



# $\mathcal{A}$ -caloric approximation and partial regularity for parabolic systems with Orlicz growth

Mikil Foss<sup>1</sup> · Teresa Isernia<sup>2</sup> · Chiara Leone<sup>3</sup> · Anna Verde<sup>3</sup>

Received: 22 March 2022 / Accepted: 21 September 2022 / Published online: 24 December 2022  
© The Author(s) 2022

## Abstract

We prove a new  $\mathcal{A}$ -caloric approximation lemma compatible with an Orlicz setting. With this result, we establish a partial regularity result for parabolic systems of the type

$$u_t - \operatorname{div} a(Du) = 0.$$

Here the growth of  $a$  is bounded by the derivative of an  $N$ -function  $\varphi$ . The primary assumption for  $\varphi$  is that  $t\varphi''(t)$  and  $\varphi'(t)$  are uniformly comparable on  $(0, \infty)$ .

**Mathematics Subject Classification** 35B65 · 35K40 · 46E30

## 1 Introduction

In this paper, we establish a partial regularity result for weak solutions to parabolic systems with general growth. By partial regularity, we mean Hölder continuity for the spatial gradient outside a closed set with zero measure. Let  $\Omega \subset \mathbb{R}^n$  be an open bounded set,  $n \geq 2$ ,  $T > 0$ , and  $N \geq 1$ ; we consider weak solutions  $u : \Omega_T \rightarrow \mathbb{R}^N$ , where  $\Omega_T = \Omega \times (-T, 0)$ , to the

---

Communicated by L. Szekelyhidi.

---

✉ Chiara Leone  
chiara.leone@unina.it

Mikil Foss  
mikil.foss@unl.edu

Teresa Isernia  
t.isernia@univpm.it

Anna Verde  
anna.verde@unina.it

<sup>1</sup> Department of Mathematics, University of Nebraska-Lincoln, 203 Avery Hall, Lincoln, NE 68588-0130, USA

<sup>2</sup> Dipartimento di Ingegneria Industriale e Scienze Matematiche, Università Politecnica delle Marche, Via Brecce Bianche 12, 60131 Ancona, Italy

<sup>3</sup> Dipartimento di Matematica “R. Caccioppoli”, Università degli Studi di Napoli Federico II Via Cinthia, Complesso Universitario di Monte S. Angelo, 80126 Napoli, Italy

following homogeneous parabolic system

$$u_t - \operatorname{div} a(Du) = 0 \quad \text{in } \Omega_T, \quad (1.1)$$

where the  $C^1$ -vector field  $a : \mathbb{R}^{Nn} \rightarrow \mathbb{R}^{Nn}$  satisfies ellipticity and growth conditions in terms of Orlicz functions. The precise structural assumptions on the vector field  $a$  will be presented later, but the principal prototype we have in mind is the parabolic  $\varphi$ -Laplacian system

$$u_t - \operatorname{div} \left( \frac{\varphi'(\mu + |Du|)}{\mu + |Du|} Du \right) = 0, \quad (1.2)$$

where  $\mu > 0$  and  $\varphi$  is an Orlicz function (see Sect. 2). In the model case, with  $\varphi(s) = \frac{1}{p}s^p$  for some  $p > 1$ , (1.2) gives the more familiar non-degenerate evolutionary  $p$ -Laplacian system:

$$u_t - \operatorname{div} ((\mu + |Du|)^{p-2} Du) = 0. \quad (1.3)$$

Hence system (1.2) (and consequently (1.1)) can be seen as a generalization of the  $p$ -Laplacian parabolic system (1.3). In particular, in addition to not requiring the system (1.1) to have the standard  $p$ -growth, we do not assume an Uhlenbeck structure as in (1.2).

The literature is rich with regularity results for parabolic systems with standard  $p$ -growth. In the paper by DiBenedetto and Friedman [14], everywhere regularity is proved. In this paper, the system has an Uhlenbeck structure:  $a(\xi) = |\xi|^{p-2}\xi$  and  $p > \frac{2n}{n+2}$ . Using a combination of the Moser and De Giorgi iteration schemes, the solution's spatial gradient is shown to be bounded and Hölder continuous on its domain. In [15], the authors extended their result to allow nonlinear forcing and introduced the intrinsic scaling which accommodates the singular ( $p < 2$ ) or degenerate ( $p > 2$ ) behavior of  $a$  in a natural way. (For a comprehensive introduction and collection of results on the subject, we refer the reader to DiBenedetto's book [13].) It is well-known that, without special structural assumptions, solutions to systems can only be expected to possess partial regularity, that is regularity on an open set of full measure. Giaquinta and Giusti [28] provided the first result in this direction. Adapting a blow-up argument, successfully used for elliptic systems, they showed partial Hölder continuity for the weak solution,  $u$ , of nondegenerate systems with  $p$ -growth ( $p \geq 2$ ). By again adapting techniques for elliptic systems, Giaquinta and Struwe proved higher integrability and partial Hölder continuity for a solution's spatial gradient,  $Du$ , provided  $a$  has quadratic growth. For a general nonlinear  $a(z, u, \xi)$  with quadratic growth, partial regularity for the spatial gradient remained an open problem until the work of Duzaar and Mingione [25]. In this transformative paper, the authors introduced the, now well-known,  $\mathcal{A}$ -caloric approximation approach to regularity theory for parabolic systems. Generalizations to problems with superquadratic or subquadratic growth were provided by Duzaar, Mingione, and Steffen [26] and Scheven [41]. Utilizing intrinsic scaling and  $p$ -caloric approximation, along with  $\mathcal{A}$ -caloric approximation, Bögelein, Duzaar, and Mingione [5] extended these results to  $p$ -growth systems, of the form (1.1), that are potentially degenerate ( $p > 2$ ) or singular  $\frac{2n}{n+2} < p < 2$ . Without any attempt for completeness, we also mention to the papers [1, 5, 6, 27, 39] where the partial Hölder continuity, either for the solution's spatial gradient or the solution itself, is established. The main goal of this paper is to extend some of these partial regularity results into the Orlicz-growth setting. In the papers cited above, the superquadratic  $p \geq 2$  and subquadratic  $p < 2$  cases require different techniques. Working in an Orlicz setting, we provide a unified treatment for both system classes.

There is a long history of interest in partial differential equations with nonstandard growth. Early existence results for both elliptic and parabolic problems were established by Donaldson [23, 24] (see also [43]). For elliptic equations and scalar-valued variational problems with

$(p, q)$ -growth, Marcellini [35, 36] developed an approximation and Moser iteration technique to prove everywhere regularity. For elliptic systems with an Uhlenbeck structure, Marcellini and Papi [38] extended this strategy to even allow problems oscillating between linear and exponential growth (see also [37]). Under general growth conditions, additional results for elliptic systems can be found, for example, in [7, 10–12, 16, 18, 21, 22] and the references therein. Regarding regularity for parabolic systems with general growth, much less work is available in the literature. Assuming an Uhlenbeck structure, the iteration strategies developed in [14, 15] have been adapted to problems with form (1.1). Assuming  $t\varphi'(t)$  and  $\varphi(t)$  are comparable on  $(0, \infty)$ , Lieberman [34] proved that a weak solution  $u$  to (1.2) has a Hölder continuous spatial gradient  $Du$ , provided  $Du$  is already known to be bounded. In [44], You removed the boundedness assumption, but under stricter growth assumptions. More recently, the boundedness of  $|Du|$  has been established under more general conditions. In [19], Diening, Scharle, and Schwarzacher assume  $t\varphi''(t)$  and  $\varphi'(t)$  to be comparable, while in [32], only a doubling property is needed to obtain the boundedness of  $u$ . For additional regularity and higher integrability results, where an Uhlenbeck or similar structure is assumed, we also mention [2, 3, 8, 20]. Without such a structural assumption, higher integrability was established by Hästö and Ok in [31]. As far as the authors are aware, the current paper is the first to establish the partial Hölder continuity of a weak solution’s spatial gradient. We now list the specific assumptions needed (see also Sect. 2).

**Assumption 1.1** Let  $\varphi \in C^1([0, \infty)) \cap C^2(0, \infty)$  be an  $N$ -function satisfying

$$0 < p_0 - 1 \leq \inf_{t>0} \frac{t\varphi''(t)}{\varphi'(t)} \leq \sup_{t>0} \frac{t\varphi''(t)}{\varphi'(t)} \leq p_1 - 1,$$

with  $\frac{2n}{n+2} < p_0 \leq p_1$ . Without loss of generality we can assume that  $p_0 < 2 < p_1$ .

With this  $\varphi$ , we consider (1.1) under the following hypotheses on the  $C^1$ -vector field  $a : \mathbb{R}^{Nn} \rightarrow \mathbb{R}^{Nn}$ :

(a<sub>1</sub>) There exists  $L > 0$  such that

$$|a(\xi)| \leq L\varphi'(1 + |\xi|)$$

holds for every  $\xi \in \mathbb{R}^{Nn}$ ;

(a<sub>2</sub>) There exists  $\nu > 0$  such that

$$Da(\xi)(\eta, \eta) \geq \nu\varphi''(1 + |\xi|)|\eta|^2$$

holds for any  $\xi, \eta \in \mathbb{R}^{Nn}$ ;

(a<sub>3</sub>) For every  $\xi \in \mathbb{R}^{Nn}$

$$|Da(\xi)| \leq L\varphi''(1 + |\xi|);$$

(a<sub>4</sub>) There exists a nondecreasing and concave function  $\omega : [0, \infty) \rightarrow [0, 1]$  with  $\omega(0) = 0$  such that

$$|Da(\xi) - Da(\eta)| \leq L\omega\left(\frac{|\xi - \eta|}{1 + |\xi| + |\eta|}\right)\varphi''(1 + |\xi| + |\eta|)$$

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

While it ensures  $t\varphi''(t)$  is comparable  $\varphi'(t)$  for  $t > 0$ , Assumption 1.1 does not imply  $\varphi$  has  $p$ -growth. It does imply  $\varphi$  and  $\varphi^*$  have the doubling property (2.1). Similar assumptions also appear in several of the works cited above.

The notion of weak solution adopted in the present paper is the following:  $u \in C^0(-T, 0, L^2(\Omega, \mathbb{R}^N)) \cap L^1(-T, 0, W^{1,1}(\Omega, \mathbb{R}^N))$  with  $\varphi(|Du|) \in L^1(-T, 0, L^1(\Omega))$  is a weak solution to (1.1) if it holds

$$\int_{\Omega_T} u \cdot \eta_t - a(Du) \cdot D\eta \, dx \, dt = 0,$$

for all  $\eta \in C_c^\infty(\Omega_T, \mathbb{R}^N)$ . Here  $Du : (-T, 0) \times \Omega \rightarrow \mathbb{R}^{Nn}$  denotes the spatial gradient of  $u$ . We can now state our regularity result.

**Theorem 1.2** *Let  $u \in C^0(-T, 0; L^2(\Omega, \mathbb{R}^N)) \cap L^1(-T, 0; W^{1,1}(\Omega, \mathbb{R}^N))$  with  $\varphi(|Du|) \in L^1(-T, 0; L^1(\Omega))$  be a weak solution to (1.1) in  $\Omega_T$  under hypotheses (a<sub>1</sub>)-(a<sub>4</sub>) and Assumption 1.1. Then for every  $\alpha \in (0, 1)$  there exists an open subset  $\Omega_0 \subseteq \Omega_T$  such that*

$$V(Du) \in C_{loc}^{0, \frac{\alpha}{2}, \alpha}(\Omega_0, \mathbb{R}^{Nn}) \quad \text{and} \quad |\Omega_T \setminus \Omega_0| = 0,$$

where  $V(\xi) = \sqrt{\frac{\varphi'(1+|\xi|)}{1+|\xi|}} \xi$ . Moreover, the singular set  $\Omega_T \setminus \Omega_0 \subseteq \Sigma_1 \cup \Sigma_2$ , where

$$\Sigma_1 = \left\{ z_0 \in \Omega_T : \liminf_{\rho \rightarrow 0} \int_{\mathcal{Q}_\rho(z_0)} |V(Du) - (V(Du))_{z_0, \rho}|^2 dz > 0 \right\},$$

$$\Sigma_2 = \left\{ z_0 \in \Omega_T : \limsup_{\rho \rightarrow 0} |(Du)_{z_0, \rho}| = +\infty \right\},$$

denoting the mean value of a function over the parabolic cylinder  $\mathcal{Q}_\rho(z_0) = \mathcal{B}_\rho(x_0) \times (t_0 - \rho^2, t_0)$  by  $(\cdot)_{z_0, \rho}$ .

Here  $z_0 = (x_0, t_0) \in \Omega \times (-T, 0)$  and  $\mathcal{B}_\rho(x_0)$  is the ball in  $\mathbb{R}^n$  with radius  $\rho$  centered at  $x_0$ . Note that the Hölder continuity of  $V(Du)$  implies the Hölder continuity of  $Du$  with a different exponent depending on  $\varphi$ .

The proof of Theorem 1.2 relies on a decay estimate for certain excess functionals, which measure in a suitable way the oscillations of the solutions. More precisely, for  $z_0 \in \Omega_T$ ,  $r > 0$ ,  $a \geq 0$ , and an affine map  $\ell : \mathbb{R}^n \rightarrow \mathbb{R}^N$ , we define the excess functional by

$$\Psi_a(z_0, r, \ell) = \int_{\mathcal{Q}_r(z_0)} \left( \left| \frac{u - \ell}{r} \right|^2 + \varphi_a \left( \left| \frac{u - \ell}{r} \right| \right) \right) dz.$$

Here  $\varphi_a$  interpolates in a certain sense between  $t^2$ , when  $t \leq a$ , and  $\varphi$ , when  $t \geq a$ . (The precise definition of the function  $\varphi_a$  is given in Sect. 2.)

In order to achieve the decay estimate, we first derive the Caccioppoli inequality which is compatible with (1.1) (see Theorem 4.1). This, in particular, allows us to control the spatial oscillations of  $u$  and oscillations in  $Du$  via the excess functional. Though  $u$  need not be differentiable, if  $\Psi_a$  is sufficiently small, then a family of smooth approximations to a spatial linearization of  $u$ , centered at  $z_0$ , can be produced. These 1st-order surrogates for  $u$  are, in fact, solutions to a constant coefficient parabolic system. Moreover, their approximation to  $u$  improves as  $r \rightarrow 0^+$  providing a decay estimate for the  $\Psi_a$ , which implies the oscillations in  $V(Du)$  decrease as  $r \rightarrow 0^+$ . The rate of decrease is fast enough to deliver the regularity of  $V(Du)$  through Campanato’s characterization of Hölder continuity.

The outline of the proof of Theorem 1.2 follows the approach developed in [26, 41]. The cornerstone to the strategy is the  $\mathcal{A}$ -caloric approximation theorem, which provides the family of approximations to  $u$ . The generalization of this theorem to something suitable for the Orlicz

setting was a significant obstacle and is the paper’s principal novelty. Our proof for this result does not require Assumption 1.1. In fact,  $\varphi$  is only assumed to be an  $N$ -function with super  $\frac{2n}{n+2}$ -growth and a doubling property near zero. Thus  $\varphi$  may have exponential or even super-exponential growth. A key difference between the  $p$ -growth and Orlicz settings is that for  $L^p$ -spaces one has

$$L^p(Q_\rho) = L^p(-\rho^2, 0; L^p(\mathcal{B}_\rho)),$$

but for the Orlicz spaces  $L^\varphi$  one only has

$$L^\varphi(Q_\rho) \subseteq L^\varphi(-\rho^2, 0; L^\varphi(\mathcal{B}_\rho)).$$

In fact, equality holds if and only if  $\varphi(t)$  is comparable to  $t^p$  (see the remarks following Proposition 1.3 in [24]). With standard growth, the proofs for the  $\mathcal{A}$ -caloric approximation theorem can take advantage of Simon’s compactness result [42] in  $L^p(-\rho^2, 0; L^p(\mathcal{B}_\rho))$  to directly obtain convergence in  $L^p(Q_\rho)$ . This, however, is not possible in the Orlicz setting. While we use Simon’s result for convergence in  $L^2$ , upgrading to convergence in  $L^\varphi(Q_\rho)$  involves a combination of approximations via convolution, sophisticated pointwise estimates, and integral bounds for the non-centered Hardy-Littlewood maximal function. With this new  $\mathcal{A}$ -caloric approximation, we prove a decay estimate for the excess function. Employing a standard iteration argument, we are able to identify the singular set with points where either the excess cannot be made sufficiently small,  $\Sigma_1$ , or the mean  $(Du)_{z_0, \rho}$  is not bounded,  $\Sigma_2$ . Finally, to prove that the singular set is negligible, we use a Poincaré-Sobolev-type inequality for solutions to (1.1) which bounds the excess of  $u$  in terms of its spatial gradient  $Du$ . The proof for this inequality is rather complicated and relies on a Gagliardo-Nirenberg inequality from [31].

The paper is organized as follows: after collecting the basic terminology and other preliminaries in Sect. 2, we present the  $\mathcal{A}$ -caloric approximation in Sect. 3. In Sect. 4, the proofs of the Caccioppoli inequalities in the parabolic setting. We detail the Poincaré-Sobolev-type inequalities in Sect. 5 and the linearization in Sect. 6. We finally establish the decay estimates and the main theorem in the last two sections.

## 2 Notation and preliminary results

Let  $\Omega \subset \mathbb{R}^n$  be a bounded domain; in the following  $\Omega_T$  will denote the parabolic cylinder  $\Omega \times (-T, 0)$ , where  $T > 0$ . If  $z \in \Omega_T$ , we denote  $z = (x, t)$  with  $x \in \Omega$  and  $t \in (-T, 0)$ . In what follows  $C$  will be often a general positive constant, possibly varying from line to line, but depending on only the structural parameters  $n, N, L/v, p_0, p_1$ , with  $1 < p_0, p_1 < \infty$  identified in Assumption 1.1 above. The notation  $Du(x, t) \equiv D_x u(x, t)$  denotes the differentiation with respect to the spatial variable  $x$ , and  $u_t$  stands for the differentiation with respect to the time variable.

With  $x_0 \in \mathbb{R}^n$ , we set

$$\mathcal{B}_r(x_0) := \{x \in \mathbb{R}^n : |x - x_0| < r\}$$

the open ball of  $\mathbb{R}^n$  with radius  $r > 0$  and center  $x_0$ . When dealing with parabolic regularity, the geometry of cylinders plays an important role. We denote the general cylinder with spatial radius  $\rho$  and time length  $\tau$  centered at  $z_0 = (x_0, t_0)$  by

$$Q_{\rho, \tau}(z_0) = \mathcal{B}_\rho(x_0) \times (t_0 - \tau, t_0),$$

and we define the standard parabolic cylinder by

$$\mathcal{Q}_\rho(z_0) = \mathcal{Q}_{\rho, \rho^2}(z_0) = \mathcal{B}_\rho(x_0) \times (t_0 - \rho^2, t_0).$$

Given a cylinder  $\mathcal{Q} = \mathcal{B} \times (s, t)$ , its parabolic boundary is

$$\partial_P \mathcal{Q} := (\mathcal{B} \times \{s\}) \cup (\partial \mathcal{B} \times [s, t]).$$

The integral averages of a function  $u$  on  $\mathcal{Q} \subset \mathbb{R}^{n+1}$  are given by

$$(u)_{\mathcal{Q}}(t) = \int_{\mathcal{Q}} u(x, t) dx, \quad (u)_{\mathcal{Q}} = \int_{\mathcal{Q}} u(x, t) dz.$$

We will denote the average  $(u)_{\mathcal{Q}_\rho(z_0)}$  by  $(u)_{z_0, \rho}$ . The parabolic metric is defined as usual by

$$dist_P(z, z_0) := \sqrt{|x - x_0|^2 + |t - t_0|}$$

whenever  $z = (x, t)$ ,  $z_0 = (x_0, t_0) \in \mathbb{R}^{n+1}$ .

We recall that a strongly elliptic bilinear form  $\mathcal{A}$  on  $\mathbb{R}^{Nn}$  with ellipticity constant  $\nu > 0$  and upper bound  $L > 0$  means that

$$\nu|\xi|^2 \leq \mathcal{A}(\xi, \xi), \quad \mathcal{A}(\xi, \tilde{\xi}) \leq L|\xi||\tilde{\xi}| \quad \forall \xi, \tilde{\xi} \in \mathbb{R}^{Nn}.$$

**Definition 2.1** We shall say that a function  $h \in L^2(t_0 - \rho^2, t_0; W^{1,2}(\mathcal{B}_\rho(x_0), \mathbb{R}^N))$  is  $\mathcal{A}$ -caloric on  $\mathcal{Q}_\rho(z_0)$  if it satisfies

$$\int_{\mathcal{Q}_\rho(z_0)} h \cdot \eta_t - \mathcal{A}(Dh, D\eta) dz = 0, \quad \text{for all } \eta \in C_c^\infty(\mathcal{Q}_\rho(z_0), \mathbb{R}^N).$$

**Remark 2.1** In the following we shall often write  $u_t$  even if a weak solution of a parabolic system may not be differentiable in the time variable. The arguments can be made rigorous by the use of a smoothing procedure in time, as for instance via Steklov averages. However, since this argument is by now quite standard, we shall abuse the notation  $u_t$  proceeding formally, without further explanation.

### 2.1 $N$ -functions

We begin recalling the notion of  $N$ -functions (see [40]).

We write  $f \sim g$ , and we say that  $f$  and  $g$  are equivalent, if there exist constants  $c_1, c_2 > 0$  such that  $c_1 g(t) \leq f(t) \leq c_2 g(t)$  for any  $t \geq 0$ . Similarly the symbol  $\lesssim$  stands for  $\leq$  up to a constant.

**Definition 2.2** A real convex function  $\varphi : [0, \infty) \rightarrow [0, \infty)$  is said to be an  $N$ -function if  $\varphi(0) = 0$  and there exists a right continuous nondecreasing derivative  $\varphi'$  satisfying  $\varphi'(0) = 0$ ,  $\varphi'(t) > 0$  for  $t > 0$  and  $\lim_{t \rightarrow \infty} \varphi'(t) = \infty$ .

An  $N$ -function  $\varphi$  satisfies the  $\Delta_2$ -condition, and we write  $\varphi \in \Delta_2$ , if there exists a constant  $c > 0$  such that

$$\varphi(2t) \leq c \varphi(t) \quad \text{for all } t \geq 0. \tag{2.1}$$

The smallest possible constant will be denoted by  $\Delta_2(\varphi)$ . Combining  $\varphi(t) \leq \varphi(2t)$  together with the  $\Delta_2$ -condition we get  $\varphi(2t) \sim \varphi(t)$ .

The *conjugate function*  $\varphi^* : [0, +\infty) \rightarrow [0, +\infty)$  of an  $N$ -function  $\varphi$  is defined by

$$\varphi^*(t) := \sup_{s \geq 0} [st - \varphi(s)] \quad \text{for all } t \geq 0.$$

It holds that  $\varphi^*$  itself is an  $N$ -function. If  $\varphi$  and  $\varphi^*$  both satisfy the  $\Delta_2$ -condition, then we will write that  $\Delta_2(\varphi, \varphi^*) := \max\{\Delta_2(\varphi), \Delta_2(\varphi^*)\} < \infty$ . Assume that  $\Delta_2(\varphi, \varphi^*) < \infty$ ; then for all  $\delta > 0$  there exists  $c_\delta$  depending only on  $\Delta_2(\varphi, \varphi^*)$  such that for all  $s, t \geq 0$  it holds the Young’s inequality

$$ts \leq \delta \varphi(t) + c_\delta \varphi^*(s).$$

In most parts of the paper we will assume that  $\varphi$  satisfies Assumption 1.1.

**Remark 2.2** We remark that the lower bound on  $p_0$  appearing in Assumption 1.1 is absolutely natural to prove regularity in such a context: one need only to consider the power case  $\varphi(t) = t^p$  (see [13]).

Under Assumption 1.1 on  $\varphi$  it follows from [12, Proposition 2.1] that  $\Delta_2(\varphi, \varphi^*) < \infty$  and

$$p_0 \leq \inf_{t > 0} \frac{t\varphi'(t)}{\varphi(t)} \leq \sup_{t > 0} \frac{t\varphi'(t)}{\varphi(t)} \leq p_1.$$

Moreover the following inequalities hold for every  $t \geq 0$ :

$$\begin{aligned} s^{p_1} \varphi(t) \leq \varphi(st) \leq s^{p_0} \varphi(t) & \quad \text{if } 0 < s \leq 1, \\ s^{p_0} \varphi(t) \leq \varphi(st) \leq s^{p_1} \varphi(t) & \quad \text{if } s \geq 1, \end{aligned} \tag{2.2}$$

as well

$$\begin{aligned} s^{\frac{p_0}{p_0-1}} \varphi^*(t) \leq \varphi^*(st) \leq s^{\frac{p_1}{p_1-1}} \varphi^*(t) & \quad \text{if } 0 < s \leq 1, \\ s^{\frac{p_1}{p_1-1}} \varphi^*(t) \leq \varphi^*(st) \leq s^{\frac{p_0}{p_0-1}} \varphi^*(t) & \quad \text{if } s \geq 1, \end{aligned} \tag{2.3}$$

and also

$$\begin{aligned} s^{p_1-1} \varphi'(t) \leq \varphi'(st) \leq s^{p_0-1} \varphi'(t) & \quad \text{if } 0 < s \leq 1, \\ s^{p_0-1} \varphi'(t) \leq \varphi'(st) \leq s^{p_1-1} \varphi'(t) & \quad \text{if } s \geq 1. \end{aligned} \tag{2.4}$$

In particular, for  $t > 0$  we have

$$\varphi(t) \sim t\varphi'(t), \quad \varphi'(t) \sim t\varphi''(t), \quad \varphi^*(\varphi'(t)) \sim \varphi(t), \quad \varphi^{-1}(t)(\varphi^*)^{-1}(t) \sim t. \tag{2.5}$$

Now, we consider a family of  $N$ -functions  $\{\varphi_a\}_{a \geq 0}$  setting, for  $t \geq 0$ ,

$$\varphi_a(t) := \int_0^t \varphi'_a(s) ds \quad \text{with} \quad \varphi'_a(t) := \varphi'(a+t) \frac{t}{a+t}.$$

The following lemma can be found in [16, Lemma 27].

**Lemma 2.1** *Let  $\varphi$  be an  $N$ -function with  $\varphi \in \Delta_2$  together with its conjugate. Then for all  $a \geq 0$  the function  $\varphi_a$  is an  $N$ -function and  $\{\varphi_a\}_{a \geq 0}$  and  $\{(\varphi_a)^*\}_{a \geq 0} \sim \{\varphi^*_{\varphi'(a)}\}_{a \geq 0}$  satisfy the  $\Delta_2$ -condition uniformly in  $a \geq 0$ .*

Let us observe that by the previous lemma  $\varphi_a(t) \sim t\varphi'_a(t)$ . Moreover, for  $t \geq a$  we have  $\varphi_a(t) \sim \varphi(t)$ , while Assumption 1.1 provides

$$C^{-1}\varphi''(a)t^2 \leq \varphi_a(t) \leq C\varphi''(a)t^2, \quad \text{for } 0 \leq t \leq a, \tag{2.6}$$

since  $\varphi' \sim t\varphi''$ . The constant  $C$  depends on only  $p_0$  and  $p_1$ . This implies also that, for all  $s \in [0, 1], a \geq 0$  and  $t \in [0, a]$ ,

$$\varphi_a(st) \leq C^2 s^2 \varphi_a(t). \tag{2.7}$$

Finally, allowing Assumption 1.1, the following relations hold uniformly with respect to  $a \geq 0$

$$\begin{aligned} \varphi_a(t) &\sim \varphi''(a+t)t^2 \sim \frac{\varphi(a+t)}{(a+t)^2}t^2 \sim \frac{\varphi'(a+t)}{a+t}t^2, \\ \varphi(a+t) &\sim \varphi_a(t) + \varphi(a). \end{aligned} \tag{2.8}$$

**Remark 2.3** It is easy to check that if  $\varphi$  satisfies Assumption 1.1, the same is true for  $\varphi_a$ , uniformly with respect to  $a \geq 0$  (with the same  $p_0$  and  $p_1$ ). This in particular means that

$$\sup_{a \geq 0} \Delta_2(\varphi_a) \leq 2^{p_1} \tag{2.9}$$

and

$$t^{p_0} \varphi_a(1) \leq \varphi_a(t) \leq t^{p_1} \varphi_a(1), \quad \text{for } t \geq a \geq 1,$$

thanks to (2.2) for  $\varphi_a$ .

Next result is a slight generalization of [16, Lemma 20].

**Lemma 2.2** *Let  $\varphi$  be an  $N$ -function with  $\Delta_2(\varphi, \varphi^*) < \infty$ ; then, uniformly in  $\xi_1, \xi_2 \in \mathbb{R}^{Nn}$  with  $|\xi_1| + |\xi_2| > 0$ , and in  $\mu \geq 0$ , it holds*

$$\frac{\varphi'(\mu + |\xi_1| + |\xi_2|)}{\mu + |\xi_1| + |\xi_2|} \sim \int_0^1 \frac{\varphi'(\mu + |\xi_\theta|)}{\mu + |\xi_\theta|} d\theta,$$

where  $\xi_\theta = \xi_1 + \theta(\xi_2 - \xi_1)$  with  $\theta \in [0, 1]$ .

**Remark 2.4** We now state the following two consequences of our structure assumptions for further reference. First, we note that the ellipticity condition  $(a_2)$  and Assumption 1.1 together with Lemma 2.2 and (2.8) imply

$$(a(\xi) - a(\xi_0)) \cdot (\xi - \xi_0) \geq c\varphi_{1+|\xi_0|}(|\xi - \xi_0|),$$

for every  $\xi, \xi_0 \in \mathbb{R}^{Nn}$ , where  $c = c(p_0, p_1, \nu)$ .

In a similar way, the growth condition  $(a_3)$  and Assumption 1.1 imply

$$|a(\xi) - a(\xi_0)| \leq c\varphi'_{1+|\xi_0|}(|\xi - \xi_0|),$$

for every  $\xi, \xi_0 \in \mathbb{R}^{Nn}$ , where  $c = c(p_0, p_1, L)$ .

The following results deal with the change of shift of  $N$ -functions  $\varphi_a$ . The first one is proved in [17, Corollary 26].

**Lemma 2.3** *Let  $\varphi$  be a  $N$ -function satisfying  $\Delta_2(\varphi, \varphi^*) < +\infty$ . Then for each  $\delta > 0$ , there exists  $c_\delta(\Delta_2(\varphi, \varphi^*)) > 0$  such that for all  $a, b \in \mathbb{R}^d$  and  $t \geq 0$*

$$\varphi_{|a|}(t) \leq c_\delta \varphi_{|b|}(t) + \delta \varphi_{|a|}(|a - b|).$$

**Lemma 2.4** *Let  $\varphi$  be a  $N$ -function satisfying Assumption 1.1; let  $M \geq 1$  and  $1 \leq a, b \leq M$  be given. Then*

$$\varphi_a(t) \leq 4^{p_1+1} M^{p_1+2} \varphi_b(t), \quad \text{for all } t \geq 0.$$



**Proof** From the definition and the fact that  $t/2\varphi'(t/2) \leq \varphi(t) \leq t\varphi'(t)$ ,

$$\varphi'_a(s) = \varphi'_b(s) \left[ \frac{\varphi'(a+s)}{\varphi'(b+s)} \left( \frac{s+b}{s+a} \right) \right] \leq \varphi'_b(s) \left[ \frac{\varphi(2(a+s))}{\varphi(b+s)} \left( \frac{s+b}{s+a} \right)^2 \right].$$

If  $s \leq M$ , then, using (2.2), we have

$$\frac{\varphi(2(a+s))}{\varphi(b+s)} \left( \frac{s+b}{s+a} \right)^2 \leq \frac{\varphi(4M)}{\varphi(1)} M^2 \leq 4^{p_1} M^{p_1+2}.$$

Otherwise,  $1 \leq a, b \leq M < s$  and

$$\frac{\varphi(2(a+s))}{\varphi(b+s)} \left( \frac{s+b}{s+a} \right)^2 \leq 4 \frac{\varphi(4s)}{\varphi(s)} \leq 4^{p_1+1}.$$

Thus

$$\varphi_a(t) = \int_0^t \varphi'_a(s) ds \leq (4^{p_1} M^{p_1+2} + 4^{p_1+1}) \int_0^t \varphi'_b(s) ds \leq 4^{p_1+1} M^{p_1+2} \varphi_b(t).$$

□

We will use the function  $V : \mathbb{R}^{Nn} \rightarrow \mathbb{R}^{Nn}$  defined by

$$V(\xi) = \sqrt{\frac{\varphi'(1+|\xi|)}{1+|\xi|}} \xi.$$

The monotonicity property of  $\varphi$  ensures that

$$|V(\xi_1) - V(\xi_2)|^2 \sim \varphi_{1+|\xi_1|}(|\xi_1 - \xi_2|) \quad \text{for any } \xi_1, \xi_2 \in \mathbb{R}^{Nn}; \tag{2.10}$$

see [16] for further properties about the  $V$ -function.

Let  $\varphi$  be an  $N$ -function that satisfies the  $\Delta_2$ -condition. The set of functions  $L^\varphi(\Omega, \mathbb{R}^N)$  is defined by

$$L^\varphi(\Omega, \mathbb{R}^N) = \left\{ u : \Omega \rightarrow \mathbb{R}^N \text{ measurable} : \int_\Omega \varphi(|u|) dx < \infty \right\}.$$

The Luxembourg norm is defined as follows:

$$\|u\|_{L^\varphi(\Omega, \mathbb{R}^N)} = \inf \left\{ \lambda > 0 : \int_\Omega \varphi \left( \frac{|u(x)|}{\lambda} \right) dx \leq 1 \right\}.$$

With this norm  $L^\varphi(\Omega, \mathbb{R}^N)$  is a Banach space.

By  $W^{1,\varphi}(\Omega, \mathbb{R}^N)$  we denote the classical Orlicz-Sobolev space, that is  $u \in W^{1,\varphi}(\Omega, \mathbb{R}^N)$  whenever  $u, Du \in L^\varphi(\Omega, \mathbb{R}^N)$ . Furthermore, by  $W_0^{1,\varphi}(\Omega, \mathbb{R}^N)$  we mean the closure of  $C_c^\infty(\Omega, \mathbb{R}^N)$  functions with respect to the norm

$$\|u\|_{W^{1,\varphi}(\Omega, \mathbb{R}^N)} = \|u\|_{L^\varphi(\Omega, \mathbb{R}^N)} + \|Du\|_{L^\varphi(\Omega, \mathbb{R}^N)}.$$

For a function  $u \in L^\varphi(Q_\rho(z_0), \mathbb{R}^N)$ , using the decomposition

$$Q_\rho^\leq(z_0) = \{z \in Q_\rho(z_0) : |u(z)| \leq a\} \quad \text{and} \quad Q_\rho^\gt(z_0) = \{z \in Q_\rho(z_0) : |u(z)| > a\},$$

as well as (2.6) and Remark 2.3, we easily get the following lemma.

**Lemma 2.5** *Let  $\varphi$  be a  $N$ -function satisfying Assumption 1.1 and let  $u \in L^\varphi(Q_\rho(z_0), \mathbb{R}^N)$ ,  $a \geq 1$ . Then*

(a)

$$\int_{\mathcal{Q}_\rho(z_0)} \varphi_a(|u|) dz \leq C\varphi''(a) \int_{\mathcal{Q}_\rho(z_0)} |u|^2 dz + \varphi_a(1) \int_{\mathcal{Q}_\rho(z_0)} |u|^{p_1} dz;$$

(b) for each  $0 \leq s \leq p_0 < 2$ ,

$$\int_{\mathcal{Q}_\rho(z_0)} |u|^s dz \leq \left( \frac{C}{\varphi''(a)} \int_{\mathcal{Q}_\rho(z_0)} \varphi_a(|u|) dz \right)^{\frac{s}{2}} + \frac{1}{\varphi_a(1)} \int_{\mathcal{Q}_\rho(z_0)} \varphi_a(|u|) dz.$$

Here  $C$  only depends on  $p_0, p_1$ .

### 2.2 Affine functions

Let  $z_0 \in \mathbb{R}^{n+1}$  and  $\rho > 0$ . Given  $u \in L^2(\mathcal{Q}_\rho(z_0), \mathbb{R}^N)$ , we denote by  $\ell_{z_0, \rho} : \mathbb{R}^n \rightarrow \mathbb{R}^N$  the unique affine function minimizing the functional

$$\ell(x) \mapsto \int_{\mathcal{Q}_\rho(z_0)} |u(x, t) - \ell(x)|^2 dz$$

amongst all affine functions  $\ell : \mathbb{R}^n \rightarrow \mathbb{R}^N$ . It is well known (see [5]) that

$$\ell_{z_0, \rho}(x) = (u)_{z_0, \rho} + P_{z_0, \rho}(x - x_0), \tag{2.11}$$

where

$$P_{z_0, \rho} = \frac{n+2}{\rho^2} \int_{\mathcal{Q}_\rho(z_0)} u(x, t) \otimes (x - x_0) dz. \tag{2.12}$$

The following lemma ensures that  $\ell_{z_0, \rho}$  is an almost minimizer of the functional

$$\ell \mapsto \int_{\mathcal{Q}_\rho(z_0)} \varphi\left(\frac{|u - \ell|}{r}\right) dz$$

amongst the affine functions  $\ell : \mathbb{R}^n \rightarrow \mathbb{R}^N$ .

**Lemma 2.6** *Let  $\varphi$  be an  $N$ -function satisfying the  $\Delta_2$ -property and let  $u \in L^\varphi(\mathcal{Q}_\rho(z_0), \mathbb{R}^N)$ . Let  $r > 0$ , then there exists a constant  $\kappa_0 = \kappa_0(n, \Delta_2(\varphi)) > 0$  such that*

$$\int_{\mathcal{Q}_\rho(z_0)} \varphi\left(\frac{|u - \ell_{z_0, \rho}|}{r}\right) dz \leq \kappa_0 \int_{\mathcal{Q}_\rho(z_0)} \varphi\left(\frac{|u - \ell|}{r}\right) dz,$$

for every affine function  $\ell : \mathbb{R}^n \rightarrow \mathbb{R}^N$ .

**Proof** Assume  $z_0 = (0, 0)$  and denote  $\ell_{z_0, \rho}$ ,  $\mathcal{Q}_\rho(z_0)$ , and  $(u)_{z_0, \rho}$  by  $\ell_\rho$ ,  $\mathcal{Q}_\rho$ , and  $(u)_\rho$ , respectively. Let us consider a generic affine function  $\ell(x) = \zeta + Ax$ , then, for  $x \in \mathcal{B}_\rho$ ,

$$|\ell - \ell_\rho| = |(u)_\rho - \zeta + (D\ell_\rho - A)x| \leq |(u)_\rho - \zeta| + \rho|D\ell_\rho - A|.$$

Now we have

$$|(u)_\rho - \zeta| = \left| \int_{\mathcal{Q}_\rho} (u - \zeta) dz \right| = \left| \int_{\mathcal{Q}_\rho} (u - \zeta - Ax) dz \right| \leq \int_{\mathcal{Q}_\rho} |u - \ell| dz,$$

and, using (2.12),

$$|D\ell_\rho - A| = \frac{n+2}{\rho^2} \left| \int_{\mathcal{Q}_\rho} (u - Ax) \otimes x dz \right| = \frac{n+2}{\rho^2} \left| \int_{\mathcal{Q}_\rho} (u - \zeta - Ax) \otimes x dz \right|$$

$$\leq \frac{n+2}{\rho} \int_{Q_\rho} |u - \ell| dz. \tag{2.13}$$

In conclusion

$$|\ell - \ell_\rho| \leq (n+3) \int_{Q_\rho} |u - \ell| dz. \tag{2.14}$$

Recalling that, by the convexity and the  $\Delta_2$ -condition,  $\varphi(s+t) \sim \varphi(s) + \varphi(t)$  for any  $s, t \geq 0$ , we have

$$\int_{Q_\rho} \varphi\left(\frac{|u - \ell_\rho|}{r}\right) dz \leq \frac{\Delta_2(\varphi)}{2} \int_{Q_\rho} \varphi\left(\frac{|u - \ell|}{r}\right) dz + \frac{\Delta_2(\varphi)}{2} \int_{Q_\rho} \varphi\left(\frac{|\ell - \ell_\rho|}{r}\right) dz.$$

Hence, using (2.14), the fact that  $\varphi$  is increasing together with Jensen’s inequality, we can infer that

$$\int_{Q_r} \varphi\left(\frac{|\ell - \ell_\rho|}{r}\right) dz \lesssim \int_{Q_\rho} \varphi\left(\frac{|u - \ell|}{r}\right) dz.$$

□

An analogous reasoning leads to another basic inequality.

**Remark 2.5** For an  $N$ -function  $\varphi$  satisfying the  $\Delta_2$ -condition, we have

$$\int_{Q_\rho(z_0)} \varphi\left(\frac{|u - (u)_{z_0, \rho}|}{r}\right) dz \leq \Delta_2(\varphi) \int_{Q_\rho(z_0)} \varphi\left(\frac{|u - u_0|}{r}\right) dz,$$

for any  $u_0 \in \mathbb{R}^N$  and for any  $r > 0$ .

Finally, we can show that  $\ell_{z_0, \rho}$  is an almost minimizer of the functional  $\ell \mapsto \int_{Q_\rho(z_0)} \varphi_{1+|D\ell|} \left(\frac{|u - \ell|}{\rho}\right) dz$  amongst the affine functions  $\ell : \mathbb{R}^n \rightarrow \mathbb{R}^N$ .

**Lemma 2.7** *Let  $\varphi$  be an  $N$ -function satisfying  $\Delta_2(\varphi, \varphi^*) < +\infty$ , and let  $u \in L^\varphi(Q_\rho(z_0), \mathbb{R}^N)$ . There exists a constant  $\kappa_1 = \kappa_1(n, \Delta_2(\varphi, \varphi^*)) > 0$  such that*

$$\int_{Q_\rho(z_0)} \varphi_{1+|D\ell_{z_0, \rho}|} \left(\frac{|u - \ell_{z_0, \rho}|}{\rho}\right) dz \leq \kappa_1 \int_{Q_\rho(z_0)} \varphi_{1+|D\ell|} \left(\frac{|u - \ell|}{\rho}\right) dz,$$

for every affine function  $\ell : \mathbb{R}^n \rightarrow \mathbb{R}^N$ .

**Proof** From Lemma 2.1, Lemma 2.6, and Lemma 2.3 we obtain

$$\begin{aligned} \int_{Q_\rho(z_0)} \varphi_{1+|D\ell_{z_0, \rho}|} \left(\frac{|u - \ell_{z_0, \rho}|}{\rho}\right) dz &\leq \kappa_0 \int_{Q_\rho(z_0)} \varphi_{1+|D\ell_{z_0, \rho}|} \left(\frac{|u - \ell|}{\rho}\right) dz \\ &\leq c_\delta \int_{Q_\rho(z_0)} \varphi_{1+|D\ell|} \left(\frac{|u - \ell|}{\rho}\right) dz + \delta \varphi_{1+|D\ell|}(|D\ell - D\ell_{z_0, \rho}|), \end{aligned}$$

using also the fact that  $\varphi_{1+|a|}(|a - b|) \sim \varphi_{1+|b|}(|a - b|)$ . Moreover, from (2.13) we infer

$$|D\ell_{z_0, \rho} - D\ell| \leq (n+2) \int_{Q_\rho(z_0)} \frac{|u - \ell|}{\rho} dz.$$

Inserting this above, applying the  $\Delta_2$ -condition, and Jensen’s inequality conclude the proof. □

In the same way you obtain the following fact.

**Remark 2.6** For an  $N$ -function  $\varphi$  satisfying  $\Delta_2(\varphi, \varphi^*) < \infty$ , we have

$$\int_{\mathcal{Q}_\rho(z_0)} \varphi_{1+|(Du)_{z_0,\rho}|} (|Du - (Du)_{z_0,\rho}|) dz \leq \kappa_2 \int_{\mathcal{Q}_\rho(z_0)} \varphi_{1+|A|} (|Du - A|) dz,$$

for any  $A \in \mathbb{R}^{Nn}$ , where  $\kappa_2 = \kappa_2(n, \Delta_2(\varphi, \varphi^*)) > 0$ .

We conclude the section with an excess-decay-estimate for weak solutions to linear parabolic systems with constant coefficients [9, Lemma 5.1]. This can be achieved along the lines of the classical proof with very minor changes, so we will consider only the main points of the proof referring for the rest to [9].

**Lemma 2.8** ( *$\mathcal{A}$ -Caloric  $\psi$ -Excess Estimate*) *Suppose that  $h \in L^1(t_0 - R^2, t_0; W^{1,1}(\mathcal{B}_R(x_0); \mathbb{R}^N))$  is  $\mathcal{A}$ -caloric, and let  $\psi : [0, \infty) \rightarrow [0, \infty)$  be an increasing function. Then  $h \in C^\infty(\mathcal{Q}_R(z_0); \mathbb{R}^N)$  and the following excess estimate holds: for each  $0 < \rho < R$ ,  $\gamma > 0$ , and  $0 < \theta < 1/4$ , we have*

$$\int_{\mathcal{Q}_{\theta\rho}(z_0)} \psi \left( \left| \frac{\gamma(h - \ell_{z_0,\theta\rho}^{(h)})}{\theta\rho} \right| \right) dz \leq \psi \left( C\theta \int_{\mathcal{Q}_\rho(z_0)} \left| \frac{\gamma(h - \ell_{z_0,\rho}^{(h)})}{\rho} \right| dz \right)$$

where  $\ell_r^{(h)}(x) := (h)_{z_0,r} + (Dh)_{z_0,r}(x - x_0)$  and  $C$  depends on only  $n, N, L/\nu$ .

**Proof** It is only necessary to prove the estimate, since the smoothness of  $h$  is already contained in [9]. As argued in [4, Remark 3.2 and Lemma 3.3], we may use (5.9) and (5.12) in [9] to show that there exists  $C' = C'(n, L/\nu) < \infty$  such that

$$\sup_{\mathcal{Q}_{\rho/2}(z_0)} |D^2w| \leq C' \int_{\mathcal{Q}_\rho(z_0)} \left| \frac{w}{\rho^2} \right| dz \quad \text{and} \quad \sup_{\mathcal{Q}_{\rho/2}(z_0)} |D^3w| \leq C' \int_{\mathcal{Q}_\rho(z_0)} \left| \frac{w}{\rho^3} \right| dz, \tag{2.15}$$

for any  $\mathcal{A}$ -caloric map  $w \in C^\infty(\mathcal{Q}_R(z_0); \mathbb{R}^N)$ . Define  $w_r = h - \ell_{z_0,r}^{(h)}$ . Then  $w_r$  is  $\mathcal{A}$ -caloric and  $w_r \in C^\infty(\mathcal{Q}_R(z_0); \mathbb{R}^N)$ , for each  $0 < r < R$ . Let  $0 < \theta < \frac{1}{4}$  and  $0 < \rho \leq R$  be given. Using (2.15), the fact that  $\mathcal{Q}_r(z_0)$  is a standard parabolic cylinder, and the fact that every derivative of  $h$  is still  $\mathcal{A}$ -caloric, for each  $(x, t) \in \mathcal{Q}_{\theta\rho}(z_0)$ , we have

$$\begin{aligned} |w_{\theta\rho}(x, t)| &\leq \theta\rho \sup_{\mathcal{Q}_{\theta\rho}(z_0)} |Dh - (Dh)_{z_0,\theta\rho}| + \theta^2\rho^2 \sup_{\mathcal{Q}_{\theta\rho}(z_0)} |\partial_t h| \\ &\leq \theta^2\rho^2 \left( \sup_{\mathcal{Q}_{\theta\rho}(z_0)} |D^2h| + \theta\rho \sup_{\mathcal{Q}_{\theta\rho}(z_0)} |\partial_t Dh| + \sup_{\mathcal{Q}_{\theta\rho}(z_0)} |\partial_t h| \right) \\ &\leq C''\theta^2\rho^2 \left( \sup_{\mathcal{Q}_{\rho/2}(z_0)} |D^2w_\rho| + \theta\rho \sup_{\mathcal{Q}_{\rho/2}(z_0)} |D^3w_\rho| \right) \\ &\leq C\theta^2 \int_{\mathcal{Q}_\rho(z_0)} |w_\rho| dz. \end{aligned}$$

Here,  $C''$  and  $C$  depend on only  $n, N$ , and  $L/\nu$ . The result follows from this and the definition of  $w_\rho$ . □

### 3 $\mathcal{A}$ -caloric approximation

To prove the partial regularity for non-degenerate parabolic systems with  $\varphi$ -growth, we shall compare the solution of our parabolic system with the solution of a linear parabolic system with constant coefficients. The comparison will be achieved by a generalization of the  $\mathcal{A}$ -caloric approximation lemma in Orlicz spaces. We emphasize that the approximation lemma requires no upper bound on the growth of  $\varphi$ .

Recall that a function  $f : [0, +\infty) \rightarrow [0, +\infty)$  is said to be almost increasing if there exists  $\Lambda \geq 1$  such that  $f(t) \leq \Lambda f(s)$  for every  $0 \leq t < s < +\infty$ . We will consider the following assumptions for the  $N$ -function  $\varphi$ , more general with respect to Assumption 1.1:

- (H1) There exists a  $p_0 > \frac{2n}{n+2}$  such that  $\frac{\varphi(t)}{t^{p_0}}$  is almost increasing,
- (H2)  $\varphi$  has a uniform doubling property near zero; i.e.  $\limsup_{t \rightarrow 0^+} \frac{\varphi(2t)}{\varphi(t)} = \Delta_0(\varphi) < \infty$ .

In general, an  $N$ -function might not satisfy assumption (H2). For example, with

$$\ell_k(t) = \frac{1}{k!} + \frac{2^k(k-1)}{k!}(t - 2^{-k}), \quad \text{for } k \in \mathbb{N},$$

the  $N$ -function

$$\varphi(t) = \begin{cases} 0, & t = 0, \\ \ell_{k+1}(t), & k \in \mathbb{N} \setminus \{1\} \text{ and } 2^{-k-1} \leq t < 2^{-k} \\ 8t^2, & 2^{-2} \leq t, \end{cases}$$

is not uniformly doubling near zero since  $(k+1)\ell_{k+1}(2^{-k-1}) = \ell_k(2^{-k})$ . For any  $N$ -function and  $a > 0$ , however, (H2) is satisfied by the shifted function  $\varphi_a(t)$ . In fact,

$$\frac{\varphi_a(2t)}{\varphi_a(t)} \leq \frac{4\varphi'(2a)}{\varphi'(a)}, \quad \text{for all } 0 < t \leq \frac{a}{2}.$$

#### 3.1 Additional notation and supporting results

For this section, we introduce some additional notation. There are also several supporting results used in the proof of the approximation lemma.

First, we require a compactness principle of Simon.

**Theorem 3.1** ([42, Theorem 6]) *Suppose that  $X \subseteq B \subseteq Y$  are Banach spaces with a compact embedding  $X \rightarrow B$ . Given  $1 < q \leq \infty$ , assume*

- $F$  is bounded in  $L^q(0, T; B) \cap L^1_{loc}(0, T; X)$ ,
- for all  $0 < t_1 < t_2 < T$ ,  $\|f(\cdot + h) - f(\cdot)\|_{L^p(t_1, t_2; Y)} \rightarrow 0$  as  $h \rightarrow 0$ , uniformly for  $f \in F$ .

Then  $F$  is relatively compact in  $L^p(0, T; B)$  for all  $1 \leq p < q$ .

We also need to work with the Orlicz norm: given a measurable  $E \subseteq \mathbb{R}^n$ ,

$$\|f\|_{L^*_\varphi(E)} = \sup \left\{ \int_E f(y)g(y)dy : \int_E \varphi^*(|g(y)|)dy \leq 1 \right\}.$$

It can be verified [33] that the Orlicz space

$$L^*_\varphi(E) = \{f \in L^1(E) : \|f\|_{L^*_\varphi(E)} < \infty\}$$

is a Banach space. The Orlicz norm is equivalent to the Luxemborg norm ([33], p. 80). Moreover, as established in [33, Lemma 9.2 and p. 75], given  $\{f_k\}_{k=1}^\infty \subseteq L_\varphi^*(E)$  and  $f \in L_\varphi^*(E)$ ,

$$\lim_{k \rightarrow \infty} \|f_k - f\|_{L_\varphi^*(E)} = 0 \implies \lim_{k \rightarrow \infty} \int_E \varphi(\lambda|f_k(x) - f(x)|)dx = 0, \quad \text{for all } \lambda > 0. \tag{3.1}$$

We will also need the following

**Definition 3.1** Given an open set  $E \subseteq \mathbb{R}^n$  and  $f \in L^1_{loc}(E)$ , the (non-centered) Hardy-Littlewood maximal operator is  $M(f) : E \rightarrow [0, \infty]$

$$Mf(x) = \sup_{\mathcal{B} \ni x} \int_{\mathcal{B} \cap E} |f(y)|dy.$$

Here the supremum is taken over all balls containing  $x$ .

It is well-known that the maximal operator is bounded on  $L^p$ , for  $p > 1$ . From [30, Corollary 4.3.3], we have

**Corollary 3.1** Let an open set  $E \subseteq \mathbb{R}^n$  and an  $N$ -function  $\varphi$  be given. If  $p > 1$  and  $\frac{\varphi(t)}{t^p}$  is almost increasing, then there exists a  $\beta > 0$  such that

$$\varphi(\beta Mf(x))^{\frac{1}{p}} \lesssim M\left(\varphi(f)^{\frac{1}{p}}\right)(x)$$

for every ball  $\mathcal{B}$ ,  $x \in \mathcal{B} \cap E$ , and  $f \in L^\varphi(E)$  satisfying  $\int_E \varphi(f)dx \leq 1$ .

Finally, as explained in the proof of the  $\mathcal{A}$ -caloric excess estimate (Lemma 2.8), we may use the regularity provided in (5.9) and (5.12) in [9] to show there is a  $C = C(L/\nu) < \infty$  such that

$$\sup_{\mathcal{Q}_{\tau R}(z_0)} (|Dw|^2 + |w|^2) \leq C \int_{\mathcal{Q}_R(z_0)} |w|^2 dz, \tag{3.2}$$

for any  $\frac{3}{4} \leq R \leq 1$ ,  $\frac{1}{2} \leq \tau \leq \frac{3}{4}$ , and  $\mathcal{A}$ -caloric map  $w \in L^1(t_0 - R^2, 0; W^{1,2}(\mathcal{B}_R(x_0); \mathbb{R}^N))$ .

### 3.2 The $\mathcal{A}$ -Caloric approximation lemma

With the preliminaries above, we can state and prove the main result for this section.

**Theorem 3.2** Suppose that (H1) and (H2) are satisfied. Let  $\varepsilon, \nu > 0$  and  $\nu < L < \infty$  be given. There exists  $\delta_0 \leq 1$ ,  $\delta_0$  depending on  $n, N, p_0, \Delta_0(\varphi), \nu, L, \varepsilon$ , and  $1 \leq K_\varphi$  with the following property: for any  $\gamma \in (0, 1/\sqrt{\omega_n}]$  (with  $\omega_n$  being the measure of the unit sphere in  $\mathbb{R}^n$ ), any bilinear form  $\mathcal{A}$  satisfying

$$\mathcal{A}(\xi, \xi) \geq \nu|\xi|^2 \quad \text{and} \quad |\mathcal{A}(\xi, \eta)| \leq L|\xi||\eta|,$$

and any approximately  $\mathcal{A}$ -caloric map  $v \in L^\infty(t_0 - \rho^2, t_0; L^2(\mathcal{B}_\rho(x_0); \mathbb{R}^N)) \cap L^1(t_0 - \rho^2, t_0; W^{1,1}(\mathcal{B}_\rho(x_0); \mathbb{R}^N))$  satisfying:

- $\varphi(|Dv|) \in L^1(t_0 - \rho^2, t_0; L^1(\mathcal{B}_\rho(x_0)))$ ,

- for some  $0 < \delta \leq \delta_0$ ,

$$\left| \int_{\mathcal{Q}_\rho(z_0)} (v \cdot \partial_t \eta - \mathcal{A}(Dv, D\eta)) dz \right| \leq \delta \sup_{\mathcal{Q}_\rho(z_0)} |D\eta|, \text{ for all } \eta \in C_c^\infty(\mathcal{Q}_\rho(z_0); \mathbb{R}^N), \tag{3.3}$$

- and

$$\sup_{t_0 - \rho^2 < t < t_0} \int_{B_\rho} \left| \frac{v}{\rho} \right|^2 dx + \int_{\mathcal{Q}_\rho(z_0)} \left( \varphi \left( \left| \frac{v}{\rho} \right| \right) + \varphi(|Dv|) \right) dz \leq \gamma^2, \tag{3.4}$$

then there exists an  $\mathcal{A}$ -caloric map  $h \in L^2(t_0 - \rho^2/4, t_0; W^{1,2}(\mathcal{B}_{\rho/2}(x_0); \mathbb{R}^N))$  such that

$$\int_{\mathcal{Q}_{\rho/2}(z_0)} \left( \left| \frac{\gamma h}{\rho/2} \right|^2 + \varphi \left( \left| \frac{\gamma h}{\rho/2} \right| \right) + \varphi(|\gamma Dh|) \right) dz \leq 2^{n+2} K_\varphi \gamma^2,$$

and

$$\int_{\mathcal{Q}_{\rho/2}(z_0)} \left( \left| \frac{v - \gamma h}{\rho/2} \right|^2 + \varphi \left( \left| \frac{v - \gamma h}{\rho/2} \right| \right) \right) dz \leq \varepsilon \gamma^2.$$

The constant  $K_\varphi$  is defined in (3.12) and depends only on  $\varphi$  and  $C(L/v)$  in (3.2).

**Proof** We translate and rescale to  $z_0 = (0, 0)$  and  $\rho = 1$ . Assuming the alternative, there exists an  $\varepsilon_0 > 0$ , sequences  $\gamma_k \in (0, 1/\sqrt{\omega_n}]$ , bilinear forms  $\mathcal{A}_k$ , and maps  $v_k \in L^\infty(-1, 0; L^2(\mathcal{B}_1; \mathbb{R}^N)) \cap L^1(-1, 0; W^{1,1}(\mathcal{B}_1; \mathbb{R}^N))$ , such that for each  $k \in \mathbb{N}$ , the following holds:

- (i)  $\varphi(|Dv_k|) \in L^1(-1, 0; L^1(\mathcal{B}_1))$ ,
- (ii)  $\int_{\mathcal{Q}_1} (v_k \cdot \partial_t \eta - \mathcal{A}_k(Dv_k, D\eta)) dz \leq \frac{1}{k} \sup_{\mathcal{Q}_1} |D\eta|$ , for all  $\eta \in C_0^\infty(\mathcal{Q}_1; \mathbb{R}^N)$ ,
- (iii)  $\sup_{t \in (-1, 0)} \int_{\mathcal{B}_1} |v_k|^2 dx + \int_{\mathcal{Q}_1} (\varphi(|v_k|) + \varphi(|Dv_k|)) dz \leq \gamma_k^2$ ,
- (iv) for any  $\mathcal{A}_k$ -caloric map  $h \in C^\infty(\mathcal{Q}_{1/2}; \mathbb{R}^N)$  satisfying

$$\int_{\mathcal{Q}_{1/2}} (4|\gamma_k h|^2 + \varphi(2|\gamma_k h|) + \varphi(|\gamma_k Dh|)) dz \leq 2^{n+2} K_\varphi \gamma_k^2,$$

we find

$$\int_{\mathcal{Q}_{1/2}} (4|v_k - \gamma_k h|^2 + \varphi(2|v_k - \gamma_k h|)) dz > \varepsilon_0 \gamma_k^2.$$

By (iii) the sequence  $\{\varphi(|Dv_k|)\}_{k=1}^\infty$  is bounded in  $L^1(-1, 0; L^1(\mathcal{B}_1; \mathbb{R}^N))$ . Assumption (H1) implies the existence of a constant  $C$  such that

$$\|v_k\|_{L^{p_0}(\mathcal{Q}_1)} + \|Dv_k\|_{L^{p_0}(\mathcal{Q}_1)} \leq C.$$

It follows that, for a non-re-labeled sequence, there exists a bilinear form  $\mathcal{A}$  and a map  $v \in L^2(\mathcal{Q}_1) \cap L^{p_0}(\mathcal{Q}_1)$  such that  $Dv \in L^{p_0}(\mathcal{Q}_1)$  and

$$\begin{cases} v_k \rightharpoonup v & \text{in } L^2(\mathcal{Q}_1; \mathbb{R}^N), \\ Dv_k \rightharpoonup Dv & \text{in } L^{p_0}(\mathcal{Q}_1; \mathbb{R}^{Nn}), \\ \mathcal{A}_k \rightarrow \mathcal{A} & \text{in bilinear forms on } \mathbb{R}^{Nn}, \\ \gamma_k \rightarrow \gamma \in [0, 1/\sqrt{\omega_n}]. \end{cases}$$

The convexity of  $\varphi$  implies

$$\int_{\mathcal{Q}_1} (|v|^2 + \varphi(|v|) + \varphi(|Dv|)) dz \leq \gamma^2. \tag{3.5}$$

Moreover, with the same argument used in [41], we conclude that  $v$  is  $\mathcal{A}$ -caloric and  $v \in C^\infty(\mathcal{Q}_1; \mathbb{R}^N)$ . From (3.2),

$$\sup_{\mathcal{Q}_{3/4}} (|v|^2 + |Dv|^2) \leq C(L/\nu) \int_{\mathcal{Q}_1} |v|^2 dz \leq C(L/\nu)\gamma^2. \tag{3.6}$$

As demonstrated in [41], given  $\ell > \frac{n+2}{2}$ , for any  $-1 < t_1 < t_2 < -h$ ,

$$\begin{aligned} \|v_k(\cdot, s) - v_k(\cdot, s+h)\|_{W^{-\ell,2}(\mathcal{B}_1)} &\leq C \left( h^{\frac{p_0-1}{p_0}} + \frac{1}{k} \right), \quad \text{for } s \in (t_1, t_2) \\ \implies \int_{t_1}^{t_2} \|v_k(\cdot, s) - v_k(\cdot, s+h)\|_{W^{-\ell,2}(\mathcal{B}_1)}^p ds &\leq C \left( h^{\frac{p(p_0-1)}{p_0}} + \frac{1}{k^p} \right), \end{aligned}$$

for any  $p \geq 1$ .

Moreover, the sequence  $\{v_k\}_{k=1}^\infty$  is uniformly bounded in  $L^\infty(-1, 0; L^2(\mathcal{B}_1; \mathbb{R}^N)) \cap L^1_{loc}(-1, 0; W^{1,p_0}(\mathcal{B}_1; \mathbb{R}^N))$ . With  $X = W^{1,p_0}(\mathcal{B}_1; \mathbb{R}^N)$ ,  $B = L^2(\mathcal{B}_1; \mathbb{R}^N)$ , and  $Y = W^{-\ell,2}(\mathcal{B}_1; \mathbb{R}^N)$ , Theorem 3.1 yields the strong convergence (for a non-relabeled subsequence)

$$v_k \rightarrow v \quad \text{in } L^p(-1, 0; L^2(\mathcal{B}_1; \mathbb{R}^N)), \tag{3.7}$$

for any  $p \geq 1$ . This entails also that  $v_k \rightarrow v$  in  $L^p(-1, 0; L^1(\mathcal{B}_1; \mathbb{R}^N))$  for any  $p \geq 1$ , so  $\|v_k - v\|_{L^1(\mathcal{B}_1; \mathbb{R}^n)} \rightarrow 0$  for almost every  $-1 < t < 0$ .

**Claim:**  $\lim_{k \rightarrow \infty} \int_{\mathcal{Q}_{3/4}} \varphi(\lambda|v_k - v|) dz = 0$ , for all  $\lambda > 0$ .

First, we observe that  $v_k(\cdot, t)$  and  $v(\cdot, t)$  belong to  $W^{1,1}(\mathcal{B}_1, \mathbb{R}^N)$ , for almost every time  $t \in (-1, 0)$ . We may therefore extend them to the whole of  $\mathbb{R}^n$  in such a way that their extensions  $\tilde{v}_k(\cdot, t)$  and  $\tilde{v}(\cdot, t)$  belong to  $W^{1,1}(\mathbb{R}^n, \mathbb{R}^N)$ , and

$$\|\tilde{v}_k(\cdot, t) - \tilde{v}(\cdot, t)\|_{W^{1,1}(\mathbb{R}^n, \mathbb{R}^N)} \leq C \|v_k(\cdot, t) - v(\cdot, t)\|_{W^{1,1}(\mathcal{B}_1, \mathbb{R}^N)}, \quad \text{for a.e. } t \in (-1, 0),$$

where the constant  $C$  depends only on  $\mathcal{B}_1$ . Let  $\sigma$  be the standard mollifier and  $\sigma_\varepsilon(x) = \frac{1}{\varepsilon^n} \sigma\left(\frac{|x|}{\varepsilon}\right)$ .

On account of (3.1), it is enough to verify

$$\lim_{k \rightarrow \infty} \sup_g \int_{\mathcal{Q}_{3/4}} |v_k - v| g dz = 0$$

where the supremum is taken over all  $g \in L^1(\mathcal{Q}_{3/4})$  such that  $\|\varphi^*(|g|)\|_{L^1(\mathcal{Q}_{3/4})} \leq 1$ . Fix  $g \in L^1(\mathcal{Q}_{3/4})$  satisfying  $\|\varphi^*(|g|)\|_{L^1(\mathcal{Q}_{3/4})} \leq 1$ . With  $0 < \varepsilon < 1$  given, for each  $-1 < t < 0$ , we let  $\tilde{v}_k * \sigma_\varepsilon$  and  $\tilde{v} * \sigma_\varepsilon$  denote the mollifications of the extended maps  $\tilde{v}_k$  and  $\tilde{v}$  in the spatial direction. We have

$$\begin{aligned} &\left| \int_{\mathcal{Q}_{3/4}} |v_k(z) - v(z)| g(z) dz \right| \\ &\leq \int_{\mathcal{Q}_{3/4}} |\tilde{v}_k(x, t) - \tilde{v}(x, t)| |g(x, t)| dx dt \end{aligned}$$



$$\begin{aligned}
 &\leq \left[ \int_{\mathcal{Q}_{3/4}} |(\tilde{v}_k * \sigma_\varepsilon)(x, t) - \tilde{v}_k(x, t)| |g(x, t)| dx dt \right. \\
 &\quad + \int_{\mathcal{Q}_{3/4}} |(\tilde{v}_k - \tilde{v}) * \sigma_\varepsilon(x, t)| |g(x, t)| dx dt \\
 &\quad \left. + \int_{\mathcal{Q}_{3/4}} |(\tilde{v} * \sigma_\varepsilon)(x, t) - v(x, t)| |g(x, t)| dx dt \right] \\
 &=: [I_{1,k} + I_{2,k} + I_{3,k}]. \tag{3.8}
 \end{aligned}$$

First, we examine  $I_{2,k}$ . Given  $-1 < t < 0$ , we use Young’s convolution inequality to write

$$\begin{aligned}
 \sup_{k \in \mathbb{N}} \|(\tilde{v}_k - \tilde{v}) * \sigma_\varepsilon(\cdot, t)\|_{L^\infty(\mathcal{B}_{3/4})} &\leq \sup_{k \in \mathbb{N}} \|\tilde{v}_k(\cdot, t) - \tilde{v}(\cdot, t)\|_{L^1(\mathcal{B}_{3/4}; \mathbb{R}^N)} \|\sigma_\varepsilon\|_{L^\infty(\mathcal{B}_{3/4})} \\
 &\leq \sup_{k \in \mathbb{N}} \left( \|\tilde{v}_k(\cdot, t)\|_{L^1(\mathcal{B}_{3/4}; \mathbb{R}^N)} + \|\tilde{v}(\cdot, t)\|_{L^1(\mathcal{B}_{3/4}; \mathbb{R}^N)} \right) \|\sigma_\varepsilon\|_{L^\infty(\mathcal{B}_{3/4})} \\
 &\leq C \sup_{k \in \mathbb{N}} \left( \|v_k(\cdot, t)\|_{L^2(\mathcal{B}_1; \mathbb{R}^N)} + \|v(\cdot, t)\|_{L^2(\mathcal{B}_1; \mathbb{R}^N)} \right) \|\sigma_\varepsilon\|_{L^\infty(\mathbb{R}^n)} \\
 &\leq C \|\sigma_\varepsilon\|_{L^\infty(\mathbb{R}^n)}.
 \end{aligned} \tag{3.9}$$

For the last two inequalities, we used (iii). On the other hand, by Young’s inequality and by the convexity of  $\varphi^*$ , for any  $\alpha \in (0, 1)$ , we have

$$\begin{aligned}
 \int_{\mathcal{Q}_{3/4}} |(\tilde{v}_k - \tilde{v}) * \sigma_\varepsilon(x, t)| |g(x, t)| dx dt &\leq \int_{\mathcal{Q}_{3/4}} \varphi \left( \frac{|(\tilde{v}_k - \tilde{v}) * \sigma_\varepsilon(x, t)|}{\alpha} \right) dx dt + \alpha \int_{\mathcal{Q}_{3/4}} \varphi^*(|g|) dx dt \\
 &\leq \int_{\mathcal{Q}_{3/4}} \varphi \left( \frac{|(\tilde{v}_k - \tilde{v}) * \sigma_\varepsilon(x, t)|}{\alpha} \right) dx dt + \alpha.
 \end{aligned}$$

Since

$$\lim_{k \rightarrow \infty} |(\tilde{v}_k - \tilde{v}) * \sigma_\varepsilon(x, t)| \rightarrow 0 \quad \text{for a.e. } (x, t) \in \mathcal{Q}_{3/4},$$

thanks to (3.9) we may use the dominated convergence theorem to conclude that

$$\lim_{k \rightarrow \infty} \sup_g I_{2,k} \leq \alpha.$$

Now, we turn to bounding  $I_{1,k}$  and  $I_{3,k}$ . The arguments for each term are similar, so we focus on  $I_{1,k}$ . We will use some results contained in [30]. As provided in [30] (p. 135), the following pointwise estimate holds for functions in  $W_{loc}^{1,1}(\mathbb{R}^n, \mathbb{R}^N)$ :

$$|\tilde{v}_k(x, \cdot) - (\tilde{v}_k * \sigma_\varepsilon)(x, \cdot)| \leq \varepsilon \int_0^1 (|D\tilde{v}_k| * \sigma_{\varepsilon\tau})(x, \cdot) d\tau,$$

almost everywhere in  $\mathcal{B}_{3/4}$ . Let  $\beta > 0$  be the constant from Corollary 3.1. Combining the bound above with [30, Lemma 4.4.6], for each  $-1 < t < 0$ , we obtain

$$|\tilde{v}_k(x, t) - (\tilde{v}_k * \sigma_\varepsilon)(x, t)| \leq \frac{2\varepsilon}{\beta} M(\beta |D\tilde{v}_k|(\cdot, t))(x), \quad \text{for a.e. } x \in \mathcal{B}_{3/4}.$$

Here  $M$  is the (non-centered Hardy-Littlewood) maximal function defined earlier. Observing that, for almost every  $t \in (-1, 0)$ , we have

$$\begin{aligned}
 M(\beta |D\tilde{v}_k(\cdot, t)|)(x) &= \sup_{\mathcal{B} \ni x} \frac{\beta}{|\mathcal{B}|} \int_{\mathcal{B} \cap \mathcal{B}_1} |D\tilde{v}_k(y, t)| dy \\
 &\leq \sup_{\mathcal{Q} \ni (x,t)} \frac{\beta}{|\mathcal{Q}|} \int_{\mathcal{Q} \cap \mathcal{Q}_1} |D\tilde{v}_k(y, s)| dy ds = M(\beta |D\tilde{v}_k|)(x, t),
 \end{aligned}$$

we deduce that

$$|\tilde{v}_k(z) - (\tilde{v}_k * \sigma_\varepsilon)(z)| \leq \frac{2\varepsilon}{\beta} M(\beta|D\tilde{v}_k|)(z), \quad \text{for a.e. } z \in Q_{3/4}.$$

Incorporating this into the definition of  $I_{1,k}$  and applying Young’s inequality, we may write

$$\begin{aligned} I_{1,k} &\leq \frac{2\varepsilon}{\beta} \int_{Q_{3/4}} M(\beta|D\tilde{v}_k|)(z)|g(z)|dz \\ &\leq \frac{2\varepsilon}{\beta} \left[ \int_{Q_{3/4}} \varphi(M(\beta|D\tilde{v}_k|)(z))dz + \int_{Q_{3/4}} \varphi^*(|g(z)|)dz \right] \\ &\leq \frac{2\varepsilon}{\beta} \left[ \int_{Q_{3/4}} \varphi(M(\beta|D\tilde{v}_k|)(z))dz + 1 \right], \end{aligned}$$

where we have used  $\|\varphi^*(|g|)\|_{L^1(Q_{3/4})} \leq 1$  in the last inequality. Recalling that  $\int_{Q_1} \varphi(|Dv_k|) dz \leq 1$  and that  $\frac{\varphi(t)}{t^{p_0}}$  is almost increasing, we can use Corollary 3.1 to infer

$$\varphi(\beta M(|Dv_k|)(z))^{\frac{1}{p_0}} \lesssim M\left(\varphi(|Dv_k|)^{\frac{1}{p_0}}\right)(z),$$

almost everywhere in  $Q_1$ . Thus

$$\int_{Q_{3/4}} \varphi(\beta M(|Dv_k|)) dz \lesssim \int_{Q_1} \left( M\left(\varphi(|Dv_k|)^{\frac{1}{p_0}}\right) \right)^{p_0} dz \lesssim \int_{Q_1} \varphi(|Dv_k|) dz \leq 1,$$

since  $M$  is bounded in  $L^{p_0}$ . We conclude that  $I_{1,k} \lesssim \varepsilon/\beta$ . A similar argument shows  $I_{3,k} \lesssim \varepsilon/\beta$ , as well.

Returning to (3.8), we have shown

$$\limsup_{k \rightarrow \infty} \int_{Q_{3/4}} |v_k(z) - v(z)|g(z)dz \lesssim \frac{\varepsilon}{\beta} + \alpha,$$

with the supremum being taken over all  $g \in L^1(Q_{3/4})$  satisfying  $\|\varphi^*(|g|)\|_{L^1(Q_{3/4})} \leq 1$ . Here  $\beta$  is independent of  $k$ . Since  $\varepsilon, \alpha > 0$  were both arbitrary, the claim is proved.

Next, we produce a sequence  $\{h_k\}_{k=1}^\infty \in C^\infty(Q_{1/2}; \mathbb{R}^N)$  of  $\mathcal{A}$ -caloric maps that will contradict (iv) for  $k$  sufficiently large.

Case 1:  $\gamma_k \rightarrow 0$ : In this case, clearly  $v = 0$ ,  $v_k \rightarrow 0$  strongly in  $L^2(Q_1)$  and  $\int_{Q_1} \varphi(2|v_k|)dz \rightarrow 0$ . Thus, we obtain a contradiction to (iv) with  $h \equiv 0$ .

Case 2:  $\gamma_k \rightarrow \gamma \in (0, 1/\sqrt{\omega_n})$ : For each  $k \in \mathbb{N}$ , let  $h_k$  be the unique solution to

$$\begin{cases} \int_{Q_{3/4}} (h_k \cdot \partial_t \eta - \mathcal{A}_k(Dh_k, D\eta)) dz = 0 & \text{for all } \eta \in C_c^\infty(Q_{3/4}; \mathbb{R}^N) \\ h_k = \gamma_k^{-1}v & \text{on } \partial_{\mathcal{P}} Q_{3/4}. \end{cases}$$

Since  $v \in C^\infty(\overline{Q}_{3/4}; \mathbb{R}^N)$  and  $\gamma_k \rightarrow \gamma > 0$ , so is each  $h_k$ . As shown in [5] and [41], we have

$$\lim_{k \rightarrow \infty} \int_{Q_{3/4}} (|v - \gamma_k h_k|^2 + |Dv - \gamma_k Dh_k|^2) dz = 0.$$

Thus

$$\gamma_k h_k \rightarrow v \text{ and } \gamma_k Dh_k \rightarrow Dv \text{ in } L^2(Q_{3/4}). \tag{3.10}$$

This implies convergence in measure for both sequences. Moreover, after taking a non-labeled subsequence if necessary, the bounds in (3.2) and (3.6) imply

$$\sup_{Q_{1/2}} (|\gamma_k h_k| + |\gamma_k Dh_k|) \leq 2^{n+1} C(L/v).$$

Thus, since  $\sup_{k \in \mathbb{N}} \|\gamma_k h_k - v\|_{L^\infty(Q_{1/2})} + \|\gamma_k Dh_k - Dv\|_{L^\infty(Q_{1/2})} \leq 2^{n+3} C(L/v)$ ,

$$\lim_{k \rightarrow \infty} \int_{Q_{1/2}} (\varphi(\lambda|\gamma_k h_k - v|) + \varphi(\lambda|\gamma_k Dh_k - Dv|)) dz = 0, \text{ for all } \lambda > 0. \tag{3.11}$$

To finish the proof, define

$$K_\varphi = 8 + \sup \left\{ \frac{\varphi(4t)}{\varphi(t)} : 0 < t \leq 2^{n+2} C(L/v) \right\}. \tag{3.12}$$

Note that  $K_\varphi$  must be finite due to assumption (H2). Using (3.5) and the convexity of  $\varphi$ ,

$$\begin{aligned} & \lim_{k \rightarrow \infty} \int_{Q_{1/2}} (4|\gamma_k h_k|^2 + \varphi(2|\gamma_k h_k|) + \varphi(|\gamma_k Dh_k|)) dz \\ & \leq \frac{1}{2} \lim_{k \rightarrow \infty} \int_{Q_{1/2}} (16|\gamma_k h_k - v|^2 + \varphi(4|\gamma_k h_k - v|) + \varphi(2|\gamma_k Dh_k - Dv|)) dz \\ & \quad + \frac{1}{2} \int_{Q_{1/2}} (16|v|^2 + \varphi(4|v|) + \varphi(2|Dv|)) dz \\ & \leq 2^{n+1} K_\varphi \int_{Q_1} (|v|^2 + \varphi(|v|) + \varphi(|Dv|)) dz \\ & \leq 2^{n+1} K_\varphi \gamma^2. \end{aligned}$$

Since  $\gamma > 0$ , for  $k$  sufficiently large, we have  $\gamma < 2\gamma_k$ , and the  $\mathcal{A}_k$ -caloric map  $h_k$  satisfies

$$\int_{Q_{1/2}} (4|\gamma_k h_k|^2 + \varphi(2|\gamma_k h_k|) + \varphi(|\gamma_k Dh_k|)) dz < 2^{n+2} K_\varphi \gamma_k^2.$$

Similarly, using the convergences provided by the claim, (3.7), (3.10), and (3.11) we conclude that

$$\lim_{k \rightarrow \infty} \int_{Q_{1/2}} (4|v_k - \gamma_k h_k|^2 + \varphi(2|v_k - \gamma_k h_k|)) dz = 0,$$

which provides the contradiction to (iv). □

**Remark 3.1** For the shifted function  $\varphi_a$ ,

$$K_{\varphi_a} \leq 8 + \max \left\{ \left( \frac{4\varphi'(2a)}{\varphi'(a)} \right)^2, \frac{\varphi_a(2^{n+4} C(L/v))}{\varphi_a(a/4)} \right\}.$$

If the function  $\varphi$  is doubling, then  $K_\varphi = 8 + \Delta_2(\varphi)^2$ .

### 4 Caccioppoli type inequality

Let us prove the following Caccioppoli inequality for standard parabolic cylinders.

**Theorem 4.1** *Let  $u \in C^0(-T, 0; L^2(\Omega, \mathbb{R}^N)) \cap L^1(-T, 0; W^{1,1}(\Omega, \mathbb{R}^N))$  be a weak solution to (1.1) satisfying  $\varphi(|Du|) \in L^1(-T, 0; L^1(\Omega))$ . Under hypotheses (a<sub>1</sub>)-(a<sub>4</sub>) and Assumption 1.1, given a standard cylinder  $\mathcal{Q}_R(z_0) \subset \Omega_T$ , with center  $z_0 = (x_0, t_0)$ , and any affine map  $\ell : \mathbb{R}^n \rightarrow \mathbb{R}^N$  and  $0 < r < R$ , we have*

$$\begin{aligned} & \sup_{t_0-r^2 < s < t_0} \int_{\mathcal{B}_r(x_0)} |u(x, s) - \ell(x)|^2 dx + \int_{\mathcal{Q}_r(z_0)} \varphi_{1+|D\ell|}(|Du - D\ell|) dz \\ & \leq c_0 \int_{\mathcal{Q}_R(z_0)} \left[ \varphi_{1+|D\ell|} \left( \left| \frac{u - \ell}{R - r} \right| \right) + \left| \frac{u - \ell}{R - r} \right|^2 \right] dz, \end{aligned}$$

where  $c_0$  depends only on  $n, N, L, \nu, p_0, p_1$ .

**Proof** For notational brevity, we put  $M = 1 + |D\ell|$ . Without loss of generality we may assume  $z_0 = (0, 0)$ . For a generic radius  $\rho$ , we denote  $\mathcal{Q}_\rho(z_0) = \mathcal{Q}_\rho$  and  $\mathcal{B}_\rho(x_0) = \mathcal{B}_\rho$ . Let us consider the function  $\eta(x, t) = \chi^{p_1}(x)\zeta^2(t)(u(x, t) - \ell(x))$  as a test function in (1.1), where  $\chi$  is a standard cutoff function between  $\mathcal{B}_r$  and  $\mathcal{B}_R$ , and  $\zeta \in C^0(\mathbb{R})$  is defined by

$$\begin{cases} \zeta(t) = 0, & t \in (-\infty, -R^2) \\ \zeta_t(t) = \frac{1}{R^2 - r^2}, & t \in (-R^2, -r^2) \\ \zeta(t) = 1, & t \in (-r^2, s) \\ \zeta_t(t) = -\frac{1}{\varepsilon}, & t \in (s, s + \varepsilon) \\ \zeta(t) = 0, & t \in (s + \varepsilon, +\infty) \end{cases}$$

for  $-r^2 < s < 0$  and  $0 < \varepsilon \leq |s|$ . We have

$$\begin{aligned} & \int_{\mathcal{Q}_R} \chi^{p_1} \zeta^2 a(Du) \cdot (Du - D\ell) dz \\ & = -p_1 \int_{\mathcal{Q}_R} \chi^{p_1-1} \zeta^2 a(Du) \cdot [D\chi \otimes (u - \ell)] dz + \int_{\mathcal{Q}_R} u \cdot \eta_t dz. \end{aligned}$$

Noting that  $\int_{\mathcal{Q}_R} a(D\ell) \cdot D\eta dz = 0$  and  $\int_{\mathcal{Q}_R} \ell \cdot \eta_t dz = 0$ , we obtain

$$\begin{aligned} I & := \int_{\mathcal{Q}_R} \chi^{p_1} \zeta^2 (a(Du) - a(D\ell)) \cdot (Du - D\ell) dz \\ & = -p_1 \int_{\mathcal{Q}_R} \chi^{p_1-1} \zeta^2 (a(Du) - a(D\ell)) \cdot [D\chi \otimes (u - \ell)] dz \\ & \quad + \int_{\mathcal{Q}_R} (u - \ell) \cdot \eta_t dz =: II + III \end{aligned} \tag{4.1}$$

The left hand side can be estimated thanks to Remark 2.4, leading to

$$I \geq c \int_{\mathcal{Q}_R} \chi^{p_1} \zeta^2 \varphi_M(|Du - D\ell|) dz \tag{4.2}$$

and

$$|II| \leq c \int_{\mathcal{Q}_R} \chi^{p_1-1} \zeta^2 \varphi'_M(|Du - D\ell|) |D\chi| |u - \ell| dz.$$

Using Young’s inequality and (2.3) together with (2.5), we derive the following bound for  $II$ :

$$\begin{aligned}
 |II| &\leq c \delta \int_{\mathcal{Q}_R} \zeta^2 \varphi_M^* (\varphi'_M (|Du - D\ell|) \chi^{p_1-1}) dz + c(\delta) \int_{\mathcal{Q}_R} \zeta^2 \varphi_M (|u - \ell| |D\chi|) dz \\
 &\leq c \delta \int_{\mathcal{Q}_R} \chi^{p_1} \zeta^2 \varphi_M (|Du - D\ell|) dz + c(\delta) \int_{\mathcal{Q}_R} \zeta^2 \varphi_M \left( \frac{|u - \ell|}{R - r} \right) dz.
 \end{aligned}
 \tag{4.3}$$

Choosing  $\delta$  sufficiently small, we can absorb the first integral of the right hand side into the left. Finally, expanding the derivative  $\eta_t$ , we may write (recalling Remark 2.1)

$$\begin{aligned}
 III &= 2 \int_{\mathcal{Q}_R} \chi^{p_1} \zeta \zeta_t |u - \ell|^2 dz + \int_{\mathcal{Q}_R} \chi^{p_1} \zeta^2 (u - \ell) \cdot (u - \ell)_t dz \\
 &= 2 \int_{\mathcal{Q}_R} \chi^{p_1} \zeta \zeta_t |u - \ell|^2 dz + \frac{1}{2} \int_{\mathcal{Q}_R} \chi^{p_1} \zeta^2 \frac{\partial}{\partial t} |u - \ell|^2 dz.
 \end{aligned}$$

So, an integration by parts yields

$$III = \int_{\mathcal{Q}_R} \chi^{p_1} \zeta \zeta_t |u - \ell|^2 dz.$$

Exploiting the definition of  $\zeta$ , we obtain

$$\begin{aligned}
 III &= \frac{1}{R^2 - r^2} \int_{-R^2}^{-r^2} \int_{\mathcal{B}_R} |u - \ell|^2 \chi^{p_1} \zeta dx dt - \frac{1}{\varepsilon} \int_s^{s+\varepsilon} \int_{\mathcal{B}_R} |u - \ell|^2 \chi^{p_1} \zeta dx dt \\
 &\leq \int_{\mathcal{Q}_R} \left| \frac{u - \ell}{R - r} \right|^2 dz - \frac{1}{\varepsilon} \int_s^{s+\varepsilon} \int_{\mathcal{B}_R} |u - \ell|^2 \chi^{p_1} \zeta dx dt,
 \end{aligned}$$

since  $R^2 - r^2 \geq (R - r)^2$ . Incorporating the above bound and the bounds for  $I$  and  $II$ , in (4.2) and (4.3), into (4.1), we deduce that

$$\begin{aligned}
 &\frac{1}{\varepsilon} \int_s^{s+\varepsilon} \int_{\mathcal{B}_R} |u - \ell|^2 \chi^{p_1} \zeta dx dt + \int_{\mathcal{Q}_R} \chi^{p_1} \zeta^2 \varphi_M (|Du - D\ell|) dz \leq c \int_{\mathcal{Q}_R} \left| \frac{u - \ell}{R - r} \right|^2 dz \\
 &+ c \int_{\mathcal{Q}_R} \varphi_M \left( \frac{|u - \ell|}{R - r} \right) dz.
 \end{aligned}$$

Recalling the definition of  $\zeta$  and  $\chi$ , we may take the limit as  $\varepsilon \rightarrow 0$  to get

$$\begin{aligned}
 &\int_{\mathcal{B}_r} |u(s, x) - \ell(x)|^2 dx + \int_{-r^2}^s \int_{\mathcal{B}_r} \varphi_M (|Du - D\ell|) dz \\
 &\leq c \int_{\mathcal{Q}_R} \left| \frac{u - \ell}{R - r} \right|^2 dz + c \int_{\mathcal{Q}_R} \varphi_M \left( \frac{|u - \ell|}{R - r} \right) dz.
 \end{aligned}$$

We use the previous inequality twice: firstly, by dropping the second term in the left hand side, and taking the supremum over  $s \in (-r^2, 0)$ ; secondly, by dropping the first term in the left hand side and letting  $s$  tend to 0. By summing up the two resulting contributions, this gives the result. □

Finally, an application of the Caccioppoli inequality in Theorem 4.1 and (2.9) produces the following

**Corollary 4.1** *Let  $u \in C^0(-T, 0; L^2(\Omega, \mathbb{R}^N)) \cap L^1(-T, 0; W^{1,1}(\Omega, \mathbb{R}^N))$  be a weak solution to (1.1) satisfying  $\varphi(|Du|) \in L^1(-T, 0; L^1(\Omega))$ . Under hypotheses (a<sub>1</sub>)-(a<sub>4</sub>) and Assumption 1.1, given any standard parabolic cylinder  $\mathcal{Q}_\rho(z_0) \subset \Omega_T$ , with center in  $z_0 = (x_0, t_0)$ , and any affine map  $\ell : \mathbb{R}^N \rightarrow \mathbb{R}^N$ , we have*

$$\begin{aligned} & \sup_{t_0 - (\frac{\rho}{2})^2 < s < t_0} \int_{\mathcal{B}_{\frac{\rho}{2}}(x_0)} \left| \frac{u(x, s) - \ell(x)}{\rho} \right|^2 dx + \int_{\mathcal{Q}_{\frac{\rho}{2}}(z_0)} \varphi_{1+|D\ell|}(|Du - D\ell|) dz \\ & \leq c_0 2^{n+p_1+2} \int_{\mathcal{Q}_\rho(z_0)} \left[ \varphi_{1+|D\ell|} \left( \left| \frac{u - \ell}{\rho} \right| \right) + \left| \frac{u - \ell}{\rho} \right|^2 \right] dz, \end{aligned}$$

where  $c_0$  depends only on  $n, N, L, \nu, p_0, p_1$ .

### 5 Poincaré type inequalities

We begin this section providing a Poincaré type inequality valid for solutions to certain parabolic-like systems. The proof follows the same lines as [5, Lemma 3.1].

**Lemma 5.1** *Let  $\psi$  be an  $N$ -function satisfying  $\Delta_2(\psi, \psi^*) < \infty$ . With  $t_1 < t_2$  and  $U \subseteq \mathbb{R}^n$ , suppose that  $\xi \in L^1(U \times (t_1, t_2), \mathbb{R}^{Nn})$  and  $w \in C^0(t_1, t_2, L^2(U, \mathbb{R}^N)) \cap L^1(t_1, t_2, W^{1,1}(U, \mathbb{R}^N))$  satisfy  $\psi(|Dw|) \in L^1(t_1, t_2, L^1(U))$  and*

$$\int_{U \times (t_1, t_2)} (w \cdot \zeta_t - \xi \cdot D\zeta) dz = 0, \quad \text{for any } \zeta \in C_c^\infty(U \times (t_1, t_2), \mathbb{R}^N). \tag{5.1}$$

Then for any parabolic cylinder  $\mathcal{Q}_\rho(z_0) \subset U \times (t_1, t_2)$ , we have

$$\int_{\mathcal{Q}_\rho(z_0)} \psi \left( \left| \frac{w - (w)_{z_0, \rho}}{\rho} \right| \right) dz \leq c_1 \left[ \int_{\mathcal{Q}_\rho(z_0)} \psi(|Dw|) dz + \psi \left( \int_{\mathcal{Q}_\rho(z_0)} |\xi| dz \right) \right]$$

with  $c_1$  depending only on  $n, N$ , and  $\Delta_2(\psi, \psi^*)$ .

**Proof** We fix a nonnegative symmetric weight function  $\eta \in C_c^\infty(\mathcal{B}_\rho(x_0))$  such that

$$\eta \geq 0, \quad \int_{\mathcal{B}_\rho(x_0)} \eta dx = 1 \quad \text{and} \quad \|\eta\|_\infty + \rho \|D\eta\|_\infty \leq c_\eta, \tag{5.2}$$

and, for  $t \in (t_0 - \rho^2, t_0)$  we denote

$$(w)_\eta(t) = \int_{\mathcal{B}_\rho(x_0)} w(x, t) \eta(x) dx,$$

as well as

$$(w)_\eta = \int_{\mathcal{Q}_\rho(z_0)} w(x, t) \eta(x) dz.$$

By the triangle inequality and the  $\Delta_2$ -condition we have

$$\begin{aligned} & \int_{\mathcal{Q}_\rho(z_0)} \psi \left( \left| \frac{w - (w)_{z_0, \rho}}{\rho} \right| \right) dz \\ & \leq c \left[ \int_{\mathcal{Q}_\rho(z_0)} \psi \left( \left| \frac{w - (w)_\eta(t)}{\rho} \right| \right) dz \right] \end{aligned}$$

$$\begin{aligned}
 & + \int_{t_0-\rho^2}^{t_0} \psi \left( \left| \frac{(w)_\eta(t) - (w)_\eta}{\rho} \right| \right) dt + \psi \left( \left| \frac{(w)_\eta - (w)_{z_0, \rho}}{\rho} \right| \right) \\
 & =: c(I + II + III),
 \end{aligned}$$

with the obvious meaning of  $I, II$ , and  $III$ . Since  $\Delta_2(\psi, \psi^*) < \infty$ , we may bound  $I$  by applying Poincaré’s inequality for vanishing  $\eta$ -mean value (see [16, Theorem 7]) slicewise with respect to  $x$ : for a.e.  $t \in (t_0 - \rho^2, t_0)$ ,

$$I \leq c(n, \Delta_2(\psi, \psi^*)) \int_{\mathcal{Q}_\rho(z_0)} \psi(|Dw|) dz.$$

To bound  $III$ , we use Jensen’s inequality followed by the triangle inequality and  $\Delta_2$ -condition to infer

$$\begin{aligned}
 III & \leq \int_{\mathcal{Q}_\rho(z_0)} \psi \left( \left| \frac{w - (w)_\eta}{\rho} \right| \right) dz \leq c \left[ \int_{\mathcal{Q}_\rho(z_0)} \psi \left( \left| \frac{w - (w)_\eta(t)}{\rho} \right| \right) dz \right. \\
 & \left. + \int_{t_0-\rho^2}^{t_0} \psi \left( \left| \frac{(w)_\eta(t) - (w)_\eta}{\rho} \right| \right) dt \right] \\
 & = c(I + II).
 \end{aligned}$$

So it remains to estimate  $II$ . For this, we recall that  $w$  is a weak solution of the parabolic system (5.1). Even if the solution  $w$ , need not be differentiable in the time variable, with Steklov averages, we may rewrite (5.1) as

$$\int_U (w_t \cdot \zeta + \xi \cdot D\zeta) dx = 0 \quad \text{for all } \zeta \in C_c^\infty(U, \mathbb{R}^N) \text{ and for a.e. } t \in (t_0 - \rho^2, t_0).$$

For  $i = 1, \dots, N$ , let  $e_i \in \mathbb{R}^N$  denote the unit vector in the  $i$ -th coordinate direction. With  $(t, \tau) \subset (t_0 - \rho^2, t_0)$ , we use (5.1) and (5.2), with  $\zeta = \eta e_i \in C_c^\infty(U; \mathbb{R}^N)$ , to write

$$\begin{aligned}
 |[(w)_\eta(t) - (w)_\eta(\tau)] \cdot e_i| & = \left| \int_t^\tau \partial_s [(w)_\eta(s)] \cdot e_i ds \right| = \left| \int_t^\tau \int_{\mathcal{B}_\rho(x_0)} w_s \cdot (\eta e_i) dx ds \right| \\
 & = \left| \int_t^\tau \int_{\mathcal{B}_\rho(x_0)} \xi \cdot (D\eta e_i) dx ds \right| \leq \|D\eta\|_\infty \int_t^\tau \int_{\mathcal{B}_\rho(x_0)} |\xi| dx ds \\
 & \leq \frac{c}{\rho} \int_t^\tau \int_{\mathcal{B}_\rho(x_0)} |\xi| dx ds \leq c \rho \int_{\mathcal{Q}_\rho(z_0)} |\xi| dz.
 \end{aligned}$$

Summing over each component, we conclude that

$$II \leq c \psi \left( \int_{\mathcal{Q}_\rho(z_0)} |\xi| dz \right).$$

Combining these estimates we obtain the desired Poincaré type inequality. □

**Remark 5.1** Suppose that  $\mathcal{A}$  is a bilinear form on  $\mathbb{R}^{Nn}$  and there is  $\Lambda < \infty$  such that  $|\mathcal{A}(\eta_1, \eta_2)| \leq \Lambda |\eta_1| |\eta_2|$ , for all  $\eta_1, \eta_2 \in \mathbb{R}^{Nn}$ . If  $h \in C^\infty(\mathcal{Q}_\rho(z_0); \mathbb{R}^N)$  is  $\mathcal{A}$ -caloric, then we may identify a  $\xi \in C^\infty(\mathcal{Q}_\rho(z_0); \mathbb{R}^{Nn})$  so that, at each  $z \in \mathcal{Q}_\rho(z_0)$ , we have  $\mathcal{A}(Dh(z), \eta) = \xi(z) \cdot \eta$ , for all  $\eta \in \mathbb{R}^{Nn}$ . Then  $|\xi| \leq \Lambda |Dh|$  and Lemma 5.1 and Jensen’s inequality imply

$$\int_{\mathcal{Q}_\rho(z_0)} \psi \left( \left| \frac{h - (h)_{z_0, \rho}}{\rho} \right| \right) dz \leq c_1^* \int_{\mathcal{Q}_\rho(z_0)} \psi(|Dh|) dz,$$

with  $\psi$  an  $N$ -function satisfying  $\Delta_2(\psi, \psi^*) < \infty$  and  $c_1^*$  depending on  $n, N, \Lambda$ , and  $\Delta_2(\psi, \psi^*)$ .

The following Poincaré type inequality for weak solutions of (1.1) is a consequence of the previous lemma.

**Theorem 5.1** *Under the assumptions (a<sub>1</sub>)-(a<sub>4</sub>) and Assumption 1.1, suppose  $u \in C^0(-T, 0; L^2(\Omega, \mathbb{R}^N)) \cap L^1(-T, 0; W^{1,1}(\Omega, \mathbb{R}^N))$  is a weak solution to (1.1) such that  $\varphi(|Du|) \in L^1(-T, 0; L^1(\Omega))$ . Let  $Q_\rho(z_0) \subset \Omega_T$  be a standard parabolic cylinder. Then, for any  $N$ -function  $\psi$  satisfying  $\Delta_2(\psi, \psi^*) < \infty$  and any  $A \in \mathbb{R}^{Nn}$ , we have*

$$\begin{aligned} & \int_{Q_\rho(z_0)} \psi \left( \left| \frac{u - (u)_{z_0, \rho} - A(x - x_0)}{\rho} \right| \right) dz \\ & \leq c_2 \left[ \int_{Q_\rho(z_0)} \psi(|Du - A|) dz + \psi \left( \int_{Q_\rho(z_0)} \varphi'_{1+|A|}(|Du - A|) dz \right) \right] \end{aligned}$$

with  $c_2$  depending on  $n, N, p_0, p_1, v, L$ , and  $\Delta_2(\psi, \psi^*)$ .

**Proof** Without loss of generality we can assume that  $z_0 = (0, 0)$ . Exploiting (1.1) and using the fact that  $\int_{\Omega_T} Ax \cdot \zeta_t dz = 0$  and  $\int_{\Omega_T} a(A) \cdot D\zeta dz = 0$  for any function  $\zeta \in C_c^\infty(\Omega_T, \mathbb{R}^N)$ , we have

$$\int_{\Omega_T} [(u - Ax) \cdot \zeta_t - (a(Du) - a(A)) \cdot D\zeta] dz = 0.$$

Therefore, we can apply the Lemma 5.1 with  $w = u - Ax$  and  $\xi = a(Du) - a(A)$ . As  $Ax$  has zero-mean on  $Q_\rho$ , we obtain:

$$\begin{aligned} & \int_{Q_\rho(z_0)} \psi \left( \left| \frac{u - (u)_\rho - Ax}{\rho} \right| \right) dz \leq c_1 \left[ \int_{Q_\rho(z_0)} \psi(|Du - A|) dz \right. \\ & \left. + \psi \left( \int_{Q_\rho(z_0)} |a(Du) - a(A)| dz \right) \right]. \end{aligned}$$

By means of Remark 2.4, we can estimate the right-hand side of the above inequality, and we get

$$\begin{aligned} & \int_{Q_\rho(z_0)} \psi \left( \left| \frac{u - (u)_\rho - Ax}{\rho} \right| \right) dz \leq c_2 \left[ \int_{Q_\rho(z_0)} \psi(|Du - A|) dz \right. \\ & \left. + \psi \left( \int_{Q_\rho(z_0)} \varphi'_{1+|A|}(|Du - A|) dz \right) \right]. \end{aligned}$$

□

We conclude this section by proving a weird Sobolev-Poincaré inequality for solutions of (1.1). The proof follows the same lines as [31, Lemma 3.4]. A special application of the inequality is required for Theorem 5.3, which we ultimately use to establish the main regularity result in Theorem 1.2.



**Theorem 5.2** *Under the assumptions (a<sub>1</sub>)-(a<sub>4</sub>) and Assumption 1.1, suppose  $u \in C^0(-T, 0; L^2(\Omega, \mathbb{R}^N)) \cap L^1(-T, 0; W^{1,1}(\Omega, \mathbb{R}^N))$  is a weak solution to (1.1) such that  $\varphi(|Du|) \in L^1(-T, 0; L^1(\Omega))$ . Let  $\mathcal{Q}_\rho(z_0) \subseteq \Omega_T$  be a standard parabolic cylinder and  $\psi$  be an  $N$ -function satisfying Assumption 1.1 (for some exponents  $1 < q_0 \leq q_1$ ). Then, for any  $A \in \mathbb{R}^{Nn}$ , any  $\theta_0 > 0$  satisfying*

$$\theta_0 q_0 \in (1, n) \quad \text{and} \quad \frac{nq_1}{nq_1 + 2q_0} \leq \theta_0 \leq 1, \tag{5.3}$$

and each  $\frac{\rho}{2} \leq r < R \leq \rho$ , we have

$$\begin{aligned} & \int_{\mathcal{Q}_r(z_0)} \psi \left( \left| \frac{u - (u)_{z_0,r} - A(x - x_0)}{r} \right| \right) dz \\ & \leq c_3 \psi(T(r, R)^{1/2})^{1-\theta_0} \left[ \int_{\mathcal{Q}_r(z_0)} \psi(|Du - A|^{\theta_0}) dz + \psi \left( \int_{\mathcal{Q}_r(z_0)} \varphi'_{1+|A|}(|Du - A|) dz \right)^{\theta_0} \right], \end{aligned} \tag{5.4}$$

with  $c_3 < \infty$  depending on  $n, N, p_0, p_1, q_0, q_1, \nu, L$ . Here

$$T(r, R) = \int_{\mathcal{Q}_R(z_0)} \left[ \left| \frac{u - (u)_{z_0,R} - A(x - x_0)}{R - r} \right|^2 + \varphi_{1+|A|} \left( \left| \frac{u - (u)_{z_0,R} - A(x - x_0)}{R - r} \right| \right) \right] dz.$$

**Proof** Without loss of generality we can assume that  $z_0 = (0, 0)$ . Suppose  $\theta_0 > 0$  satisfies (5.3). We use the Gagliardo-Nirenberg inequality (see [31, Lemma 2.13] with  $(\psi, \gamma, \theta, p, q_1, q_2) = (\psi^{\frac{1}{\theta_0}}, q_0, \theta_0, \theta_0 q_0, \frac{q_1}{\theta_0}, 2)$ ) to get

$$\int_{B_r} \psi \left( \left| \frac{f}{r} \right| \right) dx \leq c \left( \int_{B_r} \left[ \psi(|Df|^{\theta_0}) + \psi \left( \left| \frac{f}{r} \right| \right)^{\theta_0} \right] dx \right) \psi \left( \left( \int_{B_r} \left| \frac{f}{r} \right|^2 dx \right)^{\frac{1}{2}} \right)^{1-\theta_0}.$$

With  $f = u - (u)_r - Ax$ , we apply the previous inequality to each time slice:

$$\begin{aligned} & \int_{\mathcal{Q}_r} \psi \left( \left| \frac{u - (u)_r - Ax}{r} \right| \right) dz \\ & \leq c \left[ \int_{\mathcal{Q}_r} \left[ \psi(|Du - A|^{\theta_0}) + \psi \left( \left| \frac{u - (u)_r - Ax}{r} \right| \right)^{\theta_0} \right] dz \right] \times \\ & \psi \left( \left( \sup_{-r^2 < t < 0} \int_{B_r} \left| \frac{u - (u)_r - Ax}{r} \right|^2 dx \right)^{\frac{1}{2}} \right)^{1-\theta_0}. \end{aligned}$$

Observe that  $\psi^{\theta_0}$  satisfies Assumption 1.1 with exponents  $1 < \theta_0 q_0 \leq \theta_0 q_1$ . It follows that  $\Delta_2(\psi^{\theta_0}, (\psi^{\theta_0})^*) \leq \max \left\{ 2^{\theta_0 q_1}, 2^{\frac{\theta_0 q_0}{\theta_0 q_0 - 1}} \right\} < \infty$ . We may therefore use the Poincaré type inequality in Theorem 5.1, with  $\psi$  replaced with  $\psi^{\theta_0}$ , to bound the second term in the right

hand side. Thus,

$$\begin{aligned} & \int_{\mathcal{Q}_r} \psi \left( \left| \frac{u - (u)_r - Ax}{r} \right| \right) dz \\ & \leq c \left[ \int_{\mathcal{Q}_r} \psi(|Du - A|)^{\theta_0} dz + \psi \left( \int_{\mathcal{Q}_r} \varphi'_{1+|A|}(|Du - A|) dz \right)^{\theta_0} \right] \times \\ & \psi \left( \left( \sup_{-r^2 < t < 0} \int_{\mathcal{B}_r} \left| \frac{u - (u)_r - Ax}{r} \right|^2 dx \right)^{\frac{1}{2}} \right)^{1-\theta_0}. \end{aligned}$$

Now, to estimate the sup-term, we apply the Caccioppoli inequality on the cylinders  $\mathcal{Q}_r$  and  $\mathcal{Q}_R$  (see Theorem 4.1). Since  $R \leq 2r$ ,

$$\begin{aligned} & \sup_{-r^2 < t < 0} \int_{\mathcal{B}_r} \left| \frac{u - (u)_r - Ax}{r} \right|^2 dx \leq \frac{c_0}{r^{n+2}} \left[ \int_{\mathcal{Q}_R} \varphi_{1+|A|} \left( \left| \frac{u - (u)_r - Ax}{R - r} \right| \right) \right. \\ & \left. + \left| \frac{u - (u)_r - Ax}{R - r} \right|^2 dz \right] \\ & \leq c_0 2^{n+2} \left[ \int_{\mathcal{Q}_R} \varphi_{1+|A|} \left( \left| \frac{u - (u)_r - Ax}{R - r} \right| \right) + \left| \frac{u - (u)_r - Ax}{R - r} \right|^2 dz \right]. \end{aligned}$$

To replace  $(u)_r$  with  $(u)_R$ , we note that

$$|(u)_r - (u)_R| = \left| \int_{\mathcal{Q}_r} [u - (u)_R - Ax] dz \right| \leq 2^{n+2} \int_{\mathcal{Q}_R} |u - (u)_R - Ax| dz.$$

The result follows from Jensen’s inequality and the  $\Delta_2$ -condition. □

To establish the main regularity result, Theorem 1.2, we will need the following inequality, which is proved using Theorem 5.2 with the special choice  $t \mapsto \psi(t) = t^2$ .

**Theorem 5.3** *Let  $u \in C^0(-T, 0; L^2(\Omega, \mathbb{R}^N)) \cap L^1(-T, 0; W^{1,1}(\Omega, \mathbb{R}^N))$  be a weak solution to (1.1) satisfying  $\varphi(|Du|) \in L^1(-T, 0; L^1(\Omega))$ . Under assumptions (a<sub>1</sub>)-(a<sub>4</sub>) and Assumption 1.1, given any standard cylinder  $\mathcal{Q}_\rho(z_0) \subseteq \Omega_T$ , with center  $z_0 = (x_0, t_0)$ , and any  $A \in \mathbb{R}^{Nn}$ , we have*

$$\begin{aligned} & \int_{\mathcal{Q}_{\rho/2}(z_0)} \left| \frac{u - (u)_{\rho/2} - A(x - x_0)}{\rho/2} \right|^2 dz \\ & \leq c_4 \left[ \int_{\mathcal{Q}_\rho(z_0)} \varphi_{1+|A|}(|Du - A|) dz + \left( \int_{\mathcal{Q}_\rho(z_0)} |Du - A|^{p_0} dz \right)^{\frac{2}{p_0}} \right. \\ & \quad \left. + \varphi_{1+|A|} \left( \int_{\mathcal{Q}_\rho(z_0)} \varphi'_{1+|A|}(|Du - A|) dz \right) + \left( \int_{\mathcal{Q}_\rho(z_0)} \varphi'_{1+|A|}(|Du - A|) dz \right)^2 \right], \end{aligned}$$

where  $c_4$  depends on  $n, N, p_0, p_1, \nu$ , and  $L$ .

**Proof** As usual, we may assume  $z_0 = (0, 0)$ . Let  $\frac{\rho}{2} \leq r < R \leq \rho$ . For convenience, put  $M = 1 + |A|$ . We use Theorem 5.2 with the  $N$ -function  $t \mapsto \psi(t) = t^2, q_0 = q_1 = 2$ , and

$\theta_0 = p_0/2$ . Observe that, since  $\frac{2n}{n+2} < p_0 < 2$ ,

$$\theta_0 q_0 = p_0 \in (1, n) \quad \text{and} \quad \frac{n}{n+2} = \frac{nq_1}{nq_1 + 2q_0} < \theta_0 \leq 1.$$

The inequality (5.4) becomes

$$\begin{aligned} \int_{\mathcal{Q}_r} \left| \frac{u - (u)_r - Ax}{r} \right|^2 dz &\leq c_3 (T(r, R))^{1 - \frac{p_0}{2}} \times \\ &\left[ \int_{\mathcal{Q}_r} |Du - A|^{p_0} dz + \left( \int_{\mathcal{Q}_r} \varphi'_M(|Du - A|) dz \right)^{p_0} \right], \end{aligned} \tag{5.5}$$

where

$$\begin{aligned} T(r, R) &= \int_{\mathcal{Q}_R} \left| \frac{u - (u)_R - Ax}{R - r} \right|^2 + \varphi_M \left( \left| \frac{u - (u)_R - Ax}{R - r} \right| \right) dz \\ &\leq \left( \frac{R}{R - r} \right)^2 \int_{\mathcal{Q}_R} \left| \frac{u - (u)_R - Ax}{R} \right|^2 dz + \left( \frac{R}{R - r} \right)^{p_1} \int_{\mathcal{Q}_R} \varphi_M \left( \left| \frac{u - (u)_R - Ax}{R} \right| \right) dz \\ &\leq \left( \frac{R}{R - r} \right)^{p_1} \left[ \int_{\mathcal{Q}_R} \left| \frac{u - (u)_R - Ax}{R} \right|^2 dz + \int_{\mathcal{Q}_R} \varphi_M \left( \left| \frac{u - (u)_R - Ax}{R} \right| \right) dz \right]. \end{aligned}$$

Here, we have taken advantage of (2.2),  $p_1 > 2$ , and  $\frac{R}{R-r} > 1$ . Now, we use Young's inequality in (5.5), with  $\frac{1}{\theta_0} = \frac{2}{p_0}$  and its conjugate  $\frac{1}{1-\theta_0} = \frac{2}{2-p_0}$ . This produces

$$\begin{aligned} \int_{\mathcal{Q}_r} \left| \frac{u - (u)_r - Ax}{r} \right|^2 dz &\leq \frac{1}{2} \int_{\mathcal{Q}_R} \left| \frac{u - (u)_R - Ax}{R} \right|^2 dz + \frac{1}{2} \int_{\mathcal{Q}_R} \varphi_M \left( \left| \frac{u - (u)_R - Ax}{R} \right| \right) dz \\ &\quad + c \left( \frac{R}{R - r} \right)^{\frac{p_1(2-p_0)}{p_0}} \left[ \int_{\mathcal{Q}_r} |Du - A|^{p_0} dz + \left( \int_{\mathcal{Q}_r} \varphi'_M(|Du - A|) dz \right)^{p_0} \right]^{\frac{2}{p_0}}. \end{aligned}$$

For the second term in the upper bound, we can use Theorem 5.1 with  $\psi(t) = \varphi_M(t)$ . The previous inequality becomes

$$\begin{aligned} \int_{\mathcal{Q}_r} \left| \frac{u - (u)_r - Ax}{r} \right|^2 dz &\leq \frac{1}{2} \int_{\mathcal{Q}_R} \left| \frac{u - (u)_R - Ax}{R} \right|^2 dz + c \int_{\mathcal{Q}_R} \varphi_M(|Du - A|) dz + \varphi_M \left( \int_{\mathcal{Q}_R} \varphi'_M(|Du - A|) dz \right) \\ &\quad + c \left( \frac{R}{R - r} \right)^{\frac{p_1(2-p_0)}{p_0}} \left[ \int_{\mathcal{Q}_r} |Du - A|^{p_0} dz + \left( \int_{\mathcal{Q}_r} \varphi'_M(|Du - A|) dz \right)^{p_0} \right]^{\frac{2}{p_0}}. \end{aligned}$$

Enlarging the domain of integration (recall that  $\frac{\rho}{2} \leq r < R \leq \rho$ ), we get

$$\begin{aligned} \int_{\mathcal{Q}_r} \left| \frac{u - (u)_r - Ax}{r} \right|^2 dz &\leq \frac{1}{2} \int_{\mathcal{Q}_\rho} \left| \frac{u - (u)_R - Ax}{R} \right|^2 dz + c \int_{\mathcal{Q}_\rho} \varphi_M(|Du - A|) dz + \varphi_M \left( \int_{\mathcal{Q}_\rho} \varphi'_M(|Du - A|) dz \right) \\ &\quad + c \left( \frac{\rho}{R - r} \right)^{\frac{p_1(2-p_0)}{p_0}} \left[ \int_{\mathcal{Q}_\rho} |Du - A|^{p_0} dz + \left( \int_{\mathcal{Q}_\rho} \varphi'_M(|Du - A|) dz \right)^{p_0} \right]^{\frac{2}{p_0}}. \end{aligned}$$

We are now in position to apply [29, Lemma 6.1] to conclude the proof. □

### 6 Linearization

We now prove a lemma that facilitates the comparison of solutions to our system (1.1) to solutions for a linear system with constant coefficients.

**Lemma 6.1** *Suppose  $u \in C^0(-T, 0; L^2(\Omega, \mathbb{R}^N)) \cap L^1(-T, 0; W^{1,1}(\Omega, \mathbb{R}^N))$  is a weak solution to (1.1) that satisfies  $\varphi(|Du|) \in L^1(-T, 0; L^1(\Omega))$ . Let a generic parabolic cylinder  $\mathcal{Q}_{\rho, \tau}(z_0) \subset \Omega_T$ , with center  $z_0 = (x_0, t_0)$ , be given. Under the hypotheses (a1)-(a4) and Assumption 1.1, given any affine map  $\ell : \mathbb{R}^n \rightarrow \mathbb{R}^N$  and any map  $\eta \in C_c^\infty(\mathcal{Q}_{\rho/2, \tau/4}(z_0), \mathbb{R}^N)$ , we have*

$$\left| \int_{\mathcal{Q}_{\rho/2, \tau/4}(z_0)} (u - \ell) \cdot \eta_t - Da(D\ell)(Du - D\ell, D\eta) dz \right| \leq c_5 \varphi_{1+|D\ell|}(1) \left\{ \omega \left( S^{\frac{1}{2}} \right)^{\frac{1}{2}} S^{\frac{1}{2}} + S \right\} \sup_{\mathcal{Q}_{\rho/2, \tau/4}(z_0)} |D\eta|$$

where  $c_5$  depends only on  $n, N, L, \nu, p_0, p_1$ . Here

$$S = \int_{\mathcal{Q}_{\rho/2, \tau/4}(z_0)} \frac{\varphi_{1+|D\ell|}(|Du - D\ell|)}{\varphi_{1+|D\ell|}(1)} dz.$$

**Proof** We may assume  $z_0 = (0, 0)$ . For convenience, we write  $\mathcal{Q}_{\rho, \tau} = \mathcal{Q}_{\rho, \tau}(z_0)$  and  $M = 1 + |D\ell|$ . Let  $\eta \in C_c^\infty(\mathcal{Q}_{\rho/2, \tau/4}, \mathbb{R}^N)$  be given. We first note:

$$\begin{aligned} & \int_{\mathcal{Q}_{\rho/2, \tau/4}} [(u - \ell) \cdot \eta_t - Da(D\ell)(Du - D\ell, D\eta)] dz \\ &= \int_{\mathcal{Q}_{\rho/2, \tau/4}} [(u - \ell) \cdot \eta_t - a(Du) \cdot D\eta] dz \\ & \quad + \int_{\mathcal{Q}_{\rho/2, \tau/4}} [(a(Du) - a(D\ell)) \cdot D\eta - Da(D\ell)(Du - D\ell, D\eta)] dz =: I + II. \end{aligned}$$

Since  $u$  is a weak solution to (1.1) and  $\int_{\mathcal{Q}_{\rho/2, \tau/4}} \ell \cdot \eta_t dz = 0$ , we infer that  $I = 0$ . On the other hand,

$$\begin{aligned} \int_{\mathcal{Q}_{\rho/2, \tau/4}} (a(Du) - a(D\ell)) \cdot D\eta dz &= \int_{\mathcal{Q}_{\rho/2, \tau/4}} \int_0^1 \frac{d}{ds} [a(D\ell + s(Du - D\ell))] \cdot D\eta ds dz \\ &= \int_{\mathcal{Q}_{\rho/2, \tau/4}} \int_0^1 Da(D\ell + s(Du - D\ell))(Du - D\ell, D\eta) ds dz, \end{aligned}$$

so that

$$II = \int_{\mathcal{Q}_{\rho/2, \tau/4}} \int_0^1 [Da(D\ell + s(Du - D\ell)) - Da(D\ell)](Du - D\ell, D\eta) ds dz.$$

Using the continuity assumption (a4),

$$\begin{aligned} |II| &\leq L \int_{\mathcal{Q}_{\rho/2, \tau/4}} \int_0^1 \omega \left( \frac{s|Du - D\ell|}{1 + |D\ell + s(Du - D\ell)| + |D\ell|} \right) \\ & \quad \varphi''(M + |D\ell + s(Du - D\ell)|) |Du - D\ell| |D\eta| ds dz \end{aligned}$$

$$\leq L \int_{\mathcal{Q}_{\rho/2, \tau/4}} \omega(|Du - D\ell|) |Du - D\ell| |D\eta| \int_0^1 \varphi''(M + |D\ell + s(Du - D\ell)|) ds dz.$$

As  $\varphi'$  is nondecreasing, Assumption 1.1, Lemma 2.2, and (2.4) yield

$$\int_0^1 \varphi''(M + |D\ell + s(Du - D\ell)|) ds \leq c \frac{\varphi'(M + |Du| + |D\ell|)}{M + |Du| + |D\ell|} \leq c \frac{\varphi'(M + |Du|)}{M + |Du|}.$$

Thus,

$$|II| \leq c \sup_{\mathcal{Q}_{\rho/2, \tau/4}} |D\eta| \int_{\mathcal{Q}_{\rho/2, \tau/4}} \omega(|Du - D\ell|) |Du - D\ell| \frac{\varphi'(M + |Du|)}{M + |Du|} dz.$$

Now, we distinguish in  $\mathcal{Q}_{\rho/2, \tau/4}$  the points where  $|Du - D\ell| \leq M$  from those where  $|Du - D\ell| > M$ . Denote by  $\mathbb{X}$  the first set and by  $\mathbb{Y}$  the second. On  $\mathbb{X}$ , we have  $|Du| \leq 2M$ , so (2.4) implies  $\varphi'(M + |Du|) \sim \varphi'(M)$ . Thus

$$\begin{aligned} |Du - D\ell| \frac{\varphi'(M + |Du|)}{M + |Du|} &\leq c \left(\frac{\varphi'(M)}{M}\right)^{\frac{1}{2}} \left(|Du - D\ell|^2 \frac{\varphi'(M + |Du - D\ell|)}{M + |Du - D\ell|}\right)^{\frac{1}{2}} \\ &\leq c \left(\frac{\varphi'(M)}{M}\right)^{\frac{1}{2}} \varphi_M(|Du - D\ell|)^{\frac{1}{2}}, \end{aligned}$$

where we have used (2.8). Moreover,

$$\begin{aligned} |Du - D\ell| &\leq c \frac{M + |Du|}{\varphi'(M + |Du|)} \left(\frac{\varphi'(M)}{M}\right)^{\frac{1}{2}} \varphi_M(|Du - D\ell|)^{\frac{1}{2}} \\ &\leq c \left(\frac{M}{\varphi'(M)}\right)^{\frac{1}{2}} \varphi_M(|Du - D\ell|)^{\frac{1}{2}}. \end{aligned}$$

It follows that

$$\begin{aligned} &\int_{\mathcal{Q}_{\rho/2, \tau/4}} \chi_{\mathbb{X}} \omega(|Du - D\ell|) |Du - D\ell| \frac{\varphi'(M + |Du|)}{M + |Du|} dz \\ &\leq c \left(\frac{\varphi'(M)}{M}\right)^{\frac{1}{2}} \int_{\mathcal{Q}_{\rho/2, \tau/4}} \omega\left(\left(\frac{M}{\varphi'(M)}\right)^{\frac{1}{2}} \varphi_M(|Du - D\ell|)^{\frac{1}{2}}\right) \varphi_M(|Du - D\ell|)^{\frac{1}{2}} dz. \end{aligned}$$

Applying Hölder’s inequality and using the concavity of  $\omega$  and the bound  $\omega \leq 1$ , we continue with

$$\begin{aligned} &\int_{\mathcal{Q}_{\rho/2, \tau/4}} \chi_{\mathbb{X}} \omega(|Du - D\ell|) |Du - D\ell| \frac{\varphi'(M + |Du|)}{M + |Du|} dz \\ &\leq c \left(\frac{\varphi'(M)}{M}\right)^{\frac{1}{2}} \left(\int_{\mathcal{Q}_{\rho/2, \tau/4}} \omega\left(\left(\frac{M}{\varphi'(M)}\right)^{\frac{1}{2}} \varphi_M(|Du - D\ell|)^{\frac{1}{2}}\right)^2 dz\right)^{\frac{1}{2}} \times \\ &\quad \left(\int_{\mathcal{Q}_{\rho/2, \tau/4}} \varphi_M(|Du - D\ell|) dz\right)^{\frac{1}{2}} \\ &\leq c \left(\frac{\varphi'(M)}{M}\right)^{\frac{1}{2}} \omega\left(\left(\frac{M}{\varphi'(M)}\right)^{\frac{1}{2}} \left(\int_{\mathcal{Q}_{\rho/2, \tau/4}} \varphi_M(|Du - D\ell|) dz\right)^{\frac{1}{2}}\right)^{\frac{1}{2}} \times \end{aligned}$$

$$\begin{aligned} & \left( \int_{\mathcal{Q}_{\rho/2, \tau/4}} \varphi_M(|Du - D\ell|) dz \right)^{\frac{1}{2}} \\ & \leq c \left( \frac{\varphi'(M)}{M} \right)^{\frac{1}{2}} \omega \left( \left( \frac{M}{\varphi'(M)} \right)^{\frac{1}{2}} S^{\frac{1}{2}} \varphi_M(1)^{\frac{1}{2}} \right)^{\frac{1}{2}} S^{\frac{1}{2}} \varphi_M(1)^{\frac{1}{2}}. \end{aligned}$$

To complete the bound on  $\mathbb{X}$ , we use (2.4) and (2.8) to deduce that

$$\frac{\varphi'(M)}{M} \sim \frac{\varphi'(M + 1)}{M + 1} \sim \varphi_M(1).$$

We conclude that

$$\int_{\mathcal{Q}_{\rho/2, \tau/4}} \chi_{\mathbb{X}} \omega(|Du - D\ell|) |Du - D\ell| \frac{\varphi'(M + |Du|)}{M + |Du|} dz \leq c \varphi_M(1) \omega \left( S^{\frac{1}{2}} \right)^{\frac{1}{2}} S^{\frac{1}{2}}. \tag{6.1}$$

Turning to the set  $\mathbb{Y}$ , we have  $1 \leq M \leq |Du - D\ell| \leq |Du - D\ell|^2$ . Recalling (2.4) and (2.8) once more, we have

$$|Du - D\ell| \frac{\varphi'(M + |Du|)}{M + |Du|} \leq c |Du - D\ell|^2 \frac{\varphi'(M + |Du - D\ell|)}{M + |Du - D\ell|} \leq c \varphi_M(|Du - D\ell|).$$

Thus

$$\begin{aligned} & \int_{\mathcal{Q}_{\rho/2, \tau/4}} \chi_{\mathbb{Y}} \omega(|Du - D\ell|) |Du - D\ell| \frac{\varphi'(M + |Du|)}{M + |Du|} dz \\ & \leq c \int_{\mathcal{Q}_{\rho/2, \tau/4}} \varphi_M(|Du - D\ell|) dz = c \varphi_M(1) S. \end{aligned} \tag{6.2}$$

Here, we have again used  $\omega \leq 1$ . The lemma follows from summing (6.1) and (6.2). To verify the claim for the parameter dependencies of the constant  $c_5$ , we review the proof and note that only the hypothesis (a<sub>4</sub>), Assumption 1.1, and properties (2.4) and (2.8) were required. □

### 7 Decay Estimates

For convenience, we recall the excess functional introduced in Section 1. Given  $z_0 \in \Omega_T$ ,  $a \geq 0, r > 0$ , and an affine map  $\ell : \mathbb{R}^n \rightarrow \mathbb{R}^N$ , define

$$\Psi_a(z_0, r, \ell) = \int_{\mathcal{Q}_r(z_0)} \left( \left| \frac{u - \ell}{r} \right|^2 + \varphi_a \left( \left| \frac{u - \ell}{r} \right| \right) \right) dz.$$

In the following lemma we provide the decay of the excess  $\Psi_{1+|D\ell_{z_0, \rho}|}(z_0, \rho, \ell_{z_0, \rho})$ . Recall that  $\ell_{z_0, \rho}$  is defined in (2.11) and denotes the time-independent affine map closest to  $u$  with respect to the  $L^2$ -norm on  $\mathcal{Q}_\rho(z_0)$ .

**Lemma 7.1** (Decay Estimate) *Suppose that hypotheses (a<sub>1</sub>)-(a<sub>4</sub>) and Assumption 1.1 hold and that  $u \in C^0(-T, 0; L^2(\Omega, \mathbb{R}^N)) \cap L^1(-T, 0; W^{1,1}(\Omega, \mathbb{R}^N))$  is a weak solution to (1.1) satisfying  $\varphi(|Du|) \in L^1(-T, 0; L^1(\Omega))$ . Let  $M_0 > 0$  and  $0 < \alpha < 1$  be given. There exist  $0 < \varepsilon, \theta < 1$  with the following property: if  $z_0 \in \Omega_T$  and  $\rho > 0$  are such that  $\mathcal{Q}_\rho(z_0) \subseteq \Omega_T$ ,*

$$|D\ell_{z_0, \rho}| \leq M_0, \quad \text{and} \quad \Psi_{1+|D\ell_{z_0, \rho}|}(z_0, \rho, \ell_{z_0, \rho}) \leq \varepsilon,$$

then

$$\Psi_{1+|D\ell_{z_0,\theta\rho}|}(z_0, \theta\rho, \ell_{z_0,\theta\rho}) \leq \theta^{2\alpha} \Psi_{1+|D\ell_{z_0,\rho}|}(z_0, \rho, \ell_{z_0,\rho})$$

and

$$|D\ell_{z_0,\theta\rho}| \leq |D\ell_{z_0,\rho}| + (n + 2) \left(\frac{1}{\theta}\right)^{n+3} \Psi_{1+|D\ell_{z_0,\rho}|}(z_0, \rho, \ell_{z_0,\rho})^{\frac{1}{2}}.$$

**Proof** For each  $0 < r \leq \rho$ , we write  $\ell_r = \ell_{z_0,r}$  and  $\Psi_a(r) = \Psi_a(z_0, r, \ell_r)$ . Define  $v = u - \ell_\rho$  and  $M = 1 + |D\ell_\rho| \leq M_0 + 1$ . Let  $\mathcal{A}$  denote the bilinear form  $D\mathcal{A}(D\ell_\rho)$ . Thus

$$\mathcal{A}(\xi, \xi) \geq \nu\varphi''(1 + |D\ell_\rho|)|\xi|^2 \quad \text{and} \quad |\mathcal{A}(\xi, \eta)| \leq L\varphi''(1 + |D\ell_\rho|)|\xi||\eta|.$$

Let us recall that  $\Delta_2(\varphi_M) \leq 2^{p_1}$ ; with  $0 < \alpha < 1$  given, define

$$0 < \theta = \theta(M, \alpha) = \min \left\{ \frac{1}{32}, \left( \frac{1}{1 + c_{6,6}} \right)^{\frac{1}{2(1-\alpha)}} \right\}, \tag{7.1}$$

and suppose that

$$\begin{aligned} \Psi_M(\rho) &\leq \varepsilon = \varepsilon(M) \\ &\leq \min \left\{ \frac{\varphi_M(1)}{c_{6,1}}, \frac{\delta^2}{c_{6,2}^2}, \frac{1}{(\omega_n + 1)c_{6,3}c_{6,4}}, \left( \frac{\theta^{n+3}}{n + 2} \right)^2, \frac{(4\theta)^{n+p_1+4}}{c_{6,3}c_{6,5} (2 + \Delta_2(\varphi_M)^3)} \right\}, \end{aligned} \tag{7.2}$$

where the precise values of the constants  $c_{6,i} > 1$  will be determined in the course of the proof. The constant  $0 < \delta < 1$  is specified by Theorem 3.2, while the constant  $C$  appearing below may change from line to line but will depend on only  $n, N, L/\nu, p_0, p_1$ .

As  $\ell_\rho$  is independent of  $t$ , the map  $v$  is a weak solution to (1.1). Our first objective is to take advantage of the  $\mathcal{A}$ -caloric approximation lemma to produce an  $\mathcal{A}$ -caloric map close to  $v$ . With  $S$  defined in Lemma 6.1, the Caccioppoli inequality in Corollary 4.1 implies

$$\begin{aligned} S &= \int_{Q_{\rho/2}} \frac{\varphi_M(|Du - D\ell_\rho|)}{\varphi_M(1)} dz \\ &\leq \frac{c_0 2^{n+p_1+2}}{\varphi_M(1)} \int_{Q_\rho} \left[ \left| \frac{v}{\rho} \right|^2 + \varphi_M \left( \left| \frac{v}{\rho} \right| \right) \right] dz \leq \frac{c_{6,1}}{\varphi_M(1)} \Psi_M(\rho) \leq 1. \end{aligned}$$

Note that  $c_{6,1} = c_0 2^{n+p_1+2}$ . Now, Lemma 6.1 delivers the bound

$$\begin{aligned} \left| \int_{Q_{\rho/2}} (v \cdot \partial_t \eta - \mathcal{A}(Dv, D\eta)) dz \right| &\leq c_5 \varphi_M(1) \left\{ \omega \left( S^{\frac{1}{2}} \right)^{\frac{1}{2}} S^{\frac{1}{2}} + S \right\} \sup_{Q_{\rho/2}} |D\eta| \\ &\leq 2c_5 c_{6,1}^{\frac{1}{2}} \varphi_M(1)^{\frac{1}{2}} \Psi_M(\rho)^{\frac{1}{2}} \sup_{Q_{\rho/2}} |D\eta| \\ &= c_{6,2} \Psi_M(\rho)^{\frac{1}{2}} \sup_{Q_{\rho/2}} |D\eta| \\ &\leq \delta \sup_{Q_{\rho/2}} |D\eta|. \end{aligned}$$

The smallness condition (7.2) was applied in the last inequality. This verifies the requirement in (3.3) of Theorem 3.2. For the other requirement in (3.4), we again use the Caccioppoli

inequality:

$$\begin{aligned} & \sup_{t_0-\rho^2/4 < t < t_0} \int_{\mathcal{B}_{\rho/2}} \left| \frac{v}{\rho/2} \right|^2 dx + \int_{\mathcal{Q}_{\rho/2}} \left( \varphi_M \left( \left| \frac{v}{\rho/2} \right| \right) + \varphi_M(|Dv|) \right) dz \\ & \leq 4c_{6,1} \Psi_M(\rho) + \Delta_2(\varphi_M) \int_{\mathcal{Q}_{\rho/2}} \varphi_M \left( \left| \frac{v}{\rho} \right| \right) dz \\ & \leq c_0 2^{n+p_1+5} \Psi_M(\rho) = c_{6,3} \Psi_M(\rho) =: \gamma^2 \leq \min\{1/\omega_n, 1\}. \end{aligned}$$

With the hypotheses of Theorem 3.2 satisfied, taking into account Remarks 2.3 and 3.1, we obtain an  $\mathcal{A}$ -caloric map  $h \in C^\infty(\mathcal{Q}_{\rho/4}; \mathbb{R}^N)$  such that

$$\int_{\mathcal{Q}_{\rho/4}} \left( \left| \frac{\gamma h}{\rho/4} \right|^2 + \varphi_M \left( \left| \frac{\gamma h}{\rho/4} \right| \right) + \varphi_M(|\gamma Dh|) \right) dz \leq 2^{n+2} K_{\varphi_M} \gamma^2 \tag{7.3}$$

and

$$\int_{\mathcal{Q}_{\rho/4}} \left( \left| \frac{v - \gamma h}{\rho/4} \right|^2 + \varphi_M \left( \left| \frac{v - \gamma h}{\rho/4} \right| \right) \right) dz \leq \varepsilon \gamma^2.$$

Recall that  $0 < \theta < 1/16$  by definition (7.1). As in Lemma 2.8, for  $0 < r \leq \rho/4$ , we define the affine map  $\ell_r^{(h)}(x) = (h)_{z_0,r} + (Dh)_{z_0,r}(x - x_0)$ . We want to produce a bound for

$$\int_{\mathcal{Q}_{\theta\rho}} \left| \frac{v - \gamma \ell_{\theta\rho}^{(h)}}{\theta\rho} \right|^2 dz + \int_{\mathcal{Q}_{\theta\rho}} \varphi_M \left( \left| \frac{v - \gamma \ell_{\theta\rho}^{(h)}}{\theta\rho} \right| \right) dz =: I + II.$$

We will focus on  $II$ . The argument for  $I$  is similar. We may write

$$\begin{aligned} II & \leq \Delta_2(\varphi_M) \int_{\mathcal{Q}_{\theta\rho}} \varphi_M \left( \left| \frac{v - \gamma h}{\theta\rho} \right| \right) + \varphi_M \left( \left| \frac{\gamma(h - \ell_{z_0,\theta\rho}^{(h)})}{\theta\rho} \right| \right) dz \\ & \leq \Delta_2(\varphi_M) \left( \frac{1}{4\theta} \right)^{p_1+n+2} \int_{\mathcal{Q}_{\rho/4}} \varphi_M \left( \left| \frac{v - \gamma h}{\rho/4} \right| \right) dz + \Delta_2(\varphi_M) \int_{\mathcal{Q}_{\theta\rho}} \varphi_M \left( \left| \frac{\gamma(h - \ell_{z_0,\theta\rho}^{(h)})}{\theta\rho} \right| \right) dz \\ & \leq \Delta_2(\varphi_M) \left( \frac{1}{4\theta} \right)^{p_1+n+2} \varepsilon \gamma^2 + \Delta_2(\varphi_M) \int_{\mathcal{Q}_{\theta\rho}} \varphi_M \left( \left| \frac{\gamma(h - \ell_{z_0,\theta\rho}^{(h)})}{\theta\rho} \right| \right) dz. \end{aligned} \tag{7.4}$$

Observe that Lemma 2.5-(b) and (7.3) imply

$$\begin{aligned} \int_{\mathcal{Q}_{\rho/4}} \left| \frac{\gamma(h - \ell_{z_0,\rho/4}^{(h)})}{\rho/4} \right|^2 dz & \leq 6 \int_{\mathcal{Q}_{\rho/4}} \left| \frac{\gamma h}{\rho/4} \right|^2 dz + 3 \left( \int_{\mathcal{Q}_{\rho/4}} |\gamma Dh| \right)^2 dz \\ & \leq 2^{n+5} K_{\varphi_M} \gamma^2 \\ & \quad + \frac{C}{\varphi''(M)} \int_{\mathcal{Q}_{\rho/4}} \varphi_M(|\gamma Dh|) dz + \frac{C}{\varphi_M(1)^2} \left( \int_{\mathcal{Q}_{\rho/4}} \varphi_M(|\gamma Dh|) dz \right)^2 \\ & \leq 2^{n+5} K_{\varphi_M} \gamma^2 + C \left( \frac{1}{\varphi''(M)} + \frac{2^{n+2} K_{\varphi_M} \gamma^2}{\varphi_M(1)^2} \right) 2^{n+2} K_{\varphi_M} \gamma^2 \\ & \leq 2^{2n+5} K_{\varphi_M}^2 \left[ 1 + C \left( \frac{1}{\varphi''(M)} + \frac{1}{\varphi_M(1)^2} \right) \right] \gamma^2 = c_{6,4} \gamma^2 \leq 1. \end{aligned}$$



We may therefore use the inequality in Lemma 2.8, Remark 2.3, (2.7), and Jensen’s inequality to deduce that

$$\int_{Q_{\theta\rho}} \varphi_M \left( \left| \frac{\gamma(h - \ell_{z_0, \theta\rho}^{(h)})}{\theta\rho} \right| \right) dz \leq C\theta^2 \int_{Q_{\rho/4}} \varphi_M \left( \left| \frac{\gamma(h - \ell_{z_0, \rho/4}^{(h)})}{\rho/4} \right| \right) dz. \tag{7.5}$$

On the other hand Jensen’s inequality and the  $\Delta_2$ -property imply

$$\begin{aligned} \int_{Q_{\rho/4}} \varphi_M \left( \left| \frac{\gamma(h - \ell_{z_0, \rho/4}^{(h)})}{\rho/4} \right| \right) dz &\leq \Delta_2(\varphi_M)^2 \int_{Q_{\rho/4}} \left( \varphi_M \left( \left| \frac{\gamma h}{\rho/4} \right| \right) + \varphi_M(|\gamma Dh|) \right) dz \\ &\leq 2^{n+2} K_{\varphi_M} \Delta_2(\varphi_M)^2 \gamma^2. \end{aligned} \tag{7.6}$$

With (7.5) and (7.6) we return to (7.4) to obtain

$$\int_{Q_{\theta\rho}} \varphi_M \left( \left| \frac{v - \gamma \ell_{\theta\rho}^{(h)}}{\theta\rho} \right| \right) dz = II \leq \Delta_2(\varphi_M)^3 \left[ \left( \frac{1}{4\theta} \right)^{p_1+n+2} \varepsilon + C2^{n+2} K_{\varphi_M} \theta^2 \right] \gamma^2.$$

We similarly obtain

$$\int_{Q_{\theta\rho}} \left| \frac{v - \gamma \ell_{\theta\rho}^{(h)}}{\theta\rho} \right|^2 dz = I \leq \left[ 2 \left( \frac{1}{4\theta} \right)^{n+4} \varepsilon + Cc_{6,4} \theta^2 \right] \gamma^2.$$

Thus, since  $\theta \leq 1/4$ ,

$$\begin{aligned} \int_{Q_{\theta\rho}} \left| \frac{v - \gamma \ell_{z_0, \theta\rho}^{(h)}}{\theta\rho} \right|^2 + \varphi_M \left( \left| \frac{v - \gamma \ell_{z_0, \theta\rho}^{(h)}}{\theta\rho} \right| \right) dz \\ \leq \left[ \left( \frac{1}{4\theta} \right)^{n+p_1+2} (2 + \Delta_2(\varphi_M)^3) \varepsilon + Cc_{6,4} (1 + \Delta_2(\varphi_M)^3) \theta^2 \right] \gamma^2. \end{aligned}$$

Let us point out that  $1 + |D\ell_{\theta\rho}| \leq M + 1$ . Indeed, from (2.13) we obtain

$$\begin{aligned} |D\ell_{\theta\rho}| &\leq |D\ell_\rho| + |D\ell_\rho - D\ell_{\theta\rho}| \\ &\leq M - 1 + (n + 2) \int_{Q_{\theta\rho}} \left| \frac{u - \ell_\rho}{\theta\rho} \right| dz \leq M - 1 + \left( \frac{n + 2}{\theta^{n+3}} \right) \int_{Q_\rho} \left| \frac{u - \ell_\rho}{\rho} \right| dz \\ &\leq M - 1 + \left( \frac{n + 2}{\theta^{n+3}} \right) \left( \int_{Q_\rho} \left| \frac{u - \ell_\rho}{\rho} \right|^2 dz \right)^{\frac{1}{2}} \\ &\leq M - 1 + \left( \frac{n + 2}{\theta^{n+3}} \right) \Psi_{1+|D\ell_\rho|}(\rho)^{\frac{1}{2}} \\ &\leq M - 1 + \left( \frac{n + 2}{\theta^{n+3}} \right) \varepsilon^{\frac{1}{2}} \leq M \end{aligned} \tag{7.7}$$

provided  $\varepsilon \leq \left( \frac{\theta^{n+3}}{n + 2} \right)^2$ .

Now, using Lemma 2.4 (with  $M$  replaced by  $M + 1$ ) and Lemma 2.6 (with  $\ell = \ell_\rho - \gamma \ell_{z_0, \theta\rho}^{(h)}$ ), and defining  $c_{6,5} = \kappa_0 4^{p_1+1} (M + 1)^{p_1+2} (\kappa_0$  from Lemma 2.6), we have

$$\Psi_{1+|D\ell_{\theta\rho}|}(z_0, \theta\rho, \ell_{\theta\rho}) = \int_{Q_{\theta\rho}} \left( \left| \frac{u - \ell_{\theta\rho}}{\theta\rho} \right|^2 + \varphi_{1+|D\ell_{\theta\rho}|} \left( \left| \frac{u - \ell_{\theta\rho}}{\theta\rho} \right| \right) \right) dz$$

$$\begin{aligned}
 &\leq 4^{p_1+1} (M+1)^{p_1+2} \int_{\mathcal{Q}_{\theta\rho}} \left( \left| \frac{u - \ell_{\theta\rho}}{\theta\rho} \right|^2 + \varphi_M \left( \left| \frac{u - \ell_{\theta\rho}}{\theta\rho} \right| \right) \right) dz \\
 &\leq c_{6,5} \int_{\mathcal{Q}_{\theta\rho}} \left( \left| \frac{u - \ell_\rho - \gamma \ell_{z_0, \theta\rho}^{(h)}}{\theta\rho} \right|^2 + \varphi_M \left( \left| \frac{u - \ell_\rho - \gamma \ell_{z_0, \theta\rho}^{(h)}}{\theta\rho} \right| \right) \right) dz \\
 &= c_{6,5} \int_{\mathcal{Q}_{\theta\rho}} \left( \left| \frac{v - \gamma \ell_{z_0, \theta\rho}^{(h)}}{\theta\rho} \right|^2 + \varphi_M \left( \left| \frac{v - \gamma \ell_{z_0, \theta\rho}^{(h)}}{\theta\rho} \right| \right) \right) dz \\
 &\leq c_{6,3} c_{6,5} \left[ \left( \frac{1}{4\theta} \right)^{n+p_1+2} (2 + \Delta_2(\varphi_M)^3) \varepsilon + C c_{6,4} (1 + \Delta_2(\varphi_M)^3) \theta^2 \right] \Psi_M(\rho) \\
 &\leq [\theta^2 + c_{6,6} \theta^2] \Psi_M(\rho),
 \end{aligned}$$

where  $c_{6,6} = C c_{6,3} c_{6,4} c_{6,5} (1 + \Delta_2(\varphi_M)^3)$ . So, under the smallness assumption that  $\Psi_M(\rho) \leq \varepsilon$ ,

$$\Psi_{1+|D\ell_{\theta\rho}|}(z_0, \theta\rho, \ell_{\theta\rho}) \leq \theta^{2\alpha} \Psi_{1+|D\ell_\rho|}(z_0, \rho, \ell_\rho).$$

Combined with (7.7), we conclude that

$$|D\ell_{\theta\rho}| \leq |D\ell_\rho| + (n+2) \left( \frac{1}{\theta} \right)^{n+3} \Psi_{1+|D\ell_\rho|}(\rho)^{\frac{1}{2}}.$$

□

In the following lemma we will iterate the excess-decay estimate from the previous lemma.

**Lemma 7.2** (Iteration Argument) *Suppose that the assumptions, for  $u$  and  $\varphi$ , in Lemma 7.1 hold. Let  $M_0 > 1$  and  $0 < \alpha < 1$  be given. There exist  $0 < \varepsilon_0 < \theta_0 < 1$  and  $c_7 = c_7(M_0, \theta_0, n, N, L/v, p_0, p_1)$  with the following property: given a standard parabolic cylinder  $\mathcal{Q}_{\rho_0}(z_0) \subseteq \Omega_T$ , if*

$$1 + |D\ell_{z_0, \rho_0}| \leq M_0 \quad \text{and} \quad \Psi_{1+|D\ell_{z_0, \rho_0}|}(z_0, \rho, \ell_{z_0, \rho_0}) \leq \varepsilon_0,$$

then for each  $j \in \mathbb{N}$ , we have the following:

- (a)  $\Psi_{1+|D\ell_{z_0, \theta_0^j \rho_0}|}(z_0, \theta_0^j \rho_0, \ell_{z_0, \theta_0^j \rho_0}) \leq \theta_0^{2j\alpha} \Psi_{1+|D\ell_{z_0, \rho_0}|}(z_0, \rho_0, \ell_{z_0, \rho_0})$ ,
- (b)  $|D\ell_{z_0, \theta_0^j \rho_0}| \leq M_0 - \theta_0^{j\alpha}$ .

Moreover,

$$\begin{aligned}
 &\int_{\mathcal{Q}_r(z_0)} \varphi_{1+|(Du)_{z_0, r}|} (|Du - (Du)_{z_0, r}|) dz \\
 &\leq c_7 \left( \frac{r}{\rho_0} \right)^{2\alpha} \Psi_{1+|D\ell_{z_0, \rho_0}|}(z_0, \rho_0, \ell_{z_0, \rho_0}), \quad \text{for all } 0 < r \leq \rho_0/2.
 \end{aligned} \tag{7.8}$$

**Proof** Parts (a) and (b) follow from an induction argument. With  $0 < \alpha < 1$ ,  $\rho = \rho_0$ , and  $M_0$  fixed, let  $0 < \varepsilon, \theta < 1$  be provided by Lemma 7.1. Put

$$\theta_0 := \theta, \quad \varepsilon_0 := \min \left\{ \varepsilon, \frac{\theta_0^{2(n+3)} (1 - \theta_0^\alpha)^2}{(n+2)^2} \right\}, \quad \text{and} \quad \rho_j = \theta_0^j \rho_0, \quad \text{for each } j \in \mathbb{N}. \tag{7.9}$$

Clearly  $|D\ell_{z_0, \rho_0}| \leq M_0$ , and the base case,  $j = 1$ , immediately follows from Lemma 7.1. With  $j \in \mathbb{N}$  given, suppose that (a) and (b) are both true for all  $k = 1, \dots, j$ . We observe that (a) implies

$$\Psi_{1+|D\ell_{z_0, \rho_j}|}(z_0, \rho_j, \ell_{z_0, \rho_j}) \leq \theta_0^{j\alpha} \Psi_{1+|D\ell_{z_0, \rho_0}|}(z_0, \rho_0, \ell_{z_0, \rho_0}) \leq \varepsilon_0 \leq \varepsilon.$$

By the inductive assumption,

$$|D\ell_{z_0, \rho_j}| \leq M_0 - \theta_0^{j\alpha} \leq M_0.$$

We may therefore use Lemma 7.1, with  $\rho$  replaced with  $\rho_j$  and the other parameters the same as in (7.9), to obtain

$$\begin{aligned} \Psi_{1+|D\ell_{z_0, \rho_{j+1}}|}(z_0, \rho_{j+1}, \ell_{z_0, \rho_{j+1}}) &\leq \theta_0^{2\alpha} \Psi_{1+|D\ell_{z_0, \rho_j}|}(z_0, \rho_j, \ell_{z_0, \rho_j}) \\ &\leq \theta_0^{2(j+1)\alpha} \Psi_{1+|D\ell_{z_0, \rho_0}|}(z_0, \rho_0, \ell_{z_0, \rho_0}). \end{aligned}$$

For part (b), we have

$$\begin{aligned} |D\ell_{z_0, \rho_{j+1}}| &\leq |D\ell_{z_0, \rho_j}| + (n+2) \left(\frac{1}{\theta_0}\right)^{n+3} \Psi_{1+|D\ell_{z_0, \rho_j}|}(z_0, \rho_j, \ell_{z_0, \rho_j})^{\frac{1}{2}} \\ &\leq M_0 - \theta_0^{j\alpha} + (n+2) \left(\frac{1}{\theta_0}\right)^{n+3} \theta_0^{j\alpha} \varepsilon_0^{\frac{1}{2}} \leq M_0 - \theta_0^{j\alpha} + (1 - \theta_0^\alpha) \theta_0^{j\alpha} \\ &= M_0 - \theta^{(j+1)\alpha}. \end{aligned}$$

By induction, we deduce (a) and (b) for all  $j \in \mathbb{N}$ .

It remains to verify (7.8). Given  $0 < r \leq \rho_0/2$ , we may select  $j \in \mathbb{N} \cup \{0\}$  such that  $\rho_{j+1} < 2r \leq \rho_j$ . Using Remark 2.6 and Corollary 4.1, we have

$$\begin{aligned} \int_{Q_r(z_0)} \varphi_{1+|(Du)_{z_0, r}|} (|Du - (Du)_{z_0, r}|) dz &\leq \kappa_2 \int_{Q_r(z_0)} \varphi_{1+|D\ell_{z_0, \rho_j}|} (|Du - D\ell_{z_0, \rho_j}|) dz \\ &\leq \frac{\kappa_2}{\theta_0^{n+2}} \int_{Q_{\rho_j/2}(z_0)} \varphi_{1+|D\ell_{z_0, \rho_j}|} (|Du - D\ell_{z_0, \rho_j}|) dz \\ &\leq c_0 2^{n+p_1+4} \left(\frac{\kappa_2}{\theta_0^{n+2}}\right) \Psi_{1+|D\ell_{z_0, \rho_j}|}(z_0, \rho_j, \ell_{z_0, \rho_j}) \\ &\leq c_0 2^{n+p_1+4} \left(\frac{\kappa_2}{\theta_0^{n+2}}\right) \theta_0^{2j\alpha} \Psi_{1+|D\ell_{z_0, \rho_0}|}(z_0, \rho_0, \ell_{z_0, \rho_0}) \\ &\leq c_7 \left(\frac{r}{\rho_0}\right)^{2\alpha} \Psi_{1+|D\ell_{z_0, \rho_0}|}(z_0, \rho_0, \ell_{z_0, \rho_0}). \end{aligned}$$

Since  $\Delta_2(\varphi, \varphi^*)$  depends on only  $p_1$  and  $p_0$ , the lemma is proved. □

### 8 Partial regularity

We are now in position to prove the main result of the paper.

**Proof of Theorem 1.2** Let  $z_0 \in \Omega_T$  be such that

$$\liminf_{\rho \rightarrow 0} \int_{Q_\rho(z_0)} |V(Du) - (V(Du))_{z_0, \rho}|^2 dz = 0,$$

and

$$\limsup_{\rho \rightarrow 0} |(Du)_{z_0, \rho}| < +\infty. \tag{8.1}$$

Using (2.10), we deduce that

$$\liminf_{\rho \rightarrow 0} \int_{\mathcal{Q}_\rho(z_0)} \varphi_{1+|(Du)_{z_0, \rho}|} (|Du - (Du)_{z_0, \rho}|) dz = 0.$$

Exploiting Lemma 2.7 and Poincaré’s inequality in Theorem 5.1, we get

$$\begin{aligned} & \int_{\mathcal{Q}_\rho(z_0)} \varphi_{1+|D\ell_{z_0, \rho}|} \left( \left| \frac{u - \ell_{z_0, \rho}}{\rho} \right| \right) dz \\ & \leq \kappa_1 \int_{\mathcal{Q}_\rho(z_0)} \varphi_{1+|(Du)_{z_0, \rho}|} \left( \left| \frac{u - (u)_{z_0, \rho} - (Du)_{z_0, \rho}(x - x_0)}{\rho} \right| \right) dz \\ & \leq \kappa_1 c_2 \left[ \int_{\mathcal{Q}_\rho(z_0)} \varphi_{1+|(Du)_{z_0, \rho}|} (|Du - (Du)_{z_0, \rho}|) dz \right. \\ & \quad \left. + \varphi_{1+|(Du)_{z_0, \rho}|} \left( \int_{\mathcal{Q}_\rho(z_0)} \varphi'_{1+|(Du)_{z_0, \rho}|} (|Du - (Du)_{z_0, \rho}|) dz \right) \right]. \end{aligned} \tag{8.2}$$

Thanks to (2.5) and Jensen’s inequality, we may write

$$\begin{aligned} & \int_{\mathcal{Q}_\rho(z_0)} \varphi'_{1+|(Du)_{z_0, \rho}|} (|Du - (Du)_{z_0, \rho}|) dz \\ & \lesssim (\varphi_{1+|(Du)_{z_0, \rho}|}^*)^{-1} \left( \int_{\mathcal{Q}_\rho(z_0)} \varphi_{1+|(Du)_{z_0, \rho}|} (|Du - (Du)_{z_0, \rho}|) dz \right). \end{aligned} \tag{8.3}$$

On the other hand, Theorem 5.3 implies

$$\begin{aligned} & \int_{\mathcal{Q}_\rho(z_0)} \left| \frac{u - \ell_{z_0, \rho}}{\rho} \right|^2 dz \\ & \leq \int_{\mathcal{Q}_\rho(z_0)} \left| \frac{u - (u)_{z_0, \rho} - (Du)_{z_0, 2\rho}(x - x_0)}{\rho} \right|^2 dz \\ & \leq c_4 \left[ \int_{\mathcal{Q}_{2\rho}(z_0)} \varphi_{1+|(Du)_{z_0, 2\rho}|} (|Du - (Du)_{z_0, 2\rho}|) dz + \left( \int_{\mathcal{Q}_{2\rho}(z_0)} |Du - (Du)_{z_0, 2\rho}|^{p_0} dz \right)^{\frac{2}{p_0}} \right. \\ & \quad + \varphi_{1+|(Du)_{z_0, 2\rho}|} \left( \int_{\mathcal{Q}_{2\rho}(z_0)} \varphi'_{1+|(Du)_{z_0, 2\rho}|} (|Du - (Du)_{z_0, 2\rho}|) dz \right) \\ & \quad \left. + \left( \int_{\mathcal{Q}_{2\rho}(z_0)} \varphi'_{1+|(Du)_{z_0, 2\rho}|} (|Du - (Du)_{z_0, 2\rho}|) dz \right)^2 \right]. \end{aligned} \tag{8.4}$$

We can use Lemma 2.5 to control the upper bound’s second integral:

$$\begin{aligned} \int_{\mathcal{Q}_{2\rho}(z_0)} |Du - (Du)_{z_0, 2\rho}|^{p_0} dz & \leq C \left( \frac{1}{\varphi''(1 + |(Du)_{z_0, 2\rho}|)} \int_{\mathcal{Q}_{2\rho}(z_0)} \varphi_{1+|(Du)_{z_0, 2\rho}|} (|Du - (Du)_{z_0, 2\rho}|) dz \right)^{\frac{p_0}{2}} \\ & \quad + C \frac{1}{\varphi_{1+|(Du)_{z_0, 2\rho}|}(1)} \int_{\mathcal{Q}_{z_0, 2\rho}} \varphi_{1+|(Du)_{z_0, 2\rho}|} (|Du - (Du)_{z_0, 2\rho}|) dz. \end{aligned} \tag{8.5}$$

Finally, from (2.13)

$$\begin{aligned}
 |D\ell_{z_0,\rho} - (Du)_{z_0,\rho}| &\leq (n+2) \int_{\mathcal{Q}_\rho(z_0)} \left| \frac{u - (u)_{z_0,\rho} - (Du)_{z_0,\rho}(x - x_0)}{\rho} \right| dz \\
 &\lesssim (\varphi_{1+|(Du)_{z_0,\rho}|})^{-1} \left( \int_{\mathcal{Q}_\rho(z_0)} \varphi_{1+|(Du)_{z_0,\rho}|} \left( \left| \frac{u - (u)_{z_0,\rho} - (Du)_{z_0,\rho}(x - x_0)}{\rho} \right| \right) dz \right), \tag{8.6}
 \end{aligned}$$

which in turn can be bounded via (8.2) and (8.3). Let  $\varepsilon_0 > 0$  be as defined in (7.9). Keeping in mind the definition of  $z_0$ , the estimates (8.2), (8.4), and (8.6), supported by (8.1), (8.3), and (8.5), imply the existence of  $M_0 > 1$  and a radius  $R_0 > 0$  such that  $|D\ell_{z_0,R_0}| < M_0 - 1$  and  $\Psi_{1+|D\ell_{z_0,R_0}|}(z_0, R_0, \ell_{z_0,R_0}) < \varepsilon_0$ . By the absolute continuity of the integrals, there exists  $R_1 < R_0$  such that, for any  $z \in \mathcal{Q}_{R_1}(z_0)$  we have

$$1 + |D\ell_{z,R_0}| < M_0 \quad \text{and} \quad \Psi_{1+|D\ell_{z,R_0}|}(z, R_0, \ell_{z,R_0}) < \varepsilon_0.$$

Applying Lemma 7.2 to each point  $z \in \mathcal{Q}_{R_1}(z_0)$ , we deduce that, for any  $r \leq R_0/2$ ,

$$\begin{aligned}
 \int_{\mathcal{Q}_r(z)} |V(Du) - (V(Du))_{z,r}|^2 dz &\sim \int_{\mathcal{Q}_r(z)} \varphi_{1+|(Du)_{z,r}|} (|Du - (Du)_{z,r}|) dz \\
 &\leq C(M, \theta_0) \frac{r^{n+2+2\alpha}}{R_0^{2\alpha}} \varepsilon_0.
 \end{aligned}$$

This means that  $V(Du)$  belongs to the parabolic Campanato space  $\mathcal{L}^{2, \frac{2\alpha}{n+2}}(\mathcal{Q}_{R_1}(z_0), \mathbb{R}^{Nn})$  and by the usual embedding we have  $V(Du) \in C^{0, \frac{\alpha}{2}, \alpha}(\mathcal{Q}_{R_1}(z_0), \mathbb{R}^{Nn})$ . □

**Remark 8.1** Note, as indicated in Sect. 1, the Hölder continuity of  $V(Du)$  implies the Hölder continuity of  $Du$  with a different exponent depending on  $\varphi$ .

**Funding** Open access funding provided by Università degli Studi di Napoli Federico II within the CRUI-CARE Agreement.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Baroni, P.: Regularity in parabolic Dini continuous systems. *Forum Math.* **23**, 1093–1112 (2011)
2. Bögelein, V., Duzaar, F.: Higher integrability for parabolic systems with non-standard growth and degenerate diffusions. *Publ. Mat.* **1**, 201–250 (2011)
3. Bögelein, V., Duzaar, F., Marcellini, P.: Parabolic systems with  $p, q$ -growth: a variational approach. *Arch. Ration. Mech. Anal.* **210**(1), 219–267 (2013)
4. Bögelein, V., Duzaar, F., Mingione, G.: The boundary regularity of non-linear parabolic systems I. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **27**(1), 201–255 (2010)
5. Bögelein, V., Duzaar, F., Mingione, G.: The regularity of general parabolic systems with degenerate diffusion. *Mem. Amer. Math. Soc.* **221**(1041), vi+143 pp (2013)
6. Bögelein, V., Foss, M., Mingione, G.: Regularity in parabolic systems with continuous coefficients. *Math. Z.* **270**(3–4), 903–938 (2012)

7. Breit, D., Verde, A.: Quasiconvex variational functionals in Orlicz-Sobolev spaces. *Ann. Mat. Pura Appl.* (4) **192**(2), 255–271 (2013)
8. Burczak, J., Kaplicky, P.: Evolutionary, symmetric  $p$ -Laplacian Interior regularity of time derivatives and its consequences. *Commun. Pure Appl. Anal.* **15**(6), 2401–2445 (2016)
9. Campanato, S.: Equazioni paraboliche del secondo ordine e spazi  $L^{2,\theta}(\Omega, \delta)$ . *Ann. Mat. Pura Appl.* **73**, 55–102 (1966)
10. Cianchi, A., Maz'ya, V.G.: Global Lipschitz regularity for a class of quasilinear elliptic equations. *Commun. Partial Diff. Equ.* **36**(1), 100–133 (2010)
11. Cianchi, A., Maz'ya, V.G.: Global boundedness of the gradient for a class of nonlinear elliptic systems. *Arch. Ration. Mech. Anal.* **212**(1), 129–177 (2014)
12. Celada, P., Ok, J.: Partial regularity for non-autonomous degenerate quasi-convex functionals with general growth. *Nonlinear Anal.* **194**, 36 (2020)
13. DiBenedetto, E.: Degenerate parabolic equations. Universitext, Springer, New York (1993)
14. DiBenedetto, E., Friedman, A.: Regularity of solutions of nonlinear degenerate parabolic systems. *J. Reine Angew. Math.* **349**, 83–128 (1984)
15. DiBenedetto, E., Friedman, A.: Hölder estimates for nonlinear degenerate parabolic systems. *J. Reine Angew. Math.* **357**, 1–22 (1985)
16. Diening, L., Ettwein, F.: Fractional estimates for non-differentiable elliptic systems with general growth. *Forum Math.* **20**, 523–556 (2008)
17. Diening, L., Kreuzer, C.: Linear convergence of an adaptive finite element method for the  $p$ -Laplacian equation. *SIAM J. Num. Anal.* **46**(2), 614–638 (2008)
18. Diening, L., Lengeler, D., Stroffolini, B., Verde, A.: Partial regularity for minimizer of quasi-convex functional with general growth. *SIAM J. Math. Anal.* **44**(5), 3594–3616 (2012)
19. Diening, L., Scharle, T., Schwarzacher, S.: Regularity for parabolic systems of Uhlenbeck type with Orlicz growth. *J. Math. Anal. Appl.* **472**(1), 46–60 (2019)
20. Diening, L., Schwarzacher, S., Stroffolini, B., Verde, A.: Parabolic Lipschitz truncation and caloric approximation. *Calc. Var. Partial Diff. Eq.* **56**(4), 1–27 (2017)
21. Diening, L., Stroffolini, B., Verde, A.: Everywhere regularity of functionals with  $\phi$ -growth. *Manuscripta Math.* **129**(4), 449–481 (2009)
22. Diening, L., Stroffolini, B., Verde, A.: The  $\phi$ -harmonic approximation and the regularity of  $\phi$ -harmonic maps. *J. Differ. Equ.* **253**(7), 1943–1958 (2012)
23. Donaldson, T.: Nonlinear elliptic boundary value problems in Orlicz-Sobolev spaces. *J. Differ. Equ.* **10**(3), 507–528 (1971)
24. Donaldson, T.: Inhomogeneous Orlicz-Sobolev spaces and nonlinear parabolic initial value problems. *J. Differ. Equ.* **16**(2), 201–256 (1974)
25. Duzaar, F., Mingione, G.: Second order parabolic systems, optimal regularity, and singular sets of solutions. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **22**(6), 705–751 (2005)
26. Duzaar, F., Mingione, G., Steffen, K.: Parabolic systems with polynomial growth and regularity. *Mem. Amer. Math. Soc.* **214**(1005), x+118 pp (2011)
27. Foss, M., Geisbauer, J.: Partial regularity for subquadratic parabolic systems with continuous coefficients. *Manuscripta Math.* **139**, 1–47 (2012)
28. Giaquinta, M., Giusti, E.: Partial regularity for the solutions to nonlinear parabolic system. *Annali di Matematica Pura ed Applicata* **97**(1), 253–266 (1973)
29. Giusti, E.: Direct methods in the Calculus of Variations, World Scientific Publishing Co., Inc., River Edge, NJ, 2003. viii+403 pp
30. Harjulehto, P., Hästö, P.: Orlicz Spaces and Generalized Orlicz Spaces. Springer, Berlin (2019)
31. Hästö, P., Ok, J.: Higher integrability for parabolic systems with Orlicz growth. *J. Diff. Equ.* **300**, 925–948 (2021)
32. Isernia, T.:  $L^\infty$ -regularity for a wide class of parabolic systems with general growth. *Proc. Amer. Math. Soc.* **146**(11), 4741–4753 (2018)
33. Krasnoselskiĭ, M. A., Rutickiĭ, Ya. B.: Convex functions and Orlicz Spaces, P. Noordhoff Ltd., Groningen, 1961. Translated from the first Russian edition by L. F. Boron
34. Lieberman, G.M.: Hölder regularity for the gradients of solutions of degenerate parabolic systems. *Ukr. Mat. Visn.* **3**(3), 352–373 (2006)
35. Marcellini, P.: Regularity of minimizers of integrals of the calculus of variations with non standard growth conditions. *Arch. Ration. Mech. Anal.* **105**(3), 267–284 (1989)
36. Marcellini, P.: Regularity and existence of solutions of elliptic equations with  $p, q$ -growth conditions. *J. Diff. Equ.* **90**, 1–30 (1991)
37. Marcellini, P.: Growth conditions and regularity for weak solutions to nonlinear elliptic pdes. *J. Math. Anal. Appl.* **501**(1), 1–32 (2021)

38. Marcellini, P., Papi, G.: Nonlinear elliptic systems with general growth. *J. Diff. Equ.* **221**, 412–443 (2006)
39. Mons, L.: Partial regularity for parabolic systems with VMO-coefficients. *Commun. Pure. Appl. Anal.* **20**, 1783–1820 (2021)
40. Rao, M., Ren, Z. D.: *Theory of Orlicz spaces. Monographs and Textbooks in Pure and Applied Mathematics 146* Marcel Dekker, Inc., New York (1991)
41. Scheven, C.: Partial regularity for subquadratic parabolic systems by  $\mathcal{A}$ -caloric approximation. *Rev. Mat. Iberoam.* **27**, 751–801 (2011)
42. Simon, L.: *Theorems on Regularity and Singularity of Energy Minimizing Maps.* Birkhäuser Verlag, Basel (1996)
43. Visek, M.I.: Solvability of the first boundary value problem for quasilinear equations with rapidly increasing coefficients in Orlicz classes. *Dokl. Akad. Nauk. SSSR* **151**, 758–761 (1963)
44. You, J.: Regularity of solutions of certain parabolic system with nonstandard growth condition. *Acta Math. Sinica (N.S.)* **14**(2), 145–160 (1998)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.