REGULARITY THEORY FOR PARABOLIC SYSTEMS WITH UHLENBECK STRUCTURE

JIHOON OK, GIOVANNI SCILLA, AND BIANCA STROFFOLINI

ABSTRACT. We establish local regularity theory for parabolic systems of Uhlenbeck type with φ-growth. In particular, we prove local boundedness of weak solutions and their gradient, and then local Hölder continuity of the gradients, providing suitable assumptions on the growth function φ. Our approach, being independent of the degeneracy of the system, allows for a unified treatment of both the degenerate and the singular case.

1. INTRODUCTION

We study local regularity theory for the following parabolic φ-Laplace system

$$\mathbf{u}_t - \operatorname{div}\left(\frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|}D\mathbf{u}\right) = 0 \quad \text{in} \quad \Omega_T = \Omega \times (0,T)\,, \tag{1.1}$$

where $\Omega \subset \mathbb{R}^n$ $(n \geqslant 2)$ is open, φ is an Orlicz function verifying suitable growth conditions (see Section 2), $\mathbf{u} = (u^1, \dots, u^N)$ is a vector-valued function of $(x,t) \in \Omega \times (0,T)$, $\mathbf{u}_t$ is the derivative of $\mathbf{u}$ for time variable t, and $D\mathbf{u} = D_x\mathbf{u}$ is the gradient of $\mathbf{u}$ for the spatial variable x. In particular, we prove the local boundedness of $\mathbf{u}$ and $D\mathbf{u}$ and the local Hölder continuity of $D\mathbf{u}$.

A special case of φ in (1.1) is the p-power function, i.e., $\varphi(t) = \frac{1}{p}t^p$ with $1 < p < \infty$. In this case, we have the elliptic and parabolic p-Laplace systems

$$\operatorname{div}\left(|D\mathbf{u}|^{p-2}D\mathbf{u}\right) = 0 \quad \text{in} \quad \Omega \qquad \text{and} \qquad \mathbf{u}_t - \operatorname{div}\left(|D\mathbf{u}|^{p-2}D\mathbf{u}\right) = 0 \quad \text{in} \quad \Omega_T\,.$$

For the elliptic p-Laplace system, Uhlenbeck [38] proved the local Hölder continuity of $D\mathbf{u}$ when $p > 2$. In [38], Uhlenbeck considered the system

$$\operatorname{div}\left(\varrho(|D\mathbf{u}|^2)D\mathbf{u}\right) = 0 \tag{1.2}$$

and assumed that ϱ satisfies a p-growth condition. Note that by setting $\varphi(s) := \int_0^s \tau\varrho(\tau^2)\,\mathrm{d}\tau$ (i.e., $\varrho(s^2) = \varphi'(s)/s$) the previous system is changed to

$$\operatorname{div}\left(\frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|}D\mathbf{u}\right) = 0, \tag{1.3}$$

which is the elliptic counterpart of (1.1) and the Euler-Lagrange system corresponding to the following autonomous and isotropic energy functional

$$\int_\Omega \varphi(|D\mathbf{u}|)\,\mathrm{d}x\,.$$

From this fact, we sometimes say that the system (1.3) or (1.1) has the *Uhlenbeck structure*. It is worth to point out that the radial structure, meaning the dependence through the modulus of the gradient, is the only one that prevent the formation of singularities

2020 *Mathematics Subject Classification.* Primary: 35K40, 46E35; Secondary: 49N60.

Key words and phrases. full regularity, φ-Laplace parabolic system, Uhlenbeck-type parabolic systems, general growth.

(even boundedness of minimizers) and allows to prove everywhere regularity results in the vectorial case, see counterexamples in [34, 36] and also [33, Section 3].

Examples of $\varphi(t)$ satisfying the conditions in the paper are t^p, $t^p\log(1+t)$, $\max\{t^p,t^q\}$, $\min\{t^p,t^q\}$, and so on. A more complicate example, which can fit experimental data, can be found in [2, Section 2.3]. Moreover, the system (1.2) is strongly concerned with stationary, irrotational flows of compressible fluids. Precisely, when $N=1$ hence $\mathbf{u}=u$, if ϱ is the density of an irrotational flow, then the gradient Du of a solution to (1.2) represents the velocity of the flow hence the solution u is called the velocity potential of the flow. At this stage, for an ideal flow (e.g. a polytropic flow) the density function ϱ depends on $|Du|^2$. We refer to [3, 18] for applications of the above system to stationary, irrotational flow of compressible fluids.

After the pioneering work of Uhlenbeck, Tolksdorf [37] obtained $C^{1,\alpha}$-regularity results for more general elliptic systems with p-growth when $1<p<\infty$. We also refer to [19, 1, 21] for everywhere $C^{1,\alpha}$-regularity results for elliptic systems with p-growth. For the parabolic p-Laplace system, DiBenedetto and Friedman [11, 12] (see also the monograph [10]) proved Hölder continuity of $D\mathbf{u}$ when $\frac{2n}{n+2}<p<\infty$ and we refer to [5, 39, 6, 8, 9, 4] for further related results for parabolic p-Laplace systems.

For a general function φ, Lieberman studied regularity theory for elliptic equations (i.e., $N=1$) with φ-growth, and around the same time Marcellini [30, 31] had considered elliptic equations with general (p,q)-growth. Full $C^{1,\alpha}$-regularity for the elliptic φ-Laplace system (1.3) was established by Marcellini and Papi [32] and by Diening, Stroffolini and Verde [17]. Marcellini and Papi proved Lipschitz regularity for local minimizers of functionals with growth conditions general enough to embrace linear and exponential ones. The conclusion then follows using the C^1-regularity of the operator, with the help of classical results. The second result, instead, is reminiscent of the Uhlenbeck proof: a nonlinear quantity $\varphi(|D\mathbf{u}|)$ is shown to be a subsolution of an elliptic equation. In addition, the authors were able to prove an excess decay estimate for $\mathbf{V}_p(D\mathbf{u})$ (see Section 2 for the definition of $\mathbf{V}_p$) which implies the Hölder continuity of $\mathbf{V}_p(D\mathbf{u})$ and hence of $D\mathbf{u}$.

On the other hand, $C^{1,\alpha}$-regularity for the parabolic φ-Laplace system (1.1) has remained an open problem. There have been partial developments in this direction. Lieberman [27] proved that if $D\mathbf{u}$ is bounded, then $D\mathbf{u}$ is Hölder continuous. Hence the local boundedness of $D\mathbf{u}$ is missing. Diening, Scharle and Schwarzacher [15] obtained the local boundedness of $D\mathbf{u}$ for (1.1) under an additional integrability condition on $D\mathbf{u}$ which is unnatural in the singular case, that is, $p<2$ in (2.6). Moreover, Isernia [26] obtained the local boundedness of $\mathbf{u}$ for (1.1).

We note that in [27] the approximation of the parabolic system (1.1) with nondegenerate systems is omitted and the weak solution is assumed to be twice differentiable with respect to the x variable. In fact, one has to consider approximate nondegenerate parabolic systems (e.g. (4.1)), and obtain uniform regularity estimates by differentiating these systems. At this stage, the twice differentiability of weak solutions of these systems with respect to the x variable is needed. However, the proof of the twice differentiability for parabolic system with φ-growth is unclear and not an easy generalization of the one for the parabolic p-Laplace system. Even in the elliptic case, a more delicate analysis is required, see [13, Section 4]. In addition, Baroni and Lindfors [2] obtained the Hölder and Lipschitz continuity of solutions to Cauchy-Dirichlet problems for parabolic equations $(N=1)$ with φ-growth, see also [29] for similar results for parabolic obstacle problems with φ-growth. For more regularity results for the parabolic system with φ-growth we refer to [7, 16, 23, 24, 25, 35].

In this paper, we establish full $C^{1,\alpha}$-regularity for the parabolic φ-Laplace system (1.1) by filling all the gaps in previous results. Let us state the main result.

1.1. **Setting of the problem and main result.** Suppose the function $\varphi : [0,\infty) \to [0,\infty)$ is an N-function satisfying Assumption 2.1. A function $\mathbf{u} = (u^1, u^2, \dots, u^N) \in C_{\rm loc}(0,T; L^2_{\rm loc}(\Omega,\mathbb{R}^N)) \cap L^{\varphi}_{\rm loc}(0,T; W^{1,\varphi}_{\rm loc}(\Omega,\mathbb{R}^N))$ is said to be a (local) *weak solution* to (1.1) if it satisfies the following weak form of (1.1):

$$-\int_{\Omega_T} \mathbf{u}\cdot\zeta_t \,\mathrm{d}z + \int_{\Omega_T} \frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|} D\mathbf{u} : D\zeta \,\mathrm{d}z = 0 \quad \text{for all} \;\; \zeta \in C^\infty_{\rm c}(\Omega_T,\mathbb{R}^N)\,,$$

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 (see e.g. the proof of Theorem 1 in [26]), one can see that the weak solution $\mathbf{u}$ to (1.1) also satisfies for every $0 < t_1 < t_2 < T$,

$$\int_{\Omega'} \mathbf{u}\cdot\zeta(x,t)\,dx \Big|_{t=t_1}^{t=t_2} + \int_{\Omega'}\int_{t_1}^{t_2} \left[-\mathbf{u}\cdot\zeta_t + \frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|} D\mathbf{u} : D\zeta \right] \mathrm{d}t\,\mathrm{d}x = 0 \tag{1.4}$$

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 $\Omega' \Subset \Omega$. We note that weak solution $\mathbf{u}$ is not weakly differentiable with respect to t. Therefore, we cannot take a test function ζ involving the weak solution directly. This technical obstacle can be overcome by using approximation via Steklov average, see [10, I. 3-(i) and II. Proposition 3.1], which is by now a standard approximation argument. Hence we will assume that $\mathbf{u}$ is differentiable, and consider test functions involving the weak solution without specific comment.

We state $C^{1,\alpha}$-regularity, which is the maximal regularity, for the weak solution $\mathbf{u}$ of the system (1.1). This result follows directly by combining Corollary 5.3 and Theorem 6.1.

Theorem 1.1. *Suppose $\varphi \in C^1([0,\infty)) \cap C^2((0,\infty))$ satisfies Assumption 2.3 with*

$$p > \frac{2n}{n+2}\,, \tag{1.5}$$

and let $\mathbf{u}$ be a weak solution to the parabolic system (1.1). Then $D\mathbf{u}$ is locally Hölder continuous. Moreover, there exist $\alpha \in (0,1)$ and $c > 0$ depending on n, N, p, q, γ_1, c such that for every $Q_{2R}(z_0) \Subset \Omega_T$ and every $0 < r \leqslant R$,

$$\operatorname*{osc}_{Q_r(z_0)} D\mathbf{u} \leqslant c\lambda \left(\max\left\{ \varphi''(\lambda)^{\frac12}, \varphi''(\lambda)^{-\frac12} \right\} \frac{r}{R} \right)^{\alpha}$$

where

$$\lambda := \left(\fint_{Q_{2R}(z_0)} \varphi(|D\mathbf{u}|)\,\mathrm{d}z + 1 \right)^{\frac{2}{(n+2)p-2n}}.$$

We remark that the condition (1.5) is essential in the regularity theory even for the parabolic p-Laplace system, without any additional integrability condition on the solution $\mathbf{u}$, see [12, 8] and also [10].

We shall introduce the strategy of our paper. We prove sequentially local L^∞-regularity of the weak solution $\mathbf{u}$ to (1.1), the local L^∞-regularity and C^α-regularity of $D\mathbf{u}$, by providing essentially sharp conditions on φ. As for the local boundedness of $\mathbf{u}$ (Theorem 3.1), we apply the Moser iteration to a suitable test function. Next, using the parabolic embedding result in Lemma 2.9, we reach the conclusion. Once the L^∞-regularity result is achieved, we prove twice differentiability of weak solutions to approximate nondegenerate systems in Lemma 4.1 by using the difference quotients and a Giaquinta-Modica type covering argument. Note that the boundedness of $\mathbf{u}$ plays an important role in the proof of Lemma 4.1 since the constant p in (2.6) can be less than 2. Then by differentiating the approximate nondegenerate system and applying Moser iteration again, we obtain L^∞ estimate for $D\mathbf{u}$ in Theorem 5.2 and Corollary 5.3. Finally, we revisit the results with the proofs in [27], and prove Theorem 1.1.

2. Preliminaries

2.1. **Notation.** We write $\mathbf{u} = (u^\alpha) = (u^1, \dots, u^N) \in \mathbb{R}^N$ and $\mathbf{Q} = (Q_i^\alpha) \in \mathbb{R}^{N\times n} = \mathbb{R}^{Nn}$ where $1 \leqslant i \leqslant n$ and $1 \leqslant \alpha \leqslant N$. For $z = (x,t) \in \mathbb{R}^n \times \mathbb{R}$, we introduce the parabolic cylinder

$$Q_r(z) := B_r(x) \times (t - r^2, t]\,, \tag{2.1}$$

where $B_r(x)$ denotes the open ball in $\mathbb{R}^n$ with center x and radius r. The symbol $\partial_{\mathrm{p}} Q_r(z)$ denotes the usual parabolic boundary of $Q_r(z)$.

Let $f : E \to [0,\infty)$ with $E \subset \mathbb{R}$. f is called almost increasing (resp. almost decreasing) if there is $L \geqslant 1$ such that $f(s) \leqslant Lf(t)$ for all $s,t \in E$ with $s \leqslant t$ (resp. $t \leqslant s$). In particular, if we can choose $L = 1$, then f is simply called increasing (resp. decreasing).

By χ^* we denote the Sobolev conjugate exponent of χ; i.e., $\chi^* := \frac{n\chi}{n-\chi}$ for $\chi < n$, while we agree that $\chi^* := 2\chi$ if $\chi \geqslant n$.

The notation $f \sim g$ means that there exists constant $c \geqslant 1$ such that $\frac{1}{c} f \leqslant g \leqslant cf$. We will use the Einstein summation convention, that is, we will omit the summation symbol for indexes that appear twice, see e.g. (2.18) and the next inequality.

2.2. **Orlicz functions.** 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 the above condition, then constants c may depend on $\varphi(1)$. Moreover, we assume that φ satisfies the following growth conditions:

Assumption 2.1. *$\varphi : [0,\infty) \to [0,\infty)$ is an N-function, and there are $1 < p \leqslant q$ such that $\frac{\varphi(t)}{t^p}$ is almost increasing and $\frac{\varphi(t)}{t^q}$ is almost decreasing for $t \in (0,\infty)$ with constant $L \geqslant 1$.*

The almost decreasing and increasing conditions in Assumption 2.1 are equivalent to the Δ_2 and ∇_2 conditions for φ, respectively. Compared with the Δ_2 type conditions, the benefit of the almost increasing/decreasing condition is that we can directly see the lower and upper bounds of an exponent factor of φ. In particular, we will prove the boundedness of the weak solution to (1.1) under the above assumption where the lower bound p will play a crucial role. We also remark that Assumption 2.1 with $L = 1$ is equivalent to the following inequality

$$1 < p \leqslant \frac{t\varphi'(t)}{\varphi(t)} \leqslant q\,, \quad t > 0\,. \tag{2.2}$$

For any $t > 0$ and $0 < c < 1 < C$ there holds

$$c^q\varphi(t) \leqslant \varphi(ct) \leqslant c^p\varphi(t) \ \text{ and } \ C^p\varphi(t) \leqslant \varphi(Ct) \leqslant C^q\varphi(t)\,. \tag{2.3}$$

The conjugate function of φ is defined as

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

From the definition, the following Young's inequality

$$st \leqslant \varphi(t) + \varphi^*(s) , \quad s,t \geqslant 0 , \tag{2.4}$$

holds true. Since the exact value of φ^* is not always explicitly computable, the estimate

$$\varphi^*\left(\frac{\varphi(t)}{t}\right) \sim \varphi^*(\varphi'(t)) \sim \varphi(t) \tag{2.5}$$

will often be useful in computations (see [22, Theorem 2.4.10]). In fact, the above relation holds true since Assumption 2.1 guarantees the Δ_2 condition for both φ and φ^*, and relevant constants depends on p, q and L.

For higher order regularity results we will consider stronger assumptions.

Assumption 2.2. $\varphi : [0,\infty) \to [0,\infty)$ *is an N-function and satisfies*

(1) $\varphi \in C^1([0,\infty)) \cap C^2((0,\infty))$
(2) *There exist* $1 < p \leqslant q$ *such that*

$$0 < p - 1 \leqslant \frac{t\varphi''(t)}{\varphi'(t)} \leqslant q - 1 , \quad t > 0 . \tag{2.6}$$

Note that Assumption 2.2 implies Assumption 2.1 with the same p and q and with the constant $L = 1$, hence we have (2.2).

We notice that in the above two assumptions we can replace p and q by $\min\{p, 2-\varepsilon\}$ and $\max\{q, 2+\varepsilon\}$ for $\varepsilon > 0$, respectively. Therefore, without loss of generality, we always assume that p and q satisfy

$$1 < p < 2 < q . \tag{2.7}$$

The next assumption is adding an Hölder type continuity on the Hessian of $\varphi(|\mathbf{Q}|)$ for $\mathbf{Q} \in \mathbb{R}^{Nn}$, denoted by $D^2_{\mathbf{Q}}\varphi(|\mathbf{Q}|)$.

Assumption 2.3. $\varphi : [0,\infty) \to [0,\infty)$ *is an N-function and satisfies Assumption* (2.2). *Furthermore, there exist positive constants* γ_1 *and* c_h *such that for every* $\mathbf{Q}, \mathbf{P} \in \mathbb{R}^{Nn}$ *with* $|\mathbf{Q} - \mathbf{P}| \leqslant \frac{1}{2}|\mathbf{Q}|$,

$$|D^2_{\mathbf{Q}}\varphi(|\mathbf{Q}|) - D^2_{\mathbf{Q}}\varphi(|\mathbf{P}|)| \leqslant c_h \left(\frac{|\mathbf{Q} - \mathbf{P}|}{|\mathbf{Q}|}\right)^{\gamma_1} \varphi''(|\mathbf{Q}|). \tag{2.8}$$

Note that if $\varphi(t) = t^p$ with $1 < p < \infty$, then it satisfies Assumption 2.3. Similar assumptions were used for proving the $C^{1,\alpha}$-regularity for minimizers of functionals with general growth in [17].

If φ satisfies Assumption 2.1, 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, the parabolic space $L^\varphi(t_1, t_2; W^{1,\varphi}(\Omega, \mathbb{R}^N))$

denotes the set of all functions $f \in L^1(t_1,t_2;W^{1,1}(\Omega,\mathbb{R}^N))$ such that $f(\cdot,t) \in W^{1,\varphi}(\Omega,\mathbb{R}^N)$ for a.e. $t \in (0,T)$ and

$$\int_{t_1}^{t_2}\int_\Omega \varphi(|Df(x,t)|)\,\mathrm{d}x\,\mathrm{d}t < \infty.$$

2.3. **Shifted N-functions and related operators.** The following definitions and results about shifted N-functions can be found in [13, 17].

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 Assumption 2.1 or 2.2 or 2.3, then φ_a also satisfies Assumption 2.1 or 2.2 or 2.3 uniformly in $a \geqslant 0$ with the same p and q. We then recall useful inequalities for shifted N-function φ_a in [13] and [17].

Lemma 2.4. [17] *Let φ satisfy Assumption 2.2. Then we have*

$$\varphi_a(t) \sim \varphi_a'(t)\,t\,; \tag{2.9}$$

$$\varphi_a(t) \sim \varphi''(a+t)t^2 \sim \frac{\varphi(a+t)}{(a+t)^2}t^2 \sim \frac{\varphi'(a+t)}{a+t}t^2\,, \tag{2.10}$$

$$\varphi(a+t) \sim [\varphi_a(t)+\varphi(a)]\,, \tag{2.11}$$

which hold uniformly with respect to $a \geqslant 0$.

Lemma 2.5. [13, Lemma 32] *Let φ satisfy Assumption 2.1. Then for all $\delta > 0$ there exists $c_\delta > 0$ depending only on p, q, L and δ, such that for all $t,u,a \geqslant 0$*

$$\begin{cases} tu \leqslant \delta\varphi(t) + c_\delta\varphi^*(u)\,, \\ t\varphi'(u) + u\varphi'(t) \leqslant \delta\varphi(t) + c_\delta\varphi(u)\,, \\ tu \leqslant \delta\varphi_a(t) + c_\delta\varphi_a^*(u)\,, \\ t\varphi_a'(u) + u\varphi_a'(t) \leqslant \delta\varphi_a(t) + c_\delta\varphi_a(u)\,. \end{cases} \tag{2.12}$$

Lemma 2.6. [13, Lemmas 24 and 29]. *Let φ satisfy Assumption 2.2.*

(1) *Uniformly in $s,t \in \mathbb{R}^n$ with $|s|+|t| > 0$*

$$\varphi''(|s|+|t|)|s-t| \sim \varphi_{|s|}'(|s-t|)\,, \tag{2.13}$$

(2) *There exists $c = c(p,q) > 0$ such that for all $s_1,s_2,t \in \mathbb{R}^n$*

$$\varphi_{|s_2|}'(|s_1-s_2|) \lesssim \varphi_{|t|}'(|s_1-t|) + \varphi_{|t|}'(|s_2-t|)\,, \tag{2.14}$$

where the hidden constants above depend only on p and q.

The following lemma (see [14, Corollary 26]) deals with the *change of shift* for N-functions.

Lemma 2.7 (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 $a,b \in \mathbb{R}^n$ and $t \geqslant 0$*

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

We next define vector valued functions $\mathbf{A},\mathbf{V} : \mathbb{R}^{Nn} \to \mathbb{R}^{Nn}$ by

$$\mathbf{A}(\mathbf{Q}) := \frac{\varphi'(|\mathbf{Q}|)}{|\mathbf{Q}|}\mathbf{Q} = D_{\mathbf{Q}}[\varphi(|\mathbf{Q}|)] \quad \text{and} \quad \mathbf{V}(\mathbf{Q}) := \sqrt{\frac{\varphi'(|\mathbf{Q}|)}{|\mathbf{Q}|}}\mathbf{Q}.$$

In particular, for $1 < p < \infty$, we denote by $\mathbf{V}_p(\mathbf{Q})$ the function $\mathbf{V}$ associated to $\varphi(t) = \frac{1}{p}t^p$; i.e., $\mathbf{V}_p(\mathbf{Q}) := |\mathbf{Q}|^{\frac{p-2}{2}}\mathbf{Q}$. With shifted N-function φ_a, we define accordingly

$$\mathbf{A}^a(\mathbf{Q}) := \frac{\varphi_a'(|\mathbf{Q}|)}{|\mathbf{Q}|}\mathbf{Q}, \quad \mathbf{V}^a(\mathbf{Q}) := \sqrt{\frac{\varphi_a'(|\mathbf{Q}|)}{|\mathbf{Q}|}}\mathbf{Q} \quad \text{and} \quad \mathbf{V}_p^a(\mathbf{Q}) := (a + |\mathbf{Q}|)^{\frac{p-2}{2}}\,\mathbf{Q}\,. \tag{2.16}$$

We further suppose that φ satisfies Assumption 2.2. Denote

$$A_{ij}^{\alpha\beta}(\mathbf{Q}) := \frac{\partial \mathbf{A}(\mathbf{Q})_j^\beta}{\partial Q_i^\alpha} = \frac{\varphi'(|\mathbf{Q}|)}{|\mathbf{Q}|}\left\{\delta_{ij}\delta^{\alpha\beta} + \left(\frac{\varphi''(|\mathbf{Q}|)|\mathbf{Q}|}{\varphi'(|\mathbf{Q}|)} - 1\right)\frac{Q_i^\alpha Q_j^\beta}{|\mathbf{Q}|^2}\right\}, \tag{2.17}$$

where $1 \leqslant i, j \leqslant n$ and $1 \leqslant \alpha, \beta \leqslant N$. Here, $\delta^{\alpha\beta}$ and δ_{ij} are the Kronecker symbols. Note that $D^2_{\mathbf{Q}}\varphi(|\mathbf{Q}|) = (A_{ij}^{\alpha\beta}(\mathbf{Q}))$. Then we see that

$$\min\{p-1, 1\}\frac{\varphi'(|\mathbf{Q}|)}{|\mathbf{Q}|}|\boldsymbol{\omega}|^2 \leqslant A_{ij}^{\alpha\beta}(\mathbf{Q})\omega_i^\alpha\omega_j^\beta \leqslant \max\{q-1, 1\}\frac{\varphi'(|\mathbf{Q}|)}{|\mathbf{Q}|}|\boldsymbol{\omega}|^2 \tag{2.18}$$

for all $\mathbf{Q}, \boldsymbol{\omega} \in \mathbb{R}^{nN}$. Moreover, since

$$[\mathbf{A}(\mathbf{P}) - \mathbf{A}(\mathbf{Q})]_j^\beta = \int_0^1 \frac{\partial}{\partial\tau}[\mathbf{A}(\tau\mathbf{P} + (1-\tau)\mathbf{Q})]_j^\beta \,\mathrm{d}\tau = \int_0^1 A_{ij}^{\alpha\beta}(\tau\mathbf{P} + (1-\tau)\mathbf{Q})(\mathbf{P}-\mathbf{Q})_i^\alpha \,\mathrm{d}\tau,$$

using the above results and [13, Lemma 20], we have that

$$\begin{aligned}(\mathbf{A}(\mathbf{P}) - \mathbf{A}(\mathbf{Q})) : (\mathbf{P} - \mathbf{Q}) &\geqslant \frac{1}{c}\left(\int_0^1 \frac{\varphi'(|\tau\mathbf{P} + (1-\tau)\mathbf{Q}|)}{|\tau\mathbf{P} + (1-\tau)\mathbf{Q}|}\,\mathrm{d}\tau\right)|\mathbf{P} - \mathbf{Q}|^2 \\ &\geqslant \frac{1}{c}\frac{\varphi'(|\mathbf{P}| + |\mathbf{Q}|)}{|\mathbf{P}| + |\mathbf{Q}|}|\mathbf{P} - \mathbf{Q}|^2,\end{aligned}$$

and

$$|\mathbf{A}(\mathbf{P}) - \mathbf{A}(\mathbf{Q})| \leqslant c\left(\int_0^1 \frac{\varphi'(|\tau\mathbf{P} + (1-\tau)\mathbf{Q}|)}{|\tau\mathbf{P} + (1-\tau)\mathbf{Q}|}\,\mathrm{d}\tau\right)|\mathbf{P} - \mathbf{Q}| \leqslant c\frac{\varphi'(|\mathbf{P}| + |\mathbf{Q}|)}{|\mathbf{P}| + |\mathbf{Q}|}|\mathbf{P} - \mathbf{Q}|. \tag{2.19}$$

Moreover, we have that

$$(\mathbf{A}(\mathbf{P}) - \mathbf{A}(\mathbf{Q})) : (\mathbf{P} - \mathbf{Q}) \sim \varphi_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|) \sim |\mathbf{V}(\mathbf{P}) - \mathbf{V}(\mathbf{Q})|^2 \tag{2.20}$$

and

$$|\mathbf{A}(\mathbf{P}) - \mathbf{A}(\mathbf{Q})| \sim \varphi'_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|)\,. \tag{2.21}$$

(see [15, Lemma 3.1]). We note that the estimates in above still hold for φ_a and the related operators $\mathbf{A}^a$ and $\mathbf{V}^a$.

From [15, Lemma 3.3], it follows that

$$|\mathbf{A}^a(\mathbf{Q}) - \mathbf{A}(\mathbf{Q})| \leqslant \varphi'_{|\mathbf{Q}|}(a)\,.$$

Applying the same argument to the N-function $\bar{\varphi}_{|\mathbf{Q}|}$ defined by $\bar{\varphi}'_{|\mathbf{Q}|}(t) := \sqrt{\varphi'_{|\mathbf{Q}|}(t)t}$, we obtain

$$|\mathbf{V}^a(\mathbf{Q}) - \mathbf{V}(\mathbf{Q})|^2 \leqslant c|\bar{\varphi}'_{|\mathbf{Q}|}(a)|^2 \sim \varphi_{|\mathbf{Q}|}(a)\,. \tag{2.22}$$

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

2.4. **Embedding.** We recall a Gagliardo-Nirenberg type inequality for Orlicz functions in [23, Lemma 2.13]. 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 2.8. *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*

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

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

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

Applying the above lemma we can obtain a parabolic embedding result for an Orlicz function φ.

Lemma 2.9. *Let $m > 0$. Suppose that φ satisfies Assumption 2.1 with*

$$\max\{1, \tfrac{mn}{n+m}\} < p \leqslant q.$$

There exists $\theta = \theta(n,m,p,q) \in (0,1)$ and $c = c(n,m,p,q,L) > 0$ such that for every

$$f \in L^\infty(t_1,t_2; L^m(B_r)) \cap L^\varphi(t_1,t_2; W^{1,\varphi}(B_r))$$

we have

$$\fint_{B_r\times[t_1,t_2]} \varphi\big(\big|\tfrac{f}{r}\big|\big)^{\frac{n+m}{n}}\,dz \leqslant c\left(\fint_{B_r}\left[\varphi(|Df|) + \varphi\big(\big|\tfrac{f}{r}\big|\big)\right]\mathrm{d}x\right)^{\frac{\theta(n+m)}{n}}$$
$$\times\, \varphi\left(\left(\operatorname*{ess\,sup}_{t\in[t_1,t_2]}\fint_{B_r}\big|\tfrac{f}{r}\big|^m\,dx\right)^{\frac{1}{m}}\right)^{\frac{(1-\theta)(n+m)}{n}}.$$

Proof. Without loss of generality, we can assume that $p < n$. If $p \geqslant n$, it is enough to consider any $\tilde{p} \in (\frac{mn}{n+m}, n)$ instead of p. Note that by (2.2) the function $\varphi^{\frac{1}{p}}$ is a weak Φ-function, and the function $\frac{\varphi^{1/p}}{t^{q/p}}$ is decreasing. Therefore, applying Lemma 2.8 with $\psi = \varphi^{\frac{1}{p}}$ and $(\gamma, p, q_1, q_2) = (p\frac{n+m}{n}, p, \frac{q}{p}, m)$, we have that for a.e. $t \in [t_1,t_2]$,

$$\fint_{B_r}\varphi\big(\big|\tfrac{f(t)}{r}\big|\big)^{\frac{n+m}{n}}\,dx \leqslant c\left(\fint_{B_r}\left[\varphi(|Df(t)|) + \varphi\big(\big|\tfrac{f(t)}{r}\big|\big)\right]dx\right)^{\frac{\theta(n+m)}{n}}$$
$$\times\,\varphi\left(\left(\fint_{B_r}\big|\tfrac{f(t)}{r}\big|^m\,dx\right)^{\frac{1}{m}}\right)^{\frac{(1-\theta)(n+m)}{n}},$$

where $f(t) = f(x,t)$ and θ satisfies

$$\frac{n}{n+m} = \frac{\theta(n-p)}{n} + \frac{(1-\theta)q}{m} \quad\iff\quad \theta = \frac{n(nm-nq-mq)}{(n+m)(nm-nq-mp)}.$$

Note that $\theta \in (0,1)$ by the assumption $\frac{mn}{n+m} < p \leqslant q$, which yields $\frac{n-p}{n} < \frac{n}{n+m} < \frac{q}{m}$, and that $\frac{\theta(n+m)}{n} \in (0,1]$. Finally integrating for t in $[t_1,t_2]$ and using Hölder's inequality when $p < q$ (i.e., $\frac{\theta(n+m)}{n} < 1$), we obtain the desired estimate. □

Remark 2.10. In the above definition, if $m = 2$, we see that

$$\varphi(|f|)^{1+\frac{2}{n}} \in L^1(B_r \times [t_1, t_2]).$$

Note that this implies

$$f \in L^{\frac{p(n+2)}{n}}(B_r \times [t_1, t_2])\,,$$

where $\frac{p(n+2)}{n} > 2$ if $p > \frac{2n}{n+2}$.

3. Local boundedness

We first prove that any weak solution $\mathbf{u}$ to (1.1) is locally bounded by using the Moser iteration technique (for similar arguments, cfr. [9, Theorem 2], and [26], where the superquadratic case is addressed). The key points in our approach are the introduction of the function ψ in (3.2) and use of the embedding result Lemma 2.9 for Orlicz functions in the parabolic setting, that measures the superquadratic or subquadratic character of the function φ.

Theorem 3.1. *Let φ satisfy Assumption 2.1 with (1.5) and let $\mathbf{u}$ be a weak solution to (1.1). Then $\mathbf{u} \in L^\infty_{\mathrm{loc}}(\Omega_T, \mathbb{R}^N)$. Moreover, there exists $c \geqslant 1$ depending on n, N, p, q and L such that for any $Q_{2r} \Subset \Omega_T$,*

$$\sup_{Q_r} \varphi(|\tfrac{\mathbf{u}}{r}|) \leqslant c \left(\fint_{Q_{2r}} \psi(|\tfrac{\mathbf{u}}{r}|)\varphi(|\tfrac{\mathbf{u}}{r}|)^{\chi_0}\, \mathrm{d}z \right)^{\frac{1}{\chi_0}} + c, \tag{3.1}$$

where

$$\psi(s) := \max\{s^2, \varphi(s)\}, \quad s > 0, \tag{3.2}$$

and $\chi_0 > 0$ is determined in (3.12) below.

Proof. We divide the proof into three steps.

Step 1. (Caccioppoli type inequality) Let $0 < r < 1$, $z_0 = (x_0, t_0) \in \mathbb{R}^n \times \mathbb{R}$ and $Q_{2r}(z_0) \Subset \Omega_T$ be defined as in (2.1). Let $\rho_1 = s_1 r$ and $\rho_2 = s_2 r$ with $1 \leqslant s_1 < s_2 \leqslant 2$, $\xi \in C_0^\infty(B_{\rho_2}(x_0))$ be such that

$$0 \leqslant \xi \leqslant 1\,, \quad \xi \equiv 1 \;\text{ in }\; B_{\rho_1}(x_0) \;\text{ and }\; |D\xi| \leqslant \frac{2}{(s_2 - s_1) r}\,, \tag{3.3}$$

and let $\eta \in C^\infty(\mathbb{R})$ be such that

$$0 \leqslant \eta \leqslant 1\,, \quad \eta \equiv 0 \;\text{ in } (-\infty, -\rho_2^2]\,, \;\eta \equiv 1 \;\text{ in }\; [-\rho_1^2, \infty)\,, \; 0 \leqslant \eta_t \leqslant \frac{2}{(s_2^2 - s_1^2) r^2}\,. \tag{3.4}$$

With fixed $\chi > 0$ to be determined later, we take

$$\zeta = \varphi(|\tfrac{\mathbf{u}}{r}|)^\chi \eta^2 \xi^q \mathbf{u}$$

as a test function in (1.4) and integrate by parts. Then for every $\tau \in (-\rho_1^2, 0]$ we have

$$0 = \underbrace{\int_{-\rho_2^2}^{\tau} \int_{B_{\rho_2}(x_0)} \mathbf{u}_t \cdot \mathbf{u}\varphi(|\tfrac{\mathbf{u}}{r}|)^\chi \eta^2 \xi^q \,\mathrm{d}x\,\mathrm{d}t}_{=:I_1}$$

$$+ \underbrace{\int_{-\rho_2^2}^{\tau} \int_{B_{\rho_2}(x_0)} \frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|} D\mathbf{u} : D[\mathbf{u}\varphi(|\tfrac{\mathbf{u}}{r}|)^\chi \eta^2 \xi^q] \,\mathrm{d}x\,\mathrm{d}t}_{=:I_2}\,.$$

Now, we estimate both the terms I_1 and I_2 separately. For what concerns I_1, setting

$$\Phi_\chi(s) := \int_0^{s^2} \varphi(\sqrt{\sigma})^\chi \,\mathrm{d}\sigma \leqslant s^2\varphi(s)^\chi, \quad s \geqslant 0, \tag{3.5}$$

we obtain

$$\begin{aligned} I_1 &= r^2 \int_{-\rho_2^2}^{\tau} \int_{B_{\rho_2}(x_0)} \left[\tfrac{1}{2}\partial_t(\Phi_\chi(|\tfrac{\mathbf{u}}{r}|)\eta^2) - \Phi_\chi(|\tfrac{\mathbf{u}}{r}|)\eta\eta_t\right] \xi^q \,\mathrm{d}x\,\mathrm{d}t \\ &\geqslant \frac{r^2}{2}\int_{B_{\rho_2}(x_0)} \Phi_\chi(|\tfrac{\mathbf{u}(x,\tau)}{r}|)\eta(\tau)^2\xi^q \,\mathrm{d}x - \frac{2}{s_2^2 - s_1^2}\int_{Q_{\rho_2}(z_0)} \Phi_\chi(|\tfrac{\mathbf{u}}{r}|)\,\mathrm{d}z\,. \end{aligned} \tag{3.6}$$

Integrating by parts and taking into account (2.2),

$$\Phi_\chi(s) = s^2\varphi(s)^\chi - \chi\int_0^{s^2} \sigma\varphi(\sqrt{\sigma})^{\chi-1}\varphi'(\sqrt{\sigma})\frac{1}{2\sqrt{\sigma}}\,\mathrm{d}\sigma \geqslant s^2\varphi(s)^\chi - \frac{q\chi}{2}\Phi_\chi(s)$$

hence

$$\frac{2}{2+q\chi}s^2\varphi(s)^\chi \leqslant \Phi_\chi(s) \leqslant s^2\varphi(s)^\chi. \tag{3.7}$$

As for I_2, we have

$$\begin{aligned} I_2 &= \int_{-\rho_2^2}^{\tau}\int_{B_{\rho_2}(x_0)} \frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|}\left(|D\mathbf{u}|^2\varphi(|\tfrac{\mathbf{u}}{r}|)^\chi + \chi r^2\varphi(|\tfrac{\mathbf{u}}{r}|)^{\chi-1}\frac{\varphi'(|\frac{\mathbf{u}}{r}|)}{|\frac{\mathbf{u}}{r}|}\frac{|D[|\frac{\mathbf{u}}{r}|^2]|^2}{4}\right)\eta^2\xi^q\,\mathrm{d}x\,\mathrm{d}t \\ &\quad + \int_{-\rho_2^2}^{\tau}\int_{B_{\rho_2}(x_0)} \frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|}\left[D\mathbf{u} : (D\xi\otimes\mathbf{u})\right]\varphi(|\tfrac{\mathbf{u}}{r}|)^\chi\eta^2 q\xi^{q-1}\,\mathrm{d}x\,\mathrm{d}t \\ &\geqslant \frac{1}{c_1}\int_{-\rho_2^2}^{\tau}\int_{B_{\rho_2}(x_0)}\left(\varphi(|D\mathbf{u}|)\varphi(|\tfrac{\mathbf{u}}{r}|)^\chi + \chi r^2\frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|}\frac{\varphi(|\frac{\mathbf{u}}{r}|)^\chi}{|\frac{\mathbf{u}}{r}|^2}|D[|\tfrac{\mathbf{u}}{r}|^2]|^2\right)\eta^2\xi^q\,\mathrm{d}x\,\mathrm{d}t \\ &\quad - \int_{-\rho_2^2}^{\tau}\int_{B_{\rho_2}(x_0)}\left[\frac{1}{2c_1}\varphi^*(\varphi'(|D\mathbf{u}|)\xi^{q-1}) + c\varphi(|D\xi||\mathbf{u}|)\right]\varphi(|\tfrac{\mathbf{u}}{r}|)^\chi\eta^2\,\mathrm{d}x\,\mathrm{d}t \\ &\geqslant \frac{1}{2c_1}\int_{-\rho_2^2}^{\tau}\int_{B_{\rho_2}(x_0)}\left(\varphi(|D\mathbf{u}|)\varphi(|\tfrac{\mathbf{u}}{r}|)^\chi + \chi r^2\frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|}\frac{\varphi(|\frac{\mathbf{u}}{r}|)^\chi}{|\frac{\mathbf{u}}{r}|^2}|D[|\tfrac{\mathbf{u}}{r}|^2]|^2\right)\eta^2\xi^q\,\mathrm{d}x\,\mathrm{d}t \\ &\quad - \frac{c}{(s_2-s_1)^q}\int_{Q_{\rho_2}(z_0)}\varphi(|\tfrac{\mathbf{u}}{r}|)^{\chi+1}\,\mathrm{d}z\,, \end{aligned} \tag{3.8}$$

where in the first inequality we have applied Young's inequality to the second integral, while in the last one we used the inequality $|D\xi| \leqslant \frac{c}{(s_2-s_1)r}$ and (2.3).

Combining the above estimates (3.6) and (3.8), we have that for every $\tau \in (-\rho_1^2, 0]$

$$\begin{aligned} &r^2\int_{B_{\rho_2}(x_0)}\Phi_\chi(|\tfrac{\mathbf{u}(x,\tau)}{r}|)\xi^q\,\mathrm{d}x \\ &+ \int_{-\rho_2^2}^{\tau}\int_{B_{\rho_2}(x_0)}\left(\varphi(|D\mathbf{u}|)\varphi(|\tfrac{\mathbf{u}}{r}|)^\chi + \chi r^2\frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|}\frac{\varphi(|\frac{\mathbf{u}}{r}|)^\chi}{|\frac{\mathbf{u}}{r}|^2}|D[|\tfrac{\mathbf{u}}{r}|^2]|^2\right)\eta^2\xi^q\,\mathrm{d}x\,\mathrm{d}t \\ &\leqslant \frac{c}{s_2^2-s_1^2}\int_{Q_{\rho_2}(z_0)}\Phi_\chi(|\tfrac{\mathbf{u}}{r}|)\,\mathrm{d}z + \frac{c}{(s_2-s_1)^q}\int_{Q_{\rho_2}(z_0)}\varphi(|\tfrac{\mathbf{u}}{r}|)^{\chi+1}\,\mathrm{d}z\,. \end{aligned}$$

Therefore, taking into account (3.7), neglecting a non-negative term in the left hand side and recalling (3.5), we have

$$
\begin{aligned}
&\sup_{-\rho_1^2<\tau<0}\int_{B_{\rho_2}(x_0)}|\mathbf{u}(x,\tau)|^2\varphi(|\tfrac{\mathbf{u}(x,\tau)}{r}|)^{\chi}\xi^q\,\mathrm{d}x+\int_{Q_{\rho_2}(z_0)}\varphi(|D\mathbf{u}|)\varphi(|\tfrac{\mathbf{u}}{r}|)^{\chi}\eta^2\xi^q\,\mathrm{d}z\\
&\leqslant\frac{c(1+\chi)}{(s_2-s_1)^2}\int_{Q_{\rho_2}(z_0)}|\tfrac{\mathbf{u}}{r}|^2\varphi(|\tfrac{\mathbf{u}}{r}|)^{\chi}\,\mathrm{d}z+\frac{c(1+\chi)}{(s_2-s_1)^q}\int_{Q_{\rho_2}(z_0)}\varphi(|\tfrac{\mathbf{u}}{r}|)^{\chi+1}\,\mathrm{d}z\,.
\end{aligned}\tag{3.9}
$$

Step 2. (Sobolev inequality) Set

$$
G(z):=r\tilde{\varphi}(|\tfrac{\mathbf{u}}{r}|)^{\chi+1}\eta^{\frac{2}{p_0}}\xi^{\frac{q}{p_0}},\quad \text{where }\ \tilde{\varphi}(s):=\varphi(s)^{\frac{1}{p_0}}\ \text{ and }\ p_0:=\frac{2n}{n+2}.
$$

Then

$$
DG=(1+\chi)\tilde{\varphi}(|\tfrac{\mathbf{u}}{r}|)^{\chi}\tilde{\varphi}'(|\tfrac{\mathbf{u}}{r}|)\tfrac{\mathbf{u}D\mathbf{u}}{|\mathbf{u}|}\eta^{\frac{2}{p_0}}\xi^{\frac{q}{p_0}}+r\tfrac{q}{p_0}\tilde{\varphi}(|\tfrac{\mathbf{u}}{r}|)^{\chi+1}\eta^{\frac{2}{p_0}}\xi^{\frac{q}{p_0}-1}D\xi\,,
$$

Now we apply Young's inequality (2.4) to the N-function $\tilde{\varphi}$, with $t=|D\mathbf{u}|$ and $s=\tilde{\varphi}'(|\frac{\mathbf{u}}{r}|)$, together with (2.5) to get

$$
\begin{aligned}
|DG|&\leqslant(1+\chi)\tilde{\varphi}(|\tfrac{\mathbf{u}}{r}|)^{\chi}\tilde{\varphi}'(|\tfrac{\mathbf{u}}{r}|)|D\mathbf{u}|\eta^{\frac{2}{p_0}}\xi^{\frac{q}{p_0}}+cr\tilde{\varphi}(|\tfrac{\mathbf{u}}{r}|)^{\chi+1}|D\xi|\\
&\leqslant c(1+\chi)\tilde{\varphi}(|\tfrac{\mathbf{u}}{r}|)^{\chi}\tilde{\varphi}(|D\mathbf{u}|)\eta^{\frac{2}{p_0}}\xi^{\frac{q}{p_0}}+c\left(1+\chi+\frac{1}{s_2-s_1}\right)\tilde{\varphi}(|\tfrac{\mathbf{u}}{r}|)^{\chi+1}.
\end{aligned}
$$

Therefore, combining with (3.9) and recalling the definition of $\tilde{\varphi}$, we have

$$
\begin{aligned}
&\sup_{-\rho_1^2<\tau<0}\int_{B_{\rho_2}(x_0)}|\mathbf{u}(x,\tau)|^2\varphi(|\tfrac{\mathbf{u}(x,\tau)}{r}|)^{\chi}\xi^q\,\mathrm{d}x+\int_{Q_{\rho_2}(z_0)}|DG|^{p_0}\,\mathrm{d}z\\
&\leqslant\frac{c(1+\chi)^2}{(s_2-s_1)^2}\int_{Q_{\rho_2}(z_0)}|\tfrac{\mathbf{u}}{r}|^2\varphi(|\tfrac{\mathbf{u}}{r}|)^{\chi}\,\mathrm{d}z+\frac{c(1+\chi)^{p_0+1}}{(s_2-s_1)^q}\int_{Q_{\rho_2}(z_0)}\varphi(|\tfrac{\mathbf{u}}{r}|)^{\chi+1}\,\mathrm{d}z\\
&\leqslant\frac{c(1+\chi)^{p_0+1}}{(s_2-s_1)^q}\int_{Q_{\rho_2}(z_0)}\left(|\tfrac{\mathbf{u}}{r}|^2+\varphi(|\tfrac{\mathbf{u}}{r}|)\right)\varphi(|\tfrac{\mathbf{u}}{r}|)^{\chi}\,\mathrm{d}z.
\end{aligned}\tag{3.10}
$$

Now, applying Hölder's inequality, the Sobolev inequality to the function $G\in W_0^{1,p_0}(B_r)$ and using (3.10), we can write

$$
\begin{aligned}
&\int_{Q_{\rho_1}(z_0)}|\tfrac{\mathbf{u}}{r}|^{\frac{2p_0}{n}}\varphi(|\tfrac{\mathbf{u}}{r}|)^{1+\chi+\frac{\chi p_0}{n}}\,\frac{\mathrm{d}z}{r^{n+2}}\\
&\leqslant\frac{1}{r^{n+2+\frac{2p_0}{n}}}\int_{-\rho_1^2}^{0}\left(\int_{B_{\rho_1}(x_0)}|\mathbf{u}|^2\varphi(|\tfrac{\mathbf{u}}{r}|)^{\chi}\,\mathrm{d}x\right)^{\frac{p_0}{n}}\left(\int_{B_{\rho_1}(x_0)}\tilde{\varphi}(|\tfrac{\mathbf{u}}{r}|)^{\frac{(1+\chi)np_0}{n-p_0}}\,\mathrm{d}x\right)^{\frac{n-p_0}{n}}\mathrm{d}t\\
&\leqslant\frac{1}{r^{(n+2)(1+\frac{p_0}{n})}}\left(\sup_{-\rho_1^2<\tau<0}\int_{B_{\rho_1}(x_0)}|\mathbf{u}(x,\tau)|^2\varphi(|\tfrac{\mathbf{u}(x,\tau)}{r}|)^{\chi}\,\mathrm{d}x\right)^{\frac{p_0}{n}}\int_{-\rho_1^2}^{0}\left(\int_{B_{\rho_2}(x_0)}|G|^{p_0^*}\,\mathrm{d}x\right)^{\frac{p_0}{p_0^*}}\mathrm{d}t\\
&\leqslant\frac{c}{r^{(n+2)(1+\frac{p_0}{n})}}\left(\sup_{-\rho_1^2<\tau<0}\int_{B_{\rho_1}(x_0)}|\mathbf{u}(x,\tau)|^2\varphi(|\tfrac{\mathbf{u}(x,\tau)}{r}|)^{\chi}\,\mathrm{d}x\right)^{\frac{p_0}{n}}\int_{Q_{\rho_2}(z_0)}|DG|^{p_0}\,\mathrm{d}z\\
&\leqslant c\left\{\frac{(1+\chi)^{p_0+1}}{(s_2-s_1)^q}\int_{Q_{\rho_2}(z_0)}\left(|\tfrac{\mathbf{u}}{r}|^2+\varphi(|\tfrac{\mathbf{u}}{r}|)\right)\varphi(|\tfrac{\mathbf{u}}{r}|)^{\chi}\,\frac{\mathrm{d}z}{r^{n+2}}\right\}^{1+\frac{p_0}{n}}.
\end{aligned}
$$

Then, since $s^2 \leqslant c\varphi(s)s^{\frac{2p_0}{n}} = c\varphi(s)s^{\frac{4}{n+2}}$ for $s \geqslant 1$ by $p > \frac{2n}{n+2}$, recalling (3.2), we have

$$
\begin{aligned}
\int_{Q_{\rho_1}(z_0)} \psi(|\tfrac{\mathbf{u}}{r}|)\varphi(|\tfrac{\mathbf{u}}{r}|)^{\chi(1+\frac{p_0}{n})} \frac{\mathrm{d}z}{|Q_{2r}|} &\leqslant c\int_{Q_{\rho_1}(z_0)} \varphi(|\tfrac{\mathbf{u}}{r}|)^{1+\chi+\frac{\chi p_0}{n}} |\tfrac{\mathbf{u}}{r}|^{\frac{2p_0}{n}} \frac{\mathrm{d}z}{|Q_{2r}|} + c \\
&\leqslant c\left\{ \frac{(1+\chi)^{p_0+1}}{(s_2-s_1)^q} \int_{Q_{\rho_2}(z_0)} \left[\psi(|\tfrac{\mathbf{u}}{r}|)\varphi(|\tfrac{\mathbf{u}}{r}|)^{\chi}+1\right] \frac{\mathrm{d}z}{|Q_{2r}|} \right\}^{1+\frac{p_0}{n}} .
\end{aligned}
\tag{3.11}
$$

Step 3. (Iteration) We first notice that by applying the Gagliardo-Nirenberg type interpolation inequality to φ with $p > \frac{2n}{n+2}$ provided by Lemma 2.9 and Remark 2.10, we have

$$
\int_{Q_{2r}} \psi(|\mathbf{u}|)\varphi(|\mathbf{u}|)^{\chi_0}\,\mathrm{d}z < \infty
$$

where

$$
\chi_0 := \min\left\{ \left(\frac{p(n+2)}{n} - 2\right)\frac{1}{q}, \frac{2}{n} \right\} > 0 . \tag{3.12}
$$

For $m = 0, 1, 2, \dots$, set $\chi_m = \chi_0\theta^m$ and

$$
J_m := \int_{Q_{r_m}} \left[\psi(|\tfrac{\mathbf{u}}{r}|)\varphi(|\tfrac{\mathbf{u}}{r}|)^{\chi_m}+1\right] \frac{\mathrm{d}z}{|Q_{2r}|}, \quad \text{where } \theta := 1+\frac{p_0}{n} \text{ and } r_m := r(1+2^{-m}) .
$$

Then, we can iterate (3.11) and write

$$
J_m \leqslant c2^{q\theta m}(1+\chi_0\theta^{m-1})^{(p_0+1)\theta} J_{m-1}^{\theta} \leqslant c_0^m J_{m-1}^{\theta}, \quad m = 1, 2, \dots,
$$

for some $c_0 \geqslant 1$ depending on n, N, p, q and L. Hence, for $m \geqslant 2$,

$$
J_m \leqslant c_0^m\left(c_0^{m-1}J_{m-2}^{\theta}\right)^{\theta} \leqslant c_0^{m+(m-1)\theta} J_{m-2}^{\theta^2} \leqslant \cdots \leqslant c_0^{\sum_{k=1}^m (m-k+1)\theta^{k-1}} J_0^{\theta^m} \leqslant (c_1 J_0)^{\theta^m}
$$

for some large $c_1 \geqslant 1$ depending on n, N, p, q and L. Consequently, setting

$$
\mathrm{d}\mu(z) := \psi(|\tfrac{\mathbf{u}(z)}{r}|)\frac{\mathrm{d}z}{|Q_{2r}|},
$$

we have

$$
\begin{aligned}
\|\varphi(|\tfrac{\mathbf{u}}{r}|)\|_{L^\infty(Q_r)} \leqslant \|\varphi(|\tfrac{\mathbf{u}}{r}|)\|_{L^\infty(Q_r;\mathrm{d}\mu)} &= \lim_{m\to\infty}\left(\int_{Q_r} \varphi(|\tfrac{\mathbf{u}}{r}|)^{\chi_0\theta^m}\,\mathrm{d}\mu\right)^{\frac{1}{\chi_0\theta^m}} \\
&\leqslant \limsup_{m\to\infty} (J_m)^{\frac{1}{\chi_0\theta^m}} \leqslant (c_1J_0)^{\frac{1}{\chi_0}} .
\end{aligned}
$$

This implies the estimate (3.1) and the proof is concluded. □

4. Approximating problems and Second order differentiability

Let $\mathbf{u}$ be a weak solution to (1.1). Then for $Q_R \Subset \Omega_T$ and a sufficiently small $\varepsilon \in (0,1)$ we consider the following non-degenerate parabolic system with Cauchy-Dirichlet boundary condition:

$$
\begin{cases}
(\mathbf{u}_\varepsilon)_t - \operatorname{div}\left(\dfrac{\varphi_\varepsilon'(|D\mathbf{u}_\varepsilon|)}{|D\mathbf{u}_\varepsilon|}D\mathbf{u}_\varepsilon\right) = 0 & \text{in } Q_R , \\
\mathbf{u}_\varepsilon = \mathbf{u} & \text{on } \partial_{\mathrm{p}} Q_R .
\end{cases}
\tag{4.1}
$$

where φ_ε is the shifted N-function with $a = \varepsilon$. We will show that the system (1.1) can be approximated by (4.1), in the sense that if $\mathbf{u}_\varepsilon$ is the weak solution to (4.1), then $D\mathbf{u}_\varepsilon$ converges to $D\mathbf{u}$ in $L^p(Q_R)$ (see Lemma 4.3).

We first prove second differentiability in the spatial variable x for each weak solution to the following non-degenerated problem without boundary condition:

$$\mathbf{w}_t - \operatorname{div}\left(\frac{\varphi'_\varepsilon(|D\mathbf{w}|)}{|D\mathbf{w}|} D\mathbf{w}\right) = 0 \quad \text{in} \quad \Omega_T\,. \tag{4.2}$$

In order to do that, we fix some notation. For a (vector-valued) function f, we introduce the notation

$$\Delta_{k,s} f(x,t) := f\left(x + se_k, t\right) - f(x,t)\,,$$

where $s \in \mathbb{R}$ and e_k with $k \in \{1, 2, \dots, n\}$ is a standard unit vector in $\mathbb{R}^n$. Moreover, we define $T_{k,s} : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ by

$$T_{k,s}(x,t) := (x + se_k, t), \qquad (x,t) \in \mathbb{R}^{n+1}. \tag{4.3}$$

Then we have the following result (cfr. [8, Theorem 6] where analogous estimates are devised for parabolic systems with p-growth, and [13, Theorem 11], [17, Lemma 5.7] for analogous arguments for elliptic systems with φ-growth).

Lemma 4.1. *Let φ satisfy Assumption 2.2 with (1.5) and (2.7), and let $\mathbf{u}_\varepsilon$ be a weak solution to (4.2) with $\varepsilon > 0$. Then*

(i) $\mathbf{V}^\varepsilon(D\mathbf{u}_\varepsilon) \in L^2_{\mathrm{loc}}(0,T; W^{1,2}_{\mathrm{loc}}(\Omega;\mathbb{R}^{Nn}))$;
(ii) $D\mathbf{u}_\varepsilon \in L^p_{\mathrm{loc}}(0,T; W^{1,p}_{\mathrm{loc}}(\Omega;\mathbb{R}^{Nn})) \cap L^\infty_{\mathrm{loc}}(0,T; L^2_{\mathrm{loc}}(\Omega;\mathbb{R}^{Nn}))$;
(iii) if, in addition, $D\mathbf{u}_\varepsilon \in L^\infty_{\mathrm{loc}}(\Omega_T;\mathbb{R}^{Nn})$, *then* $D\mathbf{u}_\varepsilon \in L^2_{\mathrm{loc}}(0,T; W^{1,2}_{\mathrm{loc}}(\Omega;\mathbb{R}^{Nn}))$ *and* $\varphi_\varepsilon(|D\mathbf{u}_\varepsilon|) \in L^2_{\mathrm{loc}}(0,T; W^{1,2}_{\mathrm{loc}}(\Omega))$.

Proof. In order to enlighten the notation, we will denote $\mathbf{u}_\varepsilon$ by $\mathbf{w}$. Note that in view of Theorem 3.1, $\mathbf{w} \in L^\infty_{\mathrm{loc}}(\Omega_T)$. Fix any $Q_{25r} = Q_{25r}(x_0,t_0) \Subset \Omega_T$ with $r \in (0,1)$, and set $\lambda := \sup_{Q_{2r}} \mathbf{w}$. Let $s, h \in \mathbb{R}$ be such that $0 < s \leqslant h < r/2$ or $-r/2 < h \leqslant s < 0$. Then $\mathbf{w}$ satisfies

$$(\Delta_{k,s}\mathbf{w})_t - \operatorname{div}\left(\Delta_{k,s}(\mathbf{A}^\varepsilon(D\mathbf{w}))\right) = 0 \text{ in } Q_{3r/2} \tag{4.4}$$

in the weak sense, where $\mathbf{A}^\varepsilon$ is defined as in (2.16) with $a = \varepsilon$.

Let $\zeta \in C^\infty\left(Q_{3r/2}\right)$ be such that $0 \leqslant \zeta \leqslant 1$, $\zeta = 0$ in $Q_{2r} \setminus Q_{3r/2}$, $\zeta \equiv 1$ in Q_r, and $r^{-1}|D\zeta| + |D^2\zeta| + |\zeta_t| \leqslant \frac{c(n)}{r^2}$. Then, testing (4.4) with the function $(\Delta_{k,s}\mathbf{w})\zeta^q$, for every $t_0 \in (-r^2, 0]$ we obtain

$$\begin{aligned} 0 &= \int_{-4r^2}^{t_0}\int_{B_{2r}} (\Delta_{k,s}\mathbf{w})_t \cdot (\Delta_{k,s}\mathbf{w})\zeta^q \,\mathrm{d}x\,\mathrm{d}t \\ &\quad + \int_{-4r^2}^{t_0}\int_{B_{2r}} \Delta_{k,s}(\mathbf{A}^\varepsilon(D\mathbf{w})) : (\zeta^q D[\Delta_{k,s}\mathbf{w}] + q\zeta^{q-1}\Delta_{k,s}\mathbf{w} \otimes D\zeta)\,\mathrm{d}x\,\mathrm{d}t. \end{aligned}$$

Note that an integration by parts with respect to the time variable gives

$$\begin{aligned} \int_{-4r^2}^{t_0}\int_{B_{2r}} (\Delta_{k,s}\mathbf{w})_t \cdot (\Delta_{k,s}\mathbf{w})\zeta^q \,\mathrm{d}x\,\mathrm{d}t &= \frac{1}{2}\int_{B_{2r}} |\Delta_{k,s}\mathbf{w}(x,t_0)|^2\zeta^q\,\mathrm{d}x \\ &\quad - \frac{q}{2}\int_{-4r^2}^{t_0}\int_{B_{2r}} |\Delta_{k,s}\mathbf{w}|^2\zeta^{q-1}\zeta_t\,\mathrm{d}x\,\mathrm{d}t. \end{aligned}$$

Hence, for every $t_0 \in [-r^2, 0]$ we have that

$$\begin{aligned}&\int_{B_r} |\Delta_{k,s}\mathbf{w}(x,t_0)|^2 \,\mathrm{d}x \\ &\qquad + 2\int_{-4r^2}^{t_0}\int_{B_{2r}} \Delta_{k,s}(\mathbf{A}^\varepsilon(D\mathbf{w})) : \big(\Delta_{k,s}(D\mathbf{w})\zeta^q + q\zeta^{q-1}\Delta_{k,s}\mathbf{w}\otimes D\zeta\big)\,\mathrm{d}x\,\mathrm{d}t \\ &\leqslant q\int_{-4r^2}^{t_0}\int_{B_{2r}} |\Delta_{k,s}\mathbf{w}|^2\zeta^{q-1}\zeta_t\,\mathrm{d}x\,\mathrm{d}t.\end{aligned}$$

We first observe that by (2.20),

$$\begin{aligned}&\Delta_{k,s}(\mathbf{A}^\varepsilon(D\mathbf{w})) : \Delta_{k,s}D\mathbf{w} \\ &= \big[\mathbf{A}^\varepsilon(D\mathbf{w}(x+se_k,t)) - \mathbf{A}^\varepsilon(D\mathbf{w}(x,t))\big] : \big[D\mathbf{w}(x+se_k,t) - D\mathbf{w}(x,t)\big] \\ &\geqslant \frac{1}{c}|\Delta_{k,s}\mathbf{V}^\varepsilon(D\mathbf{w})|^2.\end{aligned}$$

Now, by (2.19) and (2.13),

$$\begin{aligned}|\Delta_{k,s}\mathbf{A}^\varepsilon(D\mathbf{w})| &= |\mathbf{A}^\varepsilon(D\mathbf{w}(x+se_k,t)) - \mathbf{A}^\varepsilon(D\mathbf{w}(x,t))| \\ &\leqslant c\frac{\varphi_\varepsilon'(|D\mathbf{w}(x+se_k,t)| + |D\mathbf{w}(x,t)|)}{|D\mathbf{w}(x+se_k,t)| + |D\mathbf{w}(x,t)|}|\Delta_{k,s}D\mathbf{w}(x,t)| \\ &\leqslant c\varphi_{\varepsilon+|D\mathbf{w}|}'(|\Delta_{k,s}D\mathbf{w}(x,t)|)\,,\end{aligned}$$

whence

$$\begin{aligned}&|\Delta_{k,s}(\mathbf{A}^\varepsilon(D\mathbf{w})) : \Delta_{k,s}\mathbf{w}\otimes D\zeta\zeta^{q-1}| \\ &\leqslant c\fint_0^s \underbrace{\varphi_{\varepsilon+|D\mathbf{w}|}'(|\Delta_{k,s}D\mathbf{w}|)|s||D\mathbf{w}(x+\tau e_k,t)||D\zeta|}_{=:J}\,\mathrm{d}\tau\,.\end{aligned}$$

Note that we shall write $\fint_0^s = \frac{1}{s}\int_0^s$ even if $s \in \mathbb{R}$ is negative. Recalling (4.3), we have

$$\mathbf{w}(x+\tau e_k,t) = T_{k,\tau}\circ\mathbf{w}(x,t)\,.$$

Now, from (2.14), Young's inequality in (2.12) and (2.21), recalling $h \in \mathbb{R}$ fixed above, for any sufficiently small $\delta > 0$ we have that

$$\begin{aligned}J &\leqslant c\left(\varphi_{\varepsilon+|D[T_{k,\tau}\circ\mathbf{w}]|}'(|\Delta_{k,s-\tau}D[T_{k,\tau}\circ\mathbf{w}]|) + \varphi_{\varepsilon+|D[T_{k,\tau}\circ\mathbf{w}]|}'(|\Delta_{k,\tau}D\mathbf{w}|)\right)|s||D[T_{k,\tau}\circ\mathbf{w}]||D\zeta| \\ &\leqslant \delta\left(\frac{|s|}{|h|}\right)^{\frac{q}{q-1}}\left\{\varphi_{\varepsilon+|D[T_{k,\tau}\circ\mathbf{w}]|}(|\Delta_{k,s-\tau}D[T_{k,\tau}\circ\mathbf{w}]|) + \varphi_{\varepsilon+|D[T_{k,\tau}\circ\mathbf{w}]|}(|\Delta_{k,\tau}D\mathbf{w}|)\right\} \\ &\quad + c_\delta\varphi_{\varepsilon+|D[T_{k,\tau}\circ\mathbf{w}]|}\left(|h||D[T_{k,\tau}\circ\mathbf{w}]||D\zeta|\right) \\ &\leqslant \delta\frac{|s|}{|h|}\Big(|\Delta_{k,s-\tau}\mathbf{V}^\varepsilon(D[T_{k,\tau}\circ\mathbf{w}])|^2 + |\Delta_{k,\tau}\mathbf{V}^\varepsilon(D\mathbf{w})|^2\Big) + c_\delta\frac{h^2}{r^2}\varphi_\varepsilon\left(|D[T_{k,\tau}\circ\mathbf{w}]|\right)\,,\end{aligned}$$

where we have used the facts that $0 < |s| \leqslant |h| < r/2$, $|D\zeta| \leqslant \frac{c}{r} \leqslant \frac{c}{|h|}$ and for $\varepsilon > 0$ and $s \in (0,1]$

$$\varphi_a^*(st) \leqslant s^{\frac{q}{q-1}}\varphi_a^*(t)$$

and

$$\varphi_{\varepsilon+t}(st) \sim \frac{\varphi'(\varepsilon+t+st)}{\varepsilon+t+st}(st)^2 \sim s^2\frac{\varphi'(\varepsilon+t)}{\varepsilon+t}t^2 \sim s^2\varphi_\varepsilon(t), \quad s \in [0,1].$$

Note that applying Fubini's Theorem and the change of variables $y = x + \tau e_k$ and $\tilde{\tau} = s - \tau$, we get

$$
\begin{aligned}
&\int_{Q_{3r/2}} \fint_0^s |\Delta_{k,s-\tau}\mathbf{V}^\varepsilon(D[T_{k,\tau}\circ \mathbf{w}])|^2 \,\mathrm{d}\tau\,\mathrm{d}x \\
&= \fint_0^s \int_{Q_{3r/2}} |\mathbf{V}^\varepsilon(D\mathbf{w}(x+\tau e_k + (s-\tau)e_k, t)) - \mathbf{V}^\varepsilon(D\mathbf{w}(x+\tau e_k, t))|^2 \,\mathrm{d}x\,\mathrm{d}\tau \\
&\leqslant \fint_0^s \int_{Q_{2r}} |\mathbf{V}^\varepsilon(D\mathbf{w}(y+\tilde{\tau} e_k, t)) - \mathbf{V}^\varepsilon(D\mathbf{w}(y, t))|^2 \,\mathrm{d}y\,\mathrm{d}\tilde{\tau}\,.
\end{aligned}
$$

Therefore, we have
(4.5)

$$
\begin{aligned}
&\sup_{t_0\in[-r^2,0]} \int_{B_r} [\Delta_{k,s}\mathbf{w}\,(x,t_0)]^2\,\mathrm{d}x + \int_{Q_r} |\Delta_{k,s}\mathbf{V}^\varepsilon(D\mathbf{w})|^2\,\mathrm{d}z \\
&\leqslant c\delta\frac{|s|}{|h|}\fint_0^s \int_{Q_{2r}} |\Delta_{k,\tau}\mathbf{V}^\varepsilon(D\mathbf{w})|^2\,\mathrm{d}x\,\mathrm{d}\tau + c(\delta)\frac{h^2}{r^2}\int_{Q_{2r}} \varphi_\varepsilon(|D\mathbf{w}|)\,\mathrm{d}z + \frac{c}{r^2}\int_{Q_{2r}} |\Delta_{k,s}\mathbf{w}|^2\zeta^{q-1}\,\mathrm{d}z.
\end{aligned}
$$

Further, we estimate the final term in the right hand side of the inequality above using the integration by parts:

$$
\begin{aligned}
&\int_{Q_{2r}} |\Delta_{k,s}\mathbf{w}|^2\zeta^{q-1}\,\mathrm{d}z = \int_0^s \int_{Q_{2r}} \mathbf{w}_{x_k}(x+\tau e_k, t)\cdot \Delta_{k,s}\mathbf{w}\zeta^{q-1}\,\mathrm{d}z\,\mathrm{d}\tau \\
&= -\int_0^s \int_{Q_{2r}} \mathbf{w}(x+\tau e_k, t)\cdot (\Delta_{k,s}\mathbf{w}_{x_k}\zeta^{q-1} + (q-1)\Delta_{k,s}\mathbf{w}\zeta^{q-2}\zeta_{x_k})\,\mathrm{d}z\,\mathrm{d}\tau \\
&\leqslant |s|\lambda \int_{Q_{3r/2}} |\Delta_{k,s}D\mathbf{w}| + |\Delta_{k,s}\mathbf{w}||D\zeta|\,\mathrm{d}z \\
&\leqslant |s|\lambda \int_{Q_{3r/2}} (\varepsilon + |D\mathbf{w}(x+se_k,t)| + |D\mathbf{w}(x,t)|)^{\frac{p-2}{2}+\frac{2-p}{2}} |\Delta_{k,s}D\mathbf{w}|\,\mathrm{d}z \\
&\qquad + \frac{c\lambda s^2}{r}\int_{Q_{3r/2}} |s^{-1}\Delta_{k,s}\mathbf{w}|\,\mathrm{d}z\,.
\end{aligned}
$$

Now applying Young's inequality and using facts that $\frac{\varphi(t)}{t^{p-1}}$ is increasing and $t^p \leqslant c(\varphi(t)+1)$, we have that for any $\tilde{\delta}\in(0,1)$, in order to reabsorb some terms to the left-hand side,

$$\begin{aligned}
\int_{Q_{2r}} |\Delta_{k,s}\mathbf{w}|^2\zeta^{q-1}\,\mathrm{d}z &\leqslant \tilde{\delta}\frac{\varphi'(\varepsilon)r^2}{\varepsilon^{p-1}}\int_{Q_{3r/2}}(\varepsilon+|D\mathbf{w}(x+se_k,t)|+|D\mathbf{w}(x,t)|)^{p-2}|\Delta_{k,s}D\mathbf{w}|^2\,\mathrm{d}z \\
&\quad + c\tilde{\delta}^{-1}\frac{\varepsilon^{p-1}}{r^2\varphi'(\varepsilon)}\lambda^2 s^2\int_{Q_{3r/2}}(\varepsilon+|D\mathbf{w}(x+se_k,t)|+|D\mathbf{w}(x,t)|)^{2-p}\,\mathrm{d}z \\
&\quad + \frac{c\lambda s^2}{r}\int_{Q_{2r}}|D\mathbf{w}|\,\mathrm{d}z \\
&\leqslant \tilde{\delta}r^2\int_{Q_{3r/2}}\frac{\varphi'(\varepsilon+|D\mathbf{w}(x+se_k,t)|+|D\mathbf{w}(x,t)|)}{\varepsilon+|D\mathbf{w}(x+se_k,t)|+|D\mathbf{w}(x,t)|}|\Delta_{k,s}D\mathbf{w}|^2\,\mathrm{d}z \\
&\quad + c\frac{\lambda s^2}{\tilde{\delta}r^2}\left(\frac{\varepsilon^{p-1}}{\varphi'(\varepsilon)}\lambda+\frac{1}{r}\right)\int_{Q_{2r}}[\varphi(|D\mathbf{w}|)+1]\,\mathrm{d}z \\
&\leqslant \tilde{\delta}r^2\int_{Q_{2r}}|\Delta_{k,s}\mathbf{V}^{\varepsilon}(D\mathbf{w})|^2\,\mathrm{d}z \\
&\quad + c\frac{s^2\lambda}{\tilde{\delta}r^2}\left(\frac{\varepsilon^{p-1}}{\varphi'(\varepsilon)}\lambda+1\right)\int_{Q_{2r}}[\varphi(|D\mathbf{w}|)+1]\,\mathrm{d}z.
\end{aligned}$$

Finally we have that for every $Q_{2r}\Subset Q_R$, $\delta,\tilde{\delta}\in(0,1)$ and $s,h\in\mathbb{R}$ with $0<s\leqslant h<r/2$ or $-r/2<h\leqslant s<0$,

$$\begin{aligned}
\int_{Q_r}|\Delta_{k,s}\mathbf{V}^{\varepsilon}(D\mathbf{w})|^2\,\mathrm{d}z &\leqslant c\delta\frac{|s|}{|h|}\fint_0^s\int_{Q_{2r}}|\Delta_{k,\tau}\mathbf{V}^{\varepsilon}(D\mathbf{w})|^2\,\mathrm{d}x\,\mathrm{d}\tau \\
&\quad + c\tilde{\delta}\int_{Q_{2r}}|\Delta_{k,s}\mathbf{V}^{\varepsilon}(D\mathbf{w})|^2\,\mathrm{d}z+\frac{h^2}{r^4}C(\delta,\tilde{\delta},\varepsilon,\lambda)\int_{Q_{2r}}[\varphi(|D\mathbf{w}|)+1]\,\mathrm{d}z\,.
\end{aligned}\tag{4.6}$$

Now, we re-absorb the first two terms on the right hand side. To do this, we first integrate both the sides of (4.6) with respect to s from 0 to h and apply Fubini's Theorem, so that

$$\begin{aligned}
&\fint_0^h\int_{Q_r}|\Delta_{k,s}\mathbf{V}^{\varepsilon}(D\mathbf{w})|^2\,\mathrm{d}z\,\mathrm{d}s \\
&\leqslant c\delta\fint_0^h\frac{|s|}{|h|}\fint_0^s\int_{Q_{2r}}|\Delta_{k,\tau}\mathbf{V}^{\varepsilon}(D\mathbf{w})|^2\,\mathrm{d}x\,\mathrm{d}\tau\,\mathrm{d}s+c\tilde{\delta}\fint_0^h\int_{Q_{2r}}|\Delta_{k,s}\mathbf{V}^{\varepsilon}(D\mathbf{w})|^2\,\mathrm{d}z\,\mathrm{d}s \\
&\quad +\frac{h^2}{r^4}C(\delta,\tilde{\delta},\varepsilon,\lambda)\int_{Q_{2r}}[\varphi(|D\mathbf{w}|)+1]\,\mathrm{d}z \\
&\leqslant c(\delta+\tilde{\delta})\fint_0^h\int_{Q_{2r}}|\Delta_{k,s}\mathbf{V}^{\varepsilon}(D\mathbf{w})|^2\,\mathrm{d}x\,\mathrm{d}s+\frac{h^2}{r^4}C(\delta,\tilde{\delta},\varepsilon,\lambda)\int_{Q_{2r}}[\varphi(|D\mathbf{w}|)+1]\,\mathrm{d}z.
\end{aligned}$$

Therefore, applying the Giaquinta-Modica type covering argument in [13, Lemma 13] we have that

$$\fint_0^h\int_{Q_r}|\Delta_{k,s}\mathbf{V}^{\varepsilon}(D\mathbf{w})|^2\,\mathrm{d}z\,\mathrm{d}s\leqslant c\frac{h^2}{r^4}C(\varepsilon,\lambda)\int_{Q_{5r}}[\varphi(|D\mathbf{w}|)+1]\,\mathrm{d}z$$

for every $Q_{5r} \Subset Q_R$ and $h \in \mathbb{R}$ with $0 < h < \frac{r}{10}$ or $-\frac{r}{10} < h < 0$. Inserting this into (4.6) with $s = h$, we have

$$\int_{Q_r} |\Delta_{k,h}\mathbf{V}^\varepsilon(D\mathbf{w})|^2\,\mathrm{d}z \leqslant c\tilde{\delta}\int_{Q_{2r}} |\Delta_{k,h}\mathbf{V}^\varepsilon(D\mathbf{w})|^2\,\mathrm{d}z + c\frac{h^2}{r^4}C(\tilde{\delta},\varepsilon,\lambda)\int_{Q_{5r}}[\varphi(|D\mathbf{w}|)+1]\,\mathrm{d}z.$$

Again applying the same covering argument, we have that for every $Q_{25r} \Subset Q_R$ and $h \in \mathbb{R}$ with $0 < |h| < \frac{r}{50}$,

$$\frac{1}{h^2}\int_{Q_r} |\Delta_{k,h}\mathbf{V}^\varepsilon(D\mathbf{w})|^2\,\mathrm{d}z \leqslant \frac{c}{r^4}C(\varepsilon,\lambda)\int_{Q_{25r}}[\varphi(|D\mathbf{w}|)+1]\,\mathrm{d}z. \tag{4.7}$$

Letting $h \to 0$ in (4.7), we then obtain (i).

The results in (ii) and (iii) are direct consequences of (4.7). Let $Q_{25r} \Subset Q_R$ and $h \in \mathbb{R}$ with $0 < |h| < \frac{r}{50}$. By Young's inequality we have that

$$\begin{aligned}
\int_{Q_r} h^{-p}|\Delta_{k,h}D\mathbf{w}|^p\,\mathrm{d}z &\leqslant ch^{-2}\int_{Q_r}(\varepsilon+|D\mathbf{w}(x+he_k)|+|D\mathbf{w}(x)|)^{p-2}|\Delta_{k,h}D\mathbf{w}|^2\,\mathrm{d}z\\
&\quad + c\int_{Q_r}(\varepsilon+|D\mathbf{w}(x+he_k)|+|D\mathbf{w}(x)|)^p\,\mathrm{d}z\\
&\leqslant ch^{-2}\frac{\varepsilon}{\varphi'(\varepsilon)}\int_{Q_r}\frac{\varphi'(\varepsilon+|D\mathbf{w}(x+he_k)|+|D\mathbf{w}(x)|)}{\varepsilon+|D\mathbf{w}(x+he_k)|+|D\mathbf{w}(x)|}|\Delta_{k,h}D\mathbf{w}|^2\,\mathrm{d}z\\
&\quad + c\int_{Q_{25r}}[\varphi(|D\mathbf{w}|)+1]\,\mathrm{d}z\\
&\leqslant \frac{c}{r^4}C(\varepsilon,\lambda)\int_{Q_{25r}}[\varphi(|D\mathbf{w}|)+1]\,\mathrm{d}z,
\end{aligned}$$

and, passing to the limit as $h \to 0$, this implies the first half of (ii). Moreover, by estimate (4.5), passing to the limit as $h \to 0$, we get the second half of (ii).

Finally we prove (iii). Set $M := \|D\mathbf{w}\| = \|D\mathbf{u}_\varepsilon\|_{L^\infty(Q_{25r},\mathbb{R}^{Nn})}$. Then, from (4.7), we have

$$\begin{aligned}
\int_{Q_r}|\Delta_{k,h}D\mathbf{w}|^2\,\mathrm{d}z &\leqslant \frac{\varepsilon+2M}{\varphi'(\varepsilon)}\int_{Q_r}\frac{\varphi'(\varepsilon+|D\mathbf{w}(x+he_k,t)|+|D\mathbf{w}(x,t)|)}{\varepsilon+|D\mathbf{w}(x+he_k,t)|+|D\mathbf{w}(x,t)|}|\Delta_{k,h}D\mathbf{w}|^2\,\mathrm{d}z\\
&\leqslant c\frac{\varepsilon+2M}{\varphi'(\varepsilon)}\int_{Q_r}|\Delta_{k,h}\mathbf{V}^\varepsilon(D\mathbf{w})|^2\,\mathrm{d}z\\
&\leqslant c\frac{\varepsilon+2M}{\varphi'(\varepsilon)}\frac{h^2}{r^4}C(\varepsilon,\lambda)\int_{Q_{25r}}[\varphi(|D\mathbf{w}|)+1]\,\mathrm{d}z.
\end{aligned}$$

This implies that $D\mathbf{w} \in L^2_{\mathrm{loc}}(-R^2,0;W^{1,2}_{\mathrm{loc}}(B_R;\mathbb{R}^{Nn}))$. Moreover, since $(\varphi_\varepsilon(|D\mathbf{w}|))_{x_k} = \varphi'_\varepsilon(|D\mathbf{w}|)\frac{D\mathbf{w}:D\mathbf{w}_{x_k}}{|D\mathbf{w}|}$,

$$\int_{Q_r}|D[\varphi_\varepsilon(|D\mathbf{w}|)]|^2\,\mathrm{d}z \leqslant c[\varphi'_\varepsilon(M)]^2\int_{Q_r}|D^2\mathbf{w}|^2\,\mathrm{d}z.$$

From this we get $\varphi_\varepsilon(|D\mathbf{w}|) \in L^2_{\mathrm{loc}}(-R^2,0;W^{1,2}_{\mathrm{loc}}(B_R))$, and the proof concludes. □

Combining (ii) of the above lemma with the parabolic Sobolev inequality, see [10, I, Proposition 3.1], we obtain the following result:

Lemma 4.2. *Let φ satisfy Assumption 2.2 with (1.5), and let $\mathbf{u}_\varepsilon$ be a weak solution to (4.2). Then $|D\mathbf{u}_\varepsilon| \in L^{\frac{p(n+2)}{n}}_{\mathrm{loc}}(\Omega_T)$.*

We end this section with the convergence results of $D\mathbf{u}_\varepsilon$ to $D\mathbf{u}$.

Lemma 4.3. *Let u be a weak solution to (1.1) and $\mathbf{u}_\varepsilon$ be the weak solution to (4.1) with $Q_R \Subset \Omega_T$. Then $D\mathbf{u}_\varepsilon$ converges to $D\mathbf{u}$ in $L^\varphi(Q_R)$.*

Proof. By virtue of (2.20), it suffices to show that

$$\lim_{\varepsilon\to 0+} \int_{Q_R} |\mathbf{V}(D\mathbf{u}_\varepsilon) - \mathbf{V}(D\mathbf{u})|^2 \,\mathrm{d}z = 0. \tag{4.8}$$

By following the proof of [15, Theorem 3.5], one has

$$\lim_{\varepsilon\to 0+} \int_{Q_R} |\mathbf{V}^\varepsilon(D\mathbf{u}_\varepsilon) - \mathbf{V}(D\mathbf{u})|^2 \,\mathrm{d}z = 0. \tag{4.9}$$

Moreover, together with (2.11), this implies that

$$\begin{aligned}
\int_{Q_R} \varphi(|D\mathbf{u}_\varepsilon|)\,\mathrm{d}z &\leqslant c\int_{Q_R} \Big[\varphi_\varepsilon(|D\mathbf{u}_\varepsilon|) + \varphi(\varepsilon)\Big]\,\mathrm{d}z \\
&\leqslant c\int_{Q_R} \Big[|\mathbf{V}^\varepsilon(D\mathbf{u}_\varepsilon)|^2 + \varphi(\varepsilon)\Big]\,\mathrm{d}z \\
&\leqslant c\int_{Q_R} \Big[|\mathbf{V}(D\mathbf{u})|^2 + \varphi(\varepsilon) + 1\Big]\,\mathrm{d}z \\
&\leqslant c\int_{Q_R} [\varphi(|D\mathbf{u}|) + 1]\,\mathrm{d}z
\end{aligned}$$

for any sufficiently small $\varepsilon > 0$. Applying (2.22), the change of shift formula (2.15), $\varphi_a(a) \leqslant c\varphi(a)$ and the preceding inequality we have that for any $\delta \in (0,1)$,

$$\begin{aligned}
\int_{Q_R} |\mathbf{V}^\varepsilon(D\mathbf{u}_\varepsilon) - \mathbf{V}(D\mathbf{u}_\varepsilon)|^2\,\mathrm{d}z &\leqslant c\int_{Q_R} \varphi_{|D\mathbf{u}_\varepsilon|}(\varepsilon)\,\mathrm{d}z \\
&\leqslant c\delta\int_{Q_R} \varphi_{|D\mathbf{u}_\varepsilon|}(|D\mathbf{u}_\varepsilon|)\,\mathrm{d}z + c_\delta\varphi(\varepsilon)|Q_R| \\
&\leqslant c\delta\int_{Q_R} \varphi(|D\mathbf{u}_\varepsilon|)\,\mathrm{d}z + c_\delta\varphi(\varepsilon)|Q_R| \\
&\leqslant c\delta\int_{Q_R} \Big[\varphi(|D\mathbf{u}|) + 1\Big]\,\mathrm{d}z + c_\delta\varphi(\varepsilon)|Q_R|\,,
\end{aligned}$$

hence

$$\limsup_{\varepsilon\to 0^+} \int_{Q_R} |\mathbf{V}^\varepsilon(D\mathbf{u}_\varepsilon) - \mathbf{V}(D\mathbf{u}_\varepsilon)|^2\,\mathrm{d}z \leqslant c\delta\int_{Q_R} \Big[\varphi(|D\mathbf{u}|) + 1\Big]\,\mathrm{d}z\,.$$

Therefore, since $\delta \in (0,1)$ is arbitrary, we have

$$\lim_{\varepsilon\to 0+} \int_{Q_R} |\mathbf{V}^\varepsilon(D\mathbf{u}_\varepsilon) - \mathbf{V}(D\mathbf{u}_\varepsilon)|^2\,\mathrm{d}z = 0.$$

This and (4.9) yield (4.8). □

5. Local boundedness of the gradient

Now, we address the problem of obtaining an L^∞-bound for $D\mathbf{u}$, by deriving uniform estimates in ε for weak solutions to the non-degenerate systems (4.1). We follow some ideas underlying the Moser iteration for the Lipschitz regularity for parabolic p-Laplace systems, which can be found in [8, Theorem 4], [9, Theorem 4] and [6, Proposition 3.1].

Note that $D\mathbf{u}_\varepsilon$ is weakly differentiable with respect to the spatial variable x by Lemma 4.1(*ii*). Hence differentiating (4.2) with respect to x_k we find

(5.1)
$$\partial_t(\mathbf{u}_\varepsilon)^\alpha_{x_k} = \left(\frac{\varphi'_\varepsilon(|D\mathbf{u}_\varepsilon|)}{|D\mathbf{u}_\varepsilon|}(\mathbf{u}_\varepsilon)^\alpha_{x_i x_k} + \left(\frac{\varphi''_\varepsilon(|D\mathbf{u}_\varepsilon|)}{|D\mathbf{u}_\varepsilon|} - \frac{\varphi'_\varepsilon(|D\mathbf{u}_\varepsilon|)}{|D\mathbf{u}_\varepsilon|^2}\right)\frac{(\mathbf{u}_\varepsilon)^\beta_{x_j}(\mathbf{u}_\varepsilon)^\beta_{x_j x_k}}{|D\mathbf{u}_\varepsilon|}(\mathbf{u}_\varepsilon)^\alpha_{x_i}\right)_{x_i}$$
$$=: \left(\bar{A}^{\alpha\beta}_{ij}(\mathbf{u}_\varepsilon)^\beta_{x_j x_k}\right)_{x_i}, \quad \alpha = 1,2,\dots,N,$$

where
$$\bar{A}^{\alpha\beta}_{ij} := (\mathbf{A}^\varepsilon)^{\alpha\beta}_{ij}(D\mathbf{u}_\varepsilon) = \frac{\partial \mathbf{A}^\varepsilon(\mathbf{Q})^\beta_j}{\partial Q^\beta_i}\bigg|_{\mathbf{Q}=D\mathbf{u}_\varepsilon}.$$

We start with obtaining a Caccioppoli type inequality for the system (5.1).

Lemma 5.1. *Let φ satisfy Assumption 2.2 with (1.5) and (2.7), and let $\mathbf{u}_\varepsilon$ be a weak solution to (4.2). Suppose $f \in C^{0,1}([0,\infty))$ is positive, increasing and satisfying $f'(s) > 0$ for a.e. s with $f(s) > 0$, and set $F(s) := \int_0^s \tau f(\tau)\,\mathrm{d}\tau$. For every $Q := B_\rho \times [t_1, t_2] \Subset \Omega_T$, $\xi \in C_0^\infty(B_\rho)$ with $0 \leqslant \xi \leqslant 1$ and $\eta \in C^\infty(\mathbb{R})$ with $0 \leqslant \eta \leqslant 1$ and $\eta_t \geqslant 0$, we have*

$$\begin{aligned}
&\int_{B_\rho} F(|D\mathbf{u}_\varepsilon(x,\tau_2)|)\xi^2\eta\,\mathrm{d}x - \int_{B_\rho} F(|D\mathbf{u}_\varepsilon(x,\tau_1)|)\xi^2\eta\,\mathrm{d}x \\
&\quad + \frac{1}{c}\int_Q \frac{\varphi'_\varepsilon(|D\mathbf{u}_\varepsilon|)}{|D\mathbf{u}_\varepsilon|}\left[f(|D\mathbf{u}_\varepsilon|)|D^2\mathbf{u}_\varepsilon|^2 + \frac{f'(|D\mathbf{u}_\varepsilon|)}{|D\mathbf{u}_\varepsilon|}|D(|D\mathbf{u}_\varepsilon|^2)|^2\right]\xi^2\eta\,\mathrm{d}z \\
&\leqslant c\int_Q \varphi'_\varepsilon(|D\mathbf{u}_\varepsilon|)\frac{f(|D\mathbf{u}_\varepsilon|)^2}{f'(|D\mathbf{u}_\varepsilon|)}|D\xi|^2\eta\,\mathrm{d}z + \int_Q F(|D\mathbf{u}_\varepsilon|)\xi^2\eta_t\,\mathrm{d}z\,.
\end{aligned} \tag{5.2}$$

Moreover, the term $\varphi'_\varepsilon(|D\mathbf{u}_\varepsilon|)\frac{f(|D\mathbf{u}_\varepsilon|)^2}{f'(|D\mathbf{u}_\varepsilon|)}$ in the above estimate can be replaced by $\varphi_\varepsilon(|D\mathbf{u}_\varepsilon|)f(|D\mathbf{u}_\varepsilon|)$.

Proof. For simplicity, we shall write $\bar\varphi = \varphi_\varepsilon$ and $\mathbf{w} = (w^\alpha) = \mathbf{u}_\varepsilon$.

We test (5.1) with $w^\alpha_{x_k} f(|D\mathbf{w}|)\xi^2\eta$ to obtain that

$$\underbrace{\int_{\tau_1}^{\tau_2}\int_{B_\rho}(w^\alpha_{x_k})_t w^\alpha_{x_k} f(|D\mathbf{w}|)\xi^2\eta\,\mathrm{d}x\mathrm{d}t}_{=:I_1} + \underbrace{\int_Q \bar{A}^{\alpha\beta}_{ij} w^\beta_{x_j x_k}[w^\alpha_{x_k} f(|D\mathbf{w}|)\xi^2\eta]_{x_i}\,\mathrm{d}z}_{=:I_2} = 0.$$

We estimate I_1 and I_2 separately. We have

$$\begin{aligned}
I_1 &= \int_{\tau_1}^{\tau_2}\int_{B_\rho}\left[\partial_t(F(|D\mathbf{w}|)\eta) - F(|D\mathbf{w}|)\eta_t\right]\xi^2\,\mathrm{d}x\,\mathrm{d}t \\
&= \int_{B_\rho} F(|D\mathbf{w}(x,\tau_2)|)\xi^2\eta\,\mathrm{d}x - \int_{B_\rho} F(|D\mathbf{w}(x,\tau_1)|)\xi^2\eta\,\mathrm{d}x - \int_Q F(|D\mathbf{w}|)\xi^2\eta_t\,\mathrm{d}z\,.
\end{aligned}$$

As for I_2, we first observe that, with (2.17), (2.18),

$$\bar{A}^{\alpha\beta}_{ij} w^\beta_{x_j x_k} w^\alpha_{x_k x_i} \geqslant (p-1)\frac{\bar\varphi'(|D\mathbf{w}|)}{|D\mathbf{w}|}|D^2\mathbf{w}|^2\,,$$

$$
\begin{aligned}
&\bar{A}^{\alpha\beta}_{ij} w^\beta_{x_jx_k} w^\alpha_{x_k} w^\kappa_{x_l} w^\kappa_{x_lx_i}\\
&= \frac{\bar\varphi'(|D\mathbf{w}|)}{|D\mathbf{w}|}\left\{\delta_{ij}\delta^{\alpha\beta} + \left(\frac{\bar\varphi''(|D\mathbf{w}|)|D\mathbf{w}|}{\bar\varphi'(|D\mathbf{w}|)} - 1\right)\frac{w^\alpha_{x_i} w^\beta_{x_j}}{|D\mathbf{w}|^2}\right\} w^\beta_{x_jx_k} w^\alpha_{x_k} w^\kappa_{x_l} w^\kappa_{x_lx_i}\\
&= \frac{\bar\varphi'(|D\mathbf{w}|)}{|D\mathbf{w}|}\left\{\frac{|D(|D\mathbf{w}|^2)|^2}{4} + \left(\frac{\bar\varphi''(|D\mathbf{w}|)|D\mathbf{w}|}{\bar\varphi'(|D\mathbf{w}|)} - 1\right)\frac{\sum_{\alpha=1}^N [Dw^\alpha \cdot D(|D\mathbf{w}|^2)]^2}{4|D\mathbf{w}|^2}\right\}\\
&\geqslant (p-1)\frac{\bar\varphi'(|D\mathbf{w}|)}{|D\mathbf{w}|}\frac{|D(|D\mathbf{w}|^2)|^2}{4}\,,
\end{aligned}
$$

and

$$
\begin{aligned}
|\bar{A}^{\alpha\beta}_{ij} w^\beta_{x_jx_k} w^\alpha_{x_k}| &= \frac{\bar\varphi'(|D\mathbf{w}|)}{|D\mathbf{w}|}\left|\left\{\delta_{ij}\delta^{\alpha\beta} + \left(\frac{\bar\varphi''(|D\mathbf{w}|)|D\mathbf{w}|}{\bar\varphi'(|D\mathbf{w}|)} - 1\right)\frac{w^\alpha_{x_i} w^\beta_{x_j}}{|D\mathbf{w}|^2}\right\} w^\beta_{x_jx_k} w^\alpha_{x_k}\right|\\
&\leqslant c\frac{\bar\varphi'(|D\mathbf{w}|)}{|D\mathbf{w}|}\big|D(|D\mathbf{w}|^2)\big|\,.
\end{aligned}
$$

Inserting these inequalities into I_2, we have

$$
\begin{aligned}
I_2 &= \int_Q \bar{A}^{\alpha\beta}_{ij} w^\beta_{x_jx_k}\left[w^\alpha_{x_kx_i} f(|D\mathbf{w}|) + w^\alpha_{x_k} w^\kappa_{x_l} w^\kappa_{x_lx_i}\frac{f'(|D\mathbf{w}|)}{|D\mathbf{w}|}\right]\xi^2\eta\,\mathrm{d}z\\
&\quad + \int_Q 2\bar{A}^{\alpha\beta}_{ij} w^\beta_{x_jx_k} w^\alpha_{x_k} f(|D\mathbf{w}|)\xi\xi_{x_i}\eta\,\mathrm{d}z\\
&\geqslant \frac{1}{c_1}\int_Q \frac{\bar\varphi'(|D\mathbf{w}|)}{|D\mathbf{w}|}\left[f(|D\mathbf{w}|)|D^2\mathbf{w}|^2 + \frac{f'(|D\mathbf{w}|)}{|D\mathbf{w}|}|D(|D\mathbf{w}|^2)|^2\right]\xi^2\eta\,\mathrm{d}z\\
&\quad - c\int_Q \frac{\bar\varphi'(|D\mathbf{w}|)}{|D\mathbf{w}|}\big|D(|D\mathbf{w}|^2)\big| f(|D\mathbf{w}|)\xi|D\xi|\eta\,\mathrm{d}z
\end{aligned}
\tag{5.3}
$$

for some constant $c_1 > 0$. Applying Young's inequality to the last integrand, we obtain

$$
\begin{aligned}
I_2 &\geqslant \frac{1}{2c_1}\int_Q \frac{\bar\varphi'(|D\mathbf{w}|)}{|D\mathbf{w}|}\left[f(|D\mathbf{w}|)|D^2\mathbf{w}|^2 + \frac{f'(|D\mathbf{w}|)}{|D\mathbf{w}|}|D(|D\mathbf{w}|^2)|^2\right]\xi^2\eta\,\mathrm{d}z\\
&\quad - c\int_Q \bar\varphi'(|D\mathbf{w}|)\frac{f(|D\mathbf{w}|)^2}{f'(|D\mathbf{w}|)}|D\xi|^2\eta\,\mathrm{d}z\,,
\end{aligned}
$$

or

$$
\begin{aligned}
I_2 &\geqslant \frac{1}{c_1}\int_Q \frac{\bar\varphi'(|D\mathbf{w}|)}{|D\mathbf{w}|}\left[f(|D\mathbf{w}|)|D^2\mathbf{w}|^2 + \frac{f'(|D\mathbf{w}|)}{|D\mathbf{w}|}|D(|D\mathbf{w}|^2)|^2\right]\xi^2\eta\,\mathrm{d}z\\
&\quad - c\int_Q \frac{\bar\varphi'(|D\mathbf{w}|)}{|D\mathbf{w}|}|D\mathbf{w}||D^2\mathbf{w}| f(|D\mathbf{w}|)\xi|D\xi|\eta\,\mathrm{d}z\\
&\geqslant \frac{1}{2c_1}\int_Q \frac{\bar\varphi'(|D\mathbf{w}|)}{|D\mathbf{w}|}\left[f(|D\mathbf{w}|)|D^2\mathbf{w}|^2 + \frac{f'(|D\mathbf{w}|)}{|D\mathbf{w}|}|D(|D\mathbf{w}|^2)|^2\right]\xi^2\eta\,\mathrm{d}z\\
&\quad - c\int_Q \bar\varphi(|D\mathbf{w}|) f(|D\mathbf{w}|)|D\xi|^2\eta\,\mathrm{d}z\,.
\end{aligned}
$$

Therefore, combining the above estimates, we get (5.2). □

Theorem 5.2. *Let φ satisfy Assumption 2.2 with* (1.5) *and* (2.7)*, and $\mathbf{u}_\varepsilon$ with $\varepsilon \in (0,1)$ be a weak solution to* (4.2)*. Then $D\mathbf{u}_\varepsilon \in L^\infty_{\mathrm{loc}}(\Omega_T, \mathbb{R}^{Nn})$. Moreover, we have that for every*

$Q_{2r}(z_0) \Subset \Omega_T$,

$$\|D\mathbf{u}_\varepsilon\|_{L^\infty(Q_r(z_0),\mathbb{R}^{Nn})} \leqslant c\left(\fint_{Q_{2r}(z_0)} \varphi_\varepsilon(|D\mathbf{u}_\varepsilon|)\,\mathrm{d}z + 1\right)^{\frac{2}{(n+2)p-2n}} \tag{5.4}$$

for some $c \geqslant 1$ depending on n, N, p and q, and independent of ε.

Proof. Step 1. (Setting and Caccioppoli type estimate) To enlighten the notation, we will write $\bar{\varphi} := \varphi_\varepsilon$ and $\mathbf{w} := \mathbf{u}_\varepsilon$. Let $Q_{2r} = Q_{2r}(z_0) \Subset Q_T$. Without loss of generality, we assume that $z_0 = (x_0, t_0) = (0, 0)$. Let $\rho_1 = s_1 r$ and $\rho_2 = s_2 r$ with $1 \leqslant s_1 < s_2 \leqslant 2$, $\xi \in C_0^\infty(B_{\rho_2})$ and $\eta \in C^\infty(\mathbb{R})$ be as in (3.3) and (3.4), respectively.

Then applying Lemma 5.1 with $f(t) = t^\chi$ where $\chi \geqslant 0$, $\rho = \rho_2$, $\tau_1 = -\rho_2^2$ and $\tau_2 \in (-\rho_1^2, 0)$, we have

$$\begin{aligned} &\sup_{-\rho_1^2<\tau<0} \int_{B_{\rho_2}} |D\mathbf{w}(x,\tau)|^{2+\chi}\xi^2\,\mathrm{d}x + \int_{Q_{\rho_2}} \frac{\bar{\varphi}'(|D\mathbf{w}|)}{|D\mathbf{w}|}|D\mathbf{w}|^\chi|D^2\mathbf{w}|^2\eta^2\xi^2\,\mathrm{d}z \\ &\leqslant \frac{c(2+\chi)}{(\rho_2-\rho_1)^2}\int_{Q_{\rho_2}} \left[|D\mathbf{w}|^2 + \bar{\varphi}(|D\mathbf{w}|)\right]|D\mathbf{w}|^\chi\,\mathrm{d}z\,. \end{aligned} \tag{5.5}$$

Step 2. (Improving inequality) We set

$$F(z) := [\bar{\varphi}(|D\mathbf{w}(x,t)|)|D\mathbf{w}(x,t)|^\chi]^{\frac{1}{2}}\eta(t)\xi(x)\,. \tag{5.6}$$

Note that, in order to enlighten the notation, we will often omit the dependence of F, u, η and ξ on the respective arguments. Differentiating (5.6) with respect to x_i we then have

$$\begin{aligned} F_{x_i} = {}& \tfrac{1}{2}\left[\bar{\varphi}(|D\mathbf{w}|)|D\mathbf{w}|^\chi\right]^{-\frac{1}{2}}\left[\bar{\varphi}'(|D\mathbf{w}|)|D\mathbf{w}|^\chi + \chi|D\mathbf{w}|^{\chi-1}\bar{\varphi}(|D\mathbf{w}|)\right]\frac{w^\alpha_{x_j}w^\alpha_{x_jx_i}}{|D\mathbf{w}|}\eta\xi \\ &+ \left[\bar{\varphi}(|D\mathbf{w}|)|D\mathbf{w}|^\chi\right]^{\frac{1}{2}}\eta\xi_{x_i}\,, \end{aligned}$$

whence, recalling the upper bound (3.3) for $|D\xi|$, we obtain

$$|DF|^2 \leqslant c(\chi+1)^2\frac{\bar{\varphi}'(|D\mathbf{w}|)}{|D\mathbf{w}|}|D\mathbf{w}|^\chi|D^2\mathbf{w}|^2\eta\xi + \frac{c}{(\rho_2-\rho_1)^2}\bar{\varphi}(|D\mathbf{w}|)|D\mathbf{w}|^\chi\,.$$

Therefore, combining with (5.5) we have

$$\begin{aligned} &\sup_{-\rho_1^2<\tau<0} \int_{B_{\rho_2}} |D\mathbf{w}(x,\tau)|^{2+\chi}\xi^2\,\mathrm{d}x + \int_{Q_{\rho_2}} |DF|^2\,\mathrm{d}z \\ &\leqslant \frac{c(1+\chi)^3}{(\rho_2-\rho_1)^2}\int_{Q_{\rho_2}} \left[|D\mathbf{w}|^2 + \bar{\varphi}(|D\mathbf{w}|)\right]|D\mathbf{w}|^\chi\,\mathrm{d}z\,. \end{aligned} \tag{5.7}$$

Now, applying Hölder's inequality, the Sobolev inequality to function $F \in W_0^{1,2}(B_r)$ and using (5.7), we can write

$$
\begin{aligned}
&\int_{Q_{\rho_1}} |D\mathbf{w}|^{\frac{4}{n}+\chi(\frac{2}{n}+1)}\bar{\varphi}(|D\mathbf{w}|)\,\mathrm{d}z \\
&\leqslant \int_{-\rho_1^2}^{0}\left(\int_{B_{\rho_1}}|D\mathbf{w}|^{2+\chi}\,\mathrm{d}x\right)^{\frac{2}{n}}\left(\int_{B_{\rho_1}}[\bar{\varphi}(|D\mathbf{w}|)|D\mathbf{w}|^{\chi}]^{\frac{n}{n-2}}\,\mathrm{d}x\right)^{\frac{n-2}{n}}\mathrm{d}t \\
&\leqslant \left(\sup_{-\rho_1^2<\tau<0}\int_{B_{\rho_1}}|D\mathbf{w}(x,\tau)|^{2+\chi}\,\mathrm{d}x\right)^{\frac{2}{n}}\int_{-\rho_1^2}^{0}\left(\int_{B_{\rho_2}}|F|^{2^*}\,\mathrm{d}x\right)^{\frac{2}{2^*}}\mathrm{d}t \\
&\leqslant \left(\sup_{-\rho_1^2<\tau<0}\int_{B_{\rho_1}}|D\mathbf{w}(x,\tau)|^{2+\chi}\,\mathrm{d}x\right)^{\frac{2}{n}}\int_{Q_{\rho_2}(z_0)}|DF|^2\,\mathrm{d}z \\
&\leqslant c\left(\frac{(1+\chi)^3r^2}{(\rho_2-\rho_1)^2}\int_{Q_{\rho_2}}\big(|D\mathbf{w}|^2+\bar{\varphi}(|D\mathbf{w}|)\big)\,|D\mathbf{w}|^{\chi}\,\mathrm{d}z\right)^{1+\frac{2}{n}}.
\end{aligned}
\tag{5.8}
$$

By Lemma 4.2, we have that $|D\mathbf{w}| \in L^{p(n+2)/n}(Q_{2r})$. Note that by under the assumption (1.5) on p, $p(n+2)/n > 2$. Therefore, it holds that

$$
\int_{Q_{2r}} \left[|D\mathbf{w}|^2+\bar{\varphi}(|D\mathbf{w}|)\right]\mathrm{d}z < \infty\,. \tag{5.9}
$$

Since $2 < \frac{4}{n} + p$ again by (1.5), setting

$$
\chi_1 := \frac{4}{n} + p - 2\,, \tag{5.10}
$$

we may improve estimate (5.8) as

$$
\begin{aligned}
&\int_{Q_{\rho_1}} (|D\mathbf{w}|^2+\bar{\varphi}(|D\mathbf{w}|))|D\mathbf{w}|^{\chi_1+\chi(1+\frac{2}{n})}\,\frac{\mathrm{d}z}{|Q_{2r}|} \\
&\qquad\leqslant c\left(\frac{(1+\chi)^3r^2}{(\rho_2-\rho_1)^2}\int_{Q_{\rho_2}}\left[(|D\mathbf{w}|^2+\bar{\varphi}(|D\mathbf{w}|))|D\mathbf{w}|^{\chi}+1\right]\frac{\mathrm{d}z}{|Q_{2r}|}\right)^{1+\frac{2}{n}}.
\end{aligned}
\tag{5.11}
$$

Step 3. (Iteration) Let s_1, s_2 such that $1 \leqslant s_1 < s_2 \leqslant 2$ be fixed. For $m = 0, 1, 2, \ldots,$ we set

$$
\chi_0 := 0 \quad \text{and} \quad \chi_m := \chi_1 + \theta\chi_{m-1} \ \ (m \geqslant 1), \quad \text{where} \ \ \theta := 1 + \frac{2}{n},
$$

and

$$
J_m := \int_{Q_{r_m}} \left[(|D\mathbf{w}|^2+\bar{\varphi}(|D\mathbf{w}|))|D\mathbf{w}|^{\chi_m}+1\right]\frac{\mathrm{d}z}{|Q_r|}, \quad \text{where} \ \ r_m := (s_1 + 2^{-m}(s_2-s_1))r\,.
$$

Note that $\chi_m = (\theta^m - 1)\frac{\chi_1 n}{2}$. Then we have from (5.11) that

$$
J_m \leqslant \frac{c4^{\theta m}\theta^{3\theta m}}{(s_2-s_1)^{2\theta}}J_{m-1}^{\theta} \leqslant \frac{c_0^m}{(s_2-s_1)^{2\theta}}J_{m-1}^{\theta}, \quad m = 1, 2, \ldots,
$$

where $c_0 \geqslant 1$ depends on n, N, p and q. Hence, for $m \geqslant 2$,

$$
\begin{aligned}
J_m &\leqslant \frac{c_0^m}{(s_2-s_1)^{2\theta}}\left(\frac{c_0^{m-1}}{(s_2-s_1)^{2\theta}}J_{m-2}^{\theta}\right)^{\theta} \\
&\leqslant \frac{c_0^{m+(m-1)\theta}}{(s_2-s_1)^{2(\theta+\theta^2)}}J_{m-2}^{\theta^2} \\
&\leqslant \cdots\cdots\cdots\cdots\cdots\cdots \\
&\leqslant \frac{c_0^{\sum_{k=1}^m (m-k+1)\theta^{k-1}}}{(s_2-s_1)^{2\sum_{k=1}^m \theta^k}}J_0^{\theta^m} \leqslant \left(\frac{c_1}{(s_2-s_1)^{\beta_0}}J_0\right)^{\theta^m}
\end{aligned}
$$

for some large $c_1, c_2, \beta_0 > 1$ depending on n, N, p and q. Consequently, setting

$$
\mathrm{d}\mu(z) := \left[|D\mathbf{w}(z)|^2 + \bar{\varphi}(|D\mathbf{w}(z)|)\right]\frac{\mathrm{d}z}{|Q_{2r}|},
$$

we have

$$
\begin{aligned}
\|D\mathbf{w}\|_{L^\infty(Q_{s_1 r},\mathbb{R}^{Nn})} &\leqslant \||D\mathbf{w}|\|_{L^\infty(Q_{s_1 r};\mathrm{d}\mu)} = \lim_{m\to\infty}\left(\int_{Q_r}|D\mathbf{w}|^{\chi_m}\,\mathrm{d}\mu\right)^{\frac{1}{\chi_m}} \leqslant \limsup_{m\to\infty} J_m^{\frac{1}{\chi_m}} \\
&\leqslant \limsup_{m\to\infty}\left(\frac{c_1}{(s_2-s_1)^{\theta_0}}J_0\right)^{\frac{\theta^i}{\chi_m}} \leqslant \frac{c}{(s_2-s_1)^{\theta_1}}J_0^{\frac{2}{n\chi_1}},
\end{aligned}
$$

where we used also the fact that $\chi_m = (\theta^m - 1)\frac{\chi_1 n}{2}$. Therefore, we have

$$
\|D\mathbf{w}\|_{L^\infty(Q_{s_1 r},\mathbb{R}^{Nn})} \leqslant \frac{c}{(s_2-s_1)^{\theta_1}}\left(\fint_{Q_{s_2 r}}\left[|D\mathbf{w}|^2 + \bar{\varphi}(|D\mathbf{w}|)\right]\,\mathrm{d}z + 1\right)^{\frac{2}{n\chi_1}}. \tag{5.12}
$$

By virtue of (5.9), this shows that $D\mathbf{w} \in L^\infty_{\mathrm{loc}}(Q_R;\mathbb{R}^{Nn})$.

Step 4. (Interpolation) Now we get rid of the term $|D\mathbf{w}|^2$ in the integrand in (5.12) by using an interpolation argument. Since $D\mathbf{w} \in L^\infty_{\mathrm{loc}}(Q_R;\mathbb{R}^{Nn})$ and $\frac{2(2-p)}{n\chi_1} < 1$ by (1.5), using Young's inequality we have that for every $1 \leqslant s_1 < s_2 \leqslant 2$,

$$
\begin{aligned}
\|D\mathbf{w}\|_{L^\infty(Q_{s_1 r},\mathbb{R}^{Nn})} &\leqslant \frac{c\|D\mathbf{w}\|_{L^\infty(Q_{s_2 r},\mathbb{R}^{Nn})}^{\frac{2(2-p)}{n\chi_1}}}{(s_2-s_1)^{\theta_1}}\left(\fint_{Q_{s_2 r}}|D\mathbf{w}|^p\,\mathrm{d}z\right)^{\frac{2}{n\chi_1}} \\
&\quad + \frac{c}{(s_2-s_1)^{\theta_1}}\left(\fint_{Q_{s_2 r}}\bar{\varphi}(|D\mathbf{w}|)\,\mathrm{d}z + 1\right)^{\frac{2}{n\chi_1}} \\
&\leqslant \frac{1}{2}\|D\mathbf{w}\|_{L^\infty(Q_{s_2 r},\mathbb{R}^{Nn})} + \frac{c}{(s_2-s_1)^{\frac{\theta_1 n\chi_1}{n\chi_1-2(2-p)}}}\left(\fint_{Q_{s_2 r}}|D\mathbf{w}|^p\,\mathrm{d}z\right)^{\frac{2}{n\chi_1-2(2-p)}} \\
&\quad + \frac{c}{(s_2-s_1)^{\theta_1}}\left(\fint_{Q_{s_2 r}}\bar{\varphi}(|D\mathbf{w}|)\,\mathrm{d}z + 1\right)^{\frac{2}{n\chi_1}} \\
&\leqslant \frac{1}{2}\|D\mathbf{w}\|_{L^\infty(Q_{s_2 r},\mathbb{R}^{Nn})} + \frac{c}{(s_2-s_1)^{\theta_2}}\left(\fint_{Q_{2r}}\bar{\varphi}(|D\mathbf{w}|) + 1\,\mathrm{d}z\right)^{\frac{2}{n\chi_1-2(2-p)}}
\end{aligned}
$$

Therefore, we can remove the first term on the right hand side (cfr. [20, Lemma 6.1]). Finally, recalling (5.10), we obtain (5.4). □

From the previous theorem and Lemma 4.3, we obtain the boundedness of the gradient of a weak solution to (1.1).

Corollary 5.3. *Let φ satisfy Assumption 2.2 with (1.5), and $\mathbf{u}$ be a weak solution to (1.1). Then $D\mathbf{u} \in L^{\infty}_{\mathrm{loc}}(\Omega_T, \mathbb{R}^{Nn})$. Moreover, we have that for every $Q_{2r} \Subset \Omega_T$,*

$$\|D\mathbf{u}\|_{L^{\infty}(Q_r(z_0),\mathbb{R}^{Nn})} \leqslant c \left(\fint_{Q_{2r}(z_0)} \varphi(|D\mathbf{u}|)\,\mathrm{d}z + 1 \right)^{\frac{2}{(n+2)p-2n}} \tag{5.13}$$

for some $c = c(n, N, p, q) \geqslant 1$.

Remark 5.4. When $\varphi(t) = t^p$ with $\frac{2n}{n+2} < p < 2$, the estimate (5.13) is exactly the same as [10, eq. (5.10)].

6. Hölder continuity of $D\mathbf{u}$ revisited

We prove local Hölder continuity for the gradient of weak solution to (1.1), where φ satisfies Assumption 2.3. We remark that the result was already obtained by Lieberman in [27] by assuming the local boundedness of $D\mathbf{u}$. In this section, we take advantage of the results of Section 5 and revisit his $C^{1,\alpha}$-regularity's proof, according to the setting of our paper. We also note that Lieberman's proof parallels the one given by Di Benedetto and Friedman [11, 12], using a measure theoretic approach. In addition, we are adapting the geometry of the cylinders accordingly, due to the growth conditions of the operator.

We define the intrinsic parabolic cylinder associated with an N-function φ as

$$Q_r^{\lambda}(x_0, t_0) := B_r(x_0) \times (t_0 - r^2/\varphi''(\lambda), t_0].$$

where $\lambda, r > 0$, and oscillation of a function $f : U \to \mathbb{R}^m$ by

$$\operatorname*{osc}_U f := \sup_{x,y\in U} |f(x) - f(y)|\,.$$

Now, we state the main result of this section.

Theorem 6.1. *Let $\varphi : [0, \infty) \to [0, \infty)$ satisfy Assumption 2.3, and let $\mathbf{u}$ be a weak solution to the parabolic system (1.1). and $Q_R(z_0) \Subset \Omega_T$. If $D\mathbf{u} \in L^{\infty}_{\mathrm{loc}}(Q_R(z_0); \mathbb{R}^{Nn})$, then $D\mathbf{u} \in C^{0,\alpha}_{loc}(Q_R(z_0); \mathbb{R}^{Nn})$ for some $\alpha \in (0, 1)$ depending on n, N, p, q, γ_1 and c_h. Moreover, any $Q_R(z_0) \Subset \Omega_T$, $r \in (0, R)$ and $\lambda \geqslant \|D\mathbf{u}\|_{L^{\infty}(Q_R(z_0);\mathbb{R}^{Nn})}$, we have*

$$\operatorname*{osc}_{Q_r(z_0)} D\mathbf{u} \leqslant c\lambda \left(\max\left\{ \varphi''(\lambda)^{\frac{1}{2}}, \varphi''(\lambda)^{-\frac{1}{2}} \right\} \frac{r}{R} \right)^{\alpha}$$

for some $c = c(n, N, p, q, \gamma_1, c_h) > 0$.

This result can be obtained by approximation via Lemma 4.3, once we obtain the analog of Theorem 6.1 for the gradients $D\mathbf{u}_\varepsilon$ of weak solutions $\mathbf{u}_\varepsilon$ to the approximating nondegenerate parabolic system (4.2), where $\varepsilon \in (0, 1]$. This will be a consequence of the following two propositions (cfr. [27, Propositions 1.3 and 1.4]) for $\mathbf{u}_\varepsilon$. Note that all estimates in this section are independent of $\varepsilon \in (0, 1]$. Thus, for simplicity, we shall write $\mathbf{u} = \mathbf{u}_\varepsilon$ and $\varphi = \varphi_\varepsilon$.

The first proposition provides an estimate on the oscillation of $D\mathbf{u}$ on subcylinders when $|D\mathbf{u}|$ is small on a small portion of the main cylinder.

Proposition 6.2. *Let $\varphi : [0, \infty) \to [0, \infty)$ satisfy Assumption 2.3, and let $\mathbf{u}$ be a weak solution to (4.2). Suppose that for some $\lambda, R > 0$, $Q_R^{\lambda}(z_0) \Subset \Omega_T$ and*

$$|D\mathbf{u}| \leqslant \lambda \quad \text{in } Q_R^{\lambda}(z_0). \tag{6.1}$$

There exist $\sigma \in (0, 2^{-(n+1)})$ and $C \geqslant 1$ depending on n, N, p, q, γ_1 and c_h. such that if

$$\big|\{|D\mathbf{u}| \leqslant (1-\sigma)\lambda\} \cap Q_R^\lambda(z_0)\big| \leqslant \sigma |Q_R^\lambda(z_0)| \,, \tag{6.2}$$

then

$$\operatorname*{osc}_{Q_r^\lambda(z_0)} D\mathbf{u} \leqslant C\left(\frac{r}{R}\right)^{\frac{3}{4}} \operatorname*{osc}_{Q_R^\lambda(z_0)} D\mathbf{u} \tag{6.3}$$

for all $r \in (0, R)$.

If (6.2) fails, the following proposition gives an estimate of how $|D\mathbf{u}|$ decreases.

Proposition 6.3. *Let $\varphi : [0, \infty) \to [0, \infty)$ satisfy Assumption 2.3, and let $\mathbf{u}$ be a weak solution to (4.2). Suppose that for some $\lambda, R > 0$, $Q_R^\lambda(z_0) \Subset \Omega_T$ and (6.1) holds. For any $\sigma \in (0, \frac{1}{2})$, there exists $\nu \in (0, 1)$ depending on n, N, p, q and σ such that if*

$$|\{|D\mathbf{u}| \leqslant (1-\sigma)\lambda\} \cap Q_R^\lambda(z_0)| > \sigma |Q_R^\lambda(z_0)| \,, \tag{6.4}$$

then

$$|D\mathbf{u}| \leqslant \nu\lambda \quad \text{in} \quad Q_{\sigma R/2}^\lambda(z_0) \,. \tag{6.5}$$

Proposition 6.2 and Proposition 6.3 will be proved in Subsection 6.1 and Subsection 6.2, respectively. In the remaining subsections, we always assume that $\varphi : [0, \infty) \to [0, \infty)$ satisfies Assumption 2.3, $\mathbf{u}$ is a weak solution to (4.2), and (6.1) holds for some $Q_R^\lambda(z_0) \Subset \Omega_T$ with $\lambda, R > 0$. In addition, without loss of generality, we further assume that assume (2.7) and $z_0 = (x_0, t_0) = (0, 0) = 0$, and write $Q_r^\lambda = Q_r^\lambda(0)$ for all $r \in (0, R]$.

6.1. **Proof of Proposition 6.2.** Before starting the proof, we recall the following weighted version of Poincaré's inequality, which is quite elementary and can be deduced, for instance, from [13, Theorem 7].

Lemma 6.4. *Suppose $f \in W^{1,p}(B_R; \mathbb{R}^m)$ and $\xi \in L^1(B_R)$ is nonnegative and satisfies $\|\xi\|_{L^1(B_R)} = 1$. Then we have*

$$\int_{B_R} \left|\frac{f - \langle f \rangle_\xi}{R}\right|^p \mathrm{d}x \leqslant c \int_{B_R} |Df|^p \,\mathrm{d}x$$

for some $c = c(n, m, p) > 0$, where $\langle f \rangle_\xi = \int f\xi \,\mathrm{d}x$.

We first derive a higher integrability result for $D\mathbf{u}$ (cfr. [27, Lemma 4.2]).

Lemma 6.5. *Let $\mathbf{P} \in \mathbb{R}^{Nn}$ satisfying $\frac{\lambda}{2} \leqslant |\mathbf{P}| \leqslant \lambda$. There exist $\gamma, c > 0$ depending on n, N, p and q such that*

$$\left(\fint_{Q_{R/2}^\lambda} |D\mathbf{u} - \mathbf{P}|^{2(1+\gamma)} \,\mathrm{d}z\right)^{\frac{1}{1+\gamma}} \leqslant c \fint_{Q_R^\lambda} |D\mathbf{u} - \mathbf{P}|^2 \,\mathrm{d}z \,.$$

Proof. Fix any $Q_{2r}^\lambda(z_1) \subset Q_R^\lambda$ with $z_1 = (x_1, t_1) \in Q_R^\lambda$ and $r \leqslant r_1 < r_2 \leqslant 2r$. We further set $r_3 = \frac{r_1 + r_2}{2}$, $r_4 = \frac{r_1 + 3r_2}{4}$ and $t_2 = t_1 - \frac{r_2^2}{\varphi''(\lambda)}$. Note that $r_1 < r_3 < r_4 < r_2$. We consider two cut-off functions. Let $\xi_0 \in C_0^\infty(B_{r_2}(x_1))$ satisfying $0 \leqslant \xi_0 \leqslant 1$, $\xi_0 = 1$ in $B_{r_4}(x_1)$ and $|D\xi_0| \leqslant 8/(r_2 - r_1)$, and set $\xi = \|\xi_0\|_{L^1(B_{r_2})}^{-1} \xi_0$. Note that $|B_r| \leqslant |B_{r_4}| \leqslant \|\xi_0\|_{L^1(B_{r_2})} \leqslant |B_{r_2}| \leqslant 2^n |B_r|$ and $\|\xi\|_{L^1(B_{r_2})} = 1$. Next, let $\zeta \in C^\infty(Q_{r_2}^\lambda(z_1))$ such that $\zeta = 0$ on $\partial_{\mathrm{p}} Q_{r_2}^\lambda(z_1)$, $\zeta = 1$ in $Q_{r_1}^\lambda(z_1)$,

$$|D\zeta|^2 + |D^2\zeta| \leqslant \frac{c}{(r_2 - r_1)^2} \quad \text{and} \quad 0 \leqslant \zeta_t \leqslant \frac{c\varphi''(\lambda)}{(r_2 - r_1)^2}.$$

Finally, define

$$\mathbf{w}(z) := \mathbf{u}(z) - \mathbf{P}(x - x_1), \quad \mathbf{W}(t) := \int_{B_{r_2}(x_1)} \mathbf{w}(x,t)\xi(x)\,\mathrm{d}x\,,$$

$$\text{and} \quad \tilde{\mathbf{w}} := \mathbf{w} - \mathbf{w}_0 \quad \text{with} \quad \mathbf{w}_0 := \frac{\varphi''(\lambda)}{r^2}\int_{Q^\lambda_{r_2}(z_1)} \mathbf{w}\xi\,\mathrm{d}z = \fint_{t_2}^{t_1} \mathbf{W}(t)\,\mathrm{d}t\,.$$

We take $\zeta^\chi \tilde{\mathbf{w}}$ with $\chi \geqslant 2$ as a test function in the weak form of (1.1) to get, for every $\tau \in I^\lambda_{r_2}(t_1)$,

$$\int_{t_2}^{\tau}\int_{B_{r_2}(z_1)} (\mathbf{w}_t \cdot \tilde{\mathbf{w}})\zeta^\chi\,\mathrm{d}x\,\mathrm{d}t + \int_{t_2}^{\tau}\int_{B_{r_2}(z_1)} (\mathbf{A}(D\mathbf{u}) - \mathbf{A}(\mathbf{P})) : D(\zeta^\chi\tilde{\mathbf{w}})\,\mathrm{d}x\,\mathrm{d}t = 0,$$

which yields

$$\begin{aligned}
&\sup_{\tau\in I^\lambda_{r_2}(t_1)} \int_{B_{r_2}(x_1)} |\tilde{\mathbf{w}}|^2\zeta^\chi\,\mathrm{d}x + \int_{Q^\lambda_{r_2}(z_1)} \frac{\varphi'(|D\mathbf{u}|+|\mathbf{P}|)}{|D\mathbf{u}|+|\mathbf{P}|}|D\mathbf{w}|^2\zeta^\chi\,\mathrm{d}z \\
&\leqslant \frac{c}{r_2 - r_1}\int_{Q^\lambda_{r_2}(z_1)} \frac{\varphi'(|D\mathbf{u}|+|\mathbf{P}|)}{|D\mathbf{u}|+|\mathbf{P}|}|D\mathbf{w}||\tilde{\mathbf{w}}|\zeta^{\chi-1}\,\mathrm{d}z + c\frac{\varphi''(\lambda)}{(r_2-r_1)^2}\int_{Q^\lambda_{r_2}(z_1)} |\tilde{\mathbf{w}}|^2\zeta^{\chi-1}\,\mathrm{d}z.
\end{aligned}$$

Here we used (2.19). Set

$$S_\chi := \sup_{\tau\in I^\lambda_{r_2}(t_1)} \int_{B_{r_2}(x_1)} |\tilde{\mathbf{w}}|^2\zeta^\chi\,\mathrm{d}x\,.$$

Applying Young's inequality to the integrand of the first integral on the right hand side and using the facts that $\chi \geqslant 2$, $|D\mathbf{u}| \leqslant \lambda$ and $\frac{\lambda}{2} \leqslant |\mathbf{P}| \leqslant \lambda$, we have

$$S_\chi + \varphi''(\lambda)\int_{Q^\lambda_{r_2}(z_1)} |D\mathbf{w}|^2\zeta^\chi\,\mathrm{d}z \leqslant c\frac{\varphi''(\lambda)}{(r_2-r_1)^2}\int_{Q^\lambda_{r_2}(z_1)} |\tilde{\mathbf{w}}|^2\zeta^{\chi-2}\,\mathrm{d}z$$

Then, when $\chi = 4$ and $\chi = 2$, we have

$$\int_{Q^\lambda_{r_1}(z_1)} |D\mathbf{w}|^2\,\mathrm{d}z \leqslant c\frac{S_2^{\frac{2}{n+2}}}{(r_2-r_1)^2}\int_{t_2}^{t_1}\left(\int_{B_{r_2}(x_1)} |\tilde{\mathbf{w}}(x,t)|^2\,\mathrm{d}x\right)^{\frac{n}{n+2}}\mathrm{d}t\,, \tag{6.6}$$

and

$$S_2 \leqslant c\frac{\varphi''(\lambda)}{(r_2-r_1)^2}\int_{Q^\lambda_{r_2}(z_1)} |\tilde{\mathbf{w}}|^2\,\mathrm{d}z \leqslant c\frac{\varphi''(\lambda)}{(r_2-r_1)^2}\int_{Q^\lambda_{r_2}(z_1)} \left[|\mathbf{w} - \mathbf{W}(t)|^2 + |\mathbf{w}_0 - \mathbf{W}(t)|^2\right]\mathrm{d}z\,.$$

By Poincaré's inequality with the weight ξ (Lemma 6.4) we see that for every $t \in I^\lambda_{r_2}(t_1)$

$$\int_{B_{r_2}} |\mathbf{w}(x,t) - \mathbf{W}(t)|^2\,\mathrm{d}x \leqslant cr^2\int_{B_{r_2}} |D\mathbf{w}(x,t)|^2\,\mathrm{d}x.$$

Moreover, by testing (4.2) with $\zeta = (\xi,\cdots,\xi)$ and using $|D\xi| \leqslant \frac{c}{r^n(r_2-r_1)}$ and (2.19) with $|D\mathbf{u}| \leqslant \lambda \leqslant 2|\mathbf{P}| \leqslant 2\lambda$, we see that for every $t_2 < \tau < \tau' < t_1$,

$$\begin{aligned}
|\mathbf{W}(\tau) - \mathbf{W}(\tau')| &= \left|\int_{\tau}^{\tau'}\int_{B_{r_2}(x_1)} (\mathbf{A}(D\mathbf{u}) - \mathbf{A}(\mathbf{P})) : D\zeta\,\mathrm{d}x\,\mathrm{d}t\right| \\
&\leqslant c\frac{r^2}{(r_2-r_1)}\fint_{Q^\lambda_{r_2}(z_1)} |D\mathbf{w}|\,\mathrm{d}z\,,
\end{aligned}$$

hence for every $t \in (t_2, t_1)$,

$$|\mathbf{w}_0 - \mathbf{W}(t)| \leqslant \sup_{t_2<\tau<\tau'<t_1} |\mathbf{W}(\tau) - \mathbf{W}(\tau')| \leqslant c\frac{r^2}{(r_2-r_1)} \fint_{Q^\lambda_{r_2}(z_1)} |D\mathbf{w}|\,\mathrm{d}z\,. \tag{6.7}$$

Therefore, combining the above estimates together with Hölder's inequality, we have

$$S_2 \leqslant c\frac{\varphi''(\lambda)}{(r_2-r_1)^2}\left(\frac{r^4}{(r_2-r_1)^2}+r^2\right)\int_{Q^\lambda_{r_2}(z_1)} |D\mathbf{w}|^2\,\mathrm{d}z \leqslant c\frac{\varphi''(\lambda)r^4}{(r_2-r_1)^4}\int_{Q^\lambda_{r_2}(z_1)} |D\mathbf{w}|^2\,\mathrm{d}z\,. \tag{6.8}$$

Moreover, by (6.7), a weighted Sobolev-Poincaré type inequality and Hölder's inequality, we also see that for every $t \in (t_2, t_1)$

$$\begin{aligned}\int_{B_{r_2}(x_1)} |\tilde{\mathbf{w}}(x,t)|^2\,\mathrm{d}x &\leqslant c\int_{B_{r_2}(x_1)} |\mathbf{w}(x,t)-\mathbf{W}(t)|^2\,\mathrm{d}x + cr^n|\mathbf{W}(t)-\mathbf{w}_0|^2 \\ &\leqslant c\left(\int_{B_{r_2}(x_1)} |D\mathbf{w}(x,t)|^{\frac{2n}{n+2}}\,\mathrm{d}x\right)^{\frac{n+2}{n}} + c\frac{r^{n+4}}{(r_2-r_1)^2}\left(\fint_{Q^\lambda_{r_2}(z_1)} |D\mathbf{w}|^{\frac{2n}{n+2}}\,\mathrm{d}z\right)^{\frac{n+2}{n}}.\end{aligned} \tag{6.9}$$

Therefore, inserting (6.8) and (6.9) into (6.6) and using Young's inequality, we have that for every $0 < r_1 < r_2 \leqslant r$,

$$\begin{aligned}\int_{Q^\lambda_{r_1}(z_1)} |D\mathbf{w}|^2\,\mathrm{d}z &\leqslant \frac{c}{(r_2-r_1)^2}\left(\frac{\varphi''(\lambda)r^4}{(r_2-r_1)^4}\int_{Q^\lambda_{r_2}(z_1)} |D\mathbf{w}|^2\,\mathrm{d}z\right)^{\frac{2}{n+2}} \\ &\qquad \times\left(\frac{r}{r_2-r_1}\right)^{\frac{2n}{n+2}}\int_{Q^\lambda_{r_2}(z_1)} |D\mathbf{w}|^{\frac{2n}{n+2}}\,\mathrm{d}z \\ &\leqslant c\left(\frac{r}{r_2-r_1}\right)^{\frac{4(n+3)}{n+2}}\frac{\varphi''(\lambda)^{\frac{2}{n+2}}}{r^2}\left(\int_{Q^\lambda_{r_2}(z_1)} |D\mathbf{w}|^2\,\mathrm{d}z\right)^{\frac{2}{n+2}}\int_{Q^\lambda_{2r}(z_1)} |D\mathbf{w}|^{\frac{2n}{n+2}}\,\mathrm{d}z \\ &\leqslant \frac{1}{2}\left(\frac{r}{r_2-r_1}\right)^{2(n+3)}\int_{Q^\lambda_{r_2}(z_1)} |D\mathbf{w}|^2\,\mathrm{d}z + c\left(\frac{\varphi''(\lambda)}{r^{n+2}}\right)^{\frac{2}{n}}\left(\int_{Q^\lambda_{2r}(z_1)} |D\mathbf{w}|^{\frac{2n}{n+2}}\,\mathrm{d}z\right)^{\frac{n+2}{n}}.\end{aligned}$$

Then we can remove the first term on the right hand side, and have

$$\fint_{Q^\lambda_{r}(z_1)} |D\mathbf{w}|^2\,\mathrm{d}z \leqslant c\left(\fint_{Q^\lambda_{2r}(z_1)} |D\mathbf{w}|^{\frac{2n}{n+2}}\,\mathrm{d}z\right)^{\frac{n+2}{n}},$$

for every $Q^\lambda_r(z_1) \subset Q^\lambda_R$. Finally, applying Gehring's Lemma (cfr. [20, Theorem 6.6]), we obtain the conclusion. □

Next we obtain an L^2-comparison estimate between $D\mathbf{u}$ and the gradient of a weak solution to a corresponding linear system with constant coefficients (cfr. [27, Lemma 4.3]). We recall (2.8) and the definition of $A^{\alpha\beta}_{ij}$ in (2.17), so that

$$\sum_{i,j,\alpha,\beta} |A^{\alpha\beta}_{ij}(\mathbf{Q}) - A^{\alpha\beta}_{ij}(\mathbf{P})| \leqslant c_h\left(\frac{|\mathbf{Q}-\mathbf{P}|}{|\mathbf{Q}|}\right)^{\gamma_1}\varphi''(|\mathbf{Q}|) \quad \text{for} \quad |\mathbf{Q}-\mathbf{P}| \leqslant \frac{1}{2}|\mathbf{Q}|. \tag{6.10}$$

Lemma 6.6. *Let* $\mathbf{P} = (P_i^\alpha) \in \mathbb{R}^{Nn}$ *satisfying* $\frac{\lambda}{2} \leqslant |\mathbf{P}| \leqslant \lambda$, *and* $\mathbf{v} = (v_i^\alpha)$ *be the weak solution to*

$$\begin{cases} (v^\alpha)_t - (A_{ij}^{\alpha\beta}(\mathbf{P}) v_{x_j}^\beta)_{x_i} = 0 & \text{in } \; Q_{R/2}^\lambda, \quad \alpha = 1, 2, \ldots, N, \\ \mathbf{v} = \mathbf{u} & \text{on } \; \partial_p Q_{R/2}^\lambda \,. \end{cases} \tag{6.11}$$

Then for every $\varepsilon_0 \in (0, \frac{1}{2})$,

$$\fint_{Q_{R/2}^\lambda} |D\mathbf{u} - D\mathbf{v}|^2 \,\mathrm{d}z \leqslant c \left[\varepsilon_0^{2\gamma_1} + \varepsilon_0^{-2\gamma} \lambda^{-2\gamma} \left(\fint_{Q_R^\lambda} |D\mathbf{u} - \mathbf{P}|^2 \,\mathrm{d}z \right)^\gamma \right] \fint_{Q_R^\lambda} |D\mathbf{u} - \mathbf{P}|^2 \,\mathrm{d}z \,,$$

for some $c = c(n, N, p, q, \gamma_1, c_h) > 0$, *where* $\gamma > 0$ *is from Lemma 6.5*

Proof. Observe the $\mathbf{u}$ satisfies

$$(u^\alpha)_t - (A_{ij}^{\alpha\beta}(\mathbf{P}) u_{x_j}^\beta)_{x_i} = -\Big(A_{ij}^{\alpha\beta}(\mathbf{P})(u_{x_j}^\beta - P_j^\beta) + [\mathbf{A}(\mathbf{P})]_i^\alpha - [\mathbf{A}(D\mathbf{u})]_i^\alpha \Big)_{x_i} =: -(H_i^\alpha)_{x_i}$$

for every $\alpha = 1, 2, \ldots, N$. By taking $u^\alpha - v^\alpha$ as a test function in the weak form of the above two equations, we have

$$\frac{1}{2} \int_{B_{R/2}} |\mathbf{u}(x,0) - \mathbf{v}(x,0)|^2 \,\mathrm{d}x + \int_{Q_{R/2}^\lambda} A_{ij}^{\alpha\beta}(\mathbf{P})(u_{x_j}^\beta - v_{x_j}^\beta)(u_{x_i}^\alpha - v_{x_i}^\alpha) \,\mathrm{d}z = \int_{Q_{R/2}^\lambda} H_i^\alpha (u_{x_i}^\alpha - v_{x_i}^\alpha) \,\mathrm{d}z \,,$$

and

$$\fint_{Q_{R/2}^\lambda} \varphi''(|\mathbf{P}|) |D\mathbf{u} - D\mathbf{v}|^2 \,\mathrm{d}z \leqslant \frac{c}{\varphi''(|\mathbf{P}|)} \fint_{Q_{R/2}^\lambda} |\mathbf{H}|^2 \,\mathrm{d}z \,,$$

where $\mathbf{H} = (H_i^\alpha)$. We note that

$$\begin{aligned} H_i^\alpha &= A_{ij}^{\alpha\beta}(\mathbf{P})(u_{x_j}^\beta - P_j^\beta) + [\mathbf{A}(\mathbf{P})]_i^\alpha - [\mathbf{A}(D\mathbf{u})]_i^\alpha \\ &= A_{ij}^{\alpha\beta}(\mathbf{P})(u_{x_j}^\beta - P_j^\beta) - \left(\int_0^1 A_{ij}^{\alpha\beta}(\tau D\mathbf{u} + (1-\tau)\mathbf{P}) \,\mathrm{d}\tau \right) (u_{x_j}^\beta - P_j^\beta) \,. \end{aligned}$$

If $|D\mathbf{u} - \mathbf{P}| \leqslant \varepsilon_0 |\mathbf{P}|$, by (6.10)

$$|\mathbf{H}| \leqslant c \left(\int_0^1 \left[\frac{|\tau(D\mathbf{u} - \mathbf{P})|}{|\mathbf{P}|} \right]^{\gamma_1} \mathrm{d}\tau \right) \varphi''(|\mathbf{P}|) |D\mathbf{u} - \mathbf{P}| \leqslant c \varepsilon_0^{\gamma_1} \varphi''(|\mathbf{P}|) |D\mathbf{u} - \mathbf{P}| \,.$$

If $|D\mathbf{u} - \mathbf{P}| > \varepsilon_0 |\mathbf{P}|$, then by (2.18) and (6.1)

$$|\mathbf{H}| \leqslant c \varphi''(|\mathbf{P}|) |D\mathbf{u} - \mathbf{P}| \leqslant c \varepsilon_0^{-\gamma} \varphi''(|\mathbf{P}|) |\mathbf{P}|^{-\gamma} |D\mathbf{u} - \mathbf{P}|^{1+\gamma} \,,$$

where γ is from Lemma 6.5. Combining the above results we have

$$\fint_{Q_{R/2}^\lambda} |D\mathbf{u} - D\mathbf{v}|^2 \,\mathrm{d}z \leqslant c \left(\varepsilon_0^{2\gamma_1} \fint_{Q_{R/2}^\lambda} |D\mathbf{u} - \mathbf{P}|^2 \,\mathrm{d}z + c \varepsilon_0^{-2\gamma} \lambda^{-2\gamma} \fint_{Q_{R/2}^\lambda} |D\mathbf{u} - \mathbf{P}|^{2(1+\gamma)} \,\mathrm{d}z \right) .$$

Therefore, applying Lemma 6.5, we obtain the desired estimate. □

Set $\bar{\mathbf{v}}(x,t) := \mathbf{v}\left(Rx, \frac{R^2}{\varphi''(\lambda)} t \right)$, where $\mathbf{v}$ is a weak solution to (6.11) with $\mathbf{P} \in \mathbb{R}^{Nn}$ satisfying $\frac{\lambda}{2} \leqslant |\mathbf{P}| \leqslant \lambda$. Then $\bar{\mathbf{v}}$ is a weak solution to

$$\bar{v}_t^\alpha - \left(\frac{A_{ij}^{\alpha\beta}(\mathbf{P})}{\varphi''(\lambda)} \bar{v}_{x_j}^\beta \right)_{x_i} = 0 \quad \text{in} \;\; Q_{1/2}, \quad \alpha = 1, 2, \ldots, N.$$

Since $\bar{L}^{-1}|\boldsymbol{\omega}|^2 \leqslant \frac{A_{ij}^{\alpha\beta}(\mathbf{P})}{\varphi''(\lambda)}\omega_i^\alpha\omega_j^\beta \leqslant \bar{L}|\boldsymbol{\omega}|^2$ for all $\boldsymbol{\omega} = (\omega_i^\alpha) \in \mathbb{R}^{Nn}$ and some $\bar{L} \geqslant 1$, by regularity theory for linear parabolic systems with constant coefficients (see, for instance, [20, XI. Theorem 6.6] with its proof), we have that for every $\rho \in (0, \frac{1}{2})$,

$$\int_{Q_\rho} |D\bar{\mathbf{v}} - (D\bar{\mathbf{v}})_{Q_\rho}|^2 \,\mathrm{d}z \leqslant c\rho^2 \int_{Q_{1/2}} |D\bar{\mathbf{v}} - (D\bar{\mathbf{v}})_{Q_{1/2}}|^2 \,\mathrm{d}z\,,$$

which implies the following estimate for $\mathbf{v}$: for every $\rho \in (0, \frac{R}{2})$,

$$\int_{Q_\rho^\lambda} |D\mathbf{v} - (D\mathbf{v})_{Q_\rho^\lambda}|^2 \,\mathrm{d}z \leqslant c\left(\frac{\rho}{R}\right)^{n+4} \int_{Q_{R/2}^\lambda} |D\mathbf{v} - (D\mathbf{v})_{Q_{R/2}^\lambda}|^2 \,\mathrm{d}z\,. \tag{6.12}$$

The estimate (6.12) is a key ingredient to obtain the following result, which provides an estimate for the decay of the mean oscillation of $D\mathbf{u}$ on each scale (cfr. [27, Lemma 4.4]).

Lemma 6.7. *Suppose*

$$|(D\mathbf{u})_{Q_R^\lambda}| \geqslant \frac{1}{2}\lambda \quad \textit{and} \quad \fint_{Q_R^\lambda} |D\mathbf{u} - (D\mathbf{u})_{Q_R^\lambda}|^2 \,\mathrm{d}z \leqslant \varepsilon\lambda^2$$

for some $\varepsilon \in (0,1)$*. Then for every* $\theta, \varepsilon_0 \in (0, \frac{1}{2})$*, we have*

$$\int_{Q_{\theta R}^\lambda} |D\mathbf{u} - (D\mathbf{u})_{Q_{\theta R}^\lambda}|^2 \,\mathrm{d}z \leqslant c_1(\varepsilon_0^{2\gamma_1} + \varepsilon_0^{-2\gamma}\varepsilon^\gamma + \theta^{n+4}) \int_{Q_R^\lambda} |D\mathbf{u} - (D\mathbf{u})_{Q_R^\lambda}|^2 \,\mathrm{d}z$$

for some $c_1 \geqslant 1$*, where* $\gamma > 0$ *is from Lemma 6.5.*

Proof. By Lemma 6.6 with $\mathbf{P} = (D\mathbf{u})_{Q_R^\lambda}$ and (6.12), we have that for every $\theta \in (0, \frac{1}{2})$,

$$\begin{aligned}\int_{Q_{\theta R}^\lambda} |D\mathbf{u} - (D\mathbf{u})_{Q_{\theta R}^\lambda}|^2 \,\mathrm{d}z &\leqslant 2\int_{Q_R^\lambda} |D\mathbf{u} - D\mathbf{v}|^2 \,\mathrm{d}z + 2\int_{Q_{\theta R}^\lambda} |D\mathbf{v} - (D\mathbf{v})_{Q_{\theta R}^\lambda}|^2 \,\mathrm{d}z \\ &\leqslant c\left[\varepsilon_0^{2\gamma_1} + \varepsilon_0^{-2\gamma}\varepsilon^\gamma\right] \int_{Q_R^\lambda} |D\mathbf{u} - (D\mathbf{u})_{Q_R^\lambda}|^2 \,\mathrm{d}z + c\theta^{n+4}\int_{Q_{R/2}^\lambda} |D\mathbf{v} - (D\mathbf{u})_{Q_R^\lambda}|^2 \,\mathrm{d}z\,.\end{aligned}$$

This concludes the proof. ∎

The counterpart of [27, Lemma 4.5] is the following result.

Lemma 6.8. *There exist small constants* $\theta, \varepsilon \in (0,1)$ *such if*

$$|(D\mathbf{u})_{Q_R^\lambda}| \geqslant \frac{3}{4}\lambda \quad \textit{and} \quad \fint_{Q_R^\lambda} |D\mathbf{u} - (D\mathbf{u})_{Q_R^\lambda}|^2 \,\mathrm{d}z \leqslant \varepsilon\lambda^2,$$

then for every $m \in \mathbb{N}$*,*

$$|(D\mathbf{u})_{Q_{R_m}^\lambda}| \geqslant \left(\frac{1}{2} + \frac{1}{2^{2+m}}\right)\lambda, \quad \textit{where} \quad R_m := \theta^m R,$$

and

$$\begin{aligned}\fint_{Q_{R_m}^\lambda} |D\mathbf{u} - (D\mathbf{u})_{Q_{R_m}^\lambda}|^2 \,\mathrm{d}z &\leqslant \theta^{\frac{3}{2}} \fint_{Q_{R_{m-1}}^\lambda} |D\mathbf{u} - (D\mathbf{u})_{Q_{R_{m-1}}^\lambda}|^2 \,\mathrm{d}z \\ &\leqslant \cdots \leqslant \theta^{\frac{3}{2}m} \fint_{Q_R^\lambda} |D\mathbf{u} - (D\mathbf{u})_{Q_R^\lambda}|^2 \,\mathrm{d}z.\end{aligned}$$

In particular, we have that for every $r \in (0, R]$*,*

$$\fint_{Q_r^\lambda} |D\mathbf{u} - (D\mathbf{u})_{Q_r^\lambda}|^2 \,\mathrm{d}z \leqslant \theta^{-n-\frac{7}{2}} \left(\frac{r}{R}\right)^{\frac{3}{2}} \fint_{Q_R^\lambda} |D\mathbf{u} - (D\mathbf{u})_{Q_R^\lambda}|^2 \,\mathrm{d}z\,.$$

Proof. Choose θ, ε and ε_0 small so that

$$\theta \leqslant \min\left\{(4c_1)^{-2}, 2^{-4/3}\right\}, \quad \varepsilon_0 \leqslant \left(\frac{\theta^{n+7/2}}{4c_1}\right)^{\frac{1}{2\gamma_1}}, \quad \varepsilon \leqslant \min\{\varepsilon_0^2[\theta^{n+7/2}/(4c_1)]^{1/\eta}, 2^{-6}\theta^{2n+4}\},$$

where the constant $c_1 \geqslant 1$ is from the preceding lemma. We prove the lemma by induction. We first obtain the desired estimates when $m = 1$. By the preceding lemma, we have

$$\fint_{Q^\lambda_{\theta R}} |D\mathbf{u} - (D\mathbf{u})_{Q^\lambda_{\theta R}}|^2 \,\mathrm{d}z \leqslant \theta^{3/2} \fint_{Q^\lambda_R} |D\mathbf{u} - (D\mathbf{u})_{Q^\lambda_R}|^2 \,\mathrm{d}z\,,$$

which is the desired second estimate when $m = 1$. Moreover,

$$|(D\mathbf{u})_{Q^\lambda_{\theta R}} - (D\mathbf{u})_{Q^\lambda_R}| \leqslant \theta^{-n-2} \fint_{Q^\lambda_R} |D\mathbf{u} - (D\mathbf{u})_{Q^\lambda_R}| \,\mathrm{d}z \leqslant \theta^{-n-2}\varepsilon^{1/2}\lambda \leqslant \frac{1}{8}\lambda\,,$$

hence

$$|(D\mathbf{u})_{Q^\lambda_{\theta R}}| \geqslant |(D\mathbf{u})_{Q^\lambda_R}| - \frac{1}{8}\lambda \geqslant \left(\frac{1}{2} + \frac{1}{2^3}\right)\lambda.$$

This is the first estimate when $m = 1$.

Next, we suppose the estimates hold for all $m \leqslant m_0$. Then from the first and the second inequalities when $m = m_0$ we see that

$$|(D\mathbf{u})_{Q^\lambda_{R_{m_0}}}| \geqslant \left(\frac{1}{2} + \frac{1}{2^{2+m_0}}\right)\lambda > \frac{1}{2}\lambda \ \text{ and } \fint_{Q^\lambda_{R_{m_0}}} |D\mathbf{u} - (D\mathbf{u})_{Q^\lambda_{R_{m_0}}}|^2 \,\mathrm{d}z \leqslant \varepsilon\theta^{\frac{3}{2}m_0}\lambda^2 < \varepsilon\lambda^2\,.$$

Therefore, applying the preceding lemma with R replaced by R_{m_0} and the second inequality when $m = m_0$, we see that

$$\fint_{Q^\lambda_{R_{m_0+1}}} |D\mathbf{u} - (D\mathbf{u})_{Q^\lambda_{R_{m_0+1}}}|^2 \,\mathrm{d}z \leqslant \theta^{3/2} \fint_{Q^\lambda_{R_{m_0}}} |D\mathbf{u} - (D\mathbf{u})_{Q^\lambda_{R_{m_0}}}|^2 \,\mathrm{d}z \leqslant \theta^{\frac{3}{2}(m_0+1)}\varepsilon\lambda^2.$$

This is the second estimate when $m = m_0 + 1$. Moreover, by the first estimate when $m = m_0$ and the upper bounds of θ and ε,

$$|(D\mathbf{u})_{Q^\lambda_{R_{m_0+1}}} - (D\mathbf{u})_{Q^\lambda_{R_{m_0}}}| \leqslant \frac{1}{\theta^{n+2}} \fint_{Q^\lambda_{R_{m_0}}} |D\mathbf{u} - (D\mathbf{u})_{Q^\lambda_{R_{m_0}}}| \,\mathrm{d}z \leqslant \frac{\theta^{\frac{3}{4}m_0}\varepsilon^{\frac{1}{2}}}{\theta^{n+2}}\lambda \leqslant \frac{1}{2^{3+m_0}}\lambda\,.$$

Therefore we have

$$|(D\mathbf{u})_{Q^\lambda_{R_{m_0+1}}}| \geqslant |(D\mathbf{u})_{Q^\lambda_{R_{m_0}}}| - \frac{1}{2^{3+m_0}}\lambda \geqslant \left(\frac{1}{2} + \frac{1}{2^{3+m_0}}\right)\lambda.$$

This is the first estimate when $m = m_0 + 1$, and the proof is concluded. □

We conclude the list of the auxiliary results needed to prove Proposition 6.2 with the following one, which corresponds to [27, Lemma 4.6]. For a $\sigma > 0$ small enough such that (6.2) holds, we have that the average of $D\mathbf{u}$ is comparable with λ and that $D\mathbf{u}$ remains close to its average.

Lemma 6.9. *For $\varepsilon > 0$, there exists $\sigma = \sigma(\varepsilon) \in (0, 2^{-(n+2)})$ such that if σ satisfies (6.2), then*

$$\frac{7}{8}\lambda \leqslant |(D\mathbf{u})_{Q^\lambda_{R/2}}| \leqslant \lambda$$

and

$$\fint_{Q^\lambda_{R/2}} |D\mathbf{u} - (D\mathbf{u})_{Q^\lambda_{R/2}}|^2 \leqslant \varepsilon\lambda^2.$$

Proof. Let $f(s) := (s-(1-2\theta)\lambda)_+$ with $\theta \in (0,1/4)$ to be determined and $F(s) := \int_0^s \tau f(\tau)\,\mathrm{d}\tau$. Note that we have that

(i) when $f(|D\mathbf{u}|) > 0$, $\frac{1}{2}\lambda < (1-2\theta)\lambda \leqslant |D\mathbf{u}| \leqslant \lambda$ (hence $2^{-q}\varphi(\lambda) \leqslant \varphi(|D\mathbf{u}|) \leqslant \varphi(\lambda)$) and $f'(|D\mathbf{u}|) = 1$;

(ii) $0 \leqslant f(|D\mathbf{u}|) \leqslant 2\theta\lambda$ and $0 \leqslant F(|D\mathbf{u}|) \leqslant 4\theta^2\lambda^3$.

Let $\xi \in C_0^\infty(B_R)$ and $\eta \in C^\infty(\mathbb{R})$ be cut-off functions such that $\xi \equiv 1$ in $B_{R/2}$, $|D\xi| \leqslant \frac{4}{R}$, $\eta \equiv 0$ in $(-\infty, -\frac{R^2}{\varphi''(\lambda)})$, $\eta \equiv 1$ in $(-\frac{R^2}{4\varphi''(\lambda)}, \infty)$ and $0 \leqslant \eta_t \leqslant \frac{8\varphi''(\lambda)}{R^2}$. Then, the Caccioppoli estimate (5.2) with $\rho = R$, $\tau_1 = -\frac{R^2}{\varphi''(\lambda)}$ and $\tau_2 = -\frac{R^2}{4\varphi''(\lambda)}$ yields

$$\begin{aligned}\int_{A((1-\theta)\lambda,R/2)} \varphi''(\lambda)\theta\lambda|D^2\mathbf{u}|^2\,\mathrm{d}z &\leqslant c\int_{Q^\lambda_{R/2}} \frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|} f(|D\mathbf{u}|)|D^2\mathbf{u}|^2\,\mathrm{d}z\\ &\leqslant cR^{-2}\int_{Q^\lambda_R}\left[\varphi'(|D\mathbf{u}|)\frac{f(|D\mathbf{u}|)^2}{f'(|D\mathbf{u}|)} + \varphi''(|D\mathbf{u}|)F(|D\mathbf{u}|)\right]\mathrm{d}z\\ &\leqslant cR^{-2}\big|\{|D\mathbf{u}| > (1-2\theta)\lambda\}\cap Q^\lambda_R\big|\theta^2\lambda\varphi(\lambda)\,,\end{aligned}$$

where $A(k,r) := \{z \in Q^\lambda_r : |D\mathbf{u}(z)| > k\}$. Note that in the second inequality, we used the fact that $\varphi''(\lambda) \leqslant 2^{q+2}\varphi''(|D\mathbf{u}|)$ when $F(|D\mathbf{u}|) > 0$. Hence we have

$$\int_{A((1-\theta)\lambda,R/2)} |D^2\mathbf{u}|^2\,\mathrm{d}z \leqslant c\,\theta R^{-2}\lambda^2\big|\{|D\mathbf{u}| > (1-2\theta)\lambda\}\cap Q^\lambda_R\big|\,. \tag{6.13}$$

Let $h_0 \in C^1(\mathbb{R})$ be increasing such that $h_0(t) = 0$ for $t \leqslant 3\lambda/4$, $h_0(t) = 1$ for $t > 7\lambda/8$, and $h_0' \leqslant 16\lambda^{-1}$, and set $\mathbf{h}(z) = h_0(|D\mathbf{u}(z)|)D\mathbf{u}(z)$. Then we have $|D\mathbf{h}|^2 \leqslant c|D^2\mathbf{u}|^2$. Let $\xi_0 \in C_0^\infty(B_{R/2})$ be a cut-off function such that $0 \leqslant \xi_0 \leqslant 1$ with $\xi_0 \equiv 1$ in $B_{R/4}$ and $|D^2\xi_0| + |D\xi_0|^2 \leqslant c/R^2$, and $\xi = \|\xi_0\|^{-1}_{L^1(B_{R/2})}\xi_0 \approx \frac{\xi_0}{R^n}$. Set

$$\mathbf{W}(t) := \int_{B_R} D\mathbf{u}(x,t)\xi(x)\,\mathrm{d}x\,,\quad \mathbf{W}_h(t) := \int_{B_R}\mathbf{h}(x,t)\xi(x)\,\mathrm{d}x$$

and

$$\mathbf{W}_0 = \frac{\varphi''(\lambda)}{R^2}\int_{-\frac{R^2}{\varphi''(\lambda)}}^0\int_{B_R} D\mathbf{u}(x,t)\xi(x)\,\mathrm{d}x\,\mathrm{d}t = \fint_{-\frac{R^2}{\varphi''(\lambda)}}^0 \mathbf{W}(t)\,\mathrm{d}t\,.$$

Note that by Lemma 6.4 with $f = \mathbf{h}(\cdot,t)$ and $p = \frac{2n}{n+1}$, we have

$$\begin{aligned}\iint_{Q^\lambda_{R/2}} |\mathbf{h}(x,t)-\mathbf{W}_h(t)|^2\,\mathrm{d}x\,\mathrm{d}t &\leqslant c\lambda^{\frac{2}{n+1}}\iint_{Q^\lambda_{R/2}}|\mathbf{h}(x,t)-\mathbf{W}_h(t)|^{\frac{2n}{n+1}}\,\mathrm{d}x\,\mathrm{d}t\\ &\leqslant c(|B_R|\lambda)^{\frac{2}{n+1}}\int_{Q^\lambda_{R/2}}|D\mathbf{h}|^{\frac{2n}{n+1}}\,\mathrm{d}z\,.\end{aligned}\tag{6.14}$$

Set

$$\Sigma_0 := A(3\lambda/4,R/2)\setminus A((1-\sigma)\lambda,R/2)\quad\text{and}\quad \Sigma := A((1-\sigma)\lambda,R/2)\,.$$

Since $|D\mathbf{h}| = 0$ on $Q^\lambda_{R/2}\setminus A(3\lambda/4,R/2)$, we have

$$\int_{Q^\lambda_{R/2}}|D\mathbf{h}|^{\frac{2n}{n+1}}\,\mathrm{d}z = \int_{\Sigma_0}|D\mathbf{h}|^{\frac{2n}{n+1}}\,\mathrm{d}z + \int_{\Sigma}|D\mathbf{h}|^{\frac{2n}{n+1}}\,\mathrm{d}z\,.$$

By Hölder's inequality, (6.2) and (6.13) with $\theta = \frac{1}{4}$, the first term on the right hand side can be estimated as

$$\int_{\Sigma_0}|D\mathbf{h}|^{\frac{2n}{n+1}}\,\mathrm{d}z \leqslant |\Sigma_0|^{\frac{1}{n+1}}\left(\int_{A(3\lambda/4,R/2)}|D\mathbf{h}|^2\,\mathrm{d}z\right)^{\frac{n}{n+1}} \leqslant c\sigma^{\frac{1}{n+1}}\lambda^{\frac{2n}{n+1}}R^{-\frac{2n}{n+1}}|Q^\lambda_R|\,.$$

Moreover, by (6.13) with $\theta = \sigma$

$$\begin{aligned}\int_{\Sigma} |D\mathbf{h}|^{\frac{2n}{n+1}}\,\mathrm{d}z &\leqslant |Q_R^\lambda|^{\frac{1}{n+1}}\left(\int_{A((1-\sigma)\lambda,R/2)} |D\mathbf{h}|^2\,\mathrm{d}z\right)^{\frac{n}{n+1}}\\ &\leqslant c\sigma^{\frac{n}{n+1}}\lambda^{\frac{2n}{n+1}}R^{-\frac{2n}{n+1}}|Q_R^\lambda|\,.\end{aligned}$$

Therefore, we have

$$\int_{Q_{R/2}^\lambda} |D\mathbf{h}|^{\frac{2n}{n+1}}\,\mathrm{d}z \leqslant c\sigma^{\frac{1}{n+1}}\lambda^{\frac{2n}{n+1}}R^{-\frac{2n}{n+1}}|Q_R^\lambda|\,,$$

hence, inserting this into (6.14),

$$\iint_{Q_{R/2}^\lambda} |\mathbf{h}(x,t) - \mathbf{W}_h(t)|^2\,\mathrm{d}x\,\mathrm{d}t \leqslant c\sigma^{\frac{1}{n+1}}\lambda^2|Q_R^\lambda|\,. \tag{6.15}$$

Note that

$$|D\mathbf{u} - \mathbf{w}_0|^2 \leqslant c\left(|D\mathbf{u}-\mathbf{h}|^2 + |\mathbf{h}-\mathbf{W}_h|^2 + |\mathbf{W}_h - \mathbf{W}|^2 + |\mathbf{W}-\mathbf{w}_0|^2\right).$$

First, by (6.2) with the definition of $\mathbf{h}$, we have

$$\int_{Q_{R/2}^\lambda} |D\mathbf{u}-\mathbf{h}|^2\,\mathrm{d}z = \int_{Q_{R/2}^\lambda\setminus\Sigma} (1-h_0(|D\mathbf{u}|))^2|D\mathbf{u}|^2\,\mathrm{d}z \leqslant \sigma\lambda^2|Q_R^\lambda|\,. \tag{6.16}$$

Moreover, by Hölder's inequality and the definition of ζ, we also have

$$\begin{aligned}\int_{Q_{R/2}^\lambda} |\mathbf{W}_h - \mathbf{W}|^2\,\mathrm{d}z &= |B_{R/2}|\int_{-R^2/(4\varphi''(\lambda))}^0 \left|\int_{B_{R/2}} (D\mathbf{u}(x,t) - \mathbf{h}(x,t))\xi(x)\,\mathrm{d}x\right|^2\mathrm{d}t\\ &\leqslant |B_{R/2}|\int_{-R^2/(4\varphi''(\lambda))}^0 \left[\int_{B_{R/2}} (D\mathbf{u}(x,t)-\mathbf{h}(x,t))^2\,\mathrm{d}x\right]\left[\int_{B_{R/2}}\xi^2\,\mathrm{d}x\right]\mathrm{d}t\\ &\leqslant c\int_{Q_{R/2}^\lambda} |D\mathbf{u}-\mathbf{h}|^2\,\mathrm{d}z \leqslant c\sigma\lambda^2|Q_R^\lambda|.\end{aligned} \tag{6.17}$$

Second, since $\mathbf{u} = \mathbf{u}_\varepsilon$ is a weak solution to (4.2), by testing (4.2) with $\zeta = (\xi_{x_i},\ldots,\xi_{x_i}) \in C_0^\infty(B_R,\mathbb{R}^N)$, $i = 1,2,\ldots,n$, we have that for every $-\frac{R^2}{\varphi''(\lambda)} < \tau < \tau' < 0$,

$$\int_{B_r} (u_{x_i}^\alpha(x,\tau') - u_{x_i}^\alpha(x,\tau))\xi(x)\,dx = \int_\tau^{\tau'}\int_{B_R} ([\mathbf{A}(D\mathbf{u})]^\alpha - [\mathbf{A}(\mathbf{W}(t)]^\alpha))\cdot D\xi_{x_i}\,\mathrm{d}z\,,$$

$\alpha = 1,2,\ldots,N$, hence, using (2.19) and the facts that $|D\mathbf{u}| \leqslant \lambda$ in Q_R^λ and $\varphi(t)/t^{p-1}$ is increasing,

$$\begin{aligned}|\mathbf{W}(\tau') - \mathbf{W}(\tau)| &\leqslant c\int_{Q_{R/2}^\lambda} |\mathbf{A}(D\mathbf{u}) - \mathbf{A}(\mathbf{W}(t))||D^2\xi|\,\mathrm{d}z\\ &\leqslant \frac{c}{R^{n+2}}\int_{Q_{R/2}^\lambda} \frac{\varphi'(|D\mathbf{u}|+|\mathbf{W}(t)|)}{|D\mathbf{u}|+|\mathbf{W}(t)|}(|D\mathbf{u}|+|\mathbf{W}(t)|)^{2-p}|D\mathbf{u}-\mathbf{W}(t)|^{p-1}\,\mathrm{d}z\\ &\leqslant c\frac{\varphi'(\lambda)}{\lambda^{p-1}R^{n+2}}\int_{Q_{R/2}^\lambda} |D\mathbf{u}-\mathbf{W}|^{p-1}\,\mathrm{d}z\\ &\leqslant c\lambda^{2-p}\left(\fint_{Q_{R/2}^\lambda} |D\mathbf{u}-\mathbf{W}|^2\,\mathrm{d}z\right)^{\frac{p-1}{2}}.\end{aligned} \tag{6.18}$$

Applying the estimates (6.15), (6.16) and (6.17), we obtain for every $-\frac{R^2}{\varphi''(\lambda)} < \tau < \tau' < 0$,

$$|\mathbf{W}(\tau') - \mathbf{W}(\tau)| \leqslant c\lambda^{2-p}\left(\sigma^{\frac{1}{n+1}}\lambda^2\right)^{\frac{p-1}{2}} \leqslant c\sigma^{\frac{p-1}{2(n+1)}}\lambda\,,$$

which also implies

$$\int_{Q_{R/2}^\lambda} |\mathbf{W} - \mathbf{w}_0|^2 \,\mathrm{d}z \leqslant c|Q_R^\lambda| \sup_{-\frac{R^2}{\varphi''(\lambda)}<\tau<\tau'<0} |\mathbf{W}(\tau') - \mathbf{W}(\tau)|^2 \leqslant c\sigma^{\frac{p-1}{n+1}}\lambda^2|Q_R^\lambda|\,. \tag{6.19}$$

Therefore, combining the results in (6.15)–(6.19), we have

$$\fint_{Q_{R/2}^\lambda} |D\mathbf{u} - (D\mathbf{u})_{Q_{R/2}^\lambda}|^2 \,\mathrm{d}z \leqslant \fint_{Q_{R/2}^\lambda} |D\mathbf{u} - \mathbf{w}_0|^2 \,\mathrm{d}z \leqslant c\sigma^{\frac{p-1}{n+1}}\lambda^2.$$

This implies the second desired estimate by choosing σ sufficiently small. Moreover, since $|D\mathbf{u}| \leqslant |D\mathbf{u} - (D\mathbf{u})_{Q_{R/2}^\lambda}| + |(D\mathbf{u})_{Q_{R/2}^\lambda}|$ and by the assumption of the lemma

$$\left|\{|D\mathbf{u}| > (1-\sigma)\lambda\} \cap Q_{R/2}^\lambda\right| > (1-\sigma)|Q_{R/2}^\lambda| > (1 - 2^{-(n+2)})|Q_{R/2}^\lambda|\,,$$

we have

$$\begin{aligned} |(D\mathbf{u})_{Q_{R/2}^\lambda}| &\geqslant \fint_{Q_{R/2}^\lambda} |D\mathbf{u}| \,\mathrm{d}z - \left(\fint_{Q_{R/2}^\lambda} |D\mathbf{u} - (D\mathbf{u})_{Q_{R/2}^\lambda}|^2 \,\mathrm{d}z\right)^{\frac{1}{2}} \\ &\geqslant \left[(1-\sigma)(1-2^{-(n+2)}) - c\sigma^{\frac{p-1}{2(n+1)}}\right]\lambda. \end{aligned}$$

Finally, by choosing σ sufficiently small we obtain the first desired estimate. □

We are now in position to prove Proposition 6.2.

Proof of Proposition 6.2. As we mentioned above, we assume $z_0 = 0$. Let $\varepsilon > 0$ be given from Proposition 6.8. With this ε, we determine σ as in Lemma 6.9 with $\frac{\varepsilon}{2}$ in place of ε, and suppose that $Q_R^\lambda = Q_R^\lambda(0)$ satisfies (6.2). Therefore, we have

$$\frac{7}{8}\lambda \leqslant |(D\mathbf{u})_{Q_{R/2}^\lambda}| \leqslant \lambda$$

and

$$\fint_{Q_{R/2}^\lambda} |D\mathbf{u} - (D\mathbf{u})_{Q_{R/2}^\lambda}|^2 \,\mathrm{d}z \leqslant \frac{\varepsilon}{2}\lambda\,.$$

Choose $z_1 = (x_1, t_1) \in Q_R^\lambda$ such that $\max\{|x_1|, \sqrt{\varphi''(\lambda)|t_1|}\} < \varepsilon' R$, where $\varepsilon' \in (0, 1/2)$ is a small constant to be determined. Then $Q_{R/2}^\lambda(z_1) \subset Q_R^\lambda$ and $|Q_{R/2}^\lambda(z_1) \setminus Q_{R/2}^\lambda|, |Q_{R/2}^\lambda \setminus Q_{R/2}^\lambda(z_1)| \leqslant c\varepsilon'|Q_R^\lambda|$ for some $c > 0$ depending n if ε' is sufficiently small. Therefore, by (6.1) and the preceding two inequalities, we have

$$\begin{aligned} &\fint_{Q_{R/2}^\lambda(z_1)} |D\mathbf{u} - (D\mathbf{u})_{Q_{R/2}^\lambda(z_1)}|^2 \,\mathrm{d}z \\ &\leqslant \frac{1}{|Q_{R/2}^\lambda|}\int_{Q_{R/2}^\lambda(z_1)\setminus Q_{R/2}^\lambda} |D\mathbf{u} - (D\mathbf{u})_{Q_{R/2}^\lambda}|^2 \,\mathrm{d}z + \fint_{Q_{R/2}^\lambda} |D\mathbf{u} - (D\mathbf{u})_{Q_{R/2}^\lambda}|^2 \,\mathrm{d}z \\ &\leqslant \left(c\varepsilon' + \frac{\varepsilon}{2}\right)\lambda^2 \leqslant \varepsilon\lambda^2 \end{aligned}$$

and

$$\begin{aligned}
|(D\mathbf{u})_{Q^\lambda_{R/2}(z_1)}| &\geqslant |(D\mathbf{u})_{Q^\lambda_{R/2}}| - |(D\mathbf{u})_{Q^\lambda_{R/2}} - (D\mathbf{u})_{Q^\lambda_{R/2}(z_1)}| \\
&\geqslant \frac{7}{8}\lambda - \frac{1}{|Q^\lambda_{R/2}|}\int_{(Q^\lambda_{R/2}(z_1)\setminus Q^\lambda_{R/2})\cup(Q^\lambda_{R/2}\setminus Q^\lambda_{R/2}(z_1))} |D\mathbf{u}|\,\mathrm{d}z \\
&\geqslant \frac{7}{8}\lambda - c\varepsilon'\lambda \geqslant \frac{3}{4}\lambda\,.
\end{aligned}$$

provided that ε' is sufficiently small. Therefore, by Lemma 6.8 with $Q^\lambda_{R/2}(z_1)$ in place of Q^λ_R, we obtain that for every $z_1 \in Q^\lambda_{R/2}$ and every $0 < r < \frac{R}{2}$ we have

$$\fint_{Q^\lambda_r(z_1)} |D\mathbf{u} - (D\mathbf{u})_{Q^\lambda_r(z_1)}|^2\,dz \leqslant c\left(\frac{r}{R}\right)^{3/2} \fint_{Q^\lambda_{R/2}(z_1)} |D\mathbf{u} - (D\mathbf{u})_{Q^\lambda_{R/2}(z_1)}|^2\,\mathrm{d}z\,. \tag{6.20}$$

Consequently, the desired oscillation decay estimate (6.3) follows from (6.20) by the standard embedding argument by Campanato in the parabolic setting (see for instance [28, Lemma 4.3]). □

6.2. **Proof of Proposition 6.3.** We start with a density result from (6.4) (cfr. [27, Lemma 6.2]).

Lemma 6.10. *Suppose (6.4) holds for some $\sigma \in (0, \frac{1}{2})$. There exists $m_1 \in \mathbb{N}$ depending on σ such that*

$$\sup_{-\frac{\sigma R^2}{2\varphi''(\lambda)} \leqslant t \leqslant 0} |\{|D\mathbf{u}(x,t)| > (1 - 2^{-m_1}\sigma)\lambda\} \cap B_{\tilde\sigma R}| \leqslant \tilde\sigma |B_{\tilde\sigma R}|\,, \quad \text{where } \tilde\sigma := \left(1 - \frac{\sigma}{2}\right)^{\frac{1}{n+2}}.$$

Proof. Step 1. We first prove that there exists $t_1 \in (-\frac{R^2}{\varphi''(\lambda)}, -\frac{\sigma R^2}{2\varphi''(\lambda)})$ such that

$$|\{x \in B_R : |D\mathbf{u}(x,t_1)| > (1-\sigma)\lambda\}| \leqslant \left(1 - \frac{\sigma}{2}\right)|B_R|\,. \tag{6.21}$$

Set

$$I := \int_{-\frac{R^2}{\varphi''(\lambda)}}^{-\frac{\sigma R^2}{2\varphi''(\lambda)}} |\{x \in B_R : |D\mathbf{u}(x,t)| > (1-\sigma)\lambda\}|\,\mathrm{d}t\,.$$

Then by (6.4),

$$I \leqslant |\{|D\mathbf{u}| > (1-\sigma)\lambda\} \cap Q^\lambda_R| \leqslant (1-\sigma)\frac{R^2}{\varphi''(\lambda)}|B_R|\,.$$

On the other hand, by the mean value theorem for integrals, there exists $t_1 \in (-\frac{R^2}{\varphi''(\lambda)}, -\frac{\sigma R^2}{2\varphi''(\lambda)})$ such that

$$I = |\{x \in B_R : |D\mathbf{u}(x,t_1)| > (1-\sigma)\lambda\}|\left(1 - \frac{\sigma}{2}\right)\frac{R^2}{\varphi''(\lambda)}\,.$$

This inequality with $\frac{1-\sigma}{1-\sigma/2} < 1 - \frac{\sigma}{2}$ yields (6.21).

Step 2. Let

$$\Psi(s) := \ln_+\left(\frac{\sigma}{1 - s/\lambda + 2^{1-m}\sigma}\right), \qquad 0 \leqslant s \leqslant \lambda\,,$$

where $g_+(s) := \max\{g(s), 0\}$ and $m \in \mathbb{N}$ with $m \geqslant 2$, and set

$$f(s) := \frac{2\Psi(s)\Psi'(s)}{s}\,, \qquad 0 \leqslant s \leqslant \lambda\,.$$

Then we have the following straightforward properties for Ψ and f:

(i) $0 \leqslant \Psi(s) \leqslant (m-1)\ln 2$ and $\Psi(s) = 0$ if and only if $0 \leqslant s \leqslant (1 - \sigma + 2^{1-m}\sigma)\lambda$;

(ii) for $(1-\sigma+2^{1-m}\sigma)\lambda < s < \lambda$, $\Psi'(s) = \frac{1}{(1+2^{1-m}\sigma)\lambda - s}$ and $\Psi''(s) = \Psi'(s)^2$;

(iii) for $(1-\sigma+2^{1-m}\sigma)\lambda < s < \lambda$, $f'(s) = \frac{2(1+\Psi(s))\Psi'(s)^2}{s} - \frac{f(s)}{s} > 0$;

(iv) $F(s) := \int_0^s \tau f(\tau)\,\mathrm{d}\tau = \Psi(s)^2$.

Choose any $\tilde{t} \in \left[-\frac{\sigma R^2}{2\varphi''(\lambda)}, 0\right]$ and let t_1 be that given from (6.21). Then from the proof of Lemma 5.1 (applying (5.3) to I_2) with $\rho = R$, $\tau_1 = t_1$, $\tau_2 = \tilde{t}$ and $\eta \equiv 1$, we deduce that

$$
\begin{aligned}
&\int_{B_R} F(|D\mathbf{u}(x,\tilde{t})|)\xi^2\,\mathrm{d}x \\
&\qquad + \int_{t_1}^{\tilde{t}}\int_{B_R} \frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|}\left(f(|D\mathbf{u}|)|D^2\mathbf{u}|^2 + \frac{f'(|D\mathbf{u}|)}{|D\mathbf{u}|}\frac{|D(|D\mathbf{u}|^2)|^2}{4}\right)\xi^2\,\mathrm{d}x\mathrm{d}t \\
&\leqslant \int_{B_R} F(|D\mathbf{u}(x,t_1)|)\xi^2\,\mathrm{d}x + c\int_{t_1}^{\tilde{t}}\int_{B_R}\frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|}\big|D(|D\mathbf{u}|^2)\big|f(|D\mathbf{u}|)\xi|D\xi|\,\mathrm{d}x\mathrm{d}t\,,
\end{aligned} \tag{6.22}
$$

where $\xi \in C_0^\infty(B_R)$ is a cut-off function satisfying that $\xi \equiv 1$ on $B_{\tilde{\sigma}R}$ and $|D\xi| \leqslant \frac{2}{(1-\tilde{\sigma})R}$. For what concerns the left hand side, applying (iii) and (iv), we have

$$
\begin{aligned}
\text{(LHS of (6.22))} \geqslant &\int_{B_R}\Psi(|D\mathbf{u}(x,\tilde{t})|)^2\xi^2\,\mathrm{d}x \\
&+ \int_{t_1}^{\tilde{t}}\int_{B_R}\frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|}\frac{(\Psi(|D\mathbf{u}|)+1)\Psi'(|D\mathbf{u}|)^2}{|D\mathbf{u}|^2}\frac{|D(|D\mathbf{u}|^2)|^2}{2}\xi^2\,\mathrm{d}x\mathrm{d}t\,.
\end{aligned}
$$

On the other hand, as for the right hand side, Young's inequality with the definition of f and (iv) yields.

$$
\begin{aligned}
\text{(RHS of (6.22))} \leqslant &\int_{B_R}\Psi(|D\mathbf{u}(x,t_1)|)^2\xi^2\,\mathrm{d}x + c\int_{t_1}^{\tilde{t}}\int_{B_R}\frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|}\Psi(|D\mathbf{u}|)|D\xi|^2\,\mathrm{d}x\mathrm{d}t \\
&+ \frac{1}{4}\int_{t_1}^{\tilde{t}}\int_{B_R}\frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|}\frac{\Psi(|D\mathbf{u}|)\Psi'(|D\mathbf{u}|)^2}{|D\mathbf{u}|^2}\big|D(|D\mathbf{u}|^2)\big|^2\xi^2\,\mathrm{d}x\mathrm{d}t\,.
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
&\int_{B_{\tilde{\sigma}R}}\Psi(|D\mathbf{u}(x,\tilde{t})|)^2\,\mathrm{d}x \\
&\quad\leqslant \int_{B_R}\Psi(|D\mathbf{u}(x,t_1)|)^2\,\mathrm{d}x + \frac{c}{(1-\tilde{\sigma})^2R^2}\int_{\tilde{t}}^{t_2}\int_{B_R}\frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|}\Psi(|D\mathbf{u}|)\,\mathrm{d}x\mathrm{d}t\,.
\end{aligned} \tag{6.23}
$$

For the left hand side, we see that

$$
\begin{aligned}
\text{(LHS of (6.23))} &\geqslant \Psi((1-2^{-m}\sigma)\lambda)^2|\{|D\mathbf{u}(x,\tilde{t})| > (1-2^{-m}\sigma)\lambda\}\cap B_{\tilde{\sigma}R}| \\
&\geqslant ((m-2)\ln 2)^2|\{|D\mathbf{u}(x,\tilde{t})| > (1-2^{-m}\sigma)\lambda\}\cap B_{\tilde{\sigma}R}|\,.
\end{aligned}
$$

On the other hand, to estimate the right hand side, we apply Step 1 and (i) to get

$$
\begin{aligned}
(\text{RHS of } (6.23)) &\leqslant ((m-1)\ln 2)^2 |\{x \in B_R : |D\mathbf{u}(x,t_1)| > (1-\sigma+2^{1-m}\sigma)\lambda\}| \\
&\quad + \frac{c(m-1)\ln 2}{(1-\sigma)(1-\tilde{\sigma})^2}|B_R| \\
&\leqslant \left\{ ((m-1)\ln 2)^2 \frac{(1-\frac{\sigma}{2})}{\tilde{\sigma}^n} + \frac{c(m-1)}{\tilde{\sigma}^n} \right\} |B_{\tilde{\sigma}R}| \\
&\leqslant \left\{ ((m-1)\ln 2)^2 \tilde{\sigma}^2 + \frac{c(m-1)}{\tilde{\sigma}^n} \right\} |B_{\tilde{\sigma}R}| \,.
\end{aligned}
$$

Therefore, combining the above results, we have

$$
|\{|D\mathbf{u}(x,\tilde{t})| > (1-2^{-m}\sigma)\lambda\} \cap B_{\tilde{\sigma}R}| \leqslant \left\{ \left(\frac{m-1}{m-2}\right)^2 \tilde{\sigma}^2 + \frac{c_*(m-1)}{\tilde{\sigma}^n(m-2)^2} \right\} |B_{\tilde{\sigma}R}|.
$$

Finally, by choosing $m \in \mathbb{N}$ sufficiently large so that

$$
\left(\frac{m-1}{m-2}\right)^2 \leqslant \frac{1+\tilde{\sigma}}{2\tilde{\sigma}} \quad \text{and} \quad \frac{m-1}{(m-2)^2} \leqslant \frac{\tilde{\sigma}^{n+1}(1-\tilde{\sigma})}{2c_*}\,,
$$

we obtain the conclusion. □

We note that $\tilde{\sigma} > \frac{\sigma}{2}$ since $\sigma \in (0, \frac{1}{2})$. Now, we are ready for proving Proposition 6.3.

Proof of Proposition 6.3. We remark that constants c in the proof depend also on σ.

Step 1. Let

$$
f(t) := \frac{(t-(1-\nu_1)\lambda)_+^{\chi-1}}{t}\,, \quad \chi \geqslant 2,
$$

where $\nu_1 \in (0, \frac{1}{2})$ is a sufficiently small constant to be determined, and set $F(t) := \int_0^t s f(s)\,ds = \frac{1}{\chi}(t-(1-\nu_1)\lambda)_+^{\chi}$. Note that, since $f'(t) = \frac{(t-(1-\nu_1)\lambda)^{\chi-2}((\chi-2)t+(1-\nu_1)\lambda)}{t^2}$ if $t > (1-\nu_1)\lambda$, we have

$$
f'(t) \geqslant \frac{(t-(1-\nu_1)\lambda)_+^{\chi-2}(1-\nu_1)}{\lambda} \geqslant \frac{(t-(1-\nu_1)\lambda)_+^{\chi-2}}{2\lambda} \quad \text{for} \quad 0 < t \leqslant \lambda,
$$

and

$$
\frac{f(t)^2}{f'(t)} = \frac{(t-(1-\nu_1)\lambda)_+^{\chi}}{(\chi-2)t+(1-\nu_1)\lambda} \leqslant \frac{(t-(1-\nu_1)\lambda)_+^{\chi}}{(1-\nu_1)\lambda} \leqslant 2\frac{(t-(1-\nu_1)\lambda)_+^{\chi}}{\lambda}\,.
$$

Let $\xi_0 \in C_0^\infty(B_{\tilde{\sigma}R})$ be a cut-off function with $0 \leqslant \xi_0 \leqslant 1$, $\xi_0 \equiv 1$ on $B_{\tilde{\sigma}R/2}$ and $|D\xi_0| \leqslant \frac{4}{\tilde{\sigma}R}$, and $\eta_0 \in C^\infty(\mathbb{R})$ with $0 \leqslant \eta_0 \leqslant 1$, $\eta_0 \equiv 0$ in $(-\infty, -\frac{\sigma R^2}{2\varphi''(\lambda)}]$, $\eta_0 \equiv 1$ in $[-\frac{\sigma R^2}{4\varphi''(\lambda)}, \infty)$ and $0 \leqslant (\eta_0)_t \leqslant \frac{8\varphi''(\lambda)}{\sigma R^2}$. Set $\xi = \xi_0^{\frac{(n+2)\chi-n}{2}}$, $\eta = \eta_0^{(n+2)\chi-n}$. Then applying (5.2) with $\rho = \tilde{\sigma}R$, $\tau_1 = t_1$ and $\tau_2 = t' \in (-\frac{\sigma R^2}{2\varphi''(\lambda)}, 0)$, and using the fact that $\frac{1}{2}\lambda \leqslant |D\mathbf{u}| \leqslant \lambda$

when $|D\mathbf{u}|>(1-\nu_1)\lambda$, we have

$$\begin{aligned}
&\frac{1}{\chi}\int_{B_{\tilde\sigma R}} w(x,t')^\chi\zeta(x,t')^{(n+2)\chi-n}\,\mathrm{d}x+\varphi''(\lambda)\int_{t_1}^{t'}\int_{B_{\tilde\sigma R}} w^{\chi-2}\frac{|D(|D\mathbf{u}|^2)|^2}{|D\mathbf{u}|^2}\xi^2\eta\,\mathrm{d}x\mathrm{d}t\\
&\leqslant c\int_{t_1}^{t'}\int_{B_{\tilde\sigma R}}\left[\varphi''(\lambda)w^\chi|D\xi|^2\eta+\frac{1}{\chi}w^\chi\xi^2\eta_t\right]\mathrm{d}x\mathrm{d}t\\
&\leqslant c\chi\int_{t_1}^{t'}\int_{B_{\tilde\sigma R}}\left[\varphi''(\lambda)w^\chi|D\xi_0|^2\zeta^{(n+2)\chi-n-2}+w^\chi(\eta_0)_t\zeta^{(n+2)\chi-n-1}\right]\mathrm{d}x\mathrm{d}t\\
&\leqslant c\chi\frac{\varphi''(\lambda)}{R^2}\int_{t_1}^{t'}\int_{B_{\tilde\sigma R}} w^\chi(\zeta^{(n+2)\chi-n-2}+\zeta^{(n+2)\chi-n-1})\,\mathrm{d}x\mathrm{d}t\,,
\end{aligned}$$

where we denote

$$w:=(|D\mathbf{u}|-(1-\nu_1)\lambda)_+\quad\text{and}\quad\zeta:=\xi_0\eta_0.$$

Moreover, since $D(w^{\frac{\chi}{2}}\xi)=\frac{\chi}{4}w^{\frac{\chi-2}{2}}\frac{D(|D\mathbf{u}|^2)}{|D\mathbf{u}|}\xi+w^{\frac{\chi}{2}}D\xi$ and $0\leqslant\zeta\leqslant 1$,

$$\begin{aligned}
\sup_{-\frac{\sigma R^2}{2\varphi''(\lambda)}<t<0}\int_{B_{\tilde\sigma R}} w^\chi\zeta^{(n+2)\chi-n}\,\mathrm{d}x+\varphi''(\lambda)\int_{-\frac{\sigma R^2}{2\varphi''(\lambda)}}^0\int_{B_{\tilde\sigma R}}|D(w^{\frac{\chi}{2}}\zeta^{\frac{(n+2)\chi-n}{2}})|^2\,\mathrm{d}x\mathrm{d}t&\\
\leqslant c\chi^3\frac{\varphi''(\lambda)}{R^2}\int_{-\frac{\sigma R^2}{2\varphi''(\lambda)}}^0\int_{B_{\tilde\sigma R}} w^\chi\zeta^{(n+2)\chi-n-2}\,\mathrm{d}x\mathrm{d}t\,.&
\end{aligned}\tag{6.24}$$

Setting

$$\bar w:=w\zeta^{n+2}\,,$$

by Hölder's inequality and Sobolev's inequality we have

$$\begin{aligned}
&\int_{-\frac{\sigma R^2}{2\varphi''(\lambda)}}^0\int_{B_{\tilde\sigma R}}\bar w^{\chi(1+\frac{2}{n})}\zeta^{-n-2}\,\mathrm{d}x\mathrm{d}t\\
&\leqslant\int_{-\frac{\sigma R^2}{2\varphi''(\lambda)}}^0\left(\int_{B_{\tilde\sigma R}}\bar w^\chi\zeta^{-n}\,\mathrm{d}x\right)^{\frac{2}{n}}\left(\int_{B_{\tilde\sigma R}}\left(\bar w^{\frac{\chi}{2}}\zeta^{\frac{-n}{2}}\right)^{\frac{2n}{n-2}}\,\mathrm{d}x\right)^{\frac{n-2}{n}}\mathrm{d}t\\
&\leqslant c\sup_{-\frac{\sigma R^2}{2\varphi''(\lambda)}<t<0}\left(\int_{B_{\tilde\sigma R}}\bar w^\chi\zeta^{-n}\,\mathrm{d}x\right)^{\frac{2}{n}}\int_{-\frac{\sigma R^2}{2\varphi''(\lambda)}}^0\int_{B_{\tilde\sigma R}}|D(\bar w^{\frac{\chi}{2}}\zeta^{\frac{-n}{2}})|^2\,\mathrm{d}x\mathrm{d}t\\
&\leqslant c\varphi''(\lambda)^{\frac{2}{n}}\left(\frac{\chi^3}{R^2}\int_{-\frac{\sigma R^2}{2\varphi''(\lambda)}}^0\int_{B_{\tilde\sigma R}}\bar w^\chi\zeta^{-n-2}\,\mathrm{d}x\mathrm{d}t\right)^{1+\frac{2}{n}}.
\end{aligned}$$

We further set $\tilde Q:=B_{\tilde\sigma R}\times(-\frac{\sigma R^2}{2\varphi''(\lambda)},0]$. Then we have

$$⨍_{\tilde Q}\bar w^{\chi(1+\frac{2}{n})}\zeta^{-n-2}\,\mathrm{d}z\leqslant c\chi^{3(1+\frac{2}{n})}\left(⨍_{\tilde Q}\bar w^\chi\zeta^{-n-2}\,\mathrm{d}z\right)^{1+\frac{2}{n}},\quad\chi\geqslant 2\,.$$

We now apply Moser's iteration. For $m=0,1,2,\dots$, we set

$$\chi_m=2\theta^m,\quad\text{where}\ \ \theta:=1+\frac{2}{n},$$

and

$$J_m:=⨍_{\tilde Q}\bar w^{\chi_m}\zeta^{-n-2}\,\mathrm{d}z$$

Then we have that for $m = 2, 3, \dots,$

$$J_m \leqslant c\chi_{m-1}^{\theta} J_{m-1}^{3\theta} \leqslant c^{1+\theta}\theta^{3\theta}\chi_{m-2}^{3\theta+3\theta^2} J_{m-2}^{\theta^2} \leqslant \cdots \leqslant c^{\frac{\theta^m-1}{\theta-1}}\theta^{\frac{3\theta}{\theta-1}\left(\frac{\theta(\theta^{m-1}-1)}{\theta-1}-1\right)} 2^{\frac{\theta(\theta^m-1)}{\theta-1}} J_0^{\theta^m} \leqslant (cJ_0)^{\theta^m},$$

which together with $\bar{w} = w\zeta^{n+2}$ yields

$$\|w\|_{L^\infty(\tilde{Q})} = \|\bar{w}\zeta^{-n-2}\|_{L^\infty(\tilde{Q})} = \lim_{m\to\infty} J_m^{\frac{1}{\chi_m}} \leqslant c\left(\fint_{\tilde{Q}} \bar{w}^2\zeta^{-n-2}\,\mathrm{d}z\right)^{\frac12} \leqslant c\left(\fint_{\tilde{Q}} w^2\,\mathrm{d}z\right)^{\frac12},$$

and by the definitions of w and $\tilde{Q} = B_{\tilde{\sigma}R} \times (-\frac{\sigma R^2}{2\varphi''(\lambda)}, 0] \supset Q^{\lambda}_{\sigma R/2}$,

$$\begin{aligned}\|(|D\mathbf{u}| - (1-\nu_1)\lambda)_+\|_{L^\infty(Q^{\lambda}_{\sigma R/2})} &\leqslant c\left(\fint_{\tilde{Q}} w^2\,\mathrm{d}z\right)^{\frac12}\\ &\leqslant c\nu_1\lambda\left(\frac{|\{(|D\mathbf{u}| - (1-\nu_1)\lambda)_+ > 0\} \cap \tilde{Q}|}{|\tilde{Q}|}\right)^{\frac12}.\end{aligned} \tag{6.25}$$

Step 2. Now, we show that the ratio of measures on the right hand side of (6.25) is sufficiently small if ν_1 is small. We follow the argument in [11, Lemma 6.4]. We start recalling the following well-known Poincaré-type inequality, see for instance [11, Lemma 6.3]: for $v \in W^{1,2}(B_r)$ and $k < l$,

$$(l-k)|\{v > l\} \cap B_r|^{1-\frac1n} \leqslant c\frac{r^n}{|B_r \setminus \{v > k\}|}\int_{\{k<v\leqslant l\}\cap B_r} |Dv|\,\mathrm{d}x\,. \tag{6.26}$$

Let m_2 be a positive number determined later, which is larger than the constant m_1 determined in Lemma 6.10, and $j \in \{m_1, m_1+1, \dots, m_2-1\}$. Putting $l = (1-\frac{\sigma}{2^{j+1}})\lambda$, $k = (1-\frac{\sigma}{2^j})\lambda$, $v = |D\mathbf{u}(\cdot,t)|$ and $r = \tilde{\sigma}R$, and applying Lemma 6.10 we have that for $t \in [-\frac{\sigma R^2}{\varphi''(\lambda)}, 0]$,

$$\begin{aligned}&\frac{\sigma\lambda}{2^{j+1}}|\{|D\mathbf{u}(x,t)| > (1-\frac{\sigma}{2^{j+1}})\lambda\} \cap B_{\tilde{\sigma}R}|\\ &\leqslant c\frac{R^{n+1}}{|B_{\tilde{\sigma}R} \setminus \{|D\mathbf{u}(x,t)| > (1-\frac{\sigma}{2^j})\lambda\}|}\int_{\{k<|D\mathbf{u}|(x,t)\leqslant l\}\cap B_{\tilde{\sigma}R}} |D[|D\mathbf{u}(x,t)|]|\,\mathrm{d}x\\ &\leqslant c\frac{R}{1-(1-\frac{\sigma}{2})^{1/(n+2)}}\int_{\{k<|D\mathbf{u}(x,t)|\leqslant l\}\cap B_{\tilde{\sigma}R}} |D[|D\mathbf{u}(x,t)|]|\,\mathrm{d}x\\ &\leqslant cR\int_{\{k<|D\mathbf{u}(x,t)|\leqslant l\}\cap B_{\tilde{\sigma}R}} |D[|D\mathbf{u}(x,t)|]|\,\mathrm{d}x.\end{aligned}$$

In addtion, by Hölder's inequality,

$$\begin{aligned}&\int_{\{k<|D\mathbf{u}(x,t)|\leqslant l\}\cap B_{\tilde{\sigma}R}} |D[|D\mathbf{u}(x,t)|]|\,\mathrm{d}x\\ &\leqslant \left(\int_{\{k<|D\mathbf{u}(x,t)|\leqslant l\}\cap B_{\tilde{\sigma}R}} \left|D\left(|D\mathbf{u}(x,t)| - \left(1-\frac{\sigma}{2^j}\right)\lambda\right)_+\right|^2 \mathrm{d}x\right)^{\frac12} |\{k < |D\mathbf{u}(x,t)| \leqslant l\} \cap B_{\tilde{\sigma}R}|^{\frac12}\end{aligned}$$

Therefore inserting this into the above estimate, integrating both the sides for t from $-\frac{\sigma R^2}{2\varphi''(\lambda)}$ to 0 and using Hölder's inequality, we have

$$\begin{aligned}&\frac{\sigma\lambda}{2^{j+1}}|\{|D\mathbf{u}|>(1-\frac{\sigma}{2^{j+1}})\lambda\}\cap\tilde{Q}|\\&\leqslant cR\left(\int_{\{k<|D\mathbf{u}|\leqslant l\}\cap\tilde{Q}}\left|D\left(|D\mathbf{u}|-\left(1-\frac{\sigma}{2^j}\right)\lambda\right)_+\right|^2\,\mathrm{d}z\right)^{\frac{1}{2}}|\{k<|D\mathbf{u}|\leqslant l\}\cap\tilde{Q}|^{\frac{1}{2}}.\end{aligned}$$

Moreover, using almost the same analysis as the one employed to derive (6.24), we also obtain

$$\begin{aligned}\int_{\{k<|D\mathbf{u}|\leqslant l\}\cap\tilde{Q}}\left|D\left(|D\mathbf{u}|-\left(1-\frac{\sigma}{2^j}\right)\lambda\right)_+\right|^2\,\mathrm{d}z&\leqslant\int_{\tilde{Q}}\left|D\left(|D\mathbf{u}|-\left(1-\frac{\sigma}{2^j}\right)\lambda\right)_+\right|^2\,\mathrm{d}z\\&\leqslant\frac{c}{(1-\tilde{\sigma})^2R^2}\int_{Q_R^\lambda}\left(|D\mathbf{u}|-\left(1-\frac{\sigma}{2^j}\right)\lambda\right)_+^2\,\mathrm{d}z\,.\end{aligned}$$

Combining the above results with the facts that $|D\mathbf{u}|\leqslant\lambda$ in Q_R^λ and $|Q_R^\lambda|\leqslant\frac{c(n)}{\sigma^n}|\tilde{Q}|$, we have

$$\left(\frac{\left|\left\{|D\mathbf{u}|>(1-\frac{\sigma}{2^{j+1}})\lambda\right\}\cap\tilde{Q}\right|}{|\tilde{Q}|}\right)^2\leqslant c\frac{\left|\left\{(1-\frac{\sigma}{2^j})\lambda<|D\mathbf{u}|\leqslant(1-\frac{\sigma}{2^{j+1}})\lambda\right\}\cap\tilde{Q}\right|}{|\tilde{Q}|}\,.$$

Then summing over j from m_1 to m_2-1,

$$(m_2-m_1)\left(\frac{|\{|D\mathbf{u}|>(1-\frac{\sigma}{2^{m_2}})\lambda\}\cap\tilde{Q}|}{|\tilde{Q}|}\right)^2\leqslant c\,,$$

which implies

$$\frac{|\{|D\mathbf{u}|>(1-v_1)\lambda\}\cap\tilde{Q}|}{|\tilde{Q}|}=\frac{|\{|D\mathbf{u}|>(1-\frac{\sigma}{2^{m_2}})\lambda\}\cap\tilde{Q}|}{|\tilde{Q}|}\leqslant\frac{c}{\sqrt{m_2-m_1}}\,,$$

where we choose

$$\nu_1=\frac{\sigma}{2^{m_2}}\,.$$

Step 3. We insert the previous estimate into (6.25) to get

$$\|(|D\mathbf{u}|-(1-\nu_1)\lambda)_+\|_{L^\infty(Q_{\sigma R/2}^\lambda)}\leqslant\frac{c_\sigma\nu_1\lambda}{(m_2-m_1)^{1/4}}$$

for some $c_\sigma>0$ depending on n,N,p,q and σ. At this stage, we determine m_2 depending on n,N,p,q and σ such that $\frac{c_\sigma}{(m_2-m_1)^{1/4}}\leqslant\frac{1}{2}$, hence $\nu_1=\frac{\sigma}{2^{m_2}}$ is also determined. Then we obtain

$$\|(|D\mathbf{u}|-(1-\nu_1)\lambda)_+\|_{L^\infty(Q_{\sigma R/2}^\lambda)}\leqslant\frac{\nu_1}{2}\lambda\,,$$

which implies (6.5) with $\nu:=1-\frac{\nu_1}{2}$ since $Q_{\sigma R/2}^\lambda\subset\tilde{Q}$. □

6.3. **Proof of Theorem 6.1.** We are now in position to prove the main result, Theorem 6.1. Arguing as in [27, Corollary 1.2], it will be a consequence of the following claim.

Claim. Suppose that $|D\mathbf{u}| \leqslant \lambda$ *in some* $Q_R^\lambda = Q_R^\lambda(z_0) \Subset \Omega_T$ *and* $\lambda > 0$*. Then, for every* $r \in (0, R)$ *it holds that*

$$\operatorname*{osc}_{Q_r^\lambda(z_0)} D\mathbf{u} \leqslant c\left(\frac{r}{R}\right)^\alpha \lambda \tag{6.27}$$

for some $\alpha \in (0,1)$.

We assume that $z_0 = (0,0)$ for simplicity. Fix $\sigma \in (0, 2^{-n-1})$ in Proposition 6.2. With this σ, choose $\nu \in (0,1)$ as in Proposition 6.3. If the assumption (6.2) holds, then (6.27) follows from Proposition 6.2. Hence, we assume that (6.2) does not hold, which means that (6.4) holds.

Choose $\theta \in (0,1)$ sufficiently small so that

$$\theta \leqslant \frac{\sigma\nu^{\frac{q}{2}-1}}{2} \leqslant \frac{\sigma}{2} \quad \text{and} \quad \theta \leqslant \left(\frac{\nu}{C}\right)^{\frac{4}{3}},$$

where $C \geqslant 1$ is from Proposition 6.2 and for $m \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, set $R_m := \theta^m R$ and $\lambda_m := \nu^m \lambda$. Then one has $Q_{R_{m+1}}^{\lambda_{m+1}} \subset Q_{\sigma R_m/2}^{\lambda_m}$ for every j. This implies $\{Q_{R_{m+1}}^{\lambda_{m+1}}\}_m$ is shrinking. Define

$$\mathcal{N} := \{m \in \mathbb{N}_0 : (6.2) \text{ holds with } R_m \text{ and } \lambda_m \text{ in place of } R \text{ and } \lambda\},$$

and

$$m_0 := \begin{cases} \min \mathcal{N} & \text{if } \ \mathcal{N} \neq \emptyset \\ \infty & \text{if } \ \mathcal{N} = \emptyset. \end{cases}$$

Then $m_0 \geqslant 1$. If $1 \leqslant m \leqslant m_0$, then by Proposition 6.4 with R_{m-1} and λ_{m-1} in place of R and λ, we have $|D\mathbf{u}| \leqslant \lambda_m = \nu^m\lambda$ in $Q_{R_m}^{\lambda_m}$ and

$$\begin{cases} \operatorname*{osc}_{Q_{R_m}^{\lambda_m}} (D\mathbf{u}) \leqslant 2\|D\mathbf{u}\|_{L^\infty(Q_{R_m}^{\lambda_m}, \mathbb{R}^{Nn})} \leqslant 2\nu^m\lambda \\ \text{for all } \ m = 0,1,2,\dots,m_0 \ \text{ when } \ m_0 < \infty, \ \text{ or } \ m \in \mathbb{N}_0 \ \text{ when } \ m_0 = \infty. \end{cases} \tag{6.28}$$

Furthermore, if $m_0 < \infty$, by Proposition 6.2 with R_{m_0} and λ_{m_0} in place of R and λ, the second condition of θ in above and (6.28) with $m = m_0$, we have

$$\operatorname*{osc}_{Q_{R_m}^{\lambda_{m_0}}} (D\mathbf{u}) \leqslant C\theta^{(m-m_0)3/4} \operatorname*{osc}_{Q_{R_{m_0}}^{\lambda_{m_0}}} (D\mathbf{u}) \leqslant 2\nu^m\lambda \quad \text{for all } \ m > m_0 \ \text{ when } \ m_0 < \infty. \tag{6.29}$$

Fix any $r \in (0, R/2]$. Let $t_1 \in (-\frac{r^2}{4\varphi''(\lambda)}, 0]$. Note that $Q_{R/2}^\lambda(0,t_1) \subset Q_R^\lambda$ and $\theta^{m+1}R/2 \leqslant r < \theta^m R/2$ for some $m \in \mathbb{N}_0$. Then applying (6.28) and (6.29) for $Q_{R/2}^\lambda(0,t_1)$ instead of Q_R^λ, we see that

$$|D\mathbf{u}(x,t_1) - D\mathbf{u}(0,t_1)| \leqslant 2\nu^m\lambda \leqslant c\left(\frac{r}{R}\right)^{\alpha_1}\lambda. \tag{6.30}$$

where $\alpha_1 = \log_\theta \nu$, for all $(x,t_1) \in Q_r^\lambda$.

Let $\xi_0 \in C_0^\infty(B_r)$ with $0 \leqslant \xi_0 \leqslant 1$, $\eta \equiv 1$ on $B_{r/2}$ and $|D^2\xi| + |D\xi|^2 \leqslant \frac{c}{r^2}$, $\xi = \|\xi_0\|_{L^1(B_r)}^{-1}\xi_0$, and

$$\mathbf{W}(t) := \int_{B_r} D\mathbf{u}(x,t)\xi(x)\,\mathrm{d}x.$$

Then by testing (4.2) with $\zeta = (\xi_{x_i}, \ldots, \xi_{x_i}) \in C_0^\infty(B_r, \mathbb{R}^N)$, $i = 1, 2, \ldots, n$ and applying the analysis in (6.18) along with the inequality $|D^2\xi| \leqslant \frac{c}{r^{n+2}}$, we have

$$|\mathbf{W}(t) - \mathbf{W}(0)| \leqslant \frac{c\varphi'(\lambda)}{\lambda^{p-1} r^{n+2}} \iint_{Q_r^\lambda} |D\mathbf{u}(y,s) - \mathbf{W}(s)|^{p-1} \,\mathrm{d}y\mathrm{d}s\,, \quad t \in \left(-\frac{r^2}{\varphi''(\lambda)}, 0\right].$$

Moreover, by (6.30) it follows that

$$|D\mathbf{u}(x,t) - \mathbf{W}(t)| \leqslant \int_{B_r} |D\mathbf{u}(x,t) - D\mathbf{u}(y,t)| \xi(y) \,\mathrm{d}y \leqslant c\left(\frac{r}{R}\right)^{\alpha_1} \lambda \quad \text{for every } (x,t) \in Q_r^\lambda\,.$$

Therefore, from the preceding two estimates we have that for every $(x,t) \in Q_r^\lambda$

$$\begin{aligned}|D\mathbf{u}(x,t) - D\mathbf{u}(0,0)| &\leqslant |D\mathbf{u}(x,t) - \mathbf{W}(t)| + |D\mathbf{u}(0,0) - \mathbf{W}(0)| + |\mathbf{W}(t) - \mathbf{W}(0)| \\ &\leqslant c\left(\frac{r}{R}\right)^{\alpha_1(p-1)} \lambda\,.\end{aligned}$$

This implies (6.27).

Acknowledgments

J. Ok was supported by the National Research Foundation of Korea by the Korean Government (NRF-2022R1C1C1004523). 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. E. Acerbi and N. Fusco, Regularity for minimizers of nonquadratic functionals: the case $1 < p < 2$, *J. Math. Anal. Appl.* **140** (1989), no. 1, 115–135.
2. P. Baroni and C. Lindfors, The Cauchy-Dirichlet problem for a general class of parabolic equations, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **34**(3) (2017), 593–624.
3. L. Bers, *athematical aspects of subsonic and transonic gas dynamics*, Surveys in Applied Mathematics, Vol. 3 John Wiley & Sons, Inc., New York; Chapman & Hall, Ltd., London 1958
4. V. Bögelein, F. Duzaar, N. Liao and C. Scheven, Gradient Hölder regularity for degenerate parabolic systems, *Nonlinear Anal.* **225** (2022), Paper No. 113119, 61 pp.
5. Y. Z. Chen, Hölder continuity of the gradient of solutions of nonlinear degenerate parabolic systems, Acta Math. Sinica (N.S.) **2** (1986), no. 4, 309–331.
6. Y. Z. Chen and E. DiBenedetto, Boundary estimates for solutions of nonlinear degenerate parabolic systems, *J. Reine Angew. Math.* **395** (1989), 102–131.
7. Y. Cho, Calderón-Zygmund theory for degenerate parabolic systems involving a generalized p-Laplacian type, *J. Evol. Equ.* **18** (2018), no. 3, 1229–1243.
8. H. Choe, Hölder regularity for the gradient of solutions of certain singular parabolic systems, *Comm. Partial Differential Equations* **16** (11) (1991), 1709–1732.
9. H. Choe, Hölder continuity for solutions of certain degenerate parabolic systems, *Nonlinear Anal. TMA* **18** (3) (1992), 235–243.
10. E. Di Benedetto, *Degenerate parabolic equations*, Universitext. Springer-Verlag, New York, 1993.
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. L. Diening and F. Ettwein, Fractional estimates for non-differentiable elliptic systems with general growth, *Forum Math.* **20** (3) (2008), 523–556.
14. 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.

15. 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.
16. 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.
17. L. Diening, B. Stroffolini and A. Verde, Everywhere regularity of functionals with φ-growth, *Manuscripta Math.* **129** (4) (2009), 449–481.
18. R. Finn and D. Gilbarg, Three-dimensional subsonic flows, and asymptotic estimates for elliptic partial differential equations, *Acta Math.* 98 (1957), 265–296.
19. M. Giaquinta and G. Modica, Remarks on the regularity of the minimizers of certain degenerate functionals, Manuscripta Math. **57** (1986), no. 1, 55–99.
20. E. Giusti, *Direct methods in the calculus of variations*, World Scientific Publishing Co., Inc., River Edge, NJ, 2003.
21. C. Hamburger, Regularity of differential forms minimizing degenerate elliptic functionals, *J. Reine Angew. Math.* **431** (1992), 7–64.
22. P. Harjulehto and P. Hästö, *Orlicz Spaces and Generalized Orlicz Spaces*, Lecture Notes in Mathematics, vol.2236, Springer, Cham, (2019).
23. P. Hästö and J. Ok, Higher integrability for parabolic systems with Orlicz growth, *J. Differ. Equ.* **300** (2021), 925–948.
24. S. Hwang and G. M. Lieberman, Hölder continuity of bounded weak solutions to generalized parabolic p-Laplacian equations I: degenerate case, *Electron. J. Differential Equations* **2015**, No. 287, 32 pp.
25. S. Hwang and G. M. Lieberman, Hölder continuity of bounded weak solutions to generalized parabolic p-Laplacian equations II: singular case, *Electron. J. Differential Equations* **2015**, No. 288, 24 pp.
26. T. Isernia, L^∞-regularity for a wide class of parabolic systems with general growth, *Proc. Amer. Math. Soc.* **146** (11) (2018), 4741–4753.
27. G. M. Lieberman, Hölder regularity for the gradients of solutions of degenerate parabolic systems, *Ukr. Mat. Visn.* **3** (3) (2006), 352–373.
28. G. M. Lieberman, Second order parabolic differential equations, World Scientific Publishing Co., Inc., River Edge, NJ, 1996.
29. C. Lindfors, Obstacle problem for a class of parabolic equations of generalized p-Laplacian type, *J. Differential Equations* **261** (2016), no. 10, 5499–5540.
30. P. Marcellini, Regularity of minimizers of integrals of the calculus of variations with nonstandard growth conditions, *Arch. Rational Mech. Anal.* **105** (1989), no. 3, 267–284.
31. P. Marcellini, Regularity and existence of solutions of elliptic equations with p, q-growth conditions, *J. Differential Equations* **90** (1991), no. 1, 1–30.
32. P. Marcellini and G. Papi, Nonlinear elliptic systems with general growth *J. Differential Equations* **221** (2006), no. 2, 412–443.
33. G. Mingione, Regularity of minima: an invitation to the dark side of the calculus of variations, *Appl. Math.* **51** (2006), no. 4, 355-426.
34. J. Nec as, Example of an irregular solution to a nonlinear elliptic system with analytic coefficients and conditions for regularity, Theor. Nonlin. Oper., Constr. Aspects. Proc. 4th Int. Summer School. Akademie-Verlag, Berlin, 1975, pp. 197–206.
35. J. Oh and J. Ok, Gradient estimates for parabolic problems with Orlicz growth and discontinuous coefficients, *Math. Methods Appl. Sci.* **45** (2022), no. 14, 8718–8736
36. V. Šverák, X. Yan, A singular minimizer of a smooth strongly convex functional in three dimensions, *Calc. Var. Partial Differ. Equ.*, **10** (2000), 213–221.
37. P. Tolksdorf, Everywhere-regularity for some quasilinear systems with a lack of ellipticity., *Ann. Mat. Pura Appl.* (4) **134** (1983), 241–266.
38. K. Uhlenbeck, Regularity for a class of non-linear elliptic systems, *Acta Math.* **138** (1977), no. 3-4, 219–240.
39. M. Wiegner, On C_α-regularity of the gradient of solutions of degenerate parabolic systems, *Ann. Mat. Pura Appl.* (4) **145** (1986), 385–405.

(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