Arch. Rational Mech. Anal. 210 (2013) 219–267
Digital Object Identifier (DOI) 10.1007/s00205-013-0646-4

Parabolic Systems with p, q-Growth: A Variational Approach

Verena Bögelein, Frank Duzaar & Paolo Marcellini

Communicated by D. Kinderlehrer

Abstract

We consider the evolution problem associated with a convex integrand $f:\mathbb{R}^{Nn} \to [0,\infty)$ satisfying a non-standard p, q-growth assumption. To establish the existence of solutions we introduce the concept of *variational solutions*. In contrast to weak solutions, that is, mappings $u\colon \Omega_T \to \mathbb{R}^n$ which solve

$$\partial_t u - \operatorname{div} Df(Du) = 0$$

weakly in Ω_T, variational solutions exist under a much weaker assumption on the gap $q - p$. Here, we prove the existence of variational solutions provided the integrand f is strictly convex and

$$\tfrac{2n}{n+2} < p \leqq q < p + 1.$$

These variational solutions turn out to be unique under certain mild additional assumptions on the data. Moreover, if the gap satisfies the natural stronger assumption

$$2 \leqq p \leqq q < p + \min\big\{1, \tfrac{4}{n}\big\},$$

we show that variational solutions are actually weak solutions. This means that solutions u admit the necessary higher integrability of the spatial derivative Du to satisfy the parabolic system in the weak sense, that is, we prove that

$$u \in L^q_{\mathrm{loc}}\big(0, T; W^{1,q}_{\mathrm{loc}}(\Omega, \mathbb{R}^N)\big).$$

1. Introduction

In this paper we are interested in the existence and regularity of solutions of parabolic systems with p, q-growth of the type

$$\partial_t u - \operatorname{div} Df(Du) = 0 \qquad \text{in } \Omega_T. \tag{1.1}$$

In the following, Ω denotes a bounded domain in $\mathbb{R}^n$ with $n \geqq 2$. For $T > 0$ we denote by $\Omega_T := \Omega \times (0, T)$ the space-time cylinder over Ω. Points in $\mathbb{R}^{n+1}$ are termed $z = (x, t)$. Differentiation with respect to the spatial variable x or x_i will be denoted by Du, respectively $\partial_{x_i} u$ or u_{x_i}, while $\partial_t u$ or u_t stands for the differentiation with respect to time. By assuming a variational structure for the diffusion term, that is, by writing div $Df(Du)$ for some given integrand $f: \mathbb{R}^{Nn} \to [0, \infty)$ instead of div $a(Du)$ with a general vector-field $a: \mathbb{R}^{Nn} \to \mathbb{R}^{Nn}$, we emphasize that we are interested in variational solutions. Throughout the paper the convex integrand f is assumed to be differentiable and to satisfy a non-standard growth and ellipticity condition; see (2.3) respectively (2.9) below.

Model equations and systems that we consider in this context are, for instance, for some exponents p, q with $\frac{2n}{n+2} < p \leqq q$, parabolic equations of the type

$$\partial_t u - \operatorname{div}\left(|Du|^{p-2}\, Du\right) - \partial_{x_n}\left(|u_{x_n}|^{q-2} u_{x_n}\right) = 0 \qquad \text{in } \Omega_T\,.$$

Here, the integrand is given by the convex function $f(\xi) = \frac{1}{p}|\xi|^p + \frac{1}{q}|\xi_n|^q$. Otherwise we could consider convex functions f such as

$$f(\xi) = |\xi|^p \log(1 + |\xi|)\,,$$

for some $p > \frac{2n}{n+2}$. In this case the convex function f satisfies, for every $\varepsilon > 0$, the growth condition $|\xi|^p \leqq f(\xi) \leqq L_\varepsilon\,(1+|\xi|)^{p+\varepsilon}$. We could also consider functions $f(\xi)$ which do not behave like a power when $|\xi| \to +\infty$. For instance, for $|\xi|$ large, the integrand could be of the type

$$f(\xi) = |\xi|^{a+b\sin(\log\log|\xi|)}\,.$$

A computation shows that such an intergrand $f(\xi)$ is a convex function for $|\xi| \geqq e$ (and therefore it can be extended to all $\xi \in \mathbb{R}^n$ as a convex function on $\mathbb{R}^n$) if a, b are positive real numbers such that $a > 1 + b\sqrt{2}$. In this case our integrand f satisfies the bounds

$$|\xi|^p \leqq f(\xi) \leqq L\,(1 + |\xi|)^q,$$

with $p = a - b$ and $q = a + b$. Of course the associated parabolic differential equation, again, has the form in (1.1) with f from above.

The stationary problem corresponding to (1.1) has been studied extensively in the past. Two papers [22,23] of the third author have been the starting point. In these papers one possible approach was demonstrated to attack problems with a non-standard p, q-growth condition. The idea behind this approach is to assume initially that minimizers, respectively weak solutions to the Euler–Lagrange equation (or more generally of weak solutions to an elliptic equation of the form div $a(x, Du) = 0$ in Ω) admit a gradient in the smaller energy space determined by the growth condition from above, that is, that $Du \in L^q_{\mathrm{loc}}$. We call such solutions weak (energy) solutions. The assumption $Du \in L^q_{\mathrm{loc}}$ allows the application of a Moser iteration scheme in order to obtain local bounds for $\|Du\|_{L^\infty}$ in terms of the L^q-norm of Du. This is possible, provided the gap $q - p$ is not too large. In a second step (via an interpolation argument) the Lipschitz-bound is improved in such a way that the L^q-norm on the right-hand side is replaced by the L^p-norm of Du. In this step the assumption concerning the gap $q - p$ has to be sharpened. These a priori estimates are then used to construct weak solutions to problems with non-standard p, q-growth conditions. This is achieved by considering regularized problems, that is, by adding the vector-field $\varepsilon|\xi|^{q-2}\xi$ to $a(x, \xi)$ respectively $Df(\xi)$. The associated Dirichlet-problem admits a solution $u_\varepsilon \in L^q$, and therefore fulfills the a proiri estimate. It is feasible that the solutions u_ε of the regularized problems sub-converge to a $W^{1,\infty}_{\mathrm{loc}}$- solution of the original p, q-growth problem. For more details we refer to [10,11,22–26]. For parabolic equations of the form $\partial_t u - \operatorname{div} a(x, Du) = 0$ in a space time cylinder Ω_T this approach exhibits a natural analogue; via so called weak energy solutions, that is, functions $u \in L^q_{\mathrm{loc}}{-}W^{1,q}_{\mathrm{loc}} \cap C^0{-}L^2$, it is possible to have all terms defined in the weak formulation. Then, the Moser iteration scheme yields a sup-estimate for the spatial gradient in terms of the local L^q-norm on parabolic cylinders, and further—again by an interpolation argument—by the local L^p-norms. For this, of course, one has to assume a certain smallness assumption for the gap $q - p$. Having the a priori estimates for weak energy solutions available, again a regularization procedure by considering solutions u_ε of the regularized parabolic equation $\partial_t v - \operatorname{div}(\varepsilon|Dv|^{q-2}Dv + a(x, Du)) = 0$ leads to a solution u of the original parabolic p, q-growth problem, which satisfies a sup-estimate for the spatial gradient exactly as the approximating functions. This approach was successfully carried out in [5]; see also [18] for gradient estimates for bounded solutions to certain anisotropic parabolic equations.

In the elliptic framework, a second approach was introduced in [15]. This approach has its origin in general existence results for variational functionals of the form

$$F(u) := \int_\Omega f(Du)\,\mathrm{d}x$$

for a convex integrand $f : \mathbb{R}^{Nn} \to [0, \infty)$ satisfying a non-standard p, q-growth condition as in (2.9), for example. Existence of minimizers can be shown by the Direct Method of the Calculus of Variations, assuming very mild assumptions on the integrand. The minimizing property has to be understood in the sense that $u \in W^{1,1}$ minimizes the variational integral F if and only if $F(u) < \infty$ and $F(u) \leqq F(v)$ for any $v \in W^{1,1}$ with $u - v \in W^{1,1}_0$. Of course minimizers admit a gradient in L^p by the coercivity of the integrand. The main concern is then to establish that minimizers, in fact, admit a gradient in L^q_{loc}. This is achieved by testing the Euler–Lagrange system with finite differences of u, which leads to a certain kind of fractional differentiability of Du. (Note that at this stage it is not at all clear that minimizers fulfill the Euler–Lagrange system and therefore one has to perform an approximation procedure.) Then, by fractional Sobolev embeddings, this yields higher integrability of Du, and this in turn can be used to improve the fractional differentiability of Du. By a finite iteration this leads to the desired higher integrability of Du. As mentioned before, this procedure has to be combined with an approximation scheme by considering $\varepsilon|\xi|^q + f(\xi)$ instead of $f(\xi)$. Minimizers of the regularized functionals are of class $W^{1,q}$ and satisfy the Euler–Lagrange system. Therefore, the above scheme is applicable to these minimizers, and when $\varepsilon \downarrow 0$ they sub-converge to a minimizer of the functional F. If the integrand is strictly convex, uniqueness is known and yields that the original minimizer also satisfies the local L^q-estimate for the gradient. The higher integrability, however, can only be derived if the gap $q - p$ is small enough. For more details with respect to this approach we refer to [8,15,19,29,30]. Moreover, for existence of solutions of a variational inequality with $p(x,t)$-growth, we refer to [27] and for the self-improving property of the integrability (that is, the higher integrability) of the spatial gradient of solutions of parabolic systems with $p(x,t)$-growth, to [1,4, 33].

The aim of the present paper is also to develop a variational approach in the parabolic setting in the spirit of a paper by Lichnevsky and Temam [17], in which the concept of variational solutions to the evolutionary minimal surface equation has been developed. The advantage of variational solutions stems from the fact that existence can be established under very mild assumptions on the convex integrand. However, in the evolutionary case the proof of existence of variational solutions is not immediate, since we cannot apply the Direct Method of the Calculus of Variations; in particular, we cannot use a minimizing sequence. After having established the existence of variational solutions, the second step is to show that these variational solutions are, in fact, weak energy solutions of the associated parabolic system if the gap $q - p$ is sufficiently small. Thus, the main effort is to prove the higher integrability property $Du \in L^q_{\mathrm{loc}}$ of the spatial gradient.

2. Results

As we explained in the introduction, in the case of stationary variational integrals with a non-standard p, q-growth condition, minimizers are already defined in the Sobolev space $W^{1,p}$. Therefore, the variational approach yields minimizers which may not satisfy the Euler–Lagrange system. Under suitable smallness assumptions on the gap $q - p$, however, it is possible to show that minimizers belong to $W^{1,q}_{\mathrm{loc}}$, which guarantees that minimizers also solve the Euler–Lagrange system. The parabolic analogue of this elliptic variational approach is, to our knowledge, not yet established. However, it would be a natural approach, since such an approach would lead to *parabolic minimizers* or *variational solutions* in the space $L^p\!-\!W^{1,p}$. Such variational solutions might not solve the parabolic system associated with the variational integral, since a priori the L^q_{loc}-regularity for the spatial gradient Du, needed for the derivation of the *parabolic Euler–Lagrange system*, might fail to hold. For these reasons it would be convenient and natural to ask, in the parabolic setting, for such a weaker notion of solutions, not requiring $Du \in L^q_{\mathrm{loc}}$. To build up a setting in which the existence of variational solutions can be established will be the main subject of the first part of the paper. The second part is then devoted to showing—imposing a stronger hypothesis for the gap $q - p$—that variational solutions are, in fact, higher integrable and thus solve the Euler–Lagrange system.

2.1. Existence and Uniqueness of Variational Solutions

Here our aim is to establish an existence and uniqueness result for parabolic Cauchy–Dirichlet problems admitting a variational structure of the type

$$\begin{cases} \partial_t u - \operatorname{div} Df(Du) = 0 & \text{in } \Omega_T, \\ u = g & \text{on } \partial_{\mathcal{P}}\Omega_T, \end{cases} \tag{2.1}$$

where $u\colon \Omega_T \subset \mathbb{R}^{n+1} \to \mathbb{R}^N$ with $n \geqq 2$ and $N \geqq 1$, can be a vector valued function with values in Euclidean N-space $\mathbb{R}^N$ and where $\partial_{\mathcal{P}}\Omega_T := [\partial\Omega \times (0, T)] \cup [\overline{\Omega} \times \{0\}]$ denotes the parabolic boundary of Ω_T. We assume that $f\colon \mathbb{R}^{Nn} \to \mathbb{R}_+$ is an integrand of class C^1 and that there exist two growth exponents p, q with

$$\tfrac{2n}{n+2} < p < q < p + 1, \tag{2.2}$$

such that f and Df fulfill the following growth and monotonicity conditions:

$$\begin{cases} 0 \leqq f(\xi) \leqq L\,(1 + |\xi|)^q, \\ \langle Df(\xi) - Df(\eta), \xi - \eta\rangle \geqq \nu\,\big(\mu^2 + |\xi|^2 + |\eta|^2\big)^{\frac{p-2}{2}} |\xi - \eta|^2, \end{cases} \tag{2.3}$$

whenever $\xi,\ \eta \in \mathbb{R}^{Nn}$ and for some $0 < \nu \leqq 1 \leqq L$ and $\mu \in [0, 1]$. Note that the lower bound on p, that is, $p > \frac{2n}{n+2}$, is a typical assumption in the regularity theory for non-linear parabolic equations and systems, see [12, Chapter V, Sections 3 and 5]. At this point it is worth mentioning that assumption $(2.3)_2$ implies the strict convexity of the integrand f. In turn, the convexity of f and the growth assumption $(2.3)_1$ on f imply the following growth property of Df:

$$|Df(\xi)| \leqq c(q)L\,(1 + |\xi|)^{q-1},$$

whenever $\xi \in \mathbb{R}^{Nn}$, cf. [22, Lemma 2.1]. For the boundary data g we suppose that the following regularity assumptions hold true:

$$\begin{cases} g \in L^{p'(q-1)}\big(0, T; W^{1,p'(q-1)}(\Omega, \mathbb{R}^N)\big) \cap C^0\big([0, T]; L^2(\Omega, \mathbb{R}^N)\big) \text{ with} \\ g_t \in L^{p'}\big(0, T; W^{-1,p'}(\Omega, \mathbb{R}^N)\big), \end{cases} \tag{2.4}$$

where, as usual, $p' := \frac{p}{p-1}$ denotes the Hölder conjugate of p. We note that $p'(q-1) > q$. In the following—despite a slight abuse of notation—by $u \in L^p(0, T; W^{1,p}_g(\Omega, \mathbb{R}^N))$ we mean that $u - g \in L^p(0, T; W^{1,p}_0(\Omega, \mathbb{R}^N))$.

To give the precise definition of a variational solution we need to introduce a weaker notion of *continuity with respect to time* than the usual one, which is used in the definition of weak solutions for standard p-growth problems. These solutions are continuous in time as mappings from $[0, T]$ to $L^2(\Omega, \mathbb{R}^N)$. Here, we need the following weaker type of continuity with respect to time.

Definition 2.1. Let X be a Banach space. A function $u \in L^\infty(0, T; X)$ belongs to the function space $C_w([0, T]; X)$ of weakly continuous functions from $[0, T]$ to X if $u(\cdot, t) \in X$ for any $t \in [0, T]$ and

$$t \mapsto \langle \psi, u(t)\rangle_X \ \text{ is continuous for any } \psi \in X'.$$

Here, $\langle\cdot,\cdot\rangle_X$ denotes the duality pairing between X' and X.

In the following definition we introduce the concept of a *variational solution* to the Cauchy–Dirichlet problem (2.1). Here we follow an idea by Lichnewsky and Temam [17], which has been used for the evolution problem for parametric minimal surfaces.

Definition 2.2. Suppose that $f : \mathbb{R}^{Nn} \to \mathbb{R}_+$ is an integrand of class C^1 satisfying the growth and monotonicity assumptions from (2.3). Furthermore, assume that the Cauchy–Dirichlet datum g fulfills (2.4). We identify a map

$$u \in L^p\big(0, T; W^{1,p}_g(\Omega, \mathbb{R}^N)\big) \cap C_w\big([0, T]; L^2(\Omega, \mathbb{R}^N)\big)$$

as a *variational solution of the Cauchy–Dirichlet problem* (2.1) if and only if $u(\cdot, 0) = g(\cdot, 0)$ and, further, the variational inequality

$$\begin{aligned}\int_0^\tau \langle v_t, v - u\rangle_{W^{1,p}(\Omega,\mathbb{R}^N)}\,\mathrm{d}t &+ \int_{\Omega_\tau} [f(Dv) - f(Du)]\,\mathrm{d}z \\ &\geqq \tfrac{1}{2}\|(v-u)(\cdot,\tau)\|^2_{L^2(\Omega)} - \tfrac{1}{2}\|(v-g)(\cdot,0)\|^2_{L^2(\Omega)} \end{aligned} \tag{2.5}$$

holds true, whenever $v \in L^p(0, T; W^{1,p}_g(\Omega, \mathbb{R}^N))$ with $v_t \in L^{p'}(0, T; W^{-1,p'}(\Omega, \mathbb{R}^N))$ and $\tau \in (0, T]$. □

It is worthwhile to note that a variational solution belonging to the parabolic space $L^q_{\mathrm{loc}}(0, T; W^{1,q}_{\mathrm{loc}}(\Omega, \mathbb{R}^N))$ is, in fact, a weak solution. Hence, the concept of variational solutions coincides with the classical one of a weak solution once the natural higher integrability is established; see Theorem 2.8. We should also mention that the testing functions $v \in L^p(0, T; W^{1,p}_g(\Omega, \mathbb{R}^N))$ with $v_t \in L^{p'}(0, T; W^{-1,p'}(\Omega, \mathbb{R}^N))$ in Definition 2.2 are of class $C^0([0, T]; L^2(\Omega, \mathbb{R}^N))$; this is a consequence of the embedding from [31, Chapter III, Proposition 1.2]. In this framework we can prove the existence of variational solutions. This is the content of

Theorem 2.3. *Suppose that $f : \mathbb{R}^{Nn} \to \mathbb{R}_+$ is an integrand of class C^1 satisfying* (2.2) *and* (2.3) *and that the Cauchy–Dirichlet datum g fulfills* (2.4). *Then, there exists a variational solution*

$$u \in L^p\big(0, T; W^{1,p}_g(\Omega, \mathbb{R}^N)\big) \cap C_w\big([0, T]; L^2(\Omega, \mathbb{R}^N)\big)$$

of the parabolic system (2.1) *with $u(\cdot, 0) = g(\cdot, 0)$.*

In contrast to the elliptic setting, uniqueness of evolutionary variational solutions is not completely obvious. The reason for this is the presence of the time derivative in the variational inequality (2.5) and the fact that the variational solution u does not necessarily admit a time derivative u_t in the appropriate parabolic space $L^{p'}(0, T; W^{-1,p'}(\Omega, \mathbb{R}^N))$. This prevents us from testing the variational inequality with the variational solution u itself. Nevertheless, if we assume that the Cauchy–Dirichlet datum admits a slightly stronger regularity condition (2.6) (note that $p/(p + 1 - q) > p'(q - 1) > q$), we obtain the following uniqueness result for variational solutions. The precise result is the following one.

Theorem 2.4. *Suppose that $f : \mathbb{R}^{Nn} \to \mathbb{R}_+$ is of class C^1 satisfying (2.2) and (2.3) and that g satisfies*

$$\begin{cases} g \in L^{\frac{p}{p+1-q}}\big(0, T; W^{1,\frac{p}{p+1-q}}(\Omega, \mathbb{R}^N)\big) \cap C^0\big([0, T]; L^2(\Omega, \mathbb{R}^N)\big), \\ g_t \in L^{p'}\big(0, T; W^{-1,p'}(\Omega, \mathbb{R}^N)\big). \end{cases} \tag{2.6}$$

Then, there exists a unique variational solution

$$u \in L^p\big(0, T; W_g^{1,p}(\Omega, \mathbb{R}^N)\big) \cap C_w\big([0, T]; L^2(\Omega, \mathbb{R}^N)\big)$$

of the parabolic system (2.1) with $u(\cdot, 0) = g(\cdot, 0)$.

2.2. Existence of Weak Solutions

In this chapter we restrict our considerations to the case of integrands f obeying a p-growth condition from below with $p \geqq 2$. We believe that results similar to the ones stated below also hold true for singular case $\frac{2n}{n+2} < p < 2$. However, since the proofs are quite technical we will focus our attention on the degenerate case $p \geqq 2$ only. We consider parabolic Cauchy–Dirichlet problems of the type

$$\begin{cases} \partial_t u - \operatorname{div} Df(Du) = 0 & \text{in } \Omega_T, \\ u = g & \text{on } \partial_{\mathcal{P}}\Omega_T, \end{cases} \tag{2.7}$$

where $f : \mathbb{R}^{Nn} \to \mathbb{R}_+$ is a C^2-integrand on $\mathbb{R}^{Nn}$. Further, we assume that for given fixed growth exponents p, q with

$$2 \leqq p < q \quad \text{and} \quad q - p < \min\big\{1\,,\, \tfrac{4}{n}\big\}, \tag{2.8}$$

the integrand f fulfills the following p, q-growth and ellipticity conditions:

$$\begin{cases} |\xi|^p \leqq f(\xi) \leqq L\,(1 + |\xi|)^q, \\ |D^2 f(\xi)| \leqq L\,(1 + |\xi|)^{q-2}, \\ \langle D^2 f(\xi)\eta, \eta\rangle \geqq \nu\,|\xi|^{p-2}|\eta|^2, \end{cases} \tag{2.9}$$

whenever $\xi, \eta \in \mathbb{R}^{Nn}$. In this chapter we can use a stronger notion of solution; these solutions could be termed weak energy solutions, since they obey the integrability property $u \in L^q_{\text{loc}}(0, T; W^{1,q}_{\text{loc}}(\Omega, \mathbb{R}^N))$, which makes the integral $\int_{\Omega_T} \langle Df(Du), D\varphi\rangle \,\mathrm{d}z$ well defined in the weak formulation. This notion of solution has already been used by the authors [5] in the case of parabolic equations with p, q-growth.

Definition 2.5. A map

$$u \in L^p\big(0, T; W_g^{1,p}(\Omega, \mathbb{R}^N)\big) \cap L^q_{\text{loc}}\big(0, T; W^{1,q}_{\text{loc}}(\Omega, \mathbb{R}^N)\big) \cap C_w\big([0, T]; L^2(\Omega, R^N)\big),$$

with $u(\cdot, 0) = g(\cdot, 0)$, is termed a *weak solution of the parabolic system* (2.7) if and only if

$$\int_{\Omega_T} u \cdot \varphi_t - \langle Df(Du), D\varphi\rangle \,\mathrm{d}z = 0$$

holds true whenever $\varphi \in C_0^\infty(\Omega_T, \mathbb{R}^N)$. $\square$

For weak solutions, that is, solutions in the sense of Definition 2.5, we have the following existence result.

Theorem 2.6. *Suppose that the integrand* $f\colon \mathbb{R}^{Nn} \to \mathbb{R}_+$ *is of class* C^2 *satisfying* (2.8) *and* (2.9) *and, further, that* g *is as in* (2.4). *Then, there exists a weak solution*

$$u \in L^p\big(0,T;W^{1,p}_g(\Omega,\mathbb{R}^N)\big)\cap L^q_{\mathrm{loc}}\big(0,T;W^{1,q}_{\mathrm{loc}}(\Omega,\mathbb{R}^N)\big)\cap C_w\big([0,T];L^2(\Omega,\mathbb{R}^N)\big),$$

with $u(\cdot,0) = g(\cdot,0)$ *of the parabolic system* (2.7). *Further, for any cylinder* $Q_R(z_o) \Subset \Omega_T$ *the quantitative estimate*

$$\begin{aligned}\int_{Q_{R/2}(z_o)} &|Du|^q \,\mathrm{d}z \\ &\leqq c\left[\sup_{t\in(t_o-R^2,t_o)}\int_{B_R(x_o)}|u(\cdot,t)|^2\,\mathrm{d}x + \int_{Q_R(z_o)}\big(|Du|^p+|u|^p+1\big)\,\mathrm{d}z\right]^{\chi}\end{aligned} \tag{2.10}$$

holds true for a constant $c = c(n,\nu,L,p,q,R)$ *and an exponent* $\chi = \chi(n,q-p) > 1$.

Remark 2.7. Here, we make some comments on the appearing exponents. Firstly, the upper bound on q in (2.8)—that is, $q < p+\frac{4}{n}$—is exactly the one leading to the L^∞ bound for the gradient Du in the case of one single equation, see [5, Theorem 1.2]. Moreover, the exponent $p+\frac{4}{n}$ is exactly the gain of integrability one gets in parabolic standard p-growth problems by the use of second spatial derivatives and the Sobolev embedding, see [14, Lemma 5.4]. On the other hand, comparing the elliptic bound in [23, Theorem 2.1], that is,

$$q < \frac{np}{n-2} \equiv p + \frac{2p}{n-2}\,,$$

with the parabolic bound

$$2 \leqq p \leqq q < p + \frac{4}{n} \equiv p + \frac{2p}{(n+2)-2}\cdot\frac{2}{p},$$

one must replace n by $n+2$ (which is due to the different scaling in time) and must take into account that the parabolic deficit $\frac{2}{p}$ shows up. In this respect, the parabolic restriction coincides with the elliptic one. □

2.3. Regularity of Variational Solutions

It is not difficult to show that any weak solution is also a variational solution (see Section 4.4). On the other hand, under the weak assumptions of Theorems 2.3 and 2.4 it cannot be expected, in general, that variational solutions obey the necessary integrability properties to be weak solutions in the sense of Definition 2.5. Indeed, this would require that the spatial gradient Du belongs to L^q_{loc}, which is, in general, not true; already in the elliptic setting this might be false, see Section 6 in [23]. To obtain the necessary higher integrability for variational solutions, we need to assume stronger regularity for the integrand as in (2.9) and we need to reduce the gap between p and q as in (2.8). Under these further assumptions we obtain the following regularity result for variational solutions.

Theorem 2.8. *Suppose that $f : \mathbb{R}^{Nn} \to [0, \infty)$ is of class C^2 satisfying (2.8) and (2.9) and that g satisfies (2.6). Then, any variational solution*

$$u \in L^p\big(0, T; W_g^{1,p}(\Omega, \mathbb{R}^N)\big) \cap C_w\big([0, T]; L^2(\Omega, \mathbb{R}^N)\big),$$

with $u(\cdot, 0) = g(\cdot, 0)$ of the parabolic system (2.1), is also a weak solution in the sense of Definition 2.5. In particular, we have

$$Du \in L^q_{\mathrm{loc}}(\Omega_T, \mathbb{R}^{Nn})$$

and the quantitative estimate (2.10) holds true.

3. Preliminaries and Notations

3.1. Notations

The spaces $L^p(\Omega, \mathbb{R}^N)$, $W^{1,p}(\Omega, \mathbb{R}^N)$ and $W_0^{1,p}(\Omega, \mathbb{R}^N)$ denote the usual Lebesgue and Sobolev spaces. For fixed $g \in L^p(0, T; W^{1,p}(\Omega, \mathbb{R}^N))$ we denote by $L^p(0, T; W_g^{1,p}(\Omega, \mathbb{R}^N))$ the affine space $g + L^p(0, T; W_0^{1,p}(\Omega, \mathbb{R}^N))$. Throughout the paper we use as parabolic cylinders the one-sided parabolic cylinders of the form

$$Q_\varrho(z_o) := B_\varrho(x_o) \times (t_o - \varrho^2, t_o).$$

Here, $B_\varrho(x_o)$ denotes the open ball of radius $\varrho > 0$ with center $x_o \in \mathbb{R}^n$. Points in space-time $\mathbb{R}^{n+1}$ are denoted by $z = (x, t)$. The *parabolic distance* of two points $z_1 = (x_1, t_1), z_2 = (x_2, t_2) \in \mathbb{R}^{n+1}$ is given by

$$d_{\mathcal{P}}(z_1, z_2) := \max\left\{|x_1 - x_2|, \sqrt{|t_1 - t_2|}\right\}.$$

3.2. Preliminaries

In order to "reabsorb" certain terms, we will use the following iteration lemma, which is a standard tool and can be found, for instance, in [16].

Lemma 3.1. *Let $0 < \vartheta < 1$, $A, B \geqq 0$ and $\alpha > 0$. Then there exists a constant $c \equiv c(\alpha, \vartheta)$ such that there holds: For any $0 < r < \varrho$ and any non-negative, bounded function $\phi\colon [r, \varrho] \to [0, \infty)$ satisfying*

$$\phi(s) \leqq \vartheta\phi(t) + A(t - s)^{-\alpha} + B \qquad \textit{for all } 0 < r \leqq s < t \leqq \varrho,$$

we have

$$\phi(r) \leqq c\left[A(\varrho - r)^{-\alpha} + B\right].$$

The next Lemma can be retrieved from [9, Lemma 2.2].

Lemma 3.2. *Let $p > 1$ and $k \in \mathbb{N}$. Then there exists a constant $c \equiv c(p)$ such that for any $\mu \geqq 0$ and $A, B \in \mathbb{R}^k$ there holds*

$$\left(\mu^2 + |A|^2\right)^{\frac{p}{2}} \leqq c\left(\mu^2 + |B|^2\right)^{\frac{p}{2}} + c\left(\mu^2 + |A|^2 + |B|^2\right)^{\frac{p-2}{2}} |B - A|^2.$$

The following algebraic fact can be retrieved from [2] in the case $\sigma < 0$. The case $\sigma > 0$ can be obtained in a similar way.

Lemma 3.3. *Let $k \in \mathbb{N}$. For every $\sigma \in (-1/2, \infty)$ there exists a constant $c = c(\sigma) \geqq 1$ such that the following estimate holds true:*

$$\begin{aligned} c^{-1}\left(\mu^2 + |A|^2 + |B|^2\right)^{\sigma} &\leqq \int_0^1 \left(\mu^2 + |A + s(B - A)|^2\right)^{\sigma} \mathrm{d}s \\ &\leqq c\left(\mu^2 + |A|^2 + |B|^2\right)^{\sigma} \end{aligned}$$

for any $\mu \geqq 0$ and $A, B \in \mathbb{R}^k$, not both zero if $\mu = 0$ and $\sigma < 0$.

As a consequence of Lemma 3.3 one can show

Lemma 3.4. *Let $k \in \mathbb{N}$ and $p > 1$. Then there exists a constant $c \equiv c(p) \geqq 1$ such that for any $A, B \in \mathbb{R}^k$ there holds*

$$\left\langle |A|^{p-2}A - |B|^{p-2}B, A - B\right\rangle \geqq c^{-1}\left(|A|^2 + |B|^2\right)^{\frac{p-2}{2}} |A - B|^2.$$

4. Existence of Variational Solutions

The proof of the existence of variational solutions is divided into several steps. We start with a standard regularization procedure.

4.1. Regularization

For $\varepsilon \in (0, 1]$ we define the regularized integrand f_ε by

$$f_\varepsilon(\xi) := f(\xi) + \varepsilon|\xi|^q \qquad \text{for} \quad \xi \in \mathbb{R}^{Nn}.$$

From the properties (2.3) of the integrand f and Lemmas 3.3 and 3.4 we deduce the following growth and ellipticity properties of f_ε:

$$\begin{cases} \langle Df_\varepsilon(\xi) - Df_\varepsilon(\eta), \xi - \eta\rangle \geqq \frac{\varepsilon}{c(q)}\left(|\xi|^2 + |\eta|^2\right)^{\frac{q-2}{2}} |\xi - \eta|^2 \\ \qquad\qquad\qquad + \nu\left(\mu^2 + |\xi|^2 + |\eta|^2\right)^{\frac{p-2}{2}} |\xi - \eta|^2, \\ |Df_\varepsilon(\xi)| \leqq (L + c(q))\,(|\xi| + 1)^{q-1}, \end{cases} \tag{4.1}$$

whenever $\xi, \eta \in \mathbb{R}^{Nn}$. This ensures that Df_ε fulfills a standard q-growth and monotonicity condition and therefore allows us to construct weak energy solutions to the associated Cauchy–Dirichlet problem with a datum g as in (2.4). In the following, by

$$u_\varepsilon \in L^q\left(0, T; W^{1,q}(\Omega, \mathbb{R}^N)\right) \cap C^0\left([0, T]; L^2(\Omega, \mathbb{R}^N)\right)$$

we denote the unique solution to the Cauchy–Dirichlet problem

$$\begin{cases} \partial_t u_\varepsilon - \operatorname{div}\big(Df_\varepsilon(Du_\varepsilon)\big) = 0 & \text{in } \Omega_T\,, \\ u_\varepsilon = g & \text{on } \partial_{\mathcal{P}}\Omega_T\,. \end{cases} \tag{4.2}$$

The existence of such weak solutions u_ε can be deduced from the classical theory, see [20]. Note, also, that we have

$$\partial_t u_\varepsilon \in L^{q'}\big(0, T; W^{-1,q'}(\Omega, \mathbb{R}^N)\big). \tag{4.3}$$

4.2. Energy Bound

We test the weak formulation of the parabolic system $(4.2)_1$ with the testing-function $\varphi(x,t) = (u_\varepsilon - g)(x,t)\chi_\theta(t)$, where $\chi_\theta \in W^{1,\infty}(\mathbb{R})$ satisfies $\chi_\theta(t) = 1$ for $-\infty < t < \tau - \theta$ for some $\tau \in (0, T)$ and $\theta \in (0, \tau)$, $\chi_\theta(t) = 0$ for $t > \tau$ and $\chi_\theta(t) = \frac{1}{\theta}(\tau - t)$ for $\tau - \theta \leqq t \leqq \tau$. We note that the following computations are formal concerning the use of the time derivative $\partial_t u_\varepsilon$. However, they can be made rigorous by the use of a mollification procedure as, for instance, by Steklov averages with respect to time. With this choice of φ the weak form of $(4.2)_1$ yields, for almost every $\tau \in (0, T)$, in the limit $\theta \downarrow 0$ that

$$\begin{aligned} &\tfrac{1}{2}\int_\Omega |u_\varepsilon - g|^2(\cdot,\tau)\,\mathrm{d}x + \int_{\Omega_\tau} \big\langle Df_\varepsilon(Du_\varepsilon) - Df_\varepsilon(Dg), Du_\varepsilon - Dg\big\rangle\,\mathrm{d}z \\ &= -\int_{\Omega_\tau} \langle Df_\varepsilon(Dg), Du_\varepsilon - Dg\rangle\,\mathrm{d}z - \int_0^\tau \langle u_\varepsilon - g, g_t\rangle_{W^{1,p}(\Omega,\mathbb{R}^N)}\,\mathrm{d}t \\ &=: \mathrm{I}_1 - \mathrm{I}_2, \end{aligned} \tag{4.4}$$

where we have abbreviated $\Omega_\tau := \Omega \times (0, \tau)$. The first term I_1 is estimated with the bound $(4.1)_2$ and Young's inequality as follows:

$$\begin{aligned} |\,\mathrm{I}_1\,| &\leqq (L + c(q)) \int_{\Omega_\tau} (|Dg| + 1)^{q-1}|Du_\varepsilon - Dg|\,\mathrm{d}z \\ &\leqq \int_{\Omega_\tau} \Big[\tfrac{\delta}{2^{p-1}}|Du_\varepsilon - Dg|^p\,\mathrm{d}z + c\,(|Dg| + 1)^{p'(q-1)}\Big]\,\mathrm{d}z \\ &\leqq \delta\int_{\Omega_\tau} |Du_\varepsilon|^p\,\mathrm{d}z + c\int_{\Omega_{t_o}} \big(|Dg|^{p'(q-1)} + 1\big)\,\mathrm{d}z, \end{aligned}$$

for a constant $c \equiv c(p, q, 1/\delta)$. For the second term, I_2, a slice-wise application of Poincaré's inequality to $(u_\varepsilon - g)(\cdot, t)$ for almost every $t \in (0, \tau)$ and Young's inequality imply that

$$\begin{aligned} |\,\mathrm{I}_2\,| &\leqq \bigg(\int_{\Omega_\tau} |Du_\varepsilon - Dg|^p + |u_\varepsilon - g|^p\,\mathrm{d}z\bigg)^{\frac{1}{p}} \|g_t\|_{L^{p'}(0,\tau;W^{-1,p'}(\Omega,\mathbb{R}^N))} \\ &\leqq c\bigg(\int_{\Omega_\tau} |Du_\varepsilon - Dg|^p\,\mathrm{d}z\bigg)^{\frac{1}{p}} \|g_t\|_{L^{p'}(0,\tau;W^{-1,p'}(\Omega,\mathbb{R}^N))} \\ &\leqq \delta\int_{\Omega_\tau} \big(|Du_\varepsilon|^p + |Dg|^p\big)\,\mathrm{d}z + c\,\|g_t\|^{p'}_{L^{p'}(0,\tau;W^{-1,p'}(\Omega,\mathbb{R}^N))} \end{aligned}$$

holds true, where $c \equiv c(p, 1/\delta, \operatorname{diam}(\Omega))$. Finally, by $(4.1)_1$ and Lemma 3.2, we obtain for the second term on the left-hand side of (4.4) the following bound from below:

$$
\begin{aligned}
&\int_{\Omega_\tau} \big\langle Df_\varepsilon(Du_\varepsilon) - Df_\varepsilon(Dg), Du_\varepsilon - Dg\big\rangle \,\mathrm{d}z \\
&\quad \geqq \nu \int_{\Omega_\tau} \big(\mu^2 + |Du_\varepsilon|^2 + |Dg|^2\big)^{\frac{p-2}{2}} |Du_\varepsilon - Dg|^2 \,\mathrm{d}z \\
&\quad \geqq \tfrac{\nu}{c(p)} \int_{\Omega_\tau} |Du_\varepsilon|^p \,\mathrm{d}z - \nu \int_{\Omega_\tau} \big(\mu^2 + |Dg|\big)^{\frac{p}{2}} \,\mathrm{d}z.
\end{aligned}
$$

Inserting the preceding estimates into (4.4) and choosing δ small enough to reabsorb $\int_{\Omega_\tau} |Du_\varepsilon|^p \,\mathrm{d}z$ into the left-hand side, we find that

$$
\int_\Omega |u_\varepsilon(\cdot,\tau)|^2 \,\mathrm{d}x + \int_{\Omega_\tau} |Du_\varepsilon|^p \,\mathrm{d}z \leqq c(\nu, L, p, q, \operatorname{diam}(\Omega))\, M
$$

holds true for almost every $\tau \in (0, T)$. Here, we have abbreviated

$$
M := \int_{\Omega_T} \big(|Dg|^{p'(q-1)} + 1\big) \,\mathrm{d}z + \|g\|^2_{L^\infty(0,T;L^2(\Omega,\mathbb{R}^N))} + \|g_t\|^{p'}_{L^{p'}(0,T;W^{-1,p'}(\Omega,\mathbb{R}^N))}.
$$

Taking the supremum over $\tau \in (0, T)$ in the first term on the left-hand side and letting $\tau \uparrow T$ in the second one, we end up with the following *energy estimate*:

$$
\sup_{t\in(0,T)} \int_\Omega |u_\varepsilon(\cdot,t)|^2 \,\mathrm{d}x + \int_{\Omega_T} |Du_\varepsilon|^p \,\mathrm{d}z \leqq c(\nu, L, p, q, \operatorname{diam}(\Omega))\, M. \tag{4.5}
$$

Moreover, applying Poincaré's inequality slice-wise to $(u_\varepsilon - g)(\cdot, t)$ for almost every $t \in (0, T)$, we also obtain a bound for the L^p-norm of u_ε, that is, we have the inequality:

$$
\begin{aligned}
\int_{\Omega_T} |u_\varepsilon|^p \,\mathrm{d}z &\leqq 2^{p-1} \int_{\Omega_T} \big(|u_\varepsilon - g|^p + |g|^p\big) \,\mathrm{d}z \\
&\leqq c \int_{\Omega_T} \big(|Du_\varepsilon - Dg|^p + |g|^p\big) \,\mathrm{d}z \\
&\leqq c(\nu, L, p, q, \operatorname{diam}(\Omega)) \left[M + \int_{\Omega_T} |g|^p \,\mathrm{d}z \right].
\end{aligned} \tag{4.6}
$$

4.3. Weak Continuity in Time

We fix $0 < t_1 < t_2 < T$ and choose a testing function $\varphi \in C_0^\infty(\Omega \times (t_1, t_2))$ in the weak form of $(4.2)_1$. Subsequently, using the bound $(4.1)_2$, Hölder's inequality (note that $q < p + 1$) and the energy bound (4.5), we obtain

$$
\begin{aligned}
&\left|\int_{\Omega\times(t_1,t_2)} u_\varepsilon \partial_t\varphi \,\mathrm{d}z\right| \\
&= \left|\int_{\Omega\times(t_1,t_2)} \langle Df_\varepsilon(Du_\varepsilon), D\varphi\rangle \,\mathrm{d}z\right| \\
&\leqq (L + c(q)) \int_{\Omega\times(t_1,t_2)} (1 + |Du_\varepsilon|)^{q-1}|D\varphi| \,\mathrm{d}z \\
&\leqq (L + c(q)) \,|\operatorname{spt}\varphi|^{\frac{p+1-q}{p}} \left(\int_{\Omega_T} (1 + |Du_\varepsilon|)^p \,\mathrm{d}z\right)^{\frac{q-1}{p}} \|D\varphi\|_{L^\infty(\Omega\times(t_1,t_2))} \\
&\leqq c \,|\operatorname{spt}\varphi|^{\frac{p+1-q}{p}} \|D\varphi\|_{L^\infty(\Omega\times(t_1,t_2))}, \qquad (4.7)
\end{aligned}
$$

where $c = c(n, \nu, L, p, q, \operatorname{diam}(\Omega), M)$. Now, for $t_1 < s_1 < s_2 < t_2$ and $\delta > 0$ small enough we define

$$
\chi_\delta(t) := \begin{cases} 0, & \text{for } t_1 \leqq t \leqq s_1 - \delta, \\ \frac{1}{\delta}(t - s_1 + \delta), & \text{for} s_1 - \delta \leqq t \leqq s_1, \\ 1, & \text{for } s_1 \leqq t \leqq s_2, \\ -\frac{1}{\delta}(t - s_2 - \delta), & \text{for } s_2 \leqq t \leqq s_2 + \delta, \\ 0, & \text{for } s_2 + \delta \leqq t \leqq t_2. \end{cases}
$$

We choose in (4.7) a testing function of the form $\varphi(x,t) := \chi_\delta(t)\psi(x)$ with $\psi \in C_0^\infty(\Omega, \mathbb{R}^N)$. Note that this choice is possible by an approximation argument. We find

$$
\begin{aligned}
&\left|\int_\Omega \frac{1}{\delta}\left(\int_{s_1-\delta}^{s_1} u_\varepsilon(x,t)\,\mathrm{d}t - \int_{s_2}^{s_2+\delta} u_\varepsilon(x,t)\,\mathrm{d}t\right)\psi(x)\,\mathrm{d}x\right| \\
&\quad \leqq c\,(s_2 - s_1 + 2\delta)^{\frac{p+1-q}{p}} \|D\psi\|_{L^\infty(\Omega)}.
\end{aligned}
$$

In the preceding inequality we pass to the limit $\delta \downarrow 0$ and obtain for almost every $t_1 < s_1 < s_2 < t_2$ that

$$
\left|\int_\Omega \big(u_\varepsilon(x,s_1) - u_\varepsilon(x,s_2)\big)\psi(x)\,\mathrm{d}x\right| \leqq c\,(s_2 - s_1)^{\frac{p+1-q}{p}} \|D\psi\|_{L^\infty(\Omega)}
$$

holds true, whenever $\psi \in C_0^\infty(\Omega, \mathbb{R}^N)$. Now, with $\ell \in \mathbb{N}$ such that $\ell > \frac{n+2}{2}$ the Sobolev inequality yields

$$
\|D\psi\|_{L^\infty(\Omega)} \leqq c(n, \ell, \Omega)\,\|\psi\|_{W^{\ell,2}(\Omega)},
$$

and hence

$$
\left|\int_\Omega \big(u_\varepsilon(x,s_1) - u_\varepsilon(x,s_2)\big)\psi(x)\,\mathrm{d}x\right| \leqq c\,(s_2 - s_1)^{\frac{p+1-q}{p}} \|\psi\|_{W^{\ell,2}(\Omega)},
$$

whenever $\psi \in C_0^\infty(\Omega, \mathbb{R}^N)$. The constant c depends on $n, \nu, L, p, q, \ell, \Omega$ and M. By the density of $C_0^\infty(\Omega, \mathbb{R}^N)$ in $W_0^{\ell,2}(\Omega, \mathbb{R}^N)$ the last inequality also continues to hold for any $\psi \in W_0^{\ell,2}(\Omega, \mathbb{R}^N)$. Taking the supremum over all $\psi \in W_0^{\ell,2}(\Omega, \mathbb{R}^N)$ satisfying $\|\psi\|_{W^{\ell,2}(\Omega)} \leqq 1$ we deduce that

$$\|u_\varepsilon(\cdot, s_1) - u_\varepsilon(\cdot, s_2)\|_{W^{-\ell,2}(\Omega)} \leqq c\,|s_1 - s_2|^{\frac{p+1-q}{p}} \tag{4.8}$$

holds true for almost every $s_1, s_2 \in (t_1, t_2)$. This is the desired weak continuity property with respect to time for u_ε; to be more precise, the mapping $t \mapsto u_\varepsilon(\cdot, t) \in W^{-\ell,2}(\Omega, \mathbb{R}^N)$ is Hölder continuous with Hölder exponent $\frac{p+1-q}{p} \in (0, \frac{1}{p})$.

4.4. The Variational Formulation

We fix $\tau \in (0, T]$. Since u_ε satisfies (4.3), the parabolic system $(4.2)_1$ can be rewritten in the form

$$\int_0^\tau \langle \partial_t u_\varepsilon, \varphi \rangle_{W^{1,q}(\Omega,\mathbb{R}^N)}\,\mathrm{d}t + \int_{\Omega_\tau} \langle Df_\varepsilon(Du_\varepsilon), D\varphi\rangle\,\mathrm{d}z = 0$$

for any testing function $\varphi \in L^q(0, \tau; W_0^{1,q}(\Omega, \mathbb{R}^N))$. The convexity of f implies that $\langle Df_\varepsilon(Du_\varepsilon), D\varphi\rangle \leqq f_\varepsilon(D\varphi + Du_\varepsilon) - f_\varepsilon(Du_\varepsilon)$ holds true, and therefore the preceding identity yields that

$$\int_0^\tau \langle \partial_t u_\varepsilon, \varphi \rangle_{W^{1,q}(\Omega,\mathbb{R}^N)}\,\mathrm{d}t + \int_{\Omega_\tau} f_\varepsilon(Du_\varepsilon + D\varphi) - f_\varepsilon(Du_\varepsilon)\,\mathrm{d}z \geqq 0,$$

whenever $\varphi \in L^q(0, \tau; W_0^{1,q}(\Omega, \mathbb{R}^N))$. We now perform the substitution $v = u_\varepsilon + \varphi$. Of course $v \in u_\varepsilon + L^q(0, \tau; W_0^{1,q}(\Omega, \mathbb{R}^N)) \subset L^q(0, \tau; W^{1,q}(\Omega, \mathbb{R}^N))$ and $v = u_\varepsilon = g$ on $\partial\Omega \times (0, \tau)$. Moreover, by (4.3) we have $\partial_t v \in L^{q'}(0, \tau; W^{-1,q'}(\Omega, \mathbb{R}^N))$, provided we also assume that $\partial_t \varphi \in L^{q'}(0, \tau; W^{-1,q'}(\Omega, \mathbb{R}^N))$. In terms of the comparison map v the last inequality can therefore be rewritten in the form

$$\begin{aligned}\int_0^\tau \langle \partial_t v, v - u_\varepsilon \rangle_{W^{1,q}(\Omega,\mathbb{R}^N)}\,\mathrm{d}t &+ \int_{\Omega_\tau} f_\varepsilon(Dv) - f_\varepsilon(Du_\varepsilon)\,\mathrm{d}z \\ &\geqq \int_0^\tau \langle \partial_t (v - u_\varepsilon), v - u_\varepsilon \rangle_{W^{1,q}(\Omega,\mathbb{R}^N)}\,\mathrm{d}t,\end{aligned}$$

where v is of the form $u_\varepsilon + \varphi$. But this means that the preceding inequality holds true for any $v \in L^q(0, \tau; W_g^{1,q}(\Omega, \mathbb{R}^N))$ with $\partial_t v \in L^{q'}(0, \tau; W^{-1,q'}(\Omega, \mathbb{R}^N))$. As a consequence of the Aubin-Lions embedding—see [31, Chapter III, Proposition 1.2]—we have that $v \in C^0([0, \tau], L^2(\Omega, \mathbb{R}^N))$. Therefore, the term on the right-hand side of the last inequality can be simplified to

$$\begin{aligned}&\int_0^\tau \langle \partial_t (v - u_\varepsilon), v - u_\varepsilon \rangle_{W^{1,q}(\Omega,\mathbb{R}^N)}\,\mathrm{d}t \\ &\quad = \tfrac{1}{2}\|(v - u_\varepsilon)(\cdot, \tau)\|^2_{L^2(\Omega)} - \tfrac{1}{2}\|(v - g)(\cdot, 0)\|^2_{L^2(\Omega)}.\end{aligned}$$

Furthermore, if the comparison function v satisfies $\partial_t v \in L^{p'}(0, \tau; W^{-1,p'}(\Omega, \mathbb{R}^N)) \subset L^{q'}(0, \tau; W^{-1,q'}(\Omega, \mathbb{R}^N))$, the first integral of the left-hand side can be written in the following form:

$$\int_0^\tau \langle \partial_t v, v - u_\varepsilon \rangle_{W^{1,q}(\Omega,\mathbb{R}^N)} \, dt = \int_0^\tau \langle \partial_t v, v - u_\varepsilon \rangle_{W^{1,p}(\Omega,\mathbb{R}^N)} \, dt.$$

Therefore, we obtain that

$$\begin{aligned} \int_0^\tau \langle \partial_t v, v - u_\varepsilon \rangle_{W^{1,p}(\Omega,\mathbb{R}^N)} \, dt + \int_{\Omega_\tau} f_\varepsilon(Dv) - f_\varepsilon(Du_\varepsilon) \, dz \\ \geqq \tfrac{1}{2} \|(v - u_\varepsilon)(\cdot, \tau)\|^2_{L^2(\Omega)} - \tfrac{1}{2} \|(v - g)(\cdot, 0)\|^2_{L^2(\Omega)} \end{aligned} \tag{4.9}$$

holds true for any comparison function $v \in L^q(0, \tau; W^{1,q}_g(\Omega, \mathbb{R}^N))$ whose time derivative satisfies $\partial_t v \in L^{p'}(0, \tau; W^{-1,p'}(\Omega, \mathbb{R}^N))$. Since the left-hand side of (4.9) is always infinite for any $v \in L^p(0, \tau; W^{1,p}_g(\Omega, \mathbb{R}^N)) \setminus L^q(0, \tau; W^{1,q}_g(\Omega, \mathbb{R}^N))$ (note that by definition of f_ε we have $f_\varepsilon(Dv) \geqq \varepsilon |Dv|^q$ and therefore $\int_{\Omega_\tau} f_\varepsilon(Dv)\, dx = \infty$ for such v) while the right-hand side is finite, the inequality (4.9) trivially holds true for $v \in L^p(0, \tau; W^{1,p}_g(\Omega, \mathbb{R}^N)) \setminus L^q(0, \tau; W^{1,q}_g(\Omega, \mathbb{R}^N))$ with $\partial_t v \in L^{p'}(0, \tau; W^{-1,p'}(\Omega, \mathbb{R}^N))$. This proves that u_ε is a variational solution in the sense of Definition 2.2. Note that, at this stage, the mappings u_ε belong to $C^0([0, T]; L^2(\Omega, \mathbb{R}^N))$. This property will, however, be lost in the limit $\varepsilon \downarrow 0$.

4.5. Passage to the Limit $\varepsilon \downarrow 0$

Here, we shall pass to the limit $\varepsilon \downarrow 0$ in the variational inequality (4.9). From (4.5) and (4.6) we know that u_ε is uniformly bounded (with respect to ε) in $L^p(0, T; W^{1,p}(\Omega, \mathbb{R}^N))$. Therefore, there exists a (not re-labelled) subsequence and a function $u \in L^p(0, T; W^{1,p}(\Omega, \mathbb{R}^N))$ such that

$$u_\varepsilon \rightharpoonup u \quad \text{weakly in } L^p\big(0, T; W^{1,p}(\Omega, \mathbb{R}^N)\big).$$

In order to conclude the weak continuity in time of the limit map u we use a compactness argument from [17], see Theorem A.2 below. From (4.5) we infer that u_ε is uniformly bounded (with respect to ε) in $L^\infty(0, T; L^2(\Omega, \mathbb{R}^N))$, while (4.8) yields that u_ε are equicontinuous as functions in $C^0([0, T]; W^{-\ell,2}(\Omega, \mathbb{R}^N))$ (more precisely, we can choose as modulus the function $\omega(t) := c\, t^{\frac{p+1-q}{p}}$). Hence, the hypotheses of Theorem A.2 are fulfilled and therefore, passing again to a (non-relabeled) subsequence, we see that the limit function u is contained in $C_w([0, T]; L^2(\Omega, \mathbb{R}^N))$ and that

$$u_\varepsilon(\cdot, t) \rightharpoonup u(\cdot, t) \quad \text{weakly in } L^2(\Omega, \mathbb{R}^N) \quad \text{for any } t \in [0, T]$$

holds true as $\varepsilon \downarrow 0$. Note that $u_\varepsilon(\cdot,0) = g(\cdot,0)$ for any $\varepsilon \in (0,1]$ implies that $u(\cdot,0) = g(\cdot,0)$. By lower semicontinuity of $w \mapsto \int_{\Omega_\tau} f(Dw)\,\mathrm{d}z$ with respect to weak convergence in $L^p(0,\tau;\, W^{1,p}(\Omega,\,\mathbb{R}^N))$ (respectively $w \mapsto \frac{1}{2}\|w-u(\cdot,\tau)\|^2_{L^2(\Omega)}$ with respect to weak convergence in $L^2(\Omega,\mathbb{R}^N)$) and the continuity of the pairing $w \mapsto \langle\,\cdot\,,\, w\rangle_{W^{1,p}(\Omega,\mathbb{R}^N)}$ with respect to weak convergence in $W^{1,p}(\Omega,\mathbb{R}^N)$, we can pass to the limit $\varepsilon \downarrow 0$ in (4.9). This leads to

$$\begin{aligned}\int_0^\tau \langle \partial_t v,\, v-u\rangle_{W^{1,p}(\Omega,\mathbb{R}^N)}\,\mathrm{d}t + \int_{\Omega_\tau} f(Dv) - f(Du)\,\mathrm{d}z \\ \geqq \tfrac{1}{2}\|(v-u)(\cdot,\tau)\|^2_{L^2(\Omega)} - \tfrac{1}{2}\|(v-g)(\cdot,0)\|^2_{L^2(\Omega)}\end{aligned}$$

for any $\tau \in (0,T]$. But this means that u is the variational solution we are looking for.

5. Uniqueness of Variational Solutions

In this chapter we shall prove the uniqueness result, Theorem 2.4. For this purpose we need some prerequisites which we will present in the following subsection and in Appendix B.

5.1. Mollification in Time

Due to their lack of regularity with respect to time, the variational solutions are not admissible as comparison functions in (2.5), since the term involving the time derivative $\partial_t v$ would not be well defined. To overcome this difficulty we shall use a certain mollification in time which has been useful in several other circumstances, as, for example, for the treatment of evolutionary obstacle problems; see [6,20,28]. The precise construction of the regularization is as follows: For $v \in L^1(\Omega_T,\mathbb{R}^N),\ v_o \in L^1(\Omega,\mathbb{R}^N)$ and $h \in (0,T]$, we define

$$[v]_h(\cdot,t) := \mathrm{e}^{-\frac{t}{h}} v_o + \frac{1}{h}\int_0^t \mathrm{e}^{\frac{s-t}{h}} v(\cdot,s)\,\mathrm{d}s, \tag{5.1}$$

for $t \in [0,T]$. One of the basic features of this mollification is that $[v]_h$ (formally) solves the differential equation

$$\partial_t [v]_h = -\tfrac{1}{h}\big([v]_h - v\big),$$

with initial condition $[v]_h(\cdot,0) = v_o$. This will allow us to pass to the limit in certain approximations since, in these approximations, the quantity $\partial_t[v]_h([v]_h - v)$ naturally arises when comparing parabolic variational inequalities. This comes from the the fact that we can only add, but not subtract, two variational inequalities. The nice feature of the regularization is that now the quantity $\partial_t[v]_h([v]_h - v)$ is non-positive. This will be crucial in the proof of the uniqueness result. The basic properties of the mollification $[\cdot]_h$ are provided in Appendix B.

5.2. Proof of the Uniqueness Result

Now, with the properties of the mollification in time at hand, we come to the proof of the uniqueness result from Theorem 2.4.

Proof of Theorem 2.4. Suppose that u_1 and u_2 are two different variational solutions to (2.1). Adding the variational inequalities (2.5) for u_1 and u_2 and taking into account the fact that $\|(v-u_i)(\cdot,\ T)\|^2_{L^2(\Omega)} \geqq 0$ for $i=1,2$ yields, for any $v \in L^p(0,T;\ W_g^{1,p}(\Omega,\ \mathbb{R}^N))$ with $v_t \in L^{p'}(0,T;\ W^{-1,p'}(\Omega,\ \mathbb{R}^N))$ (note that this implies $v \in C^0([0,T];\ L^2(\Omega,\mathbb{R}^N))$), that

$$\int_{\Omega_T} [f(Du_1)+f(Du_2)]\,\mathrm{d}z \leqq 2\int_{\Omega_T} f(Dv)\,\mathrm{d}z + 2\int_0^T \langle v_t, v-w\rangle_{W^{1,p}(\Omega,\ \mathbb{R}^N)}\,\mathrm{d}t + \|(v-g)(\cdot,\ 0)\|^2_{L^2(\Omega)}.$$

Here we have abbreviated $w := \frac{u_1+u_2}{2}$. At this point we would like to choose the comparison map $v = w$ in the previous inequality. However, this is not allowed, since we do not know that w_t belongs to $L^{p'}(0,T;\ W^{-1,p'}(\Omega,\mathbb{R}^N))$. For this reason we introduce, with the mollification $[\cdot]_h$ from (5.1), the time-regularized functions

$$v_h := [w-g]_h + g, \quad \text{for } h \in (0,T].$$

Since $(w-g)(\cdot,\ 0) = 0$, we choose $v_o \equiv (w-g)(\cdot,\ 0) = 0$ in the definition. This makes sense, since the mollification should admit the same initial conditions as the function $u-g$ itself. By Lemma B.2 we have $v_h \in L^p(0,T;\ W_g^{1,p}(\Omega,\ \mathbb{R}^N)) \cap C^0([0,T];\ L^2(\Omega,\ \mathbb{R}^N))$ and $v_h(\cdot,0) = g(\cdot,\ 0)$. Moreover, from Lemma B.3 we know that $\partial_t v_h \in L^{p'}(0,T;\ W^{-1,p'}(\Omega,\ \mathbb{R}^N))$. Therefore, we are allowed to choose $v = v_h$ as the comparison function in the last inequality. This choice leads to

$$\int_{\Omega_T} [f(Du_1)+f(Du_2)]\,\mathrm{d}z \leqq 2\int_{\Omega_T} f(Dv_h)\,\mathrm{d}z + 2\int_0^T \langle \partial_t v_h, v_h-w\rangle_{W^{1,p}(\Omega,\mathbb{R}^N)}\,\mathrm{d}t =: 2(\mathrm{I}_h+\mathrm{II}_h), \quad (5.2)$$

with the obvious meaning of I_h and II_h. In order to treat I_h we first rewrite the integrand in the form

$$f(Dv_h) = f\big(D[w]_h + D(g-[g]_h)\big) =: f\big(D[w]_h + A_h\big),$$

with the obvious abbreviation $A_h := D(g-[g]_h)$. Here, we recall that $w(\cdot,\ 0) = g(\cdot,\ 0)$, and therefore $[w-g]_h = [w]_h - [g]_h$ with $[w]_h$ and $[g]_h$ defined according to (5.1) with $v_o = g(\cdot,\ 0)$. Next, we apply the mean value theorem to infer, for a function $\mu(x,t) \in [0,1]$, that

$$f(Dv_h) = f\big(D[w]_h\big) + \big\langle Df\big(D[w]_h + \mu A_h\big), A_h\big\rangle.$$

Since $q < p+1$, the second term on the right-hand side can be estimated with the growth condition for Df from (2.3) and an application of Hölder's inequality as follows:

$$\begin{aligned}
&\left|\int_{\Omega_T}\big\langle Df\big(D[w]_h+\mu A_h\big), A_h\big\rangle\,\mathrm{d}z\right| \\
&\leqq L\int_{\Omega_T}|A_h|\big(1+|D[w]_h+\mu A_h|\big)^{q-1}\,\mathrm{d}z \\
&\leqq c\bigg[\int_{\Omega_T}|A_h|^{\frac{p}{p+1-q}}\,\mathrm{d}z\bigg]^{\frac{p+1-q}{p}}\bigg[\int_{\Omega_T}1+|D[w]_h|^p+|A_h|^p\,\mathrm{d}z\bigg]^{\frac{q-1}{p}},
\end{aligned}$$

for a constant $c=c(p,q,L)$. Now we use the hypotheses on the datum g to control the integrals involving A_h. The first integral converges to 0 as $h\downarrow 0$ by Lemma B.2 (ii), since $g\in L^{\frac{p}{p+1-q}}(0,T;W^{1,\frac{p}{p+1-q}}(\Omega,\mathbb{R}^N))$. More precisely, we have

$$\bigg[\int_{\Omega_T}|A_h|^{\frac{p}{p+1-q}}\,\mathrm{d}z\bigg]^{\frac{p+1-q}{p}}=\bigg[\int_{\Omega_T}|D(g-[g]_h)|^{\frac{p}{p+1-q}}\,\mathrm{d}z\bigg]^{\frac{p+1-q}{p}}\to 0$$

as $h\downarrow 0$. The second integral can be bounded uniformly with respect to h, again, by an application of Lemma B.2 (i), since $w,g\in L^p(0,T;W^{1,p}(\Omega,\mathbb{R}^N))$. Here we have the estimate

$$\begin{aligned}
&\bigg[\int_{\Omega_T}|D[w]_h|^p+|A_h|^p\,\mathrm{d}z\bigg]^{\frac{1}{p}} \\
&\leqq c\bigg[\int_{\Omega_T}|Dg|^p+|Du_1|^p+|Du_2|^p\,\mathrm{d}z+h\int_{\Omega}|Dg(\cdot,0)|^p\,\mathrm{d}x\bigg]^{\frac{1}{p}},
\end{aligned}$$

where the constant c depends only p. Joining the last two inequalities therefore implies

$$\lim_{h\downarrow 0}\left|\int_{\Omega_T}\big\langle Df\big(D[w]_h+\mu A_h\big), A_h\big\rangle\,\mathrm{d}z\right|=0.$$

It remains to treat the integral involving $f(D[w]_h)$. For this we observe that

$$\frac{1}{h(1-\mathrm{e}^{-\frac{t}{h}})}\int_0^t\mathrm{e}^{\frac{s-t}{h}}\,\mathrm{d}s=1.$$

This allows us to interpret the mollification $[w]_h$—modulo a multiplicative factor—as a mean with respect to the measure $\mathrm{e}^{\frac{s-t}{h}}\,\mathrm{d}s$. Therefore, we rewrite $f(D[w]_h)$ according to this interpretation and afterwards use the convexity of f and Jensen's inequality. This procedure yields a pointwise bound of the term in question as follows:

$$\begin{aligned}
f\big(D[w]_h(\cdot,t)\big)&=f\bigg(\mathrm{e}^{-\frac{t}{h}}Dg(\cdot,0)+\frac{1-\mathrm{e}^{-\frac{t}{h}}}{h(1-\mathrm{e}^{-\frac{t}{h}})}\int_0^t Dw(\cdot,t)\mathrm{e}^{\frac{s-t}{h}}\,\mathrm{d}s\bigg) \\
&\leqq \mathrm{e}^{-\frac{t}{h}}f(Dg(\cdot,0))+\big(1-\mathrm{e}^{-\frac{t}{h}}\big)f\bigg(\frac{1}{h(1-\mathrm{e}^{-\frac{t}{h}})}\int_0^t Dw(\cdot,t)\mathrm{e}^{\frac{s-t}{h}}\,\mathrm{d}s\bigg) \\
&\leqq \mathrm{e}^{-\frac{t}{h}}f(Dg(\cdot,0))+\frac{1-\mathrm{e}^{-\frac{t}{h}}}{h(1-\mathrm{e}^{-\frac{t}{h}})}\int_0^t f\big(Dw(\cdot,t)\big)\mathrm{e}^{\frac{s-t}{h}}\,\mathrm{d}s \\
&=\big[f(Dw)\big]_h(\cdot,t),
\end{aligned}$$

where $[f(Dw)]_h$ is defined according to (5.1) with $v_o = f(Dg(\cdot, 0))$. Since $f(Dw) \in L^1(0, T; L^1(\Omega)) = L^1(\Omega_T)$ and $f(Dg(\cdot, 0)) \in L^1(\Omega)$, we have by Lemma B.2 (i) the uniform bound

$$\begin{aligned}
&\big\| \big[f(Dw)\big]_h \big\|_{L^1(\Omega_T)} \\
&\quad \leqq \|f(Dw)\|_{L^1(\Omega_T)} + h\|f(Dg(\cdot, 0))\|_{L^1(\Omega)} \\
&\quad \leqq \tfrac{1}{2}\|f(Du_1)\|_{L^1(\Omega_T)} + \tfrac{1}{2}\|f(Du_2)\|_{L^1(\Omega_T)} + h\|f(Dg(\cdot, 0))\|_{L^1(\Omega)} < \infty,
\end{aligned}$$

where we used for the last inequality the fact that u_1 and u_2 are variational solutions, which implies, in particular, that $\|f(Du_i)\|_{L^1(\Omega_T)} < \infty$ for $i = 1, 2$. Since the second term, that is, the term $h\|f(Dg(\cdot, 0))\|_{L^1(\Omega)}$, vanishes in the limit $h \downarrow 0$, a variant of the dominated convergence theorem implies that

$$\lim_{h\downarrow 0} \int_{\Omega_T} f\big([Dw]_h\big)\,\mathrm{d}z = \int_{\Omega_T} f(Dw)\,\mathrm{d}z = \int_{\Omega_T} f\big(\tfrac{1}{2}Du_1 + \tfrac{1}{2}Du_2\big)\,\mathrm{d}z \tag{5.3}$$

holds true. This finishes the treatment of the first term appearing on the right-hand side of (5.2); it remains to consider the second term II_h. We rewrite II_h as follows:

$$\mathrm{II}_h = \mathrm{II}_h^{(1)} + \mathrm{II}_h^{(2)} + \mathrm{II}_h^{(3)},$$

where we have abbreviated

$$\begin{aligned}
\mathrm{II}_h^{(1)} &:= \int_{\Omega_T} \partial_t[w]_h([w]_h - w)\,\mathrm{d}z, \\
\mathrm{II}_h^{(2)} &:= \int_{\Omega_T} \partial_t[w]_h \cdot (g - [g]_h)\,\mathrm{d}z, \\
\mathrm{II}_h^{(3)} &:= \int_0^T \langle \partial_t g - \partial_t[g]_h, [w - g]_h - (w - g)\rangle_{W^{1,p}(\Omega)}\,\mathrm{d}t.
\end{aligned}$$

Due to Lemma B.2 (iv) we know that the first term is non-negative, since

$$\mathrm{II}_h^{(1)} = -\tfrac{1}{h}\int_{\Omega_T} \big|[w]_h - w\big|^2\,\mathrm{d}z \leqq 0.$$

Using the fact that $[w]_h \to w$ in $L^p(0, T; W^{1,p}(\Omega, \mathbb{R}^N))$, and the uniform bound for the norm of $\partial_t[g]_h$ in $L^{p'}(0, T; W^{-1,p'}(\Omega, \mathbb{R}^N))$ by $\|\partial_t g\|_{L^{p'}(0,T;W^{-1,p'}(\Omega,\mathbb{R}^N))}$ (see Lemma B.3), we obtain in the limit $h \downarrow 0$ that there holds

$$\begin{aligned}
\mathrm{II}_h^{(2)} &= -\frac{1}{h}\int_{\Omega_T} ([w]_h - w)\cdot(g - [g]_h)\,\mathrm{d}z \\
&= -\int_{\Omega_T} ([w]_h - w)\cdot \partial_t[g]_h\,\mathrm{d}z \\
&= -\int_0^T \langle \partial_t[g]_h, [w]_h - w\rangle_{W^{1,p}(\Omega)}\,\mathrm{d}t \to 0.
\end{aligned}$$

The same reasoning can be used to treat the term $\mathrm{II}_h^{(3)}$, since $[w-g]_h \to w-g$ strongly in $L^p(0,T;W^{1,p}(\Omega,\mathbb{R}^N))$. Therefore, we have

$$\lim_{h\downarrow 0} \mathrm{II}_h^{(3)} = 0.$$

Altogether we have established that

$$\limsup_{h\downarrow 0} \mathrm{II}_h \leqq 0. \tag{5.4}$$

With (5.3) and (5.4) at hand we can pass in (5.2) to the limit $h \downarrow 0$ to obtain that

$$\begin{aligned}\int_{\Omega_T} [f(Du_1)+f(Du_2)]\,\mathrm{d}z &\leqq 2\int_{\Omega_T} f\big(\tfrac12 Du_1 + \tfrac12 Du_2\big)\,\mathrm{d}z \\ &< \int_{\Omega_T} [f(Du_1)+f(Du_2)]\,\mathrm{d}z.\end{aligned}$$

In the last step we used the strict convexity of f and the assumption that $u_1 \neq u_2$. Thus we arrived with the preceding inequality at the desired contradiction. This proves the uniqueness of variational solutions. □

6. A Local L^q-Estimate for the Spatial Gradient

In this Section we prove quantitative, local interior L^q-estimates for the spatial gradient Du of weak solutions in terms of their local L^p-norm. These estimates shall be proved as a priori estimates, in the sense that we initially assume $Du \in L^q_{\mathrm{loc}}(\Omega_T,\mathbb{R}^{Nn})$. Therefore, they are not directly applicable to variational solutions. Later on, these a priori estimates will be applied in an approximation scheme, that is, an approximating sequence of solutions to regularized problems for which the higher integrability assumption is known to hold true. Before we start with the proof of the higher integrability estimate, we shall provide the necessary facts about parabolic function spaces.

6.1. Parabolic Function Spaces

The first Lemma is a parabolic version of Sobolev's inequality, which follows from Gagliardo–Nirenber's inequality. The statement can be found, for instance, in [12, Chapter I, Proposition 3.1].

Lemma 6.1. *Let* $\sigma \geqq 1$, $Q_\varrho(z_o) \equiv B_\varrho(x_o)\times(t_o-\varrho^2,t_o) \subset \mathbb{R}^{n+1}$ *and*

$$u \in L^\sigma\big(t_o-\varrho^2,t_o;W^{1,\sigma}(B_\varrho(x_o))\big) \cap C^0\big(t_o-\varrho^2 t_o;L^2(B_\varrho(x_o))\big).$$

Then, $u \in L^{\frac{\sigma(n+2)}{n}}(Q_\varrho(z_o))$, *and there exists a constant* $c=c(n,\sigma)$ *such that*

$$\begin{aligned}&-\!\!\!\!\!\!\int_{Q_\varrho(z_o)} \Big|\frac{u}{\varrho}\Big|^{\frac{\sigma(n+2)}{n}}\,\mathrm{d}z \\ &\quad\leqq c -\!\!\!\!\!\!\int_{Q_\varrho(z_o)} \Big(|Du|^\sigma + \Big|\frac{u}{\varrho}\Big|^\sigma\Big)\,\mathrm{d}z \Big(\sup_{t\in(t_o-\varrho^2,t_o)} -\!\!\!\!\!\!\int_{B_\varrho(x_o)} |u(\cdot,t)|^2\,\mathrm{d}x\Big)^{\frac{\sigma}{n}}.\end{aligned}$$

The following results are concerned with elliptic and parabolic fractional Sobolev spaces. Since embeddings of such spaces provide certain higher integrability properties, they will play a crucial role in the proof of the a priori estimate. We first recall their definitions. We say that $f \in W^{k,p}(\Omega, \mathbb{R}^k)$ with $1 \leqq p < \infty$, $k \in \mathbb{N}_0$ belongs to the fractional Sobolev space $W^{k+\alpha, p}(\Omega, \mathbb{R}^k)$, with $\alpha \in (0, 1)$, if the Gagliardo semi-norm $[D^\beta f]_{\alpha,p;\Omega}$ of any weak partial derivative $D^\beta f$ of order $|\beta| = k$ is finite; here, the Gagliardo semi-norm is defined by

$$[D^\beta f]^p_{\alpha, p;\Omega} := \int_\Omega \int_\Omega \frac{|D^\beta f(x) - D^\beta f(y)|^p}{|x-y|^{n+\alpha p}} \,\mathrm{d}x \,\mathrm{d}y$$

for any multiindex $\beta \in \mathbb{N}_0^n$ with $|\beta| = k$. Endowing $W^{k+\alpha, p}(\Omega, \mathbb{R}^k)$ with the norm

$$\|f\|_{W^{k+\alpha, p}(\Omega)} := \|f\|_{W^{k,p}(\Omega)} + \sum_{|\beta|=k} [D^\beta f]_{\alpha,p;\Omega},$$

$W^{k+\alpha, p}(\Omega, \mathbb{R}^k)$ becomes a Banach space. Later on, we will need a Gagliardo–Nirenberg inequality for fractional Sobolev spaces as stated in Lemma 6.4. This inequality will be a consequence of an interpolation inequality from [7, Corollary 3.2] and a fractional Sobolev inequality, which can be found, for instance, in [13, Theorem 6.5].

Lemma 6.2. (Fractional interpolation) *Assume that* $0 \leqq \alpha_1 < \alpha_2 < \infty$, $1 < p_1, p_2 < \infty$ *and* $\theta \in (0, 1)$. *Let* α *and* p *denote the convex combinations of* α_1, α_2 *(respectively* $\frac{1}{p_1}, \frac{1}{p_2}$*), that is,*

$$\alpha = \theta\alpha_1 + (1-\theta)\alpha_2, \quad \frac{1}{p} = \frac{\theta}{p_1} + \frac{1-\theta}{p_2}. \tag{6.1}$$

Finally, suppose that $f \in W^{\alpha_1,p_1}(\mathbb{R}^n) \cap W^{\alpha_2,p_2}(\mathbb{R}^n)$. *Then,* $f \in W^{\alpha,p}(\mathbb{R}^n)$*, and there exists a constant* $c = c(\alpha_1, \alpha_2, p_1, p_2, \theta)$ *such that*

$$\|f\|_{W^{\alpha,p}(\mathbb{R}^n)} \leqq c\, \|f\|^\theta_{W^{\alpha_1,p_1}(\mathbb{R}^n)} \|f\|^{1-\theta}_{W^{\alpha_2,p_2}(\mathbb{R}^n)}.$$

Lemma 6.3. (Fractional Sobolev embedding) *Let* $\alpha \in (0, 1)$ *and* $p \in [1, \infty)$ *such that* $\alpha p < n$. *Then, for any* $f \in W^{\alpha,p}(\mathbb{R}^n)$ *we have* $f \in L^{\frac{np}{n-\alpha p}}(\mathbb{R}^n)$*, and there exists a constant* $c = c(n, s, \alpha)$ *such that*

$$\|f\|_{L^{\frac{np}{n-\alpha p}}(\mathbb{R}^n)} \leqq c\, [f]_{\alpha,p;\mathbb{R}^n}.$$

With these versions of interpolation and embedding in fractional Sobolev spaces we are able to establish the *fractional Gagliardo–Nirenberg inequality*, which is suited for our purposes. We note that for Besov-spaces a similar inequality has been proved in [21, Corollary 2]. For the sake of completeness we provide a proof in the setting of fractional Sobolev-spaces.

Lemma 6.4. (Fractional Gagliardo–Nirenberg inequality) *Let $B_\varrho(x_o) \subset \mathbb{R}^n$, with $\varrho \leqq 1$ and $\lambda, \mu, \theta \in (0,1)$, $1 < p, r < s < \infty$, such that*

$$-\frac{n}{s} \leqq \theta\Big(\lambda - \frac{n}{p}\Big) - (1-\theta)\Big(1 - \mu + \frac{n}{r}\Big) \tag{6.2}$$

holds true. Suppose that $f \in W^{1+\lambda,p}(B_\varrho(x_o)) \cap W^{\mu,r}(B_\varrho(x_o))$. Then, $Df \in L^s(B_\vartheta(x_o))$ for any radius $0 < \vartheta < \varrho$, and there exists a constant $c = c(n,\mu,\lambda,r,p,s,\theta,1/(\varrho-\vartheta))$ such that

$$\|Df\|_{L^s(B_\vartheta(x_o))} \leqq c\,\|f\|^{\theta}_{W^{1+\lambda,p}(B_\varrho(x_o))}\|f\|^{1-\theta}_{W^{\mu,r}(B_\varrho(x_o))}.$$

Proof. We choose a cut-off function $\eta \in C^2_0(B_\varrho(x_o),[0,1])$ such that $\eta \equiv 1$ on $B_\vartheta(x_o)$ and $\|\eta\|_\infty + (\varrho-\vartheta)\|D\eta\|_\infty + (\varrho-\vartheta)^2\|D^2\eta\|_\infty \leqq c$.

Next, we choose α according to

$$\frac{n+\alpha s}{ns} = \frac{\theta}{p} + \frac{1-\theta}{r}. \tag{6.3}$$

Then, by (6.2) we have

$$\begin{aligned}\alpha &= \frac{\theta n}{p} + \frac{(1-\theta)n}{r} - \frac{n}{s}\\ &\leqq \frac{\theta n}{p} + \frac{(1-\theta)n}{r} + \theta\Big(\lambda - \frac{n}{p}\Big) - (1-\theta)\Big(1-\mu+\frac{n}{r}\Big)\\ &= \theta\lambda - (1-\theta)(1-\mu) < 1.\end{aligned} \tag{6.4}$$

Since $p, r < s$, we also have that

$$\alpha = \theta\Big(\frac{n}{p} - \frac{n}{s}\Big) + (1-\theta)\Big(\frac{n}{r} - \frac{n}{s}\Big) > 0.$$

Therefore, we can apply Lemma 6.3 with $(\alpha, \frac{ns}{n+\alpha s})$ instead of (α, p) to $f\eta$ to infer that

$$\|Df\|_{L^s(B_\vartheta(x_o))} \leqq \|D(f\eta)\|_{L^s(\mathbb{R}^n)} \leqq c(n,s,\theta)\,[D(f\eta)]_{\alpha,\frac{ns}{n+\alpha s};\mathbb{R}^n}.$$

By our choice of α from (6.3) we are allowed to apply Lemma 6.2 with $(1+\alpha, 1+\lambda, \mu, \frac{ns}{n+\alpha s}, p, r)$ instead of $(\alpha, \alpha_1, \alpha_2, p, p_1, p_2)$ to the right-hand side. Note that at this stage we have taken into account that $1+\alpha \leqq \theta(1+\lambda) + (1-\theta)\mu$, which is a consequence of (6.4) and therefore of (6.2). We also mention that we can always apply Lemma 6.2 to values of α (in our case $1+\alpha$) for which $\leqq$ holds in $(6.1)_1$. This is a consequence of the embedding $W^{\theta(1+\lambda)+(1-\theta)\mu,p}(B_\varrho(x_o)) \hookrightarrow W^{1+\alpha,p}(B_\varrho(x_o))$ and the fact that ηf has its support in $B_\varrho(x_o)$; this is relevant only in the case that $1+\alpha < \theta(1+\lambda)+(1-\theta)\mu$. The application of Lemma 6.2, therefore, yields that

$$\begin{aligned}\|Df\|_{L^s(B_r(x_o))} &\leqq c\,\|f\eta\|^{\theta}_{W^{1+\lambda,p}(\mathbb{R}^n)}\|f\eta\|^{1-\theta}_{W^{\mu,r}(\mathbb{R}^n)}\\ &\leqq c\,\|f\|^{\theta}_{W^{1+\lambda,p}(B_\varrho(x_o))}\|f\|^{1-\theta}_{W^{\mu,r}(B_\varrho(x_o))}\end{aligned}$$

holds true for a constant $c = c(n,\mu,\lambda,r,p,s,\theta,1/(\varrho-\vartheta))$. □

We also need a parabolic version of fractional Sobolev spaces. We say that $u \in L^p(0, T; W^{k,p}(\Omega, \mathbb{R}^k))$ with $1 \leqq p < \infty$, $k \in \mathbb{N}_0$, $\alpha \in (0, 1)$ belongs to the parabolic fractional Sobolev space $L^p(0, T; W^{k+\alpha,p}(\Omega, \mathbb{R}^k))$ if the parabolic Gagliardo semi-norm

$$[D^\beta u]^p_{\alpha,0,p;\Omega_T} := \int_0^T \int_\Omega \int_\Omega \frac{|D^\beta u(x,t) - D^\beta u(y,t)|^p}{|x-y|^{n+\alpha p}} \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}t$$

is finite for any multiindex $\beta \in \mathbb{N}_0^n$ with $|\beta| = k$. As in the time independent elliptic setting, $L^p(0, T; W^{k+\alpha,p}(\Omega, \mathbb{R}^k))$ becomes a Banach-space with the norm

$$\|u\|_{k+\alpha,0,p;\Omega_T} := \|u\|_{L^p(0,T;W^{k,p}(\Omega,\mathbb{R}^k))} + \sum_{|\beta|=k} [D^\beta f]_{\alpha,0,p;\Omega_T}.$$

The next lemma provides an embedding result for the fractional parabolic Sobolev spaces and is an immediate consequence of the fractional Gagliardo–Nirenberg inequality from Lemma 6.4.

Lemma 6.5. (Parabolic fractional Sobolev inequality) *Let $B_\varrho(z_o) \times (t_1, t_2) \subset \mathbb{R}^{n+1}$ be a general space-time cylinder with $\varrho \leqq 1$ and $\lambda, \mu \in (0, 1)$, $1 < p, r < s < \infty$ parameters such that*

$$(s-p)\Big(1 - \mu + \frac{n}{r}\Big) \leqq \lambda p. \tag{6.5}$$

Further, assume that $u \in L^p(t_1, t_2; W^{1+\lambda,p}(B_\varrho(x_o))) \cap L^\infty(t_1, t_2; W^{\mu,r}(B_\varrho(x_o)))$. Then, $Du \in L^s(B_\vartheta(x_o) \times (t_1, t_2))$ for any $0 < \vartheta < \varrho$ and, moreover, the quantitative estimate

$$\|Du\|_{L^s(B_\vartheta(x_o)\times(t_1,t_2))} \leqq c\, \|u\|^{\frac{p}{s}}_{L^p(t_1,t_2;W^{1+\lambda,p}(B_\varrho(x_o)))} \sup_{t\in(t_1,t_2)} \|u(\cdot,t)\|^{\frac{s-p}{s}}_{W^{\mu,r}(B_\varrho(x_o))}$$

holds true with a constant $c = c(n, \mu, \lambda, r, p, s, 1/(\varrho - \vartheta))$.

Proof. For almost every $t \in (t_1, t_2)$ we have $u(\cdot, t) \in W^{1+\lambda,p}(B_\varrho(x_o)) \cap W^{\mu,r}(B_\varrho(x_o))$. Moreover, assumption (6.5) implies that hypothesis (6.2) in Lemma 6.4 is fulfilled for the choice $\theta = \frac{p}{s}$. Applying Lemma 6.4 slicewise to $u(\cdot, t)$ on $B_\varrho(x_o)$ we obtain

$$\begin{aligned}
&\int_{t_1}^{t_2} \int_{B_\vartheta(x_o)} |Du|^s \, \mathrm{d}x \, \mathrm{d}t \\
&\leqq c \int_{t_1}^{t_2} \|u(\cdot,t)\|^{\theta s}_{W^{1+\lambda,p}(B_\varrho(x_o))} \|u(\cdot,t)\|^{(1-\theta)s}_{W^{\mu,r}(B_\varrho(x_o))} \, \mathrm{d}t \\
&= c \int_{t_1}^{t_2} \|u(\cdot,t)\|^{p}_{W^{1+\lambda,p}(B_\varrho(x_o))} \|u(\cdot,t)\|^{s-p}_{W^{\mu,r}(B_\varrho(x_o))} \, \mathrm{d}t \\
&\leqq c \int_{t_1}^{t_2} \|u(\cdot,t)\|^{p}_{W^{1+\lambda,p}(B_\varrho(x_o))} \, \mathrm{d}t \sup_{t\in(t_1,t_2)} \|u(\cdot,t)\|^{s-p}_{L^{\mu,r}(B_\varrho(x_o))},
\end{aligned}$$

where c stands for the constant from Lemma 6.4 and therefore depends on n, μ, λ, r, p, s and $1/(\varrho - \vartheta)$. This proves the assertion of the lemma. □

Finally, we need an elliptic and parabolic version of the embedding from Nikolskii spaces, which are defined via finite differences, into fractional Sobolev spaces. The first part of the following Lemma is a consequence of [3, 7.73], while the second one is taken from [14, Proposition 2.9].

Lemma 6.6. *Let* $k \in \mathbb{N}$, $\widetilde{\Omega} \Subset \Omega$, $\theta \in (0,1)$ *and* $0 \leqq t_1 < t_2 \leqq T$.

(i) *Assume that* $u \in L^\infty(0,T; L^2(\Omega,\mathbb{R}^k))$ *satisfies*

$$\sup_{t\in(t_1,t_2)} \int_{\widetilde{\Omega}} |u(x+he_i,t) - u(x,t)|^2 \,\mathrm{d}x \leqq M|h|^{2\theta}$$

for every $i \in \{1,\dots,n\}$ *and* $h \in \mathbb{R}$ *with* $|h| \leqq \min\{\operatorname{dist}(\widetilde{\Omega},\partial\Omega), A\}$*, where* $A, M > 0$*. Then for every* $\alpha \in (0,\theta)$ *and* $\mathcal{O} \Subset \widetilde{\Omega}$ *there exists a constant* $c = c(n,\theta,\alpha,A,\operatorname{dist}(\mathcal{O},\partial\widetilde{\Omega}),\operatorname{dist}(\Omega,\widetilde{\Omega}))$ *such that*

$$\sup_{t\in(t_1,t_2)} [u(\cdot,t)]^2_{\alpha,p;\mathcal{O}} \equiv \sup_{t\in(t_1,t_2)} \int_{\mathcal{O}}\int_{\mathcal{O}} \frac{|u(x,t)-u(y,t)|^2}{|x-y|^{n+2\alpha}} \,\mathrm{d}x\,\mathrm{d}y \leqq c\,M.$$

(ii) *Assume that* $u \in L^p(\Omega_T,\mathbb{R}^k)$ *satisfies*

$$\int_{t_1}^{t_2}\int_{\widetilde{\Omega}} |u(x+he_i,t) - u(x,t)|^p \,\mathrm{d}x\,\mathrm{d}t \leqq M|h|^{\theta p}$$

for every $i \in \{1,\dots,n\}$ *and* $h \in \mathbb{R}$ *with* $|h| \leqq \min\{\operatorname{dist}(\widetilde{\Omega},\partial\Omega), A\}$*, where* $A, M > 0$*. Then for every* $\gamma \in (0,\theta)$ *and* $\mathcal{O} \Subset \widetilde{\Omega}$ *there exists a constant* $c = c(n,\theta,\gamma,A,\operatorname{dist}(\mathcal{O},\partial\widetilde{\Omega}),\operatorname{dist}(\Omega,\widetilde{\Omega}))$ *such that*

$$[u]^p_{\alpha,0,p;\mathcal{O}\times(t_1,t_2)} \equiv \int_{t_1}^{t_2}\int_{\mathcal{O}}\int_{\mathcal{O}} \frac{|u(x,t)-u(y,t)|^p}{|x-y|^{n+p\gamma}} \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t \leqq c\,M.$$

With these prerequisites at hand we can now start with the proof of the L^q-estimate for the gradient.

6.2. Caccioppoli Inequality for Finite Differences

The first step in most proofs of higher integrability for the spatial gradient is usually a *Caccioppoli inequality*, that is, an inequality of the type of a reverse Poincaré inequality. Since for systems of the type considered here we do not know that second spatial derivatives exist, we need a version for finite differences of Du. By $\tau_{h,i}[v]$ with $i \in \{1,\dots,n\}$ we denote the finite difference of a function v in the spatial direction e_i with increment h, that is, for $v \in L^1(\Omega_T)$ we define

$$\tau_{h,i}[v](x,t) := v(x+he_i,t) - v(x,t)$$

for $(x,t) \in \Omega_T$ such that also $(x+he_i,t) \in \Omega_T$. Then we have the following

Lemma 6.7. (Caccioppoli inequality) *Let*

$$u \in L^p\big(0, T; W^{1,p}(\Omega, \mathbb{R}^N)\big) \cap L^q_{\mathrm{loc}}\big(0, T; W^{1,q}_{\mathrm{loc}}(\Omega, \mathbb{R}^N)\big) \cap C^0\big([0, T]; L^2(\Omega, \mathbb{R}^N)\big)$$

be a weak solution to (2.7) *where the structural conditions* (2.9) *are in force. Then, for any parabolic cylinder* $Q_\varrho(z_o) \Subset \Omega_T$, *any* $0 < r < \varrho$, *any* $0 < |h| <$ $\operatorname{dist}(B_\varrho(x_o), \partial\Omega)$ *and any* $i \in \{1, \dots, n\}$ *there holds*

$$\begin{aligned} &\sup_{t \in (t_o - r^2, t_o)} \int_{B_r(x_o)} |\tau_{h,i}[u](\cdot, t)|^2 \,\mathrm{d}x + \int_{Q_r(z_o)} |\tau_{h,i}[Du]|^p \,\mathrm{d}z \\ &\quad \leqq \frac{c}{(\varrho - r)^2} \int_{Q_\varrho(z_o)} \big(|Du| + |\tau_{h,i}[Du]|\big)^{q-2} |\tau_{h,i}[u]|^2 + |\tau_{h,i}[u]|^2 \,\mathrm{d}z, \end{aligned}$$

with a constant $c = c(\nu, L, p, q)$.

Proof. Without loss of generality, we assume that $z_o = 0$ and write Q_ϱ instead of $Q_\varrho(0)$. In the weak formulation of (2.7), that is, in

$$\int_{\Omega_T} u \cdot \varphi_t - \langle Df(Du), D\varphi \rangle \,\mathrm{d}z = 0 \qquad \text{for all } \varphi \in C_0^\infty(\Omega_T, \mathbb{R}^N),$$

we replace φ by $\tau_{-h,i}[\varphi]$ with $0 < |h| \ll 1$ and perform an “integration by parts for finite differences”. This leads us to

$$\int_{\Omega_T} \tau_{h,i}[u] \cdot \varphi_t - \big\langle \tau_{h,i}[Df(Du)], D\varphi \big\rangle \,\mathrm{d}z = 0$$

for all $\varphi \in C_0^\infty(\Omega_T, \mathbb{R}^N)$ and $|h|$ small enough. In the following, we shall proceed formally concerning the use of the time derivative $\partial_t u$. However, the arguments can be made rigorous by the use of a mollifying procedure with respect to time, as for instance by Steklov averages. Since these arguments are standard we shall omit them and proceed formally. In the preceding identity we choose the testing function $\varphi(x, t) = \tau_{h,i}[u](x, t)\eta^2(x)\zeta(t)\chi_\theta(t)$, where $\eta \in C_0^1(B_\varrho, [0, 1])$, $\zeta \in W^{1,\infty}(\mathbb{R}, [0, 1])$ and $\chi_\theta \in W^{1,\infty}(\mathbb{R}, [0, 1])$ are cut-off functions. The spatial cut-off function η satisfies $\eta \equiv 1$ on B_r, $|D\eta| \leqq 2/(\varrho - r)$, while ζ is defined by

$$\zeta(t) := \begin{cases} 0 & \text{for } t \in (-\infty, -\varrho^2] \\ \frac{1}{\varrho^2 - r^2}(t + \varrho^2) & \text{for } t \in [-\varrho^2, -r^2) \\ 1 & \text{for } t \in [-r^2, \infty) \end{cases}$$

and χ_θ is given by

$$\chi_\theta(t) := \begin{cases} 1 & \text{for } t \in (-\infty, \tau - \theta] \\ \frac{1}{\theta}(\tau - t) & \text{for } t \in (\tau - \theta, \tau] \\ 0 & \text{for } t \in (\tau, 0] \end{cases}, \tag{6.6}$$

for some $\tau \in (-r^2, 0)$ and $\theta \in (0, r^2 + \tau)$. With this choice of the testing function φ we obtain from the last identity

$$\begin{aligned} &-\int_{Q_\varrho} \tau_{h,i}[u] \cdot \partial_t(\tau_{h,i}[u]\zeta\chi_\theta)\eta^2 \,\mathrm{d}z + \int_{Q_\varrho} \big\langle \tau_{h,i}[Df(Du)], \tau_{h,i}[Du]\big\rangle \eta^2 \zeta\chi_\theta \,\mathrm{d}z \\ &\quad = -\int_{Q_\varrho} \big\langle \tau_{h,i}[Df(Du)], \nabla\eta^2 \otimes \tau_{h,i}[u]\big\rangle \zeta\chi_\theta \,\mathrm{d}z. \end{aligned} \tag{6.7}$$

For the first term on the left-hand side we compute for almost every $\tau \in (-r^2, 0)$ that

$$\begin{aligned} &-\int_{Q_\varrho} \tau_{h,i}u \cdot \partial_t(\tau_{h,i}[u]\zeta\chi_\theta)\eta^2 \,\mathrm{d}z = \int_{Q_\varrho} \partial_t\tau_{h,i}[u] \cdot \tau_{h,i}[u]\eta^2\zeta\chi_\theta \,\mathrm{d}z \\ &= \tfrac{1}{2}\int_{Q_\varrho} \partial_t|\tau_{h,i}[u]|^2\eta^2\zeta\chi_\theta \,\mathrm{d}z = -\tfrac{1}{2}\int_{Q_\varrho} |\tau_{h,i}[u]|^2\eta^2\partial_t(\zeta\chi_\theta) \,\mathrm{d}z \\ &= -\frac{1}{2(\varrho^2 - r^2)}\int_{-\varrho^2}^{-r^2}\int_{B_\varrho} |\tau_{h,i}[u]|^2\eta^2 \,\mathrm{d}x\,\mathrm{d}t + \frac{1}{2\theta}\int_{\tau-\theta}^{\tau}\int_{B_\varrho} |\tau_{h,i}[u]|^2\eta^2 \,\mathrm{d}x\,\mathrm{d}t \\ &\to -\frac{1}{2(\varrho^2 - r^2)}\int_{-\varrho^2}^{-r^2}\int_{B_\varrho} |\tau_{h,i}[u]|^2\eta^2 \,\mathrm{d}x\,\mathrm{d}t + \tfrac{1}{2}\int_{B_\varrho} |\tau_{h,i}[u](\cdot,\tau)|^2\eta^2 \,\mathrm{d}x, \end{aligned}$$

in the limit $\theta \downarrow 0$. Passing also in the other terms in (6.7) to the limit $\theta \downarrow 0$ and taking into account that $1/(\varrho^2 - r^2) \leqq 1/(\varrho - r)^2$, we infer

$$\begin{aligned} &\tfrac{1}{2}\mathrm{I} + \mathrm{II} := \tfrac{1}{2}\int_{B_\varrho} |\tau_{h,i}[u](\cdot,\tau)|^2\eta^2 \,\mathrm{d}x + \int_{Q_\varrho^\tau} \big\langle \tau_{h,i}[Df(Du)], \tau_{h,i}[Du]\big\rangle \eta^2\zeta \,\mathrm{d}z \\ &\leqq -2\int_{Q_\varrho^\tau} \big\langle \tau_{h,i}[Df(Du)], \nabla\eta \otimes \tau_{h,i}[u]\big\rangle \eta\zeta \,\mathrm{d}z + \frac{1}{2(\varrho - r)^2}\int_{Q_\varrho} |\tau_{h,i}[u]|^2\eta^2 \,\mathrm{d}z \\ &=: \mathrm{III} + \mathrm{IV}, \end{aligned} \tag{6.8}$$

where we introduced the shorthand notion $Q_\varrho^\tau := B_\varrho \times (-\varrho^2, \tau)$. We rewrite the term II in the following way:

$$\mathrm{II} = \int_{Q_\varrho^\tau}\int_0^1 \big\langle D^2 f(Du + s\tau_{h,i}[Du])\tau_{h,i}[Du], \tau_{h,i}[Du]\big\rangle \eta^2\zeta \,\mathrm{d}s\,\mathrm{d}z.$$

Similarly, in the term III we rewrite the finite difference $\tau_{h,i}[Df(Du)]$ and then use the Cauchy–Schwartz inequality for the symmetric bilinear form $(\sigma, \widetilde{\sigma}) \mapsto \langle D^2 f(Du)\sigma, \widetilde{\sigma}\rangle$. This leads us to the estimate

$$\begin{aligned} \mathrm{III} &= -2\int_{Q_\varrho^\tau}\int_0^1 \big\langle D^2 f(Du + s\tau_{h,i}[Du])\tau_{h,i}[Du], \nabla\eta \otimes \tau_{h,i}[u]\big\rangle \eta\zeta \,\mathrm{d}s\,\mathrm{d}z \\ &\leqq \tfrac{1}{2}\mathrm{II} + 2\int_{Q_\varrho^\tau}\int_0^1 \big\langle D^2 f(Du + s\tau_{h,i}[Du])\nabla\eta \otimes \tau_{h,i}[u], \nabla\eta \otimes \tau_{h,i}[u]\big\rangle \zeta \,\mathrm{d}s\,\mathrm{d}z. \end{aligned}$$

We use the preceding inequality in (6.8) and reabsorb the integral $\frac{1}{2}$II on the left-hand side. Subsequently, we use assumption $(2.9)_2$ and the choice of η, that is, $|D\eta| \leqq 2/(\varrho - r)$ to infer that

$$
\begin{aligned}
\mathrm{I}+\mathrm{II} &\leqq 4 \int_{Q_\varrho^\tau} \int_0^1 \langle D^2 f(Du + s\tau_{h,i}[Du])\nabla\eta \otimes \tau_{h,i}[u], \nabla\eta \otimes \tau_{h,i}u\rangle \zeta \,\mathrm{d}s\,\mathrm{d}z + 2\,\mathrm{IV} \\
&\leqq 4L \int_{Q_\varrho^\tau} \int_0^1 \big(1 + |Du + s\tau_{h,i}[Du]|^{q-2}\big)\big|\nabla\eta \otimes \tau_{h,i}[u]\big|^2 \,\mathrm{d}s\,\mathrm{d}z + 2\,\mathrm{IV} \\
&\leqq \frac{c(q,L)}{(\varrho - r)^2} \int_{Q_\varrho^\tau} \Big[1 + \big(|Du|^2 + |\tau_{h,i}[Du]|^2\big)^{\frac{q-2}{2}}\Big] |\tau_{h,i}[u]|^2 \,\mathrm{d}z + 2\,\mathrm{IV}.
\end{aligned}
$$

With the help of $(2.9)_3$, Lemma 3.3 and the fact that we restrict ourselves to the case $p \geqq 2$, we estimate the term II from below as follows:

$$
\begin{aligned}
\mathrm{II} &\geqq \nu \int_{Q_\varrho^\tau} \int_0^1 |Du + s\tau_{h,i}[Du]|^{p-2} |\tau_{h,i}[Du]|^2 \eta^2 \zeta \,\mathrm{d}s\,\mathrm{d}z \\
&\geqq \tfrac{\nu}{c} \int_{Q_\varrho^\tau} \big(|Du(x,t)|^2 + |Du(x + he_i, t)|^2\big)^{\frac{p-2}{2}} \big|\tau_{h,i}[Du]\big|^2 \eta^2 \zeta \,\mathrm{d}z \\
&\geqq \tfrac{\nu}{c} \int_{Q_\varrho^\tau} |\tau_{h,i}[Du]|^p \eta^2 \zeta \,\mathrm{d}z
\end{aligned}
$$

for a constant $c = c(p)$. Inserting this above we see that

$$
\begin{aligned}
&\int_{B_\varrho} |\tau_{h,i}[u](\cdot,\tau)|^2 \eta^2 \,\mathrm{d}x + \int_{Q_\varrho^\tau} |\tau_{h,i}[Du]|^p \eta^2 \zeta \,\mathrm{d}z \\
&\quad\leqq \frac{c}{(\varrho - r)^2} \int_{Q_\varrho} \big(|Du| + |\tau_{h,i}[Du]|\big)^{q-2} |\tau_{h,i}[u]|^2 + |\tau_{h,i}[u]|^2 \,\mathrm{d}z
\end{aligned}
$$

holds true with a constant $c = c(\nu, L, p, q)$. Note that the preceding inequality holds for almost every $\tau \in (-r^2, 0)$. Therefore, we can use it in two different directions: In the first term on the left-hand side we take the supremum over $\tau \in (-r^2, 0)$, while for the second one we let $\tau \uparrow 0$. Proceeding in this way and taking into account the properties of the cut-off functions η, ζ, particularly that $\eta \equiv 1$ in B_r and $\zeta \equiv 1$ in $(-r^2, 0)$, we conclude the desired Caccioppoli inequality. □

6.3. Quantitative Higher Integrability

In this section we provide a first quantitative higher integrability estimate. The structure of the estimate is as follows: for any exponent $\sigma \in [p, q)$ we prove that there exists $\mathcal{S}(\sigma) > \sigma$ such that for any integrability exponent $s < \mathcal{S}(\sigma)$ the local L^s -energy of Du can be bounded in terms of the local L^σ-energy of Du. The precise result is as follows

Lemma 6.8. (Improvement of integrability) *Let*

$$u \in L^p\big(0,T;\, W^{1,p}(\Omega,\mathbb{R}^N)\big) \cap L^q_{\mathrm{loc}}\big(0,T;\, W^{1,q}_{\mathrm{loc}}(\Omega,\mathbb{R}^N)\big) \cap C^0\big([0,T]; L^2(\Omega,R^N)\big)$$

be a weak solution of (2.7) *where the structural conditions* (2.9) *are in force. Then, for any* $\sigma \in [p,q)$, *any cylinder* $Q_\varrho(z_o) \Subset \Omega_T$, *any radius* $r \in [\varrho/2,\varrho)$ *and any*

$$s < \mathcal{S}(\sigma) := p + \frac{4n^2 - 2n(n+2)(q-\sigma)}{n^3 + (n^2-4)(q-\sigma)}, \tag{6.9}$$

the quantitative higher integrability estimate

$$\int_{Q_r(z_o)} |Du|^s \,\mathrm{d}z \leqq c\left(\int_{Q_\varrho(z_o)} |Du|^\sigma \,\mathrm{d}z + M_{z_o,\varrho}\right)^{\frac{n+2}{n}} \tag{6.10}$$

holds true with a constant $c = c(n,\nu,L,p,q,\sigma,s,\varrho,r)$. *Here, we have abbreviated*

$$M_{z_o,\varrho} := 1 + \sup_{t\in(t_o-\varrho^2,t_o)} \int_{B_\varrho(x_o)} |u(\cdot,t)|^2 \,\mathrm{d}x + \int_{Q_\varrho(z_o)} |u|^p \,\mathrm{d}z. \tag{6.11}$$

Proof. We choose two radii ϱ_1, ϱ_2 such that $\frac{\varrho+r}{2} \leqq \varrho_1 < \varrho_2 \leqq \frac{3\varrho+r}{4}$ and consider increments $0 < |h| < \frac14(\varrho - r)$. These choices ensure that $|h| < \varrho - \varrho_2$, and therefore $x + he_i \in B_\varrho(x_o)$, whenever $x \in B_{\varrho_2}(x_o)$ and $i \in \{1,\dots,n\}$. Next, we define

$$\alpha := 2 - \frac{(n+2)(q-\sigma)}{n} \in (0,2). \tag{6.12}$$

Note that $\alpha > 0$ by our hypothesis (2.8). We now apply the Caccioppoli type inequality from Lemma 6.7 with (ϱ_1,ϱ_2) instead of (r,ϱ), Hölder's inequality and a standard estimate for finite differences to obtain, for any $i \in \{1,\dots,n\}$, that

$$\begin{aligned}
&\sup_{t\in(t_o-\varrho_1^2,t_o)} \int_{B_{\varrho_1}(x_o)} |\tau_{h,i}[u](\cdot,t)|^2 \,\mathrm{d}x + \int_{Q_{\varrho_1}(z_o)} |\tau_{h,i}[Du]|^p \,\mathrm{d}z \\
&\quad\leqq \frac{c}{(\varrho_2-\varrho_1)^2} \int_{Q_{\varrho_2}(z_o)} \big(1 + |Du| + |\tau_{h,i} Du|\big)^{q-2} |\tau_{h,i} u|^\alpha |\tau_{h,i} u|^{2-\alpha} \,\mathrm{d}z \\
&\quad\leqq \frac{c}{(\varrho_2-\varrho_1)^2} \left(\int_{Q_{\varrho_2}(z_o)} \big(1 + |Du| + |\tau_{h,i}[Du]|\big)^\sigma\right)^{\frac{q-2}{\sigma}} \\
&\qquad\times \left(\int_{Q_{\varrho_2}(z_o)} |\tau_{h,i}[u]|^\sigma \,\mathrm{d}z\right)^{\frac{\alpha}{\sigma}} \\
&\qquad\times \left(\int_{Q_{\varrho_2}(z_o)} |\tau_{h,i}[u]|^{\frac{\sigma(2-\alpha)}{\sigma-\alpha-(q-2)}} \,\mathrm{d}z\right)^{\frac{\sigma-\alpha-(q-2)}{\sigma}} \\
&\quad\leqq \frac{c\,|h|^\alpha}{(\varrho_2-\varrho_1)^2} \left(\int_{Q_\varrho(z_o)} (1 + |Du|)^\sigma\right)^{\frac{q+\alpha-2}{\sigma}} \\
&\qquad\times \left(\int_{Q_{\varrho_2}(z_o)} |\tau_{h,i}[u]|^{\frac{\sigma(2-\alpha)}{\sigma-\alpha-(q-2)}} \,\mathrm{d}z\right)^{\frac{\sigma-\alpha-(q-2)}{\sigma}}
\end{aligned} \tag{6.13}$$

holds with a constant $c = c(\nu, L, p, q)$. It now remains to estimate the last term on the right-hand side of (6.13). This will be achieved by an application of the parabolic Sobolev inequality from Lemma 6.1. We observe that by our choice of α in (6.12) we have

$$\frac{\sigma(2-\alpha)}{\sigma-\alpha-(q-2)} = \frac{\sigma(n+2)}{2}.$$

Therefore, applying Lemma 6.1 leads us to

$$\begin{aligned}
&\int_{Q_{\varrho_2}(z_o)} |\tau_{h,i}[u]|^{\frac{\sigma(2-\alpha)}{\sigma-\alpha-(q-2)}}\,\mathrm{d}z = \int_{Q_{\varrho_2}(z_o)} |\tau_{h,i}[u]|^{\frac{\sigma(n+2)}{2}}\,\mathrm{d}z \\
&\leqq c\int_{Q_{\varrho_2}(z_o)} |D[\tau_{h,i}u]|^{\sigma} + \left|\frac{\tau_{h,i}[u]}{\varrho_2}\right|^{\sigma}\mathrm{d}z \\
&\quad\times\left(\sup_{t\in(t_o-\varrho_2^2,t_o)}\int_{B_{\varrho_2}(x_o)} |\tau_{h,i}[u](\cdot,t)|^2\,\mathrm{d}x\right)^{\frac{\sigma}{n}},
\end{aligned}$$

for a constant $c = c(n, \sigma)$. Recalling that $|h| < \varrho - \varrho_2 < \varrho_2$ we can estimate the first integral on the right-hand side of the preceding inequality by

$$\begin{aligned}
\int_{Q_{\varrho_2}(z_o)} |D[\tau_{h,i}u]|^{\sigma} + \left|\frac{\tau_{h,i}[u]}{\varrho_2}\right|^{\sigma}\mathrm{d}z &\leqq \int_{Q_{\varrho}(z_o)} |Du|^{\sigma} + \left(\tfrac{|h|}{\varrho_2}\right)^{\sigma}|Du|^{\sigma}\,\mathrm{d}z \\
&\leqq 2\int_{Q_{\varrho}(z_o)} |Du|^{\sigma}\,\mathrm{d}z.
\end{aligned}$$

Inserting this into the second last inequality and joining the result with (6.13), we get

$$\begin{aligned}
&\sup_{t\in(t_o-\varrho_1^2,t_o)}\int_{B_{\varrho_1}(x_o)} |\tau_{h,i}[u](\cdot,t)|^2\,\mathrm{d}x + \int_{Q_{\varrho_1}(z_o)} |\tau_{h,i}[Du]|^p\,\mathrm{d}z \\
&\leqq \frac{c\,|h|^{\alpha}}{(\varrho_2-\varrho_1)^2}\int_{Q_{\varrho}(z_o)} (1+|Du|)^{\sigma}\,\mathrm{d}z \\
&\quad\times\left(\sup_{t\in(t_o-\varrho_2^2,t_o)}\int_{B_{\varrho_2}(x_o)} |\tau_{h,i}[u](\cdot,t)|^2\,\mathrm{d}x\right)^{\frac{\sigma-\alpha-(q-2)}{n}},
\end{aligned}$$

where $c = c(n, \nu, L, p, q, \sigma)$. To the right-hand side we apply Young's inequality and take into account that the Hölder conjugate of $n/(\sigma-\alpha-(q-2))$ is given by

$$\frac{n}{n-\sigma+\alpha+q-2} = \frac{n^2}{n^2-2(q-\sigma)}.$$

This leads us to

$$\sup_{t\in(t_o-\varrho_1^2,t_o)}\int_{B_{\varrho_1}(x_o)} |\tau_{h,i}[u](\cdot,t)|^2\,\mathrm{d}x + \int_{Q_{\varrho_1}(z_o)} |\tau_{h,i}[Du]|^p\,\mathrm{d}z$$

$$\leqq \tfrac{1}{2} \sup_{t\in(t_o-\varrho_2^2,t_o)} \int_{B_{\varrho_2}(x_o)} |\tau_{h,i}[u](\cdot,t)|^2\,\mathrm{d}x + c\left(\frac{|h|^\alpha}{(\varrho_2-\varrho_1)^2}\int_{Q_\varrho(z_o)} (|Du|+1)^\sigma\,\mathrm{d}z\right)^{\frac{n^2}{n^2-2(q-\sigma)}},$$

where $c = c(n, \nu, L, p, q, \sigma)$. For the application of Young's inequality we need to have that $2(q-\sigma) < n^2$, which is a consequence of (2.8). With the help of Lemma 3.1 we can reabsorb the sup-term from the right into the left-hand side, yielding that

$$\sup_{t\in(t_o-r_1^2,t_o)} \int_{B_{r_1}(x_o)} |\tau_{h,i}[u](\cdot,t)|^2\,\mathrm{d}x + \int_{Q_{r_1}(z_o)} |\tau_{h,i}[Du]|^p\,\mathrm{d}z \leqq c\left(\frac{|h|^\alpha}{(\varrho-r)^2}\int_{Q_\varrho(z_o)} (1+|Du|)^\sigma\,\mathrm{d}z\right)^{\frac{n^2}{n^2-2(q-\sigma)}},$$

where $r_1 := \frac{\varrho+r}{2}$. We also abbreviate $r_o := \frac{\varrho+3r}{4}$, so that $r < r_o < r_1$. Now, the embedding of parabolic Nikolskii spaces into fractional Sobolev spaces from Lemma 6.6 implies that

$$u \in L^p\big(t_o-r_o^2, t_o; W^{1+\lambda,p}(B_{r_o}(x_o),\mathbb{R}^N)\big) \cap L^\infty\big(t_o-r_o^2, t_o; W^{\mu,2}(B_{r_o}(x_o),\mathbb{R}^N)\big)$$

holds for any

$$\lambda < \frac{\alpha n^2}{p(n^2-2(q-\sigma))} =: \tilde\lambda \quad\text{and}\quad \mu < \frac{\alpha n^2}{2(n^2-2(q-\sigma))} =: \tilde\mu. \tag{6.14}$$

Note that $p\tilde\lambda = 2\tilde\mu$ and $\tilde\lambda \leqq \tilde\mu < 1$. Furthermore, the Lemma also implies the quantitative estimates

$$[Du]^p_{\lambda,0,p;Q_{r_o}(x_o)} \leqq c\left(\int_{Q_\varrho(z_o)} (1+|Du|)^\sigma\,\mathrm{d}z\right)^{\frac{n^2}{n^2-2(q-\sigma)}}$$

and

$$\sup_{t\in(t_o-r_o^2,t_o)} [u(\cdot,t)]^2_{\mu,2;B_{r_o}(x_o)} \leqq c\left(\int_{Q_\varrho(z_o)} (1+|Du|)^\sigma\,\mathrm{d}z\right)^{\frac{n^2}{n^2-2(q-\sigma)}}.$$

Note that $c = c(n, \nu, L, p, q, \sigma, \mu, \lambda, \varrho, r)$. These bounds allow us to estimate the L^p–$W^{1+\lambda,p}$-norm of u on $Q_{r_o}(z_o)$ as follows:

$$\begin{aligned}
&\|u\|_{L^p(t_o-r_o^2,t_o;W^{1+\lambda,p}(B_{r_o}(x_o)))} \\
&\quad = \|u\|_{L^p(Q_{r_o}(z_o))} + \|Du\|_{L^p(Q_{r_o}(z_o))} + [Du]_{\lambda,0,p;Q_{r_o}(z_o)} \\
&\quad \leqq c\left(\|u\|_{L^p(Q_\varrho(z_o))} + \|Du\|_{L^\sigma(Q_\varrho(z_o))} + \|Du\|_{L^\sigma(Q_\varrho(z_o))}^{\frac{\sigma n^2}{p(n^2-2(q-\sigma))}} + 1\right) \\
&\quad \leqq c\left(\int_{Q_\varrho(z_o)} |Du|^\sigma\,\mathrm{d}z + M_{z_o,\varrho}\right)^{\frac{n^2}{p(n^2-2(q-\sigma))}}.
\end{aligned}$$

Similarly, for the $L^\infty{-}W^{\mu,2}$-norm of u we obtain

$$\begin{aligned}
&\sup_{t\in(t_o-r_o^2,t_o)} \|u(\cdot,t)\|_{W^{\mu,2}(B_{r_o}(x_o))}\\
&\quad= \sup_{t\in(t_o-r_o^2,t_o)} \|u(\cdot,t)\|_{L^2(B_{r_o}(x_o))} + \sup_{t\in(t_o-r_o^2,t_o)} [u(\cdot,t)]_{\mu,2;B_{r_o}(x_o)}\\
&\quad\leqq c\bigg(\sup_{t\in(t_o-\varrho^2,t_o)} \|u(\cdot,t)\|_{L^2(B_\varrho(x_o))} + \|Du\|_{L^\sigma(Q_\varrho(z_o))}^{\frac{\sigma n^2}{2(n^2-2(q-\sigma))}} + 1\bigg)\\
&\quad\leqq c\bigg(\int_{Q_\varrho(z_o)} |Du|^\sigma\,\mathrm{d}z + M_{z_o,\varrho}\bigg)^{\frac{n^2}{2(n^2-2(q-\sigma))}}.
\end{aligned}$$

We now want to apply Lemma 6.5. For this we first need to check that hypothesis (6.5) is satisfied. By the definitions of $\mathcal{S}(\sigma),\ \tilde\lambda,\ \alpha$, the fact that $\tilde\mu = p\tilde\lambda/2$ and since $s < \mathcal{S}(\sigma)$, we have

$$\begin{aligned}
&(s-p)\Big(1-\tilde\mu+\frac{n}{2}\Big)\frac{1}{\tilde\lambda p} < (\mathcal{S}(\sigma)-p)\Big(1-\tilde\mu+\frac{n}{2}\Big)\frac{1}{\tilde\lambda p}\\
&= \frac{4n^2-2n(n+2)(q-\sigma)}{n^3+(n^2-4)(q-\sigma)}\cdot\Big(1-\frac{\tilde\lambda p}{2}+\frac{n}{2}\Big)\frac{1}{\tilde\lambda p}\\
&= 2\cdot\frac{2n^2-n(n+2)(q-\sigma)}{n^3+(n^2-4)(q-\sigma)}\cdot\Big(\frac{n+2}{2\tilde\lambda p}-\frac{1}{2}\Big)\\
&= 2\cdot\frac{2n^2-n(n+2)(q-\sigma)}{n^3+(n^2-4)(q-\sigma)}\cdot\Big(\frac{(n+2)(n^2-2(q-\sigma))}{2\alpha n^2}-\frac{1}{2}\Big)\\
&= \frac{2n^2-n(n+2)(q-\sigma)}{n^3+(n^2-4)(q-\sigma)}\cdot\Big(\frac{(n+2)(n^2-2(q-\sigma))}{\alpha n^2}-1\Big)\\
&= \frac{2n^2-n(n+2)(q-\sigma)}{n^3+(n^2-4)(q-\sigma)}\cdot\Big(\frac{(n+2)(n^2-2(q-\sigma))}{2n^2-n(n+2)(q-\sigma)}-1\Big)\\
&= \frac{(n+2)(n^2-2(q-\sigma))-2n^2+n(n+2)(q-\sigma)}{n^3+(n^2-4)(q-\sigma)}\\
&= \frac{n^3-2(n+2)(q-\sigma)+n(n+2)(q-\sigma)}{n^3+(n^2-4)(q-\sigma)} = 1.
\end{aligned}$$

Therefore, we can choose λ and μ according to (6.14) in such a way that (6.5) is satisfied with $r=2$, that is, that

$$(s-p)\Big(1-\mu+\frac{n}{2}\Big) \leqq \lambda p$$

holds true. This is possible since the preceding strict inequality is valid with $\tilde\mu$ and $\tilde\lambda$. This fixes λ, μ in dependence on n, p, q, σ and s. The fractional embedding from Lemma 6.5 therefore implies $Du \in L^s(Q_r(z_o), \mathbb{R}^{Nn})$ for any integrability exponent s as in (6.9). Moreover, the following quantitative higher integrability estimate holds:

$$\begin{aligned}\int_{Q_r(z_o)} |Du|^s \, \mathrm{d}z &\leqq c \, \|u\|^p_{L^p(t_o - r_o^2, t_o; W^{1+\lambda,p}(B_{r_o}(x_o)))} \sup_{t \in (t_o - r_o^2, t_o)} \|u(\cdot, t)\|^{s-p}_{W^{\mu,2}(B_{r_o}(x_o))} \\ &\leqq c \left(\int_{Q_\varrho(z_o)} |Du|^\sigma \, \mathrm{d}z + M_{\varrho, z_o} \right)^{\frac{n^2}{n^2 - 2(q-\sigma)} \cdot (1 + \frac{s-p}{2})},\end{aligned}$$

for a constant $c = c(n, \nu, L, p, q, \sigma, s, \varrho, r)$. The assumption $s < \mathcal{S}(\sigma)$ allows us to bound the exponent on the right-hand side as follows:

$$\begin{aligned}\frac{n^2}{n^2 - 2(q-\sigma)} \cdot \left(1 + \frac{s-p}{2}\right) &< \frac{n^2}{n^2 - 2(q-\sigma)} \cdot \left(1 + \frac{\mathcal{S}(\sigma) - p}{2}\right) \\ &= \frac{n^2}{n^2 - 2(q-\sigma)} \cdot \frac{(n+2)(n^2 - 2(q-\sigma))}{n^3 + (n^2 - 4)(q-\sigma)} \\ &= \frac{n^2(n+2)}{n^3 + (n^2-4)(q-\sigma)} \leqq \frac{n+2}{n}.\end{aligned}$$

Since $M_{z_o,\varrho} \geqq 1$ by definition, we can replace the exponent on the right-hand side of the second last inequality by $\frac{n+2}{n}$. This gives

$$\int_{Q_r(z_o)} |Du|^s \, \mathrm{d}z \leqq c \left(\int_{Q_\varrho(z_o)} |Du|^\sigma \, \mathrm{d}z + M_{\varrho, z_o} \right)^{\frac{n+2}{n}}$$

and proves the claim of the lemma. □

Remark 6.9. We note that the constant in the higher integrability estimate (6.10) blows up, that is, $c \uparrow \infty$, when $\sigma \uparrow q$ or $s \uparrow \mathcal{S}(\sigma)$ or $r \uparrow \varrho$. □

6.4. Uniform Improvement of the Integrability Exponent

Our aim in this section is to ensure that Lemma 6.8 in fact yields a uniform integrability improvement, in the sense that there exists a constant $\varepsilon_o = \varepsilon_o(n, q-p) > 0$ such that

$$\mathcal{S}(\sigma) \geqq \sigma + \varepsilon_o \quad \text{for any } \sigma \in [p, q). \tag{6.15}$$

The quantity $\mathcal{S}(\sigma)$ was defined in (6.9). We let $\delta := \sigma - p \in [0, q-p)$. We observe that (6.15) is equivalent to

$$\begin{aligned}&4n^2 - 2n(n+2)(q-p-\delta) - \delta\big(n^3 + (n^2-4)(q-p-\delta)\big) \\ &\geqq \varepsilon_o\big(n^3 + (n^2-4)(q-\sigma)\big).\end{aligned}$$

In the following we prove that there exists $\varepsilon_1 = \varepsilon_1(n, q-p) > 0$ such that

$$g(\delta) := 4n^2 - 2n(n+2)(q-p-\delta) - \delta\big(n^3 + (n^2-4)(q-p-\delta)\big) \geqq \varepsilon_1, \tag{6.16}$$

for any $\delta \in [0, q-p)$. Once (6.16) has been established, (6.15) can be concluded with $\varepsilon_o = \varepsilon_1/(2n^3)$, since $q - \sigma < 1$. To prove (6.16) we distinguish between the cases $n = 2, 3$ and $n \geqq 4$. If $n \in \{2, 3\}$, assumption (2.8) reads as $q - p < 1$ and therefore $\varepsilon := 1 - (q - p) > 0$. We now set $\varepsilon_1 := 2n(n + 2)\varepsilon$. Then we have

$$\begin{aligned}\varepsilon_1 &= 2n(n+2)\varepsilon \\ &\leqq 2n(n+2)\varepsilon + 2n^2 - 4n + (n^2 + 4n + 4 - n^3)\delta + (n^2 - 4)(\varepsilon + \delta)\delta \\ &= 4n^2 - 2n(n+2)(1 - \varepsilon - \delta) - \delta\big(n^3 + (n^2 - 4)(1 - \varepsilon - \delta)\big) = g(\delta).\end{aligned}$$

Here we used, in turn, the fact that $2n^2 - 4n + (n^2 + 4n + 4 - n^3)\delta \geqq 0$ for $n \in \{2, 3\}$ and $\delta \in [0, 1)$. This proves (6.16) with $\varepsilon_1 = 2n(n + 2)(1 - (q - p))$.

Next we consider the case $n \geqq 4$, where (2.8) turns into $q - p < \frac{4}{n}$. We consider the real valued function g defined in (6.16) for $\delta \in [0, q - p]$ and compute

$$g'(\delta) = 2n(n + 2) - n^3 - (n^2 - 4)(q - p) + 2(n^2 - 4)\delta.$$

Then, $g'(\delta) \leqq g'(q - p) < 0$ on $[0, q - p]$, so that g is strictly decreasing. This implies $g(\delta) > g(q - p) = 4n^2$ for any $\delta \in [0, q - p)$ and this establishes, as above, (6.16) for the choice $\varepsilon_1 = 4n^2$. Note that this implies that (6.15) holds true with the constant $\varepsilon_o = \frac{n+2}{n^2}(1 - (q - p))$ if $n = 2, 3$ (respectively $\varepsilon_o = 2/n$ if $n \geqq 4$).

6.5. Iteration

In this section we iterate the higher integrability estimate from Lemma 6.8 in order to obtain the local L^q-estimate for the gradient Du in terms of the local L^p-norm. This is possible, since the improvement in integrability in estimate (6.10) from Lemma 6.8 is uniform, as we have shown in (6.15). The precise result is as follows:

Proposition 6.10. *Let* $2 \leqq p < q$ *satisfy* (2.8) *and suppose that*

$$u \in L^p\big(0, T; W^{1,p}(\Omega, \mathbb{R}^N)\big) \cap L^q_{\mathrm{loc}}\big(0, T; W^{1,q}_{\mathrm{loc}}(\Omega, \mathbb{R}^N)\big) \cap C^0\big([0,T]; L^2(\Omega, \mathbb{R}^N)\big)$$

is a weak solution of the parabolic system (2.7) *where the structural conditions* (2.9) *are in force. Then, there exists a constant* $\chi = \chi(n, q - p) > 1$ *such that, for any cylinder* $Q_R(z_o) \Subset \Omega_T$*, there holds*

$$\int_{Q_{R/2}(z_o)} |Du|^q \,\mathrm{d}z \leqq c\left(\int_{Q_R(z_o)} |Du|^p \,\mathrm{d}z + M_{z_o,R}\right)^{\chi}, \tag{6.17}$$

where $M_{z_o,R}$ *is defined in* (6.11) *and* $c = c(n, \nu, L, p, q, R)$.

Proof. We consider a cylinder $Q_R(z_o) \Subset \Omega_T$ and define, for $i \in \{0, \ldots, I\}$, radii r_i and integrability exponents σ_i by

$$r_i := \frac{R}{2} + \frac{R}{2^{i+1}} \quad \text{and} \quad \sigma_i := p + i\,\frac{q - p}{I},$$

where

$$I := \left\lceil \frac{q - p}{\varepsilon_o} \right\rceil$$

and $\varepsilon_o = \varepsilon_o(n, q-p) > 0$ is the constant from (6.15). Then, $p = \sigma_o < \sigma_1 < \cdots < \sigma_{I-1} < \sigma_I = q$ and, moreover, $\sigma_{i+1} < \mathcal{S}(\sigma_i)$ for any $i \in \{0\dots, I-1\}$. Therefore, we can apply Lemma 6.8 and conclude that

$$\int_{Q_{r_{i+1}}(z_o)} |Du|^{\sigma_{i+1}}\,\mathrm{d}z \leqq c\left(\int_{Q_{r_i}(z_o)} |Du|^{\sigma_i}\,\mathrm{d}z + M_{z_o,R}\right)^{\frac{n+2}{n}}$$

holds for any $i \in \{0\dots, I-1\}$. Here, we also used that $M_{z_o,r_i} \leqq M_{z_o,R}$. Joining these estimates and taking into account that $M_{z_o,R} \geqq 1$, $r_I \geqq R/2$ and $r_o = R$, we find that

$$\begin{aligned}
\int_{Q_{R/2}(z_o)} |Du|^q\,\mathrm{d}z &\leqq \int_{Q_{r_I}(z_o)} |Du|^q\,\mathrm{d}z \\
&\leqq c\left(\int_{Q_{r_{I-1}}(z_o)} |Du|^{\sigma_{I-1}}\,\mathrm{d}z + M_{z_o,R}\right)^{\frac{n+2}{n}} \\
&\leqq c\left[\left(\int_{Q_{r_{I-2}}(z_o)} |Du|^{\sigma_{I-2}}\,\mathrm{d}z + M_{z_o,R}\right)^{\frac{n+2}{n}} + M_{z_o,R}\right]^{\frac{n+2}{n}} \\
&\leqq c\left(\int_{Q_{r_{I-2}}(z_o)} |Du|^{\sigma_{I-2}}\,\mathrm{d}z + M_{z_o,R}\right)^{(\frac{n+2}{n})^2} \\
&\leqq c\left(\int_{Q_R(z_o)} |Du|^p\,\mathrm{d}z + M_{z_o,R}\right)^{(\frac{n+2}{n})^I}.
\end{aligned}$$

This proves the asserted estimate (6.17) with $\chi = \chi(n, q-p) = \left(\frac{n+2}{n}\right)^I > 1$ and a constant c depending on n, ν, L, p, q and R. □

7. Proof of Theorems 2.6 and 2.8

Here, we start with the proof of the existence of weak solutions stated in Theorem 2.6. This will be achieved by constructing an approximating sequence of solutions to a regularized problem and then passing to the limit. The passage to the limit will be achieved thanks to the L^q-bound from Proposition 6.10.

Proof of Theorem 2.6. We shall proceed in several steps.
Step 1. Regularization. For $\varepsilon \in (0, 1]$ we define the regularized integrand f_ε by

$$f_\varepsilon(\xi) := f(\xi) + \varepsilon|\xi|^q \qquad \text{for } \xi \in \mathbb{R}^{Nn}.$$

From the properties (2.9) of the integrand f we infer the following growth and ellipticity properties for the regularized integrand f_ε:

$$\begin{cases}
\varepsilon|\xi|^q + |\xi|^p \leqq f_\varepsilon(\xi) \leqq (L+1)\,(1+|\xi|^q)\,, \\
|D^2 f_\varepsilon(\xi)| \leqq (L + q(q-1))\left(1+|\xi|^{q-2}\right), \\
\langle D^2 f_\varepsilon(\xi)\eta, \eta\rangle \geqq \nu\,|\xi|^{p-2}|\eta|^2 + \varepsilon q|\xi|^{q-2}|\eta|^2\,,
\end{cases} \tag{7.1}$$

whenever $\xi, \eta \in \mathbb{R}^{Nn}$. In particular, for every fixed $\varepsilon \in (0, 1]$, Df_ε satisfies standard q-growth conditions. Therefore, from the classical theory, see [20], we conclude the existence of a weak solution

$$u_\varepsilon \in L^q\big(0, T; W^{1,q}(\Omega, \mathbb{R}^N)\big) \cap C^0\big([0, T]; L^2(\Omega, \mathbb{R}^N)\big)$$

to the parabolic Cauchy–Dirichlet problem

$$\begin{cases} \partial_t u_\varepsilon - \operatorname{div}\big(Df_\varepsilon(Du_\varepsilon)\big) = 0 & \text{in } \Omega_T\,, \\ u_\varepsilon = g & \text{on } \partial_{\mathcal{P}}\Omega_T\,. \end{cases} \tag{7.2}$$

Step 2. Uniform bounds and weak convergence. In the following, we want to pass to the limit $\varepsilon \downarrow 0$. We first recall some facts from the proof of Theorem 2.3 (note that the parabolic system (7.2) coincides with (4.2)). Therefore, the solutions u_ε satisfy the energy bounds (4.5) and (4.6). Moreover, they admit the weak continuity property from (4.8). By the arguments of Section 4.5 we infer the existence of a function

$$u \in L^p\big(0, T; W_g^{1,p}(\Omega,\, \mathbb{R}^N)\big) \cap C_w\big([0, T]; L^2(\Omega,\, \mathbb{R}^N)\big),$$

with $u(\cdot,\ 0) = g(\cdot,\ 0)$ and a (not relabelled) subsequence, such that

$$\begin{cases} u_\varepsilon \rightharpoonup u & \text{weakly in } L^p(\Omega_T, \mathbb{R}^N), \\ Du_\varepsilon \rightharpoonup Du & \text{weakly in } L^p(\Omega_T, \mathbb{R}^{Nn}), \\ u_\varepsilon(\cdot,\ t) \rightharpoonup u(\cdot,\ t) & \text{weakly in } L^2(\Omega, \mathbb{R}^N) \text{for any } t \in [0, T]. \end{cases}$$

Next, we observe that Lemma 6.1 (more precisely, a version valid for the space-time cylinder Ω_T, cf. [12, Chapter I, Proposition 3.1]) yields

$$\int_{\Omega_T} |u_\varepsilon|^{\frac{p(n+2)}{n}}\,\mathrm{d}z \leqq c \int_{\Omega_T} \big(|Du_\varepsilon|^p + |u_\varepsilon|^p\big)\,\mathrm{d}z \left(\sup_{t\in(0,T)} \int_\Omega |u_\varepsilon(\cdot,\ t)|^2\,\mathrm{d}x\right)^{\frac{p}{n}},$$

for a constant $c = c(n, p, \operatorname{diam}(\Omega))$. From (4.5) and (4.6) we know that the right-hand side is uniformly bounded with respect to ε. Since $q < p + \frac{4}{n} \leqq p + \frac{2p}{n} = \frac{p(n+2)}{n}$ by (2.8) and the assumption $p \geqq 2$, the preceding inequality ensures that u_ε is bounded in $L^q(\Omega_T, \mathbb{R}^N)$ uniformly with respect to ε. Moreover, since f_ε satisfies (2.9) with $L + q(q-1)$ instead of L, we are allowed to apply Proposition 6.10 to u_ε. This yields the existence of constants $\chi = \chi(n, q-p)$ and $c = c(n, \nu,\ L, p, q, R)$ such that, for any cylinder $Q_R(z_o) \Subset \Omega_T$, there holds

$$\int_{Q_{R/2}(z_o)} |Du_\varepsilon|^q\,\mathrm{d}z \leqq c\Bigg[\sup_{t\in(t_o-R^2,t_o)} \int_{B_R(x_o)} |u_\varepsilon(\cdot, t)|^2\,\mathrm{d}x + \int_{Q_R(z_o)} \big(1 + |u_\varepsilon|^p + |Du_\varepsilon|^p\big)\,\mathrm{d}z\Bigg]^\chi. \tag{7.3}$$

By (4.5) and (4.6) we therefore conclude that Du_ε is uniformly bounded in L^q_{loc} $(\Omega_T, \mathbb{R}^{Nn})$. Together with the L^q-bound for u_ε from above, we conclude that

$u \in L^q_{\mathrm{loc}}(0, T; W^{1,q}_{\mathrm{loc}}(\Omega, \mathbb{R}^N))$ and that we can extract a further (not relabeled) subsequence, such that

$$\begin{cases} u_\varepsilon \rightharpoonup u & \text{weakly in } L^q(Q_o, \mathbb{R}^N) \text{ for any } Q_o \Subset \Omega_T. \\ Du_\varepsilon \rightharpoonup Du & \text{weakly in } L^q(Q_o, \mathbb{R}^{Nn}) \text{for any } Q_o \Subset \Omega_T. \end{cases} \tag{7.4}$$

At this point, we start again from the weak formulation of $(7.2)_1$ with a general testing function $\varphi \in C^\infty_0(\Omega_T, \mathbb{R}^N)$. Using the bound $|Df_\varepsilon(w)| \leqq c(q)L(1 + |w|^{q-1})$ (which follows from the convexity of f_ε and $(7.1)_1$), Hölder's inequality, (7.3) and the energy bound (4.5), we obtain that

$$\begin{aligned} \left| \int_{\Omega_T} u_\varepsilon \cdot \varphi_t \, \mathrm{d}z \right| &= \left| \int_{\Omega_T} \langle Df_\varepsilon(Du_\varepsilon), D\varphi \rangle \, \mathrm{d}z \right| \leqq c \int_{\Omega_T} (1 + |Du_\varepsilon|)^{q-1} |D\varphi| \, \mathrm{d}z \\ &\leqq c \bigg(\int_{\operatorname{spt} \varphi} (1 + |Du_\varepsilon|)^q \, \mathrm{d}z \bigg)^{\frac{q-1}{q}} \|D\varphi\|_{L^q(\Omega_T)} \leqq c \, \|D\varphi\|_{L^q(\Omega_T)}, \end{aligned}$$

for a constant c depending only on L, q, $|\operatorname{spt} \varphi|$ and $\sup_{0<\varepsilon\leqq 1} \|Du_\varepsilon\|_{L^q(\operatorname{spt}\varphi)}$. Since $\varphi \in C^\infty_0(\Omega_T, \mathbb{R}^N)$—and in particular $\operatorname{spt} \varphi \Subset \Omega_T$—was arbitrary, this shows that $\partial_t u_\varepsilon$ is uniformly bounded in $L^{q'}(t_1, t_2; W^{-1,q'}(\mathcal{O}, \mathbb{R}^N))$ for any $0 < t_1 < t_2 < T$ and $\mathcal{O} \Subset \Omega$. By the lower semicontinuity of the norm, this also implies that $u_t \in L^{q'}(t_1, t_2; W^{-1,q'}(\mathcal{O}, \mathbb{R}^N))$. Finally, the embedding

$$\begin{aligned} &\big\{ v \in L^q\big(t_1, t_2; W^{1,q}(\mathcal{O}, \mathbb{R}^N)\big) : v_t \in L^{q'}\big(t_1, t_2; W^{-1,q'}(\mathcal{O}, \mathbb{R}^N)\big) \big\} \\ &\quad \hookrightarrow C^0\big([t_1, t_2]; L^2(\mathcal{O}, \mathbb{R}^N)\big) \end{aligned}$$

guarantees that $u \in C^0([t_1, t_2]; L^2(\mathcal{O}, \mathbb{R}^N))$.

Step 3. Strong convergence. Due to (4.5) and (4.8) we can, on the one hand, apply [32, Theorem 6] with

$$(X, B, Y, p, q) = \big(W^{1,p}(\Omega, \mathbb{R}^N), L^2(\Omega, \mathbb{R}^N), W^{-\ell,2}(\Omega, \mathbb{R}^N), 2, \infty \big)$$

to conclude that (u_ε) is relatively compact in $L^2(\Omega_T, \mathbb{R}^N)$ and, on the other hand, we can apply [32, Theorem 5] with

$$(X, B, Y, p) = \big(W^{1,q}(\mathcal{O}, \mathbb{R}^N), L^p(\mathcal{O}, \mathbb{R}^N), W^{-\ell,2}(\mathcal{O}, \mathbb{R}^N), q \big),$$

yielding that (u_ε) is relatively compact in $L^p(\mathcal{O} \times (t_1, t_2), \mathbb{R}^N)$ for any $\mathcal{O} \times (t_1, t_2) \Subset \Omega_T$. Together, we infer the existence of a further, still not relabeled, subsequence such that

$$u_\varepsilon \to u \quad \text{strongly in } L^2(\Omega_T, \mathbb{R}^N) \text{ and } L^q(Q_o, \mathbb{R}^N) \quad \text{for any } Q_o \Subset \Omega_T. \tag{7.5}$$

In the following we will use this result to show that we, indeed, have strong convergence of Du_ε in $L^p_{\mathrm{loc}}(\Omega_T, \mathbb{R}^N)$, that is, that

$$Du_\varepsilon \to Du \quad \text{strongly in } L^p(Q_o, \mathbb{R}^{Nn}) \quad \text{for any} Q_o \Subset \Omega_T \tag{7.6}$$

and, further, that

$$u_\varepsilon(\cdot, t) \to u(\cdot, t) \quad \text{strongly in } L^2(\mathcal{O}, \mathbb{R}^N) \text{ for any } \mathcal{O} \Subset \Omega \text{ and any } t \in (t_1, t_2). \tag{7.7}$$

For this aim we consider $Q \equiv \mathcal{O} \times (t_1, t_2) \Subset \Omega_T$ with $\mathcal{O} \Subset \Omega$ and $0 < t_1 < t_2 < T$ and $\varphi \in C_0^\infty(\Omega_T, \mathbb{R}^N)$ with $\operatorname{spt}\varphi \subset Q$. From the weak form of (7.2)$_1$ we infer that

$$\begin{aligned} &\int_Q \big[(u_\varepsilon - u)\cdot \varphi_t - \langle Df_\varepsilon(Du_\varepsilon) - Df(Du), D\varphi\rangle\big]\,\mathrm{d}z \\ &\quad = \int_Q \langle Df(Du), D\varphi\rangle\,\mathrm{d}z + \int_{t_1}^{t_2} \langle u_t, \varphi\rangle_{W^{1,q}(\mathcal{O},\mathbb{R}^N)}\,\mathrm{d}t. \end{aligned} \tag{7.8}$$

Since $u_t \in L^{q'}(t_1, t_2; W^{-1,q'}(\mathcal{O}, \mathbb{R}^N))$, the last integral on the right-hand side is finite. In this identity we now formally choose the testing function $\varphi = \chi_\theta \psi(u_\varepsilon - u)$, where $\psi \in C_0^\infty(\Omega_T, [0,1])$ has support $\operatorname{spt}\psi \subset Q$ and χ_θ is defined according to (6.6) with some $\tau \in (0, T)$ and $\theta \in (0, \tau)$. Note that this choice of testing function can be made rigorous by an approximation argument. For the first term on the left-hand side of (7.8) we obtain in the limit $\theta \downarrow 0$ that there holds:

$$\begin{aligned} \int_Q (u_\varepsilon - u)\cdot\varphi_t\,\mathrm{d}z &= \tfrac12 \int_Q |u_\varepsilon - u|^2 \partial_t(\chi_\theta\psi)\,\mathrm{d}z \\ &= -\tfrac{1}{2\theta}\int_{\tau-\theta}^{\tau}\int_\Omega |u_\varepsilon - u|^2\psi\,\mathrm{d}z + \frac12\int_Q |u_\varepsilon - u|^2\chi_\theta\partial_t\psi\,\mathrm{d}z \\ &\to -\tfrac12\int_{\mathcal{O}} |(u_\varepsilon - u)(\cdot,\tau)|^2\psi\,\mathrm{d}x + \tfrac12\int_{Q_\tau} |u_\varepsilon - u|^2\partial_t\psi\,\mathrm{d}z. \end{aligned}$$

Here, we have abbreviated $Q_\tau := \mathcal{O} \times (t_1, \tau)$. Therefore, passing to the limit $\theta \downarrow 0$ in (7.8) we get

$$\begin{aligned} &\tfrac12\int_{\mathcal{O}} |(u_\varepsilon - u)(\cdot,\tau)|^2\psi\,\mathrm{d}x + \int_{Q_\tau}\langle Df_\varepsilon(Du_\varepsilon) - Df(Du), D\varphi_\varepsilon\rangle\,\mathrm{d}z \\ &\quad = \tfrac12\int_{Q_\tau} |u_\varepsilon - u|^2\partial_t\psi\,\mathrm{d}z - \int_{Q_\tau}\langle Df(Du), D\varphi_\varepsilon\rangle\,\mathrm{d}z - \int_{t_1}^{\tau}\langle u_t, \varphi_\varepsilon\rangle_{W^{1,q}(\mathcal{O},\mathbb{R}^N)}\,\mathrm{d}t \\ &\quad =: \mathrm{I}_\varepsilon + \mathrm{II}_\varepsilon + \mathrm{III}_\varepsilon, \end{aligned} \tag{7.9}$$

where we have set, for short, $\varphi_\varepsilon := \psi(u_\varepsilon - u)$. Next, we decompose the second term on the left-hand side of (7.9) as follows:

$$\begin{aligned} &\int_{Q_\tau}\langle Df_\varepsilon(Du_\varepsilon) - Df(Du), D\varphi_\varepsilon\rangle\,\mathrm{d}z \\ &\quad = \int_{Q_\tau}\langle Df(Du_\varepsilon) - Df(Du), Du_\varepsilon - Du\rangle\psi\,\mathrm{d}z && (=: \mathrm{IV}_\varepsilon^{(1)}) \\ &\qquad + \int_{Q_\tau}\langle Df(Du_\varepsilon) - Df(Du), D\psi\otimes(u_\varepsilon - u)\rangle\,\mathrm{d}z && (=: \mathrm{IV}_\varepsilon^{(2)}) \\ &\qquad + \varepsilon\, q\int_{Q_\tau} |Du_\varepsilon|^{q-2}Du_\varepsilon\cdot D\varphi_\varepsilon\,\mathrm{d}z && (=: \mathrm{IV}_\varepsilon^{(3)}). \end{aligned}$$

For the first term we use (7.1)$_3$ and Lemma 3.3 to derive the following lower bound:

$$\begin{aligned}\mathrm{IV}_\varepsilon^{(1)} &= \int_{Q_\tau}\int_0^1 \big\langle D^2 f\big(Du + s(Du_\varepsilon - Du)\big)(Du_\varepsilon - Du), Du_\varepsilon - Du\big\rangle \psi \,\mathrm{d}s\,\mathrm{d}z \\ &\geqq \nu \int_{Q_\tau}\int_0^1 |Du + s(Du_\varepsilon - Du)|^{p-2}|Du_\varepsilon - Du|^2\psi\,\mathrm{d}s\,\mathrm{d}z \\ &\geqq \tfrac{\nu}{c(p)}\int_{Q_\tau}\big(|Du|^2 + |Du_\varepsilon - Du|^2\big)^{\frac{p-2}{2}}|Du_\varepsilon - Du|^2\psi\,\mathrm{d}z.\end{aligned}$$

Combining this inequality with (7.9) we get

$$\begin{aligned}&\tfrac12\int_{\mathcal{O}}|(u_\varepsilon - u)(\cdot,\tau)|^2\psi\,\mathrm{d}x + \tfrac{\nu}{c(p)}\int_{Q_\tau}\big(|Du|^2 + |Du_\varepsilon - Du|^2\big)^{\frac{p-2}{2}} \\ &\quad |Du_\varepsilon - Du|^2\psi\,\mathrm{d}z \leqq |\mathrm{I}_\varepsilon| + |\mathrm{II}_\varepsilon| + |\mathrm{III}_\varepsilon| + |\mathrm{IV}_\varepsilon^{(2)}| + |\mathrm{IV}_\varepsilon^{(3)}|.\end{aligned} \tag{7.10}$$

In the following we will show that the terms on the right-hand side of (7.10) converge to zero in the limit $\varepsilon \downarrow 0$. For the first term this is a consequence of (7.5), that is, we have $|\mathrm{I}_\varepsilon| \to 0$ as $\varepsilon \downarrow 0$. For the term $\mathrm{IV}_\varepsilon^{(2)}$ we obtain, with the help of $|Df(\xi)| \leqq c(q)L(1+|\xi|^{q-1})$, the uniform L^q-bound for Du_ε on Q and the strong convergence $u_\varepsilon \to u$ in $L^q(Q,\mathbb{R}^N)$ from (7.5), that

$$\begin{aligned}&|\mathrm{IV}_\varepsilon^{(2)}| \leqq c(q)\,L\int_Q \big[(|Du_\varepsilon|+1)^{q-1} + (|Du|+1)^{q-1}\big]|D\psi||u_\varepsilon - u|\,\mathrm{d}z \\ &\leqq c(q)\,L\|D\psi\|_{L^\infty}\bigg(\int_Q\big(|Du_\varepsilon|^q + |Du|^q + 1\big)\,\mathrm{d}z\bigg)^{\frac{q-1}{q}}\bigg(\int_Q|u_\varepsilon - u|^q\,\mathrm{d}z\bigg)^{\frac1q} \\ &\to 0 \qquad \text{as } \varepsilon\downarrow 0.\end{aligned}$$

By the uniform boundedness of u_ε and Du_ε in $L^q(Q)$, we get for the term $\mathrm{IV}_\varepsilon^{(3)}$ that

$$\begin{aligned}|\mathrm{IV}_\varepsilon^{(3)}| &\leqq \varepsilon\, q\int_Q |Du_\varepsilon|^{q-1}\big(|Du_\varepsilon - Du|\psi + |u_\varepsilon - u||D\psi|\big)\,\mathrm{d}z \\ &\leqq c(q)\varepsilon\,\big(\|\psi\|_{L^\infty} + \|D\psi\|_{L^\infty}\big)\int_{Q_{t_o}}\big(|Du_\varepsilon|^q + |Du|^q + |u_\varepsilon - u|^q\big)\,\mathrm{d}z \\ &\to 0 \qquad \text{as } \varepsilon\downarrow 0.\end{aligned}$$

Finally, from (7.4) we obtain that $\varphi_\varepsilon \rightharpoonup 0$ weakly in $L^q(t_1,t_2;W^{1,q}(\mathcal{O},\mathbb{R}^N))$. Since $Df(Du) \in L^{q'}(Q,\mathbb{R}^{Nn})$ and $u_t \in L^{q'}(t_1,t_2;W^{-1,q'}(\mathcal{O},\mathbb{R}^N))$, this also implies that $|\mathrm{II}_\varepsilon|$ and $|\mathrm{III}_\varepsilon|$ converge to zero in the limit $\varepsilon\downarrow 0$. Altogether, we have shown that

$$\begin{aligned}\lim_{\varepsilon\downarrow 0}\bigg[&\tfrac12\int_{\mathcal{O}}|(u_\varepsilon - u)(\cdot,\tau)|^2\psi\,\mathrm{d}x \\ &+\int_Q\big(|Du|^2 + |Du_\varepsilon - Du|^2\big)^{\frac{p-2}{2}}|Du_\varepsilon - Du|^2\psi\,\mathrm{d}z\bigg] = 0.\end{aligned} \tag{7.11}$$

Since spt $\psi \subset Q \Subset \Omega_T$ and $\tau \in (t_1, t_2)$ are arbitrary, this already implies claims (7.6) and (7.7).

Step 4. Passage to the limit. It only remains to establish that u is a weak solution to (2.7). Starting from the weak form of $(7.2)_2$, that is, from

$$\int_{\Omega_T} u_\varepsilon \cdot \varphi_t - \langle Df_\varepsilon(Du_\varepsilon), D\varphi\rangle \, \mathrm{d}z = 0 \qquad \forall \varphi \in C_0^\infty(\Omega_T, \mathbb{R}^N),$$

we use (7.5) and (7.6) to pass to the limit $\varepsilon \downarrow 0$. We conclude that

$$\int_{\Omega_T} u \cdot \varphi_t - \langle Df(Du), D\varphi\rangle \, \mathrm{d}z = 0 \qquad \forall \varphi \in C_0^\infty(\Omega_T, \mathbb{R}^N)$$

holds true. Moreover, for any cylinder $Q_R(z_o) \Subset \Omega_T$ we have $Du_\varepsilon \to Du$ almost everywhere on $Q_R(z_o)$, which, together with (7.5), (7.3), (7.6) and (7.7), shows that

$$\begin{aligned}
&\int_{Q_{R/2}(z_o)} |Du|^q \, \mathrm{d}z \\
&= \lim_{\varepsilon \downarrow 0} \int_{Q_{R/2}(z_o)} |Du_\varepsilon|^q \, \mathrm{d}z \\
&\leqq c \Bigg[\sup_{t \in (t_o - R^2, t_o)} \int_{B_R(x_o)} |u_\varepsilon(\cdot, t)|^2 \, \mathrm{d}x + \int_{Q_R(z_o)} \big(|Du_\varepsilon|^p + |u_\varepsilon|^p + 1\big) \, \mathrm{d}z \Bigg]^\chi \\
&= c \Bigg[\sup_{t \in (t_o - R^2, t_o)} \int_{B_R(x_o)} |u(\cdot, t)|^2 \, \mathrm{d}x + \int_{Q_R(z_o)} \big(|Du|^p + |u|^p + 1\big) \, \mathrm{d}z \Bigg]^\chi .
\end{aligned}$$

This completes the proof of Theorem 2.6. □

Finally, we prove the regularity result for variational solutions from Theorem 2.8.

Proof of Theorem 2.8. Let u be a variational solution to (2.1) under the assumptions (2.6), (2.8) and (2.9). Firstly, we observe that the assumptions of Theorem 2.6 are satisfied and, therefore, there exists a weak solution

$$\tilde{u} \in L^p\big(0, T; W_g^{1,p}(\Omega, \mathbb{R}^N)\big) \cap L^q_{\mathrm{loc}}\big(0, T; W^{1,q}_{\mathrm{loc}}(\Omega, \mathbb{R}^N)\big) \cap C_w\big([0, T]; L^2(\Omega, \mathbb{R}^N)\big)$$

of the parabolic Cauchy–Dirichlet problem (2.1) satisfying (2.10). By the argument from Section 4.4, we infer that $\tilde{u}$ is also a variational solution to (2.1). Therefore, the uniqueness result from Theorem 2.4 ensures that $\tilde{u} = u$, and hence the variational solution u is indeed a weak solution satisfying the desired properties. □

Appendix A: A Compactness Result

For the sake of completeness we state and prove in Theorem A.2, below, a probably not so well known compactness result from [17] which we used in the last step of the proof of Theorem 2.3. The argument is based on the following result from [34, Theorem 2.1].

Theorem A.1. *Any function* $u \in L^\infty(0,T;L^2(\Omega,\mathbb{R}^N)) \cap C_w([0,T];W^{-\ell,2}(\Omega,\mathbb{R}^N))$*, with* $\ell \in \mathbb{N}$*, is weakly continuous as a mapping* $[0,T] \ni t \mapsto u(\cdot,t) \in L^2(\Omega,\mathbb{R}^N)$*; this means that* $u \in C_w([0,T];L^2(\Omega,\mathbb{R}^N))$*. Moreover, there holds*

$$\|u(\cdot,t)\|_{L^2(\Omega)} \leqq \|u\|_{L^\infty(0,T;L^2(\Omega))} \quad \text{for any } t \in [0,T].$$

The previous theorem can be used to establish the following compactness result.

Theorem A.2. *Let* $\ell \in \mathbb{N}$*. Suppose that*

$$(u_i)_{i\in\mathbb{N}} \subset L^\infty\big(0,T;L^2(\Omega,\mathbb{R}^N)\big) \cap C^0\big([0,T];W^{-\ell,2}(\Omega,\mathbb{R}^N)\big)$$

is a sequence of maps satisfying

(i) $(u_i)_{i\in\mathbb{N}}$ *is bounded in* $L^\infty(0,T;L^2(\Omega,\mathbb{R}^N))$*, and moreover*
(ii) $\|u_i(\cdot,t_2)-u_i(\cdot,t_1)\|_{W^{-\ell,2}(\Omega)} \leqq \omega(|t_2-t_1|)$ *for any* $i \in \mathbb{N}$ *and* $t_1,t_2 \in [0,T]$ *for some non-decreasing modulus* $\omega\colon [0,T] \to \mathbb{R}_+$ *with* $\lim_{s\downarrow 0}\omega(s) = 0 = \omega(0)$.

Then, there exists a function $u \in C_w([0,T];L^2(\Omega,\mathbb{R}^N))$ *and a (non-relabelled) subsequence* $(u_i)_{i\in\mathbb{N}}$ *such that* $u_i(\cdot,t) \rightharpoonup u(\cdot,t)$ *weakly in* $L^2(\Omega,\mathbb{R}^N)$ *in the limit* $i \to \infty$ *for any* $t \in [0,T]$.

Proof. Due to hypotheses (i) and (ii) we can apply the compactness result [32, Theorem 5] to the sequence $(u_i)_{i\in\mathbb{N}}$ with $p=\infty$, $B = W^{-\ell,2}(\Omega,\mathbb{R}^N)$ and $X = L^2(\Omega,\mathbb{R}^N)$ to infer the existence of a function $u \in C^0([0,T];W^{-\ell,2}(\Omega,\mathbb{R}^N))$ and a subsequence—still denoted by $(u_i)_{i\in\mathbb{N}}$—such that $u_i \to u$ strongly in $C^0([0,T];W^{-\ell,2}(\Omega,\mathbb{R}^N))$. Further, by (i) we are allowed to pass to another subsequence—still denoted by u_i—so that $u_i \rightharpoonup^* u$ weakly* in $L^\infty(0,T;L^2(\Omega,\mathbb{R}^N))$. At this stage we apply Theorem A.1 and conclude that $u, u_i \in C_w([0,T];L^2(\Omega,\mathbb{R}^N))$ for any $i \in \mathbb{N}$. Moreover, for any $t \in [0,T]$ we have

$$\sup_{i\in\mathbb{N}} \|u_i(\cdot,t)\|_{L^2(\Omega)} \leqq \sup_{i\in\mathbb{N}} \|u_i\|_{L^\infty(0,T;L^2(\Omega))} =: M < \infty \tag{A.1}$$

and

$$\|u(\cdot,t)\|_{L^2(\Omega)} \leqq \|u\|_{L^\infty(0,T;L^2(\Omega))} \leqq M, \tag{A.2}$$

where we used the lower semicontinuity of $\|\cdot\|_{L^\infty(0,T;L^2(\Omega))}$ with respect to weak* convergence.

Next, we consider $t \in [0,T]$, $\psi \in L^2(\Omega,\mathbb{R}^N)$ and $\delta > 0$. Since $W^{\ell,2}(\Omega,\mathbb{R}^N)$ is dense in $L^2(\Omega,\mathbb{R}^N)$, there exists $\psi_\delta \in W^{\ell,2}(\Omega,\mathbb{R}^N)$ such that $\|\psi-\psi_\delta\|_{L^2(\Omega)} \leqq \delta$. Using (A.1) and (A.2), we find that

$$\begin{aligned}
&\left|\int_\Omega \big(u_i(\cdot,t)-u(\cdot,t)\big)\psi \,\mathrm{d}x\right| \\
&\leqq \left|\int_\Omega \big(u_i(\cdot,t)-u(\cdot,t)\big)\big(\psi-\psi_\delta\big)\,\mathrm{d}x\right| + \left|\int_\Omega \big(u_i(\cdot,t)-u(\cdot,t)\big)\psi_\delta\,\mathrm{d}x\right| \\
&\leqq \delta\big(\|u_i(\cdot,t)\|_{L^2(\Omega)} + \|u(\cdot,t)\|_{L^2(\Omega)}\big) + \|u_i(\cdot,t)-u(\cdot,t)\|_{W^{-\ell,2}(\Omega)}\|\psi_\delta\|_{W^{\ell,2}(\Omega)} \\
&\leqq 2\delta M + \|u_i(\cdot,t)-u(\cdot,t)\|_{W^{-\ell,2}(\Omega)}\|\psi_\delta\|_{W^{\ell,2}(\Omega)}.
\end{aligned}$$

In order to pass to the limit $i \to \infty$ on the right-hand side we recall that u_i converges to u strongly in $u \in C^0([0, T]; W^{-\ell,2}(\Omega, \mathbb{R}^N))$. This implies that the second term on the right-hand side of the preceding inequality vanishes in the limit $i \to \infty$. But this implies

$$\limsup_{i\to\infty} \left| \int_\Omega \big(u_i(\cdot, t) - u(\cdot, t)\big)\psi \,\mathrm{d}x \right| \leqq 2\delta M.$$

Since $\delta > 0$ was arbitrary, we can pass in this inequality to the limit $\delta \downarrow 0$, yielding that

$$\lim_{i\to\infty} \int_\Omega \big(u_i(\cdot, t) - u(\cdot, t)\big)\psi \,\mathrm{d}x = 0$$

holds whenever $\psi \in L^2(\Omega, \mathbb{R}^N)$, that is, $u_i(\cdot, t) \rightharpoonup u(\cdot, t)$ weakly in $L^2(\Omega, \mathbb{R}^N)$. Since $t \in [0, T]$ is arbitrary, we have proved the desired weak convergence for any $t \in [0, T]$. This finishes the proof of the theorem. □

Appendix B: Mollification in Time

Here, we will provide the basic properties of the mollification in time $[\cdot]_h$ defined in (5.1). The following lemma, which considers the mollification in a more general setting, will be useful later.

Lemma B.1. *Let X be a Banach space and assume that $v_o \in X$ and, moreover, $v \in L^r(0, T; X)$ for some $1 \leqq r \leqq \infty$. Then, the mollification in time defined by*

$$[v]_h(t) := \mathrm{e}^{-\frac{t}{h}} v_o + \tfrac{1}{h} \int_0^t e^{\frac{s-t}{h}} v(s)\, ds, \tag{B.1}$$

for $h \in (0, T]$ and $t \in [0, T]$ belongs to $L^r(0, T; X)$ and

$$\|[v]_h\|_{L^r(0,t_o;X)} \leqq \|v\|_{L^r(0,t_o;X)} + \Big[\tfrac{h}{r}\big(1 - \mathrm{e}^{-\frac{t_o r}{h}}\big)\Big]^{\frac{1}{r}} \|v_o\|_X \tag{B.2}$$

for any $t_o \in (0, T]$. Moreover, we have $\partial_t[v]_h \in L^r(0, T; X)$ and

$$\partial_t[v]_h = -\tfrac{1}{h}\big([v]_h - v\big).$$

Proof. For $t \in (0, T)$ we decompose

$$\|[v]_h(t)\|_X \leqq \mathrm{e}^{-\frac{t}{h}} \|v_o\|_X + \tfrac{1}{h} \int_0^t \mathrm{e}^{\frac{s-t}{h}} \|v(s)\|_X \,\mathrm{d}s =: \mathrm{I}_h(\cdot, t) + \mathrm{II}_h(\cdot, t),$$

with the obvious meaning of I_h and II_h. For $t \in (0, T)$ and $r > 1$, Hölder's inequality implies for the term II_h that there holds:

$$
\begin{aligned}
\mathrm{II}_h(t) &= \int_0^t \Big(\tfrac{1}{h}\mathrm{e}^{\frac{s-t}{h}}\Big)^{\frac{1}{r'}} \Big(\tfrac{1}{h}\mathrm{e}^{\frac{s-t}{h}}\Big)^{\frac{1}{r}} \|v(s)\|_X \,\mathrm{d}s \\
&\leqq \bigg(\int_0^t \tfrac{1}{h}\mathrm{e}^{\frac{s-t}{h}}\,\mathrm{d}s\bigg)^{\frac{1}{r'}} \bigg(\int_0^t \tfrac{1}{h}\mathrm{e}^{\frac{s-t}{h}} \|v(s)\|_X^r \,\mathrm{d}s\bigg)^{\frac{1}{r}} \\
&= \big(1-\mathrm{e}^{-\frac{t}{h}}\big)^{\frac{1}{r'}} \bigg(\int_0^t \tfrac{1}{h}\mathrm{e}^{\frac{s-t}{h}} \|v(s)\|_X^r \,\mathrm{d}s\bigg)^{\frac{1}{r}} \\
&\leqq \bigg(\int_0^t \tfrac{1}{h}\mathrm{e}^{\frac{s-t}{h}} \|v(s)\|_X^r \,\mathrm{d}s\bigg)^{\frac{1}{r}}.
\end{aligned}
$$

For the case $r = 1$ the last inequality trivially holds. We take this to the power r and integrate over $(0, t_o)$, where $t_o \in (0, T]$. By Fubini's theorem we obtain

$$
\begin{aligned}
\int_0^{t_o} \|\mathrm{II}_h(t)\|_X^r \,\mathrm{d}t &\leqq \int_0^{t_o}\int_0^t \tfrac{1}{h}\mathrm{e}^{\frac{s-t}{h}} \|v(s)\|_X^r \,\mathrm{d}s\,\mathrm{d}t \\
&= \int_0^{t_o}\int_s^{t_o} \tfrac{1}{h}\mathrm{e}^{\frac{s-t}{h}}\,\mathrm{d}t\, \|v(s)\|_X^r \,\mathrm{d}s \\
&= \int_0^{t_o} \Big(1-\mathrm{e}^{\frac{s-T}{h}}\Big) \|v(s)\|_X^r \,\mathrm{d}s \leqq \int_0^{t_o} \|v(s)\|_X^r \,\mathrm{d}s.
\end{aligned}
$$

The first term is treated as follows:

$$
\int_0^{t_o} \|\mathrm{I}_h(t)\|_X^r \,\mathrm{d}z \leqq \|v_o\|_X^r \int_0^{t_o} \mathrm{e}^{-\frac{tr}{h}}\,\mathrm{d}t = \tfrac{h}{r}\Big(1-\mathrm{e}^{-\frac{t_o r}{h}}\Big)\|v_o\|_X^r.
$$

Combining the previous estimates, we have shown

$$
\|[v]_h\|_{L^r(0,t_o;X)} \leqq \|v\|_{L^r(0,t_o;X)} + \Big[\tfrac{h}{r}\big(1-\mathrm{e}^{-\frac{t_o r}{h}}\big)\Big]^{\frac{1}{r}} \|v_o\|_X,
$$

whenever $t_o \in (0, T]$. This proves $[v]_h \in L^r(0, T; X)$ together with the L^r-estimate. In order to prove the assertion concerning the time derivative of $[v]_h$ we perform a direct computation of $\partial_t[v]_h$, which yields for almost every $t \in (0, T)$ that

$$
\begin{aligned}
\partial_t[v]_h(t) &= \frac{\mathrm{d}}{\mathrm{dt}}\bigg[\mathrm{e}^{-\frac{t}{h}}v(0) + \frac{\mathrm{e}^{-\frac{t}{h}}}{h}\int_0^t \mathrm{e}^{\frac{s}{h}}v(s)\,\mathrm{d}s\bigg] \\
&= -\frac{\mathrm{e}^{-\frac{t}{h}}}{h}v(0) - \frac{\mathrm{e}^{-\frac{t}{h}}}{h^2}\int_0^t \mathrm{e}^{\frac{s}{h}}v(s)\,\mathrm{d}s + \tfrac{1}{h}v(t) \\
&= -\tfrac{1}{h}\big([v]_h(t) - v(t)\big)
\end{aligned}
$$

holds true. We note that for this computation it suffices to have $v \in L^1(0, T; X)$ and $v(0) \in X$. Since we have already shown that the right-hand side belongs to $L^r(0, T; X)$, we conclude that $\partial_t[v]_h \in L^r(0, T; X)$. This finishes the proof of the lemma. □

The following lemma provides the basic properties of the mollification $[\cdot]_h$ in Sobolev spaces.

Lemma B.2. *Suppose that $v_o \in L^1(\Omega, \mathbb{R}^N)$ and, moreover, $v \in L^1(0, T; L^1(\Omega, \mathbb{R}^N)) = L^1(\Omega_T, \mathbb{R}^N)$. Then, the mollification $[v]_h$ defined in (5.1) admits the following properties:*

(i) *If $v \in L^p(\Omega_T, \mathbb{R}^N)$ and $v_o \in L^p(\Omega, \mathbb{R}^N)$ with $p \geqq 1$, then also $[v]_h \in L^p(\Omega_T, \mathbb{R}^N)$, and the following quantitative estimate holds true:*

$$\|[v]_h\|_{L^p(\Omega_T)} \leqq \|v\|_{L^p(\Omega_T)} + h^{\frac{1}{p}} \|v_o\|_{L^p(\Omega)}.$$

Moreover, $[v]_h \to v$ in $L^p(\Omega_T, \mathbb{R}^N)$ as $h \downarrow 0$.

(ii) *If $v \in L^p(0, T; W^{1,p}(\Omega, \mathbb{R}^N))$ and $v_o \in W^{1,p}(\Omega, \mathbb{R}^N)$ with $p \geqq 1$, then also $[v]_h \in L^p(0, T; W^{1,p}(\Omega, \mathbb{R}^N))$, and the following quantitative estimate holds true:*

$$\|[v]_h\|_{L^p(0,T;W^{1,p}(\Omega))} \leqq \|v\|_{L^p(0,T;W^{1,p}(\Omega))} + h^{\frac{1}{p}} \|v_o\|_{W^{1,p}(\Omega)}.$$

Moreover, $[v]_h \to v$ in $L^p(0, T; W^{1,p}(\Omega, \mathbb{R}^N))$ as $h \downarrow 0$.

(iii) *If $v \in L^\infty(0, T; L^2(\Omega, \mathbb{R}^N))$ and $v_o \in L^2(\Omega, \mathbb{R}^N)$, then $[v]_h \in C^0([0, T]; L^2(\Omega, \mathbb{R}^N))$ and $[v]_h \to v$ in $L^2(\Omega_T, \mathbb{R}^N)$ as $h \downarrow 0$; in particular, there holds $[v]_h(\cdot, 0) = v_o$.*

(iv) *If $v \in L^\infty(0, T; L^2(\Omega, \mathbb{R}^N))$ and $v_o \in L^2(\Omega, \mathbb{R}^N)$, then also $\partial_t[v]_h \in L^\infty(0, T; L^2(\Omega, \mathbb{R}^N))$. Moreover, we have*

$$\partial_t[v]_h = -\tfrac{1}{h}\big([v]_h - v\big).$$

(v) *If $v \in L^p(0, T; W_0^{1,p}(\Omega, \mathbb{R}^N))$ and $v_o \in W_0^{1,p}(\Omega, \mathbb{R}^N)$, then $[v]_h \in L^p(0, T; W_0^{1,p}(\Omega, \mathbb{R}^N))$.*

Proof. We start with the **proof of (i)**. The assertion $[v]_h \in L^p(\Omega_T, \mathbb{R}^N)$ and the L^p-bound for $[v]_h$ directly follow by an application of Lemma B.1 with the choice $r = p$ and $X = L^p(\Omega, \mathbb{R}^N)$.

It remains to establish $[v]_h \to v$ in $L^p(\Omega_T, \mathbb{R}^N)$ as $h \downarrow 0$. For this we approximate $v \in L^p(\Omega_T, \mathbb{R}^N)$ by continuous functions $\tilde{v}$ with compact support $\operatorname{spt} \tilde{v} \Subset \Omega_T$; more precisely, for given $\varepsilon > 0$ we find $\tilde{v} \in C^0(\Omega_T, \mathbb{R}^N)$ with $\operatorname{spt} \tilde{v} \Subset \Omega_T$ and $\|v - \tilde{v}\|_{L^p(\Omega_T)} \leqq \varepsilon$. We construct $[\tilde{v}]_h$ according to (5.1) with v replaced by $\tilde{v}$ and observe that the difference

$$[v]_h - [\tilde{v}]_h = \frac{1}{h} \int_0^t \mathrm{e}^{\frac{s-t}{h}} \big(v(\cdot, s) - \tilde{v}(\cdot, s)\big)\, \mathrm{d}s$$

equals $[v - \tilde{v}]_h$ as defined in (B.1) with initial datum $v_o \equiv 0$. Therefore, using estimate (B.2) from Lemma B.1 with $v - \tilde{v}$ instead of v, $v_o = 0$, $X = L^p(\Omega, \mathbb{R}^N)$ and $t_o = T$, we get

$$\|[v]_h - [\tilde{v}]_h\|_{L^p(\Omega_T)} \leqq \|v - \tilde{v}\|_{L^p(\Omega_T)} \leqq \varepsilon,$$

and therefore

$$\begin{aligned}\|[v]_h - v\|_{L^p(\Omega_T)} &\leqq \|[v]_h - [\tilde{v}]_h\|_{L^p(\Omega_T)} + \|[\tilde{v}]_h - \tilde{v}\|_{L^p(\Omega_T)} + \|\tilde{v} - v\|_{L^p(\Omega_T)} \\ &\leqq 2\,\varepsilon + \|[\tilde{v}]_h - \tilde{v}\|_{L^p(\Omega_T)}.\end{aligned}$$

At this stage it remains to show that $[\tilde{v}]_h \to \tilde{v}$ in $L^p(\Omega_T, \mathbb{R}^N)$ as $h \downarrow 0$ for $\tilde{v} \in C^0(\Omega_T)$ with $\operatorname{spt} \tilde{v} \Subset \Omega_T$; this can be inferred as follows: Taking into account that

$$\frac{1}{h(1-e^{-\frac{t}{h}})} \int_0^t e^{\frac{s-t}{h}}\, ds = 1,$$

we first rewrite $\tilde{v}(\cdot, t) - [\tilde{v}]_h(\cdot, t)$ and obtain for $t \in (0, T)$ that

$$[\tilde{v}]_h(\cdot, t) - \tilde{v}(\cdot, t) = e^{-\frac{t}{h}}(v_o - \tilde{v}(\cdot,\ t)) + \frac{1}{h} \int_0^t e^{\frac{s-t}{h}} \big(\tilde{v}(\cdot, s) - \tilde{v}(\cdot, t)\big)\, ds. \tag{B.3}$$

Since $\operatorname{spt} \tilde{v} \Subset \Omega_T$, we find $0 < \delta_o < T$ such that $\tilde{v}(\cdot,\ t) = 0$ whenever $0 \leqq t \leqq \delta_o$. We now estimate the $L^p(\Omega_T)$-norm of both terms appearing on the right-hand side of the last identity separately. For the first one we observe that for $0 < t \leqq \delta_o$ it simplifies to $e^{-\frac{t}{h}} v_o$. This yields

$$\begin{aligned}
&\int_0^T \int_\Omega |e^{-\frac{t}{h}}(v_o - \tilde{v}(\cdot,\ t))|^p\, dx\, dt \\
&= \int_0^{\delta_0} \int_\Omega |e^{-\frac{t}{h}} v_o|^p\, dx\, dt + \int_{\delta_0}^T \int_\Omega |e^{-\frac{t}{h}}(v_o - \tilde{v}(\cdot,\ t))|^p\, dx\, dt \\
&\leqq h \int_\Omega |v_o|^p\, dx + e^{-\frac{p\delta_o}{h}} \int_{\delta_0}^T \int_\Omega |v_o - \tilde{v}(\cdot,\ t)|^p\, dx\, dt \\
&\leqq h \int_\Omega |v_o|^p\, dx + e^{-\frac{p\delta_o}{h}} \big[T \|v_o\|_{L^p(\Omega)} + \|\tilde{v}\|_{L^p(\Omega_T)}\big]^p.
\end{aligned}$$

Next, we observe that the second term on the right-hand side of (B.3) vanishes if $0 < t \leqq \delta_o$. Moreover, since $\tilde{v}$ is uniformly continuous we know that for any $\varepsilon > 0$ there exists $\delta(\varepsilon) \in (0, \delta_o]$ such that $|\tilde{v}(x, t) - \tilde{v}(x, s)| \leqq \varepsilon$ whenever $x \in \Omega$ and $|t - s| \leqq \delta$. Using this in the estimate of the second term, we find

$$\begin{aligned}
&\int_0^T \int_\Omega \left| \frac{1}{h} \int_0^t e^{\frac{s-t}{h}} \big(\tilde{v}(\cdot, s) - \tilde{v}(\cdot, t)\big)\, ds \right|^p dx\, dt \\
&= \int_{\delta_o}^T \int_\Omega \left| \frac{1}{h} \int_0^t e^{\frac{s-t}{h}} \big(\tilde{v}(\cdot, s) - \tilde{v}(\cdot, t)\big)\, ds \right|^p dx\, dt \\
&= \int_{\delta_o}^T \int_\Omega \left| \frac{1}{h} \int_0^{t-\delta} e^{\frac{s-t}{h}} \big(\tilde{v}(\cdot, t) - \tilde{v}(\cdot, s)\big)\, ds \right. \\
&\qquad \left. + \frac{1}{h} \int_{t-\delta}^t e^{\frac{s-t}{h}} \big(\tilde{v}(\cdot, t) - \tilde{v}(\cdot, s)\big)\, ds \right|^p dx\, dt \\
&\leqq T|\Omega| \big[2e^{-\frac{\delta}{h}} \|\tilde{v}\|_{L^\infty(\Omega_T)} + \varepsilon\big]^p.
\end{aligned}$$

Combing the previous estimates, we infer that

$$\begin{aligned}
&\int_0^T \int_\Omega \big|[\tilde{v}]_h(x, t) - \tilde{v}(x, t)\big|^p\, dx\, dt \\
&\leqq h\|v_o\|^p_{L^p(\Omega)} + e^{-\frac{p\delta_o}{h}} \big[T \|v_o\|_{L^p(\Omega)} + \|\tilde{v}\|_{L^p(\Omega_T)}\big]^p \\
&\quad + T|\Omega| \big[2e^{-\frac{\delta}{h}} \|\tilde{v}\|_{L^\infty(\Omega_T)} + \varepsilon\big]^p.
\end{aligned}$$

Here, we first let $\varepsilon \downarrow 0$ (which is possible, since $\varepsilon > 0$ was arbitrary) and then $h \downarrow 0$. The right-hand side then converges to 0, proving that $[\tilde{v}]_h \to \tilde{v}$ in $L^p(\Omega_T, \mathbb{R}^N)$ as $h \downarrow 0$. Altogether, we have shown the desired L^p-convergence and this finishes the proof of (i).

The **proof of (ii)** is an easy consequence of (i), since

$$D[v]_h(\cdot, t) = \mathrm{e}^{-\frac{t}{h}} Dv_o(\cdot) + \tfrac{1}{h} \int_0^t \mathrm{e}^{\frac{s-t}{h}} Dv(\cdot, s)\, \mathrm{d}s.$$

Now, by (i), the assumptions $Dv \in L^p(\Omega_T)$ and $Dv_o \in L^p(\Omega)$ guarantee that $[Dv]_h \in L^p(\Omega_T)$ and $[Dv]_h \to Dv$ in $L^p(\Omega_T)$. Therefore, we have $[v]_h \in L^p(0, T; W^{1,p}(\Omega))$ and $[v]_h \to v$ in $L^p(0, T; W^{1,p}(\Omega))$. The bound for the $L^p{-}W^{1,p}$ norm of $[v]_h$ follows immediately from the estimate in (i) applied to Dv.

Now we come to the **proof of (iii)**. The assertion that $[v]_h \to v$ in $L^2(\Omega_T)$ is a consequence of (i), since $L^\infty(0, T; L^2(\Omega)) \subset L^2(\Omega_T)$. Therefore, we only have to establish the continuity of $[v]_h$ with respect to time. For $0 \leqq t_1 < t_2 \leqq T$ we rewrite the difference of $[v]_h(\cdot, t)$ at $t = t_2$ and $t = t_1$ as follows:

$$\begin{aligned}
&[v]_h(\cdot, t_2) - [v]_h(\cdot, t_1) \\
&\quad = \Big(\mathrm{e}^{-\frac{t_2}{h}} - \mathrm{e}^{-\frac{t_1}{h}}\Big)\Big[v_o + \tfrac{1}{h} \int_0^{t_1} \mathrm{e}^{\frac{s}{h}} v(\cdot, s)\, \mathrm{d}s\Big] + \mathrm{e}^{-\frac{t_2}{h}} \tfrac{1}{h} \int_{t_1}^{t_2} \mathrm{e}^{\frac{s}{h}} v(\cdot, s)\, \mathrm{d}s.
\end{aligned}$$

For the estimate of the right-hand side terms we use the following elementary inequality with suitable choices of τ_1 and τ_2. This inequality follows by an application of the Cauchy–Schwarz inequality:

$$\bigg|\int_{\tau_1}^{\tau_2} \mathrm{e}^{\frac{s}{h}} v(\cdot, s)\, \mathrm{d}s\bigg|^2 \leqq \int_{\tau_1}^{\tau_2} \mathrm{e}^{\frac{2s}{h}}\, \mathrm{d}s \int_{\tau_1}^{\tau_2} v(\cdot, s)^2\, \mathrm{d}s = \tfrac{h}{2}\Big(\mathrm{e}^{\frac{2\tau_2}{h}} - \mathrm{e}^{\frac{2\tau_1}{h}}\Big) \int_{\tau_1}^{\tau_2} v^2(\cdot, s)\, \mathrm{d}s.$$

We apply the preceding inequality with the choices $(\tau_1, \tau_2) = (0, t_1)$ (respectively (t_1, t_2)) and obtain

$$\begin{aligned}
&\big\|[v]_h(\cdot, t_2) - [v]_h(\cdot, t_1)\big\|^2_{L^2(\Omega)} \\
&\quad \leqq 3\Big(\mathrm{e}^{-\frac{t_2}{h}} - \mathrm{e}^{-\frac{t_1}{h}}\Big)^2 \|v_o\|^2_{L^2(\Omega)} + \tfrac{3}{2h}\Big(1 - \mathrm{e}^{\frac{2(t_1 - t_2)}{h}}\Big) \int_{t_1}^{t_2} \int_\Omega v^2(x, t)\, \mathrm{d}x\, \mathrm{d}s \\
&\qquad + \tfrac{3}{2h}\Big(1 - \mathrm{e}^{-\frac{2t_1}{h}}\Big)\Big(1 - \mathrm{e}^{\frac{2(t_1 - t_2)}{h}}\Big) \int_0^{t_1} \int_\Omega v^2(x, t)\, \mathrm{d}x\, \mathrm{d}s \\
&\quad \leqq 3\Big(\mathrm{e}^{-\frac{t_2}{h}} - \mathrm{e}^{-\frac{t_1}{h}}\Big)^2 \|u_o\|^2_{L^2(\Omega)} + \tfrac{3}{2h}\Big(1 - \mathrm{e}^{\frac{2(t_1 - t_2)}{h}}\Big) \|v\|^2_{L^2(\Omega_T)}.
\end{aligned}$$

Therefore, the right-hand side converges to 0 whenever $t_2 - t_1 \to 0$ and this yields the first assertion in (iii). By the same reasoning we infer that

$$\int_\Omega |[v]_h(x, t) - v_o|^2\, \mathrm{d}x \ \leqq 2\big(1 - \mathrm{e}^{-\frac{t}{h}}\big)^2 \int_\Omega |v_o|^2\, \mathrm{d}x + 2 \int_\Omega \int_0^t |v(x, s)|^2\, \mathrm{d}s\, \mathrm{d}x,$$

and, as this implies, that $[v]_h(\cdot, t) \to v_o$ as $t \downarrow 0$.

Proof of (iv): Claim (iv) directly follows from Lemma B.1 applied with $r = 2$ and $X = L^2(\Omega, \mathbb{R}^N)$.

Finally, the proof of **proof of (v)** is standard (compare the analogous result for Steklov-averages). □

To treat terms involving the time derivative of v or $[v]_h$, we need to define what is meant by $\partial_t v \in L^{p'}(0, T; W^{-1,p'}(\Omega))$ for a function $v \in L^p(0, T; W^{1,p}(\Omega))$ with $p > \frac{2n}{n+2}$. The reason we restrict our consideration to this values of p comes from the fact that we have the following inclusions

$$W^{1,p}(\Omega) \hookrightarrow L^2(\Omega) \cong (L^2(\Omega))' \hookrightarrow W^{-1,p'}(\Omega),$$

with continuous injections. This allows us to interpret the mapping $t \mapsto v(\cdot, t) \in W^{1,p}(\Omega)$ as map from $(0, T)$ into $W^{-1,p'}(\Omega)$, that is, a curve in $W^{-1,p'}(\Omega)$. If we denote the embedding from $W^{1,p}$ to $W^{-1,p'}$ by $\mathbb{J}$, we easily see that

$$\int_0^T \|\mathbb{J}(v(t))\|^p_{W^{-1,p'}} \, dt \leqq c \int_0^T \|v(t)\|^p_{W^{1,p}} \, dt = c \, \|v\|^p_{L^p(0,T;W^{1,p}(\Omega))}.$$

In particular $v \in L^1(0, T; W^{-1,p'}(\Omega))$, more precisely, the map $t \mapsto \mathbb{J}(v(\cdot, t))$ is in this space. In the following, we use the shorthand notion $v(t)$ instead of $\mathbb{J}(v(t))$.

Now, we let X be a Banach space with norm $\| \cdot \|_X$.

In the space $L^1(0, T; X)$ it is possible to define the weak time derivative $\partial_t v(t)$ of $v(t)$ as follows: $w \in L^1(0, T; X)$ is called a weak time derivative if there holds:

$$\int_0^T w(t)\psi(t) \, dt = -\int_0^T v(t)\psi'(t) \, dt \qquad \forall \, \psi \in C_0^\infty(0, T).$$

In case such a w exists, we write $\partial_t v \equiv w$. Finally, we say that $v(t) \in L^r(0, T; X)$ has a weak time derivative $\partial_t v \in L^r(0, T; X)$ if the preceding identity holds and, moreover,

$$\int_0^T \|\partial_t v(t)\|^r_X \, dt < \infty,$$

that is, $\partial_t v \in L^r(0, T; X)$.

We apply this with $X = W^{-1,p'}(\Omega)$ and $r = p'$. This makes it clear what is meant when we say $v \in L^p(0, T; W^{1,p}(\Omega))$ admits a distributional time derivative $\partial_t v \in L^{p'}(0, T; W^{-1,p'}(\Omega))$.

The next Lemma is concerned with the time derivative of the mollification $[v]_h$.

Lemma B.3. *Let X be a Banach space. Assume that $v \in L^r(0, T; X)$ with $\partial_t v \in L^r(0, T; X)$ (note that this implies $v \in C^{0,r'}([0, T]; X)$ where $r' := \frac{r}{r-1}$ when $r > 1$ and $v \in C^0([0, T]; X)$ when $r = 1$; in particular, $v(0) \in X$). Then, for the mollification in time defined by*

$$[v]_h(t) := e^{-\frac{t}{h}} v(0) + \frac{1}{h} \int_0^T e^{\frac{s-t}{h}} v(s) \, ds,$$

the time derivative $\partial_t[v]_h$ *can be computed by*

$$\partial_t[v]_h(t) = \tfrac{1}{h}\int_0^t e^{\frac{s-t}{h}}\,\partial_s v(s)\,ds$$

and, moreover, its $L^r(0,T;X)$*-norm is bounded independently of* h *by the* $L^r(0,T;X)$*-norm of* $\partial_t v$*, that is, we have*

$$\|\partial_t[v]_h\|_{L^r(0,T;X)} \leqq \|\partial_t v\|_{L^r(0,T;X)}.$$

Proof. From Lemma B.1 we know that $\partial_t[v]_h \in L^r(0,T;X)$ and for almost every $t \in (0,T)$ we have that

$$\partial_t[v]_h(t) = -\tfrac{1}{h}\big([v]_h(t) - v(t)\big) = \tfrac{1}{h}\left(v(t) - e^{-\frac{t}{h}}v(0) - \frac{e^{-\frac{t}{h}}}{h}\int_0^t e^{\frac{s}{h}}v(s)\,ds\right).$$

Now, using the assumption $\partial_t v \in L^r(0,T;X)$—which in particular implies $v \in C^{0,r'}([0,T];X)$—we may continue the preceding computation as follows:

$$\begin{aligned}
\partial_t[v]_h(t) &= \tfrac{1}{h}\left(v(t) - e^{-\frac{t}{h}}v(0) - e^{-\frac{t}{h}}\int_0^t \frac{d}{ds}e^{\frac{s}{h}}\,v(s)\,ds\right)\\
&= \tfrac{1}{h}\left(v(t) - e^{-\frac{t}{h}}v(0) - v(t) + e^{-\frac{t}{h}}v(0) + e^{-\frac{t}{h}}\int_0^t e^{\frac{s}{h}}\,\partial_s v(s)\,ds\right)\\
&= \tfrac{1}{h}\int_0^t e^{\frac{s-t}{h}}\,\partial_s v(s)\,ds.
\end{aligned}$$

This proves the first assertion of the lemma. Applying Lemma B.1 with (v_o, v) replaced by $(0, \partial_s v)$ in the right-hand side we conclude that

$$\|\partial_t[v]_h\|_{L^r(0,T;X)} = \left\|\tfrac{1}{h}\int_0^t e^{\frac{s-t}{h}}\,\partial_s v(s)\,ds\right\|_{L^r(0,T;X)} \leqq \|\partial_t v\|_{L^r(0,T;X)}.$$

This proves the asserted $L^r(0,T;X)$-estimate for the time derivative and finishes the proof of the Lemma. □

References

1. Acerbi, E., Mingione, G., Seregin, G.A.: Regularity results for parabolic systems related to a class of non-newtonian fluids. *Ann. Inst. Henri Poincaré Anal. Non Linéaire*, **21**(1), 25–60 (2004)
2. Acerbi, E., Fusco, N.: Regularity for minimizers of nonquadratic functionals: the case $1 < p < 2$. *J. Math. Anal. Appl.* **140**(1), 115–135 (1989)
3. Adams, R.A.: Sobolev Spaces. In: *Pure and Applied Mathematics*, vol. 65. Academic Press, New York, 1975
4. Bögelein, V., Duzaar, F.: Higher integrability for parabolic systems with non-standard growth and degenerate diffusions, *Publ. Math.* **55**(1), 201–250 (2011)
5. Bögelein, V., Duzaar, F., Marcellini, P.: Parabolic equations with p, q-growth. *J. Math. Pures Appl.* (2013). doi:10.1016/j.matpur.2013.01.012
6. Bögelein, V., Duzaar, F., Mingione, G.: Degenerate problems with irregular obstacles, *J. Reine Angew. Math.* **650**, 107–160 (2011)

7. BREZIS, H., MIRONESCU, P.: Gagliardo–Nirenberg, composition and products in fractional Sobolev spaces. *J. Evol. Equ.* **1**(4), 387–404 (2001)
8. CAROZZA, M., KRISTENSEN, J., PASSARELLI DI NAPOLI, A.: Higher differentiability of minimizers of convex variational integrals. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **28**(3), 395–411 (2011)
9. CUPINI, G., FUSCO, N., PETTI, R.: Hölder continuity of local minimizers. *J. Math. Anal. Appl.* **235**(2), 578–597 (1999)
10. CUPINI, G., MARCELLINI, P., MASCOLO, E.: Regularity under sharp anisotropic general growth conditions. *Discr. Contin. Dyn. Syst. Ser. B* **11**, 66–86 (2009)
11. CUPINI, G., MARCELLINI, P., MASCOLO, E.: Local boundedness of solutions to quasilinear elliptic systems. *Manuscripta Math.* **137**, 287–315 (2012)
12. DIBENEDETTO, E.: *Degenerate parabolic equations*. Springer-Verlag, Universitytext xv, 387, New York, NY, 1993
13. DI NEZZA, E., PALATUCCI, G., VALDINOCI, E.: Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.***136**(5), 521–573 (2012)
14. DUZAAR, F., MINGIONE, G., STEFFEN, K.: Parabolic systems with polynomial growth and regularity. *Mem. Am. Math. Soc.* **214**, 1005 (2011)
15. ESPOSITO, L., LEONETTI, F., MINGIONE, G.: Higher integrability for minimizers of integral functionals with p, q growth. *J. Differ. Equ.* **157**(2), 414–438 (1999)
16. GIUSTI, E.: *Direct Methods in the Calculus of Variations*. World Scientific Publishing Company, Tuck Link, Singapore, 2003
17. LICHNEWSKY, A., TEMAM, R.: Pseudosolutions of the time-dependent minimal surface problem. *J. Differ. Equ.*, **30**(3), 340–364 (1978)
18. LIEBERMAN, G.M.: Gradient estimates for a new class of degenerate elliptic and parabolic equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **21**(4), 497–522 (1994)
19. LIEBERMAN, G.M.: Gradient estimates for anisotropic elliptic equations. *Adv. Differ. Equ.* **10**(7), 767–812 (2005)
20. LIONS, J.L.: *Quelques méthodes de résolution des problémes aux limites non linéaires*. Dunod, Paris, 1969
21. MACHIHARA, S., OZAWA, T.: Interpolation inequalities in Besov spaces. *Proc. Am. Math. Soc.* **131**(5), 1553–1556 (2003)
22. MARCELLINI, P.: Regularity of minimizers of integrals of the calculus of variations with nonstandard growth conditions. *Arch. Rational Mech. Analysis* **105**(3), 267–284 (1989)
23. MARCELLINI, P.: Regularity and existence of solutions of elliptic equations with p, q-growth conditions. *J. Differ. Equ.* **90**(1), 1–30 (1991)
24. MARCELLINI, P.: Regularity for elliptic equations with general growth conditions. *J. Differ. Equ.* **105**, 296–333 (1993)
25. MARCELLINI, P.: Everywhere regularity for a class of elliptic systems without growth conditions. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **23**(1), 1–25 (1996)
26. MARCELLINI, P., PAPI, G.: Nonlinear elliptic systems with general growth. *J. Differ. Equ.* **221**, 412–443 (2006)
27. MASHIYEV, R.A., BUHRII, O.M.: Existence of solutions of the parabolic variational inequality with variable exponent of nonlinearity. *J. Math. Anal. Appl.* **377**(2), 450–463 (2011)
28. NAUMANN, J.: *Einführung in die Theorie parabolischer Variationsungleichungen*. Teubner Verlagsgesellschaft, Leipzig, 1984
29. SCHMIDT, T.: Regularity theorems for degenerate quasiconvex energies with p, q-growth. *Adv. Calc. Var.* **1**(3), 241–270 (2008)
30. SCHMIDT, T.: Regularity of relaxed minimizers of quasiconvex variational integrals with p, q-growth. *Arch. Rational Mech. Anal.* **193**(2), 311–337 (2009)
31. SHOWALTER, R.E.: *Monotone operators in Banach space and nonlinear partial differential equations*. In: Mathematical Surveys and Monographs, vol. 49. American Mathematical Society, Providence, 1997
32. SIMON, J.: Compact sets in the space $L^p(0, T; B)$. *Ann. Mat. Pura Appl., IV.* **146**, 65–96 (1987)

33. Zhikov, V.V., Pastukhova, S.E.: On the property of higher integrability for parabolic systems of variable order of nonlinearity. *Mat. Zametki* **87**(2), 179–200 (2010); *translation in Math. Notes* **87**, 169–188 (2010)
34. Strauss, W.: On continuity of functions with values in various Banach spaces. *Pacific J. Math.* **19**, 543–551 (1966)

Department Mathematik
Universität Erlangen–Nürnberg
Cauerstrasse 11, 91058 Erlangen, Germany.
e-mail: boegelein@math.fau.de
e-mail: duzaar@math.fau.de

and

Dipartimento di Matematica "U.Dini"
Università di Firenze
Viale Morgagni 67/A, 50134 Firenze, Italy.
e-mail: marcellini@math.unifi.it

(*Received January 3, 2013 / Accepted April 13, 2013*)
Published online May 22, 2013 – © *Springer-Verlag Berlin Heidelberg* (*2013*)