PARTIAL REGULARITY FOR DEGENERATE PARABOLIC SYSTEMS WITH GENERAL GROWTH VIA CALORIC APPROXIMATIONS

JIHOON OK, GIOVANNI SCILLA, AND BIANCA STROFFOLINI

Abstract. We establish a partial regularity result for solutions of parabolic systems with general φ-growth, where φ is an Orlicz function. In this setting we can develop a unified approach that is independent of the degeneracy of system and relies on two caloric approximation results: the φ-caloric approximation, which was introduced in [19], and an improved version of the $\mathcal{A}$-caloric approximation, which we prove without using the classical compactness method.

1. Introduction

The aim of this paper is to prove partial regularity for solutions of the following autonomous parabolic system:

$$\partial_t \mathbf{u} - \operatorname{div} \mathbf{A}(D\mathbf{u}) = \mathbf{0} \quad \text{in} \;\; \Omega \times (0,T], \tag{1.1}$$

where $\mathbf{A} \in C(\mathbb{R}^{Nn}, \mathbb{R}^{Nn}) \cap C^1(\mathbb{R}^{Nn} \setminus \{\mathbf{0}\}, \mathbb{R}^{Nn})$ is modeled as the φ-Laplacian, see Assumption (A). We would like to point out that, already in the stationary case, the best result we can expect for non-radial systems is the $C^{1,\alpha}$-regularity outside a set of Lebesgue measure zero, see the survey [39] and references therein.

In this direction, a powerful tool is the comparison and closeness with suitable smooth maps, for which excess decay estimates are available. The first use of a compactness argument for approximately harmonic maps goes back to De Giorgi, in the context of regularity of minimal surfaces in geometric measure theory, see [22, 44]. De Giorgi's Lemma states that there is a rigidity behaviour of approximately harmonic maps, in the sense that they are close to harmonic ones. This Lemma has been generalized to strongly elliptic bilinear forms, the so-called $\mathcal{A}$-harmonic approximation, in [28] for applications to the boundary regularity of minimizing elliptic currents. For elliptic systems and related quasiconvex functionals of p-growth, we refer to for instance [28, 23, 24, 1, 30, 2, 7]. Here partial regularity results are proved, using a two-scale approach. As long as the excess functional is small compared to the gradient average in a ball, one can linearize the system via the $\mathcal{A}$-harmonic approximation Lemma. When, instead, the system is degenerate, one compares with the p-Laplacian via the p-harmonic approximation, see [25] and also [3]. The final partial regularity result is then achieved with an exit time argument. In a more general setting, the $\mathcal{A}$-harmonic approximation in Orlicz spaces and the φ-harmonic approximation were proven in [15] and [21], respectively, by utilizing a refined Lipschitz truncation argument. Using these approximation results, in recent years, partial regularity for elliptic systems or quasiconvex functionals with general growth has been studied in [10, 43, 35, 31, 40].

Regularity results for the evolutive p-Laplacian system

$$\partial_t \mathbf{u} - \operatorname{div}\left(|D\mathbf{u}|^{p-2} D\mathbf{u}\right) = \mathbf{0} \quad \text{in} \;\; \Omega_T$$

Key words and phrases. degenerate parabolic systems, partial regularity, caloric approximation, Orlicz growth.

MSC(2020): Primary: 35K40, 35B65; Secondary: 35K92, 46E30.

were established by DiBenedetto and Friedman in [11, 12]. Their key idea was to look at the system in a new geometry, that, in a sense, reduces the system to the classical heat system. Roughly speaking, if the average of the gradient of a solution is locally comparable with λ, the system looks like

$$\partial_t \mathbf{u} = \operatorname{div}(\lambda^{p-2} D\mathbf{u})\,.$$

This suggests to consider a "new metric" in which the scaling is homogeneous and to consider "balls" centered at the point $z_0 = (x_0, t_0)$ with respect to this metric; i.e., the cylinders:

$$Q_r^\lambda(z_0) := B_\rho(x_0) \times (t_0 - \lambda^{2-p} r^2, t_0 + \lambda^{2-p} r^2)\,.$$

This method is known as the *intrinsic scaling* method. As for the evolutive Uhlenbeck system with general φ-growth

$$\partial_t \mathbf{u} - \operatorname{div}\left(\frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|} D\mathbf{u}\right) = \mathbf{0} \quad \text{in} \;\; \Omega_T\,, \tag{1.2}$$

everywhere $C^{1,\alpha}$-regularity is finally established in our recent paper [41] using cylinders that are intrinsic with respect to the function φ, namely,

$$Q_r^\lambda(z_0) := B_r(x_0) \times \left(t_0 - \frac{\lambda^2}{\varphi(\lambda)} r^2, t_0 + \frac{\lambda^2}{\varphi(\lambda)} r^2\right)\,.$$

We also refer to [37, 18] for related regularity results for the system (1.2).

Partial regularity for parabolic systems using caloric approximations has been studied in [26, 4, 5, 27, 6, 42]. In particular, in [6] Bögelein, Duzaar and Mingione obtained the $\mathcal{A}$-caloric and p-caloric approximations, and using these proved partial Hölder continuity of the gradient of weak solutions to the degenerate parabolic system (1.1) with standard p-growth. Let us review the proof of partial regularity in [6]. By assuming a smallness condition on the relevant excess at some scale, it is possible to linearize the system at the gradient average in the nondegenerate case, and compare the solution of the original system with one of the linearized system. The comparison argument ensures the decay estimate for the excess at smaller scales. In the degenerate regime, one compares with a suitable p-Laplace evolutive system via the p-caloric approximation. At this stage, one can proceed using the intrinsic cylinders á la Di Benedetto. Finally, the degenerate and nondegenerate regimes are matching together keeping track of the so called "switching radius".

In this paper we consider degenerate parabolic systems with general φ-growth and obtain partial Hölder continuity of the gradient of weak solutions. Our result covers a large class of systems whose degeneracy need not to be determined. In particular, we extend the results of the subquadratic and superquadratic systems obtained in [6].

We emphasize that our method, inspired by [6], deals with systems in a unified way, without any distinction between the superquadratic and subquadratic cases. The main tools are some caloric approximations in the Orlicz setting. In the nondegenerate regime, we prove a new version of the $\mathcal{A}$-caloric approximation which is more suitable to our setting. We would like to mention a recent related result in [29], where the authors obtained an $\mathcal{A}$-caloric approximation using the classical compactness method, and, applying this, they proved partial Hölder regularity for nondegenerate parabolic systems with general growth. The proof of our version, stated in Theorem 3.8, relies on a parabolic duality argument and an improved parabolic Lipschitz truncation, along the lines of [15] for the elliptic case. Consequently, we can obtain closeness with comparing mappings in terms of gradients directly, which is sharper than the one considered in [29]. Moreover, we underline explicitly how the constant δ in Theorem 3.8 depends only on p and q instead of φ. As for the degenerate regime, we use the φ-caloric approximation lemma proven in [19]. Now, we state the main theorem of our paper.

Theorem 1.1. *Let $\mathbf{u}$ be a weak solution to (1.1) where $\mathbf{A}$ satisfies Assumption (A). There exist $U \subset \Omega_T$ with $|\Omega_T \setminus U| = 0$ and $\alpha \in (0,1)$ such that $D\mathbf{u} \in C^{\alpha,\frac{\alpha}{2}}_{\mathrm{loc}}(U)$. Moreover, we have $(\Omega_T \setminus U) \subset (\Sigma_1 \cup \Sigma_2)$, where*

$$\Sigma_1 := \left\{ z_0 \in \Omega_T : \liminf_{r \to 0^+} \fint_{Q_r(z_0)} |\mathbf{V}(D\mathbf{u}) - (\mathbf{V}(D\mathbf{u}))_{Q_r(z_0)}|^2 \,\mathrm{d}z > 0 \right\} \tag{1.3}$$

and

$$\Sigma_2 := \left\{ z_0 \in \Omega_T : \limsup_{r \to 0^+} |(D\mathbf{u})_{Q_r(z_0)}| = \infty \right\} . \tag{1.4}$$

Overview of the paper. In Section 2, we fix the basic notation and collect some definitions and results on Orlicz functions. In Section 3, we prove the $\mathcal{A}$-caloric approximation and recall the φ-caloric one. In Section 4, we obtain the Caccioppoli inequality and the higher integrability. In Sections 5 and 6, we consider the nondegenerate and degenerate regimes, respectively. In Section 7, we perform the iteration procedure and then prove the main theorem, Theorem 1.1.

2. Preliminaries

2.1. **Notation.** For $z_0 = (x_0, t_0) \in \mathbb{R}^n \times \mathbb{R}$ and $r, \tau > 0$, we define

$$Q_{r,\tau}(z_0) := B_r(x_0) \times (t_0 - \tau, t_0 + \tau) ,$$

where $B_r(x_0)$ is the open ball in $\mathbb{R}^n$ centered at x_0 with radius r. In particular, we write $Q_r(z_0) := Q_{r,r^2}(z)$ which is the usual parabolic cylinder. Moreover, for a given function $\varphi : (0,\infty) \to (0,\infty)$ and $\lambda > 0$, we write $Q_r^\lambda(z_0) := Q_{r,\tau}(z_0)$ with $\tau = \frac{\lambda^2}{\varphi(\lambda)} r^2$. If the center z_0 is the origin or is not important, we omit writing the center of the cylinders. The notation $f \lesssim g$ or $f \sim g$ means that there exists constant $c \geqslant 1$ such that $f \leqslant cg$ or $\frac{1}{c} f \leqslant g \leqslant cf$. We write the average of a function f on $Q_r(z_0)$ and on $Q_r^\lambda(z_0)$ as

$$(f)_r = (f)_{Q_r(z_0)} := \fint_{Q_r(z_0)} f \,\mathrm{d}z \quad \text{and} \quad (f)_r^\lambda = (f)_{Q_r^\lambda(z_0)} := \fint_{Q_r^\lambda(z_0)} f \,\mathrm{d}z ,$$

respectively.

2.2. **Orlicz functions and related operators.** In this paper, $\varphi : [0,\infty) \to [0,\infty)$ is always an N-function, that is, $\varphi(0) = 0$, there exists a right continuous derivative φ' of φ, φ' is increasing with $\varphi'(0) = 0$ and $\varphi'(t) > 0$ when $t > 0$. For simplicity, we shall assume that

$$\varphi(1) = 1 .$$

Note that if we do not assume this condition, constants c in this paper may depend on $\varphi(1)$. We say that φ satisfies the Δ_2 condition denoted by $\Delta(\varphi) < \infty$ if there exists a positive constant $K =: \Delta(\varphi)$ such that $\varphi(2t) \leqslant K\varphi(t)$ for all $t > 0$. The conjugate function of φ is defined as

$$\varphi^*(t) := \sup_{s \geqslant 0} (st - \varphi(s)) . \tag{2.1}$$

From the definition, the following Young's inequality

$$st \leqslant \varphi(t) + \varphi^*(s) , \quad s, t \geqslant 0 , \tag{2.2}$$

holds true. From now on we always assume that φ and φ^* satisfy the Δ_2 condition and this is denoted by $\Delta(\varphi,\varphi^*)<\infty$, where $\Delta(\varphi,\varphi^*)$ denotes the relevant constants K. We note that the exact value of φ^* is not always explicitly computable and instead the estimate

$$\varphi^*\left(\frac{\varphi(t)}{t}\right)\sim\varphi^*(\varphi'(t))\sim\varphi(t) \tag{2.3}$$

will often be useful in computations (see [32, Theorem 2.4.10]).

If φ satisfies $\Delta(\varphi,\varphi^*)$, we define the Orlicz space $L^\varphi(\Omega,\mathbb{R}^N)$ as the set of all measurable functions $f:\Omega\to\mathbb{R}^N$ such that

$$\int_\Omega \varphi(|f(x)|)\,\mathrm{d}x<\infty,$$

and the Orlicz-Sobolev space $W^{1,\varphi}(\Omega,\mathbb{R}^N)$ as the set of all $f\in L^\varphi(\Omega,\mathbb{R}^N)\cap W^{1,1}(\Omega,\mathbb{R}^N)$ such that

$$\int_\Omega \varphi(|Df(x)|)\,\mathrm{d}x<\infty.$$

$L^\varphi(\Omega,\mathbb{R}^N)$ and $W^{1,\varphi}(\Omega,\mathbb{R}^N)$ are endowed with the usual Luxembourg type norms. Then they are reflexive Banach spaces. Moreover, for an interval I in $\mathbb{R}$, the parabolic space $L^\varphi(I;W^{1,\varphi}(\Omega,\mathbb{R}^N))$ or $L^\varphi(I;W^{1,\varphi}_0(\Omega,\mathbb{R}^N))$ denotes the set of all functions $f:\Omega\times I\to\mathbb{R}^N$ such that $f(\cdot,t)\in W^{1,\varphi}(\Omega,\mathbb{R}^N)$ or $f(\cdot,t)\in W^{1,\varphi}_0(\Omega,\mathbb{R}^N)$ for a.e. $t\in I$ and

$$\int_I\int_\Omega \varphi(|Df(x,t)|)\,\mathrm{d}x\,\mathrm{d}t<\infty.$$

We recall Jensen type inequality in [33, Lemma 2.2].

Lemma 2.1. *If $\psi:[0,\infty)\to[0,\infty]$ is increasing with $\psi(0)=0$ and satisfies that $\psi(t)/t\leqslant L\psi(s)/s$ for every $0\leqslant t\leqslant s$ with constant $L\geqslant 1$, then*

$$\psi\left(\frac{1}{L^2}\fint_U |f|\,dz\right)\leqslant \fint_U \psi(|f|)\,dz.$$

Now we further assume for φ that

$$\frac{2n}{n+2}<p\leqslant\frac{t\varphi''(t)}{\varphi'(t)}+1\leqslant q,\quad \text{for all } t\in(0,\infty). \tag{2.4}$$

Without loss of generality, we always assume that $p<2<q$. Note that this implies

$$1<p\leqslant\frac{t\varphi'(t)}{\varphi(t)}\leqslant q\,,\quad t>0\,. \tag{2.5}$$

and hence the Δ_2 conditions of φ and φ^*. Then we define vector valued functions $\mathbf{V}:\mathbb{R}^{Nn}\to\mathbb{R}^{Nn}$ by

$$\mathbf{V}(\mathbf{Q}):=\sqrt{\frac{\varphi'(|\mathbf{Q}|)}{|\mathbf{Q}|}}\mathbf{Q}. \tag{2.6}$$

Then we recall equivalent relations in [14, Lemmas 3 and 20] and [18, Lemma 3.1]:

$$\frac{\varphi'(|\mathbf{P}|+|\mathbf{Q}|)}{|\mathbf{P}|+|\mathbf{Q}|}|\mathbf{P}-\mathbf{Q}|^2\sim|\mathbf{V}(\mathbf{P})-\mathbf{V}(\mathbf{Q})|^2\sim\varphi_{|\mathbf{Q}|}(|\mathbf{P}-\mathbf{Q}|)\,, \tag{2.7}$$

$$\frac{\varphi'(|\mathbf{P}|+|\mathbf{Q}|)}{|\mathbf{P}|+|\mathbf{Q}|}\sim\int_0^1\frac{\varphi'(|\tau\mathbf{P}+(1-\tau)\mathbf{Q}|)}{|\tau\mathbf{P}+(1-\tau)\mathbf{Q}|}\,\mathrm{d}\tau\,, \tag{2.8}$$

and

$$|\mathbf{A}(\mathbf{P}) - \mathbf{A}(\mathbf{Q})| \sim \varphi'_{|\mathbf{Q}|}(|\mathbf{P} - \mathbf{Q}|)\,. \tag{2.9}$$

Moreover, by the same proof of [16, Lemma A.2], we have that for every $\mathbf{g} \in L^{\varphi}(Q_r; \mathbb{R}^{Nn})$,

$$\fint_{Q_r} |\mathbf{V}(\mathbf{g}) - (\mathbf{V}(\mathbf{g}))_{Q_r}|^2 \,\mathrm{d}z \sim \fint_{Q_r} |\mathbf{V}(\mathbf{g}) - \mathbf{V}((\mathbf{g})_{Q_r})|^2 \,\mathrm{d}z \tag{2.10}$$

Finally we recall the following Young type inequality from [34, Proposition 3.8 (3)]: for every $\varepsilon \in (0,1)$

$$\varphi(|\mathbf{P} - \mathbf{Q}|) \leqslant \varepsilon\,(\varphi(|\mathbf{P}|) + \varphi(|\mathbf{Q}|)) + c\varepsilon^{-1}|\mathbf{V}(\mathbf{P}) - \mathbf{V}(\mathbf{Q})|^2\,. \tag{2.11}$$

Note that all constants concerned with the relation $\sim$ and c in above depend only on p and q.

2.3. **Shifted N-functions.** The following definitions and results about shifted N-functions can be found in [14, 20].

For an N-function φ and for $a \geqslant 0$, we define the shifted N-function φ_a by

$$\varphi_a(t) := \int_0^t \frac{\varphi'(a+s)s}{a+s}\,\mathrm{d}s \quad \left(\text{i.e., } \varphi'_a(t) = \frac{\varphi'(a+t)}{a+t}t\right).$$

We note that if φ satisfies (2.4) then φ_a also satisfies (2.4) uniformly in $a \geqslant 0$ with the same p and q.

Under assumption (2.4) on φ, we have the following relations (see, e.g., [10, Proposition 2.3] and [20]), which hold uniformly with respect to $a \geqslant 0$:

$$\varphi_a(t) \sim \varphi'_a(t)\,t\,; \tag{2.12}$$

$$\varphi_a(t) \sim \varphi''(a+t)t^2 \sim \frac{\varphi(a+t)}{(a+t)^2}t^2 \sim \frac{\varphi'(a+t)}{a+t}t^2\,, \tag{2.13}$$

$$\varphi(a+t) \sim [\varphi_a(t) + \varphi(a)]\,. \tag{2.14}$$

The following lemma (see [17, Corollary 26]) deals with the *change of shift* for N-functions.

Lemma 2.2 (change of shift)**.** *Let φ be an N-function with $\Delta_2(\varphi), \Delta_2(\varphi^*) < \infty$. Then for any $\eta > 0$ there exists $c_\eta > 0$, depending only on η and $\Delta_2(\varphi)$, such that for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$ and $t \geqslant 0$*

$$\varphi_{|\mathbf{a}|}(t) \leqslant c_\eta \varphi_{|\mathbf{b}|}(t) + \eta\varphi_{|\mathbf{a}|}(|\mathbf{a} - \mathbf{b}|)\,. \tag{2.15}$$

2.4. **Assumption and weak solution.** We state the assumption of the main theorem, Theorem 1.1, and the definition of weak solution to (1.1).

Assumption (A). The operator $\mathbf{A}$ verifies the following assumptions with constants $0 < \nu \leqslant 1 \leqslant L$ and an N-function $\varphi \in C^1([0,\infty)) \cap C^2((0,\infty))$ satisfying (2.4).

(A1) (φ-growth condition)

$$|\mathbf{A}(\mathbf{P})| + |D\mathbf{A}(\mathbf{P})||\mathbf{P}| \leqslant L\varphi'(|\mathbf{P}|)\,, \tag{2.16}$$
$$\big[D\mathbf{A}(\mathbf{P})(\mathbf{a} \otimes \mathbf{b})\big] : (\mathbf{a} \otimes \mathbf{b}) \geqslant \nu\varphi''(|\mathbf{P}|)|\mathbf{a}||\mathbf{b}|\,,$$

for all $\mathbf{P} \in \mathbb{R}^{Nn} \setminus \{\mathbf{0}\}$, $\mathbf{a} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^N$.

(A2) (Off diagonal condition on $\mathbf{A}$ and φ)

$$|D\mathbf{A}(\mathbf{P}) - D\mathbf{A}(\mathbf{Q})| + |D^2(\varphi(|\mathbf{P}|)) - D^2(\varphi(|\mathbf{Q}|))| \leqslant L\left(\frac{|\mathbf{P} - \mathbf{Q}|}{|\mathbf{P}|}\right)^{\gamma}\varphi''(|\mathbf{P}|) \tag{2.17}$$

for some $\gamma \in (0,1)$, all $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{Nn}$ with $|\mathbf{P} - \mathbf{Q}| \leqslant \frac{1}{2}|\mathbf{P}|$, $\mathbf{a} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^N$.

(A3) (Almost φ-isotropic condition near the origin) For every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that

$$\left| \mathbf{A}(\mathbf{P}) - \frac{\varphi'(|\mathbf{P}|)}{|\mathbf{P}|}\mathbf{P} \right| \leqslant \varepsilon\varphi'(|\mathbf{P}|) \tag{2.18}$$

for all $\mathbf{P} \in \mathbb{R}^{Nn}$ with $|\mathbf{P}| \leqslant \delta$.

We note that the assumption (i) implies the following monotonicity

$$(\mathbf{A}(\mathbf{P}) - \mathbf{A}(\mathbf{Q})) : (\mathbf{P} - \mathbf{Q}) \geqslant \tilde{\nu}\varphi''(|\mathbf{P}| + |\mathbf{Q}|)|\mathbf{P} - \mathbf{Q}|^2, \quad \mathbf{P}, \mathbf{Q} \in \mathbb{R}^{Nn}. \tag{2.19}$$

for some $\tilde{\nu} = \tilde{\nu}(\nu, L) > 0$.

A function $\mathbf{u} = (u^1, u^2, \dots, u^N) \in C_{\mathrm{loc}}(0, T; L^2_{\mathrm{loc}}(\Omega, \mathbb{R}^N)) \cap L^{\varphi}_{\mathrm{loc}}(0, T; W^{1,\varphi}_{\mathrm{loc}}(\Omega, \mathbb{R}^N))$ is said to be a (local) *weak solution* to (1.1) if it satisfies the following weak form of (1.1):

$$-\int_{\Omega_T} \mathbf{u} \cdot \zeta_t \,\mathrm{d}z + \int_{\Omega_T} \mathbf{A}(D\mathbf{u}) : D\zeta \,\mathrm{d}z = 0 \quad \text{for all} \;\; \zeta \in C^{\infty}_{\mathrm{c}}(\Omega_T, \mathbb{R}^N)\,,$$

where "$\cdot$" and ":" are the Euclidean inner products in $\mathbb{R}^N$ and $\mathbb{R}^{Nn}$, respectively. By the density of smooth functions in Orlicz-Sobolev spaces and a standard approximation argument one can see that the weak solution $\mathbf{u}$ to (1.1) also satisfies for every $0 < t_1 < t_2 \leqslant T$,

$$\int_{\Omega'} \mathbf{u} \cdot \zeta(x, t) \,\mathrm{d}x \bigg|_{t=t_1}^{t=t_2} + \int_{\Omega'} \int_{t_1}^{t_2} [-\mathbf{u} \cdot \zeta_t + \mathbf{A}(D\mathbf{u}) : D\zeta] \,\mathrm{d}t \,\mathrm{d}x = 0$$

for all $\zeta \in W^{1,2}([t_1, t_2]; L^2(\Omega', \mathbb{R}^N)) \cap L^{\varphi}([t_1, t_2]; W^{1,\varphi}_0(\Omega', \mathbb{R}^N))$ and $\Omega' \Subset \Omega$.

Remark 2.3. The weak solution $\mathbf{u}$ to (1.1) is not weakly differentiable in t. Consequently, we are unable to employ a test function ζ that directly involves the weak solution. However, this problem can be overcome by utilizing an approximation method known as the Steklov average, as described in [13, I. 3-(i) and II. Proposition 3.1]. This technique has become a standard approach for addressing such problems. Henceforth, we shall assume that $\mathbf{u}$ is differentiable and proceed to consider test functions that involve the weak solution without further explicit clarification.

3. $\mathcal{A}$-caloric and φ-caloric approximations

In this section, we introduce two caloric approximations. They play a crucial role in the proof of partial regularity. φ-caloric approximation was obtained in [19] and we just recall it. On the other hand, we derive a new version of $\mathcal{A}$-caloric approximation with gradient estimates by using parabolic duality and Lipschitz truncation. In Sections 3.1 and 3.2, we obtain auxiliary lemmas for $\mathcal{A}$-caloric approximation.

3.1. **Regularity estimates for linear systems with constant coefficients.** We introduce Lipschitz estimates for $\mathcal{A}$-caloric maps and Calderón-Zygmund estimates for parabolic linear systems with constant coefficient $\mathcal{A}$. Let $\mathcal{A} = (\mathcal{A}^{\alpha\beta}_{ij}) \in \mathbb{R}^{N^2n^2}$ satisfy the *Legendre-Hadamard condition*: for every $\mathbf{a} = (a^{\alpha}) \in \mathbb{R}^N$ and $\mathbf{b} = (b_i) \in \mathbb{R}^n$,

$$\mathcal{A}(\mathbf{a} \otimes \mathbf{b}) \,:\, (\mathbf{a} \otimes \mathbf{b}) = \mathcal{A}^{\alpha\beta}_{ij} a^{\alpha} a^{\beta} b_i b_j \geqslant \nu|\mathbf{a}|^2|\mathbf{b}|^2$$

for some $\nu > 0$. Then a weak solution $\mathbf{h} : Q_r \to \mathbb{R}^N$ to the linear system with coefficient $\mathcal{A}$

$$\mathbf{h}_t - \operatorname{div}(\mathcal{A}D\mathbf{h}) = \mathbf{0} \quad \text{in} \;\; Q_r$$

is called an $\mathcal{A}$*-caloric map*. By standard regularity theory, see for instance [9], $\mathbf{v} \in C^\infty(Q_r, \mathbb{R}^N)$, and in particular we have the following Lipschitz estimate and excess decay estimate, which will be used in Section 5.

Lemma 3.1. *Suppose* $\mathbf{h} \in C^\infty(Q_r, \mathbb{R}^N)$ *is an* $\mathcal{A}$*-caloric map in* Q_r*. Then we have that*

$$\sup_{Q_{r/2}} |D\mathbf{h}| \leqslant c \fint_{Q_r} |D\mathbf{h}| \,\mathrm{d}z\,. \tag{3.1}$$

Moreover, for every $\theta \in (0,1)$,

$$\fint_{Q_{\theta r}} |D\mathbf{h} - (D\mathbf{h})_{\theta r}| \,\mathrm{d}z \leqslant c\theta \fint_{Q_r} |D\mathbf{h} - (D\mathbf{h})_r| \,\mathrm{d}z\,. \tag{3.2}$$

Proof. In view of [9, (5.9)–(5.12)], one can have

$$\sup_{Q_{r/2}}(|\mathbf{h}| + r|D\mathbf{h}| + r^2|D^2\mathbf{h}|) \leqslant c \fint_{Q_r} |\mathbf{h}| \mathrm{d}z\,.$$

Since every $\mathbf{h}_{x_i}$, $i = 1, 2, \dots, n$, is also $\mathcal{A}$-harmonic, (3.1) directly follows from the previous inequality. Let

$$\boldsymbol{\ell}(x) := (D\mathbf{h})_r x.$$

Suppose $\theta \in (0, 1/2]$. We note from the mean value theorem for $D\mathbf{h}$ in $Q_{\theta\rho}$ that

$$\begin{aligned}
\sup_{Q_{\theta\rho}} |D\mathbf{h} - (D\mathbf{h})_{\theta r}| &\leqslant 2\theta r \sup_{Q_{\theta r}} |D^2\mathbf{h}| + 2\theta^2 r^2 \sup_{z\in Q_{\theta r}} |[D\mathbf{h}]_t| \\
&= 2\theta r \sup_{Q_{\theta r}} |D^2(\mathbf{h} - \boldsymbol{\ell})| + 2\theta^2 r^2 \sup_{Q_{\theta r}} |D \operatorname{div}[\mathcal{A}D(\mathbf{h} - \boldsymbol{\ell})]| \\
&\leqslant 2\theta r \sup_{Q_{r/2}} |D^2(\mathbf{h} - \boldsymbol{\ell})| + 2\theta^2 r^2 \sup_{Q_{r/2}} |D^3(\mathbf{h} - \boldsymbol{\ell})| \\
&\leqslant 2\theta \left(r \sup_{Q_{r/2}} |D^2(\mathbf{h} - \boldsymbol{\ell})| + r^2 \sup_{Q_{r/2}} |D^3(\mathbf{h} - \boldsymbol{\ell})| \right).
\end{aligned}$$

Since every $(\mathbf{h} - \boldsymbol{\ell})_{x_i}$, $i = 1, 2, \dots, n$, is also $\mathcal{A}$-caloric in Q_r, we have from (3.1) that

$$\sup_{Q_{\theta r}} |D\mathbf{h} - (D\mathbf{h})_{\theta r}| \leqslant c\theta \fint_{Q_r} |D(\mathbf{h} - \boldsymbol{\ell})| \,\mathrm{d}z\,,$$

which implies (3.2). □

We introduce the parabolic Calderón-Zygmund estimates for an N-function ψ with $\Delta(\psi, \psi^*) < \infty$. We shall assume that $\psi(1) = 1$ without loss of generality. In the next lemma, if $\psi(\tau) = \tau^p$, $1 < p < \infty$, the estimate (3.3) is well known, see for instance [36] and references therein. Furthermore, for general Orlicz functions it can be obtained by applying a standard interpolation argument for linear operators as in [15, Theorem 18]. We also refer to [8] for more general results for parabolic Calderón-Zygmund estimates in Orlicz spaces.

Lemma 3.2. *(Calderón-Zygmund estimates) Let* ψ *be an* N*-function with* $\Delta(\psi, \psi^*) < \infty$ *and* $\mathbf{G} \in L^\psi(Q_r, \mathbb{R}^{Nn})$ *where* $Q_r = B_r \times (-r^2, r^2)$*. There exists a unique weak solution* $\mathbf{u} \in L^\psi(-r^2, r^2; L^\psi(B_r))$ *with* $\mathbf{w}_t \in (L^\psi((-r^2, r^2); W_0^{1,\psi}(\Omega)))'$ *to the system*

$$\begin{cases} \partial_t \mathbf{w} - \operatorname{div}(\mathcal{A}D\mathbf{w}) = -\operatorname{div}\mathbf{G} & \textit{in } \ Q_r, \\ \mathbf{w} = \mathbf{0} \quad \textit{on } \ \partial_{\mathrm{p}} Q_r, \end{cases}$$

and we have the estimates

$$\|D\mathbf{w}\|_{L^\psi(Q_r)} \leqslant c\|\mathbf{G}\|_{L^\psi(Q_r)}\,,$$

and

$$\int_{Q_r} \psi(|D\mathbf{w}|)\,dz \leqslant c\int_{Q_r} \psi(|\mathbf{G}|)\,dz\,, \tag{3.3}$$

where the constant $c > 0$ depends on $n, N, \nu, |\mathcal{A}|$ and $\Delta(\psi, \psi^*)$.

Remark 3.3. Analogous estimates as above can be inferred for the weak solution $\mathbf{v}$ to

$$\begin{cases} \partial_t \mathbf{v} + \operatorname{div}(\mathcal{A}^T D\mathbf{v}) = -\operatorname{div}\mathbf{G} \quad \text{in } \ Q_r, \\ \mathbf{v} = \mathbf{0} \quad \text{on } \ (\partial B_r \times \{-r^2 < t \leqslant r^2\}) \cup (B_r \times \{t = r^2\}), \end{cases} \tag{3.4}$$

by considering the reflecting function $\widetilde{\mathbf{v}}(x,t) = \mathbf{v}(x,-t)$. Here, $\mathcal{A}^T$ is the transpose of $\mathcal{A}$, and note that if $\mathcal{A}$ satisfies the Legendre-Hadamard condition then so does $\mathcal{A}^T$.

We estimate the gradient of a function in $L^\psi(-r^2, r^2; W_0^{1,\psi}(B_r))$ in terms of functions in the dual space L^{ψ^*}.

Lemma 3.4. *For every $\mathbf{w} \in L^\psi(-r^2, r^2; W_0^{1,\psi}(B_r))$, we have*

$$\int_{Q_r} \psi(|D\mathbf{w}|)\,dz \leqslant \sup_{\mathbf{G}\in L^{\psi^*}\cap C^\infty(Q_r,\mathbb{R}^{Nn})} \left(\int_{Q_r} \mathbf{w}\cdot(\mathbf{v_G})_t - \langle \mathcal{A}D\mathbf{w}, D\mathbf{v}_G\rangle - \psi^*(|\mathbf{G}|)\,\mathrm{d}z \right),$$

where $\mathbf{v_G}$ is the weak solution to (3.4).

Proof. We note from the definition of the conjugate function in (2.1) that

$$\psi(t) = \psi^{**}(t) = \sup_{s\geqslant 0}\,(st - \psi^*(s)) = t\psi'(t) - \psi^*(\psi'(t))\,.$$

Hence, denoting $\mathbf{G_w} := \psi'(|D\mathbf{w}|)\frac{D\mathbf{w}}{|D\mathbf{w}|}$, we have $\mathbf{G_w} \in L^{\psi^*}(Q_r, \mathbb{R}^{Nn})$ by (2.3), and

$$\begin{aligned} \int_{Q_r} \psi(|D\mathbf{w}|)\,\mathrm{d}z &= \int_{Q_R} |D\mathbf{w}|\psi'(|D\mathbf{w}|) - \psi^*(\psi'(|D\mathbf{w}|))\,\mathrm{d}z \\ &= \int_{Q_r} \langle \mathbf{G_w}, D\mathbf{w}\rangle - \psi^*(|\mathbf{G_w}|)\,\mathrm{d}z \\ &\leqslant \sup_{\mathbf{G}\in L^{\psi^*}(Q_r,\mathbb{R}^{Nn})} \left(\int_{Q_r} \langle \mathbf{G}, D\mathbf{w}\rangle - \psi^*(|\mathbf{G}|)\,\mathrm{d}z \right) \\ &= \sup_{\mathbf{G}\in L^{\psi^*}\cap C^\infty(Q_r,\mathbb{R}^{Nn})} \left(\int_{Q_r} \langle \mathbf{G}, D\mathbf{w}\rangle - \psi^*(|\mathbf{G}|)\,\mathrm{d}z \right). \end{aligned}$$

For each $\mathbf{G} \in L^{\psi^*}(Q_r, \mathbb{R}^{Nn}) \cap C^\infty(Q_r, \mathbb{R}^{Nn})$, let $\mathbf{v}_G$ be the weak solution to (3.4). Then we see that $\mathbf{v_G} \in C^\infty(Q_r, \mathbb{R}^N)$ since $\mathcal{A}$ is constant, and by testing (3.4) with $\mathbf{w}$ we have

$$\int_{Q_r} \langle \mathbf{G}, D\mathbf{w}\rangle\,\mathrm{d}z = \int_{Q_r} (\mathbf{v_G})_t\cdot\mathbf{w} - \langle \mathcal{A}^T D\mathbf{v_G}, D\mathbf{w}\rangle\,\mathrm{d}z = \int_{Q_r} \mathbf{w}\cdot(\mathbf{v_G})_t - \langle \mathcal{A}D\mathbf{w}, D\mathbf{v_G}\rangle\,\mathrm{d}z\,.$$

This concludes the proof. ∎

3.2. **Parabolic Lipschitz truncation.** We recall the parabolic Lipschitz truncations introduced in [19] and their main properties, in the particular case when the scaling quantity α therein is equal to 1.

Let $\mathbf{v} \in L^{\psi}(-r^1, r^1; W_0^{1,1}(B_r, \mathbb{R}^{Nn}))$ and $\mathbf{G} \in L^1(Q_r, \mathbb{R}^{Nn})$ satisfy the system

$$\mathbf{v}_t = \operatorname{div} \mathbf{G} \quad \text{in } \ Q_r, \qquad \mathbf{v} = 0 \quad \text{on } \ \partial_{\mathrm{p}} Q_r, \tag{3.5}$$

in the distributional sense. We take as "bad set" a superlevel set for the maximal function of the spatial gradient and of the time derivative in the following way:

$$\mathcal{O}_\lambda := \{\mathcal{M}(\chi_{Q_r} \nabla \mathbf{v}) > \lambda\} \cup \{\mathcal{M}(\chi_{Q_r} \mathbf{G}) > \lambda\}, \quad \lambda > 0\,,$$

where $\mathcal{M}$ is the parabolic maximal operator defined by

$$\mathcal{M}(f)(\tilde{z}) := \sup_{Q_\rho(z_0)\,:\,\tilde{z} \in Q_\rho(z_0)} \fint_{Q_\rho(z_0)} |f| \,\mathrm{d}z\,.$$

Then we have the following properties, which can be inferred by [19, Theorem 2.3] with $\alpha = 1$, $\mathcal{M} := \mathcal{M}^1$ and $\mathcal{O}_\lambda := \mathcal{O}_\lambda^1$.

Lemma 3.5. *Let $\mathbf{v} \in L^{\psi}(-r^2, r^2; W_0^{1,\psi}(B_r, \mathbb{R}^{Nn}))$ satisfy the system (3.5) in the distributional sense, and let $\lambda > 0$. Then there exists $\mathbf{v}_\lambda \in L^1(-r^2, r^2; W_0^{1,1}(B_r))$ with $|D\mathbf{v}_\lambda| \in L^{\psi}(Q_r)$ such that*

(1) $\mathbf{v}_\lambda = \mathbf{v}$ on $(\mathcal{O}_\lambda)^c$.
(2) $\mathcal{M}(D\mathbf{v}_\lambda) \leqslant c\lambda$.
(3) we have

$$\int_{Q_r} \psi(|D(\mathbf{v}_\lambda - \mathbf{v})|) \,\mathrm{d}z \leqslant c \int_{Q_r} \psi(|D\mathbf{v}|) \,\mathrm{d}z + c\psi(\lambda)|\mathcal{O}_\lambda|\,.$$

Here the constants $c > 0$ depend on n, N and $\Delta_2(\psi, \psi^*)$.

We note from [19, Section 2.3] that the function $\mathbf{v}_\lambda$ is determined by

$$\mathbf{v}_\lambda := \mathbf{v} - \sum_i \zeta_i(\mathbf{v} - \mathbf{v}_i), \quad \text{where } \mathbf{v}_i := \begin{cases} (\mathbf{v})_{\zeta_i} & \text{if } \ \frac{3}{4}Q_i \subset B_r \times (-r^2, 3r^2)\,, \\ \mathbf{0} & \text{otherwise,} \end{cases} \tag{3.6}$$

where we extend $\mathbf{v}$ and $\mathbf{G}$ to $B_r \times (r^2, 3r^2)$ by $\mathbf{v}(x, 2r^2 - t)$ and $-\mathbf{G}(x, 2r^2 - t)$ and to the outside of $B_r \times (-r^2, 3r^2)$ by zeros, hence this extended $\mathbf{v}$ satisfies the system $\mathbf{v}_t = \operatorname{div} \mathbf{G}$ in $B_r \times (-\infty, \infty)$ in the sense of distributions, and $\{Q_i\}_{i=1}^\infty$ is a parabolic Whitney covering of $\mathcal{O}_\lambda$ such that $Q_j = Q_{r_i}(z_i)$,

(W1) $\bigcup_i \frac{1}{2} Q_i = \mathcal{O}_\lambda$,
(W2) for all $j \in \mathbb{N}$ we have $8Q_i \subset \mathcal{O}_\lambda$ and $16Q_i \cap (\mathbb{R}^{m+1} \setminus \mathcal{O}_\lambda) \neq \emptyset$,
(W3) if $Q_i \cap Q_j \neq \emptyset$ then $\frac{1}{2} r_j \leqslant r_i \leqslant 2\, r_j$,
(W4) $\frac{1}{4} Q_i \cap \frac{1}{4} Q_j = \emptyset$ for all $i \neq j$,
(W5) each $x \in \mathcal{O}_\lambda$ belongs to at most 120^{n+2} of the sets $4Q_i$.

Here, $\kappa Q_i := Q_{\kappa r_i}$ for $\kappa > 0$, and $\{\zeta_i\} \subset C_0^\infty(\mathbb{R}^{n+1})$ is a partition of unity with respect to $\{Q_i\}$ that satisfies

(P1) $\chi_{\frac{1}{2} Q_i} \leqslant \zeta_i \leqslant \chi_{\frac{3}{4} Q_i}$
(P2) $\|\zeta_i\|_\infty + r_i \|D\zeta_i\|_\infty + r_i^2 \|D^2 \zeta_i\|_\infty + r_i^2 \|(\zeta_i)_t\|_\infty \leqslant c$.
(P3) For each $j \in \mathbb{N}$ we define $A_j := \{i \,:\, \frac{3}{4} Q_j \cap \frac{3}{4} Q_i \neq \emptyset\}$. Then $\sum_{i \in A_j} \zeta_i = 1$ on $\frac{3}{4} Q_j$.

The following result provides an upper bound for the measure of the bad set $\mathcal{O}_\lambda$, see [19, Lemma 4.1] with $\alpha = 1$.

Lemma 3.6. *Let* $\mathbf{v} \in L^{\psi}(-r^2, r^2; W_0^{1,\psi}(B_r))$ *and* $\mathbf{G} \in L^{\psi^*}(Q_r)$ *satisfy* (3.5) *in the distribution sense. Set* $\gamma > 0$ *such that*

$$\psi(\gamma) := \fint_{Q_r} \psi(|D\mathbf{v}|)\,\mathrm{d}z + \fint_{Q_r} \psi(|\mathbf{G}|)\,\mathrm{d}z\,.$$

Then, for every $m_0 \in \mathbb{N}$*, there exists* $\lambda \in [\gamma, 2^{m_0}\gamma]$ *such that*

$$|\mathcal{O}_\lambda| \leqslant c\frac{\psi(\gamma)}{m_0\psi(\lambda)}|Q_r|$$

for some $c > 0$ *depending on* n, N *and* $\Delta_2(\psi, \psi^*)$.

We end this subsection presenting a Poincaré-type inequality. Note that the following lemma is irrelevant to the above setting.

Lemma 3.7. *Let* $\mathbf{w} \in L^{\psi}(-r^2, r^2; W_0^{1,1}(B_r, \mathbb{R}^{Nn}))$ *and* $\mathbf{H} \in L^1(Q_r, \mathbb{R}^{Nn})$ *satisfy the system*

$$\mathbf{w}_t = \operatorname{div}\mathbf{H} \quad \text{in} \;\; Q_r, \qquad \mathbf{w} = \mathbf{0} \quad \text{on} \;\; \partial_{\mathrm{p}} Q_r,$$

in the distributional sense. Extend $\mathbf{w}$ *and* $\mathbf{H}$ *by* $\mathbf{w}(x, 2r^2 - t)$ *and* $-\mathbf{H}(x, 2r^2 - t)$ *to* $B_r \times (r^2, 3r^2)$ *and by zero outside* $B_r \times (-r^2, 3r^2)$*. For any parabolic cylinder* $Q_\rho = Q_\rho(z)$ *in* $\mathbb{R}^{n+1}$ *and any* $\zeta \in C_0^\infty(\frac{3}{4}Q_\rho)$ *with* $\zeta \geqslant 0$ *with* $\|\zeta\|_{L^\infty(\frac{3}{4}Q_\rho)} \leqslant c_0|\frac{3}{4}Q_\rho|^{-1}\|\zeta\|_{L^1(\frac{3}{4}Q_\rho)}$*, set*

$$\overline{\mathbf{w}} := \begin{cases} (\mathbf{w})_\zeta & \text{if } \; \frac{3}{4}Q_\rho \subset B_r \times (-r^2, 3r^2)\,, \\ \mathbf{0} & \text{otherwise.} \end{cases}$$

Then we have

$$\fint_{\frac{3}{4}Q_\rho} \psi\left(\frac{\mathbf{w} - \overline{\mathbf{w}}}{\rho}\right)\mathrm{d}z \leqslant c\fint_{Q_\rho} \psi(|D\mathbf{w}|)\,\mathrm{d}z + c\psi\left(\fint_{Q_\rho} |\mathbf{H}|\,\mathrm{d}z\right)$$

for some $c > 0$ *depending on* $n, N, \Delta_2(\psi, \psi^*)$ *and* c_0.

Proof. The proof is almost the same as the one of [19, Lemma 2.11] with replacing [19, Lemma 2.8] by [19, Lemma 2.9]. In fact, if $\frac{3}{4}Q_\rho \subset B_r \times (-r^2, 3r^2)$, then the inequality follows directly from [19, Lemma 2.9].

If $\frac{3}{4}Q_\rho \not\subset B_r \times (-r^2, 3r^2)$ and $\frac{4}{5}Q_\rho \subset B_r \times (-\infty, \infty)$, choose $\tilde{\zeta} \in C_0^\infty(\frac{4}{5}Q_\rho)$ with $\tilde{\zeta} \geqslant 0$, $\operatorname{supp}(\tilde{\zeta}) \subset \frac{4}{5}Q_\rho \setminus (B_r \times (-r^2, 3r^2))$ and $\|\zeta\|_{L^\infty(\frac{4}{5}Q_\rho)} \leqslant c(n)|\frac{4}{5}Q_\rho|^{-1}\|\zeta\|_{L^1(\frac{4}{5}Q_\rho)}$. Then, since $\mathbf{w} \equiv \mathbf{0}$ in $\operatorname{supp}(\tilde{\zeta})$, we have $(\mathbf{w})_{\tilde{\zeta}} = \mathbf{0}$ hence again by [19, Lemma 2.9]

$$\fint_{\frac{3}{4}Q_\rho} \psi\left(\frac{\mathbf{w} - \overline{\mathbf{w}}}{\rho}\right)\mathrm{d}z = \fint_{\frac{3}{4}Q_\rho} \psi\left(\frac{\mathbf{w} - (\mathbf{w})_{\tilde{\zeta}}}{\rho}\right)\mathrm{d}z \leqslant c\fint_{\frac{4}{5}Q_\rho} \psi(|D\mathbf{w}|)\,\mathrm{d}z + c\psi\left(\fint_{\frac{4}{5}Q_\rho} |\mathbf{H}|\,\mathrm{d}z\right).$$

Finally, if $\frac{4}{5}Q_\rho \subset B_r \times (-\infty, \infty)$, then there exists $c(n) > 0$ such that $\frac{|B_r|}{|\{\mathbf{w}(x,t)=\mathbf{0}\} \cap B_r\}|} \leqslant c(n)$ for a.e. time slice of Q_ρ. Therefore by the Poincaré inequality for the space variable, see [14, Theorem 7], we have

$$\fint_{\frac{3}{4}Q_\rho} \psi\left(\frac{\mathbf{w} - \overline{\mathbf{w}}}{\rho}\right)\mathrm{d}z \leqslant \fint_{Q_\rho} \psi\left(\frac{\mathbf{w}}{\rho}\right)\mathrm{d}z \leqslant c\fint_{Q_\rho} \psi(|D\mathbf{w}|)\,\mathrm{d}z\,.$$

The proof is concluded. $\square$

3.3. **Caloric approximations.** We first obtain the $\mathcal{A}$-caloric approximation. As a novelty with respect to previous caloric type approximations, we do not need to restrict the choice of the test functions $\boldsymbol{\zeta}$ in $C_0^\infty(Q_r)$, but we only assume them to be zero on the lateral boundary. This allows us to choose as test functions also the solutions of suitable linear systems.

Theorem 3.8. *($\mathcal{A}$-caloric approximation) Let $\mu, \sigma, C_0 > 0$ and ψ be an N-function with $\Delta_2(\psi,\psi^*) < \infty$. For every $\varepsilon \in (0,1)$, there exists $\delta > 0$ depending on σ, C_0, $\Delta_2(\psi,\psi^*)$ and ε such that if $\mathbf{u} \in L^1(-r^2, r^2; W^{1,\psi^{1+\sigma}}(B_r, \mathbb{R}^N))$ and $\mathbf{H} \in L^{\psi^{1+\sigma}}(Q_r, \mathbb{R}^{Nn})$ satisfy*

$$\partial_t \mathbf{u} = \operatorname{div} \mathbf{H} \quad \textit{in} \ \ Q_r,$$

in the distributional sense, with the inequality

$$\left(\fint_{Q_r} \psi(|D\mathbf{u}|)^{1+\sigma} + \psi(|\mathbf{H}|)^{1+\sigma} \, \mathrm{d}z \right)^{\frac{1}{1+\sigma}} \leqslant C_0 \psi(\mu), \tag{3.7}$$

and for every $\boldsymbol{\zeta} \in C^\infty(Q_r; \mathbb{R}^N)$ with $\boldsymbol{\zeta} = \mathbf{0}$ on $\partial B_r \times (-r^2, r^2)$,

$$\frac{1}{|Q_r|} \left| \int_{Q_r} \mathbf{u} \cdot \boldsymbol{\zeta}_t - \langle \mathcal{A} D\mathbf{u}, D\boldsymbol{\zeta} \rangle \, dz - \left[\int_{B_r} \mathbf{u} \cdot \boldsymbol{\zeta}_t \, \mathrm{d}x \right]_{t=-r^2}^{t=r^2} \right| \leqslant \delta \mu \| D\boldsymbol{\zeta} \|_{L^\infty(Q_r, \mathbb{R}^{Nn})}, \tag{3.8}$$

then

$$\fint_{Q_r} \psi(|D\mathbf{u} - D\mathbf{h}|) \, \mathrm{d}z \leqslant \varepsilon \psi(\mu), \tag{3.9}$$

where $\mathbf{h}$ is the weak solution to

$$\begin{cases} \partial_t \mathbf{h} - \operatorname{div}(\mathcal{A} D\mathbf{h}) = 0 & \textit{in} \ \ Q_r, \\ \mathbf{h} = \mathbf{u} & \textit{on} \ \ \partial_{\mathrm{p}} Q_r. \end{cases}$$

Proof. It will suffice to prove the assertion in the case $\mu = 1$ with $\psi(1) = 1$, as the general case can be obtained by scaling argument with the functions $\tilde{\mathbf{u}} = \mu^{-1}\mathbf{u}$, $\tilde{\mathbf{H}} = \mu^{-1}\mathbf{H}$ and $\tilde{\psi}(\tau) = \frac{\psi(\mu\tau)}{\psi(\mu)}$. Set $\mathbf{w} := \mathbf{u} - \mathbf{h}$. Then $\mathbf{w}$ satisfies

$$\begin{cases} \partial_t \mathbf{w} - \operatorname{div}(\mathcal{A} D\mathbf{w}) = -\operatorname{div}\left(\mathcal{A} D\mathbf{u} + \mathbf{H}\right) & \text{in} \ \ Q_r, \\ \mathbf{w} = \mathbf{0} & \text{on} \ \ \partial_{\mathrm{p}} Q_r, \end{cases} \tag{3.10}$$

in the distributional sense. Moreover, by applying Lemma 3.2 to the N-function $\psi^{1+\sigma}$ and (3.7), we see that

$$\fint_{Q_r} \psi(|D\mathbf{w}|)^{1+\sigma} \, \mathrm{d}z \leqslant c \fint_{Q_r} \psi(|\mathcal{A} D\mathbf{u} + \mathbf{H}|)^{1+\sigma} \, \mathrm{d}z \leqslant c\,. \tag{3.11}$$

We will apply the inequality in Lemma 3.4. Fix any $\mathbf{G} \in L^{\psi^*}(Q_r, \mathbb{R}^{Nn}) \cap C^\infty(Q_r, \mathbb{R}^{Nn})$ and consider the weak solution $\mathbf{v}_{\mathbf{G}}$ to (3.4). Note that, by Lemma 3.2 with Remark 3.3 and ψ^* in place of ψ,

$$\int_{Q_r} \psi^*(|D\mathbf{v}_{\mathbf{G}}|) \, \mathrm{d}z \leqslant c \int_{Q_r} \psi^*(|\mathbf{G}|) \, \mathrm{d}z\,. \tag{3.12}$$

Moreover, $\mathbf{v}_{\mathbf{G}} \in C^\infty(Q_r, \mathbb{R}^N)$ since $\mathbf{G} \in C^\infty(Q_r, \mathbb{R}^{Nn})$. To enlighten the notation, from now on we will simply denote $\mathbf{v}_{\mathbf{G}}$ by $\mathbf{v}$.

Choose $\gamma \in [0, \infty)$ such that

$$\psi^*(\gamma) = \int_{Q_r} \psi^*(|D\mathbf{v}|) \, \mathrm{d}z + \int_{Q_r} \psi^*(|\mathcal{A}^T D\mathbf{v} + \mathbf{G}|) \, \mathrm{d}z\,. \tag{3.13}$$

Then, with (3.12) and the subadditivity of ψ^* we have

$$\psi^*(\gamma) \leqslant c \int_{Q_r} \psi^*(|\mathbf{G}|)\, \mathrm{d}z\,.$$

Let $m_0 \in \mathbb{N}$ be large enough, to be determined later. Then by Lemma 3.5(1) and Lemma 3.6 with ψ^* in place of ψ, there exists $\lambda \in [\gamma, 2^{m_0}\gamma]$ such that $\{\mathbf{v} \neq \mathbf{v}_\lambda\} \subset \mathcal{O}_\lambda$ and

$$\frac{|\mathcal{O}_\lambda|}{|Q_r|} \leqslant \frac{c\psi^*(\gamma)}{m_0\psi^*(\lambda)} \leqslant \frac{c}{m_0}\,, \tag{3.14}$$

where $\mathbf{v}_\lambda$ is the parabolic Lipschitz truncation of $\mathbf{v}$ provided by Lemma 3.5. Note that $\mathbf{v}$ is zero on the top, but not on the base, of the cylinder Q_r. Hence we apply the Lipschitz truncation and related results in the previous subsection to the function $\mathbf{v}(x,-t)$. Accordingly, $\mathbf{v}$ and $\tilde{\mathbf{G}} := -\mathcal{A}^T D\mathbf{v} - \mathbf{G}$ are extended to $B_r \times (-3r^2, -r^2)$ by $\mathbf{v}(x,t) = \mathbf{v}(x, -2r^2 - t)$ and $\tilde{\mathbf{G}}(x,t) = -\tilde{\mathbf{G}}(x, -2r^2 - t)$ and to the outside of $B_r \times (-3r^2, r^2)$ by zeros, hence from (3.4) $\mathbf{v}$ satisfies the system $\mathbf{v}_t = \operatorname{div} \tilde{\mathbf{G}}$ in $B_R \times (-\infty, \infty)$ in the distribution sense.

Then we observe that

$$\begin{aligned}
\int_{Q_r} \mathbf{w}\cdot\mathbf{v}_t - \langle \mathcal{A}D\mathbf{w}, D\mathbf{v}\rangle\, \mathrm{d}z &= \int_{Q_r} \mathbf{w}\cdot(\mathbf{v}_\lambda)_t - \langle \mathcal{A}D\mathbf{w}, D\mathbf{v}_\lambda\rangle\, \mathrm{d}z \\
&\quad + \int_{Q_r} \mathbf{w}\cdot(\mathbf{v}-\mathbf{v}_\lambda)_t\, \mathrm{d}z - \int_{Q_r} \langle \mathcal{A}D\mathbf{w}, D(\mathbf{v}-\mathbf{v}_\lambda)\rangle\, \mathrm{d}z \\
&=: I_1 + I_2 - I_3\,.
\end{aligned} \tag{3.15}$$

For I_1, since $\mathbf{w} = \mathbf{u} + \mathbf{h}$,

$$\begin{aligned}
I_1 &= \int_{Q_r} \mathbf{w}\cdot(\mathbf{v}_\lambda)_t - \langle \mathcal{A}D\mathbf{w}, D\mathbf{v}_\lambda\rangle\, \mathrm{d}z - \left[\int_{Q_r} \mathbf{w}\cdot\mathbf{v}_\lambda\, \mathrm{d}x\right]_{t=-r^2}^{t=r^2} \\
&= \int_{Q_r} \mathbf{u}\cdot(\mathbf{v}_\lambda)_t - \langle \mathcal{A}D\mathbf{u}, D\mathbf{v}_\lambda\rangle\, \mathrm{d}z - \left[\int_{Q_r} \mathbf{u}\cdot\mathbf{v}_\lambda\, \mathrm{d}x\right]_{t=-r^2}^{t=r^2}.
\end{aligned}$$

Then by (3.8), Lemma 3.5(2) and Young's inequality we have that for any $\kappa_1 \in (0,1)$,

$$\frac{1}{|Q_r|} I_1 \leqslant \delta \|D\mathbf{v}_\lambda\|_{L^\infty(Q_r,\mathbb{R}^{Nn})} \leqslant c\delta\lambda \leqslant c_{\kappa_1}\psi(\delta) + \kappa_1\psi^*(\lambda) \leqslant c_{\kappa_1}\psi(\delta) + \kappa_1\psi^*(2^{m_0}\gamma).$$

We next estimate I_3. By Young's inequality, Hölder's inequality and Lemma 3.5(3) with ψ^* in place of ψ and Lemma 3.6, we have that for any $\kappa_2 \in (0,1)$

$$\begin{aligned}
|I_3| &\leqslant c_{\kappa_2} \int_{\mathcal{O}_\lambda \cap Q_r} \psi(|D\mathbf{w}|)\, \mathrm{d}z + \kappa_2 \int_{\mathcal{O}_\lambda \cap Q_r} \psi^*(|D(\mathbf{v}-\mathbf{v}_\lambda)|)\, \mathrm{d}z \\
&\leqslant c_{\kappa_2} \int_{\mathcal{O}_\lambda \cap Q_r} \psi(|D\mathbf{w}|)\, \mathrm{d}z + c\kappa_2 \int_{Q_r} \psi^*(|D\mathbf{v}|)\, \mathrm{d}z + c\kappa_2 |\mathcal{O}_\lambda| \psi^*(\lambda) \\
&\leqslant c_{\kappa_2} \left(\int_{Q_r} \psi(|D\mathbf{w}|)^{1+\sigma}\, \mathrm{d}z\right)^{\frac{1}{1+\sigma}} |\mathcal{O}_\lambda|^{\frac{\sigma}{1+\sigma}} + c\kappa_2 \int_{Q_r} \psi^*(|D\mathbf{v}|)\, \mathrm{d}z + c\kappa_2 |\mathcal{O}_\lambda| \psi^*(\lambda)\,,
\end{aligned}$$

hence applying (3.11) and (3.14)

$$\frac{1}{|Q_r|}|I_3| \leqslant c_{\kappa_2} m_0^{-\frac{\sigma}{1+\sigma}} + c\kappa_2 \fint_{Q_r} \psi^*(|D\mathbf{v}|)\, \mathrm{d}z + \frac{c\kappa_2}{m_0}\psi^*(\gamma)\,.$$

Finally, we estimate I_2. Recall the parabolic Whitney covering $\{Q_i\}_{i=1}^\infty$ of $\mathcal{O}_\lambda$ and the partition of unity $\{\zeta_i\}_{i=1}^\infty \subset C_0^\infty(\frac{3}{4}Q_i)$ with respect to $\mathbf{v}$ and the definition of $\mathbf{v}_\lambda$ in (3.6). In addition, we extend $\mathbf{w}$ and $\tilde{\mathbf{H}} := \mathcal{A}(D\mathbf{w} - D\mathbf{u} - \mathbf{H})$ to $B_r \times (r^2, 3r^2)$ by $\mathbf{w}(x,t) = \mathbf{w}(x, 2r^2 - t)$ and $\tilde{\mathbf{H}}(x,t) = -\tilde{\mathbf{H}}(x, 2r^2 - t)$ and to the outside of $B_r \times (-r^2, 3r^2)$ by zeros, hence from (3.10) $\mathbf{w}$ satisfies the system $\mathbf{w}_t = \operatorname{div} \tilde{\mathbf{H}}$ in $B_R \times (-\infty, \infty)$ in the distribution sense. With this extended $\mathbf{w}$, we set

$$\mathbf{w}_i := \begin{cases} (\mathbf{w})_{\zeta_i} & \text{if } \frac{3}{4}Q_i \subset B_r \times (-r^2, 3r^2)\,, \\ \mathbf{0} & \text{otherwise.} \end{cases}$$

Then, since $\mathbf{w}_i$'s are constants and $\mathbf{v}$ solves (3.4), we have

$$\begin{aligned} I_2 &\leqslant c \sum_{\frac{3}{4}Q_i \cap Q_r \neq \emptyset} \left| \int_{\frac{3}{4}Q_i \cap Q_r} \mathbf{w} \cdot [\zeta_i(\mathbf{v} - \mathbf{v}_i)]_t \,\mathrm{d}z \right| \\ &= c \sum_{\frac{3}{4}Q_i \cap Q_r \neq \emptyset} \left| \int_{\frac{3}{4}Q_i \cap Q_r} (\mathbf{w} - \mathbf{w}_i) \cdot [\zeta_i(\mathbf{v} - \mathbf{v}_i)]_t \,\mathrm{d}z \right| \\ &= c \sum_{\frac{3}{4}Q_i \cap Q_r \neq \emptyset} \left| \int_{\frac{3}{4}Q_i \cap Q_r} (\mathbf{w} - \mathbf{w}_i) \cdot [(\mathbf{v} - \mathbf{v}_i)(\zeta_i)_t + \mathbf{v}_t \zeta_i] \,\mathrm{d}z \right| \\ &= c \sum_{\frac{3}{4}Q_i \cap Q_r \neq \emptyset} \left| \int_{\frac{3}{4}Q_i \cap Q_r} (\mathbf{w} - \mathbf{w}_i) \cdot (\mathbf{v} - \mathbf{v}_i)(\zeta_i)_t + \langle (\mathcal{A}^T D\mathbf{v} + \mathbf{G}), D[(\mathbf{w} - \mathbf{w}_i)\zeta_i] \rangle \,\mathrm{d}z \right| \\ &\leqslant c \sum_{\frac{3}{4}Q_i \cap Q_r \neq \emptyset} \int_{\frac{3}{4}Q_i \cap Q_r} \frac{|\mathbf{w} - \mathbf{w}_i|}{r_i} \frac{|\mathbf{v} - \mathbf{v}_i|}{r_i} + (|D\mathbf{v}| + |\mathbf{G}|) \left(|D\mathbf{w}| + \frac{|\mathbf{w} - \mathbf{w}_i|}{r_i} \right) \mathrm{d}z\,. \end{aligned}$$

Moreover, by Young's inequality we have that for any $\kappa_3 \in (0,1)$,

$$\begin{aligned} I_2 &\leqslant c_{\kappa_3} \sum_{\frac{3}{4}Q_i \cap Q_r \neq \emptyset} \int_{\frac{3}{4}Q_i} \psi\left(\frac{|\mathbf{w} - \mathbf{w}_i|}{r_i} \right) + \psi(|D\mathbf{w}|) \,\mathrm{d}z \\ &\quad + \kappa_3 \sum_{\frac{3}{4}Q_i \cap Q_r \neq \emptyset} \int_{\frac{3}{4}Q_i} \psi^*\left(\frac{|\mathbf{v} - \mathbf{v}_i|}{r_i} \right) + \psi^*(|D\mathbf{v}| + |\mathbf{G}|) \,\mathrm{d}z\,. \end{aligned}$$

Then, applying Lemma 3.7 to $\mathbf{w}$ and $\mathbf{v}$ with the extensions of $\mathbf{w}$, $\mathbf{v}$, $\mathbf{H}$ and $\mathbf{G}$, we have that

$$\begin{aligned} \int_{\frac{3}{4}Q_i} \psi\left(\frac{|\mathbf{w} - \mathbf{w}_i|}{r_i} \right) \mathrm{d}z &\leqslant c \int_{Q_i} \psi(|D\mathbf{w}|) \,\mathrm{d}z + \int_{Q_i} \psi(|\mathcal{A}D\mathbf{w} - \mathcal{A}D\mathbf{u} - \mathbf{H}|) \,\mathrm{d}z \\ &\leqslant c \int_{Q_i} \psi(|D\mathbf{w}| + |D\mathbf{u}| + |\mathbf{H}|) \,\mathrm{d}z\,, \end{aligned}$$

and

$$\begin{aligned} \int_{\frac{3}{4}Q_i} \psi^*\left(\frac{|\mathbf{v} - \mathbf{v}_i|}{r_i} \right) \mathrm{d}z &\leqslant c \int_{Q_i} \psi^*(|D\mathbf{v}|) \,\mathrm{d}z + \int_{Q_i} \psi^*(|-\mathcal{A}^T D\mathbf{v} - \mathbf{G}|) \,\mathrm{d}z \\ &\leqslant c \int_{Q_i} \psi^*(|D\mathbf{v}| + |\mathbf{G}|) \,\mathrm{d}z\,. \end{aligned}$$

Using these inequalities, the fact that $\sum_i \chi_{Q_i} \leqslant c(n)$ and considering the extension of functions, we estimate I_2 as

$$\begin{aligned} I_2 &\leqslant c_{\kappa_3} \int_{\mathcal{O}_\lambda} \psi(|D\mathbf{w}| + |D\mathbf{u}| + |\mathbf{H}|)\,\mathrm{d}z + c\kappa_3 \int_{\mathcal{O}_\lambda} \psi^*(|D\mathbf{v}| + |\mathbf{G}|)\,\mathrm{d}z \\ &\leqslant c_{\kappa_3} \left(\int_{\mathcal{O}_\lambda} \psi(|D\mathbf{w}| + |D\mathbf{u}| + |\mathbf{H}|)^{1+\sigma}\,\mathrm{d}z \right)^{\frac{1}{1+\sigma}} |\mathcal{O}_\lambda|^{\frac{\sigma}{1+\sigma}} + c\kappa_3 \int_{\mathcal{O}_\lambda} \psi^*(|D\mathbf{v}| + |\mathbf{G}|)\,\mathrm{d}z \\ &\leqslant c_{\kappa_3} \left(\int_{Q_r} \psi(|D\mathbf{w}| + |D\mathbf{u}| + |\mathbf{H}|)^{1+\sigma}\,\mathrm{d}z \right)^{\frac{1}{1+\sigma}} |\mathcal{O}_\lambda|^{\frac{\sigma}{1+\sigma}} + c\kappa_3 \int_{Q_r} \psi^*(|D\mathbf{v}| + |\mathbf{G}|)\,\mathrm{d}z\,. \end{aligned}$$

Therefore, by Hölder's inequality and the estimates (3.7), (3.11), (3.12) and (3.13), we have

$$\frac{1}{|Q_r|} I_2 \leqslant c_{\kappa_3} \left(\frac{|\mathcal{O}_\lambda|}{|Q_r|} \right)^{\frac{\sigma}{1+\sigma}} + c\kappa_3 ⨍_{Q_r} \psi(|D\mathbf{v}| + |\mathbf{G}|)\,\mathrm{d}z \leqslant c_{\kappa_3} m_0^{-\frac{\sigma}{1+\sigma}} + c\kappa_3 ⨍_{Q_r} \psi^*(|\mathbf{G}|)\,\mathrm{d}z\,.$$

Inserting the estimates for I_1, I_2 and I_3 into (3.15), we have

$$\begin{aligned} ⨍_{Q_r} \mathbf{w}\cdot\mathbf{v}_t - \langle \mathcal{A} D\mathbf{w}, D\mathbf{v}\rangle\,\mathrm{d}z &\leqslant c_{\kappa_1}\psi(\delta) + (c_{\kappa_2} + c_{\kappa_3}) m_0^{-\frac{\sigma}{1+\sigma}} \\ &\qquad + (c_{m_0}\kappa_1 + c\kappa_2 + c\kappa_3) ⨍_{Q_r} \psi^*(|\mathbf{G}|)\,\mathrm{d}z\,. \end{aligned}$$

We choose κ_2, κ_3 small so that $c\kappa_2 + c\kappa_3 \leqslant \frac{1}{2}$, then m_0 large so that $(c_{\kappa_2} + c_{\kappa_3}) m_0^{-\frac{\sigma}{1+\sigma}} \leqslant \frac{\varepsilon}{2}$, then κ_1 small so that $c_{m_0}\kappa_1 \leqslant \frac{1}{2}$, and then δ small so that $c_{\kappa_1}\psi(\delta) \leqslant \frac{\varepsilon}{2}$. Then we have

$$⨍_{Q_r} \mathbf{w}\cdot\mathbf{v}_t - \langle \mathcal{A} D\mathbf{w}, D\mathbf{v}\rangle\,\mathrm{d}z \leqslant \varepsilon + ⨍_{Q_r} \psi^*(|\mathbf{G}|)\,\mathrm{d}z\,.$$

Since $\mathbf{G}$ is an arbitrary function in $L^\psi(Q_r, \mathbb{R}^{Nn}) \cap C^\infty(Q_r, \mathbb{R}^{Nn})$, by Lemma 3.4 we deduce (3.9). This concludes the proof. $\square$

The φ-caloric approximation has been proved in [19, Theorem 4.2]. It states that every "almost φ-caloric" function has a φ-caloric function "close enough".

Theorem 3.9. *(φ-caloric approximation) Let $\gamma_1, \gamma_2 \in (0,1)$, $\gamma_3 \geqslant 1$, $I := (t^-, t^+)$. Suppose φ be an N-function with $\Delta_2(\varphi, \varphi^*) < \infty$ with $\varphi(1) = 1$. Then for every $\varepsilon > 0$ there exists $\delta > 0$ depending on $n, N, \Delta_2(\varphi, \varphi^*), \gamma_1, \gamma_2, \gamma_2$ and ε such that the following holds: if $\mathbf{u} \in L^\varphi(I, W^{1,\varphi}(B))$ satisfying $\mathbf{u}_t = \operatorname{div}\mathbf{G}$ in the distribution sense is almost φ-caloric in the sense that for all $\boldsymbol{\zeta} \in C_0^\infty(Q)$,*

$$\left| ⨍_Q \mathbf{u}\cdot\boldsymbol{\zeta}_t + \frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|} \langle D\mathbf{u}, D\boldsymbol{\zeta}\rangle\,\mathrm{d}z \right| \leqslant \delta \Big[⨍_Q \varphi(|D\mathbf{u}|) + \varphi^*(|\mathbf{G}|)\,\mathrm{d}z + \varphi(\|D\boldsymbol{\zeta}\|_\infty) \Big]\,,$$

then there exists a φ-caloric function $\mathbf{h}$ such that $\mathbf{h} = \mathbf{u}$ on $\partial_{\mathrm{p}} Q$ and

$$\begin{aligned} \left(⨍_I \left(⨍_B \left(\frac{|\mathbf{u} - \mathbf{h}|^2}{|t^+ - t^-|} \right)^{\gamma_2} \mathrm{d}x \right)^{\frac{\gamma_3}{\gamma_2}} \mathrm{d}t \right)^{\frac{1}{\gamma_3}} &+ \left(⨍_Q |\mathbf{V}(D\mathbf{u}) - \mathbf{V}(D\mathbf{h})|^{2\gamma_1}\,\mathrm{d}z \right)^{\frac{1}{\gamma_1}} \\ &\leqslant \varepsilon ⨍_Q \varphi(|D\mathbf{u}|) + \varphi^*(|\mathbf{G}|)\,\mathrm{d}z\,. \end{aligned}$$

Note that a first version of the p-caloric approximation method was developed by Bögelein-Duzaar-Mingione [6] by using a contradiction argument. They used it to show a partial regularity result for solutions of parabolic systems of p-growth; that is, almost everywhere $\nabla u \in C^{\alpha}$ for some $\alpha > 0$. We wish to quickly point out the improvements of the approximation lemma here with respect to the one in [6]. The proof is done directly by a comparison argument and the parabolic Lipschitz truncation. This direct approach allows for showing the closeness both in $L^{2\gamma_2}(L^{2\gamma_3})$ and $L^{\varphi^{\gamma_1}}(W^{1,\varphi^{\gamma_1}})$ norms (the last closeness is via the function $\mathbf{V}$ in (2.6)).

4. Caccioppoli type inequality and Higher integrability

Let $\mathbf{u}$ be a weak solution to (1.1). We always assume that φ and $\mathbf{A}$ satisfies Assumption (A). Let $\boldsymbol{\ell} : \mathbb{R}^n \to \mathbb{R}^N$ be any fixed linear map of the form

$$\boldsymbol{\ell}(x) := \mathbf{P}(x - x_0) + \mathbf{b}, \quad x \in \mathbb{R}^n, \tag{4.1}$$

where $\mathbf{P} \in \mathbb{R}^{Nn}$, $x_0 \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^N$, and set

$$\mathbf{u}_{\boldsymbol{\ell}} := \mathbf{u} - \boldsymbol{\ell}\,. \tag{4.2}$$

In this section we will obtain the higher integrability of not only $D\mathbf{u}$ but also of $D\mathbf{u}_{\ell}$. We follow the argument in [33].

We first recall a Gagliardo-Nirenberg type inequality for Orlicz functions, see Lemma 4.1, which has been proved in [33, Lemma 2.13]. In order to do that, we fix some notation. A function $\varphi : [0, \infty) \to [0, \infty)$ is said to be a *weak Φ-function* if it is increasing with $\varphi(0) = 0$, $\lim_{t\to 0^+} \varphi(t) = 0$, $\lim_{t\to +\infty} \varphi(t) = +\infty$ and such that the map $t \to \frac{\varphi(t)}{t}$ is almost increasing. Note that every N-function is a weak Φ-function.

Lemma 4.1. *Assume that $\psi : [0, \infty) \to [0, \infty)$ is a weak Φ-function and such that $t \mapsto \frac{\psi(t)}{t^{q_1}}$ is almost decreasing with constant $L \geqslant 1$ for some $q_1 \geqslant 1$. For $p \in [1, n)$ and $q_2 > 0$ we have*

$$\left(\fint_{B_r} \psi(|\tfrac{f}{r}|)^{\gamma}\,\mathrm{d}x \right)^{\frac{1}{\gamma}} \leqslant c\left(\fint_{B_r} \left[\psi(|Df|)^p + \psi(|\tfrac{f}{r}|)^p\right]\,\mathrm{d}x \right)^{\frac{\theta}{p}} \psi\left(\left(\fint_{B_r} |\tfrac{f}{r}|^{q_2}\,\mathrm{d}x \right)^{\frac{1}{q_2}} \right)^{1-\theta}$$

for some $c = c(n, L, q_1, q_2) > 0$, provided that $\theta \in (0, 1)$ and γ satisfies

$$\frac{1}{\gamma} \geqslant \frac{\theta}{p^*} + \frac{(1-\theta)q_1}{q_2}\,.$$

We start with a Caccioppoli type inequality for $\mathbf{u}_{\boldsymbol{\ell}}$.

Lemma 4.2 (Caccioppoli inequality for $\mathbf{u}_{\boldsymbol{\ell}}$)**.** *Let $\mathbf{u}$ be a weak solution to (1.1). For every pair of concentric cylinders $Q_{r_1,\tau_1}(z_0) \subset Q_{r_2,\tau_2}(z_0) \Subset \Omega_T$ with $z_0 = (x_0, t_0)$, $0 < r_1 < r_2$ and $0 < \tau_1 < \tau_2$, and $\mathbf{b} \in \mathbb{R}^N$, we have*

$$\begin{aligned} &\sup_{t \in I_{\tau_1}(t_0)} \int_{B_{r_1}(x_0)} |\mathbf{u}_{\boldsymbol{\ell}}(t) - \mathbf{b}|^2\,\mathrm{d}x + \int_{Q_{r_1,\tau_1}(z_0)} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)\,\mathrm{d}z \\ &\qquad \leqslant c \int_{Q_{r_2,\tau_2}(z_0)} \left[\frac{|\mathbf{u}_{\boldsymbol{\ell}} - \mathbf{b}|^2}{\tau_2 - \tau_1} + \varphi_{|D\boldsymbol{\ell}|}\left(\left| \frac{\mathbf{u}_{\boldsymbol{\ell}} - \mathbf{b}}{r_2 - r_1} \right| \right) \right] \mathrm{d}z \end{aligned} \tag{4.3}$$

for some $c = c(n, N, p, q, L, \nu) > 0$, where $\mathbf{u}(t) = \mathbf{u}(x, t)$ and $I_{\tau}(t_0) = (t_0 - \tau, t_0 + \tau)$.

Proof. We assume without loss of generality that the center of Q_{r_1,τ_1} and Q_{r_2,τ_2} is the origin. Let $\xi \in C_0^1(B_R)$ with $\xi \equiv 1$ in B_r and $|D\xi| \leqslant 2/(r_2 - r_1)$ and $\eta \in C^1(\mathbb{R})$ with $\eta \equiv 0$ in $(-\infty, -\tau_2]$, $\eta \equiv 1$ in $[-\tau_1, \infty)$ and $0 \leqslant \eta' \leqslant 2/(\tau_2 - \tau_1)$. Using

$$\boldsymbol{\zeta}(x, t) := \xi(x)^q \eta(t)^2 (\mathbf{u}_{\boldsymbol{\ell}}(x, t) - \mathbf{b})$$

as a test function in (1.1), we have that for $s \in I_{\tau_1}$,

$$\int_{-\tau_2}^{s} \int_{B_{r_2}} [\partial_t \mathbf{u} \cdot \boldsymbol{\zeta} + \mathbf{A}(D\mathbf{u}) : D\boldsymbol{\zeta}] \,\mathrm{d}x \,\mathrm{d}t \;= 0\,.$$

Moreover, since $\partial_t \mathbf{b} = \partial_t \boldsymbol{\ell} = \operatorname{div} \mathbf{A}(D\boldsymbol{\ell}) = \mathbf{0}$, we further have

$$\int_{-\tau_2}^{s} \int_{B_{r_2}} [\partial_t (\mathbf{u}_{\boldsymbol{\ell}} - \mathbf{b}) \cdot \boldsymbol{\zeta} + (\mathbf{A}(D\mathbf{u}) - \mathbf{A}(D\boldsymbol{\ell})) : D\boldsymbol{\zeta}] \,\mathrm{d}x \,\mathrm{d}t \;= 0\,.$$

Note that

$$\begin{aligned}
\int_{-\tau_2}^{s} \int_{B_{r_2}} \partial_t \mathbf{u} \cdot \boldsymbol{\zeta} \,\mathrm{d}x \,\mathrm{d}t &= \int_{-\tau_2}^{s} \int_{B_{r_2}} \frac{1}{2} \partial_t [\xi^q \eta^2 |\mathbf{u}_{\boldsymbol{\ell}} - \mathbf{b}|^2] - \xi^q \eta \eta' |\mathbf{u}_{\boldsymbol{\ell}} - \mathbf{b}|^2 \,\mathrm{d}x \,\mathrm{d}t \\
&= \frac{1}{2} \int_{B_{r_2}} \xi^q \eta(s)^2 |\mathbf{u}_{\boldsymbol{\ell}}(s) - \mathbf{b}|^2 \,\mathrm{d}x - \int_{-\tau_2}^{s} \int_{B_{r_2}} \xi^q \eta \eta' |\mathbf{u}_{\boldsymbol{\ell}} - \mathbf{b}|^2 \,\mathrm{d}x \,\mathrm{d}t,
\end{aligned}$$

where $\mathbf{u}_{\boldsymbol{\ell}}(s) = \mathbf{u}_{\boldsymbol{\ell}}(x,s)$. Then, applying (2.16), (2.19), (2.7) and (2.9) we have that for every $s \in I_{\tau_1}$,

$$\begin{aligned}
&\frac{1}{2} \int_{B_{r_2}} \xi^q \eta(s)^2 |\mathbf{u}_{\boldsymbol{\ell}}(s) - \mathbf{b}|^2 \,\mathrm{d}x + \frac{1}{c} \int_{\sigma}^{s} \int_{B_{r_2}} \xi^q \eta^2 \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}) \,\mathrm{d}x \,\mathrm{d}t \\
&\leqslant -q \int_{-\tau_2}^{s} \int_{B_{r_2}} \xi^{q-1} \eta^2 (\mathbf{A}(D\mathbf{u}) - \mathbf{A}(D\boldsymbol{\ell})) : D\xi \otimes (\mathbf{u}_{\boldsymbol{\ell}} - \mathbf{b}) \,\mathrm{d}x \,\mathrm{d}t \\
&\quad + \int_{-\tau_2}^{s} \int_{B_{r_2}} \xi^q \eta \eta' |\mathbf{u}_{\boldsymbol{\ell}} - \mathbf{b}|^2 \,\mathrm{d}x \,\mathrm{d}t \\
&\leqslant c \int_{-\tau_2}^{s} \int_{B_{r_2}} \xi^{q-1} \eta^2 \varphi'_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|) \Big| \frac{\mathbf{u}_{\boldsymbol{\ell}} - \mathbf{b}}{r_2 - r_1} \Big| \,\mathrm{d}x \,\mathrm{d}t + \int_{-\tau_2}^{s} \int_{B_{r_2}} \frac{|\mathbf{u}_{\boldsymbol{\ell}} - \mathbf{b}|^2}{\tau_2 - \tau_1} \,\mathrm{d}x \,\mathrm{d}t.
\end{aligned}$$

Finally, applying Young's inequality (2.2) with (2.3) and using the properties the cut-off functions ξ and η we have the estimate (4.3). $\square$

Considering the intrinsic cylinders with the N-functions φ and $\varphi_{|D\boldsymbol{\ell}|}$, defined by

$$Q_r^\lambda(z_0) := B_r(x_0) \times I_r^\lambda(t_0), \quad \text{where } \; I_r^\lambda(t_0) := \left(t_0 - \frac{\lambda^2}{\varphi(\lambda)} r^2, t_0 + \frac{\lambda^2}{\varphi(\lambda)} r^2 \right),$$

and

$$\mathcal{Q}_r^\lambda(z_0) := B_r(x_0) \times \mathcal{I}_r^\lambda(t_0), \quad \text{where } \; \mathcal{I}_r^\lambda(t_0) := \left(t_0 - \frac{\lambda^2}{\varphi_{|D\boldsymbol{\ell}|}(\lambda)} r^2, t_0 + \frac{\lambda^2}{\varphi_{|D\boldsymbol{\ell}|}(\lambda)} r^2 \right),$$

respectively, we obtain the following Caccioppoli-type estimates.

Corollary 4.3. *Let $\mathbf{u}$ be a weak solution to (1.1). For every $Q_R^\lambda \Subset \Omega_T$ or $\mathcal{Q}_R^\lambda \Subset \Omega_T$ with $\lambda, R > 0$ and $r \in (0,R)$ and $\mathbf{b} \in \mathbb{R}^N$, we have*

$$\begin{aligned}
&\sup_{t \in I_r^\lambda} \int_{B_r} |\mathbf{u}_{\boldsymbol{\ell}}(t) - \mathbf{b}|^2 \,\mathrm{d}x + \int_{Q_r^\lambda} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|) \,\mathrm{d}z \\
&\qquad \leqslant c \int_{Q_R^\lambda} \left[\frac{\varphi(\lambda)}{\lambda^2} \Big| \frac{\mathbf{u}_{\boldsymbol{\ell}} - \mathbf{b}}{R - r} \Big|^2 + \varphi_{|D\boldsymbol{\ell}|}\Big(\Big| \frac{\mathbf{u}_{\boldsymbol{\ell}} - \mathbf{b}}{R - r} \Big| \Big) \right] \mathrm{d}z
\end{aligned} \tag{4.4}$$

and

$$\sup_{t\in\mathcal{I}_r^\lambda}\int_{B_r}|\mathbf{u}_{\boldsymbol\ell}(t)-\mathbf{b}|^2\,\mathrm{d}x+\int_{\mathcal{Q}_r^\lambda}\varphi_{|D\boldsymbol\ell|}(|D\mathbf{u}_{\boldsymbol\ell}|)\,\mathrm{d}z \leqslant c\int_{\mathcal{Q}_R^\lambda}\left[\frac{\varphi_{|D\boldsymbol\ell|}(\lambda)}{\lambda^2}\left|\frac{\mathbf{u}_{\boldsymbol\ell}-\mathbf{b}}{R-r}\right|^2+\varphi_{|D\boldsymbol\ell|}\left(\left|\frac{\mathbf{u}_{\boldsymbol\ell}-\mathbf{b}}{R-r}\right|\right)\right]\mathrm{d}z \tag{4.5}$$

for some $c=c(n,N,p,q,L,\nu)>0$*, where* $\mathbf{u}(t)=\mathbf{u}(x,t)$*.*

In the remaining part of this section, we obtain higher integrability estimates for $\varphi_{|D\boldsymbol\ell|}(|D\mathbf{u}_{\boldsymbol\ell}|)$. Note that the higher integrability of $\varphi(|D\mathbf{u}|)$ (i.e., the case $\boldsymbol\ell=\mathbf{0}$) is proved in [33]. We then follow the argument therein.

As a first key tool, we introduce a Sobolev–Poincaré type inequality, Lemma 4.4. The proof of (4.6) could be obtained with minor modifications as in [33, Lemma 3.4], just replacing $\mathbf{u}$ and φ by $\mathbf{u}_{\boldsymbol\ell}$ and $\varphi_{|D\boldsymbol\ell|}$, respectively, and modifying the estimate for $|\langle\mathbf{u}_{\boldsymbol\ell}\rangle_\xi(t)-(\mathbf{u}_{\boldsymbol\ell})_\rho^\lambda|$ (see (4.10) below) according to assumption (2.9). However, for the reader's convenience, we prefer to provide a detailed proof. Note that the simplified version of the Poincaré inequality in (4.8) can be also found in [19, Lemma 2.9].

Let $\xi\in C_0^\infty(B_\rho)$ satisfy $0\leqslant\xi\leqslant 1$, $\xi\equiv 1$ in $B_{\rho/2}$, $|D\xi|\leqslant\frac{4}{\rho}$. Note that $2^{-n}|B_\rho|\leqslant\|\xi\|_1\leqslant|B_\rho|$. Define

$$(f)_\rho^\lambda:=\frac{1}{\|\xi\|_1}\fint_{\mathcal{I}_\rho^\lambda}\int_{B_\rho}f\xi\,\mathrm{d}x\,\mathrm{d}t\quad\text{and}\quad\langle f\rangle_\xi(t):=\frac{1}{\|\xi\|_1}\int_{B_\rho}f(x,t)\xi\,\mathrm{d}x\ \ \text{for } t\in\mathcal{I}_\rho^\lambda.$$

Lemma 4.4. *Let* $\mathbf{u}$ *be a weak solution to* (1.1)*. For an N-function* ψ *satisfying* (2.5) *with* $1\leqslant p_1\leqslant q_1$ *in place of* $1<p\leqslant q$ *,* $\mathcal{Q}_{4\rho}^\lambda\Subset\Omega_T$ *with* $\lambda>0$ *and* $\rho\leqslant r<R\leqslant 4\rho$*, we have*

$$\fint_{\mathcal{Q}_r^\lambda}\psi\left(\left|\frac{\mathbf{u}_{\boldsymbol\ell}-(\mathbf{u}_{\boldsymbol\ell})_\rho^\lambda}{r}\right|\right)\mathrm{d}z\leqslant c\psi(A_0)+c\psi\big(\mathcal{T}(r,R)^{\frac12}\big)^{(1-\theta_0)}\fint_{\mathcal{Q}_r^\lambda}\psi(|D\mathbf{u}_{\boldsymbol\ell}|)^{\theta_0}\,\mathrm{d}z, \tag{4.6}$$

for some $c=c(n,N,p,q,p_1,q_1,\theta_0,L,\nu,\Lambda)>0$ *provided that*

$$\theta_0p_1\in[1,n)\quad\textit{and}\quad\frac{nq_1}{nq_1+2p_1}\leqslant\theta_0\leqslant 1.$$

Here $(\mathbf{u}_{\boldsymbol\ell})_\rho^\lambda$ *it the average of* $\mathbf{u}$ *on* $\mathcal{Q}_\rho^\lambda$*,*

$$A_0:=\frac{\lambda^2}{\varphi_{|D\boldsymbol\ell|}(\lambda)}\fint_{\mathcal{Q}_r^\lambda}\varphi'_{|D\boldsymbol\ell|}(|D\mathbf{u}_{\boldsymbol\ell}|)\,\mathrm{d}z, \tag{4.7}$$

$$\mathcal{T}(r,R):=\fint_{\mathcal{Q}_R^\lambda}\left[\left|\frac{\mathbf{u}_{\boldsymbol\ell}-(\mathbf{u}_{\boldsymbol\ell})_\rho^\lambda}{R-r}\right|^2+\frac{\lambda^2}{\varphi_{|D\boldsymbol\ell|}(\lambda)}\varphi_{|D\boldsymbol\ell|}\left(\left|\frac{\mathbf{u}_{\boldsymbol\ell}-(\mathbf{u}_{\boldsymbol\ell})_\rho^\lambda}{R-r}\right|\right)\right]\mathrm{d}z+A_0^2.$$

In particular, when $\theta_0=p_1=1$*, we have*

$$\fint_{\mathcal{Q}_r^\lambda}\psi\left(\left|\frac{\mathbf{u}_{\boldsymbol\ell}-(\mathbf{u}_{\boldsymbol\ell})_\rho^\lambda}{r}\right|\right)\mathrm{d}z\leqslant c\fint_{\mathcal{Q}_r^\lambda}\psi(|D\mathbf{u}_{\boldsymbol\ell}|)\,\mathrm{d}z+c\psi(A_0). \tag{4.8}$$

Proof. The triangle inequality implies

$$\begin{aligned}\fint_{\mathcal{Q}_r^\lambda}\psi\left(\left|\frac{\mathbf{u}_{\boldsymbol\ell}-(\mathbf{u}_{\boldsymbol\ell})_\rho^\lambda}{r}\right|\right)\mathrm{d}z&=\fint_{\mathcal{Q}_r^\lambda}\psi\left(\left|\frac{\mathbf{u}_{\boldsymbol\ell}(z)-\langle\mathbf{u}_{\boldsymbol\ell}\rangle_\xi(t)+\langle\mathbf{u}_{\boldsymbol\ell}\rangle_\xi(t)-(\mathbf{u}_{\boldsymbol\ell})_\rho^\lambda}{r}\right|\right)\mathrm{d}z\\&\leqslant c\fint_{\mathcal{I}_r^\lambda}\psi\left(\left|\frac{\langle\mathbf{u}_{\boldsymbol\ell}\rangle_\xi(t)-(\mathbf{u}_{\boldsymbol\ell})_\rho^\lambda}{r}\right|\right)\mathrm{d}t+c\fint_{\mathcal{Q}_r^\lambda}\psi\left(\left|\frac{\mathbf{u}_{\boldsymbol\ell}(z)-\langle\mathbf{u}_{\boldsymbol\ell}\rangle_\xi(t)}{r}\right|\right)\mathrm{d}z.\end{aligned} \tag{4.9}$$

We start with an estimate of the first term in the right hand side above. By the definition of $\langle \mathbf{u}_{\boldsymbol\ell}\rangle_\xi$ and using the weak formulation of (1.1) with test-function $\boldsymbol\zeta(x,t) := (\xi(x),\dots,\xi(x))$, we find from (2.9) that

$$
\begin{aligned}
|\langle \mathbf{u}_{\boldsymbol\ell}\rangle_\xi(t) - (\mathbf{u}_{\boldsymbol\ell})^\lambda_\rho| &\leqslant \sup_{\tau\in\mathcal{I}^\lambda_r} |\langle \mathbf{u}_{\boldsymbol\ell}\rangle_\xi(t) - \langle \mathbf{u}_{\boldsymbol\ell}\rangle_\xi(\tau)| = \sup_{\tau\in\mathcal{I}^\lambda_r}\left|\int_\tau^t \partial_t \langle \mathbf{u}_{\boldsymbol\ell}\rangle_\xi(s)\,\mathrm{d}s\right| \\
&= \sup_{\tau\in\mathcal{I}^\lambda_r}\left|\int_\tau^t \frac{1}{\|\eta\|_1}\int_{B_r} \partial_t \mathbf{u}_{\boldsymbol\ell}(x,s)\xi(x)\,\mathrm{d}x\,\mathrm{d}s\right| \\
&= \sup_{\tau\in\mathcal{I}^\lambda_r}\left|\int_\tau^t \frac{1}{\|\eta\|_1}\int_{B_r} \partial_t \mathbf{u}(x,s)\xi(x)\,\mathrm{d}x\,\mathrm{d}s\right| \\
&\approx \sup_{\tau\in\mathcal{I}^\lambda_r}\left|\int_\tau^t ⨍_{B_r} (\mathbf{A}(D\mathbf{u}) - \mathbf{A}(D\boldsymbol\ell))D\xi\,\mathrm{d}x\,\mathrm{d}s\right| \\
&\leqslant \frac{cr\lambda^2}{\varphi_{|D\boldsymbol\ell|}(\lambda)} ⨍_{\mathcal{Q}^\lambda_r} \varphi'_{|D\boldsymbol\ell|}(|D\mathbf{u}_{\boldsymbol\ell}|)\,\mathrm{d}z = rA_0.
\end{aligned}
\tag{4.10}
$$

We next estimate the second term in (4.9). From the Gagliardo–Nirenberg type inequality in Lemma 4.1 with $(\psi,\gamma,p,q_1,q_2) := (\psi^{1/p_1}, p_1, \theta_0 p_1, \frac{q_1}{p_1}, 2)$ we conclude that

$$
⨍_{B_r} \psi\big(\big|\tfrac{f}{r}\big|\big)\,\mathrm{d}x \leqslant c\left(⨍_{B_r}\Big[\psi(|Df|)^{\theta_1} + \psi\big(\big|\tfrac{f}{r}\big|\big)^{\theta_0}\Big]\,\mathrm{d}x\right)\ \psi\left(\left[⨍_{B_r}\big|\tfrac{f}{r}\big|^2\,\mathrm{d}x\right]^{\frac12}\right)^{1-\theta_0}
\tag{4.11}
$$

provided $\theta_0 p_1 \in [1,n)$ and

$$
\frac{1}{p_1} \geqslant \frac{\theta_0}{(\theta_0 p_1)^*} + \frac{1-\theta_0}{2}\frac{q_1}{p_1} = \frac{1}{p_1} - \frac{\theta_0}{n} + \frac{1-\theta_0}{2}\frac{q_1}{p_1}.
$$

This can be written as $\theta_0 \geqslant \frac{nq_1}{nq_1+2p_1}$. Applying (4.11) with $f := \mathbf{u}_{\boldsymbol\ell} - \langle \mathbf{u}_{\boldsymbol\ell}\rangle_\xi$ on each time slice gives

$$
\begin{aligned}
⨍_{\mathcal{Q}^\lambda_r} \psi\Big(\Big|\frac{\mathbf{u}_{\boldsymbol\ell}(z) - \langle \mathbf{u}_{\boldsymbol\ell}\rangle_\xi(t)}{r}\Big|\Big)\,\mathrm{d}z &\leqslant c\left(⨍_{\mathcal{Q}^\lambda_r}\left[\psi(|D\mathbf{u}_{\boldsymbol\ell}|)^{\theta_0} + \psi\left(\frac{\mathbf{u}_{\boldsymbol\ell} - \langle \mathbf{u}_{\boldsymbol\ell}\rangle_\xi}{r}\right)^{\theta_0}\right]\mathrm{d}z\right) \\
&\qquad \times \psi\left(\left(\sup_{t\in\mathcal{I}^\lambda_r} ⨍_{B_r}\left|\frac{\mathbf{u}_{\boldsymbol\ell} - \langle \mathbf{u}_{\boldsymbol\ell}\rangle_\xi}{r}\right|^2\mathrm{d}x\right)^{\frac12}\right)^{1-\theta_0}.
\end{aligned}
\tag{4.12}
$$

Note that for each time slice of $\mathcal{Q}^\lambda_r$, since $\theta_0 p_1 \geqslant 1$, we can apply the weighted Poincaré inequality [14, Theorem 7], so that

$$
⨍_{B_r} \psi\Big(\Big|\frac{\mathbf{u}_{\boldsymbol\ell}(x,t) - \langle \mathbf{u}_{\boldsymbol\ell}\rangle_\xi(t)}{r}\Big|\Big)^{\theta_0}\mathrm{d}x \leqslant c ⨍_{B_r} \psi(|D\mathbf{u}_{\boldsymbol\ell}(x,t)|)^{\theta_0}\,\mathrm{d}x\,.
$$

Finally, from the Caccioppoli inequality in (4.5) with $\mathbf{b} := (\mathbf{u}_{\boldsymbol\ell})^\lambda_\rho$ and (4.10) we conclude that

$$
\begin{aligned}
&\sup_{t\in\mathcal{I}^\lambda_r} ⨍_{B_r}\left|\frac{\mathbf{u}_{\boldsymbol\ell} - \langle \mathbf{u}_{\boldsymbol\ell}\rangle_\xi}{r}\right|^2\mathrm{d}x \\
&\leqslant c\sup_{t\in\mathcal{I}^\lambda_r} ⨍_{B_r}\left|\frac{\mathbf{u}_{\boldsymbol\ell}(x,t) - (\mathbf{u}_{\boldsymbol\ell})^\lambda_\rho}{r}\right|^2\mathrm{d}x + c\sup_{t\in\mathcal{I}^\lambda_r}\left|\frac{(\mathbf{u}_{\boldsymbol\ell})^\lambda_\rho - \langle \mathbf{u}_{\boldsymbol\ell}\rangle_\xi(t)}{r}\right|^2 \\
&\leqslant c ⨍_{\mathcal{Q}^\lambda_R}\left[\left|\frac{\mathbf{u}_{\boldsymbol\ell} - (\mathbf{u}_{\boldsymbol\ell})^\lambda_\rho}{R-r}\right|^2 + \frac{\lambda^2}{\varphi_{|D\boldsymbol\ell|}(\lambda)}\varphi_{|D\boldsymbol\ell|}\left(\left|\frac{\mathbf{u}_{\boldsymbol\ell} - (\mathbf{u}_{\boldsymbol\ell})^\lambda_\rho}{R-r}\right|\right)\right]\mathrm{d}z + cA_0^2\,.
\end{aligned}
$$

Therefore, inserting the above two estimates into (4.12) and combining with (4.9)–(4.10), we complete the proof of (4.6). □

The next two lemmas show that the right hand side of the estimate in Lemma 4.4 can be controlled by suitable quantities when we are in suitable intrinsic cylinders.

Lemma 4.5. *Let the assumptions of Lemma 4.4 be in force, and assume additionally that*

$$\fint_{Q^\lambda_{4\rho}} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)\,\mathrm{d}z \leqslant \varphi_{|D\boldsymbol{\ell}|}(\lambda)\,.$$

Then, for some $c = c(n, N, p, q, p_1, q_1, \theta_0, L, \nu, \Lambda) > 0$,

$$\fint_{Q^\lambda_{2\rho}} \psi\left(\left|\frac{\mathbf{u}_{\boldsymbol{\ell}} - (\mathbf{u}_{\boldsymbol{\ell}})^\lambda_\rho}{\rho}\right|\right)\mathrm{d}z \leqslant c\psi(A_0) + c\psi(\lambda)^{1-\theta_0} \fint_{Q^\lambda_{2\rho}} \psi(|D\mathbf{u}_{\boldsymbol{\ell}}|)^{\theta_0}\,\mathrm{d}z,$$

where A_0 is that of (4.7) with $r = 2\rho$.

Proof. The proof is exactly the same as the one of [33, Lemma 3.9] with $\mathbf{u}_{\boldsymbol{\ell}}$ and $\varphi_{|D\boldsymbol{\ell}|}$ in place of $\mathbf{u}$ and φ, respectively. □

Finally, we obtain a reverse Hölder inequality for $\varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)$. The proof is almost the same as that of [33, Lemma 3.12]. The main difference is the use of the Caccioppoli estimate (4.4) in place of the usual one [33, Lemma 3.1].

Lemma 4.6. *Let $\mathbf{u}$ be a weak solution to (1.1) and $Q^\lambda_{4\rho} \Subset \Omega_I$ with $\lambda, \rho > 0$. Suppose that*

$$\varphi_{|D\boldsymbol{\ell}|}(\lambda) \leqslant \fint_{Q^\lambda_{\rho}} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)\,\mathrm{d}z \quad \text{and} \quad \fint_{Q^\lambda_{4\rho}} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)\,\mathrm{d}z \leqslant \varphi_{|D\boldsymbol{\ell}|}(\lambda). \tag{4.13}$$

Then there exist $\theta = \theta(n, p, q) \in (0, 1)$ and $c = c(n, N, p, q, L, \nu, \Lambda) > 0$ such that

$$\fint_{Q^\lambda_{\rho}} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)\,\mathrm{d}z \leqslant c\left(\fint_{Q^\lambda_{4\rho}} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)^\theta\,\mathrm{d}z\right)^{\frac{1}{\theta}}. \tag{4.14}$$

Proof. We denote $p_0 := \frac{2n}{n+2}$, and recall A_0 in (4.7). Arguing as in [33, Lemma 2.9] we have that for every $\delta \in (0, 1)$ and $\theta_0 \in (1 - \frac{1}{q}, 1]$,

$$A_0 \leqslant \begin{cases} \delta\lambda + c_\delta \varphi^{-1}_{|D\boldsymbol{\ell}|}\left(\left(\fint_{Q^\lambda_{2\rho}} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)^{\theta_0}\,\mathrm{d}z\right)^{\frac{1}{\theta_0}}\right), \\ c\lambda\,. \end{cases} \tag{4.15}$$

By the Caccioppoli inequality (4.4) with $\mathbf{b} := (\mathbf{u}_{\boldsymbol{\ell}})^\lambda_\rho$, we find that

$$\fint_{Q^\lambda_{\rho}} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)\,\mathrm{d}z \leqslant c\frac{\varphi_{|D\boldsymbol{\ell}|}(\lambda)}{\lambda^2} \fint_{Q^\lambda_{2\rho}} \left|\frac{\mathbf{u}_{\boldsymbol{\ell}} - (\mathbf{u}_{\boldsymbol{\ell}})^\lambda_\rho}{\rho}\right|^2 \mathrm{d}z + c\fint_{Q^\lambda_{2\rho}} \varphi_{|D\boldsymbol{\ell}|}\left(\left|\frac{\mathbf{u}_{\boldsymbol{\ell}} - (\mathbf{u}_{\boldsymbol{\ell}})^\lambda_\rho}{\rho}\right|\right)\mathrm{d}z. \tag{4.16}$$

We then estimate the two integrals in the right hand side.

By Lemma 4.5 for $\psi := \varphi_{|D\boldsymbol{\ell}|}$, considering also (4.15) and the classical Young's inequality for conjugate exponents $\frac{1}{\theta_0}, \frac{1}{1-\theta_0}$, we have that for any $\delta \in (0, 1)$

$$\begin{aligned} \fint_{Q^\lambda_{2\rho}} \varphi_{|D\boldsymbol{\ell}|}\left(\left|\frac{\mathbf{u}_{\boldsymbol{\ell}} - (\mathbf{u}_{\boldsymbol{\ell}})^\lambda_\rho}{\rho}\right|\right)\mathrm{d}z &\leqslant c\varphi_{|D\boldsymbol{\ell}|}(A_0) + c\varphi_{|D\boldsymbol{\ell}|}(\lambda)^{(1-\theta_0)} \fint_{Q^\lambda_{2\rho}} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)^{\theta_0}\,\mathrm{d}z \\ &\leqslant c_\delta \left(\fint_{Q^\lambda_{2\rho}} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)^{\theta_0}\,\mathrm{d}z\right)^{\frac{1}{\theta_0}} + c\delta\varphi_{|D\boldsymbol{\ell}|}(\lambda)\,. \end{aligned} \tag{4.17}$$

An analogous argument in the case $\psi(t) := t^2$ and $\theta_0 := \frac{p_0}{2}$, shows that for any $\delta \in (0,1)$

$$\left(\fint_{\mathcal{Q}^\lambda_{2\rho}} \left|\frac{\mathbf{u}_{\boldsymbol{\ell}} - (\mathbf{u}_{\boldsymbol{\ell}})^\lambda_\rho}{\rho}\right|^2 \mathrm{d}z\right)^{\frac{1}{2}} \leqslant cA_0 + c\left(\lambda^{2-p_0} \fint_{\mathcal{Q}^\lambda_{2\rho}} |D\mathbf{u}_{\boldsymbol{\ell}}|^{p_0}\,\mathrm{d}z\right)^{\frac{1}{2}}$$
$$\leqslant c_\delta \left(\fint_{\mathcal{Q}^\lambda_{2\rho}} |D\mathbf{u}_{\boldsymbol{\ell}}|^{p_0}\,\mathrm{d}z\right)^{\frac{1}{p_0}} + cA_0 + \delta\lambda.$$

In particular, we also have

$$\left(\fint_{\mathcal{Q}^\lambda_{2\rho}} \left|\frac{\mathbf{u}_{\boldsymbol{\ell}} - (\mathbf{u}_{\boldsymbol{\ell}})^\lambda_\rho}{\rho}\right|^2 \mathrm{d}z\right)^{\frac{1}{2}} \leqslant c\lambda.$$

Now, multiplying the previous two inequalities and using Young's inequality (2.2), (2.3), (2.12), the Jensen inequality in Lemma 2.1 with $\psi(t) := \varphi_{|D\boldsymbol{\ell}|}(t^{\frac{1}{p_0}})^{\theta_0}$ with $\theta_0 \in (0,1)$ sufficiently close to 1 and (4.15), we obtain that for any $\delta \in (0,1)$

(4.18)

$$\frac{\varphi_{|D\boldsymbol{\ell}|}(\lambda)}{\lambda^2} \fint_{\mathcal{Q}^\lambda_{2\rho}} \left|\frac{\mathbf{u}_{\boldsymbol{\ell}} - (\mathbf{u}_{\boldsymbol{\ell}})^\lambda_\rho}{\rho}\right|^2 \mathrm{d}z \leqslant c\varphi'_{|D\boldsymbol{\ell}|}(\lambda)\left[c_\delta\left(\fint_{\mathcal{Q}^\lambda_{2\rho}} |D\mathbf{u}_{\boldsymbol{\ell}}|^{p_0}\,\mathrm{d}z\right)^{\frac{1}{p_0}} + A_0 + \delta\lambda\right]$$
$$\leqslant c_\delta \varphi_{|D\boldsymbol{\ell}|}\left(\left(\fint_{\mathcal{Q}^\lambda_{2\rho}} |D\mathbf{u}_{\boldsymbol{\ell}}|^{p_0}\,\mathrm{d}z\right)^{\frac{1}{p_0}}\right) + c_\delta\varphi_{|D\boldsymbol{\ell}|}(A_0) + c\delta\varphi_{|D\boldsymbol{\ell}|}(\lambda)$$
$$\leqslant c_\delta \left(\fint_{\mathcal{Q}^\lambda_{2\rho}} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)^{\theta_0}\,\mathrm{d}z\right)^{\frac{1}{\theta_0}} + c\delta\varphi_{|D\boldsymbol{\ell}|}(\lambda).$$

Finally, inserting (4.17) and (4.18) into (4.16), we find that

$$\fint_{\mathcal{Q}^\lambda_{\rho}} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)\,\mathrm{d}z \leqslant c_\delta \left(\fint_{\mathcal{Q}^\lambda_{2\rho}} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)^{\theta_0}\,\mathrm{d}z\right)^{\frac{1}{\theta_0}} + c\delta\varphi_{|D\boldsymbol{\ell}|}(\lambda)\,.$$

Choosing δ so small that $c\delta = \frac{1}{2}$ and absorbing the term in the left-hand side by (4.13) we obtain the reverse Hölder inequality (4.14). $\square$

Finally, by arguing exactly as in [33, Section 4] with $\varphi_{|D\boldsymbol{\ell}|}$ and $\mathbf{u}_{\boldsymbol{\ell}}$ in place of φ and $\mathbf{u}$, respectively, we have the following higher integrability result for $D\mathbf{u}_{\boldsymbol{\ell}}$.

Theorem 4.7. *Let $\mathbf{u}$ be a local weak solution to (1.1). There exists $\sigma = \sigma(n,N,p,q,L,\nu) > 0$ such that $\varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|) \in L^{1+\sigma}_{loc}(\Omega_T)$ with the following estimate: for any $Q_{4\rho} \Subset \Omega_T$,*

$$\fint_{Q_\rho} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)^{1+\sigma}\,\mathrm{d}z \leqslant c\left[(\varphi_{|D\boldsymbol{\ell}|}\circ\mathcal{D}^{-1})\left(\fint_{Q_{2\rho}} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)\,\mathrm{d}z\right)\right]^\sigma \fint_{Q_{2\rho}} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)\,\mathrm{d}z$$

for some $c = c(n,N,p,q,L,\nu,\Lambda) > 0$, where $\mathcal{D}(t) := \min\{t^2, \varphi_{|D\boldsymbol{\ell}|}(t)^{\frac{n+2}{2}}t^{-n}\}$ and $\mathcal{D}^{-1}$ is the inverse of $\mathcal{D}$.

Moreover, by a scaling argument, we have the following homogeneous higher integrability result in intrinsic parabolic cylinders with φ.

Corollary 4.8. *Let $\mathbf{u}$ be a local weak solution to (1.1). There exists $\sigma = \sigma(n,N,p,q,L,\nu) > 0$ such that if $Q^\lambda_{4\rho} \Subset \Omega_T$ and*

$$\fint_{Q^\lambda_{4\rho}} \varphi(|D\mathbf{u}|)\,\mathrm{d}z \leqslant \varphi(\lambda) \quad \textit{and} \quad |D\boldsymbol{\ell}| \leqslant \lambda, \tag{4.19}$$

then

$$\left(\fint_{Q_\rho^\lambda} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)^{1+\sigma}\,\mathrm{d}z \right)^{\frac{1}{1+\sigma}} \leqslant c\varphi(\lambda) \tag{4.20}$$

for some $c = c(n, N, p, q, L, \nu) > 0$.

Proof. Let

$$\tilde{\mathbf{u}}(x,t) := \frac{1}{\lambda}\mathbf{u}(x, t\lambda^2/\varphi(\lambda)), \quad \tilde{\mathbf{A}}(\mathbf{P}) := \frac{\lambda \mathbf{A}(\lambda \mathbf{P})}{\varphi(\lambda)}, \quad \tilde{\varphi}(\tau) := \frac{\varphi(\lambda\tau)}{\varphi(\lambda)}, \quad \tilde{\boldsymbol{\ell}} := \frac{1}{\lambda}\boldsymbol{\ell}. \tag{4.21}$$

Note that $\tilde{\mathbf{A}}$ satisfies the same properties of $\mathbf{A}$ listed in Assumption (A), with $\tilde{\varphi}$ in place of φ. Then $\tilde{\mathbf{u}}$ is a weak solution to

$$\partial_t \tilde{\mathbf{u}} - \operatorname{div} \tilde{\mathbf{A}}(D\tilde{\mathbf{u}}) = \mathbf{0} \quad \text{in} \ \ Q_{4\rho}\,.$$

Moreover, by (4.19), we have

$$\fint_{Q_{4\rho}} \tilde{\varphi}(|D\tilde{\mathbf{u}}|)\,\mathrm{d}z \leqslant 1 \quad \text{and} \quad |D\tilde{\boldsymbol{\ell}}| \leqslant 1\,,$$

whence, taking into account (2.14),

$$\fint_{Q_{4\rho}} \tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(|D\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}}|)\,\mathrm{d}z \leqslant c \fint_{Q_{4\rho}} \tilde{\varphi}(|D\tilde{\mathbf{u}}| + |D\tilde{\boldsymbol{\ell}}|)\,\mathrm{d}z \leqslant c\,.$$

Therefore, by Theorem 4.7 we have

$$\left(\fint_{Q_\rho} \tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(|D\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}}|)^{1+\sigma}\,\mathrm{d}z \right)^{\frac{1}{1+\sigma}} \leqslant c\,. \tag{4.22}$$

In addition, since by (2.13) and (4.21)

$$\tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(\tau) \sim \frac{\varphi'(|D\boldsymbol{\ell}| + \lambda\tau)}{\varphi(\lambda)(|D\boldsymbol{\ell}| + \lambda\tau)}(\lambda\tau)^2\,,$$

we have, taking into account also (2.12),

$$\tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(|D\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}}|) \sim \frac{\varphi'(|D\boldsymbol{\ell}| + |D\mathbf{u}_{\boldsymbol{\ell}}|)}{\varphi(\lambda)(|D\boldsymbol{\ell}| + |D\mathbf{u}_{\boldsymbol{\ell}}|)}|D\mathbf{u}_{\boldsymbol{\ell}}|^2 \sim \frac{\varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)}{\varphi(\lambda)}\,.$$

Therefore, inserting the above estimate in (4.22) we obtain (4.20). □

5. Nondegenerate regime

In this section we consider the *nondegenerate* regime, which means that the average of the gradient of solution is relatively greater than the relevant excess, see for instance (5.6). In this regime, we apply the $\mathcal{A}$-caloric approximation.

We first show that the solution $\mathbf{u}$ to (1.1) is an almost weak solution of a linear system with constant coefficients.

Lemma 5.1. *Let* $\mathbf{u}$ *be a weak solution to* (1.1) *and* $Q_r^\lambda \Subset \Omega_T$*. Then for every* $\boldsymbol{\zeta} \in C^\infty(Q_r^\lambda;\mathbb{R}^N)$ *with* $\boldsymbol{\zeta} = \mathbf{0}$ *on* $\partial B_r \times I_r^\lambda$*, we have*

$$\begin{aligned} &\frac{1}{Q_r^\lambda}\left| \int_{Q_r^\lambda} \mathbf{u}_{\boldsymbol{\ell}} \cdot \boldsymbol{\zeta}_t - D\mathbf{A}(D\boldsymbol{\ell})\langle D\mathbf{u}_{\boldsymbol{\ell}}, D\boldsymbol{\zeta}\rangle\,\mathrm{d}z - \left[\int_{B_r} \mathbf{u}_{\boldsymbol{\ell}} \cdot \boldsymbol{\zeta}\,\mathrm{d}x \right]_{t=-r^2/\varphi''(\lambda)}^{t=r^2/\varphi''(\lambda)} \right| \\ &\leqslant c\varphi'(|D\boldsymbol{\ell}|)\,(\mu^\gamma + \mu)\,\mu\,\|D\boldsymbol{\zeta}\|_{L^\infty(Q_r^\lambda;\mathbb{R}^{Nn})}\,, \end{aligned} \tag{5.1}$$

where $\boldsymbol{\ell}$ *and* $\mathbf{u}_{\boldsymbol{\ell}}$ *are from* (4.1) *and* (4.2) *and*

$$\mu := \left(\frac{1}{\varphi(|D\boldsymbol{\ell}|)} \fint_{Q_r^\lambda} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|) \,\mathrm{d}z \right)^{\frac{1}{2}} .$$

Proof. It is enough to consider $\boldsymbol{\zeta} \in C^\infty(Q_r^\lambda; \mathbb{R}^N)$ with $\|D\boldsymbol{\zeta}\|_{L^\infty(Q_r^\lambda;\mathbb{R}^{Nn})} \leqslant 1$ by linearity. From the weak form of (1.1) and the fact that $\boldsymbol{\ell}_t = \operatorname{div}(\mathbf{A}(D\boldsymbol{\ell})) = \mathbf{0}$, we observe that

$$\begin{aligned}
&\fint_{Q_r^\lambda} \mathbf{u}_{\boldsymbol{\ell}} \cdot \boldsymbol{\zeta}_t - D\mathbf{A}(D\boldsymbol{\ell})\langle D\mathbf{u}_{\boldsymbol{\ell}}, D\boldsymbol{\zeta}\rangle \,\mathrm{d}z - \left[\int_{B_r} \mathbf{u}_{\boldsymbol{\ell}} \cdot \boldsymbol{\zeta} \,\mathrm{d}x \right]_{t=-r^2\lambda^2/\varphi(\lambda)}^{t=r^2\lambda^2/\varphi(\lambda)} \\
&= \fint_{Q_r^\lambda} \mathbf{u} \cdot \boldsymbol{\zeta}_t - D\mathbf{A}(D\boldsymbol{\ell})\langle D\mathbf{u}_{\boldsymbol{\ell}}, D\boldsymbol{\zeta}\rangle \,\mathrm{d}z - \left[\int_{B_r} \mathbf{u} \cdot \boldsymbol{\zeta} \,\mathrm{d}x \right]_{t=-r^2/\lambda^2\varphi(\lambda)}^{t=r^2\lambda^2/\varphi(\lambda)} \\
&= \fint_{Q_r^\lambda} \langle \mathbf{A}(D\mathbf{u}) - \mathbf{A}(D\boldsymbol{\ell}), D\boldsymbol{\zeta}\rangle - D\mathbf{A}(D\boldsymbol{\ell})\langle D\mathbf{u}_{\boldsymbol{\ell}}, D\boldsymbol{\zeta}\rangle \,\mathrm{d}z \\
&= \fint_{Q_r^\lambda} \int_0^1 \langle [D\mathbf{A}(sD\mathbf{u}_{\boldsymbol{\ell}} + D\boldsymbol{\ell}) - D\mathbf{A}(D\boldsymbol{\ell})] D\mathbf{u}_{\boldsymbol{\ell}}, D\boldsymbol{\zeta}\rangle \,\mathrm{d}s \,\mathrm{d}z \\
&\leqslant \fint_{Q_r^\lambda} \left[\int_0^1 |D\mathbf{A}(sD\mathbf{u}_{\boldsymbol{\ell}} + D\boldsymbol{\ell}) - D\mathbf{A}(D\boldsymbol{\ell})| \,\mathrm{d}s \right] |D\mathbf{u}_{\boldsymbol{\ell}}||D\boldsymbol{\zeta}| \,\mathrm{d}z \,.
\end{aligned} \tag{5.2}$$

Set $S_1 = \{z \in Q_r^\lambda : |D\mathbf{u}_{\boldsymbol{\ell}}(z)| > \frac{1}{2}|D\boldsymbol{\ell}|\}$ and $S_2 = \{z \in Q_r^\lambda : |D\mathbf{u}_{\boldsymbol{\ell}}(z)| \leqslant \frac{1}{2}|D\boldsymbol{\ell}|\}$.

If $z \in S_1$, using (2.8) and the fact that $|D\mathbf{u}| + |D\boldsymbol{\ell}| \leqslant |D\mathbf{u}_{\boldsymbol{\ell}}| + 2|D\boldsymbol{\ell}| \leqslant 5|D\mathbf{u}_{\boldsymbol{\ell}}|$,

$$\begin{aligned}
&\int_0^1 |D\mathbf{A}(sD\mathbf{u}_{\boldsymbol{\ell}}(z) + D\boldsymbol{\ell}) - D\mathbf{A}(D\boldsymbol{\ell})| \,\mathrm{d}s \\
&\leqslant c \int_0^1 \varphi''(|sD\mathbf{u}(z) + (1-s)D\boldsymbol{\ell}|) \,\mathrm{d}s + c\varphi''(|D\boldsymbol{\ell}|) \\
&\leqslant c \frac{\varphi'(|D\mathbf{u}(z)| + |D\boldsymbol{\ell}|)}{|D\mathbf{u}(z)| + |D\boldsymbol{\ell}|} + c \frac{\varphi'(|D\mathbf{u}(z)| + |D\boldsymbol{\ell}|)}{|D\boldsymbol{\ell}|} \\
&\leqslant \frac{c}{|D\boldsymbol{\ell}|} \varphi'(|D\mathbf{u}_{\boldsymbol{\ell}}(z)|) \leqslant \frac{c}{|D\boldsymbol{\ell}|} \varphi'(|D\mathbf{u}_{\boldsymbol{\ell}}(z)| + |D\boldsymbol{\ell}|) \frac{5|D\mathbf{u}_{\boldsymbol{\ell}}(z)|}{|D\mathbf{u}_{\boldsymbol{\ell}}(z)| + 2|D\boldsymbol{\ell}|} \leqslant \frac{c}{|D\boldsymbol{\ell}|} \varphi'_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}(z)|) \,,
\end{aligned}$$

hence

$$\fint_{Q_r^\lambda} \left[\int_0^1 |D\mathbf{A}(sD\mathbf{u}_{\boldsymbol{\ell}} + D\boldsymbol{\ell}) - D\mathbf{A}(D\boldsymbol{\ell})| \,\mathrm{d}s \right] |D\mathbf{u}_{\boldsymbol{\ell}}| \chi_{S_1} \,\mathrm{d}z \leqslant c\varphi'(|D\boldsymbol{\ell}|)\mu^2 \,. \tag{5.3}$$

On the other hand, if $z \in S_2$, applying (2.17) with $\mathbf{P} = D\boldsymbol{\ell}$ and $\mathbf{Q} = sD\mathbf{u}_{\boldsymbol{\ell}}(z) + D\boldsymbol{\ell}$

$$\int_0^1 |D\mathbf{A}(sD\mathbf{u}_{\boldsymbol{\ell}}(z) + D\boldsymbol{\ell}) - D\mathbf{A}(D\boldsymbol{\ell})| \,\mathrm{d}s \leqslant c \left(\frac{|D\mathbf{u}_{\boldsymbol{\ell}}|}{|D\boldsymbol{\ell}|} \right)^\gamma \varphi''(|D\boldsymbol{\ell}|)$$

and

$$\begin{aligned}
\frac{|D\mathbf{u}_{\boldsymbol{\ell}}(z)|^2}{|D\boldsymbol{\ell}|^2} &\leqslant \frac{\varphi'(|\mathbf{u}_{\boldsymbol{\ell}}(z)| + |D\boldsymbol{\ell}|)}{\varphi'(|D\boldsymbol{\ell}|)|D\boldsymbol{\ell}|} \frac{|D\mathbf{u}_{\boldsymbol{\ell}}(z)|^2}{|D\boldsymbol{\ell}|} \\
&\leqslant \frac{\varphi'(|D\mathbf{u}_{\boldsymbol{\ell}}(z)| + |D\boldsymbol{\ell}|)}{\varphi'(|D\boldsymbol{\ell}|)|D\boldsymbol{\ell}|} \frac{|D\mathbf{u}_{\boldsymbol{\ell}}(z)|^2}{|D\mathbf{u}_{\boldsymbol{\ell}}| + \frac{1}{2}|D\boldsymbol{\ell}|} \leqslant c \frac{\varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)}{\varphi(|D\boldsymbol{\ell}|)} \,.
\end{aligned}$$

Then using these estimates and the fact that $\omega(\cdot) \leqslant 1$ and applying Hölder's inequality and Jensen's inequality to the concave function $\tau \mapsto \omega(\tau^{1/2})$, we have

$$\begin{aligned}
&\fint_{Q_r^\lambda} \left[\int_0^1 |D\mathbf{A}(sD\mathbf{u}_{\boldsymbol{\ell}} + D\boldsymbol{\ell}) - D\mathbf{A}(D\boldsymbol{\ell})|\,\mathrm{d}s\right] |D\mathbf{u}_{\boldsymbol{\ell}}|\chi_{S_2}\,\mathrm{d}z \\
&\leqslant c\varphi'(|D\boldsymbol{\ell}|) \fint_{Q_r^\lambda} \frac{|D\mathbf{u}_{\boldsymbol{\ell}}(z)|}{|D\boldsymbol{\ell}|} \left(\frac{|D\mathbf{u}_{\boldsymbol{\ell}}(z)|}{|D\boldsymbol{\ell}|}\right)^\gamma \chi_{S_2}\,\mathrm{d}z \\
&\leqslant c\varphi'(|D\boldsymbol{\ell}|) \left[\fint_{Q_r^\lambda} \frac{|D\mathbf{u}_{\boldsymbol{\ell}}(z)|^2}{|D\boldsymbol{\ell}|^2}\,\mathrm{d}z\right]^{\frac12} \left[\fint_{Q_r^\lambda} \left(\frac{|D\mathbf{u}_{\boldsymbol{\ell}}(z)|}{|D\boldsymbol{\ell}|}\right)^{2\gamma}\mathrm{d}z\right]^{\frac12} \\
&\leqslant c\varphi'(|D\boldsymbol{\ell}|) \left[\fint_{Q_r^\lambda} \frac{\varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)}{\varphi(|D\boldsymbol{\ell}|)}\,\mathrm{d}z\right]^{\frac12} \left[\fint_{Q_r^\lambda} \frac{\varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)}{\varphi(|D\boldsymbol{\ell}|)}\,\mathrm{d}z\right]^{\frac{\gamma}{2}} \\
&\leqslant c\varphi'(|D\boldsymbol{\ell}|)\mu^{1+\gamma}\,.
\end{aligned} \tag{5.4}$$

Therefore, plugging (5.3) and (5.4) into (5.2) we obtain (5.1). □

Now, we derive an excess decay estimate in the non-degenerate regime.

Lemma 5.2. *Let $Q_{2r}^\lambda = Q_{2r}^\lambda(z_0) \Subset \Omega_T$, $\beta \in (0,1)$, and $\mathbf{u}$ be a weak solution to (1.1). Suppose that*

$$\frac{\lambda}{2K} \leqslant |(D\mathbf{u})_{2r}^\lambda| \leqslant 2K\lambda \tag{5.5}$$

for some $K > 0$. There exist small $\delta_0, \theta \in (0,1)$ depending on $n, N, p, q, L, \nu, K, \gamma$ and β such that if

$$\fint_{Q_r^\lambda} \varphi_{|(D\mathbf{u})_r^\lambda|}(|D\mathbf{u} - (D\mathbf{u})_r^\lambda|)\,\mathrm{d}z \leqslant \delta_0\varphi(|(D\mathbf{u})_r^\lambda|) \tag{5.6}$$

then

$$\fint_{Q_{\theta r}^\lambda} \varphi_{|(D\mathbf{u})_{\theta r}^\lambda|}(|D\mathbf{u} - (D\mathbf{u})_{\theta r}^\lambda|)\,\mathrm{d}z \leqslant \theta^{2\beta} \fint_{Q_r^\lambda} \varphi_{|(D\mathbf{u})_r^\lambda|}(|D\mathbf{u} - (D\mathbf{u})_r^\lambda|)\,\mathrm{d}z\,. \tag{5.7}$$

Proof. For simplicity, we assume that $z_0 = (x_0, t_0) = (0,0)$. We fix the linear function

$$\boldsymbol{\ell}(x) := (D\mathbf{u})_r^\lambda\, x + (\mathbf{u})_r^\lambda, \quad x \in \mathbb{R}^n.$$

Then we have $D\boldsymbol{\ell} = (D\mathbf{u})_r^\lambda$ and

$$\fint_{Q_{2r}^\lambda} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)\,\mathrm{d}z = \fint_{Q_{2r}^\lambda} \varphi_{|(D\mathbf{u})_r^\lambda|}(|D\mathbf{u} - (D\mathbf{u})_r^\lambda|)\,\mathrm{d}z\,.$$

We divide the proof into three steps.

Step 1. (Scaling) We first observe from (2.14), (5.5) and (5.6) that

$$\fint_{Q_r^\lambda} \varphi(|D\mathbf{u}|)\,\mathrm{d}z \leqslant 2^{n+2} \fint_{Q_{2r}^\lambda} \varphi(|D\mathbf{u}|)\,\mathrm{d}z \leqslant c \fint_{Q_{2r}^\lambda} \varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)\,\mathrm{d}z + c\varphi(|D\mathbf{u}_{\boldsymbol{\ell}}|) \leqslant c\varphi(\lambda)\,.$$

Now, we consider the following scaled functions:

$$\tilde{\mathbf{u}}(x,t) := \frac{1}{\lambda}\mathbf{u}(x, t\lambda^2/\varphi(\lambda)), \quad \tilde{\mathbf{A}}(\mathbf{P}) := \frac{\lambda\mathbf{A}(\lambda\mathbf{P})}{\varphi(\lambda)}, \quad \tilde{\varphi}(\tau) := \frac{\varphi(\lambda\tau)}{\lambda\varphi'(\lambda)}, \quad \tilde{\boldsymbol{\ell}} := \frac{1}{\lambda}\boldsymbol{\ell}\,.$$

Then, by a direct computation, we have

$$\tilde{\varphi}_{\frac{a}{\lambda}}(t) = \frac{\varphi_a(\lambda t)}{\lambda\varphi'(\lambda)} \quad \text{and} \quad \tilde{\varphi}'_{\frac{a}{\lambda}}(t) = \frac{\varphi'_a(\lambda t)}{\varphi'(\lambda)}\,,$$

whence

$$\tilde{\varphi}'_{|D\tilde{\boldsymbol{\ell}}|}(\tau) = \frac{\varphi'_{|D\boldsymbol{\ell}|}(\lambda\tau)}{\varphi'(\lambda)} = \frac{\varphi'(|D\boldsymbol{\ell}|+\lambda\tau)}{\varphi'(\lambda)(|D\boldsymbol{\ell}|+\lambda\tau)}(\lambda\tau)\,.$$

In particular, taking into account (5.5),

$$\tilde{\varphi}'_{|D\tilde{\boldsymbol{\ell}}|}(1) \sim \tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(1) \sim 1\,.$$

Moreover, $\tilde{\mathbf{u}}$ is a weak solution to

$$\partial_t\tilde{\mathbf{u}} - \operatorname{div}\tilde{\mathbf{A}}(D\tilde{\mathbf{u}}) = \mathbf{0} \quad \text{in } \; Q_r\,,$$

where $\tilde{\mathbf{A}}$ satisfies the same properties of $\mathbf{A}$ listed in Assumption (A), with $\tilde{\varphi}$ in place of φ, and satisfies

$$\text{(5.8)}\quad \frac{1}{2K} \leqslant |D\tilde{\boldsymbol{\ell}}| = |(D\tilde{\mathbf{u}})_r| \leqslant 2K, \quad \frac{1}{c} \leqslant \fint_{Q_r}\tilde{\varphi}(|D\tilde{\mathbf{u}}|)\,\mathrm{d}z = \frac{1}{\lambda\varphi'(\lambda)}\fint_{Q_r^\lambda}\varphi(|D\mathbf{u}|)\,\mathrm{d}z \leqslant c\,,$$

and by (5.6)

$$\text{(5.9)}\qquad \mu := \left(\frac{1}{\varphi(|D\boldsymbol{\ell}|)}\fint_{Q_r^\lambda}\varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)\,\mathrm{d}z\right)^{\frac{1}{2}} \leqslant \sqrt{\delta_0} \leqslant 1\,,$$

where $\delta_0 \leqslant 1$ will be determined later. Note that the last inequality also yields

$$\text{(5.10)}\qquad \fint_{Q_r}\tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(|D\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}}|)\,\mathrm{d}z = \frac{1}{\lambda\varphi'(\lambda)}\fint_{Q_r^\lambda}\varphi_{|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)\,\mathrm{d}z \sim \mu^2 \leqslant \delta_0\,.$$

Moreover, by Corollary 4.8 and the estimate (5.1) we also have that

$$\text{(5.11)}\qquad \left(\fint_{Q_{r/2}}\tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(|D\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}}|)^{1+\sigma_0}\,\mathrm{d}z\right)^{\frac{1}{1+\sigma_0}} \leqslant c\mu^2 \leqslant c\delta_0$$

for some $\sigma_0 > 0$ and this implies that

$$\begin{aligned}\frac{1}{|Q_r|}\left|\int_{Q_r}\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}}\cdot\boldsymbol{\zeta}_t - D\tilde{\mathbf{A}}(D\tilde{\boldsymbol{\ell}})\langle D\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}}, D\boldsymbol{\zeta}\rangle\,\mathrm{d}z - \int_{B_r}\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}}\cdot\boldsymbol{\zeta}\,\mathrm{d}x\right|&\\ \leqslant c\left(\left(\sqrt{\delta_0}\right)^{\gamma} + \sqrt{\delta_0}\right)\mu\sup_{Q_r}|D\boldsymbol{\zeta}|\,,&\end{aligned}$$

for all $\boldsymbol{\zeta} \in C^\infty(Q_r)$ with $\boldsymbol{\zeta} = \mathbf{0}$ on $\partial B_r \times (-r^2, r^2)$.

Step 2. ($\mathcal{A}$-caloric approximation) Observe that

$$\partial_t\tilde{\mathbf{u}}_{\boldsymbol{\ell}} = \partial_t\tilde{\mathbf{u}} = \operatorname{div}\tilde{\mathbf{A}}(D\tilde{\mathbf{u}}) =: \operatorname{div}\mathbf{H} \quad \text{in } \; Q_r\,,$$

i.e., $\mathbf{H} := \tilde{\mathbf{A}}(D\tilde{\mathbf{u}})$, in the distributional sense, and writing $\tilde{\varphi}^*_{|D\tilde{\boldsymbol{\ell}}|} := (\tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|})^*$

$$\text{(5.12)}\qquad \fint_{Q_r}\tilde{\varphi}^*_{|D\tilde{\boldsymbol{\ell}}|}(|\mathbf{H}|)\,\mathrm{d}z = \fint_{Q_r}\tilde{\varphi}^*_{|D\tilde{\boldsymbol{\ell}}|}(|\tilde{\mathbf{A}}(D\tilde{\mathbf{u}})|)\,\mathrm{d}z \leqslant c\fint_{Q_r}\tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(|D\tilde{\mathbf{u}}|)\,\mathrm{d}z \leqslant c\mu^2\,.$$

Set

$$p_1 := \min\left\{p, \frac{q}{q-1}\right\} \quad \text{and} \quad p_0 := \frac{1+p_1}{2}.$$

Then we see from the Jensen inequality in Lemma 2.1 with $\psi(t) := \tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(t^{\frac{1}{p_1}})$, (5.10) and (5.12) that

$$\left(\fint_{Q_r}|D\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}}|^{p_1}\,\mathrm{d}z\right)^{\frac{1}{p_1}} \leqslant c\tilde{\varphi}^{-1}_{|D\tilde{\boldsymbol{\ell}}|}\left(\fint_{Q_r}\tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(|D\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}}|)\,\mathrm{d}z\right) \leqslant c\tilde{\varphi}^{-1}_{|D\tilde{\boldsymbol{\ell}}|}(\mu^2)$$

and, arguing as before with $\psi(t) := \tilde{\varphi}^*_{|D\tilde{\boldsymbol{\ell}}|}(t^{1/p_1})$,

$$\left(\fint_{Q_r} |\mathbf{H}|^{p_1}\,\mathrm{d}z\right)^{\frac{1}{p_1}} \leqslant c(\tilde{\varphi}^*_{|D\tilde{\boldsymbol{\ell}}|})^{-1}\left(\fint_{Q_r} \tilde{\varphi}^*_{|D\tilde{\boldsymbol{\ell}}|}(\mathbf{H})\,\mathrm{d}z\right) \leqslant c(\tilde{\varphi}^*_{|D\tilde{\boldsymbol{\ell}}|})^{-1}(\mu^2)\,.$$

Furthermore, we notice from the first inequality in (5.8) and the fact that $\mu \in (0,1)$ that

$$\frac{\tilde{\varphi}'(|D\tilde{\boldsymbol{\ell}}|+\mu)}{|D\tilde{\boldsymbol{\ell}}|+\mu} \sim 1$$

and for $\tau_1 \geqslant 0$ satisfying that $\tilde{\varphi}'_{|D\tilde{\boldsymbol{\ell}}|}(\tau_1) = \mu$,

$$\tau_1 \lesssim 1 \quad \text{hence} \quad \tau_1 \sim \frac{\tilde{\varphi}'(|D\tilde{\boldsymbol{\ell}}|+\tau_1)}{|D\tilde{\boldsymbol{\ell}}|+\tau_1}\tau_1 = \mu\,,$$

which imply

$$\mu^2 \sim \tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(\mu) \quad \text{and} \quad \mu^2 \sim \tau_1\mu = (\tilde{\varphi}'_{|D\tilde{\boldsymbol{\ell}}|})^{-1}(\mu)\mu = (\tilde{\varphi}^*_{|D\tilde{\boldsymbol{\ell}}|})'(\mu)\mu \sim \tilde{\varphi}^*_{|D\tilde{\boldsymbol{\ell}}|}(\mu)\,. \tag{5.13}$$

Collecting the previous estimates, we then have

$$\left(\fint_{Q_r} |D\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}}|^{p_1}\,\mathrm{d}z + \fint_{Q_r} |\mathbf{H}|^{p_1}\,\mathrm{d}z\right)^{\frac{1}{p_1}} \leqslant c\mu\,. \tag{5.14}$$

Therefore, by Theorem 3.8 with, in particular, $\mathcal{A} := \tilde{\mathbf{A}}(D\tilde{\boldsymbol{\ell}})$, $\psi(\tau) := \tau^{p_0}$ and $\sigma := \frac{p_1}{p_0}-1$ (i.e., $p_0(1+\sigma) = p_1$), for $\varepsilon \in (0,1)$ to be determined small later, there exists small $\delta_0 > 0$ depending on $n, N, L, \nu, p, q, \gamma$ and ε such that

$$\fint_{Q_r} |D\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}} - D\mathbf{h}|^{p_0}\,\mathrm{d}z \leqslant \varepsilon\mu^{p_0}\,, \tag{5.15}$$

where $\mathbf{h}$ is the weak solution to

$$\begin{cases} \partial_t \mathbf{h} - \operatorname{div}(\mathcal{A}D\mathbf{h}) = 0 & \text{in} \ \ Q_r, \\ \mathbf{h} = \tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}} \quad \text{on} \ \ \partial_{\mathrm{p}} Q_r\,. \end{cases}$$

We note from (3.1), (5.15), (5.14) and (5.13) that

$$\left(\fint_{Q_{r/2}} \tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(|D\mathbf{h}|)^{1+\sigma_0}\,\mathrm{d}z\right)^{\frac{1}{1+\sigma_0}} \leqslant c\tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}\left(\fint_{Q_r} |D\mathbf{h}|\,\mathrm{d}z\right) \leqslant c\tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(\mu) \leqslant c\mu^2\,. \tag{5.16}$$

Therefore, by Hölder's inequality, the Jensen inequality in Lemma 2.1 with $\psi^{-1}(t) := \tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(t)^{\frac{1}{q}}$ and the estimates (5.11), (5.15), (5.16) and (5.13), we have that with $\kappa_0 \in (0,1)$ satisfying $\frac{\kappa_0}{q} + (1-\kappa_0)(1+\sigma_0) = 1$,

$$\begin{aligned}
&\fint_{Q_{r/2}} \tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(|D\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}} - D\mathbf{h}|)\,\mathrm{d}z \\
&\leqslant \left(\fint_{Q_{r/2}} \tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(|D\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}} - D\mathbf{h}|)^{\frac{1}{q}}\,\mathrm{d}z\right)^{\kappa_0}\left(\fint_{Q_{r/2}} \tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(|D\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}} - D\mathbf{h}|)^{1+\sigma_0}\,\mathrm{d}z\right)^{1-\kappa_0} \\
&\leqslant c\tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}\left(\left[\fint_{Q_{r/2}} |D\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}} - D\mathbf{h}|^{p_0}\,\mathrm{d}z\right]^{\frac{1}{p_0}}\right)^{\frac{\kappa_0}{q}}\mu^{2(1-\kappa_0)(1+\sigma)} \\
&\leqslant c\varepsilon^{\frac{p\kappa_0}{p_0 q}}\tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(\mu)^{\frac{\kappa_0}{q}}\mu^{2(1-\kappa_0)(1+\sigma)} \leqslant c\varepsilon^{\frac{p\kappa_0}{p_0 q}}(\mu^2)^{\frac{\kappa_0}{q}+(1-\kappa_0)(1+\sigma)} = c\varepsilon^{\frac{p\kappa_0}{p_0 q}}\mu^2\,.
\end{aligned} \tag{5.17}$$

Moreover, by (3.1) in Lemma 3.1 with $\rho = r/2$, (5.16) and (5.9), we also have

$$\sup_{Q_{r/4}} |D\mathbf{h}| \leqslant c(\tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|})^{-1}\left(\fint_{Q_{r/2}} \tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(|D\mathbf{h}|)\,\mathrm{d}z\right) \leqslant c\mu \leqslant c\sqrt{\delta_0}.$$

Note that we choose δ_0 small so that

$$\sup_{Q_{r/4}} |D\mathbf{h}| \leqslant c\mu \leqslant \frac{1}{4K}\,. \tag{5.18}$$

Step 3. (Decay estimate) Let $\theta \in (0, 1/8)$ to be determined later and recall function $\mathbf{V}$ corresponding to the N-function $\tilde{\varphi}$ defined as in (2.6). We first observe from (5.8) and (5.18) that

$$\frac{1}{8}|(D\tilde{\mathbf{u}})_r| \leqslant \frac{K}{4} \leqslant |(D\tilde{\mathbf{u}})_r| - |(D\mathbf{h})_{\theta r}| \leqslant |(D\tilde{\mathbf{u}})_r + (D\mathbf{h})_{\theta r}| \leqslant |(D\tilde{\mathbf{u}})_r| + \frac{1}{4K} \leqslant \frac{9}{4}|(D\tilde{\mathbf{u}})_r|\,.$$

Then, using (2.7), (2.10) and the preceding estimate,

$$\begin{aligned}
\fint_{Q_{\theta r}} \tilde{\varphi}_{|(D\tilde{\mathbf{u}})_{\theta r}|}(|D\tilde{\mathbf{u}} - (D\tilde{\mathbf{u}})_{\theta r}|)|)\,\mathrm{d}z &\sim \fint_{Q_{\theta r}} |\mathbf{V}(D\tilde{\mathbf{u}}) - \mathbf{V}((D\tilde{\mathbf{u}})_{\theta r})|^2\,\mathrm{d}z \\
&\sim \fint_{Q_{\theta r}} |\mathbf{V}(D\tilde{\mathbf{u}}) - (\mathbf{V}(D\tilde{\mathbf{u}}))_{\theta r}|^2\,\mathrm{d}z \\
&\leqslant \fint_{Q_{\theta r}} |\mathbf{V}(D\tilde{\mathbf{u}}) - \mathbf{V}((D\tilde{\mathbf{u}})_r + (D\mathbf{h})_{\theta r})|^2\,\mathrm{d}z \\
&\sim \fint_{Q_{\theta r}} \tilde{\varphi}_{|(D\tilde{\mathbf{u}})_r + (D\mathbf{h})_{\theta r}|}(|D\tilde{\mathbf{u}} - (D\tilde{\mathbf{u}})_r - (D\mathbf{h})_{\theta r}|)\,\mathrm{d}z \\
&\sim \fint_{Q_{\theta r}} \tilde{\varphi}_{|(D\tilde{\mathbf{u}})_r|}(|D\tilde{\mathbf{u}} - (D\tilde{\mathbf{u}})_r - (D\mathbf{h})_{\theta r}|)\,\mathrm{d}z\,.
\end{aligned}$$

Moreover, by (5.8) and (5.18) we have

$$\tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(\theta|(D\mathbf{h})_{r/4}|) \sim \frac{\varphi'(|(D\tilde{\mathbf{u}})_r| + \theta|(D\mathbf{h})_{r/4}|)}{|(D\tilde{\mathbf{u}})_r| + \theta|(D\mathbf{h})_{r/4}|}\theta^2|(D\mathbf{h})_{r/4}|^2 \sim \theta^2|(D\mathbf{h})_{r/4}|^2 \lesssim \theta^2\mu^2\,,$$

Using the above two estimates, (3.2) in Lemma 3.1 with $\rho = r/2$ and (5.17), we obtain

$$\begin{aligned}
&\fint_{Q_{\theta r}} \tilde{\varphi}_{|(D\tilde{\mathbf{u}})_{\theta r}|}(|D\tilde{\mathbf{u}} - (D\tilde{\mathbf{u}})_{\theta r}|)\,\mathrm{d}z \\
&\leqslant c\fint_{Q_{\theta r}} \tilde{\varphi}_{|(D\tilde{\mathbf{u}})_r|}(|D\tilde{\mathbf{u}} - (D\tilde{\mathbf{u}})_r - (D\mathbf{h})_{\theta r}|)\,\mathrm{d}z \\
&\leqslant c\fint_{Q_{\theta r}} \tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(|D\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}} - D\mathbf{h}|)\,\mathrm{d}z + c\fint_{Q_{\theta r}} \tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(|D\mathbf{h} - (D\mathbf{h})_{\theta r})|)\,\mathrm{d}z \\
&\leqslant c\theta^{-(n+2)}\fint_{Q_{r/2}} \tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}(|D\tilde{\mathbf{u}}_{\tilde{\boldsymbol{\ell}}} - D\mathbf{h}|)\,\mathrm{d}z + c\tilde{\varphi}_{|D\tilde{\boldsymbol{\ell}}|}\left(\theta\fint_{Q_{r/4}} |D\mathbf{h} - (D\mathbf{h})_{r/4}|\,\mathrm{d}z\right) \\
&\leqslant c\theta^{-(n+2)}\varepsilon^{\frac{p\kappa_0}{p_0 q}}\mu^2 + c\theta^2\mu^2.
\end{aligned}$$

Finally, choosing θ small so that $c\theta^{1-\beta} \leqslant \frac{1}{2}$ and then ε small so that $c\theta^{-(n+2)}\varepsilon^{\frac{p\kappa_0}{p_0 q}} \leqslant \frac{1}{2}\theta^{\beta+1}$, we obtain

$$\fint_{Q_{\theta r}} \tilde{\varphi}_{|(D\tilde{\mathbf{u}})_{\theta r}|}(|D\tilde{\mathbf{u}} - (D\tilde{\mathbf{u}})_{\theta r}|)|)\,\mathrm{d}z \leqslant \theta^{1+\beta}\mu^2\,,$$

whence, by scaling back,

$$\fint_{Q_{\theta r}} \varphi_{|(D\mathbf{u})^\lambda_{\theta r}|}(|D\mathbf{u} - (D\mathbf{u})^\lambda_{\theta r}|)\,\mathrm{d}z \leqslant \theta^{1+\beta}\lambda\varphi'(\lambda)\mu^2\,.$$

This estimate, together with (5.10), yields (5.7) by choosing θ sufficiently small depending on $n, N, p, q, L, \nu, K, \gamma$ and β. This concludes the proof. □

From the previous lemma, we obtain decay estimates for $D\mathbf{u}$ in the nondegenerate regime.

Lemma 5.3. *Let $\lambda > 0$, $\beta \in (0,1)$, $Q^\lambda_{2R} = Q^\lambda_{2R}(z_0) \Subset \Omega_T$ and $\mathbf{u}$ be a weak solution to (1.1). Suppose*

$$\frac{\lambda}{K_0} \leqslant |(D\mathbf{u})^\lambda_{2R}| \leqslant K_0\lambda$$

for some $K_0 > 0$. There exists small $\delta_1 \in (0,1)$ depending on $n, N, p, q, L, \nu, \beta, \gamma$ and K_0 such that if

$$\fint_{Q^\lambda_R} \varphi_{|(D\mathbf{u})^\lambda_R|}(|D\mathbf{u} - (D\mathbf{u})^\lambda_R|)\,\mathrm{d}z \leqslant \delta_1\varphi(|(D\mathbf{u})^\lambda_R|) \tag{5.19}$$

then the limit

$$\Gamma_{z_0} := \lim_{r\to 0^+} (D\mathbf{u})_{Q_r(z_0)} \tag{5.20}$$

exists with

$$\frac{\lambda}{2K_0} \leqslant |\Gamma_{z_0}| \leqslant 2K_0\lambda\,, \tag{5.21}$$

and for every $r \in (0,R)$,

$$\fint_{Q^\lambda_r(z_0)} \varphi(|D\mathbf{u} - \Gamma_{z_0}|)\,\mathrm{d}z \leqslant c\left(\frac{r}{R}\right)^{\beta_1}\varphi(\lambda) \tag{5.22}$$

for some $c = c(n, N, p, q, L, \nu, K_0, \beta, \gamma) > 0$ and $\beta_1 = \beta_1(p, q, \beta) > 0$.

Proof. For simplicity, we shall omit to write the center z_0. We recall the parameters θ and δ_0 from Lemma 5.2. However, we notice that for any smaller θ and δ_0 satisfying additional conditions, (5.6) and (5.7) still hold. Then we choose $\delta_1 \leqslant \delta_0$. We divide the proof into two steps.

Step 1. We shall prove by induction that for every $i \in \mathbb{N}$,

$$\fint_{Q^\lambda_{\theta^i R}} \varphi_{|(D\mathbf{u})^\lambda_{\theta^i R}|}(|D\mathbf{u} - (D\mathbf{u})^\lambda_{\theta^i R}|)\,\mathrm{d}z \leqslant \theta^{2\beta i}\fint_{Q^\lambda_R} \varphi_{|(D\mathbf{u})^\lambda_R|}(|D\mathbf{u} - (D\mathbf{u})^\lambda_R|)\,\mathrm{d}z\,, \tag{5.23}$$

and

$$\tfrac{1}{2}|(D\mathbf{u})^\lambda_R| \leqslant \left[1 - \tfrac{1}{4}\sum_{k=0}^{i-1}2^{-k}\right]|(D\mathbf{u})^\lambda_R| \leqslant |(D\mathbf{u})^\lambda_{\theta^i R}| \leqslant \left[1 + \tfrac{1}{4}\sum_{k=0}^{i-1}2^{-k}\right]|(D\mathbf{u})^\lambda_R| \leqslant \tfrac{3}{2}|(D\mathbf{u})^\lambda_R|\,. \tag{5.24}$$

Suppose $i = 1$. Then (5.7) yields (5.23). In order to prove (5.24), we first observe from (2.14) and (5.19) that

$$\fint_{Q^\lambda_R} \varphi(|D\mathbf{u}|)\,\mathrm{d}z \leqslant c\fint_{Q^\lambda_R} \varphi_{|(D\mathbf{u})^\lambda_R|}(|D\mathbf{u} - (D\mathbf{u})^\lambda_R|)\,\mathrm{d}z + c\varphi(|(D\mathbf{u})^\lambda_R|) \leqslant c\varphi(|(D\mathbf{u})^\lambda_R|).$$

and, applying (2.11) with $\varepsilon=\delta_0^{\frac12}$,

$$\fint_{Q_R^\lambda}\varphi(|D\mathbf{u}-(D\mathbf{u})_R^\lambda|)\,\mathrm{d}z$$
$$\leqslant \fint_{Q_R^\lambda}\delta_0^{\frac12}[\varphi(|D\mathbf{u}|)+\varphi(|(D\mathbf{u})_R^\lambda|)]+c\delta_0^{-\frac12}\varphi_{|(D\mathbf{u})_R^\lambda|}(|D\mathbf{u}-(D\mathbf{u})_R^\lambda|)\,\mathrm{d}z\leqslant c\delta_0^{\frac12}\varphi(|(D\mathbf{u})_R^\lambda|)\,.$$

Hence we have

$$\begin{aligned}|(D\mathbf{u})_{\theta R}^\lambda-(D\mathbf{u})_R^\lambda| &\leqslant \theta^{-\frac{n+2}{p}}\varphi^{-1}\left(\fint_{Q_R^\lambda}\varphi(|D\mathbf{u}-(D\mathbf{u})_R^\lambda|)\,\mathrm{d}z\right)\\ &\leqslant c\theta^{-\frac{n+2}{p}}\varphi^{-1}(\delta_0^{\frac12}\varphi(|(D\mathbf{u})_R^\lambda|))\leqslant c\theta^{-\frac{n+2}{p}}\delta_0^{\frac{1}{2q}}|(D\mathbf{u})_R^\lambda|\leqslant\frac14|(D\mathbf{u})_R^\lambda|\,,\end{aligned}$$

after choosing sufficiently small $\delta_0=\delta_0(n,N,p,q,L,\nu,K_0,\beta)$, which implies (5.24).

Suppose that (5.23) and (5.24) hold for $i=1,2,\dots,j-1$ for some $j\geqslant 2$. Then, using (5.23), (5.24) with $i=j-1$ and (5.19), and choosing θ such that $\theta^{2\beta}\leqslant 2^{-q}$, we have

$$\begin{aligned}\fint_{Q_{\theta^{j-1}R}^\lambda}\varphi_{|(D\mathbf{u})_{\theta^{j-1}R}^\lambda|}(|D\mathbf{u}-(D\mathbf{u})_{\theta^{j-1}R}^\lambda|)\,\mathrm{d}z &\leqslant \theta^{2\beta(j-1)}\delta_1\varphi(2|(D\mathbf{u})_{\theta^{j-1}R}^\lambda|)\\ &\leqslant \delta_0\varphi(|(D\mathbf{u})_{\theta^{j-1}R}^\lambda|)\,.\end{aligned}\tag{5.25}$$

Since $\frac{K_0}{2}\leqslant|(D\mathbf{u})_{\theta^{j-1}R}^\lambda|\leqslant 2K_0$, applying Lemma 5.2 with $r=\theta^{j-1}R$ we obtain

$$\begin{aligned}\fint_{Q_{\theta^jR}^\lambda}\varphi_{|(D\mathbf{u})_{\theta^jR}^\lambda|}(|D\mathbf{u}-(D\mathbf{u})_{\theta^jR}^\lambda|)\,\mathrm{d}z &\leqslant \theta^{2\beta}\fint_{Q_{\theta^{j-1}R}^\lambda}\varphi_{|(D\mathbf{u})_{\theta^{j-1}R}^\lambda|}(|D\mathbf{u}-(D\mathbf{u})_{\theta^{j-1}R}^\lambda|)\,\mathrm{d}z\\ &\leqslant \theta^{2\beta j}\fint_{Q_R^\lambda}\varphi_{|(D\mathbf{u})_R^\lambda|}(|D\mathbf{u}-(D\mathbf{u})_R^\lambda|)\,\mathrm{d}z\,,\end{aligned}$$

which proves (5.23) with $i=j$. We next prove (5.24) with $i=j$. As in the case $i=1$, we have from (5.25) that

$$\fint_{Q_{\theta^{j-1}R}^\lambda}\varphi(|D\mathbf{u}-(D\mathbf{u})_{\theta^{j-1}R}^\lambda|)\,\mathrm{d}z\leqslant c\theta^{\beta(j-1)}\delta_1^{\frac12}\varphi(|(D\mathbf{u})_{\theta^{j-1}R}^\lambda|)\,,$$

and hence

$$\begin{aligned}|(D\mathbf{u})_{\theta^jR}^\lambda-(D\mathbf{u})_{\theta^{j-1}R}^\lambda| &\leqslant \theta^{-\frac{n+2}{p}}\varphi^{-1}\left(\fint_{Q_{\theta^{j-1}R}^\lambda}\varphi(|D\mathbf{u}-(D\mathbf{u})_{\theta^{j-1}R}^\lambda|)\,\mathrm{d}z\right)\\ &\leqslant c\theta^{-\frac{n+2}{p}}\varphi^{-1}(\delta_1^{\frac12}\theta^{\beta(j-1)}\varphi(|(D\mathbf{u})_{\theta^{j-1}R}^\lambda|))\\ &\leqslant c\theta^{-\frac{n+2}{p}}\delta_1^{\frac{1}{2q}}\theta^{\frac{\beta(j-1)}{q}}|(D\mathbf{u})_{\theta^{j-1}R}^\lambda|\,.\end{aligned}$$

Therefore, choosing θ such that $\theta^{\frac{\beta}{q}}\leqslant 2^{-1}$ and δ_1 such that $c\theta^{-\frac{n+2}{p}}\delta_1^{\frac{1}{2q}}\leqslant 1/4$, we obtain

$$|(D\mathbf{u})_{\theta^jR}^\lambda-(D\mathbf{u})_{\theta^{j-1}R}^\lambda|\leqslant\theta^{\frac{\beta(j-1)}{q}}|(D\mathbf{u})_{\theta^{j-1}R}^\lambda|\leqslant\frac14 2^{-(j-1)}|(D\mathbf{u})_{\theta^{j-1}R}^\lambda|\,,\tag{5.26}$$

which, together with (5.24) with $i=j-1$, implies (5.24) with $i=j$.

Step 2. We first note that $\{(D\mathbf{u})_{\theta^iR}^\lambda\}_{i\in\mathbb{N}}$ is a Cauchy sequence. Indeed, by (5.26) and (5.24), we have that any $i,j\in\mathbb{N}$ with $i<j$

$$|(D\mathbf{u})_{\theta^jR}^\lambda-(D\mathbf{u})_{\theta^iR}^\lambda|\leqslant\sum_{k=i}^{j-1}\theta^{\frac{\beta k}{q}}|(D\mathbf{u})_{\theta^kR}^\lambda|\leqslant\frac14\sum_{k=i}^{j-1}2^{-k}|(D\mathbf{u})_{\theta^kR}^\lambda|\leqslant\frac34 2^{-i}|(D\mathbf{u})_R^\lambda|\,.\tag{5.27}$$

Therefore, set

$$\Gamma_0 := \lim_{i\to\infty} (D\mathbf{u})^{\lambda}_{\theta^i R}\,.$$

Then (5.24) implies (5.21). We shall prove (5.22). Note that by (5.23), (5.19), (5.27) and (5.24),

$$\begin{aligned}\fint_{Q^{\lambda}_{\theta^i R}} \varphi(|D\mathbf{u}-\Gamma_0|)\,\mathrm{d}z &\leqslant c\fint_{Q^{\lambda}_{\theta^i R}} \varphi(|D\mathbf{u}-(D\mathbf{u})^{\lambda}_{\theta^i R}|)\,\mathrm{d}z + c\varphi(|(D\mathbf{u})^{\lambda}_{\theta^i R}-\Gamma_0|)\\ &\leqslant c\theta^{\beta i}\varphi(\lambda) + c\varphi\left(\sum_{k=i}^{\infty}\theta^{\frac{\beta k}{q}}\lambda\right) \leqslant c\theta^{\frac{p\beta i}{q}}\varphi(\lambda)\end{aligned}$$

Therefore, for every $r<\theta R$ with $i\in\mathbb{N}$ satisfying $\theta^{i+1}R\leqslant r\leqslant \theta^i R$, we have

$$\fint_{Q^{\lambda}_{r}} \varphi(|D\mathbf{u}-\Gamma_0|)\,\mathrm{d}z \leqslant \frac{1}{\theta^{n+2}}\fint_{Q^{\lambda}_{\theta^i R}} \varphi(|D\mathbf{u}-\Gamma_0|)\,\mathrm{d}z \leqslant c\theta^{\frac{p\beta i}{q}}\varphi(\lambda) \leqslant c\left(\frac{r}{R}\right)^{\frac{p\beta}{q}}\varphi(\lambda),$$

which implies (5.22).

It remains to prove (5.20). For $0<r\leqslant \min\{1,(\varphi(\lambda)/\lambda^2)^{-\frac12}\}\theta R$, set

$$\rho := \max\{1,(\varphi(\lambda)/\lambda^2)^{\frac12}\}\,r \leqslant \theta R\,.$$

Then we have $Q_r\subset Q^{\lambda}_{\rho}$ and by (5.22)

$$\begin{aligned}|(D\mathbf{u})_{Q_r}-\Gamma_0| &\leqslant \frac{\rho^{n+2}\lambda^2}{\varphi(\lambda)r^{n+2}}\fint_{Q^{\lambda}_{\rho}} |D\mathbf{u}-\Gamma_0|\,\mathrm{d}z\\ &\leqslant c\frac{\max\{1,(\varphi(\lambda)/\lambda^2)^{-\frac{n+2}{2}}\}}{(\varphi(\lambda)/\lambda^2)}\varphi^{-1}\left(\left(\frac{\rho}{R}\right)^{\frac{p\beta}{q}}\varphi(\lambda)\right)\\ &\leqslant c\frac{\max\{1,(\varphi(\lambda)/\lambda^2)^{-\frac{n+2}{2}}\}}{\varphi(\lambda)/\lambda^2}\left(\frac{r}{R}\max\{1,(\varphi(\lambda)/\lambda^2)^{-\frac12}\}\right)^{\frac{p\beta}{q^2}}\lambda \ \longrightarrow\ 0\end{aligned}$$

as $r\to 0$, which implies (5.20). Therefore, the proof is completed. □

6. Degenerate regime

We consider the *degenerate* regime, which means that the average of the gradient of solution is relatively smaller than the relevant excess function, see for instance (6.12). In this regime, we apply the φ-caloric approximation. We start by investigating regularity results for φ-caloric maps. We refer to [41] for regularity results for φ-caloric maps.

Let $\mathbf{h}$ be a weak solution to

$$\partial_t\mathbf{h} - \operatorname{div}\left(\frac{\varphi'(|D\mathbf{h}|)}{|D\mathbf{h}|}D\mathbf{h}\right) = \mathbf{0} \quad\text{in}\ \ Q^{\lambda}_R. \tag{6.1}$$

Then by [41, Corollary 5.3] with the scaling argument used in the proof of Corollary 4.8, we have

$$\sup_{Q^{\lambda}_{R/2}} \varphi(|D\mathbf{h}|) \leqslant c\varphi(\lambda)\,. \tag{6.2}$$

for some $c>0$ depending on n,N,p,q and $\tilde{c}$ if $\mathbf{h}$ satisfies

$$\fint_{Q^{\lambda}_R} \varphi(|D\mathbf{h}|)\,\mathrm{d}z \leqslant \tilde{c}\varphi(\lambda)\,.$$

Moreover, from [41, Section 6], we have the following result concerned with $C^{1,\alpha}$-regularity for $\mathbf{h}$.

Lemma 6.1. *Let* $\mathbf{h}$ *be a weak solution to* (6.1) *in* $Q_R^\lambda = Q_R^\lambda(z_0)$ *with*

$$\sup_{Q_R^\lambda} |D\mathbf{h}| \leqslant \lambda \tag{6.3}$$

for some $\lambda > 0$. *Then there exist* $\alpha_1 \in (0,1)$ *depending on* n, N, p, q, *a switching radius* $r_s \in [0, R]$ *and* $\lambda_r > 0$ *for each* $r \in (0, R]$, *such that*

$$\lambda_r = \lambda_{r_s} \quad \text{if} \ \ r \in (0, r_s] \quad \text{and} \quad \left(\frac{r}{R}\right)^{\alpha_1} \lambda \leqslant \lambda_r \leqslant 2\left(\frac{r}{R}\right)^{\alpha_1} \lambda \quad \text{if} \ \ r \in (r_s, R]\,, \tag{6.4}$$

$$\sup_{Q_r^{\lambda_r}} |D\mathbf{h}| \leqslant \lambda_r \quad \text{for all} \ \ r \in (0, R]\,, \tag{6.5}$$

and

$$\fint_{Q_r^{\lambda_r}} \varphi_{|(D\mathbf{h})_r^{\lambda_r}|}(|D\mathbf{h} - (D\mathbf{h})_r^{\lambda_r}|)\,\mathrm{d}z \leqslant c\left(\frac{r}{r_s}\right)^{3/4} \varphi(\lambda_r) \quad \text{if} \ \ r \in (0, r_s]. \tag{6.6}$$

Moreover, we also have

$$|(D\mathbf{h})_r^{\lambda_r}| \geqslant C_s^{-1}\lambda_r \quad \text{and} \quad \operatorname*{osc}_{Q_r^{\lambda_r}} D\mathbf{h} \leqslant c\left(\frac{r}{r_s}\right)^{3/4} \lambda_r \quad \text{if} \ \ r \in (0, r_s]\,, \tag{6.7}$$

for some $C_s > 1$ *and* $c > 0$ *depending on* n, N, p, q.

Proof. For each $i = 0, 1, 2, \dots$, we inductively define

$$\lambda_0 := \lambda, \quad \lambda_{i+1} := \nu\lambda_i \quad \text{and} \quad r_0 := R, \quad r_{i+1} = \tilde{\sigma} r_i,$$

where $\sigma \in (0,1)$ is from [41, Proposition 6.2], $\nu \in (0,1)$ is from [41, Proposition 6.3] corresponding to the preceding σ, and

$$\tilde{\sigma} := \min\left\{ \frac{\sigma p^{\frac12} \nu^{\frac{q-2}{2}}}{2q^{\frac12}}, \nu^{\frac{4q}{3}} \right\} < \frac{\sigma}{2}\,. \tag{6.8}$$

Note that we may assume that $\nu > 1/2$. Then we have

$$\frac{r_{i+1}^2 \lambda_{i+1}^2}{\varphi(\lambda_{i+1})} \leqslant \frac{q\tilde{\sigma}^2 \lambda_{i+1}}{\varphi'(\lambda_{i+1})} r_i^2 \leqslant \frac{q\tilde{\sigma}^2 \lambda_i}{\varphi'(\lambda_i)\nu^{q-2}} r_i^2 \leqslant \frac{q\tilde{\sigma}^2}{p\nu^{q-2}} \frac{\lambda_i^2 r_i^2}{\varphi(\lambda_i)} = \left(\frac{\sigma}{2}\right)^2 \frac{r_i^2 \lambda_i^2}{\varphi(\lambda_i)} < \frac{r_i^2 \lambda_i^2}{\varphi(\lambda_i)}\,,$$

hence $Q_{r_{i+1}}^{\lambda_{i+1}} \subset Q_{\frac{\sigma}{2} r_i}^{\lambda_i} \subset Q_{r_i}^{\lambda_i}$. Moreover, we have

$$\lambda_i = \nu^i \lambda_0 = \tilde{\sigma}^{i\frac{\ln\nu}{\ln\tilde{\sigma}}} \lambda_0 = \left(\frac{r_i}{R}\right)^{\alpha_1} \lambda_0\,, \quad \text{where} \ \ \alpha_1 := \frac{\ln\nu}{\ln\tilde{\sigma}} \leqslant \frac{3}{4q}\,,$$

hence for every $r \in [r_{i+1}, r_i]$,

$$\left(\frac{r}{R}\right)^{\alpha_1} \lambda_0 \leqslant \lambda_i = \left(\frac{r_{i+1}}{R}\right)^{\alpha_1} \tilde{\sigma}^{-\alpha_1} \lambda_0 = \left(\frac{r_{i+1}}{R}\right)^{\alpha_1} \nu^{-1} \lambda_0 \leqslant 2\left(\frac{r}{R}\right)^{\alpha_1} \lambda_0\,. \tag{6.9}$$

Now we consider the inequality

$$|\{D\mathbf{h} \leqslant (1-\sigma)\lambda_i\}| \cap Q_{r_i}^{\lambda_i}| > \sigma |Q_{r_i}^{\lambda_i}| \quad \text{for} \ \ i = 0, 1, 2, \dots\,. \tag{6.10}$$

Then we have the following three cases:

(i) (6.10) does not hold when $i = 0$.

(ii) There exists $n_0 \in \mathbb{N}$ such that (6.10) holds when $i = 0, 1, 2, \dots, n_0 - 1$, but not when $i = n_0$.

(iii) (6.10) holds for every i.

If the case (i) holds, then by [41, Proposition 6.2] we have for every $r \in (0, R]$

$$\operatorname*{osc}_{Q_r^\lambda} D\mathbf{h} \leqslant c\left(\frac{r}{R}\right)^{\frac{3}{4}} \operatorname*{osc}_{Q_R^\lambda} D\mathbf{h}\,,$$

whence, with (6.3),

$$\begin{aligned}
&\fint_{Q_r^{\lambda_0}} \varphi_{|(D\mathbf{h})_r^{\lambda_0}|}(|D\mathbf{h}-(D\mathbf{h})_r^{\lambda_0}|)\,\mathrm{d}z\\
&\leqslant c\fint_{Q_r^{\lambda_0}} \varphi'(|(D\mathbf{h})_r^{\lambda_0}|+|D\mathbf{h}-(D\mathbf{h})_r^{\lambda_0}|)|D\mathbf{h}-(D\mathbf{h})_r^{\lambda_0}|\,\mathrm{d}z\\
&\leqslant c\varphi'(\lambda_0)\left(\frac{r}{R}\right)^{\frac{3}{4}}\lambda_0 \leqslant c\varphi(\lambda_0)\left(\frac{r}{R}\right)^{\frac{3}{4}},
\end{aligned}$$

which implies the inequalities (6.4), (6.5), (6.6) and the second inequality in (6.7) with $\lambda_r = \lambda_0 = \lambda$ for all $r \in (0, R]$ and $r_s = R$.

If the case (ii) holds, then by [41, Proposition 6.2] with $R = r_i$, $i = 0, \dots, n_0 - 1$ and [41, Proposition 6.3] with $R = r_{n_0}$, we have that

$$\sup_{Q_{r_i}^{\lambda_i}} |D\mathbf{h}| \leqslant \lambda_i \quad \text{for all} \quad i = 0, 1, 2, \dots, n_0, \tag{6.11}$$

and for every $r \in (0, r_{n_0}]$

$$\operatorname*{osc}_{Q_r^{\lambda_{n_0}}} D\mathbf{h} \leqslant c\left(\frac{r}{r_{n_0}}\right)^{\frac{3}{4}} \operatorname*{osc}_{Q_{r_{n_0}}^{\lambda}} D\mathbf{h}\,,$$

hence

$$\fint_{Q_r^{\lambda_{n_0}}} \varphi_{|(D\mathbf{h})_r^{\lambda_{n_0}}|}(|D\mathbf{h}-(D\mathbf{h})_r^{\lambda_{n_0}}|)\,\mathrm{d}z \leqslant c\varphi(\lambda_{n_0})\left(\frac{r}{r_{n_0}}\right)^{3/4}.$$

Therefore, choosing

$$\lambda_r = \begin{cases} \lambda_i & \text{when } \ r \in (r_{i+1}, r_i] \ \text{ and } \ i = 0, 1, \dots, n_0 - 1\,, \\ \lambda_{n_0} & \text{when } \ r \in (0, r_{n_{i_0}}]\,, \end{cases} \qquad \text{and} \quad r_s = r_{n_0}\,.$$

we have (6.4) (see (6.9)) and (6.5), (6.6) and and the second inequality in (6.7).

If the case (iii) holds, we have (6.11) for all $i = 0, 1, 2, \dots,$ which implies the desired estimates with $r_s = 0$ and $\lambda_r = \lambda_i$ for every $r \in (r_{i+1}, r_i]$ and $i = 0, 1, 2, \dots$.

Finally, we are left to prove the first inequality in (6.7). Note that, since in case (iii) $r_s = 0$, there is nothing to prove. Then we consider the cases (i) and (ii) where $r_s > 0$. By [41, Lemma 6.8] with $R = \frac{r_s}{2}$, together with [41, Lemma 6.9] with $R = r_s$, we have

$$|(D\mathbf{h})_{Q^{\lambda_{r_s}}_{\theta^j r_s/2}}| \geqslant \frac{1}{2}\lambda,$$

for all $j \in \mathbb{N}_0$ and for some sufficiently small $\theta \in (0, , 1)$ depending on n, N, p, q. If $r \in (\frac{r_s}{2}, r_s]$ then $|(D\mathbf{h})_{Q_r^{\lambda_{r_s}}}| \geqslant \frac{1}{2^{n+2}}|(D\mathbf{u})_{Q^{\lambda_r}_{r_s/2}}| \geqslant \lambda/2^{n+3}$. If $r \in (\frac{\theta^{j+1} r_s}{2}, \frac{\theta^j r_s}{2}]$, $|(D\mathbf{h})_{Q_r^{\lambda_{r_s}}}| \geqslant \frac{1}{\theta^{n+2}}|(D\mathbf{h})_{Q^{\lambda_{r_s}}_{\frac{\theta^{j+1} r_s}{2}}}| \geqslant \frac{\theta^{n+2}}{2}\lambda$. Thus, we obtain the first inequality in (6.7) with $C_1 = \max\{2^{n+3}, 2\theta^{-(n+2)}\}$. □

Remark 6.2. We list some remarks about the previous lemma.

(1) The numbers r_s and λ_r may depend on $\mathbf{h}$ and the center z_0 of Q_R^λ.
(2) Recalling the constants σ and $\tilde{\sigma}$ in the first part of the proof, one can see that if $r \in (0, \tilde{\sigma}R]$ then $Q_r^{\lambda_r} \subset Q^\lambda_{\frac{\sigma}{2}R}$.

The following lemma will be crucially used in the iteration process in Section 7.

Lemma 6.3. *Let $M_1 \geqslant 1$, $\lambda \in (0,1]$, $\chi, \chi_1 \in (0,1]$ and $\alpha_1 \in (0, \frac{3}{4q})$ be given in Lemma 6.1. There exist large constants $K_1, C_1 \geqslant 1$ depending on $n, N, p, q, \nu, L, \gamma, M_1$ and α_1 such that the following holds: for every $\vartheta \in (0, C_1^{-1}]$ there exists large $K \geqslant 1$ depending on $n, N, p, q, \nu, L, \gamma, M_1, \alpha_1$ and ϑ and small $\varepsilon_1 \in (0,1)$ depending on $n, N, p, q, \nu, L, \gamma, M_1, \alpha_1, \vartheta$ and χ, such that the following holds: if*

$$|(D\mathbf{u})_r^\lambda| \leqslant M_1 \lambda,$$

$$\chi \varphi(|(D\mathbf{u})_r^\lambda|) \leqslant \fint_{Q_r^\lambda} \varphi_{|(D\mathbf{u})_r^\lambda|}(|D\mathbf{u} - (D\mathbf{u})_r^\lambda|)\, \mathrm{d}z \quad \textit{or} \quad |(D\mathbf{u})_r^\lambda| \leqslant \frac{\lambda}{K} \tag{6.12}$$

and

$$\fint_{Q_r^\lambda} \varphi_{|(D\mathbf{u})_r^\lambda|}(|D\mathbf{u} - (D\mathbf{u})_r^\lambda|)\, \mathrm{d}z \leqslant \min\{\varphi(\lambda), \varepsilon_1\},$$

then there exists

$$\lambda_1 \in [\vartheta^{\alpha_1}\lambda, C_1\lambda]$$

such that $Q_{\vartheta r}^{\lambda_1} \subset Q_{\frac{\sigma}{2}r}^\lambda$ with $\sigma \in (0,1)$ from Lemma 6.1,

$$\fint_{Q_{\vartheta r}^{\lambda_1}} \varphi_{|(D\mathbf{u})_{\vartheta r}^{\lambda_1}|}(|D\mathbf{u} - (D\mathbf{u})_{\vartheta r}^{\lambda_1}|)\, \mathrm{d}z \leqslant \varphi(\lambda_1) \quad \textit{and} \quad |(D\mathbf{u})_{2\vartheta r}^{\lambda_1}| \leqslant \lambda_1 . \tag{6.13}$$

Additionally, for each $\alpha \in (0, \alpha_1)$ there exists small $\vartheta_1 \in (0, C_1^{-1}]$ depending on $n, N, p, q, \nu, L, M_1, \alpha_1$, α and χ_1 such that for every $\vartheta \in (0, \vartheta_1]$ if

$$\chi_1 \varphi(|(D\mathbf{u})_{\vartheta r}^{\lambda_1}|) \leqslant \fint_{Q_{\vartheta r}^{\lambda_1}} \varphi_{|(D\mathbf{u})_{\vartheta r}^{\lambda_1}|}(|D\mathbf{u} - (D\mathbf{u})_{\vartheta r}^{\lambda_1}|)\, \mathrm{d}z \quad \textit{or} \quad |(D\mathbf{u})_{\vartheta r}^{\lambda_1}| \leqslant \frac{\lambda_1}{K_1}, \tag{6.14}$$

then

$$\lambda_1 \leqslant \vartheta^\alpha \lambda. \tag{6.15}$$

Proof. We first observe that, from all the assumptions and (2.14), we get

$$\begin{aligned} \fint_{Q_r^\lambda} \varphi(|D\mathbf{u}|)\, \mathrm{d}z &\leqslant c \left(\fint_{Q_r^\lambda} \varphi_{|(D\mathbf{u})_r^\lambda|}(|D\mathbf{u} - (D\mathbf{u})_r^\lambda|)\, \mathrm{d}z + \varphi(|(D\mathbf{u})_r^\lambda|) \right) \\ &\leqslant \begin{cases} c(1 + M_1^q)\varphi(\lambda), \\ c\left\{(1 + \chi^{-1})\varepsilon_1 + K^{-q}\right\}. \end{cases} \end{aligned} \tag{6.16}$$

We divide the proof into two steps.

Step 1. (φ-caloric approximation) We show that $\mathbf{u}$ is an almost φ-caloric mapping. Namely, for every $\boldsymbol{\zeta} \in C_0^\infty(Q_r^\lambda)$

$$\left| \fint_{Q_r^\lambda} \mathbf{u} \cdot \boldsymbol{\zeta}_t - \frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|} \langle D\mathbf{u}, D\boldsymbol{\zeta} \rangle\, \mathrm{d}z \right| \leqslant c\varepsilon_0 \left(\fint_{Q_r^\lambda} \varphi(|D\mathbf{u}|)\, \mathrm{d}z + \varphi(\|D\boldsymbol{\zeta}\|_\infty) \right), \tag{6.17}$$

where $\varepsilon_0 > 0$ is a sufficiently small constant determined later. From the weak form of (1.1) we have

$$\left| \fint_{Q_r^\lambda} \mathbf{u} \cdot \boldsymbol{\zeta}_t - \frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|} \langle D\mathbf{u}, D\boldsymbol{\zeta} \rangle\, \mathrm{d}z \right| = \left| \fint_{Q_r^\lambda} \langle \mathbf{A}(D\mathbf{u}) - \frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|} D\mathbf{u}, D\boldsymbol{\zeta} \rangle\, \mathrm{d}z \right| .$$

Choose $\delta = \delta(\varepsilon_0)$ such that (2.18) holds for $\varepsilon = \varepsilon_0$, and set $E_1 := \{z \in Q_r^\lambda : |D\mathbf{u}(z)| \leqslant \delta\}$ and $E_2 = Q_r^\lambda \setminus E_1$. Then one has from the Young inequality (2.2) and (2.3) that

$$\begin{aligned}\frac{1}{|Q_r^\lambda|}\left|\int_{E_1} \langle \mathbf{A}(D\mathbf{u}) - \frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|} D\mathbf{u}, D\boldsymbol{\zeta}\rangle \,\mathrm{d}z\right| &\leqslant \varepsilon_0 \fint_{Q_r^\lambda} \varphi'(|D\mathbf{u}|)\|D\boldsymbol{\zeta}\|_\infty \,\mathrm{d}z \\ &\leqslant c\varepsilon_0 \left(\fint_{Q_r^\lambda} \varphi(|D\mathbf{u}|)\,\mathrm{d}z + \varphi(\|D\boldsymbol{\zeta}\|_\infty)\right).\end{aligned}$$

On the other hand, by (2.16), the Young inequality (2.2), (2.5) and (6.16),

$$\begin{aligned}&\frac{1}{|Q_r^\lambda|}\left|\int_{E_2} \langle \mathbf{A}(D\mathbf{u}) - \frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|} D\mathbf{u}, D\boldsymbol{\zeta}\rangle \,\mathrm{d}z\right| \\ &\leqslant \frac{1}{|Q_r^\lambda|}\int_{E_2} \left(|\mathbf{A}(D\mathbf{u})| + \varphi'(|D\mathbf{u}|)\right) \|D\boldsymbol{\zeta}\|_\infty \,\mathrm{d}z \\ &\leqslant \frac{c\|D\boldsymbol{\zeta}\|_\infty}{\delta(\varepsilon_0)} \fint_{Q_r^\lambda} \varphi(|D\mathbf{u}|)\,\mathrm{d}z \\ &\leqslant c\varphi^*\left(\frac{1}{\delta(\varepsilon_0)\varepsilon_0^{1/p}} \fint_{Q_r^\lambda} \varphi(|D\mathbf{u}|)\,\mathrm{d}z\right) + \varepsilon_0\varphi(\|D\boldsymbol{\zeta}\|_\infty) \\ &\leqslant \frac{c}{\delta(\varepsilon_0)^q\varepsilon_0^{q/p}}\varphi^*\left(\fint_{Q_r^\lambda} \varphi(|D\mathbf{u}|)\,\mathrm{d}z\right) + \varepsilon_0\varphi(\|D\boldsymbol{\zeta}\|_\infty) \\ &\leqslant \frac{c}{\delta(\varepsilon_0)^q\varepsilon_0^{q/p}}\psi\left((1+\chi^{-1})\varepsilon_1 + K^{-q}\right) \fint_{Q_r^\lambda} \varphi(|D\mathbf{u}|)\,\mathrm{d}z + \varepsilon_0\varphi(\|D\boldsymbol{\zeta}\|_\infty),\end{aligned}$$

where $\psi(t) := \frac{\varphi^*(t)}{t}$. Therefore, choosing $\varepsilon_1 = \varepsilon_1(\chi, \varepsilon_0)$ sufficiently small and $K = K(\varepsilon_0)$ sufficiently large, we obtain (6.17).

Therefore, by Theorem 3.9 applied with $\mathbf{G} := \mathbf{A}(D\mathbf{u})$ and $\gamma_1 := \frac{1}{2}$, we have that for a constant $\varepsilon > 0$ to be determined later, there exists $\varepsilon_0 = \varepsilon_0(\varepsilon)$ such that

$$\left(\fint_{Q_r^\lambda} |\mathbf{V}(D\mathbf{u}) - \mathbf{V}(D\mathbf{h})|\,\mathrm{d}z\right)^2 \leqslant \varepsilon \fint_{Q_r^\lambda} \varphi(|D\mathbf{u}|)\,\mathrm{d}z \leqslant c\varepsilon\varphi(\lambda), \tag{6.18}$$

where $\mathbf{h}$ is the weak solution to

$$\begin{cases}\partial_t \mathbf{h} - \operatorname{div}\left(\frac{\varphi'(|D\mathbf{h}|)}{|D\mathbf{h}|} D\mathbf{h}\right) = \mathbf{0} & \text{in } Q_r^\lambda, \\ \mathbf{h} = \mathbf{u} & \text{in } \partial_{\mathrm{p}} Q_r^\lambda.\end{cases}$$

The existence and uniqueness of the solution $\mathbf{h}$ follows from the theory of monotone operators or by utilizing the Galerkin approximation method, see for instance [38]. Here from higher integrability result in (4.20) with $\boldsymbol{\ell} = \mathbf{0}$ and the Lipschitz estimate for $\mathbf{h}$ in (6.2) we have

$$\left(\fint_{Q_{r/2}^\lambda} \varphi(|D\mathbf{u}|)^{1+\sigma}\,\mathrm{d}z\right)^{\frac{1}{1+\sigma}} \leqslant c\varphi(\lambda) \tag{6.19}$$

and

$$\sup_{Q_{r/2}^\lambda} \varphi(|D\mathbf{h}|) \leqslant c\varphi(\lambda), \tag{6.20}$$

since

$$\fint_{Q_r^\lambda} \varphi(|D\mathbf{h}|)\,\mathrm{d}z \leqslant c\fint_{Q_r^\lambda} \varphi(|D\mathbf{u}|)\,\mathrm{d}z \leqslant c\varphi(\lambda)$$

by a standard energy estimate of the above system. Moreover, with the change of shift formula (2.15) and the triangle inequality, we get

$$\begin{aligned}\fint_{Q^\lambda_{r/2}} \varphi^{1+\sigma}_{|D\mathbf{h}|}(|D\mathbf{u}-D\mathbf{h}|)\,\mathrm{d}z &\leqslant c\left(c_\eta \fint_{Q^\lambda_{r/2}} \varphi^{1+\sigma}(|D\mathbf{u}-D\mathbf{h}|)\,\mathrm{d}z + c\eta\varphi^{1+\sigma}(|D\mathbf{h}|)\right)\\ &\leqslant \tilde c\left(\fint_{Q^\lambda_{r/2}} \varphi^{1+\sigma}(|D\mathbf{u}|)\,\mathrm{d}z + \varphi^{1+\sigma}(|D\mathbf{h}|)\right),\end{aligned}$$

whence

$$\left(\fint_{Q^\lambda_{r/2}} \varphi^{1+\sigma}_{|D\mathbf{h}|}(|D\mathbf{u}-D\mathbf{h}|)\,\mathrm{d}z\right)^{\frac{1}{1+\sigma}} \leqslant c\left(\left(\fint_{Q^\lambda_{r/2}} \varphi^{1+\sigma}(|D\mathbf{u}|)\,\mathrm{d}z\right)^{\frac{1}{1+\sigma}} + \varphi(|D\mathbf{h}|)\right).$$

This, together with (6.19) and (6.20), implies

$$\left(\fint_{Q^\lambda_{r/2}} |\mathbf{V}(D\mathbf{u})-\mathbf{V}(D\mathbf{h})|^{2(1+\sigma)}\,\mathrm{d}z\right)^{\frac{1}{1+\sigma}} \leqslant c\varphi(\lambda)\,. \tag{6.21}$$

Therefore, by interpolation, choosing $\tau\in(0,1)$ such that $\frac{1-\tau}{2}+\tau(1+\sigma)=1$; i.e., $\tau=\frac{1}{1+2\sigma}$, we have, with (6.18), (6.21) and the last inequality in (6.20),

$$\begin{aligned}&\fint_{Q^\lambda_{r/2}} |\mathbf{V}(D\mathbf{u})-\mathbf{V}(D\mathbf{h})|^2\,\mathrm{d}z\\ &\leqslant \left(\fint_{Q^\lambda_{r/2}} |\mathbf{V}(D\mathbf{u})-\mathbf{V}(D\mathbf{h})|\,\mathrm{d}z\right)^{1-\tau}\left(\fint_{Q^\lambda_{r/2}} |\mathbf{V}(D\mathbf{u})-\mathbf{V}(D\mathbf{h})|^{2(1+\sigma)}\,\mathrm{d}z\right)^{\tau}\\ &\leqslant c\varepsilon^{\frac{1-\tau}{2}}\varphi(\lambda)\,.\end{aligned} \tag{6.22}$$

Moreover, by applying (2.11), with (6.21) and again by (6.20), we also have

$$\begin{aligned}&\varphi\left(\fint_{Q^\lambda_{r/2}} |D\mathbf{u}-D\mathbf{h}|\,\mathrm{d}z\right) \leqslant \fint_{Q^\lambda_{r/2}} \varphi(|D\mathbf{u}-D\mathbf{h}|)\,\mathrm{d}z\\ &\leqslant \varepsilon^{\frac{1-\tau}{4}}\fint_{Q^\lambda_{r/2}} [\varphi(|D\mathbf{u}|)+\varphi(|D\mathbf{h}|)]\,\mathrm{d}z + c\varepsilon^{-\frac{1-\tau}{4}}\fint_{Q^\lambda_{r/2}} |\mathbf{V}(D\mathbf{u})-\mathbf{V}(D\mathbf{h})|^2\,\mathrm{d}z\\ &\leqslant c\varepsilon^{\frac{1-\tau}{4}}\varphi(\lambda)\,,\end{aligned}$$

hence

$$\fint_{Q^\lambda_{r/2}} |D\mathbf{u}-D\mathbf{h}|\,\mathrm{d}z \leqslant c\varepsilon^{\frac{1-\tau}{4q}}\lambda\,. \tag{6.23}$$

Step 2. Let $0<\theta\leqslant 1$. Applying Lemma 6.1 with $R=r/2$ and writing $\lambda_\theta:=\lambda_{\theta r}$ for each $\theta\in(0,1]$ and $\theta_s:=\frac{2r_s}{r}$, we have that

$$\lambda_\theta=\lambda_{\theta_s} \ \text{ if } \ \theta\in(0,\theta_s] \quad \text{and} \quad \theta^{\alpha_1}\lambda\leqslant\lambda_\theta\leqslant 2\theta^{\alpha_1}\lambda \ \text{ if } \ \theta\in(\theta_s,1]\,, \tag{6.24}$$

$$\sup_{Q^{\lambda_\theta}_{\theta r}} |D\mathbf{h}|\leqslant\lambda_\theta \quad \text{for all} \quad \theta\in(0,1]\,, \tag{6.25}$$

$$\fint_{Q^{\lambda_\theta}_{\theta r}} \varphi_{|(D\mathbf{h})^{\lambda_\theta}_{\theta r}|}(|D\mathbf{h}-(D\mathbf{h})^{\lambda_\theta}_{\theta r}|)\,\mathrm{d}z \leqslant c\left(\frac{\theta}{\theta_s}\right)^{3/4}\varphi(\lambda_\theta) \quad \text{if } \ \theta\in(0,\theta_s] \tag{6.26}$$

and

$$|(D\mathbf{h})^{\lambda_\theta}_{\theta r}| \geqslant \tfrac{1}{C_s}\lambda_\theta \quad \text{and} \quad \operatorname*{osc}_{Q^{\lambda_\theta}_{\theta r}} D\mathbf{h} \leqslant c\Big(\frac{\theta}{\theta_s}\Big)^{\alpha_1}\lambda_\theta \quad \text{if} \quad \theta \in (0,\theta_s]\,. \tag{6.27}$$

Moreover, we have from Remark 6.2 (2) that $Q^{\lambda_\theta}_{\theta r} \subset Q^{\lambda}_{\frac{\sigma}{2}r}$.

For $0 < \theta \leqslant \tilde\sigma$ with $\tilde\sigma \in (0,1)$ as in (6.8) and a large constant $C_0 > 1$ to be determined later, set

$$C_1 := \max\{2C_0, 2C_0^{\frac{2-p}{2}}\tilde\sigma^{-1}\}.$$

For $\vartheta \in (0, C_1^{-1}]$ we set

$$\theta := 2C_0^{\frac{2-p}{2}}\vartheta \in (0,\tilde\sigma] \quad \text{and then} \quad \lambda_1 := C_0\lambda_\theta\,. \tag{6.28}$$

Note that $Q^{\lambda_1}_{2\vartheta r} \subset Q^{\lambda_\theta}_{\theta r} \subset Q^{\lambda}_{\frac{\sigma}{2}r}$ since $\frac{(2\vartheta)^2\lambda_1^2}{\varphi(\lambda_1)} \leqslant \frac{4C_0^{2-p}\vartheta^2\lambda_\theta^2}{\varphi(\lambda_\theta)} = \frac{\theta^2\lambda_\theta^2}{\varphi(\lambda_\theta)}$ and that by (6.24)

$$\vartheta^{\alpha_1}\lambda \leqslant \theta^{\alpha_1}\lambda \leqslant \lambda_\theta \leqslant \lambda_1 \leqslant 2\theta^{\alpha_1}C_0\lambda \leqslant 2C_0\lambda \leqslant C_1\lambda\,.$$

We then prove (6.13) for $\vartheta \in (0, C_1^{-1}]$ with choosing ε and C_0. Note that using (2.7), (2.10) and (6.24), we have

$$\begin{aligned}
&\fint_{Q^{\lambda_1}_{\vartheta r}} \varphi_{|(D\mathbf{u})^{\lambda_1}_{\vartheta r}|}(|D\mathbf{u} - (D\mathbf{u})^{\lambda_1}_{\vartheta r}|)\,\mathrm{d}z \leqslant c\fint_{Q^{\lambda_1}_{\vartheta r}} |\mathbf{V}(D\mathbf{u}) - \mathbf{V}((D\mathbf{u})^{\lambda_1}_{\vartheta r})|^2\,\mathrm{d}z\\
&\leqslant c\fint_{Q^{\lambda_1}_{\vartheta r}} |\mathbf{V}(D\mathbf{u}) - (\mathbf{V}(D\mathbf{u}))^{\lambda_1}_{\vartheta r}|^2\,\mathrm{d}z \leqslant c\fint_{Q^{\lambda_1}_{\vartheta r}} |\mathbf{V}(D\mathbf{u}) - (\mathbf{V}(D\mathbf{h}))^{\lambda_1}_{\vartheta r}|^2\,\mathrm{d}z\\
&\leqslant c\fint_{Q^{\lambda_1}_{\vartheta r}} |\mathbf{V}(D\mathbf{u}) - \mathbf{V}(D\mathbf{h})|^2\,\mathrm{d}z + c\fint_{Q^{\lambda_1}_{\vartheta r}} |\mathbf{V}(D\mathbf{h}) - (\mathbf{V}(D\mathbf{h}))^{\lambda_1}_{\vartheta r}|^2\,\mathrm{d}z\\
&\leqslant c\frac{\varphi(\lambda_1)\lambda_\theta^2\vartheta^{n+2}}{\varphi(\lambda_\theta)\lambda_1^2\theta^{n+2}}\theta^{-(n+2)-(q-2)\alpha_1}\fint_{Q^{\lambda}_{r}} |\mathbf{V}(D\mathbf{u}) - \mathbf{V}(D\mathbf{h})|^2\,\mathrm{d}z + c\sup_{Q^{\lambda_\theta}_{\theta r}}\varphi(|D\mathbf{h}|)\,.
\end{aligned}$$

For the second term on the right hand side, by (6.25) and (6.28), we have

$$\sup_{Q^{\lambda_\theta}_{\theta r}}\varphi(|D\mathbf{h}|) \leqslant \varphi(C_0^{-1}\lambda_1) \leqslant C_0^{-p}\varphi(\lambda_1)\,.$$

As for the first term on the right hand side, by (6.28) and (6.22), we have

$$\begin{aligned}
&\frac{\varphi(\lambda_1)\lambda_\theta^2\vartheta^{n+2}}{\varphi(\lambda_\theta)\lambda_1^2\theta^{n+2}}\theta^{-(n+2)-(q-2)\alpha_1}\fint_{Q^{\lambda}_{r}} |\mathbf{V}(D\mathbf{u}) - \mathbf{V}(D\mathbf{h})|^2\,\mathrm{d}z\\
&\leqslant cC_0^{q-2+\frac{(2-p)(2n+4+(q-2)\alpha_1)}{2}}\vartheta^{-(n+2)-(q-2)\alpha_1}\fint_{Q^{\lambda}_{r}} |\mathbf{V}(D\mathbf{u}) - \mathbf{V}(D\mathbf{h})|^2\,\mathrm{d}z\\
&\leqslant cC_0^{q-2+\frac{(2-p)(2n+4+(q-2)\alpha_1)}{2}-p}\vartheta^{-(n+2)-(q-1)\alpha_1}\varepsilon^{\frac{1-\tau}{2}}\varphi(\lambda_1)\,.
\end{aligned}$$

Therefore, choosing C_0 large and then $\varepsilon = \varepsilon(\vartheta)$ small, we obtain the first estimate in (6.13). Moreover, in a similar way with Jensen's inequality, we also have

$$\begin{aligned}
\varphi\left(|(D\mathbf{u})^{\lambda_1}_{2\vartheta r}|\right) &\leqslant \fint_{Q^{\lambda_1}_{2\vartheta r}} \varphi(|D\mathbf{u}|)\,\mathrm{d}z \leqslant c\fint_{Q^{\lambda_1}_{2\vartheta r}} |\mathbf{V}(D\mathbf{u})|^2\,\mathrm{d}z\\
&\leqslant c\fint_{Q^{\lambda_1}_{2\vartheta r}} |\mathbf{V}(D\mathbf{u}) - \mathbf{V}(D\mathbf{h})|^2\,\mathrm{d}z + c\fint_{Q^{\lambda_1}_{2\vartheta r}} |\mathbf{V}(D\mathbf{h})|^2\,\mathrm{d}z\\
&\leqslant cC_0^{q-2+\frac{(2-p)(2n+4+(q-2)\alpha_1)}{2}-p}\vartheta^{-(n+2)-(q-1)\alpha_1}\varepsilon^{\frac{1-\tau}{2}}\varphi(\lambda_1) + cC_0^{-p}\varphi(\lambda_1) \leqslant \varphi(\lambda_1)\,,
\end{aligned}$$

provided that C_0 is large enough and then $\varepsilon=\varepsilon(\vartheta)$ is small enough, which together with the monotonicity of φ implies the second estimate in (6.13).

We next prove (6.15) under the condition (6.14), by choosing ϑ sufficiently small. Recall the definition of θ and we distinguish between two cases. If $\theta\in[\theta_s,1)$, by (6.24) and (6.28), we have that for each $\alpha\in(0,\alpha_1)$

$$\lambda_1=C_0\lambda_\theta\leqslant 2C_0\theta^{\alpha_1}\lambda=2C_0\Big(2C_0^{\frac{2-p}{2}}\Big)^{\alpha_1}\vartheta^{\alpha_1-\alpha}\vartheta^{\alpha}\lambda\leqslant\vartheta^{\alpha}\lambda,$$

where we chose ϑ small to satisfy the last inequality. Therefore, we obtain (6.15) without considering (6.14). On the other hand, let $\theta\in(0,\theta_s)$. Note that, as $\vartheta<\theta<\theta_s$, there hold $\lambda_{\theta_s}=\lambda_\theta=\lambda_\vartheta$ and $Q^{\lambda_1}_{\vartheta r}\cup Q^{\lambda_\vartheta}_{\vartheta r}\subset Q^{\lambda_\theta}_{\theta r}$. Hence, by (2.11) and (6.23)

$$\begin{aligned}
|(D\mathbf{u})^{\lambda_1}_{\vartheta r}-(D\mathbf{h})^{\lambda_\theta}_{\theta r}|&\leqslant c\fint_{Q^{\lambda_\theta}_{\theta r}}|D\mathbf{u}-(D\mathbf{h})^{\lambda_\theta}_{\theta r}|\,\mathrm{d}z\\
&\leqslant c\fint_{Q^{\lambda_\theta}_{\theta r}}|D\mathbf{u}-D\mathbf{h}|\,\mathrm{d}z+c\fint_{Q^{\lambda_\theta}_{\theta r}}|D\mathbf{h}-(D\mathbf{h})^{\lambda_\theta}_{\theta r}|\,\mathrm{d}z\\
&\leqslant c\theta^{-n-2-\alpha_1(q-2)}\fint_{Q^{\lambda}_{r/2}}|D\mathbf{u}-D\mathbf{h}|\,\mathrm{d}z+c\left(\frac{\theta}{\theta_s}\right)^{\alpha_1}\lambda_\theta\\
&\leqslant c\vartheta^{-n-2-\alpha_1(q-2)}\varepsilon^{\frac{1-\tau}{4q}}\lambda+c\left(\frac{\theta}{\theta_s}\right)^{\alpha_1}\lambda_\theta\\
&\leqslant c\vartheta^{-\kappa_1}\varepsilon^{\kappa_2}\lambda_1+c\left(\frac{\theta}{\theta_s}\right)^{\alpha_1}\lambda_1\,,
\end{aligned}$$

where $\kappa_1=-n-2-\alpha_1(q-2)-\alpha_1$ and $\kappa_2=\frac{1-\tau}{4q}$. Then using (6.27) and choosing $\varepsilon=\varepsilon(\vartheta)$ sufficiently small we have

$$\begin{aligned}
|(D\mathbf{u})^{\lambda_1}_{\vartheta r}|&\geqslant|(D\mathbf{h})^{\lambda_\theta}_{\theta r}|-|(D\mathbf{u})^{\lambda_1}_{\vartheta r}-(D\mathbf{h})^{\lambda_\theta}_{\theta r}|\\
&\geqslant\left(C_0^{-1}C_s^{-1}-c\vartheta^{-\kappa_1}\varepsilon^{\kappa_2}-c\left(\frac{\theta}{\theta_s}\right)^{\alpha_1}\right)\lambda_1\\
&\geqslant\left(\frac{1}{2C_0C_s}-c\left(\frac{\theta}{\theta_s}\right)^{\alpha_1}\right)\lambda_1\,.
\end{aligned}$$

Moreover, if the first condition in (6.14) holds, then by (2.7) and (6.26) we have

$$\begin{aligned}
|(D\mathbf{u})^{\lambda_1}_{\vartheta r}|&\leqslant\chi_1^{-1}\fint_{Q^{\lambda_1}_{\vartheta r}}\varphi_{|(D\mathbf{u})^{\lambda_1}_{\vartheta r}|}(|D\mathbf{u}-(D\mathbf{u})^{\lambda_1}_{\vartheta r}|)\,\mathrm{d}z\leqslant c\chi_1^{-1}\fint_{Q^{\lambda_1}_{\vartheta r}}|\mathbf{V}(D\mathbf{u})-(\mathbf{V}(D\mathbf{u}))^{\lambda_1}_{\vartheta r}|^2\,\mathrm{d}z\\
&\leqslant c\chi_1^{-1}\fint_{Q^{\lambda_\theta}_{\theta r}}|\mathbf{V}(D\mathbf{u})-(\mathbf{V}(D\mathbf{u}))^{\lambda_\theta}_{\theta r}|^2\,\mathrm{d}z\leqslant c\chi_1^{-1}\Big(\frac{\theta}{\theta_s}\Big)^{3/4}\varphi(\lambda_\theta)\,,
\end{aligned}$$

which implies that

$$\left(\frac{1}{2C_0C_s}-c\left(\frac{\theta}{\theta_s}\right)^{\alpha_1}\right)\lambda_1\leqslant c\chi_1^{-q}\left(\frac{\theta}{\theta_s}\right)^{\frac{3}{4q}}\lambda_\theta\leqslant c\chi_1^{-q}\left(\frac{\theta}{\theta_s}\right)^{\alpha_1}\lambda_1,$$

hence

$$\theta_s^{\alpha_1}\leqslant c\chi_1^{-q}\theta^{\alpha_1}\,. \tag{6.29}$$

If the second condition in (6.14) holds, then choosing $K_1\geqslant 4C_0C_s$,

$$\left(\frac{1}{2C_0C_s}-c\left(\frac{\theta}{\theta_s}\right)^{\alpha_1}\right)\lambda_1\leqslant\frac{\lambda_1}{K_1}\leqslant\frac{\lambda_1}{4C_0C_s}\,,$$

which implies (6.29) again. Consequently, we have

$$\lambda_1 = C_0\lambda_\theta = 2C_0\theta_s^{\alpha_1}\lambda \leqslant c\chi_1^{-q}\theta^{\alpha_1}\lambda \leqslant c\chi_1^{-q}\vartheta^{\alpha_1-\alpha}\vartheta^\alpha\lambda \leqslant \vartheta^\alpha\lambda\,,$$

Therefore, choosing ϑ_1 sufficiently small depending on χ_1, we get

$$\lambda_1 \leqslant \vartheta_1^\alpha\lambda\,,$$

whenever $\vartheta \in (0,\vartheta_1]$. □

7. Iteration and Proof of Theorem 1.1

With the following result we will combine the degenerate and the nondegenerate regimes by an inductive iteration scheme. Roughly speaking, as long as on an iteration scale the degenerate case holds, we shall apply on this scale Lemma 6.3. On the other hand, when on an iteration scale the nondegenerate regime occurs we can apply Lemma 5.3 which provides a suitable excess decay estimate. Either this happens at a certain scale $\vartheta^m R$, $m<\infty$, or we go on by iterating the excess improvement from the degenerate case on each scale (i.e, $m=\infty$), thus obtaining the desired excess decay estimate.

Lemma 7.1. *Let $\mathbf{u}$ be a weak solution to (1.1), $M_0 \geqslant 1$ and $\alpha\in(0,\alpha_1)$. There exist small $\vartheta,\varepsilon_2\in(0,1)$ and large $K_2\geqslant 1$ depending on n,N,p,q,ν,L,α_1 and M_0 such that the following holds: suppose $Q_{2R}(z_0)\Subset\Omega_T$,*

$$|(D\mathbf{u})_{Q_{2R}(z_0)}| \leqslant M_0\,, \tag{7.1}$$

and

$$\fint_{Q_R(z_0)} \varphi_{|(D\mathbf{u})_{Q_R(z_0)}|}(|D\mathbf{u}-(D\mathbf{u})_{Q_R(z_0)}|)\,\mathrm{d}z \leqslant \varepsilon_2\,. \tag{7.2}$$

Then the limit

$$\Gamma_{z_0} := \lim_{r\to0^+}(D\mathbf{u})_{Q_r(z_0)}$$

exists and there exist $m\in\mathbb{N}_0\cup\{\infty\}$ and positive numbers λ_j with $j\in\{0,1,2,\dots,m\}$ such that $\lambda_0=1$,

$$\vartheta^{\alpha_1}\lambda_{j-1}\leqslant\lambda_j\leqslant\vartheta^\alpha\lambda_{j-1}\quad\text{for } 1\leqslant j<m\,,\quad \vartheta^{\alpha_1}\lambda_{m-1}\leqslant\lambda_m\leqslant C_1\lambda_{m-1}\quad\text{for } 0<m<\infty\,, \tag{7.3}$$

$$Q^{\lambda_j}_{\vartheta^j R}(z_0)\subset Q^{\lambda_{j-1}}_{\sigma_1\vartheta^{j-1}R}(z_0)\qquad\text{for }\ 1\leqslant j\leqslant m\ \text{ with }\ j<\infty\,, \tag{7.4}$$

$$\frac{\lambda_m}{2K_2}\leqslant|\Gamma_{z_0}|\leqslant 2K_2\lambda_m\ \text{ if }\ m<\infty\quad\text{and}\quad\Gamma_{z_0}=\mathbf{0}\ \text{ if }\ m=\infty\,, \tag{7.5}$$

$$\fint_{Q^{\lambda_j}_{\vartheta^j R}(z_0)}\varphi(|D\mathbf{u}-\Gamma_{z_0}|)\,\mathrm{d}z\leqslant c\varphi(\lambda_j)\qquad\text{for }\ 0\leqslant j\leqslant m\ \text{ with }\ j<\infty\,, \tag{7.6}$$

and, if $m<\infty$,

$$\fint_{Q^{\lambda_m}_r(z_0)}\varphi(|D\mathbf{u}-\Gamma_{z_0}|)\,\mathrm{d}z\leqslant c\left(\frac{r}{\vartheta^m R}\right)^{\alpha_2}\varphi(\lambda_m)\quad\text{for all }\ 0<r\leqslant\vartheta^m R\,. \tag{7.7}$$

Proof. Step 1. (Choice of parameters) Without loss of generality, we may assume that $z_0=0$. We start by fixing the parameters. Let $M_1:=2^{n+2}M_0$. First, we fix K_1 and C_1 as in Lemma 6.3, and then choose $\delta_1=\delta_1(K_1)$ in Lemma 5.3 with $K_0=\max\{M_1,2^{n+2}K_1\}$, and set $\chi_1:=\delta_1(K_1)$ in Lemma 6.3. Then we next determine ϑ_1 as in Lemma 6.3, and then K as in Lemma 6.3 with $\vartheta\leqslant\vartheta_1$. Then choose $\delta_1=\delta_1(K)$ in Lemma 5.3 with

$K_0 = \max\{M_1, 2^{n+2}K\}$, and set $\chi := \delta_1(K)$ in Lemma 6.3. We then determine ε_1 as in Lemma 6.3. Note that there exists $j_* \in \mathbb{N}$ such that

$$\vartheta^{\alpha p(j_*+1)} \leqslant \varepsilon_1 \,. \tag{7.8}$$

With this j_*, we determine

$$\varepsilon_2 = C_2^{-1}\vartheta^{(n+4-p)(j_*+1)}\varepsilon_1 \,, \tag{7.9}$$

where $C_2 > 1$ is a large constant to be determined in (7.14). Finally, set

$$K_2 = \max\{M_1, 2^{n+2}K_1, 2^{n+2}K\} \,.$$

We next choose λ_j inductively. Set $\lambda_0 = 1$. For some $j \in \mathbb{N}_0$ and $\lambda_j \in (0,1]$, we consider the following condition:

$$\chi|(D\mathbf{u})^{\lambda_j}_{\vartheta^j R}| \leqslant \fint_{Q^{\lambda_j}_{\vartheta^j R}} \varphi_{|(D\mathbf{u})^{\lambda_j}_{\vartheta^j R}|}(|D\mathbf{u} - (D\mathbf{u})^{\lambda_j}_{\vartheta^j R}|)\,\mathrm{d}z \quad \text{or} \quad |(D\mathbf{u})^{\lambda_j}_{\vartheta^j R}| \leqslant \frac{\lambda_j}{K} \,. \tag{7.10}$$

If this condition holds, then by Lemma 6.3 with $\lambda_j = \lambda$ there exists $\lambda_{j+1} \in [\vartheta^{\alpha_1}\lambda_j, C_1\lambda_j]$ such that $Q^{\lambda_{j+1}}_{\vartheta^{j+1}R} \subset Q^{\lambda_j}_{\frac{\sigma}{2}\vartheta^j R}$,

$$\fint_{Q^{\lambda_{j+1}}_{\vartheta^{j+1} R}} \varphi_{|(D\mathbf{u})^{\lambda_{j+1}}_{\vartheta^{j+1} R}|}(|D\mathbf{u} - (D\mathbf{u})^{\lambda_{j+1}}_{\vartheta^{j+1} R}|)\,\mathrm{d}z \leqslant \varphi(\lambda_{j+1}) \quad \text{and} \quad |(D\mathbf{u})^{\lambda_{j+1}}_{2\vartheta^{j+1} R}| \leqslant \lambda_{j+1} \,. \tag{7.11}$$

Then we have two cases:

$$\chi_1|(D\mathbf{u})^{\lambda_{j+1}}_{\vartheta^{j+1} R}| \leqslant \fint_{Q^{\lambda_{j+1}}_{\vartheta^{j+1} R}} \varphi_{|(D\mathbf{u})^{\lambda_{j+1}}_{\vartheta^{j+1} R}|}(|D\mathbf{u} - (D\mathbf{u})^{\lambda_{j+1}}_{\vartheta^{j+1} R}|)\,\mathrm{d}z \quad \text{or} \quad |(D\mathbf{u})^{\lambda_{j+1}}_{\vartheta^{j+1} R}| \leqslant \frac{\lambda_{j+1}}{K_1} \,, \tag{7.12}$$

and the other case. Set $\widetilde{\mathbb{N}} := \{j \in \mathbb{N}_0 : (7.10) \text{ with } j \text{ does not holds.}\}$ and $m_1 \in \mathbb{N}_0 \cup \{\infty\}$ such that

$$\begin{cases} m_1 = \min \widetilde{\mathbb{N}} & \text{if } \widetilde{\mathbb{N}} \neq \emptyset \,, \\ m_1 = \infty & \text{if } \widetilde{\mathbb{N}} = \emptyset \,. \end{cases}$$

Step 2. (Nondegenerate decay) If $m_1 = 0$, then the lemma with $m = 0$ follows directly from Lemma 5.3 with $R = r$, $K_0 = 2^{n+2}\max\{M_1, K_1\}$ and $\delta_1 = \delta_1(K)$.

We next suppose $m_1 \in \mathbb{N} \cup \{\infty\}$. If there exists $m_0 \in \mathbb{N}$ with $m_0 \leqslant m_1 + 1$ such that (7.12) holds for all $j < m_0 - 1$ but not $j = m_0 - 1$ then we choose $m = m_0$. On the other hand, if $m_1 \in \mathbb{N}$ and (7.12) holds for all $j \leqslant m_1 - 1$, then we choose $m = m_1$. We note that if all the assumptions of Lemma 6.3, with $\theta^j r$ in place of r, hold and $\lambda_j \leqslant \vartheta^{\alpha j}$, then $\lambda_{j+1} \leqslant \vartheta^{\alpha(j+1)}$,

$$\fint_{Q^{\lambda_{j+1}}_{\vartheta^{j+1} R}} \varphi_{|(D\mathbf{u})^{\lambda_{j+1}}_{\vartheta^{j+1} R}|}(|D\mathbf{u} - (D\mathbf{u})^{\lambda_{j+1}}_{\vartheta^{j+1} R}|)\,\mathrm{d}z \leqslant \varphi(\lambda_{j+1}) \leqslant \varphi(\vartheta^{\alpha(j+1)}) \leqslant \vartheta^{\alpha p(j+1)} \tag{7.13}$$

and

$$|(D\mathbf{u})^{\lambda_{j+1}}_{\vartheta^{j+1} R}| \leqslant 2^{n+2}|(D\mathbf{u})^{\lambda_{j+1}}_{2\vartheta^{j+1} R}| \leqslant 2^{n+2}\lambda_{j+1} \leqslant M_1\lambda_{j+1} \,.$$

Moreover if $j \geqslant j_*$, the estimate (7.13) together with (7.8) implies

$$\fint_{Q^{\lambda_{j+1}}_{\vartheta^{j+1} R}} \varphi_{|(D\mathbf{u})^{\lambda_{j+1}}_{\vartheta^{j+1} R}|}(|D\mathbf{u} - (D\mathbf{u})^{\lambda_{j+1}}_{\vartheta^{j+1} R}|)\,\mathrm{d}z \leqslant \varepsilon_1 \,,$$

while if $j < j_*$, then by (7.9), taking into account (2.7) and (2.10),

$$
\begin{aligned}
&\fint_{Q^{\lambda_{j+1}}_{\vartheta^{j+1}R}} \varphi_{|(D\mathbf{u})^{\lambda_{j+1}}_{\vartheta^{j+1}R}|}(|D\mathbf{u}-(D\mathbf{u})^{\lambda_{j+1}}_{\vartheta^{j+1}R}|)\,\mathrm{d}z \\
&\leqslant c\fint_{Q^{\lambda_{j+1}}_{\vartheta^{j+1}R}} |\mathbf{V}(D\mathbf{u})-(\mathbf{V}(D\mathbf{u}))^{\lambda_{j+1}}_{\vartheta^{j+1}R}|^2\,\mathrm{d}z \\
&\leqslant c\vartheta^{-(n+4-p)(j+1)}\fint_{Q_R} |\mathbf{V}(D\mathbf{u})-(\mathbf{V}(D\mathbf{u}))_R|^2\,\mathrm{d}z \\
&\leqslant c\vartheta^{-(n+4-p)(j_*+1)}\fint_{Q_R} \varphi_{|(D\mathbf{u})_R|}(|D\mathbf{u}-(D\mathbf{u})_R|)\,\mathrm{d}z \\
&\leqslant C_2\vartheta^{-(n+4-p)(j_*+1)}\varepsilon_2 \leqslant \varepsilon_1\,.
\end{aligned}
\tag{7.14}
$$

Therefore, we can inductively apply Lemma 6.3 with $r=\theta^j R$ for $j=0,1,\dots,m-1$.

Moreover, by the second inequality in (7.11) with $j=m-1$ and the reverse inequality of the second one in (7.10) when $m=m_1$, or (7.12) when $m<m_1$, we have either

$$
\lambda_m \geqslant |(D\mathbf{u})^{\lambda_m}_{2\vartheta^m R}| \geqslant \frac{1}{2^{n+2}}|(D\mathbf{u})^{\lambda_m}_{\vartheta^m R}| \geqslant \frac{1}{2^{n+2}K}\lambda_m \geqslant \frac{1}{K}\lambda_m\,,
$$

or

$$
\lambda_m \geqslant |(D\mathbf{u})^{\lambda_m}_{2\vartheta^m R}| \geqslant \frac{1}{2^{n+2}}|(D\mathbf{u})^{\lambda_m}_{\vartheta^m R}| \geqslant \frac{1}{2^{n+2}K}\lambda_m \geqslant \frac{1}{K_1}\lambda_m\,.
$$

Therefore, applying Lemma 5.3 with $R=\vartheta^m r$, $K_0=\max\{M_1,K\}$ and $\delta_1=\delta_1(K)$ when $m=m_1$, or with $R=\vartheta^m r$, $K_0=\max\{M_1,K_1\}$ and $\delta_1=\delta_1(K_1)$ when $m<m_1$, we can get all the estimates except (7.6).

Thus, we are left to prove (7.6) for $j\leqslant m-1$. Note that the case $j=m$ follows from (7.7). We first observe that if $j\leqslant m-1$, we have (7.11) which, together with (2.14), implies

$$
\fint_{Q^{\lambda_j}_{\vartheta^j R}} \varphi(|D\mathbf{u}-(D\mathbf{u})^{\lambda_j}_{\vartheta^j R}|)\,\mathrm{d}z \leqslant \fint_{Q^{\lambda_j}_{\vartheta^j R}} \varphi(|D\mathbf{u}-(D\mathbf{u})^{\lambda_j}_{\vartheta^j R}|+|(D\mathbf{u})^{\lambda_j}_{\vartheta^j R}|)\,\mathrm{d}z \leqslant c\varphi(\lambda_j)\,,
\tag{7.15}
$$

whenever $1\leqslant j\leqslant m$. Moreover, by the same argument, we also have (7.15) when $j=0$ from (7.1) and (7.2). From this estimate we have that

$$
\begin{aligned}
\fint_{Q^{\lambda_j}_{\vartheta^j R}(z_0)} \varphi(|D\mathbf{u}-\Gamma_0|)\,\mathrm{d}z &\leqslant c\fint_{Q^{\lambda_j}_{\vartheta^j R}} \varphi(|D\mathbf{u}-(D\mathbf{u})^{\lambda_j}_{\vartheta^j R}|)\,\mathrm{d}z + c\varphi(|(D\mathbf{u})^{\lambda_j}_{\vartheta^j r}-\Gamma_0|) \\
&\leqslant c\varphi(\lambda_j)+c\varphi(|(D\mathbf{u})^{\lambda_j}_{\vartheta^j R}-\Gamma_0|)\,.
\end{aligned}
$$

Moreover, by (7.15) and (7.3),

$$
\begin{aligned}
|(D\mathbf{u})^{\lambda_j}_{\vartheta^j R} - \Gamma_0| &\leqslant \sum_{k=j}^{m-1} |(D\mathbf{u})^{\lambda_k}_{\vartheta^k R} - (D\mathbf{u})^{\lambda_{k+1}}_{\vartheta^{k+1} R}| + |(D\mathbf{u})^{\lambda_m}_{\vartheta^m R} - \Gamma_0| \\
&\leqslant \sum_{k=j}^{m-1} \varphi^{-1}\left(\frac{|Q^{\lambda_k}_{\vartheta^k R}|}{|Q^{\lambda_{k+1}}_{\vartheta^{k+1} R}|} \fint_{Q^{\lambda_k}_{\vartheta^k R}} \varphi(|D\mathbf{u} - (D\mathbf{u})^{\lambda_k}_{\vartheta^k R}|)\,\mathrm{d}z \right) + \varphi^{-1}\left(\fint_{Q^{\lambda_m}_{\vartheta^m R}} \varphi(|D\mathbf{u} - \Gamma_0|)\,\mathrm{d}z \right) \\
&\leqslant \sum_{k=j}^{m-1} \varphi^{-1}\left(\vartheta^{-(n+2)} \frac{\varphi(\lambda_{k+1})\lambda_k^2}{\varphi(\lambda_k)\lambda_{k+1}^2} \varphi(\lambda_k) \right) + c\lambda_m \\
&\leqslant \sum_{k=j}^{m-1} \varphi^{-1}\left(\vartheta^{-(n+2+2\alpha_1+q)} \varphi(\lambda_k) \right) + c\lambda_m \\
&\leqslant c \sum_{k=j}^{m} \lambda_k \leqslant c\lambda_j \sum_{k=j}^{m} \vartheta^{\alpha(k-j)} \leqslant c\lambda_j \,.
\end{aligned}
$$

Therefore, combining the preceding two estimates, we obtain (7.6).

Step 3. (Degenerate decay) Suppose $m_1 = \infty$ and (7.12) holds for all $j \in \mathbb{N}$. Then we choose $m = \infty$ and by Lemma 6.3 with $r = \theta^j R$, we have the first inequality in (7.3) and (7.4). We next prove the remaining results, namely, the second condition in (7.5) and (7.6) with $\Gamma_0 = \mathbf{0}$. Observe that applying Lemma 6.3 inductively, we have that for all $j \in \mathbb{N}_0$,

$$
\fint_{Q^{\lambda_j}_{\vartheta^j R}} \varphi_{|(D\mathbf{u})^{\lambda_j}_{\vartheta^j R}|}(|D\mathbf{u} - (D\mathbf{u})^{\lambda_j}_{\vartheta^j R}|)\,\mathrm{d}z \leqslant \varphi(\lambda_j) \quad \text{and} \quad |(D\mathbf{u})^{\lambda_j}_{2\vartheta^j R}| \leqslant \lambda_j \leqslant \vartheta^{\alpha j} \,, \tag{7.16}
$$

hence

$$
\lim_{j\to\infty} (D\mathbf{u})^{\lambda_j}_{\vartheta^j R} = \mathbf{0} \,.
$$

Fix any $r \in (0, R)$. Since

$$
Q^{\lambda_j}_{\vartheta^j R} \subset Q^{\lambda_0}_{\tilde{\sigma}_1^j R} = Q_{\sigma_1^j R} \,, \qquad j \in \mathbb{N}_0
$$

by (7.4), there exists $j \in \mathbb{N}_0$ such that

$$
Q_r \subset Q^{\lambda_j}_{\vartheta^j R} \quad \text{and} \quad Q_r \not\subset Q^{\lambda_{j+1}}_{\vartheta^{j+1} R} \,,
$$

which implies that

$$
\min\left\{ \vartheta^{j+1} R, \frac{\vartheta^{j+1} R}{\sqrt{\varphi(\lambda_{j+1})/\lambda_{j+1}^2}} \right\} < r \leqslant \min\left\{ \vartheta^j R, \frac{\vartheta^j R}{\sqrt{\varphi(\lambda_j)/\lambda_{j+1}^2}} \right\} .
$$

By (7.16) and (7.15) and the inequality $\frac{2n}{n+2} < p < 2$, we have

$$
\begin{aligned}
|(D\mathbf{u})_r| &\leqslant \fint_{Q_r} |D\mathbf{u} - (D\mathbf{u})^{\lambda_j}_{\vartheta^j R}| \,\mathrm{d}z + |(D\mathbf{u})^{\lambda_j}_{\vartheta^j R}| \\
&\leqslant \varphi^{-1}\left(\frac{|Q^{\lambda_j}_{\vartheta^j R}|}{|Q_r|} \fint_{Q^{\lambda_j}_{\vartheta^j R}} \varphi(|D\mathbf{u} - (D\mathbf{u})^{\lambda_j}_{\vartheta^j R}|) \,\mathrm{d}z \right) + |(D\mathbf{u})^{\lambda_j}_{\vartheta^j R}| \\
&\leqslant c\varphi^{-1}\left(\frac{(\vartheta^j R)^{n+2}}{r^{n+2}} \lambda_j^2 \right) + \lambda_j \\
&\leqslant c\varphi^{-1}\left(\vartheta^{-(n+2)} \max\{1, (\varphi(\lambda_{j+1})/\lambda_{j+1}^2)^{\frac{n+2}{2}}\} \lambda_j^2 \right) + \lambda_j \\
&\leqslant c\varphi^{-1}\left(\max\{\lambda_j^2, \lambda_j^2 (\varphi(\lambda_j)/\lambda_{j+1}^2)^{\frac{n+2}{2}}\} \right) + \lambda_j \\
&\leqslant c\varphi^{-1}\left(\max\{\lambda_j^2, \lambda_j^{\frac{n+2}{2}p - n}\} \right) + \lambda_j \leqslant c\lambda_j^{\frac{1}{q}(\frac{n+2}{2}p-n)} .
\end{aligned}
$$

Moreover, since

$$
\frac{r}{R} \geqslant \vartheta^{j+1} \min\left\{1, (\varphi(\lambda_{j+1})/\lambda_{j+1}^2)^{-1/2}\right\} \geqslant \vartheta^{j+1} \lambda_{j+1}^{\frac{2-p}{2}} \geqslant c\lambda_j^{\frac{1}{\alpha_1} + \frac{2-p}{2}}
$$

by (7.3) we have

$$
|(D\mathbf{u})_r| \leqslant c \left(\frac{r}{R}\right)^{\frac{1}{q}(\frac{n+2}{2}p-n)/(\frac{1}{\alpha_1}+\frac{2-p}{2})} \longrightarrow 0 \quad \text{as} \quad r \to 0 ,
$$

hence

$$
\lim_{r\to 0^+} (D\mathbf{u})_r = \mathbf{0} =: \Gamma_0 .
$$

Finally, by (7.15) and the second inequality in (7.16) we have

$$
\fint_{Q^{\lambda_j}_{\vartheta^j R}} \varphi(|D\mathbf{u}|) \,\mathrm{d}z \leqslant c \fint_{Q^{\lambda_j}_{\vartheta^j R}} \varphi(|D\mathbf{u} - (D\mathbf{u})^{\lambda_j}_{\vartheta^j R}|) \,\mathrm{d}z + c\varphi(|(D\mathbf{u})^{\lambda_j}_{\vartheta^j R}|) \leqslant c\varphi(\lambda_j) ,
$$

which proves (7.6) with $m = \infty$. □

Now we prove the partial Hölder continuity result for $D\mathbf{u}$.

Proof of Theorem 1.1. By the parabolic Lebesgue differentiation theorem, we may assume that $D\mathbf{u}(z) = \lim_{r\to 0^+} (D\mathbf{u})_{Q_r(z)}$ if the limit exists. Fix $z_0 \in \Omega_T \setminus (\Sigma_1 \cup \Sigma_2)$, where Σ_1 and Σ_2 are defined in (1.3) and (1.4), respectively. Hence we have

$$
\liminf_{r\to 0^+} \fint_{Q_r(z_0)} \varphi_{|(D\mathbf{u})_{Q_r(z_0)}|}(|D\mathbf{u} - (D\mathbf{u})_{Q_r(z_0)}|) \,\mathrm{d}z = 0 ,
$$

$$
\limsup_{r\to 0^+} |(D\mathbf{u})_{Q_r(z_0)}| < \infty .
$$

From these results and the absolute continuity of the integral, one can find $R > 0$ such that $Q_{2R}(z_0) \Subset \Omega_T$ and for every $\tilde{z} \in Q_R(z_0)$,

$$
\fint_{Q_R(\tilde{z})} \varphi_{|(D\mathbf{u})_{Q_R(\tilde{z})}|}(|D\mathbf{u} - (D\mathbf{u})_{Q_R(\tilde{z})}|) \,\mathrm{d}z \leqslant \varepsilon_2
$$

with ε_2 as in (7.9), and

$$
|(D\mathbf{u})_{Q_R(\tilde{z})}| \leqslant M_0 ,
$$

for some $M_0 < \infty$. Then, in view of Lemma 7.1, we have that for each $\tilde{z} \in Q_R(z_0)$,

$$
\Gamma_{\tilde{z}} := \lim_{r\to 0^+} (D\mathbf{u})_{Q_r(\tilde{z})}
$$

exists and there exist $m_{\tilde{z}} \in \mathbb{N}_0 \cup \{\infty\}$ and positive numbers $\lambda_{\tilde{z},j}$ with $j \in \{0, 1, \ldots, m_{\tilde{z}}\}$ such that

$$\begin{cases} \lambda_{\tilde{z},0} = 1\,, \\ \vartheta^{\alpha_1}\lambda_{\tilde{z},j-1} \leqslant \lambda_{\tilde{z},j} \leqslant \vartheta^{\alpha}\lambda_{\tilde{z},j-1} \quad \text{for } \ 1 \leqslant j < m_{\tilde{z}}\,, \\ \vartheta^{\alpha_1}\lambda_{\tilde{z},m_{\tilde{z}}-1} \leqslant \lambda_{\tilde{z},m_{\tilde{z}}} \leqslant C_1\lambda_{\tilde{z},m_{\tilde{z}}-1} \quad \text{for } \ 0 < m_{\tilde{z}} < \infty\,, \end{cases} \tag{7.17}$$

$$\frac{\lambda_{\tilde{z},m_{\tilde{z}}}}{2K_2} \leqslant |\Gamma_{\tilde{z}}| \leqslant 2K_2\lambda_{\tilde{z},m_{\tilde{z}}}\,,$$

$$\fint_{Q^{\lambda_{\tilde{z},j}}_{\vartheta^j R}(\tilde{z})} \varphi(|D\mathbf{u} - \Gamma_{\tilde{z}}|)\,\mathrm{d}z \leqslant c\varphi(\lambda_{\tilde{z},j}) \qquad \text{for } \ 0 \leqslant j \leqslant m_{\tilde{z}} \ \text{ with } \ j < \infty\,, \tag{7.18}$$

and, if $m_{\tilde{z}} < \infty$,

$$\fint_{Q^{\lambda_{\tilde{z},m_{\tilde{z}}}}_{r}(\tilde{z})} \varphi(|D\mathbf{u} - \Gamma_{\tilde{z}}|)\,\mathrm{d}z \leqslant c\left(\frac{r}{\vartheta^{m_{\tilde{z}}}R}\right)^{\alpha_2}\varphi(\lambda_{\tilde{z},m_{\tilde{z}}}) \quad \text{for all } \ 0 < r \leqslant \vartheta^{m_{\tilde{z}}}R\,. \tag{7.19}$$

We shall prove that the mapping $z \mapsto \Gamma_z = D\mathbf{u}(z)$ from $Q_{R/2}(z_0)$ to $\mathbb{R}^{Nn}$ is Hölder continuous with respect to the parabolic distance in $Q_{R/2}(z_0) \subset \mathbb{R}^{n+1}$ and the Euclidean distance in $\mathbb{R}^{Nn}$. For $\tilde{z} \in Q_{R/2}(z_0)$ and $r \in (0, R)$, we first suppose $Q_r(\tilde{z}) \subset Q^{\lambda_{\tilde{z},m_{\tilde{z}}}}_{\vartheta^{m_{\tilde{z}}}R}(\tilde{z})$. Note that in this case $m_{\tilde{z}} < \infty$. Then define $\rho > 0$ as

$$\rho := \max\left\{1, \sqrt{\varphi(\lambda_{m_{\tilde{z}}})/\lambda^2_{m_{\tilde{z}}}}\right\} r\,,$$

so that $Q_r(\tilde{z}) \subset Q^{\lambda_{\tilde{z},m_{\tilde{z}}}}_{\rho}(\tilde{z})$. Hence by (7.19), the inequality $\frac{2n}{n+2} < p < 2$ and (7.17), we have

$$\begin{aligned}
\fint_{Q_r(\tilde{z})} \varphi(|D\mathbf{u} - \Gamma_{\tilde{z}}|)\,\mathrm{d}z &\leqslant \frac{|Q^{\lambda_{\tilde{z},m_{\tilde{z}}}}_{\rho}(\tilde{z})|}{|Q_r(\tilde{z})|}\fint_{Q^{\lambda_{\tilde{z},m_{\tilde{z}}}}_{\rho}(\tilde{z})} \varphi(|D\mathbf{u} - \Gamma_{\tilde{z}}|)\,\mathrm{d}z \\
&\leqslant c\frac{\rho^{n+2}\lambda^2_{\tilde{z},m_{\tilde{z}}}}{r^{n+2}\varphi(\lambda_{\tilde{z},m_{\tilde{z}}})}\left(\frac{\rho}{\vartheta^{m_{\tilde{z}}}R}\right)^{\alpha_2}\varphi(\lambda_{\tilde{z},m_{\tilde{z}}}) \\
&= c\left(\frac{\rho}{\vartheta^{m_{\tilde{z}}}R}\right)^{\alpha_2}\lambda^2_{\tilde{z},m_{\tilde{z}}}\max\left\{1, (\varphi(\lambda_{\tilde{z},m_{\tilde{z}}})/\lambda^2_{\tilde{z},m_{\tilde{z}}})^{\frac{n+2}{2}}\right\} \\
&\leqslant c\left(\frac{\rho}{\vartheta^{m_{\tilde{z}}}R}\right)^{\alpha_3}\max\{\lambda^2_{\tilde{z},m_{\tilde{z}}}, \lambda^{\frac{n+2}{2}p-n}_{\tilde{z},m_{\tilde{z}}}\} \\
&\leqslant c\left(\frac{\max\left\{1, \sqrt{\varphi(\lambda_{m_{\tilde{z}}})/\lambda^2_{m_{\tilde{z}}}}\right\} r}{\lambda^{1/\alpha}_{\tilde{z},m_{\tilde{z}}}R}\right)^{\alpha_3}\lambda^{\frac{n+2}{2}p-n}_{\tilde{z},m_{\tilde{z}}} \\
&\leqslant c\left(\frac{r}{R}\right)^{\alpha_3}\lambda^{-(\frac{2-p}{2}+\frac{1}{\alpha})\alpha_3+\alpha(\frac{n+2}{2}p-n)}_{\tilde{z},m_{\tilde{z}}} \leqslant c\left(\frac{r}{R}\right)^{\alpha_3}\,,
\end{aligned}$$

where

$$0 < \alpha_3 \leqslant \min\left\{\alpha_2, \frac{1}{2}\frac{\alpha(\frac{n+2}{2}p - n)}{\frac{2-p}{2} + \frac{1}{\alpha}}\right\}\,.$$

We next suppose $Q_r(\tilde{z}) \not\subset Q^{\lambda_{\tilde{z},m_{\tilde{z}}}}_{\vartheta^{m_{\tilde{z}}}R}(\tilde{z})$. Then there exists $0 \leqslant j < m_{\tilde{z}}$ such that

$$Q_r(\tilde{z}) \not\subset Q^{\lambda_{\tilde{z},l+1}}_{\vartheta^{l+1}R}(\tilde{z}) \quad \text{and} \quad Q_r(z_0) \subset Q^{\lambda_{\tilde{z},l}}_{\vartheta^{l}R}(\tilde{z})\,,$$

which implies

$$\vartheta^{j+1}R < \max\left\{1, \sqrt{\varphi(\lambda_{\tilde{z},j+1})/\lambda^2_{\tilde{z},j+1}}\right\} r \quad \text{and} \quad \vartheta^{j}R \geqslant \max\left\{1, \sqrt{\varphi(\lambda_{\tilde{z},j})/\lambda^2_{\tilde{z},j+1}}\right\} r\,. \tag{7.20}$$

Then by (7.18), the first inequality in (7.20), the inequality $\frac{2n}{n+2} < p < 2$ and (7.17), we have

$$\fint_{Q_r(\tilde{z})} \varphi(|D\mathbf{u} - \Gamma_{\tilde{z}}|)\,\mathrm{d}z \leqslant \frac{|Q^{\lambda_{\tilde{z},j}}_{\vartheta^j R}(\tilde{z})|}{|Q_r(\tilde{z})|} \fint_{Q^{\lambda_{\tilde{z},j}}_{\vartheta^j R}(\tilde{z})} \varphi(|D\mathbf{u} - \Gamma_{\tilde{z}}|)\,\mathrm{d}z$$
$$\leqslant c\frac{(\vartheta^j R)^{n+2}\lambda^2_{\tilde{z},j}}{r^{n+2}\varphi(\lambda_{\tilde{z},j})}\varphi(\lambda_{\tilde{z},j})$$
$$\leqslant c\lambda^2_{\tilde{z},j}\max\{1, (\varphi(\lambda_{\tilde{z},j+1})/\lambda^2_{\tilde{z},j+1})^{\frac{n+2}{2}}\} \leqslant \lambda_{\tilde{z},j+1}^{\frac{n+2}{2}p-n}\,.$$

Moreover, by the first inequality in (7.20) again we have

$$\frac{r}{R} > \frac{\vartheta^{j+1}}{\max\left\{1, \sqrt{\varphi(\lambda_{\tilde{z},j+1})/\lambda^2_{\tilde{z},j+1}}\right\}} \geqslant c\lambda_{\tilde{z},j+1}^{\frac{1}{\alpha}}\min\left\{1, \lambda_{\tilde{z},j+1}^{1-\frac{p}{2}}\right\} \geqslant c\lambda_{\tilde{z},j+1}^{1+\frac{1}{\alpha}-\frac{p}{2}}\,.$$

Therefore, combining the above results, we have that

$$\fint_{Q_r(\tilde{z})} \varphi(|D\mathbf{u} - \Gamma_{\tilde{z}}|)\,\mathrm{d}z \leqslant c\left(\frac{r}{R}\right)^{\alpha_3}$$

for some small $\alpha_3 \in (0,1)$.

Now, let $z_1, z_2 \in Q_{R/2}(z_0)$ be any two points with $z_1 \neq z_2$ and $r := d_p(z_1, z_2)$, where

$$d_p(z_1, z_2) := \max\{|x_1 - x_2|, \sqrt{|t_1 - t_2|}\}\,.$$

Set $Q := Q_r(z_1) \cap Q_r(z_2)$. Note that $c(n)|Q_r| \leqslant |Q| \leqslant |Q_r|$. Therefore,

$$|\Gamma_{z_1} - \Gamma_{z_2}| \leqslant \fint_Q |D\mathbf{u} - \Gamma_{z_1}|\,\mathrm{d}z + \fint_Q |D\mathbf{u} - \Gamma_{z_1}|\,\mathrm{d}z$$
$$\leqslant c\fint_{Q_r(z_1)} |D\mathbf{u} - \Gamma_{z_1}|\,\mathrm{d}z + c\fint_{Q_r(z_2)} |D\mathbf{u} - \Gamma_{z_1}|\,\mathrm{d}z$$
$$\leqslant c\varphi^{-1}\left(\left(\frac{r}{R}\right)^{\alpha_3}\right) \leqslant c\left(\frac{r}{R}\right)^{\alpha_3/q} = c\left(\frac{d_p(z_1,z_2)}{R}\right)^{\alpha_3/q}\,,$$

which implies the Hölder continuity of $z \mapsto \Gamma_z = D\mathbf{u}(z)$ with respect to the parabolic metric on $Q_{R/2}(z_0)$ with Hölder exponent α_3/q. Since $z_0 \in \Omega_T \setminus (\Sigma_1 \cup \Sigma_2)$ was an arbitrary point and both Σ_1 and Σ_2 are $\mathcal{L}^{n+1}$-null sets, the proof is complete. □

Acknowledgments

J. Ok was supported by the National Research Foundation of Korea by the Korean Government (NRF-2022R1C1C1004523) and by GNAMPA visiting program. G. Scilla has been supported by the Italian Ministry of Education, University and Research through the MIUR – PRIN project 2017BTM7SN "Variational methods for stationary and evolution problems with singularities and interfaces". The research of B. Stroffolini was supported by PRIN Project 2017TEXA3H "Gradient flows, Optimal Transport and Metric Measure Structures".

References

1. L. Beck, Partial regularity for weak solutions of nonlinear elliptic systems: the subquadratic case, *Manuscripta Math.* **123** (2007), no.4, 453–491.
2. L. Beck, Boundary regularity for elliptic problems with continuous coefficients J. Convex Anal. 16 (2009), no. 1, 287–320.

3. V. Bögelein, Partial regularity for minimizers of discontinuous quasi-convex integrals with degeneracy, *J. Differential Equations* **252** (2012), no. 2, 1052–1100.
4. V. Bögelein, F.Duzaar and G.Mingione, The boundary regularity of non-linear parabolic systems. I, *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **27** (1) (2010) 201–255.
5. V. Bögelein, F.Duzaar and G.Mingione, The boundary regularity of non-linear parabolic systems. II, *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **27** (1) (2010) 145–200.
6. V. Bögelein, F.Duzaar and G.Mingione, The regularity of general parabolic systems with degenerate diffusion, Mem. Amer. Math. Soc. 221(1041), vi+143 pp (2013) .
7. V. Bögelein, F. Duzaar, J. Habermann, and C. Scheven, Partial Hölder continuity for discontinuous elliptic problems with VMO-coefficients, *Proc. Lond. Math. Soc.* (3) **103** (2011), no. 3, 371–404.
8. S. Byun, J. Ok, D. K. Palagachev and L. G. Softova, Parabolic systems with measurable coefficients in weighted Orlicz spaces, *Commun. Contemp. Math.* **18** (2016), no.2, 1550018, 19 pp.
9. S. Campanato, Equazioni paraboliche del secondo ordine e spazi $L^{2,\theta}(\Omega,\delta)$, *Ann. Mat. Pura Appl.* (4) **73** (1966), 55–102.
10. P. Celada and J. Ok, Partial regularity for non-autonomous degenerate quasi-convex functionals with general growth, *Nonlinear Anal.* **194** (2020), 111473, 36 pp.
11. E. Di Benedetto and A. Friedman, *Regularity of solutions of nonlinear degenerate parabolic systems*, *Journal für die Reine Angew. Math* **349** (1984), 83–128.
12. E. Di Benedetto and A. Friedman, Hölder estimates for nonlinear degenerate parabolic systems, *Journal für die Reine Angew. Math* **357** (1985), 1–22, (addendum, ibid. 217–220).
13. E. DiBenedetto, *Degenerate parabolic equations*, Universitext. Springer-Verlag, New York, 1993.
14. L. Diening and F. Ettwein, Fractional estimates for non-differentiable elliptic systems with general growth, *Forum Math.* **20** (3) (2008), 523–556.
15. L. Diening, D. Lengeler, B. Stroffolini and A. Verde, Partial regularity for minimizers of quasi-convex functionals with general growth, *SIAM J. Math. Anal.* **44** (2012), no. 5, 3594–3616.
16. L. Diening, P. Kaplický, and S. Schwarzacher, BMO estimates for the p-Laplacian, *Nonlinear Anal.* **75** (2) (2012), 637–650.
17. L. Diening and C. Kreuzer, Linear convergence of an adaptive finite element method for the p-laplacian equation, *SIAM J. Numer. Anal.* **46** (2008), no. 2, 614–638.
18. L. Diening, T. Scharle, and S. Schwarzacher, Regularity for parabolic systems of Uhlenbeck type with Orlicz growth, *J. Math. Anal. Appl.* **472** (1) (2019), 46–60.
19. L. Diening, S. Schwarzacher, B. Stroffolini, and A. Verde, Parabolic Lipschitz truncation and caloric approximation, *Calc. Var. Partial Differential Equations* **56** (2017), no. 4, Paper No. 120, 27 pp.
20. L. Diening, B. Stroffolini and A. Verde, Everywhere regularity of functionals with φ-growth, *Manuscripta Math.* **129** (4) (2009), 449–481.
21. L. Diening, B. Stroffolini and A. Verde, The φ–harmonic approximation and the regularity of φ–harmonic maps, *J. Differ. Equ.* **253** (7) (2012) 1943–1958.
22. E. De Giorgi, *Frontiere orientate di misura minima*, Seminario di Matematica della Scuola Normale Superiore di Pisa, 1960–1961, 1–56.
23. F. Duzaar, A. Gastel, J.F. Grotowski, Partial regularity for almost minimizers of quasi-convex integrals, SIAM J. Math. Anal. 32 (3) (2000) 665–687.
24. F. Duzaar, J.F. Grotowski, M, Kronz regularity of almost minimizers of quasi-convex variational integrals with subquadratic growth, Ann. Mat. Pura Appl. (4) 184 (4) (2005) 421–448.
25. F. Duzaar and G. Mingione, The p–harmonic approximation and the regularity of p–harmonic maps, *Calc. Var. Partial Differential Equations* **20** (3) (2004) 235–256.
26. F. Duzaar and G. Mingione, Second order parabolic systems, optimal regularity, and singular sets of solutions, *Ann. Institut Poincaré Nonlinéaire* **22** (6) (2005) 705–751.
27. F. Duzaar, G. Mingione and K. Steffen, Parabolic systems with polynomial growth and regularity, *Mem. Amer. Math. Soc.* **214** (2011), no. 1005, x+118 pp.
28. F. Duzaar and K. Steffen, Optimal interior and boundary regularity for almost minimizers to elliptic variational integrals, *J. Reine Angew. Math.* **546** (2002) 73–138.
29. M. Foss, T. Isernia, C. Leone and A. Verde $\mathcal{A}$-caloric approximation and partial regularity for parabolic systems with Orlicz growth, *Calc. Var. Partial Differential Equations* **62** (2023), no. 2, Paper No. 51.
30. M. Foss and G. Mingione, Partial continuity for elliptic problems, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **25** (3) (2008) 471–503.

31. C. Goodrich, G. Scilla and B. Stroffolini, Partial regularity for minimizers of discontinuous quasiconvex integrals with general growth, *Proc. Roy. Soc. Edinburgh Sect. A* **152** (2022), no. 5, 1191–1232.
32. P. Harjulehto and P. Hästö, *Orlicz Spaces and Generalized Orlicz Spaces*, Lecture Notes in Mathematics, vol.2236, Springer, Cham, (2019).
33. P. Hästö and J. Ok, Higher integrability for parabolic systems with Orlicz growth, *J. Differ. Equ.* **300** (2021), 925–948.
34. P. Hästö and J. Ok, Maximal regularity for local minimizers of non-autonomous functionals, *J. Eur. Math. Soc. (JEMS)* **24** (2022), no. 4, 1285–1334.
35. T. Isernia, C. Leone and A. Verde, Partial regularity result for non-autonomous elliptic systems with general growth, *Commun. Pure Appl. Anal.* **20** (2021), no. 12, 4271–4305.
36. D. Kim, S. Ryu and K. Woo, Parabolic equations with unbounded lower-order coefficients in Sobolev spaces with mixed norms, *J. Evol. Equ.* **22** (2022), no.1, Paper No. 9, 40 pp.
37. G. M. Lieberman, Hölder regularity for the gradients of solutions of degenerate parabolic systems, *Ukr. Mat. Visn.* **3** (3) (2006), 352–373.
38. J. L. Lions, Quelques éthodes de résolution des problémes aux limites non linéaires, Dunod; Gauthier-Villars, Paris, 1969.
39. G. Mingione, Regularity of minima: an invitation to the dark side of the calculus of variations, *Appl. Math.* **51** (2006), no. 4, 355-426.
40. J. Ok, G. Scilla and B. Stroffolini, Boundary partial regularity for minimizers of discontinuous quasiconvex integrals with general growth, *Commun. Pure Appl. Anal.* **21** (2022), no. 12, 4173–4214.
41. J. Ok, G. Scilla and B. Stroffolini, Regularity theory for parabolic systems with Uhlenbeck structure, arXiv:2302.05649.
42. C. Scheven, Partial regularity for subquadratic parabolic systems by $\mathcal{A}$-caloric approximation, *Rev. Mat. Iberoam.* **27** (2011), 751–801.
43. B. Stroffolini, Partial regularity results for quasimonotone elliptic systems with general growth, *Z. Anal. Anwend.* **39** (2020), no. 3, 315–347.
44. L. Simon, *Lectures on Regularity and Singularities of Harmonic Maps*, Birkhäuser-Verlag, Basel, 1996.

(Jihoon Ok) Department of Mathematics, Sogang University, 35 Baekbeom-ro, Mapo-gu, Seoul 04107, Republic of Korea.
Email address: `jihoonok@sogang.ac.kr`

(Giovanni Scilla) Department of Mathematics and Applications "R. Caccioppoli", University of Naples Federico II, Via Cintia, Monte S. Angelo, 80126 Naples, Italy
Email address: `giovanni.scilla@unina.it`

(Bianca Stroffolini) Department of Mathematics and Applications "R. Caccioppoli", University of Naples Federico II, Via Cintia, Monte S. Angelo, 80126 Naples, Italy
Email address: `bstroffo@unina.it`