. Amer. Math. Soc.* DOI: 10.1090/S0065-9266-2011-00614-3.

[23] Da Prato, G.: Spazi $\mathcal{L}^{p,\vartheta}(\Omega,\delta)$ e loro proprietà. *Ann. Mat. Pura Appl.* **69** (1965), 383–392.

[24] Giaquinta, M.: *Multiple integrals in the calculus of variations and nonlinear elliptic systems.* Annals of Mathematics Studies **105**. Princeton University Press, Princeton, NJ, 1983.

[25] Giaquinta, M. and Struwe, M.: An optimal regularity result for a class of quasilinear parabolic systems. *Manuscripta Math.* **36** (1981/82), no. 2, 223–239.

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

[27] Hao, W., Leonardi, S. and Nečas, J.: An example of irregular solution to a nonlinear Euler-Lagrange elliptic system with real analytic coefficients. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **23** (1996), no.1, 57–67.

[28] John, O. and Stará, J.: On the regularity of weak solutions to parabolic systems in two spatial dimensions. *Comm. Partial Differential Equations* **23** (1998), no. 7-8, 1159–1170.

[29] Kronz, M.: Partial regularity results for minimizers of quasiconvex functionals of higher order. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **19** (2002), no. 1, 81–112.

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

[31] Lions, J.: *Quelques méthodes de résolution des problèmes aux limites non linéaires.* Dunod, Gauthier-Villars, Paris, 1969.

[32] Misawa, M.: Partial regularity results for evolutional $p$-Laplacian systems with natural growth. *Manuscripta Math.* **109** (2002), no.4, 419–454.

[33] Misawa, M.: Local Hölder regularity of gradients for evolutional $p$-Laplacian systems. *Ann. Mat. Pura Appl.* **181** (2002), no. 4, 389–405.

[34] Naumann, J., Wolf, J. and Wolff, M.: On the Hölder continuity of weak solutions to nonlinear parabolic systems in two space dimensions. *Comment. Math. Univ. Carolin.* **39** (1998), no.2, 237–255.

[35] Nečas, J. and Sverák, V.: On regularity of solutions of nonlinear parabolic systems. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **18** (1991), no. 1, 1–11.

[36] Scheven, C.: Regularity for subquadratic parabolic systems: Higher integrability and dimension estimates. *Proc. Roy. Soc. Edinburgh Sect. A* **140** (2010), 1269–1308.

[37] Simon, J.: Compact sets in the space $L^p(0,T;B)$. *Ann. Mat. Pura Appl.* **146** (1987), 65–96.

[38] Simon, J.: *Theorems on regularity and singularity of energy minimizing maps.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1996.

[39] Stará, J. and John, O.: Some (new) counterexamples of parabolic systems. *Comment. Math. Univ. Carolin.* **36** (1995), no.3, 503–510.

[40] Stará, J., John, O. and Malý, J.: Counterexample to the regularity of weak solution of the quasilinear parabolic system. *Comment. Math. Univ. Carolin.* **27** (1986), no. 1, 123–136.

[41] Struwe, M.: On the Hölder continuity of bounded weak solutions of quasilinear parabolic systems. *Manuscripta Math.* **35** (1981), no. 1-2, 125–145.

[42] Struwe, M.: A counterexample in regularity theory for parabolic systems. *Czechoslovak Math. J.* **34 (109)** (1984), no. 2, 183–188.

[43] Sverák, V. and Yan, X.: Non-Lipschitz minimizers of smooth uniformly convex functionals. *Proc. Natl. Acad. Sci. USA* **99** (2002), no. 24, 15269–15276.

*Recibido*: 5 de noviembre de 2009
*Revisado*: 11 de febrero de 2010

Christoph Scheven
Department Mathematik
Friedrich-Alexander-Universität Erlangen
Bismarckstr. $1\frac{1}{2}$
91054 Erlangen, Germany
scheven@mi.uni-erlangen.de