# PARTIAL REGULARITY FOR PARABOLIC SYSTEMS OF DOUBLE PHASE TYPE

JIHOON OK, GIOVANNI SCILLA, AND BIANCA STROFFOLINI

ABSTRACT. We study partial regularity for nondegenerate parabolic systems of double phase type, where the growth function is given by $H(z, s) = s^p + a(z)s^q$, $z = (x, t) \in \Omega_T$, with $\frac{2n}{n+2} < p \leqslant q$ and $a(z)$ a nonnegative $C^{0,\alpha,\frac{\alpha}{2}}$-continuous function for some $\alpha \in (0, 1]$. As the main result we prove that if $q < \min\{p + \frac{\alpha p}{n+2}, p + 1\}$ the spatial gradient of any weak solution is locally Hölder continuous, except on a set of measure zero.

## CONTENTS

## 1. INTRODUCTION

Double phase energies of the form

$$(1.1) \qquad \int_\Omega H(x, D\mathbf{u})\, \mathrm{d}x\,, \quad H(x, s) := s^p + a(x)s^q\,,\ 1 < p < q\,,$$

have been originally introduced by Zhikov [64, 67] in the context of homogenization theory and Elasticity, with the aim of modeling strongly anisotropic materials, and then studied as an example exhibiting Lavrentiev phenomenon ([65, 66], see also [36, 37, 3]). In this framework, the coefficient $a(\cdot)$ plays a crucial role, since it describes the geometry of a composite made of two different materials exhibiting power-type hardening with exponents $p$ and $q$, respectively. The double phase setting naturally interpolates between two different growth conditions and thus provides a unifying model capturing nonuniform behaviors within a single variational framework. We further refer to [14, 41] for the application of the double phase problems to image restoration.

Equations and systems related to (1.1) have been the object of intensive study over the last decade, particularly in connection with regularity theory. For scalar minimizers of (1.1), sharp and essentially optimal regularity results were obtained in the works of Baroni, Colombo and Mingione

[16, 17, 4, 5], where they proved that, under the condition $q \leqslant p + \frac{\alpha p}{n}$, the gradient of minimizers is locally Hölder continuous. For further regularity results related to elliptic double phase type problems, we refer to [1, 2, 9, 10, 18, 19, 20, 21, 24, 25, 56, 59] and the references therein. We also mention [22, 23, 43, 42, 44, 45] for regularity results concerning more general energies, including certain double phase type functionals.

Regularity for elliptic systems of the form

$$\text{(1.2)} \qquad\qquad\qquad \operatorname{div} \mathbf{A}(x, D\mathbf{u}) = \mathbf{0} \quad \text{in } \Omega,$$

where $\Omega \subset \mathbb{R}^n$ $(n \geqslant 2)$ is a bounded and open set, $\mathbf{u} = (u^1, \ldots, u^N) : \Omega \to \mathbb{R}^N$ with $N > 1$, and the operator $\mathbf{A}(x, \boldsymbol{\xi})$ satisfies standard growth and ellipticity conditions, has been actively studied. It is well known that, without structural assumptions, solutions to systems exhibit only partial regularity, i.e., they are regular outside a set of measure zero, see the survey [55] and references therein. An efficient and widely used technique is the so-called $\mathcal{A}$-harmonic approximation, which involves linearizing the functional or partial differential equation around the average gradient in a ball, see [33, 28]. This method enables one to locally approximate solutions of nonlinear problems by solutions of linearized equations, facilitating regularity estimates. In contrast to the $\mathcal{A}$-harmonic approximation, which is suitable when the gradient is bounded away from zero, the regime where the gradient is small requires approximation schemes that reflect the degeneracy or singularity of the underlying operator, see [34, 32, 13], respectively for the power case and the Orlicz one.

More recently, partial regularity for nondegenerate systems of the type (1.2) with operators $\mathbf{A}(x, \boldsymbol{\xi})$ associated with the double phase function $H(x, s) = s^p + a(x)s^q$ has been investigated in [57, 58, 62]. Here "nondegenerate" refers to the condition $|D_{\boldsymbol{\xi}} \mathbf{A}(x, 0)| \sim 1$. Those works deal only with the superquadratic case, namely $p \geqslant 2$. The general (possibly degenerate) case of double phase systems was recently treated in [61], where partial Hölder regularity of the gradient of weak solutions was obtained without the superquadratic assumption, thus providing a unified approach independent of the exponent $p$.

For the corresponding parabolic systems

$$\text{(1.3)} \qquad\qquad\qquad \mathbf{u}_t - \operatorname{div} \mathbf{A}(z, D\mathbf{u}) = \mathbf{0} \quad \text{in } \Omega_T,$$

where $z = (x, t)$ and $\Omega_T = \Omega \times (0, T)$ is a space-time cylinder in $\mathbb{R}^{n+1}$, partial regularity in the $p$-growth case has been studied in [35, 8]. Moreover, in recent years, partial regularity in the Orlicz growth setting has also been investigated in [38, 60].

In contrast to the elliptic setting, where the picture of regularity results is rich and appears to be almost completely understood, the corresponding parabolic theory for double phase equations and systems remains much less developed. Chlebicka, Gwiazda and Zatorska-Goldstein [15] proved the existence of weak solutions to parabolic equations with Musielak–Orlicz growth, and Kim, Kinnunen and Särkiö [50] proved existence and uniqueness results for weak solutions to parabolic systems with superquadratic double phase growth. Concerning regularity theory, Kim, Kinnunen, Moring and Särkiö [47, 48, 49, 52] have established higher integrability for the gradient as well as Calderón–Zygmund type estimates for weak solutions. These results were obtained under the conditions

$$\text{(1.4)} \qquad q \leqslant p + \frac{2\alpha}{n+2} \quad \text{when } p \geqslant 2, \qquad q \leqslant p + \frac{\alpha}{n+2} \frac{p(n+2) - 2n}{2} \quad \text{when } \frac{2n}{n+2} < p < 2.$$

Moreover, in [49] an additional technical assumption $D\mathbf{u} \in L^q(\Omega_T)$ was required; however, this can be removed by applying a double phase parabolic Lipschitz truncation, see [50]. We further refer to [51] for results on the Hölder continuity of bounded weak solutions to parabolic equations, [12] for parabolic double phase obstacle problems, and [6, 63] for related work on parabolic $(p, q)$-growth problems.

To the best of our knowledge, partial regularity for parabolic systems of double phase type has not yet been explored. In this paper, we address this problem for homogeneous nondegenerate systems of the form (1.3), where the nonlinearity $\mathbf{A} : \Omega_T \times \mathbb{R}^{N \times n} \to \mathbb{R}^{N \times n}$ satisfies the following growth and ellipticity conditions with respect to the function $H(z, s)$:

$$\text{(A1)} \qquad\qquad |\mathbf{A}(z, \boldsymbol{\xi})| + (1 + |\boldsymbol{\xi}|) |D_{\boldsymbol{\xi}} \mathbf{A}(z, \boldsymbol{\xi})| \leqslant L \, H'(z, 1 + |\boldsymbol{\xi}|),$$

(A2) $$[D_{\boldsymbol{\xi}}\mathbf{A}(z,\boldsymbol{\xi})\boldsymbol{\lambda}]:\boldsymbol{\lambda}\geqslant\nu\,H''(z,1+|\boldsymbol{\xi}|)\,|\boldsymbol{\lambda}|^2\,,$$

for every $z=(x,t)\in\Omega_T$, $\boldsymbol{\xi}\in\mathbb{R}^{N\times n}$, and $\boldsymbol{\lambda}\in\mathbb{R}^{N\times n}$, and for some $0<\nu\leqslant L$. Here, ":" denotes the Euclidean inner product in $\mathbb{R}^{N\times n}$, and the function $H:\Omega_T\times[0,\infty)\to[0,\infty)$ is the so-called double phase function defined by

(1.5) $$H(z,s):=s^p+a(z)\,s^q\,,\quad\text{where}\quad\frac{2n}{n+2}<p\leqslant q,\ \ 0\leqslant a(z)\leqslant L\ \text{for}\ z\in\Omega_T\,.$$

We further assume that $a\in C^{0,\alpha,\frac{\alpha}{2}}(\Omega_T)$ for some $\alpha\in(0,1]$, that is,

(1.6) $$|a(z_1)-a(z_2)|\leqslant L\big(|x_1-x_2|^\alpha+|t_1-t_2|^{\frac{\alpha}{2}}\big)$$

for every $z_1=(x_1,t_1)$ and $z_2=(x_2,t_2)$ in $\Omega_T$, and

(1.7) $$q<\min\left\{p+\frac{\alpha p}{n+2},\,p+1\right\}.$$

Moreover, $H'(z,s)$ and $H''(z,s)$ denote the first and second derivatives with respect to the $s$-variable, respectively. Note that in the region where $a(z)=0$, we have $H(z,s)=s^p$, so that $H$ corresponds to the $p$-phase, while in the region where $a(z)>0$, the function $H$ corresponds to the $(p,q)$-phase. We also note that (A2) implies the following monotonicity property: for every $z\in\Omega_T$ and $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2\in\mathbb{R}^{N\times n}$,

(1.8) $$\big(\mathbf{A}(z,\boldsymbol{\xi}_1)-\mathbf{A}(z,\boldsymbol{\xi}_2)\big):(\boldsymbol{\xi}_1-\boldsymbol{\xi}_2)\geqslant\tilde{\nu}\,H''(z,1+|\boldsymbol{\xi}_1|+|\boldsymbol{\xi}_2|)\,|\boldsymbol{\xi}_1-\boldsymbol{\xi}_2|^2$$

for some $\tilde{\nu}>0$ depending only on $p$, $q$, and $\nu$.

For system (1.3), we establish the partial $C^{1,\alpha}$-regularity theory; see Section 1.1 for the main theorem. We outline the novelty of our result and the main ideas of the proof. Our analysis is focused on non-degenerate systems, which represent the initial step in the study of partial regularity theory. However, even in this context, the proof of partial regularity necessitates the introduction of significantly novel concepts. In the nondegenerate setting, a natural condition on the exponents $p$ and $q$ in the definition of $H$ might be

(1.9) $$q\leqslant p+\frac{p\alpha}{n+2}\,,$$

as this allows one to avoid the use of intrinsic parabolic cylinders. We point out that, even in the elliptic case, in order to obtain regularity in the borderline situation $q=p+\frac{p\alpha}{n+2}$, higher integrability plays an essential role. In the parabolic setting, it seems that the conditions in (1.4) are essentially required to obtain higher integrability, and are more restrictive than (1.9). Therefore, assuming the strict inequality in (1.9) is quite natural.

A Caccioppoli type inequality was obtained in [50] for weak solutions $\mathbf{u}$ to the double phase parabolic systems satisfying the natural integrability condition $H(\cdot,|D\mathbf{u}|)\in L^1_{\mathrm{loc}}(\Omega_T)$, assuming the strong condition (1.4). On the other hand, in this paper, we are able to prove it under the optimal condition (1.9) by means of a mollification argument of convolution type in both space variable $z$ and the time variable $t$, originally devised for the Euclidean metric in [40] and later adapted to the parabolic setting in [46].

Thanks to our recent $\mathcal{A}$-caloric approximation result in [60], we can define the *excess functional* directly in terms of the gradient $D\mathbf{u}$ of a weak solution $\mathbf{u}$ to (1.3):

(1.10) $$\Phi(z_0,\rho,\mathbf{u})=\fint_{Q_\rho(z_0)}H^-_{1+|(D\mathbf{u})_{z_0,\rho}|}\big(|D\mathbf{u}-(D\mathbf{u})_{z_0,\rho}|\big)\,\mathrm{d}z+\rho^{\frac{\beta_1}{2}}\,,$$

where $H^-_\sigma$ denotes the shifted $N$-function of $H^-(s):=s^p+\inf_{z\in Q_\rho(z_0)}a(z)s^q$ and $\beta_1$ is defined in (4.5).

Then a linearization procedure combined with $\mathcal{A}$-caloric approximation yields local comparison estimates for $D\mathbf{u}$ with the gradient of a suitably chosen smooth $\mathcal{A}$-caloric function. We note that the additional assumption $q<p+1$ is imposed in order to apply the $\mathcal{A}$-caloric approximation to the function $H^-$, rather than $H$. This condition could, in principle, be relaxed if a higher integrability result for functions of the form $H_{1+|\mathbf{Q}|}(z,|D\mathbf{u}-\mathbf{Q}|)$ were to be established. However, as previously noted, this would likely require the conditions in (1.4). On the other hand, the

condition $q < p + 1$ is less restrictive than (1.4), and naturally arises in the regularity theory for parabolic problems with $(p, q)$-growth; see, for instance, [6, 7].

We then establish a decay estimate for the excess functional (1.10) in Lemma 5.2. Finally, by iterating this decay at smaller scales, as shown in Lemma 5.3, we obtain the partial regularity result in the standard manner, as stated in Section 6: the weak solution is locally Hölder continuous, except on a set of measure zero; see Theorem 1.2.

As a final remark, we highlight an important aspect of our approach. In most papers on double phase problems, different methods have been adopted for the $p$-phase and the $(p, q)$-phase. However, in this paper, the arguments of approximation, comparison, and iteration do not distinguish between the phases. This allows us to provide simpler and more unified proofs.

*Outline of the paper:* The paper is organized as follows. In Section 1.1, we present our main result, Theorem 1.2. In Section 2, we fix the necessary notation, and recall some basic facts and elementary inequalities about Orlicz and the generalized Orlicz function $H$. In Section 3, we obtain a Caccioppoli type estimate by using a new mollification argument. In Section 4, we collect two main ingredients in order to get the decay estimates: we recall regularity estimates for the solutions of systems with constant coefficients and the $\mathcal{A}$-caloric approximation, see Section 4.1, and then we setup the linearization, Section 4.2. Section 5 provides decay estimates for the excess functional (Lemma 5.2), which can be iterated on shrinking parabolic cylinders (Lemma 5.3). Finally, in Section 6, we prove our main result.

1.1. **Main result.** We introduce the main result of the paper. First, we list the required additional assumptions on $\mathbf{A}(z, \boldsymbol{\xi})$: we assume the following continuity condition on $z \mapsto \mathbf{A}(z, \boldsymbol{\xi})$

$$
\begin{aligned}
\text{(A3)} \quad |\mathbf{A}(z_1, \boldsymbol{\xi}) - \mathbf{A}(z_2, \boldsymbol{\xi})| &\leqslant L(|x_1 - x_2| + \sqrt{|t_1 - t_2|})^{\beta_0} \left(H'(z_1, 1 + |\boldsymbol{\xi}|) + H'(z_2, 1 + |\boldsymbol{\xi}|)\right) \\
&\quad + L|a(z_1) - a(z_2)|(1 + |\boldsymbol{\xi}|)^{q-1}
\end{aligned}
$$

for all $z_1 = (x_1, t_2)$ and $z_2 = (x_2, t_2)$ in $\Omega_T$ and for some $\beta_0 \in (0, 1)$, and the off diagonal condition on $z \mapsto D_{\boldsymbol{\xi}} \mathbf{A}(z, \boldsymbol{\xi})$

$$
\text{(A4)} \quad |D_{\boldsymbol{\xi}} \mathbf{A}(z, \boldsymbol{\xi}_1) - D_{\boldsymbol{\xi}} \mathbf{A}(z, \boldsymbol{\xi}_2)| \leqslant L \left(\frac{|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2|}{1 + |\boldsymbol{\xi}_1|}\right)^{\gamma} H''(z, 1 + |\boldsymbol{\xi}_1|) \quad \text{for some} \quad \gamma \in (0, 1),
$$

for all $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in \mathbb{R}^{N \times n}$ with $2|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2| \leqslant 1 + |\boldsymbol{\xi}_1|$.

We then specify the notion of weak solution to (1.3).

**Definition 1.1.** *A function* $\mathbf{u} = (u^1, u^2, \ldots, u^N) \in C_{\mathrm{loc}}(0, T; L^2_{\mathrm{loc}}(\Omega, \mathbb{R}^N)) \cap L^p_{\mathrm{loc}}(0, T; W^{1,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^N))$ *with* $H(\cdot, |D\mathbf{u}|) \in L^1_{\mathrm{loc}}(\Omega_T)$ *is said to be a* weak solution *to the system* (1.3) *if it satisfies the following weak form of* (1.3)*:*

$$
\text{(1.11)} \quad - \iint_{\Omega_T} \mathbf{u} \cdot \boldsymbol{\zeta}_t \, \mathrm{d}z + \iint_{\Omega_T} \mathbf{A}(z, D\mathbf{u}) : D\boldsymbol{\zeta} \, \mathrm{d}z = 0 \quad \text{for all} \quad \boldsymbol{\zeta} \in C^{\infty}_{\mathrm{c}}(\Omega_T, \mathbb{R}^N),
$$

*where "·" is the Euclidean inner product in* $\mathbb{R}^N$*.*

Finally, we are in position to state the main result of the paper.

**Theorem 1.2.** *Let* $H : \Omega_T \times [0, \infty) \to [0, \infty)$ *be defined as in* (1.5) *complying with* (1.6) *and* (1.7)*, and* $\mathbf{A} : \Omega_T \times \mathbb{R}^{N \times n} \to \mathbb{R}^{N \times n}$ *comply with* (A1)–(A4)*. If* $\mathbf{u} \in C_{\mathrm{loc}}(0, T; L^2_{\mathrm{loc}}(\Omega, \mathbb{R}^N)) \cap L^p_{\mathrm{loc}}(0, T; W^{1,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^N))$ *with* $H(\cdot, |D\mathbf{u}|) \in L^1_{\mathrm{loc}}(\Omega_T)$ *is a weak solution to* (1.3)*, then there exist* $\beta = \beta(n, p, q, \alpha, \beta_0) \in (0, 1)$ *and an open subset* $\Omega_0 \subset \Omega_T$ *such that* $|\Omega_T \setminus \Omega_0| = 0$ *and*

$$
D\mathbf{u} \in C^{0,\beta,\frac{\beta}{2}}_{\mathrm{loc}} \left(\Omega_0; \mathbb{R}^{N \times n}\right).
$$

*Moreover,* $\Omega_T \setminus \Omega_0 \subset \Sigma_1 \cup \Sigma_2$ *where*

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

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

where $H^-_{Q_r(z_0)}(s)$ and $\mathbf{V}_{H^-_{Q_r(z_0)}}(\mathbf{P})$ are defined as in (2.4) and (2.12) with $\varphi(s) = H^-_{Q_r(z_0)}(s)$, respectively.

**Remark 1.3.** *Notice that the set $\Sigma_1$ has measure zero, since it is contained in the set of (parabolic) non-Lebesgue points of $\mathbf{V}_H(\cdot, D\mathbf{u})$ (see (2.16) for the definition of $\mathbf{V}_H$) in $L^2$-space. For details, see [61, Remark 1].*

## 2. Preliminaries

### 2.1. Basic notation.
A space-time cylinder in $\mathbb{R}^{n+1}$ is denoted by

$$Q_{r,\tau}(z_0) = B_r(x_0) \times (t_0 - \tau, t_0 + \tau), \quad \text{where} \quad z_0 = (x_0, t_0) \in \mathbb{R}^{n+1} \quad \text{and} \quad r, \tau > 0.$$

In particular, if $\tau = r^2$, we write $Q_r(z_0) := Q_{r,r^2}(z_0)$ and call it a parabolic cylinder.

Let $\boldsymbol{\ell} : \mathbb{R}^n \to \mathbb{R}^N$ be any linear map of the form

$$(2.1) \qquad \boldsymbol{\ell}(x) := \mathbf{P}(x - x_0) + \mathbf{b}, \quad x \in \mathbb{R}^n,$$

where $\mathbf{P} \in \mathbb{R}^{N \times n}$, $x_0 \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^N$. If $\mathbf{u}$ is a weak solution to (1.3), we set

$$(2.2) \qquad \mathbf{u}_{\boldsymbol{\ell}}(z) := \mathbf{u}(z) - \boldsymbol{\ell}(x), \quad z = (x, t) \in \Omega_T.$$

Let $H : \Omega_T \times [0, \infty) \to [0, \infty)$ be given in (1.5) with (1.6). For $Q_r(z_0) \Subset \Omega_T$, let $z^-_{z_0,r}, z^+_{z_0,r} \in \overline{Q_r(z_0)}$ be such that

$$(2.3) \qquad a^-_{z_0,r} := a(z^-_{z_0,r}) = \inf_{z \in Q_r(z_0)} a(z) \quad \text{and} \quad a^+_{z_0,r} := a(z^+_{z_0,r}) = \sup_{z \in Q_r(z_0)} a(z).$$

Then we write

$$(2.4) \qquad H^-_{Q_r(z_0)}(s) := H(z^-_{z_0,r}, s) \quad \text{and} \quad H^+_{Q_r(z_0)}(s) := H(z^+_{z_0,r}, s).$$

### 2.2. Orlicz functions and operators.
We recall basic notation and properties about Orlicz functions. The following definitions and results can be found, e.g., in [53, 54].

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 $\varphi'$ of $\varphi$ such that $\varphi'(0) = 0$, $\varphi'$ is increasing, and $\varphi'(s) > 0$ when $s > 0$. We say that $\varphi$ satisfies the $\Delta_2$ condition, denoted by $\Delta(\varphi) < \infty$, if there exists a positive constant $K =: \Delta(\varphi)$ such that $\varphi(2s) \leqslant K\varphi(s)$ for all $s > 0$. The conjugate function of $\varphi$ is defined as

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

From the definition of $\varphi^*$, the following Young's inequality

$$s\tilde{s} \leqslant \varphi(\tilde{s}) + \varphi^*(s), \quad s, \tilde{s} \geqslant 0,$$

holds true. From now on, we always assume that $\varphi$ and $\varphi^*$ satisfy the $\Delta_2$ condition and this is indicated by $\Delta(\varphi, \varphi^*) < \infty$, where $\Delta(\varphi, \varphi^*)$ denotes the constants $\Delta(\varphi)$ and $\Delta(\varphi^*)$. We note that the exact value of $\varphi^*$ is not always explicitly computable and instead the estimate

$$(2.5) \qquad \varphi^*\left(\frac{\varphi(s)}{s}\right) \sim \varphi^*(\varphi'(s)) \sim \varphi(s)$$

will often be useful in computations (see [40, Theorem 2.4.10]). Moreover, if $\varphi \in C^1([0, \infty))$ and $\varphi'$ is strictly increasing, then (2.5) implies that

$$(2.6) \qquad (\varphi^*)'(s) \sim (\varphi')^{-1}(s), \quad s \geqslant 0.$$

We recall the notion of "*almost convexity*" and a Jensen type inequality (see [40, Lemma 4.3.1]).

**Lemma 2.1.** *If $\psi : [0, \infty) \to [0, \infty]$ is increasing with $\psi(0) = 0$ and $\frac{\psi(s)}{s}$ is almost increasing; i.e., $\frac{\psi(s)}{s} \leqslant L\frac{\psi(\tilde{s})}{\tilde{s}}$ for every $0 < s \leqslant \tilde{s}$ with constant $L \geqslant 1$, then the following Jensen's type inequality holds:*

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

Now we further assume for $\varphi$ that $\varphi \in C^1([0, \infty)) \cap C^2((0, \infty))$ and

$$(2.7) \qquad 1 < p \leqslant \frac{s\varphi''(s)}{\varphi'(s)} + 1 \leqslant q \quad \text{for all} \ \ s > 0.$$

Note that this implies

$$1 < p \leqslant \frac{s\varphi'(s)}{\varphi(s)} \leqslant q \quad \text{for all} \ \ s > 0\,,$$

and hence the $\Delta_2$ conditions of $\varphi$ and $\varphi^*$.

For an $N$-function $\varphi$ and for $\sigma \geqslant 0$, we define the *shifted N-function* $\varphi_\sigma$ of $\varphi$ by

$$\varphi_\sigma(s) := \int_0^s \frac{\varphi'(\sigma + \tilde{s})\tilde{s}}{\sigma + \tilde{s}}\, \mathrm{d}\tilde{s}\,, \quad \text{that is,} \ \ \varphi_\sigma'(s) = \frac{\varphi'(\sigma + s)}{\sigma + s}s\,.$$

The condition (2.7) implies that

$$(2.8) \qquad \tilde{p} := \min\{p, 2\} \leqslant \frac{s\varphi_\sigma''(s)}{\varphi_\sigma'(s)} + 1 \leqslant \max\{q, 2\} =: \tilde{q} \quad \text{for all} \ \ s > 0 \ \ \text{and} \ \ \sigma \geqslant 0\,.$$

Moreover, we have the following relations (see, e.g., [13, Proposition 2.3] and [27, 31]), which hold uniformly with respect to $\sigma \geqslant 0$:

$$(2.9) \qquad \varphi_\sigma(s) \sim \varphi_\sigma'(s)\, s\,;$$

$$(2.10) \qquad \varphi_\sigma(s) \sim \varphi''(\sigma + s)s^2 \sim \frac{\varphi(\sigma + s)}{(\sigma + s)^2}s^2 \sim \frac{\varphi'(\sigma + s)}{\sigma + s}s^2\,;$$

$$(2.11) \qquad \varphi(s) \leqslant \varphi(\sigma + s) \sim \varphi_\sigma(s) + \varphi(\sigma)\,.$$

We define vector valued functions $\mathbf{V}_\varphi : \mathbb{R}^{N \times n} \to \mathbb{R}^{N \times n}$ by

$$(2.12) \qquad \mathbf{V}_\varphi(\mathbf{Q}) := \sqrt{\frac{\varphi_1'(|\mathbf{Q}|)}{|\mathbf{Q}|}}\mathbf{Q} = \sqrt{\frac{\varphi'(1 + |\mathbf{Q}|)}{1 + |\mathbf{Q}|}}\mathbf{Q}.$$

In particular, we write

$$\mathbf{V}_p(\mathbf{Q}) := \mathbf{V}_\varphi(\mathbf{Q}) \quad \text{when} \ \ \varphi(s) = s^p\,.$$

Then we recall equivalent relations in [27, Lemmas 3 and 20] and [30, Lemma 3.1] with $\varphi_1$ in place of $\varphi$:

$$(2.13) \qquad \frac{\varphi'(1 + |\mathbf{P}| + |\mathbf{Q}|)}{1 + |\mathbf{P}| + |\mathbf{Q}|}|\mathbf{P} - \mathbf{Q}|^2 \sim |\mathbf{V}_\varphi(\mathbf{P}) - \mathbf{V}_\varphi(\mathbf{Q})|^2 \sim \varphi_{1+|\mathbf{Q}|}(|\mathbf{P} - \mathbf{Q}|)\,,$$

$$(2.14) \qquad \frac{\varphi'(1 + |\mathbf{P}| + |\mathbf{Q}|)}{1 + |\mathbf{P}| + |\mathbf{Q}|} \sim \int_0^1 \frac{\varphi'(1 + |\tau\mathbf{P} + (1 - \tau)\mathbf{Q}|)}{1 + |\tau\mathbf{P} + (1 - \tau)\mathbf{Q}|}\, \mathrm{d}\tau\,.$$

Moreover, by the same proof of [29, Lemma A.2], we have that for every $\mathbf{g} \in L^\varphi(Q_r; \mathbb{R}^{N \times n})$,

$$(2.15) \qquad \fint_{Q_r} |\mathbf{V}(\mathbf{g}) - (\mathbf{V}(\mathbf{g}))_{Q_r}|^2\, \mathrm{d}z \sim \fint_{Q_r} |\mathbf{V}(\mathbf{g}) - \mathbf{V}((\mathbf{g})_{Q_r})|^2\, \mathrm{d}z\,.$$

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

### 2.3. Generalized Orlicz function $H$. 
For basic definitions and properties about generalized Orlicz functions and the associated spaces, we refer to the monograph [40].

Let $H : \Omega_T \times [0, \infty) \to [0, \infty)$ be the function given in (1.5). Note that for each $z \in \Omega_T$, $\varphi(s) := H(z, s)$ is an $N$-function satisfying that $\Delta_2(\varphi, \varphi^*) < \infty$. With this $H$, we define the generalized Orlicz space $L^H(\Omega_T, \mathbb{R}^N)$ as the set of all measurable functions $f : \Omega_T \to \mathbb{R}^N$ such that

$$\iint_{\Omega_T} H(z, |f(z)|)\, \mathrm{d}z = \int_0^T \int_\Omega H(x, t, |f(x, t)|)\, \mathrm{d}x\, \mathrm{d}t < \infty,$$

endowed with the usual Luxembourg type norm

$$\|f\|_{L^H(\Omega_T)} := \inf\left\{\lambda > 0 : \iint_{\Omega_T} H\left(z, \frac{|f(z)|}{\lambda}\right)\, \mathrm{d}z \leqslant 1\right\}\,.$$

We define

$$(2.16) \qquad \mathbf{V}_H(z, \mathbf{Q}) := \sqrt{\frac{H'(z, 1 + |\mathbf{Q}|)}{1 + |\mathbf{Q}|}} \mathbf{Q}, \quad z \in \Omega_T.$$

Then it follows from (2.13) that for every $z \in \Omega_T$,

$$
\begin{aligned}
(2.17) \qquad |\mathbf{V}_H(z, \mathbf{P}) - \mathbf{V}_H(z, \mathbf{Q})|^2 &\sim H_{1+|\mathbf{Q}|}(z, |\mathbf{P} - \mathbf{Q}|) \\
&\sim \varphi_{p, 1+|\mathbf{Q}|}(|\mathbf{P} - \mathbf{Q}|) + a(z)\varphi_{q, 1+|\mathbf{Q}|}(|\mathbf{P} - \mathbf{Q}|) \\
&\sim |\mathbf{V}_p(\mathbf{P}) - \mathbf{V}_p(\mathbf{Q})|^2 + a(z)|\mathbf{V}_q(\mathbf{P}) - \mathbf{V}_q(\mathbf{Q})|^2 \\
&\geqslant |\mathbf{V}_p(\mathbf{P}) - \mathbf{V}_p(\mathbf{Q})|^2,
\end{aligned}
$$

where $\varphi_{p,\sigma}(s)$ is the shifted $N$ function of the power function $s^p$ with shift $\sigma \geqslant 0$. Moreover, combining (2.14) and (A1), following [30, Lemma 3.1] we can prove that for every $z \in \Omega_T$,

$$(2.18) \qquad |\mathbf{A}(z, \mathbf{P}) - \mathbf{A}(z, \mathbf{Q})| \sim H'_{1+|\mathbf{Q}|}(z, |\mathbf{P} - \mathbf{Q}|).$$

We also notice that for the shifted $N$-functions $H_\sigma^\pm$ and the corresponding $V_{H_\sigma^\pm}$ function the equivalence (2.15) still holds true. This is related to the set of Lebesgue points.

We end the section with the following parabolic Sobolev-Poincaré type inequality for shifted $N$-functions of $H$.

**Lemma 2.2** (Sobolev-Poincaré inequality). *Let $H$ be defined as in (1.5) with (1.6) and (1.9), and let $\sigma \geqslant 0$ and $0 < r \leqslant 1$. Then there exists $\theta = \theta(n, p, q) \in (0, 1)$ such that for $Q_r = Q_r(z_0) \subset \Omega_T$ with $z_0 = (x_0, t_0)$, if $\mathbf{w} \in L^1(t_0 - r^2, t_0 + r^2; W^{1,1}(B_r(x_0); \mathbb{R}^N))$ with $\|H_\sigma^-(|D\mathbf{w}|)\|_{L^1(Q_r)} < \infty$ satisfies that $\mathbf{w}(\cdot, t) = \mathbf{0}$ on $\partial B_r(x_0)$ in the trace sense for a.e. $t \in (t_0 - r^2, t_0 + r^2)$, then we have*

$$
\fint\!\!\!\!\!\!\int_{Q_r} H_\sigma^+ \left( \frac{|\mathbf{w}|}{r} \right) \, dz \leqslant c_P \left( \fint\!\!\!\!\!\!\int_{Q_r} H_\sigma^-(|D\mathbf{w}|)^\theta \, dz \right)^{\frac{1}{\theta}} + c_P \left( r^{\alpha - \frac{(q-p)(n+2)}{p}} + r^\alpha \sigma^{q-p} \right) \sigma^p,
$$

*for some $c_P = c_P(n, p, q, L) \geqslant 1$, where $H_\sigma^\pm$ denote the shifted $N$-functions of $H_{Q_r}^\pm$ (see (2.4)) with shift $\sigma \geqslant 0$.*

*Proof.* Recall $z_{0,r}^\pm$ from (2.3) and, for ease of notation, set $z^\pm := z_{0,r}^\pm$. First, by (2.10) and (1.6) we get

$$
\begin{aligned}
H_\sigma^+(s) &\sim (\sigma + s)^{p-2} s^2 + a(z^+)(\sigma + s)^{q-2} s^2 \\
&\lesssim (\sigma + s)^{p-2} s^2 + a(z^-)(\sigma + s)^{q-2} s^2 + r^\alpha (\sigma + s)^{q-2} s^2 \\
&\lesssim H_\sigma^-(s) + r^\alpha (\sigma^q + s^q).
\end{aligned}
$$

Now, for each fixed time $t \in (t_0 - r^2, t_0 + r^2)$, we can estimate, by using Sobolev-Poincaré inequalities for the shifted $N$-function $H_\sigma^-(s)$ (see, e.g., [27, Theorem 7]) and for the function $s \mapsto s^q$,

$$
\begin{aligned}
\fint\!\!\!\!\!\!\int_{Q_r} H_\sigma \left( z, \frac{|\mathbf{w}|}{r} \right) \, dz &\lesssim \fint\!\!\!\!\!\!\int_{Q_r} H_\sigma^- \left( \frac{|\mathbf{w}|}{r} \right) \, dz + r^\alpha \fint\!\!\!\!\!\!\int_{Q_r} \left( \frac{|\mathbf{w}|}{r} \right)^q \, dz + r^\alpha \delta^q \\
&\lesssim \left( \fint\!\!\!\!\!\!\int_{Q_r} H_\sigma^-(|D\mathbf{w}|)^{\tilde{\theta}} \, dz \right)^{\frac{1}{\tilde{\theta}}} + r^\alpha \left( \fint\!\!\!\!\!\!\int_{Q_r} |D\mathbf{w}|^{q_*} \, dz \right)^{\frac{q}{q_*}} + r^\alpha \sigma^q,
\end{aligned}
$$

where $\tilde{\theta} \in (0, 1)$ depends on $n, p, q$ and $q_* := \max\{1, \frac{nq}{n+q}\} \leqslant \max\{1, q\frac{n}{n+1}\} < p$. Set $\theta := \max\{\tilde{\theta}, q_*/p\} \in (0, 1)$. Note that, by (2.11) applied to the $N$-function $s \mapsto s^p$, we have

$$(2.19) \qquad s^p \lesssim (\delta + \sigma)^{p-2} s^2 + \sigma^p \lesssim H_\sigma^-(s) + \sigma^p.$$

Then using Hölder's inequality, (2.19), (1.9) and the assumption that $\|D\mathbf{w}\|_{L^p(Q_r;\mathbb{R}^{Nn})} \leqslant 1$, we obtain

$$
r^\alpha \left( \iint_{Q_r} |D\mathbf{w}|^{q_*} \, \mathrm{d}z \right)^{\frac{q}{q_*}} \leqslant r^\alpha \left( \iint_{Q_r} |D\mathbf{w}|^{p\theta} \, \mathrm{d}z \right)^{\frac{1}{\theta}} \left( \iint_{Q_r} |D\mathbf{w}|^p \, \mathrm{d}z \right)^{\frac{q-p}{p}}
$$

$$
\leqslant r^{\alpha - \frac{(q-p)(n+2)}{p}} \left( \iint_{Q_r} |D\mathbf{w}|^{p\theta} \, \mathrm{d}z \right)^{\frac{1}{\theta}}
$$

$$
\leqslant c \left( \iint_{Q_r} H_\sigma^-(|D\mathbf{w}|)^\theta \, \mathrm{d}z \right)^{\frac{1}{\theta}} + cr^{\alpha - \frac{(q-p)(n+2)}{p}} \sigma^p \, .
$$

This completes the proof. $\qquad\qquad\square$

## 3. Caccioppoli inequality via mollification

In this section, we derive a Caccioppoli type estimate for $\mathbf{u}_{\boldsymbol{\ell}} = \mathbf{u} - \boldsymbol{\ell}$ with the shifted function $H_{1+|D\boldsymbol{\ell}|}(z,s)$, where $\mathbf{u}$ is the weak solution to (1.3) and $\boldsymbol{\ell}$ is any linear function (2.1). We emphasize that in the definition of weak solution, we only assume that $H(\cdot,|D\mathbf{u}|) \in L^1_{\mathrm{loc}}(\Omega_T)$. Therefore, a more general regularization argument than the Steklov average is needed. To this end, we use mollification in $\mathbb{R}^{n+1}$ in the parabolic setting. We refer to [40, Chapter 4] and [46].

Let $\tilde{\varrho}_m \in C_0^\infty(\mathbb{R}^m)$ be a standard mollifier in $\mathbb{R}^m$ defined by $\tilde{\varrho}_m(y) = \gamma_m \exp(-\frac{1}{1-|y|^2})$ if $|y| < 1$ and $\tilde{\varrho}_m(y) = 0$ if $|y| \geqslant 1$, where the constant $\gamma_m > 0$ is chosen such that $\|\tilde{\varrho}_m\|_1 = 1$. Then we set $\varrho(z) = \varrho(x,t) := \tilde{\varrho}_n(x)\tilde{\varrho}_1(t)$ and, for $h > 0$, $\varrho_h(z) = \varrho_h(x,t) := h^{-n-2}\tilde{\varrho}_n(h^{-1}x)\tilde{\varrho}_1(h^{-2}t)$. For $f \in L^1_{\mathrm{loc}}(\Omega_T)$, $z \in \Omega_T$ and $h > 0$ such that $Q_h(z) \Subset \Omega_T$, we define

$$
f_h(z) := \iint_{Q_h(0)} f(z - \tilde{z})\varrho_h(\tilde{z}) \, \mathrm{d}\tilde{z} \, .
$$

Let $H$ be defined as in (1.5). We observe that the conditions (1.6) and (1.9) imply that for both $H$ and $H^*$ holds the parabolic version of the (A1) condition in [40, Definition 4.1.1], where the ball $B$ is replaced by the parabolic cylinder $Q_r$ in the definition of the usual (A1) condition; see [40, Proposition 7.2.2 and Lemma 4.1.7]. Therefore, the convergence of mollification in the spaces $L^H$ and $L^{H^*}$ in the parabolic setting follows as in [40, Theorem 4.4.7]. We also refer to [46, Proposition 4.1] and the discussion therein for further details. Precisely, if the function $H$ defined in (1.5) satisfies (1.6) and (1.9), then for every $f \in L^H_{\mathrm{loc}}(\Omega_T)$ and $g \in L^{H^*}_{\mathrm{loc}}(\Omega_T)$ we have

$$
(3.1) \qquad f_h \longrightarrow f \quad \text{in } L^H_{\mathrm{loc}}(\Omega_T) \quad \text{and} \quad g_h \longrightarrow g \quad \text{in } L^{H^*}_{\mathrm{loc}}(\Omega_T), \qquad \text{as } h \to 0^+.
$$

**Lemma 3.1.** *Let $\mathbf{u}$ be a weak solution to (1.3) satisfying (A1), (A2) and (A3) with (1.5), (1.6) and (1.9). For every pair of concentric cylinders $Q_{r_1,\rho_1}(z_0) \subset Q_{r_2,\rho_2}(z_0) \Subset \Omega_T$ with $z_0 = (x_0,t_0)$, $0 < r_1 < r_2 \leqslant 1$ and $0 < \rho_1 < \rho_2 \leqslant 1$, and linear map $\boldsymbol{\ell}$, we have*

$$
\operatorname*{ess\,sup}_{t\in(t_0-\rho_1,t_0+\rho_1)} \int_{B_{r_1}(x_0)} |\mathbf{u}_{\boldsymbol{\ell}}(x,t)|^2 \, \mathrm{d}x + \iint_{Q_{r_1,\rho_1}(z_0)} H_{1+|D\boldsymbol{\ell}|}(z,|D\mathbf{u}_{\boldsymbol{\ell}}|) \, \mathrm{d}z
$$

$$
(3.2) \qquad \leqslant \hat{c} \iint_{Q_{r_2,\rho_2}(z_0)} \left[ \frac{|\mathbf{u}_{\boldsymbol{\ell}}|^2}{\rho_2 - \rho_1} + H_{1+|D\boldsymbol{\ell}|}\left(z, \frac{|\mathbf{u}_{\boldsymbol{\ell}}|}{r_2 - r_1}\right) \right] \mathrm{d}z
$$

$$
+ \hat{c}|Q_{r_2,\rho_2}| \max \left\{ C_1^{\frac{\tilde{p}}{\tilde{p}-1}}, C_1^{\frac{\tilde{q}}{\tilde{q}-1}} \right\} H(z^+, 1 + |D\boldsymbol{\ell}|) \, .
$$

*for some $\hat{c} = \hat{c}(n,N,p,q,L,\nu) > 0$, where*

$$
(3.3) \qquad\qquad C_1 := (r_2 + \sqrt{\rho_2})^{\beta_0} + (r_2 + \sqrt{\rho_2})^\alpha (1 + |D\boldsymbol{\ell}|)^{q-p}
$$

*and $1 < \tilde{p} \leqslant \tilde{q}$ are given in (2.8).*

*Proof.* Let $h_0 \in (0,1)$ sufficiently small so that

$$
h \leqslant \min \left\{ \frac{1}{4}\operatorname{dist}(Q_{r_2,\rho_2}(z_0), \partial(\Omega_T)), \frac{\rho_2 - \rho_1}{4} \right\},
$$

and choose $h \in (0, h_0)$. Let $z^- \in \overline{Q_{r_2,\rho_2}(z_0)}$ such that $a(z^-) = \inf\limits_{z \in Q_{r_2,\rho_2}(z_0)} a(z)$. Since $\partial_t \boldsymbol{\ell} =$ div $\mathbf{A}(z^-, D\boldsymbol{\ell}) = \mathbf{0}$, we obtain from the weak formulation (1.11) that

$$(3.4) \qquad -\iint_{Q_{r_2,\rho_2}(z_0)} [\mathbf{u}_{\boldsymbol{\ell}}]_h \cdot \partial_t \boldsymbol{\zeta} \, dz + \iint_{Q_{r_2,\rho_2}(z_0)} [\mathbf{A}(\cdot, D\mathbf{u}) - \mathbf{A}(z^-, D\mathbf{u})]_h : D\boldsymbol{\zeta} \, dz = 0$$

for all $\boldsymbol{\zeta} \in C_c^\infty(Q_{r_2,\rho_2}(z_0), \mathbb{R}^N)$, where $[\cdot]_h$ is the mollification stated in above. Note that by approximation the test function $\boldsymbol{\zeta}$ can be chosen as a Lipschitz continuous function in $Q_{r_2,\rho_2}(z_0)$ with zero value on $\partial(Q_{r_2,\rho_2}(z_0))$, and $\mathbf{A}(\cdot, D\mathbf{u}), \mathbf{A}(z^-, D\mathbf{u}) \in L_{\mathrm{loc}}^{H^*}(\Omega_T, \mathbb{R}^{N \times n})$.

Let $\xi \in C_0^\infty(B_{r_2}(x_0))$ be a cut-off function with

$$(3.5) \qquad 0 \leqslant \xi \leqslant 1, \quad \xi \equiv 1 \text{ in } B_{r_1}(x_0) \quad \text{and} \quad \|D\xi\|_{L^\infty} \leqslant \frac{2}{r_2 - r_1},$$

and $\eta \in C_0^\infty(I_{\rho_2 - h_0}(t_0))$ with

$$(3.6) \qquad 0 \leqslant \eta \leqslant 1, \quad \eta \equiv 1 \text{ in } I_{\rho_1}(t_0) \quad \text{and} \quad \|\eta'\|_{L^\infty} \leqslant \frac{3}{\rho_2 - \rho_1}.$$

In addition, for $t_* \in I_{\rho_1}(t_0)$ and $\delta \in (0, h_0)$, define $\eta_\delta : \mathbb{R} \to \mathbb{R}$ as

$$(3.7) \qquad \eta_\delta(t) = \begin{cases} 1, & t \in (-\infty, t_* - \delta), \\ 1 - \dfrac{t - t_* + \delta}{\delta}, & t \in [t_* - \delta, t_*], \\ 0, & t \in (t_*, \infty). \end{cases}$$

Then we can choose $\boldsymbol{\zeta} = [\mathbf{u}_{\boldsymbol{\ell}}]_h \xi^{\tilde{q}} \eta^2 \eta_\delta$ as a test function in (3.4), so that

$$(3.8) \begin{aligned} 0 = & \iint_{Q_{r_2,\rho_2}(z_0)} \partial_t [\mathbf{u}_{\boldsymbol{\ell}}]_h \cdot \boldsymbol{\zeta} \, dz + \iint_{Q_{r_2,\rho_2}(z_0)} [\mathbf{A}(\cdot, D\mathbf{u}) - \mathbf{A}(\cdot, D\boldsymbol{\ell})]_h : D\boldsymbol{\zeta} \, dz \\ & + \iint_{Q_{r_2,\rho_2}(z_0)} [\mathbf{A}(\cdot, D\boldsymbol{\ell}) - \mathbf{A}(z^-, D\boldsymbol{\ell})]_h : D\boldsymbol{\zeta} \, dz \\ =: & \, J_1 + J_2 + J_3 \, . \end{aligned}$$

We start by estimating $J_1$. Integration by parts gives

$$\begin{aligned} J_1 = & \iint_{Q_{r_2,\rho_2}(z_0)} \tfrac{1}{2} \partial_t(|[\mathbf{u}_{\boldsymbol{\ell}}]_h|^2) \xi^{\tilde{q}} \eta^2 \eta_\delta \, dz \\ = & -\iint_{Q_{r_2,\rho_2}(z_0)} |[\mathbf{u}_{\boldsymbol{\ell}}]_h|^2 \xi^{\tilde{q}} \eta \eta' \eta_\delta \, dz + \tfrac{1}{2} \int_{t_* - \delta}^{t_*} \int_{B_{r_2}(x_0)} |[\mathbf{u}_{\boldsymbol{\ell}}]_h|^2 \xi^{\tilde{q}} \eta^2 \, dx \, dt \, . \end{aligned}$$

Then since $\mathbf{u}_{\boldsymbol{\ell}} \in L_{\mathrm{loc}}^2(\Omega_T)$, letting $h \to 0^+$ and then $\delta \to 0^+$ and using (3.6) and (3.7), we obtain

$$(3.9) \begin{aligned} \lim_{\delta \to 0^+} \lim_{h \to 0^+} J_1 = & -\int_{t_0 - \rho_2}^{t^*} \int_{B_{r_2}(x_0)} |\mathbf{u}_{\boldsymbol{\ell}}|^2 \xi^{\tilde{q}} \eta \eta' \, dx \, dt + \tfrac{1}{2} \int_{B_{r_2}(x_0)} |\mathbf{u}_{\boldsymbol{\ell}}(x, t^*)|^2 \xi^{\tilde{q}} \, dx \\ \geqslant & -c \iint_{Q_{r_2,\rho_2}(z_0)} \frac{|\mathbf{u}_{\boldsymbol{\ell}}|^2}{\rho_2 - \rho_1} \, dz + \tfrac{1}{2} \int_{B_{r_2}(x_0)} |\mathbf{u}_{\boldsymbol{\ell}}(x, t^*)|^2 \xi^{\tilde{q}} \, dx \, . \end{aligned}$$

Now we estimate $J_2$. Since $\mathbf{A}(\cdot, D\mathbf{u}) - \mathbf{A}(\cdot, D\boldsymbol{\ell}) \in L_{\text{loc}}^{H^*}(\Omega_T)$ and $D\mathbf{u}_{\boldsymbol{\ell}}, \mathbf{u}_{\boldsymbol{\ell}} \in L_{\text{loc}}^{H}(\Omega_T)$, by (3.1),

(3.10)
$$
\begin{aligned}
\lim_{\delta \to 0^+} \lim_{h \to 0^+} J_2 &= \lim_{\delta \to 0^+} \lim_{h \to 0^+} \Bigg[ \iint_{Q_{r_2,\rho_2}(z_0)} [\mathbf{A}(\cdot, D\mathbf{u}) - \mathbf{A}(\cdot, D\boldsymbol{\ell})]_h : [D\mathbf{u}_{\boldsymbol{\ell}}]_h \xi^{\tilde{q}} \eta^2 \eta_\delta \, \mathrm{d}z \\
&\qquad + \tilde{q} \iint_{Q_{r_2,\rho_2}(z_0)} [\mathbf{A}(\cdot, D\mathbf{u}) - \mathbf{A}(\cdot, D\boldsymbol{\ell})]_h : D\xi \otimes [\mathbf{u}_{\boldsymbol{\ell}}]_h \xi^{\tilde{q}-1} \eta^2 \eta_\delta \, \mathrm{d}z \Bigg] \\
&= \int_{t_0-\rho_2}^{t^*} \int_{B_{r_2}(x_0)} [\mathbf{A}(z, D\mathbf{u}) - \mathbf{A}(z, D\boldsymbol{\ell})] : D\mathbf{u}_{\boldsymbol{\ell}} \, \xi^{\tilde{q}} \eta^2 \, \mathrm{d}x \, \mathrm{d}t \\
&\qquad + \tilde{q} \int_{t_0-\rho_2}^{t^*} \int_{B_{r_2}(x_0)} [\mathbf{A}(z, D\mathbf{u}) - \mathbf{A}(z, D\boldsymbol{\ell})] : D\xi \otimes \mathbf{u}_{\boldsymbol{\ell}} \, \xi^{\tilde{q}-1} \eta^2 \, \mathrm{d}x \, \mathrm{d}t .
\end{aligned}
$$

Applying (1.8) with $\boldsymbol{\xi}_1 = D\mathbf{u}$, $\boldsymbol{\xi}_2 = D\boldsymbol{\ell}$ and (2.10) to get

$$
\begin{aligned}
\int_{t_0-\rho_2}^{t^*} &\int_{B_{r_2}(x_0)} [\mathbf{A}(z, D\mathbf{u}) - \mathbf{A}(z, D\boldsymbol{\ell})] : D\mathbf{u}_{\boldsymbol{\ell}} \xi^{\tilde{q}} \eta^2 \, \mathrm{d}x \, \mathrm{d}t \\
&\geqslant \tilde{\nu} \int_{t_0-\rho_2}^{t^*} \int_{B_{r_2}(x_0)} H''(z, 1 + |D\mathbf{u}| + |D\boldsymbol{\ell}|) |D\mathbf{u}_{\boldsymbol{\ell}}|^2 \xi^{\tilde{q}} \eta^2 \, \mathrm{d}x \, \mathrm{d}t \\
&\geqslant \tilde{c} \int_{t_0-\rho_2}^{t_*} \int_{B_{r_2}(x_0)} H_{1+|D\boldsymbol{\ell}|}(z, |D\mathbf{u}_{\boldsymbol{\ell}}|) \xi^{\tilde{q}} \eta^2 \, \mathrm{d}x \, \mathrm{d}t
\end{aligned}
$$

for some small $\tilde{c} > 0$. To estimate the second term on the right hand side in (3.10), we use (2.18) and (3.5), Young's inequality applied to $\varphi(s) = H_{1+|D\boldsymbol{\ell}|}(z, s)$ for any fixed $z$ with (2.5) to conclude that for every $\varepsilon > 0$,

$$
\begin{aligned}
\int_{t_0-\rho_2}^{t_*} &\int_{B_{r_2}(x_0)} [\mathbf{A}(z, D\mathbf{u}) - \mathbf{A}(z, D\boldsymbol{\ell})] : D\xi \otimes \mathbf{u}_{\boldsymbol{\ell}} \xi^{\tilde{q}-1} \eta^2 \, \mathrm{d}x \, \mathrm{d}t \\
&\geqslant -c \int_{t_0-\rho_2}^{t_*} \int_{B_{r_2}(x_0)} H'_{1+|D\boldsymbol{\ell}|}(z, |D\mathbf{u}_{\boldsymbol{\ell}}|) \frac{|\mathbf{u}_{\boldsymbol{\ell}}|}{r_2 - r_1} \xi^{\tilde{q}-1} \eta^2 \, \mathrm{d}x \, \mathrm{d}t \\
&\geqslant -\varepsilon \int_{t_0-\rho_2}^{t_*} \int_{B_{r_2}(x_0)} H_{1+|D\boldsymbol{\ell}|}(z, |D\mathbf{u}_{\boldsymbol{\ell}}|) \xi^{\tilde{q}} \eta^2 \, \mathrm{d}x \, \mathrm{d}t - c(\varepsilon) \iint_{Q_{r_2,\rho_2}(z_0)} H_{1+|D\boldsymbol{\ell}|} \left( z, \frac{|\mathbf{u}_{\boldsymbol{\ell}}|}{r_2 - r_1} \right) \eta^2 \, \mathrm{d}z .
\end{aligned}
$$

Therefore, it follows that

(3.11)
$$
\begin{aligned}
\lim_{\delta \to 0^+} \lim_{h \to 0^+} J_2 &\geqslant \frac{\tilde{c}}{2} \int_{t_0-\rho_2}^{t_*} \int_{B_{r_2}(x_0)} H_{1+|D\boldsymbol{\ell}|}(z, |D\mathbf{u}_{\boldsymbol{\ell}}|) \xi^{\tilde{q}} \eta^2 \, \mathrm{d}x \, \mathrm{d}t \\
&\qquad - c \iint_{Q_{r_2,\rho_2}(z_0)} H_{1+|D\boldsymbol{\ell}|} \left( z, \frac{|\mathbf{u}_{\boldsymbol{\ell}}|}{r_2 - r_1} \right) \mathrm{d}z.
\end{aligned}
$$

We finally estimate $J_3$. Similarly to the estimation of $J_2$, we have

$$
\lim_{\delta \to 0^+} \lim_{h \to 0^+} J_3 = \int_{t_0-\rho_2}^{t^*} \int_{B_{r_2}(x_0)} [\mathbf{A}(z, D\boldsymbol{\ell}) - \mathbf{A}(z^-, D\boldsymbol{\ell})] : [D\mathbf{u}_{\boldsymbol{\ell}} \xi^{\tilde{q}} + \tilde{q} D\xi \otimes \mathbf{u}_{\boldsymbol{\ell}} \xi^{\tilde{q}-1}] \eta^2 \, \mathrm{d}x \, \mathrm{d}t .
$$

By (A3) and (1.6), for $z \in Q_{r_2,\rho_2}(z_0)$,

$$
\begin{aligned}
|\mathbf{A}(z, D\boldsymbol{\ell}) - \mathbf{A}(z^-, D\boldsymbol{\ell})| &\leqslant c(r_2 + \sqrt{\rho_2})^{\beta_0} H'(z, 1 + |D\boldsymbol{\ell}|) + c(r_2 + \sqrt{\rho_2})^{\alpha}(1 + |D\boldsymbol{\ell}|)^{q-1} \\
&\leqslant cC_1 H'_{1+|D\boldsymbol{\ell}|}(z, 1 + |D\boldsymbol{\ell}|),
\end{aligned}
$$

where $C_1$ is given in (3.3). Plugging this in the previous identity and using Young's inequality, we obtain

$$
\lim_{\delta \to 0^+} \lim_{h \to 0^+} J_3 \geqslant -\frac{\tilde{c}}{4} \int_{t_0-\rho_2}^{t^*} \int_{B_{r_2}(x_0)} H_{1+|D\boldsymbol{\ell}|}(z, |D\mathbf{u}_{\boldsymbol{\ell}}|) \xi^{\tilde{q}} \eta^2 \, dx \, dt
$$

(3.12)
$$
- c \iint_{Q_{r_2,\rho_2}(z_0)} H_{1+|D\boldsymbol{\ell}|}\left(z, \frac{|\mathbf{u}_{\boldsymbol{\ell}}|}{r_2-r_1}\right) dz
$$

$$
- c|Q_{r_2,\rho_2}| \max\left\{ C_1^{\frac{\tilde{p}}{\tilde{p}-1}}, C_1^{\frac{\tilde{q}}{\tilde{q}-1}} \right\} H(z^+, 1+|D\boldsymbol{\ell}|).
$$

Therefore, combining (3.8), (3.9), (3.11) and (3.12) yields

$$
\int_{B_{r_2}(x_0)} |\mathbf{u}_{\boldsymbol{\ell}}(x,t^*)|^2 \xi^{\tilde{q}} \, dx + \int_{t_0-\rho_2}^{t^*} \int_{B_{r_2}(x_0)} H_{1+|D\boldsymbol{\ell}|}(z, |D\mathbf{u}_{\boldsymbol{\ell}}|) \xi^{\tilde{q}} \eta^2 \, dx \, dt
$$

$$
\leqslant c \iint_{Q_{r_2,\rho_2}(z_0)} \left[ \frac{|\mathbf{u}_{\boldsymbol{\ell}}|^2}{\rho_2-\rho_1} + H_{1+|D\boldsymbol{\ell}|}\left(z, \frac{|\mathbf{u}_{\boldsymbol{\ell}}|}{r_2-r_1}\right) \right] dz
$$

$$
+ c|Q_{r_2,\rho_2}| \max\left\{ C_1^{\frac{\tilde{p}}{\tilde{p}-1}}, C_1^{\frac{\tilde{q}}{\tilde{q}-1}} \right\} H(z^+, 1+|D\boldsymbol{\ell}|).
$$

Since $t_* \in (t_0 - \rho_1, t_0 + \rho_1)$ is arbitrary, considering (3.5) and (3.6), we obtain (3.2). $\qquad\square$

## 4. $\mathcal{A}$-CALORIC APPROXIMATION AND LINEARIZATION

The $\mathcal{A}$-caloric approximation, as stated in [60, Theorem 3.8], is a crucial tool in the nondegenerate setting. It will be used in Section 4.2 to establish a comparison estimate between the gradient of a weak solution to the system (1.3) and the gradient of a smooth $\mathcal{A}$-caloric map (see Section 4.1). This comparison is possible provided that $\mathbf{u}$ can be shown to be an almost solution of a suitable $\mathcal{A}$-caloric system.

### 4.1. Regularity estimates for linear systems with constant coefficients.
We first introduce an $\mathcal{A}$-caloric map and it regularity estimates. Let $\mathcal{A} = (\mathcal{A}_{ij}^{\alpha\beta}) \in \mathbb{R}^{N^2 n^2}$ satisfy the *Legendre-Hadamard condition*: for every $\mathbf{a} = (a^1, \ldots, a^N) \in \mathbb{R}^N$ and $\mathbf{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n$,

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

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

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

is called an *$\mathcal{A}$-caloric map*. By standard regularity theory (see [11, 26]), $\mathbf{h} \in C^\infty(Q_r, \mathbb{R}^N)$. Moreover, one can derive the following estimates (see, e.g., [11, 26, (5.9)–(5.12)]), which will be used later.

**Lemma 4.1.** *Suppose $\mathbf{h} \in C^\infty(Q_r, \mathbb{R}^N)$ is an $\mathcal{A}$-caloric map in $Q_r$. Then we have that*

$$
(4.1) \qquad \sup_{Q_{r/2}} \left( |\mathbf{h}| + r|D\mathbf{h}| + r^2|D^2\mathbf{h}| \right) \leqslant c \iint_{Q_r} |\mathbf{h}| \, dz,
$$

*and*

$$
(4.2) \qquad \sup_{Q_{r/2}} |D\mathbf{h}| \leqslant c \iint_{Q_r} |D\mathbf{h}| \, dz.
$$

*Here, the constants $c > 0$ depend on $n, N, \nu$ and $|\mathcal{A}|$.*

We recall the following $\mathcal{A}$-caloric approximation result, proved in [60, Theorem 3.8].

**Theorem 4.2.** *($\mathcal{A}$-caloric approximation) Let $\mu, \sigma_0, C_0 > 0$ and $\psi$ be an $N$-function with $\Delta_2(\psi, \psi^*) < \infty$. For every $\varepsilon \in (0, 1)$, there exists $\delta > 0$ depending on $\sigma_0, C_0, \Delta_2(\psi, \psi^*)$ and $\varepsilon$ such that if $\mathbf{u} \in L^1(-r^2, r^2, W^{1,1}(B_r; \mathbb{R}^N))$ with $\psi^{1+\sigma_0}(|D\mathbf{u}|) \in L^1(Q_r)$ and $\psi^{1+\sigma_0}(\mathbf{H}) \in L^1(Q_r)$ satisfy*

$$
\partial_t \mathbf{u} = \mathrm{div}\, \mathbf{H} \quad \text{in} \quad Q_r,
$$

*in the distributional sense, and the inequality*

$$\left( \iint_{Q_r} \left[ \psi(|D\mathbf{u}|)^{1+\sigma_0} + \psi(|\mathbf{H}|)^{1+\sigma_0} \right] \mathrm{d}z \right)^{\frac{1}{1+\sigma_0}} \leqslant C_0 \psi(\mu),$$

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

$$\frac{1}{|Q_r|} \left| \iint_{Q_r} \mathbf{u} \cdot \boldsymbol{\zeta}_t - [\mathcal{A}D\mathbf{u} : D\boldsymbol{\zeta}] \, \mathrm{d}z - \left[ \int_{B_r} \mathbf{u} \cdot \boldsymbol{\zeta}_t \, \mathrm{d}x \right]_{t=-r^2}^{t=r^2} \right| \leqslant \delta\mu \|D\boldsymbol{\zeta}\|_{L^\infty(Q_r, \mathbb{R}^{N\times n})},$$

*then*

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

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

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

4.2. **Linearization.** To take advantage of the comparison estimates in Theorem 4.2, we demonstrate that the weak solution $\mathbf{u}$ to (1.3) is an 'almost' weak solution of a linear system with constant coefficients. We first note that, since $\|H(\cdot, |D\mathbf{u}|)\|_{L^1(\Omega_T)} < \infty$ by the definition of the weak solution to (1.3), there exists small $R_0 \in (0, 1)$ such that

$$(4.3) \qquad \iint_{Q_r(z_0)} H(z, 1 + |D\mathbf{u}|) \, \mathrm{d}z \leqslant 1 \quad \text{for every} \ \ Q_r(z_0) \Subset \Omega_T \ \ \text{with} \ \ r \in (0, R_0].$$

**Lemma 4.3.** *Let* $\mathbf{u}$ *be a weak solution to* (1.3) *satisfying* (A1), (A2), (A3) *and* (A4) *with* (1.5), (1.6) *and* (1.7), $\boldsymbol{\ell}$ *be any linear map defined in* (2.1), *and* $Q_r = Q_r(z_0) \Subset \Omega_T$ *with* $z_0 = (x_0, t_0)$ *and* $r \in (0, R_0]$. *Then for every* $\boldsymbol{\zeta} \in C^\infty(\overline{Q_r}; \mathbb{R}^N)$ *with* $\boldsymbol{\zeta} = \mathbf{0}$ *on* $\partial B_r \times I_r$ *and* $\|D\boldsymbol{\zeta}\|_{L^\infty(Q_r; \mathbb{R}^{N\times n})} \leqslant 1$, *we have*

$$\frac{1}{|Q_r|} \left| \iint_{Q_r} \mathbf{u}_{\boldsymbol{\ell}} \cdot \boldsymbol{\zeta}_t - \langle D_{\boldsymbol{\xi}}\mathbf{A}(z^-, D\boldsymbol{\ell})D\mathbf{u}_{\boldsymbol{\ell}} \,|\, D\boldsymbol{\zeta} \rangle \, \mathrm{d}z - \left[ \int_{B_r} \mathbf{u}_{\boldsymbol{\ell}} \cdot \boldsymbol{\zeta} \, \mathrm{d}x \right]_{t=t_0-r^2}^{t=t_0+r^2} \right|$$

$$(4.4) \qquad \leqslant cr^{\beta_1} \left\{ \left( (H^-)' \circ (H^-)^{-1} \right) \left( \iint_{Q_r} H^-_{1+|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|) \, \mathrm{d}z \right) + (H^-)'(1 + |D\boldsymbol{\ell}|) \right\}$$

$$+ c(H^-)'(1 + |D\boldsymbol{\ell}|) \left\{ \iint_{Q_r} \frac{H^-_{1+|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)}{H^-(1 + |D\boldsymbol{\ell}|)} \, \mathrm{d}z + \left( \iint_{Q_r} \frac{H^-_{1+|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)}{H^-(1 + |D\boldsymbol{\ell}|)} \, \mathrm{d}z \right)^{\frac{1+\gamma}{2}} \right\}$$

*for some* $c = c(n, N, p, q, L, \nu)$, *where* $\beta_1 = \beta_1(n, p, q, \alpha, \beta_0) > 0$ *is defined as*

$$(4.5) \qquad\qquad \beta_1 := \min \left\{ \beta_0, \alpha - (n+2)\frac{q-p}{p} \right\},$$

*and* $\beta_0$, $\gamma$, $\mathbf{u}_{\boldsymbol{\ell}}$ *and* $H^\pm := H^\pm_{Q_r}$ *are from* (A3), (A4), (2.2) *and* (2.4), *respectively.*

*Proof.* Throughout the proof, we will denote $z^-_{z_0, r}$ by $z^-$. Note that by a similar approximation used in the proof of Lemma 3.1, we have from (1.11) that

$$-\iint_{Q_r} \mathbf{u} \cdot \boldsymbol{\zeta}_t \, \mathrm{d}z + \left[ \int_{B_r} \mathbf{u} \cdot \boldsymbol{\zeta} \, \mathrm{d}x \right]_{t=t_0-r^2}^{t=t_0+r^2} + \iint_{Q_r} \mathbf{A}(z, D\mathbf{u}) : D\boldsymbol{\zeta} \, \mathrm{d}z = 0.$$

From this and the fact that $\boldsymbol{\ell}_t = \mathrm{div}(\mathbf{A}(z^-, D\boldsymbol{\ell})) = \mathbf{0}$, we obtain

$$
\frac{1}{|Q_r|}\left|\iint_{Q_r} \mathbf{u}_{\boldsymbol{\ell}} \cdot \boldsymbol{\zeta}_t - \langle D_{\boldsymbol{\xi}}\mathbf{A}(z^-, D\boldsymbol{\ell})D\mathbf{u}_{\boldsymbol{\ell}} \,|\, D\boldsymbol{\zeta}\rangle \,\mathrm{d}z - \left[\int_{B_r} \mathbf{u}_{\boldsymbol{\ell}} \cdot \boldsymbol{\zeta}\,\mathrm{d}x\right]_{t=-r^2}^{t=r^2}\right|
$$

$$
= \frac{1}{|Q_r|}\left|\iint_{Q_r} \mathbf{u} \cdot \boldsymbol{\zeta}_t - \langle D_{\boldsymbol{\xi}}\mathbf{A}(z^-, D\boldsymbol{\ell})D\mathbf{u}_{\boldsymbol{\ell}} \,|\, D\boldsymbol{\zeta}\rangle \,\mathrm{d}z - \left[\int_{B_r} \mathbf{u} \cdot \boldsymbol{\zeta}\,\mathrm{d}x\right]_{t=-r^2}^{t=r^2}\right|
$$

$$
= \left|\fint\!\!\!\!\fint_{Q_r} \langle \mathbf{A}(z, D\mathbf{u}) - \mathbf{A}(z^-, D\boldsymbol{\ell}) \,|\, D\boldsymbol{\zeta}\rangle - \langle D_{\boldsymbol{\xi}}\mathbf{A}(z^-, D\boldsymbol{\ell})D\mathbf{u}_{\boldsymbol{\ell}} \,|\, D\boldsymbol{\zeta}\rangle \,\mathrm{d}z\right|
$$

(4.6)
$$
= \left|\fint\!\!\!\!\fint_{Q_r} \langle \mathbf{A}(z, D\mathbf{u}) - \mathbf{A}(z^-, D\mathbf{u}) \,|\, D\boldsymbol{\zeta}\rangle \,\mathrm{d}z\right.
$$

$$
\left. + \fint\!\!\!\!\fint_{Q_r}\int_0^1 \langle[D_{\boldsymbol{\xi}}\mathbf{A}(z^-, sD\mathbf{u}_{\boldsymbol{\ell}} + D\boldsymbol{\ell}) - D_{\boldsymbol{\xi}}\mathbf{A}(z^-, D\boldsymbol{\ell})]D\mathbf{u}_{\boldsymbol{\ell}}, D\boldsymbol{\zeta}\rangle \,\mathrm{d}s\,\mathrm{d}z\right|
$$

$$
\leqslant \fint\!\!\!\!\fint_{Q_r} |\mathbf{A}(z, D\mathbf{u}) - \mathbf{A}(z^-, D\mathbf{u})|\,\mathrm{d}z
$$

$$
+ \fint\!\!\!\!\fint_{Q_r} \left[\int_0^1 |D_{\boldsymbol{\xi}}\mathbf{A}(z^-, sD\mathbf{u}_{\boldsymbol{\ell}} + D\boldsymbol{\ell}) - D_{\boldsymbol{\xi}}\mathbf{A}(z^-, D\boldsymbol{\ell})|\,\mathrm{d}s\right]|D\mathbf{u}_{\boldsymbol{\ell}}|\,\mathrm{d}z
$$

$$
=: I_1 + I_2\,.
$$

We first estimate $I_1$. By (A3) and the definition of $z^-$,

(4.7)
$$
I_1 \leqslant cr^{\beta_0}\fint\!\!\!\!\fint_{Q_r} H'(z, 1 + |D\mathbf{u}|)\,\mathrm{d}z + c\fint\!\!\!\!\fint_{Q_r}(a(z) - a(z^-))(1 + |D\mathbf{u}|)^{q-1}\,\mathrm{d}z
$$

$$
\leqslant cr^{\beta_0}\fint\!\!\!\!\fint_{Q_r}(H^-)'(1 + |D\mathbf{u}|)\,\mathrm{d}z + cr^{\alpha}\fint\!\!\!\!\fint_{Q_r}(1 + |D\mathbf{u}|)^{q-1}\,\mathrm{d}z.
$$

By Jensen's inequality and the fact that $((H^-)^*)^{-1} \approx (H^-)' \circ (H^-)^{-1}$, the first integral is estimated as

$$
\fint\!\!\!\!\fint_{Q_r}(H^-)'(1 + |D\mathbf{u}|)\,\mathrm{d}z \leqslant ((H^-)^*)^{-1}\left(\fint\!\!\!\!\fint_{Q_r}((H^-)^* \circ (H^-)')(1 + |D\mathbf{u}|)\,\mathrm{d}z\right)
$$

$$
\sim \left((H^-)' \circ (H^-)^{-1}\right)\left(\fint\!\!\!\!\fint_{Q_r} H^-(1 + |D\mathbf{u}|)\,\mathrm{d}z\right).
$$

For the second integral, by (1.6), (1.7), (4.3) and Jensen's inequality for the function $t \mapsto H^-(t^{1/p})$,

$$
r^{\alpha}\fint\!\!\!\!\fint_{Q_r}(1 + |D\mathbf{u}|)^{q-1}\,\mathrm{d}z \leqslant r^{\alpha}\left(\fint\!\!\!\!\fint_{Q_r}(1 + |D\mathbf{u}|)^p\,\mathrm{d}z\right)^{\frac{q-1}{p}}
$$

$$
\leqslant r^{\alpha - \frac{q-p}{p}(n+2)}\left(\fint\!\!\!\!\fint_{Q_r}(1 + |D\mathbf{u}|)^p\,\mathrm{d}z\right)^{\frac{p-1}{p}}
$$

$$
\leqslant r^{\alpha - \frac{q-p}{p}(n+2)}(H^-)^{-1}\left(\fint\!\!\!\!\fint_{Q_r} H^-(1 + |D\mathbf{u}|)\,\mathrm{d}z\right)^{p-1}
$$

$$
\leqslant r^{\alpha - \frac{q-p}{p}(n+2)}\left((H^-)' \circ (H^-)^{-1}\right)\left(\fint\!\!\!\!\fint_{Q_r} H^-(1 + |D\mathbf{u}|)\,\mathrm{d}z\right).
$$

Therefore, inserting the above estimates into (4.7) and using (2.11) for $\varphi(t) := H^-(t)$, $\sigma := 1 + |D\boldsymbol{\ell}|$ and $s = |D\mathbf{u}_{\boldsymbol{\ell}}|$, we have

(4.8)
$$
I_1 \leqslant cr^{\beta_1}\left((H^-)' \circ (H^-)^{-1}\right)\left(\fint\!\!\!\!\fint_{Q_r} H^-(1 + |D\mathbf{u}|)\,\mathrm{d}z\right)
$$

$$
\leqslant cr^{\beta_1}\left((H^-)' \circ (H^-)^{-1}\right)\left(\fint\!\!\!\!\fint_{Q_r} H^-_{1+|D\boldsymbol{\ell}|}(|D\mathbf{u}_{\boldsymbol{\ell}}|)\,\mathrm{d}z + H^-(1 + |D\boldsymbol{\ell}|)\right),
$$

where $\beta_1 > 0$ is defined in (4.5).

We next estimate $I_2$ by decomposing $Q_r$ into $F_1 = \{z \in Q_r : 2|D\mathbf{u}_{\boldsymbol{\ell}}(z)| > 1 + |D\boldsymbol{\ell}|\}$ and $F_2 = \{z \in Q_r : 2|D\mathbf{u}_{\boldsymbol{\ell}}(z)| \leqslant 1 + |D\boldsymbol{\ell}|\}$, together with their characteristic functions $\chi_{F_1}$ and $\chi_{F_2}$. For $z \in F_1$, the inequality

$$(4.9) \qquad |D\mathbf{u}(z)| + 1 + |D\boldsymbol{\ell}| \leqslant |D\mathbf{u}_{\boldsymbol{\ell}}(z)| + 2(1 + |D\boldsymbol{\ell}|) \leqslant 5|D\mathbf{u}_{\boldsymbol{\ell}}(z)|$$

holds. Using this along with (A1), (2.14) with $\varphi(t) = H_1^-(t)$, (2.9), (2.10) and (4.9), we can estimate:

$$\int_0^1 |D_{\boldsymbol{\xi}}\mathbf{A}(z^-, sD\mathbf{u}_{\boldsymbol{\ell}}(z) + D\boldsymbol{\ell}) - D_{\boldsymbol{\xi}}\mathbf{A}(z^-, D\boldsymbol{\ell})| \, ds$$

$$\leqslant c\int_0^1 (H^-)''(1 + |sD\mathbf{u}(z) + (1-s)D\boldsymbol{\ell}|) \, ds + c(H^-)''(1 + |D\boldsymbol{\ell}|)$$

$$\leqslant c\int_0^1 \frac{(H_1^-)'(|sD\mathbf{u}(z) + (1-s)D\boldsymbol{\ell}|)}{|sD\mathbf{u}(z) + (1-s)D\boldsymbol{\ell}|} \, ds + c(H^-)''(1 + |D\boldsymbol{\ell}|)$$

$$\leqslant c\frac{(H^-)'(1 + |D\mathbf{u}(z)| + |D\boldsymbol{\ell}|)}{1 + |D\mathbf{u}(z)| + |D\boldsymbol{\ell}|} + c\frac{(H^-)'(1 + |D\mathbf{u}(z)| + |D\boldsymbol{\ell}|)}{1 + |D\boldsymbol{\ell}|}$$

$$\leqslant \frac{c}{1 + |D\boldsymbol{\ell}|}(H^-)'(|D\mathbf{u}_{\boldsymbol{\ell}}(z)|)$$

$$\leqslant \frac{c}{1 + |D\boldsymbol{\ell}|}(H^-)'(|D\mathbf{u}_{\boldsymbol{\ell}}(z)| + 1 + |D\boldsymbol{\ell}|)\frac{5|D\mathbf{u}_{\boldsymbol{\ell}}(z)|}{|D\mathbf{u}_{\boldsymbol{\ell}}(z)| + 1 + |D\boldsymbol{\ell}|} \leqslant \frac{c}{1 + |D\boldsymbol{\ell}|}(H_{1+|D\boldsymbol{\ell}|}^-)'(|D\mathbf{u}_{\boldsymbol{\ell}}(z)|).$$

Hence, by (2.9),

$$(4.10) \qquad \fint\!\!\!\!\fint_{Q_r} \chi_{F_1}(z)\left[\int_0^1 |D_{\boldsymbol{\xi}}\mathbf{A}(z^-, sD\mathbf{u}_{\boldsymbol{\ell}} + D\boldsymbol{\ell}) - D_{\boldsymbol{\xi}}\mathbf{A}(z^-, D\boldsymbol{\ell})| \, ds\right]|D\mathbf{u}_{\boldsymbol{\ell}}| \, dz$$

$$\leqslant \frac{c}{1 + |D\boldsymbol{\ell}|}\fint\!\!\!\!\fint_{Q_r} H_{1+|D\boldsymbol{\ell}|}^-(|D\mathbf{u}_{\boldsymbol{\ell}}|) \, dz.$$

On the other hand, if $z \in F_2$, applying (A4) with $z = z^-$, $\boldsymbol{\xi}_1 = D\boldsymbol{\ell}$ and $\boldsymbol{\xi}_2 = sD\mathbf{u}_{\boldsymbol{\ell}}(z) + D\boldsymbol{\ell}$ we have

$$\int_0^1 |D_{\boldsymbol{\xi}}\mathbf{A}(z^-, sD\mathbf{u}_{\boldsymbol{\ell}}(z) + D\boldsymbol{\ell}) - D_{\boldsymbol{\xi}}\mathbf{A}(z^-, D\boldsymbol{\ell})| \, ds \leqslant c\left(\frac{|D\mathbf{u}_{\boldsymbol{\ell}}(z)|}{1 + |D\boldsymbol{\ell}|}\right)^\gamma (H^-)''(1 + |D\boldsymbol{\ell}|)$$

and, since $(H^-)'$ is increasing, using also (2.9) and (2.10), we get

$$\frac{|D\mathbf{u}_{\boldsymbol{\ell}}(z)|^2}{(1 + |D\boldsymbol{\ell}|)^2} \leqslant \frac{(H^-)'(|D\mathbf{u}_{\boldsymbol{\ell}}(z)| + 1 + |D\boldsymbol{\ell}|)}{(H^-)'(1 + |D\boldsymbol{\ell}|)(1 + |D\boldsymbol{\ell}|)} \cdot \frac{|D\mathbf{u}_{\boldsymbol{\ell}}(z)|^2}{1 + |D\boldsymbol{\ell}|}$$

$$\leqslant \frac{(H^-)'(|D\mathbf{u}_{\boldsymbol{\ell}}(z)| + 1 + |D\boldsymbol{\ell}|)}{H^-(1 + |D\boldsymbol{\ell}|)} \cdot \frac{|D\mathbf{u}_{\boldsymbol{\ell}}(z)|^2}{|D\mathbf{u}_{\boldsymbol{\ell}}| + \frac{1}{2}(1 + |D\boldsymbol{\ell}|)} \leqslant c\frac{H_{1+|D\boldsymbol{\ell}|}^-(|D\mathbf{u}_{\boldsymbol{\ell}}|)}{H^-(1 + |D\boldsymbol{\ell}|)}.$$

Then combining these estimates, using the definition of $F_2$, and applying Hölder's inequality, we have

$$\fint\!\!\!\!\fint_{Q_r} \chi_{F_2}(z)\left[\int_0^1 |D_{\boldsymbol{\xi}}\mathbf{A}(z^-, sD\mathbf{u}_{\boldsymbol{\ell}} + D\boldsymbol{\ell}) - D_{\boldsymbol{\xi}}\mathbf{A}(z^-, D\boldsymbol{\ell})| \, ds\right]|D\mathbf{u}_{\boldsymbol{\ell}}| \, dz$$

$$(4.11) \qquad \leqslant c(H^-)'(1 + |D\boldsymbol{\ell}|)\fint\!\!\!\!\fint_{Q_r} \chi_{F_2}(z)\left(\frac{|D\mathbf{u}_{\boldsymbol{\ell}}|}{1 + |D\boldsymbol{\ell}|}\right)^{1+\gamma} \, dz$$

$$\leqslant c(H^-)'(1 + |D\boldsymbol{\ell}|)\left(\fint\!\!\!\!\fint_{Q_r} \frac{H_{1+|D\boldsymbol{\ell}|}^-(|D\mathbf{u}_{\boldsymbol{\ell}}|)}{H^-(1 + |D\boldsymbol{\ell}|)} \, dz\right)^{\frac{1+\gamma}{2}}.$$

Therefore, collecting (4.10) and (4.11), we obtain

$$(4.12) \qquad I_2 \leqslant c(H^-)'(1 + |D\boldsymbol{\ell}|)\left\{\fint\!\!\!\!\fint_{Q_r} \frac{H_{1+|D\boldsymbol{\ell}|}^-(|D\mathbf{u}_{\boldsymbol{\ell}}|)}{H^-(1 + |D\boldsymbol{\ell}|)} \, dz + \left(\fint\!\!\!\!\fint_{Q_r} \frac{H_{1+|D\boldsymbol{\ell}|}^-(|D\mathbf{u}_{\boldsymbol{\ell}}|)}{H^-(1 + |D\boldsymbol{\ell}|)} \, dz\right)^{\frac{1+\gamma}{2}}\right\}.$$

Finally, plugging (4.8) and (4.12) into (4.6), we get (4.4). This concludes the proof. $\qquad\square$

## 5. DECAY ESTIMATE

For $z_0 = (x_0, t_0) \in \Omega_T$ and $\rho > 0$ such that $Q_\rho(z_0) \Subset \Omega_T$, we define the excess functional

$$(5.1) \qquad \Phi(z_0, \rho, \mathbf{u}) = \iint_{Q_\rho(z_0)} H^-_{1+|(D\mathbf{u})_{z_0,\rho}|} \left( |D\mathbf{u} - (D\mathbf{u})_{z_0,\rho}| \right) \, dz + \rho^{\frac{\beta_1}{2}} \,,$$

where $H^-_{1+|(D\mathbf{u})_{z_0,\rho}|} := (H^-_{Q_\rho(z_0)})_{1+|(D\mathbf{u})_{z_0,\rho}|}$ and $\beta_1$ is given in (4.5), and linear map

$$(5.2) \qquad \boldsymbol{\ell}_{z_0,\rho,\mathbf{u}}(x) := \left( \frac{n+2}{\rho^2} \iint_{Q_\rho(z_0)} \mathbf{u}(x,t) \otimes (x - x_0) \, dx \, dt \right) (x - x_0) + (\mathbf{u})_{Q_\rho(z_0)} \,.$$

Note that $\boldsymbol{\ell}_{z_0,\rho,\mathbf{u}}$ is the minimizer of the functional

$$F(\boldsymbol{\ell}) = \iint_{Q_\rho(z_0)} |\mathbf{u} - \boldsymbol{\ell}|^2 \, dz \,,$$

where $\boldsymbol{\ell} = \boldsymbol{\ell}(x)$ is any affine function from $\mathbb{R}^n$ to $\mathbb{R}^N$, and

$$(5.3) \quad |D\boldsymbol{\ell}_{z_0,\rho,\mathbf{u}} - D\boldsymbol{\ell}| = \frac{n+2}{\rho^2} \left| \iint_{Q_\rho(z_0)} (\mathbf{u} - \boldsymbol{\ell}) \otimes (x - x_0) \, dx \, dt \right| \leqslant (n+2) \iint_{Q_\rho(z_0)} \frac{|\mathbf{u} - \boldsymbol{\ell}|}{\rho} \, dz \,,$$

see [8, Section 2.5]. The following Lemma collects some results of (almost) minimality proven for function $\boldsymbol{\ell}_{z_0,\rho,\mathbf{u}}$ in [38, Lemma 2.6, Lemma 2.7 and Remark 2.6].

**Lemma 5.1.** *Let $\varphi$ be an $N$-function satisfying $\Delta_2(\varphi, \varphi^*) < \infty$, and let $\mathbf{u} \in L^\varphi(Q_\rho(z_0), \mathbb{R}^N)$. There exists a constant $c = c(n, \Delta_2(\varphi, \varphi^*)) > 0$ such that for every affine function $\boldsymbol{\ell} : \mathbb{R}^n \to \mathbb{R}^N$,*

$$(5.4) \qquad \iint_{Q_\rho(z_0)} \varphi \left( \frac{|\mathbf{u} - \boldsymbol{\ell}_{z_0,\rho}|}{\rho} \right) \, dz \leqslant c \iint_{Q_\rho(z_0)} \varphi \left( \frac{|\mathbf{u} - \boldsymbol{\ell}|}{\rho} \right) \, dz \,.$$

*Moreover, if $D\mathbf{u} \in L^\varphi(Q_\rho(z_0), \mathbb{R}^{N \times n})$, then for every $\mathbf{Q} \in \mathbb{R}^{N \times n}$,*

$$(5.5) \qquad \iint_{Q_\rho(z_0)} \varphi_{1+|(D\mathbf{u})_{z_0,\rho}|} \left( |D\mathbf{u} - (D\mathbf{u})_{z_0,\rho}| \right) \, dz \leqslant c \iint_{Q_\rho(z_0)} \varphi_{1+|\mathbf{Q}|} \left( |D\mathbf{u} - \mathbf{Q}| \right) \, dz \,.$$

Now, we prove a decay estimate for the excess $\Phi(z_0, \rho, \mathbf{u})$.

**Lemma 5.2.** *Let $\mathbf{u}$ be a weak solution to (1.3), $K \geqslant 1$, and $\beta \in (0, \frac{\beta_1}{4})$ where $\beta_1$ is given in (4.5). There exist small $r_0 \in (0, R_0]$ and $\delta_0, \tau \in (0, 1)$ depending on $n, N, p, q, \alpha, L, \nu, \gamma, K$ and $\beta$ such that if $Q_{2r}(z_0) \Subset \Omega_T$ with $r \in (0, r_0]$,*

$$(5.6) \qquad |(D\mathbf{u})_{z_0,r}| \leqslant K$$

*and*

$$(5.7) \qquad \Phi(z_0, r, \mathbf{u}) \leqslant \delta_0 \,,$$

*then*

$$(5.8) \qquad |(D\mathbf{u})_{z_0,\tau r}| \leqslant |(D\mathbf{u})_{z_0,r}| + c_1 \tau^{-n-2} \Phi(z_0, r, \mathbf{u})^{\frac{1}{2}}$$

*for some $c_1 = c_1(n) > 0$, and*

$$(5.9) \qquad \Phi(z_0, \tau r, \mathbf{u}) \leqslant \tau^{2\beta} \Phi(z_0, r, \mathbf{u}) \,.$$

*Proof. Step 1. (Setting)* All constants $c$, including implicit constants, may depend also on $K$. For simplicity, we write $Q_\rho := Q_\rho(z_0)$ for any $\rho > 0$, $z^- := z^-_{z_0,r}$, $H^\pm_\sigma := (H^\pm_{Q_r})_\sigma$, $H^{\pm,*}_\sigma := (H^\pm_\sigma)^*$, $\boldsymbol{\ell}_{\tau r} := \boldsymbol{\ell}_{z_0,\tau r,\mathbf{u}}$ (see (5.2)), and

$$\tilde{\boldsymbol{\ell}}(x) := (D\mathbf{u})_{z_0,r}(x - x_0) + (\mathbf{u})_{z_0,r} \,.$$

Note that, by (5.6),

$$(5.10) \qquad 1 + |D\tilde{\boldsymbol{\ell}}| \sim 1 \,.$$

Set

$$(5.11) \qquad \mu := \left( \iint_{Q_r} H^-_{1+|D\tilde{\ell}|}(|D\mathbf{u}_{\tilde{\ell}}|) \, dz + r^{\frac{\beta_1}{2}} \right)^{1/2} = \Phi(z_0, r, \mathbf{u})^{\frac{1}{2}} \leqslant \delta_0^{\frac{1}{2}} \leqslant 1 \,.$$

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

$$\partial_t \mathbf{u}_{\tilde{\ell}} = \partial_t \mathbf{u} = \operatorname{div} \mathbf{G} \quad \text{in} \quad Q_r, \quad \text{where} \quad \mathbf{G} := \mathbf{A}(z, D\mathbf{u}) - \mathbf{A}(z^-, D\tilde{\ell}) \,,$$

in the distributional sense, and that

$$\begin{aligned}
|\mathbf{G}| &\leqslant |\mathbf{A}(z, D\mathbf{u}) - \mathbf{A}(z^-, D\mathbf{u})| + |\mathbf{A}(z^-, D\mathbf{u}) - \mathbf{A}(z^-, D\tilde{\ell})| \\
(5.12) \qquad &\leqslant c(a(z) - a(z^-))(1 + |D\mathbf{u}|)^{q-1} + cr^{\beta_0} H'(z, 1 + |D\mathbf{u}|) + (H^-_{1+|D\tilde{\ell}|})'(|D\mathbf{u}_{\tilde{\ell}}|) \\
&\leqslant c(a(z) - a(z^-))(1 + |D\mathbf{u}|)^{q-1} + cr^{\beta_0}(H^-)'(1 + |D\mathbf{u}|) + (H^-_{1+|D\tilde{\ell}|})'(|D\mathbf{u}_{\tilde{\ell}}|) \,.
\end{aligned}$$

Indeed, the first term above $|\mathbf{A}(z, D\mathbf{u}) - \mathbf{A}(z^-, D\mathbf{u})|$ can be estimated by (A3), while by using (2.18) with $\mathbf{P} = D\mathbf{u}$ and $\mathbf{Q} = D\tilde{\ell}$, we obtain

$$|\mathbf{A}(z^-, D\mathbf{u}) - \mathbf{A}(z^-, D\tilde{\ell})| \lesssim (H^-_{1+|D\tilde{\ell}|})'(|D\mathbf{u}_{\tilde{\ell}}|) \,.$$

Let

$$(5.13) \qquad p_1 := \min \left\{ \frac{p}{q-1}, p, \frac{q}{q-1}, 2 \right\} > 1 \quad \text{and} \quad p_0 := \frac{1+p_1}{2}.$$

Then one can see that both functions $\psi(s) = H^{-,*}_\sigma(s^{1/p_1})$ and $\psi(s) = H^-_\sigma(s^{1/p_1})$ with $\sigma \geqslant 0$ satisfy the assumption in Lemma 2.1. We first estimate, using (5.12),

$$\begin{aligned}
\iint_{Q_r} |\mathbf{G}|^{p_1} \, dz &\leqslant c \iint_{Q_r} \left[ (a(z) - a(z^-))(1 + |D\mathbf{u}|)^{q-1} \right]^{p_1} \, dz + c \iint_{Q_r} \left[ r^{\beta_0}(H^-)'(1 + |D\mathbf{u}|) \right]^{p_1} \, dz \\
&\quad + c \iint_{Q_r} (H^-_{1+|D\tilde{\ell}|})'(|D\mathbf{u}_{\tilde{\ell}}|)^{p_1} \, dz \\
&=: II_1 + II_2 + II_3 \,.
\end{aligned}$$

By Hölder's inequality, (1.6), (4.3), (2.11), (5.10), (4.5) and (5.11), we obtain that

$$\begin{aligned}
(II_1)^{\frac{1}{p_1}} &\leqslant cr^\alpha \left( \iint_{Q_r} (1 + |D\mathbf{u}|)^p \, dz \right)^{\frac{q-1}{p}} \\
&\leqslant cr^\alpha \left( \iint_{Q_r} H^-(1 + |D\mathbf{u}|) \, dz \right)^{\frac{q-1}{p}} \\
&\leqslant cr^\alpha \left( \iint_{Q_r} H^-_{1+|D\tilde{\ell}|}(|D\mathbf{u}_{\tilde{\ell}}|) \, dz + H^-(1 + |D\tilde{\ell}|) \right)^{\frac{q-1}{p}} \leqslant cr^\alpha \leqslant c\mu.
\end{aligned}$$

Similarly, in addition, using the Jensen type inequality in Lemma 2.1 with $\psi(s) = H^{-,*}_\sigma(s^{1/p_1})$ and using $H^{-,*}_\sigma \circ (H^-_\sigma)' \sim H^-_\sigma$, we also obtain that

$$\begin{aligned}
(II_2)^{\frac{1}{p_1}} &\leqslant cr^{\beta_0}(H^{-,*})^{-1} \left( \iint_{Q_r} H^-(1 + |D\mathbf{u}|) \, dz \right) \\
&\leqslant cr^{\beta_0}(H^{-,*})^{-1} \left( \iint_{Q_r} H^-_{1+|D\tilde{\ell}|}(|D\mathbf{u}_{\tilde{\ell}}|) \, dz + H^-(1 + |D\tilde{\ell}|) \right) \leqslant cr^{\beta_0} \leqslant c\mu \,,
\end{aligned}$$

and

$$(II_3)^{\frac{1}{p_1}} \leqslant (H^{-,*}_{1+|D\tilde{\ell}|})^{-1} \left( \iint_{Q_r} H^-_{1+|D\tilde{\ell}|}(|D\mathbf{u}_{\tilde{\ell}}|) \, dz \right) \leqslant c(H^{-,*}_{1+|D\tilde{\ell}|})^{-1}(\mu^2) \,.$$

Therefore,

$$\left( \iint_{Q_r} |\mathbf{G}|^{p_1} \, dz \right)^{\frac{1}{p_1}} \leqslant c \left( \mu + (H^{-,*}_{1+|D\tilde{\ell}|})^{-1}(\mu^2) \right) \,.$$

Moreover, using the Jensen type inequality in Lemma 2.1 with $\psi(s) = H_\sigma^-(s^{1/p_1})$, we also obtain

$$\left( \fint\!\!\!\!\fint_{Q_r} |D\mathbf{u}_{\tilde{\boldsymbol{\ell}}}|^{p_1} \, dz \right)^{\frac{1}{p_1}} \leqslant c(H_{1+|D\tilde{\boldsymbol{\ell}}|}^-)^{-1} \left( \fint\!\!\!\!\fint_{Q_r} H_{1+|D\tilde{\boldsymbol{\ell}}|}^-(|D\mathbf{u}_{\tilde{\boldsymbol{\ell}}}|) \, dz \right) \leqslant c(H_{1+|D\tilde{\boldsymbol{\ell}}|}^-)^{-1}(\mu^2) \,.$$

We further observe from (5.10), (2.9) and (2.10) that for any $s \in (0,1]$,

$$(H_{1+|D\tilde{\boldsymbol{\ell}}|}^\pm)'(s) = \frac{(H^\pm)'(1+|D\tilde{\boldsymbol{\ell}}|+s)}{1+|D\tilde{\boldsymbol{\ell}}|+s} s \sim s \,,$$

which together with (2.9) and (2.6) implies

$$(5.14) \qquad s^2 \sim H_{1+|D\tilde{\boldsymbol{\ell}}|}^\pm(s) \quad \text{and} \quad s^2 \sim ((H_{1+|D\tilde{\boldsymbol{\ell}}|}^\pm)')^{-1}(s)s \sim (H_{1+|D\tilde{\boldsymbol{\ell}}|}^{\pm,*})'(s)s \sim H_{1+|D\tilde{\boldsymbol{\ell}}|}^{\pm,*}(s) \,.$$

Finally, applying (5.14) to $s = \mu$, we obtain

$$(5.15) \qquad \left( \fint\!\!\!\!\fint_{Q_r} |D\mathbf{u}_{\tilde{\boldsymbol{\ell}}}|^{p_1} \, dz + \fint\!\!\!\!\fint_{Q_r} |\mathbf{G}|^{p_1} \, dz \right)^{\frac{1}{p_1}} \leqslant c\mu \,.$$

Therefore, since the estimate (4.4) with $\boldsymbol{\ell} = \tilde{\boldsymbol{\ell}}$, together with (5.10), (5.11) and (5.14) for $s = \mu$, yields

$$\frac{1}{|Q_r|} \left| \iint_{Q_r} \mathbf{u}_{\tilde{\boldsymbol{\ell}}} \cdot \boldsymbol{\zeta}_t - \langle D_{\boldsymbol{\xi}}\mathbf{A}(z^-, D\tilde{\boldsymbol{\ell}}) D\mathbf{u}_{\tilde{\boldsymbol{\ell}}} \,|\, D\boldsymbol{\zeta} \rangle \, dz - \left[ \int_{B_r} \mathbf{u}_{\tilde{\boldsymbol{\ell}}} \cdot \boldsymbol{\zeta} \, dx \right]_{t=-r^2}^{t=r^2} \right|$$
$$\leqslant c \left( r^{\frac{\beta_1}{2}} + \delta_0^{\frac{\gamma}{2}} \right) \mu \|D\boldsymbol{\zeta}\|_{L^\infty(B_r, \mathbb{R}^{N\times n})} \,,$$

we may apply Theorem 4.2 to $\mathcal{A} := D_{\boldsymbol{\xi}}\mathbf{A}(z^-, D\tilde{\boldsymbol{\ell}})$ (this satisfies the Legendre-Hadamard condition by (A1) and (A2) and (5.10)), $\psi(s) := s^{p_0}$ and $\sigma_0 := \frac{p_1}{p_0} - 1$, where $p_1$ and $p_0$ are given in (5.13). Thus, for $\varepsilon \in (0,1)$ to be determined small later, there exist small $\delta_0, r > 0$ depending on $n, N, L, \nu, p, q, \gamma, K$ and $\varepsilon$ such that

$$(5.16) \qquad \fint\!\!\!\!\fint_{Q_r} |D\mathbf{u}_{\tilde{\boldsymbol{\ell}}} - D\mathbf{h}|^{p_0} \, dz \leqslant \varepsilon \mu^{p_0} \,,$$

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

$$(5.17) \qquad \begin{cases} \partial_t \mathbf{h} - \operatorname{div}(\mathcal{A}D\mathbf{h}) = \mathbf{0} & \text{in } Q_r, \\ \mathbf{h} = \mathbf{u}_{\tilde{\boldsymbol{\ell}}} & \text{on } \partial_{\mathrm{p}}Q_r \,. \end{cases}$$

Note that by the Lipschitz estimate (4.2), (5.16), (5.15), and (5.14) we have that

$$(5.18) \qquad \begin{aligned} \fint\!\!\!\!\fint_{Q_{r/2}} H_{1+|D\tilde{\boldsymbol{\ell}}|}^-(|D\mathbf{h}|) \, dz &\leqslant c H_{1+|D\tilde{\boldsymbol{\ell}}|}^- \left( \fint\!\!\!\!\fint_{Q_r} |D\mathbf{h}| \, dz \right) \\ &\leqslant c H_{1+|D\tilde{\boldsymbol{\ell}}|}^- \left( \left( \fint\!\!\!\!\fint_{Q_r} |D\mathbf{u}_{\tilde{\boldsymbol{\ell}}}|^{p_0} \, dz \right)^{\frac{1}{p_0}} + \mu \right) \\ &\leqslant c H_{1+|D\tilde{\boldsymbol{\ell}}|}^-(\mu) \leqslant c\mu^2 \,. \end{aligned}$$

Therefore, recalling $1 < \tilde{p} \leqslant \tilde{q}$ given in (2.8) and $\theta \in (0,1)$ given in Lemma 2.2 and additionally assumed that $\tilde{q}\theta > 1$ and $\tilde{p}\theta > 2n/(n+2)$, letting $\kappa_0 \in (0,1)$ such that $\theta = \frac{1-\kappa_0}{\tilde{q}} + \kappa_0$ (i.e., $\kappa_0 = \frac{\tilde{q}\theta-1}{\tilde{q}-\theta}$), and using Hölder's inequality, the Jensen type inequality in Lemma 2.1 with $\psi^{-1}(s) :=$

$[H^-_{1+|D\tilde{\ell}|}(s)]^{1/\tilde{q}}$, the estimates (5.16) and (5.18), (5.11), and (5.14), we obtain

$$
\begin{aligned}
(5.19) \quad & \left( \fint\!\!\!\!\!\fint_{Q_{r/2}} \left[ H^-_{1+|D\tilde{\ell}|}(|D\mathbf{u}_{\tilde{\ell}} - D\mathbf{h}|) \right]^\theta \, \mathrm{d}z \right)^{\frac{1}{\theta}} \\
& \leqslant \left( \fint\!\!\!\!\!\fint_{Q_{r/2}} \left[ H^-_{1+|D\tilde{\ell}|}(|D\mathbf{u}_{\tilde{\ell}} - D\mathbf{h}|) \right]^{\frac{1}{\tilde{q}}} \, \mathrm{d}z \right)^{\frac{1-\kappa_0}{\theta}} \left( \fint\!\!\!\!\!\fint_{Q_{r/2}} H^-_{1+|D\tilde{\ell}|}(|D\mathbf{u}_{\tilde{\ell}} - D\mathbf{h}|) \, \mathrm{d}z \right)^{\frac{\kappa_0}{\theta}} \\
& \leqslant c \left[ H^-_{1+|D\tilde{\ell}|} \left( \left( \fint\!\!\!\!\!\fint_{Q_{r/2}} |D\mathbf{u}_{\tilde{\ell}} - D\mathbf{h}|^{p_0} \, \mathrm{d}z \right)^{\frac{1}{p_0}} \right) \right]^{\frac{1-\kappa_0}{\tilde{q}\theta}} \mu^{\frac{2\kappa_0}{\theta}} \\
& \leqslant c\varepsilon^{\frac{\tilde{p}(1-\kappa_0)}{p_0\tilde{q}\theta}} H^-_{1+|D\tilde{\ell}|}(\mu)^{\frac{1-\kappa_0}{\tilde{q}\theta}} \mu^{2\frac{\kappa_0}{\theta}} \leqslant c\varepsilon^{\frac{\tilde{p}(1-\kappa_0)}{p_0\tilde{q}\theta}} \mu^2 \,.
\end{aligned}
$$

Therefore, by Lemma 2.2 and (5.11) with (4.5),

$$
\begin{aligned}
\fint\!\!\!\!\!\fint_{Q_{r/2}} H^+_{1+|D\tilde{\ell}|} \left( \frac{|\mathbf{u}_{\tilde{\ell}} - \mathbf{h}|}{r} \right) \, \mathrm{d}z & \leqslant c \left( \fint\!\!\!\!\!\fint_{Q_{r/2}} \left[ H^-_{1+|D\tilde{\ell}|}(|D\mathbf{u}_{\tilde{\ell}} - D\mathbf{h}|) \right]^\theta \, \mathrm{d}z \right)^{1/\theta} \\
& \quad + c \left( r^{\alpha - \frac{(q-p)(n+2)}{p}} + r^\alpha (1 + |D\tilde{\ell}|)^{q-p} \right) (1 + |D\tilde{\ell}|)^p \\
& \leqslant c\varepsilon^{\frac{\tilde{p}(1-\kappa_0)}{p_0\tilde{q}\theta}} \mu^2 + r^{\alpha - \frac{(q-p)(n+2)}{p}} \\
& \leqslant c \left( \varepsilon^{\frac{\tilde{p}(1-\kappa_0)}{p_0\tilde{q}\theta}} + r^{\frac{1}{2}[\alpha - \frac{(q-p)(n+2)}{p}]} \right) \mu^2 \,.
\end{aligned}
$$

Moreover, by Sobolev inequality for each time $t$, the Jensen type inequality in Lemma 2.1 with $\psi(s) = H^-_{1+|D\tilde{\ell}|}(s^{\frac{n+2}{2n}})$, since $p > \frac{2n}{n+2}$, and (5.19) with (5.14), we also have

$$
\begin{aligned}
\fint\!\!\!\!\!\fint_{Q_{r/2}} \left| \frac{\mathbf{u}_{\tilde{\ell}} - \mathbf{h}}{r} \right|^2 \, \mathrm{d}z & \leqslant c \left( \fint\!\!\!\!\!\fint_{Q_{r/2}} |D\mathbf{u}_{\tilde{\ell}} - D\mathbf{h}|^{\frac{2n}{n+2}} \, \mathrm{d}z \right)^{\frac{n+2}{n}} \\
& \leqslant c \left[ (H^-_{1+|D\tilde{\ell}|})^{-1} \left( \left( \fint\!\!\!\!\!\fint_{Q_{r/2}} [H^-_{1+|D\tilde{\ell}|}(|D\mathbf{u}_{\tilde{\ell}} - D\mathbf{h}|)]^\theta \, \mathrm{d}z \right)^{\frac{1}{\theta}} \right) \right]^2 \\
& \leqslant c \left[ (H^-_{1+|D\tilde{\ell}|})^{-1} \left( \varepsilon^{\frac{\tilde{p}(1-\kappa_0)}{p_0\tilde{q}\theta}} \mu^2 \right) \right]^2 \\
& \leqslant c\varepsilon^{\frac{2(\tilde{p}-1)(1-\kappa_0)}{p_0\tilde{q}\theta}} \mu^2 \,.
\end{aligned}
$$

Therefore, collecting the last two estimates we obtain

$$
(5.20) \qquad \fint\!\!\!\!\!\fint_{Q_{r/2}} \left[ \left| \frac{\mathbf{u}_{\tilde{\ell}} - \mathbf{h}}{r} \right|^2 + H^+_{1+|D\tilde{\ell}|} \left( \frac{|\mathbf{u}_{\tilde{\ell}} - \mathbf{h}|}{r} \right) \right] \, \mathrm{d}z \leqslant c\tilde{\varepsilon}\mu^2 \,,
$$

where

$$
\tilde{\varepsilon} := \varepsilon^{\frac{\tilde{p}(1-\kappa_0)}{p_0\tilde{q}\theta}} + r^{\frac{1}{2}[\alpha - \frac{(q-p)(n+2)}{p}]} + \varepsilon^{\frac{2(\tilde{p}-1)(1-\kappa_0)}{p_0\tilde{q}\theta}} \,.
$$

*Step 3 (Decay estimates).* Let $\tau \in (0, \frac{1}{4})$ be a small constant determined later. We shall show that

$$
(5.21) \qquad \fint\!\!\!\!\!\fint_{Q_{2\tau r}} \left| \frac{\mathbf{u}_{\ell_{2\tau r}}}{2\tau r} \right|^2 \, \mathrm{d}z + \fint\!\!\!\!\!\fint_{Q_{2\tau r}} H^+_{1+|D\ell_{2\tau r}|} \left( \frac{|\mathbf{u}_{\ell_{2\tau r}}|}{2\tau r} \right) \, \mathrm{d}z \leqslant c_2 \left( \tau^{-(n+2+\tilde{q})}\tilde{\varepsilon} + \tau^2 \right) \mu^2 \,,
$$

for some constant $c_2 \geqslant 1$, where $\ell_{2\tau r} := \ell_{z_0, 2\tau r, \mathbf{u}}$ is the linear function defined in (5.2).

Let $\mathbf{h}$ be the $\mathcal{A}$-caloric map of (5.17), and define

$$
\tilde{\ell}_{\rho, \mathbf{h}} := \tilde{\ell}_{z_0, \rho, \mathbf{h}}(x) := (D\mathbf{h})_{Q_\rho(z_0)}(x - x_0) + \mathbf{h}(z_0), \quad z_0 = (x_0, t_0) \,.
$$

Observe that for $z \in Q_{r/2}$,

$$\mathbf{h}(z) - \mathbf{h}(z_0) = \left( \int_0^1 D\mathbf{h}(sz + (1-s)z_0) \, \mathrm{d}s \right) (x - x_0) + \left( \int_0^1 \partial_t \mathbf{h}(sz + (1-s)z_0) \, \mathrm{d}s \right) (t - t_0),$$

from which together with (4.1) with $\mathbf{h}$ in place of $D_i \mathbf{h}$, $i = 1, 2, \ldots, n$, (4.2), (5.15) and (5.16)

$$
\begin{aligned}
\mathbf{h}(z) - \mathbf{h}(z_0) &\leqslant \frac{r}{2} \|D\mathbf{h}\|_{L^\infty(Q_{r/2})} + \frac{r^2}{4} \|\mathbf{h}_t\|_{L^\infty(Q_{r/2})} \\
(5.22) \qquad &\leqslant c \left( r \|D\mathbf{h}\|_{L^\infty(Q_{r/2})} + r^2 \|D^2\mathbf{h}\|_{L^\infty(Q_{r/2})} \right) \\
&\leqslant cr \fint_{Q_r} |D\mathbf{h}| \, \mathrm{d}z \leqslant cr \left( \fint_{Q_r} |D\mathbf{h} - D\mathbf{u}_{\tilde{\ell}}| \, \mathrm{d}z + \fint_{Q_r} |D\mathbf{u}_{\tilde{\ell}}| \, \mathrm{d}z \right) \leqslant cr\mu.
\end{aligned}
$$

Similarly, we also obtain that

$$
\begin{aligned}
\sup_{Q_{2\tau r}} |\mathbf{h} - \tilde{\ell}_{2\tau r, \mathbf{h}}| &\leqslant 2\tau r \sup_{Q_{2\tau r}} |D\mathbf{h} - (D\mathbf{h})_{Q_{2\tau r}}| + (2\tau r)^2 \sup_{Q_{2\tau r}} |\partial_t \mathbf{h}| \\
&\leqslant (2\tau r)^2 \sup_{Q_{2\tau r}} |D^2\mathbf{h}| + (2\tau r)^3 \sup_{Q_{2\tau r}} |D\partial_t \mathbf{h}| + (2\tau r)^2 \sup_{Q_{2\tau r}} |\partial_t \mathbf{h}| \\
(5.23) \qquad &\leqslant c\tau^2 r \left( r \sup_{Q_{r/2}} |D^2\mathbf{h}| + r^2 \sup_{Q_{r/2}} |D^3\mathbf{h}| \right) \\
&\leqslant c\tau^2 r \left( \iint_{Q_r} |D\mathbf{h}|^{p_0} \, \mathrm{d}z \right)^{1/p_0} \leqslant c\tau^2 r\mu.
\end{aligned}
$$

Let us first consider the second integral in (5.21). Observe that by (5.3) with $\rho = 2\tau r$ and $\ell = \tilde{\ell} + \mathbf{h}(z_0)$, (5.20), (5.22) and (5.11),

$$
\begin{aligned}
|D\ell_{2\tau r}| &\leqslant |D\tilde{\ell}| + (n+2) \iint_{Q_{2\tau r}} \frac{|\mathbf{u}_{\tilde{\ell}} - \mathbf{h}(z_0)|}{2\tau r} \, \mathrm{d}z \\
&\leqslant |D\tilde{\ell}| + \frac{c}{\tau^{n+3}} \left( \iint_{Q_{r/2}} \left| \frac{\mathbf{u}_{\tilde{\ell}} - \mathbf{h}}{r} \right|^2 \, \mathrm{d}z + \iint_{Q_{r/2}} \left| \frac{\mathbf{h} - \mathbf{h}(z_0)}{r} \right|^2 \, \mathrm{d}z \right)^{1/2} \\
&\leqslant |D\tilde{\ell}| + c\tau^{-n-3}\mu \\
&\leqslant |D\tilde{\ell}| + c\tau^{-n-3}\delta_0^{\frac{1}{2}} \leqslant K + 1.
\end{aligned}
$$

In the last inequality, we have chosen $\delta_0 > 0$ sufficiently small such that $c(\frac{\tau}{2})^{-n-3}\delta_0^{1/2} \leqslant 1$ with $\tau > 0$ that will be determined in the end of the proof. Then it follows that $1 + |D\ell_{2\tau r}| \sim 1 \sim 1 + |D\tilde{\ell}|$, and hence $H^+_{1+|D\ell_{2\tau r}|}(s) \sim H^+_{1+|D\tilde{\ell}|}(s)$. Moreover, by (5.4) with $\varphi = H^+_{1+|D\tilde{\ell}|}$ and $\ell = \tilde{\ell} + \tilde{\ell}_{2\tau r, \mathbf{h}}$,

$$
\begin{aligned}
\iint_{Q_{2\tau r}} H^+_{1+|D\ell_{2\tau r}|} & \left( \frac{|\mathbf{u}_{\ell_{2\tau r}}|}{2\tau r} \right) \, \mathrm{d}z \leqslant c \iint_{Q_{2\tau r}} H^+_{1+|D\tilde{\ell}|} \left( \frac{|\mathbf{u}_{\ell_{2\tau r}}|}{2\tau r} \right) \, \mathrm{d}z \\
&\leqslant c \iint_{Q_{2\tau r}} H^+_{1+|D\tilde{\ell}|} \left( \frac{|\mathbf{u}_{\tilde{\ell}} - \tilde{\ell}_{2\tau r, \mathbf{h}}|}{2\tau r} \right) \, \mathrm{d}z \\
(5.24) \qquad &\leqslant c \iint_{Q_{2\tau r}} H^+_{1+|D\tilde{\ell}|} \left( \frac{|\mathbf{u}_{\tilde{\ell}} - \mathbf{h}|}{2\tau r} \right) \, \mathrm{d}z + c \iint_{Q_{2\tau r}} H^+_{1+|D\tilde{\ell}|} \left( \frac{|\mathbf{h} - \tilde{\ell}_{2\tau r, \mathbf{h}}|}{2\tau r} \right) \, \mathrm{d}z \\
&\leqslant c\tau^{-(n+2+\tilde{q})} \iint_{Q_{r/2}} H^+_{1+|D\tilde{\ell}|} \left( \frac{|\mathbf{u}_{\tilde{\ell}} - \mathbf{h}|}{r} \right) \, \mathrm{d}z + c \iint_{Q_{2\tau r}} H^+_{1+|D\tilde{\ell}|} \left( \frac{|\mathbf{h} - \tilde{\ell}_{2\tau r, \mathbf{h}}|}{2\tau r} \right) \, \mathrm{d}z.
\end{aligned}
$$

Combining the estimates (5.24), (5.20), (5.23) and (5.14) yields

$$\iint_{Q_{2\tau r}} H^+_{1+|D\ell_{2\tau r}|} \left( \frac{|\mathbf{u}_{\ell_{2\tau r}}|}{2\tau r} \right) \, \mathrm{d}z \leqslant c\tau^{-(n+2+\tilde{q})} \tilde{\varepsilon}\mu^2 + cH^+_{1+|D\tilde{\ell}|}(\tau\mu) \leqslant c\tau^{-(n+2+\tilde{q})} \tilde{\varepsilon}\mu^2 + c\tau^2 \mu^2.$$

In a similar same way, we can estimate the first integral in (5.21) as follows:

$$\iint_{Q_{2\tau r}} \left| \frac{\mathbf{u}_{\boldsymbol{\ell}_{2\tau r}}}{2\tau r} \right|^2 \mathrm{d}z \leqslant c \iint_{Q_{2\tau r}} \left| \frac{\mathbf{u}_{\tilde{\boldsymbol{\ell}}} - \tilde{\boldsymbol{\ell}}_{2\tau r, \mathbf{h}}}{2\tau r} \right|^2 \mathrm{d}z$$

$$\leqslant c \iint_{Q_{2\tau r}} \left| \frac{\mathbf{u}_{\tilde{\boldsymbol{\ell}}} - \mathbf{h}}{2\tau r} \right|^2 \mathrm{d}z + \iint_{Q_{2\tau r}} \left| \frac{\mathbf{h} - \tilde{\boldsymbol{\ell}}_{2\tau r, \mathbf{h}}}{2\tau r} \right|^2 \mathrm{d}z$$

$$\leqslant c \tau^{-(n+4)} \tilde{\varepsilon} \mu^2 + c \tau^2 \mu^2.$$

Therefore, we obtain (5.21).

Now, we prove the inequality (5.9). By (5.1), (5.5) for $\mathbf{Q} = D\boldsymbol{\ell}_{2\tau r}$, the Caccioppoli inequality (3.2) for $\boldsymbol{\ell} = \boldsymbol{\ell}_{2\tau r}$, and (5.21), we obtain

$$\Phi(z_0, \tau r, \mathbf{u}) = \iint_{Q_{\tau r}} H^-_{1+|(D\mathbf{u})_{z_0,\tau r}|} \left( |D\mathbf{u} - (D\mathbf{u})_{z_0,\tau r}| \right) \mathrm{d}z + (\tau r)^{\frac{\beta_1}{2}}$$

$$\leqslant c \iint_{Q_{\tau r}} H_{1+|D\boldsymbol{\ell}_{2\tau r}|}(z, |D\mathbf{u}_{\boldsymbol{\ell}_{2\tau r}}|) \mathrm{d}z + (\tau r)^{\frac{\beta_1}{2}}$$

$$\leqslant c \left\{ \iint_{Q_{2\tau r}} \left| \frac{\mathbf{u}_{\boldsymbol{\ell}_{2\tau r}}}{2\tau r} \right|^2 + H^+_{1+|D\boldsymbol{\ell}_{2\tau r}|} \left( \frac{|\mathbf{u}_{\boldsymbol{\ell}_{2\tau r}}|}{2\tau r} \right) \mathrm{d}z + (\tau r)^{\beta_1} + (\tau r)^{\frac{\beta_1}{2}} \right\}$$

$$\leqslant c_3 \left( \tau^{-(n+2+\tilde{q})} \tilde{\varepsilon} + \tau^2 \right) \mu^2 + c_3 \tau^{\frac{\beta_1}{2}} r^{\frac{\beta_1}{2}}$$

$$\leqslant \tau^{2\beta} \Phi(z_0, r, \mathbf{u}),$$

where we chose $\tau \in (0, \frac{1}{4})$ and $\tilde{\varepsilon} \in (0, 1)$ such that

$$c_3 \tau^{\frac{\beta_1}{2} - 2\beta} \leqslant \frac{1}{2}, \quad 4c_3 \tau^{2-2\beta} \leqslant \frac{1}{4}, \quad \text{and} \quad c_3 \tau^{-(n+2+\tilde{q})} \tilde{\varepsilon} \leqslant \frac{1}{4} \tau^{2\beta}.$$

(Then $\delta_0$ is also fixed.) This corresponds to (5.9).

Finally, using Jensen's inequality, (5.11) and (5.14), we obatin

$$|(D\mathbf{u})_{z_0,\tau r}| \leqslant |(D\mathbf{u})_{z_0,\tau r} - D\tilde{\boldsymbol{\ell}}| + |D\tilde{\boldsymbol{\ell}}|$$

$$\leqslant \tau^{-(n+2)} \iint_{Q_r} |D\mathbf{u}_{\tilde{\boldsymbol{\ell}}}| \, \mathrm{d}z + |D\tilde{\boldsymbol{\ell}}|$$

$$\leqslant c \tau^{-(n+2)} (H^-_{1+|D\tilde{\boldsymbol{\ell}}|})^{-1} \left( \iint_{Q_r} H^-_{1+|D\tilde{\boldsymbol{\ell}}|} (|D\mathbf{u}_{\tilde{\boldsymbol{\ell}}}|) \, \mathrm{d}z \right) + |D\tilde{\boldsymbol{\ell}}|$$

$$\leqslant c \tau^{-(n+2)} (H^-_{1+|D\tilde{\boldsymbol{\ell}}|})^{-1} (\mu^2) + |D\tilde{\boldsymbol{\ell}}|$$

$$\leqslant c \tau^{-(n+2)} \Phi(z_0, r, \mathbf{u})^{1/2} + |D\tilde{\boldsymbol{\ell}}|$$

$$\leqslant c_1 \tau^{-(n+2)} \delta_0^{1/2} + |D\tilde{\boldsymbol{\ell}}|$$

for some constant $c_1 \geqslant 0$, which implies (5.8). This completes the proof. □

The excess decay estimate of Lemma 5.2 can be iterated on each scale, as expressed by the following lemma.

**Lemma 5.3** (Iteration). *Let $\mathbf{u}$ be a weak solution to* (1.3), $K \geqslant 1$, *and* $\beta \in (0, \frac{\beta_1}{4})$ *where* $\beta_1$ *is given in* (4.5). *There exist small* $r_0 \in (0, R_0]$ *and* $\delta_1, \tau_0 \in (0, 1)$ *depending on* $n, N, p, q, \alpha, L, \nu, \gamma, K$ *and* $\beta$ *such that if* $Q_r(z_0) \Subset \Omega_T$ *with* $r \in (0, r_0]$,

$$1 + |(D\mathbf{u})_{z_0,r}| \leqslant K \quad \text{and} \quad \Phi(z_0, r, \mathbf{u}) \leqslant \delta_1,$$

*then for each* $j \in \mathbb{N} \cup \{0\}$ *we have*

(i) $\Phi(z_0, \tau^j r, \mathbf{u}) \leqslant \tau^{2\beta j} \Phi(z_0, r, \mathbf{u})$,

(ii) $|(D\mathbf{u})_{z_0,\tau^j r}| \leqslant K - \tau^{\beta j}$,

*and moreover, for every $0 < \rho \leqslant r$*

$$(5.25) \qquad \fint\kern-0.9em\fint_{Q_\rho(z_0)} H^-_{1+|(D\mathbf{u})_{z_0,\rho}|}(|D\mathbf{u} - (D\mathbf{u})_{z_0,\rho}|)\,\mathrm{d}z \leqslant \bar{c}\left(\frac{\rho}{r}\right)^{2\beta}\Phi(z_0, r, \mathbf{u})$$

*for some $\bar{c} = \bar{c}(n, N, p, q, \alpha, L, \nu, \gamma, K, \beta)$.*

*Proof.* Recall the constants $r_0$, $\delta_0$, $\tau$ and $c_1$ determined in Lemma 5.2, and set

$$(5.26) \qquad \delta_1 := \left[\min\left\{\delta_0, c_1^{-1}\tau^{n+2}\left(1 - \tau^\beta\right)\right\}\right]^2.$$

Note that the inequalities in $(i)$ and $(ii)$ are trivial when $j = 0$. Then, given $j_0 \geqslant 0$, we assume that the inequalities in $(i)$ and $(ii)$ hold for every $j = 0, 1, \ldots, j_0$. By the definition of $\delta_1$ in (5.26), $(i)$ and $(ii)$ for $j = j_0$, we deduce that

$$|(D\mathbf{u})_{z_0,\tau^{j_0}r}| \leqslant K \quad \text{and} \quad \Phi(z_0, \tau^{j_0}r, \mathbf{u}) \leqslant \Phi(z_0, r, \mathbf{u}) \leqslant \delta_1 \leqslant \delta_0,$$

so we can apply Lemma 5.2 with $\tau^{j_0}r$ in place of $r$ to obtain

$$\Phi(z_0, \tau^{j_0}r, \mathbf{u}) \leqslant \tau^{2\beta}\Phi(z_0, \tau^{j_0+1}r, \mathbf{u}) \leqslant \tau^{2\beta(j_0+1)}\Phi(z_0, r, \mathbf{u}),$$

which proves $(i)$ for $j = j_0 + 1$. For what concerns $(ii)$, we have from (5.8) with $\tau^{j_0}r$ in place of $r$ and $(i)$ for $j = j_0$ that

$$|(D\mathbf{u})_{z_0,\tau^{j_0+1}r}| \leqslant |(D\mathbf{u})_{z_0,\tau^{j_0}r}| + c_1\tau^{-n-2}\Phi(z_0, \tau_{j_0}r, \mathbf{u})^{\frac{1}{2}}$$
$$\leqslant K - \tau^{\beta j_0} + c_1\tau^{-n-2}\tau^{\beta j_0}\delta_1^{\frac{1}{2}}$$
$$\leqslant K - \tau^{\beta j_0} + \tau^{\beta j_0}\left(1 - \tau^\beta\right) = K - \tau^{\beta(j_0+1)},$$

which corresponds to $(ii)$ for $j = j_0 + 1$. Therefore, by induction, assertions $(i)$ and $(ii)$ hold for every integer $j \geqslant 1$. We now prove (5.25). Let $\rho \in (0, r]$. There exists $j \geqslant 0$ such that $\tau^{j+1}r < \rho \leqslant \tau^j r$. Using (5.5) and $(i)$ above, we can write

$$\fint\kern-0.9em\fint_{Q_\rho} H^-_{1+|(D\mathbf{u})_{z_0,\rho}|}(|D\mathbf{u} - (D\mathbf{u})_{z_0,\rho}|)\,\mathrm{d}z \leqslant c\fint\kern-0.9em\fint_{Q_\rho} H^-_{1+|(D\mathbf{u})_{z_0,\tau^j r}|}(|D\mathbf{u} - (D\mathbf{u})_{z_0,\tau^j r}|)\,\mathrm{d}z$$
$$\leqslant \frac{c}{\tau^{n+2}}\fint\kern-0.9em\fint_{Q_{\tau^j r}} H^-_{1+|(D\mathbf{u})_{z_0,\tau^j r}|}(|D\mathbf{u} - (D\mathbf{u})_{z_0,\tau_1^j r}|)\,\mathrm{d}z$$
$$\leqslant \frac{c}{\tau^{n+2}}\Phi(z_0, \tau^j r, \mathbf{u})$$
$$\leqslant \frac{c}{\tau^{n+2}}\tau^{2\beta j}\Phi(z_0, r, \mathbf{u})$$
$$\leqslant \bar{c}\left(\frac{\rho}{r}\right)^{2\beta}\Phi(z_0, r, \mathbf{u}),$$

then the proof is concluded. $\qquad\square$

## 6. Proof of Theorem 1.2

This section is entirely devoted to the proof of the main result, the partial regularity result of Theorem 1.2.

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

$$\liminf_{r\to 0^+}\fint\kern-0.9em\fint_{Q_r(z_0)}|\mathbf{V}_{H^-_{Q_r(z_0)}}(D\mathbf{u}) - (\mathbf{V}_{H^-_{Q_r(z_0)}}(D\mathbf{u}))_{z_0,r}|^2\,\mathrm{d}z = 0$$

and

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

In particular, taking into account (2.17), the first limit implies

$$\liminf_{r\to 0}\fint\kern-0.9em\fint_{Q_r} H^-_{1+|(D\mathbf{u})_{z_0,r}|}(|D\mathbf{u} - (D\mathbf{u})_{z_0,r}|)\,\mathrm{d}z = 0,$$

where $H^-_{1+|(D\mathbf{u})_{z_0,r}|}$ is the shifted $N$-function of $H^-_{Q_r}$ with shift $1 + |(D\mathbf{u})_{z_0,r}|$. Let

$$K := 2(2 + M_{z_0}),$$

and, with $\delta_1$ and $r_0$ in Lemma 5.3,

$$\varepsilon := \frac{\delta_1}{4},$$

and

$$\bar{r} := \min\left\{ \left(\frac{\delta_1}{4}\right)^{\frac{2}{\beta_1}}, r_0 \right\}.$$

Then we can find $r \leqslant \bar{r}$ small enough such that $Q_{2r}(z_0) \Subset \Omega_T$ and

$$\iint_{Q_r(z_0)} H^-_{1+|(D\mathbf{u})_{z_0,r}|}(|D\mathbf{u} - (D\mathbf{u})_{z_0,r}|)\, \mathrm{d}z < \varepsilon \qquad \text{and} \qquad |(D\mathbf{u})_{z_0,r}| < M_{z_0} + 1,$$

which implies that

$$\Phi(z_0, r, \mathbf{u}) \leqslant \frac{\delta_1}{2} \quad \text{and} \quad 1 + |(D\mathbf{u})_{z_0,r}| \leqslant \frac{K}{2}.$$

From the absolute continuity of the integrals, there exists $\tilde{r} \in (0, r]$ such that for every $\tilde{z} \in Q_{\tilde{r}}(z_0) \Subset \Omega_T$,

$$\Phi(\tilde{z}, r, \mathbf{u}) \leqslant \delta_1 \quad \text{and} \quad 1 + |(D\mathbf{u})_{\tilde{z},r}| \leqslant K.$$

Therefore, by applying Lemma 5.3 to each $Q_r(\tilde{z}) \Subset \Omega_T$, in place of $Q_r(z_0)$, with $\tilde{z} \in Q_{\tilde{r}}(z_0)$, and recalling (2.17), we deduce that for each $\beta \in (0, \frac{\beta_1}{4})$,

$$\iint_{Q_\rho(\tilde{z})} \frac{|\mathbf{V}_p(D\mathbf{u}) - (\mathbf{V}_p(D\mathbf{u}))_{\tilde{z},\rho}|^2}{\rho^{2\beta}}\, \mathrm{d}z \leqslant \frac{c}{\rho^{2\beta}} \iint_{Q_\rho(\tilde{z})} |\mathbf{V}_{H^-_{Q_\rho(\tilde{z})}}(D\mathbf{u}) - (\mathbf{V}_{H^-_{Q_\rho(\tilde{z})}}(D\mathbf{u}))_{\tilde{z},\rho}|^2\, \mathrm{d}z$$

$$\leqslant \frac{c}{\rho^{2\beta}} \iint_{Q_\rho(\tilde{z})} H^-_{1+|(D\mathbf{u})_{\tilde{z},\rho}|}(|D\mathbf{u} - (D\mathbf{u})_{\tilde{z},\rho}|)\, \mathrm{d}z$$

$$\leqslant \frac{c\delta_1}{r^{2\beta}},$$

for every $\rho \in (0, r]$ and $\tilde{z} \in Q_{\tilde{r}}(z_0)$. This implies $\mathbf{V}_p(D\mathbf{u}) \in C^{0,\beta,\frac{\beta}{2}}(Q_{\tilde{r}}(z_0), \mathbb{R}^{N \times n})$, and concludes the proof. $\square$

## ACKNOWLEDGMENTS.

## REFERENCES

1. S. Baasandorj and S. Byun, Sun-Sig, Regularity for Orlicz phase problems, *Mem. Amer. Math. Soc.* **308** (2025), no. 1556, v+131 pp.

2. S. Baasandorj, S. Byun, Sun-Sig and W. Kim, Self-improving properties of very weak solutions to double phase systems, *Trans. Amer. Math. Soc.* **376** (2023), no. 12, 8733–8768.

3. A. K. Balci, L. Diening and M. Surnachev, New examples on Lavrentiev gap using fractals, *Calc. Var. Partial Differential Equations* **59** (2020), no. 5, Paper No. 180, 34 pp.

4. P. Baroni, M. Colombo and G. Mingione, Harnack inequalities for double phase functionals *Nonlinear Anal.* **121** (2015), 206–222.

5. P. Baroni, M. Colombo and G. Mingione, Regularity for general functionals with double phase, *Calc. Var. Partial Differential Equations* **57** (2018), no. 2, Paper No. 62, 48 pp.

6. V. Bögelein, F. Duzaar and P. Marcellini, Parabolic Systems with $(p, q)$-growth: a variational approach, *Arch. Ration. Mech. Anal.*, **210**, (2013) 219–267.

7. V. Bögelein, F. Duzaar and P. Marcellini, Parabolic equations with $p, q$-growth, *J. Math. Pures Appl.* (9) **100** (2013), no. 4, 535–563.
8. V. Bögelein, F. Duzaar and G. Mingione, The regularity of general parabolic systems with degenerate diffusion, *Mem. Amer. Math. Soc.* **221** (2013), no. 1041, vi+143 pp.
9. S. Byun and J. Oh, Global gradient estimates for non-uniformly elliptic equations, *Calc. Var. Partial Differential Equations* **56** (2017), no. 2, Paper No. 46, 36 pp.
10. S. Byun, J. Ok and K. Song, Hölder regularity for weak solutions to nonlocal double phase problems, *J. Math. Pures Appl.* (9) **168** (2022), 110–142.
11. S. Campanato, Equazioni paraboliche del secondo ordine e spazi $L^{2,\theta}(\Omega, \delta)$, *Ann. Mat. Pura Appl.* (4) **73** (1966), 55–102.
12. S. Carl and P. Winkert, Parabolic double phase obstacle problems, *Nonlinear Anal. Real World Appl.* **80** (2024), Paper No. 104169, 14 pp.
13. P. Celada and J. Ok, Partial regularity for non-autonomous degenerate quasi-convex functionals with general growth, *Nonlinear Anal.* **194** (2020), 111473, 36 pp.
14. A. Charkaoui, A. Ben-Loghfyry and S. Zeng, Shengda, A novel parabolic model driven by double phase flux operator with variable exponents: application to image decomposition and denoising, *Comput. Math. Appl.* **174** (2024), 97–141.
15. I. Chlebicka, P. Gwiazda and A. Zatorska-Goldstein, Parabolic equation in time and space dependent anisotropic Musielak-Orlicz spaces in absence of Lavrentiev's phenomenon, *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **36** (2019), no. 5, 1431–1465.
16. M. Colombo and G. Mingione, Regularity for double phase variational problems, *Arch. Ration. Mech. Anal.* **215** (2015), no. 2, 443–496.
17. M. Colombo and G. Mingione, Bounded minimisers of double phase variational integrals *Arch. Ration. Mech. Anal.* **218** (2015), no. 1, 219–273.
18. M. Colombo and G. Mingione, Calderón-Zygmund estimates and non-uniformly elliptic operators *J. Funct. Anal.* **270** (2016), no. 4, 1416–1478.
19. C. De Filippis, F. De Filippis and M. Piccinini, Bounded minimizers of double phase problems at nearly linear growth, ArXiv:2411.14325 (2024).
20. C. De Filippis and G. Mingione, A borderline case of Calderón-Zygmund estimates for nonuniformly elliptic problems, *St. Petersburg Math. J.* **31** (2020), no. 3, 455–477.
21. C. De Filippis and G. Mingione, Regularity for double phase problems at nearly linear growth, *Arch. Ration. Mech. Anal.* **247** (2023), no. 5, Paper No. 85, 50 pp.
22. C. De Filippis and G. Mingione, Nonuniformly elliptic Schauder theory, *Invent. Math.* **234** (2023), no. 3, 1109–1196.
23. C. De Filippis and G. Mingione, The sharp growth rate in nonuniformly elliptic Schauder theory, *Duke Math. J.* **174** (2025), no. 9, 1775–1848.
24. C. De Filippis and J. Oh, Regularity for multi-phase variational problems, *J. Differential Equations* **267** (2019), no. 3, 1631–1670.
25. C. De Filippis and G. Palatucci, Hölder regularity for nonlocal double phase equations, *J. Differential Equations* **267** (2019), no. 1, 547–586.
26. E. DiBenedetto, *Degenerate parabolic equations*, Universitext. Springer-Verlag, New York, 1993.
27. L. Diening and F. Ettwein, Fractional estimates for non-differentiable elliptic systems with general growth, *Forum Math.* **20** (3) (2008), 523–556.
28. L. Diening, D. Lengeler, B. Stroffolini and A. Verde, Partial regularity for minimizers of quasi-convex functionals with general growth, *SIAM J. Math. Anal.* **44** (2012), no. 5, 3594–3616.
29. L. Diening, P. Kaplický, and S. Schwarzacher, BMO estimates for the $p$-Laplacian, *Nonlinear Anal.* **75** (2) (2012), 637–650.
30. 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.
31. L. Diening, B. Stroffolini and A. Verde, Everywhere regularity of functionals with $\varphi$-growth, *Manuscripta Math.* **129** (4) (2009), 449–481.
32. L. Diening, B. Stroffolini and A. Verde, The $\varphi$–harmonic approximation and the regularity of $\varphi$–harmonic maps, *J. Differ. Equ.* **253** (7) (2012) 1943–1958.
33. F. Duzaar and J. Grotowski, Optimal interior partial regularity for nonlinear elliptic systems: the method of A-harmonic approximation, *Manuscripta Math.* **103** (2000), no. 3, 267–298.
34. F. Duzaar and G. Mingione, Regularity for degenerate elliptic problems via $p$-harmonic approximation, *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **21** (5) (2004) 735–766.
35. F. Duzaar, G. Mingione and K. Steffen, Parabolic systems with polynomial growth and regularity, *Mem. Amer. Math. Soc.* **214** (2011), no. 1005, x+118 pp.
36. L. Esposito, F. Leonetti and G. Mingione, Sharp regularity for functionals with $(p, q)$ growth., *J. Differential Equations* **204** (2004), no. 1, 5–55.
37. I. Fonseca, J. Malý, and G. Mingione, Scalar minimizers with fractal singular sets, *Arch. Ration. Mech. Anal.* **172** (2004), no. 2, 295–307.

38. M. Foss, T. Isernia, C. Leone and A. Verde, $\mathcal{A}$-caloric approximation and partial regularity for parabolic systems with Orlicz growth, *Calc. Var. Partial Differential Equations* **62** (2023), no. 2, Paper No. 51.

39. M. Foss and G. Mingione, Partial continuity for elliptic problems, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **25** (3) (2008) 471–503.

40. P. Harjulehto and P. Hästö, Orlicz Spaces and Generalized Orlicz Spaces, *Lecture Notes in Mathematics*, vol.2236, Springer, Cham, (2019).

41. P. Harjulehto and P. Hästö, Double phase image restoration, *J. Math. Anal. Appl.* **501** (2021), no. 1, Paper No. 123832, 12 pp.

42. P. Hästö, M. Lee and J. Ok, Mean oscillation conditions for nonlinear equation and regularity results, Preprint 2025, arXiv:2504.02159.

43. P. Hästö and J. Ok, Maximal regularity for local minimizers of non-autonomous functionals, *J. Eur. Math. Soc. (JEMS)* **24** (2022), no. 4, 1285–1334.

44. P. Hästö and J. Ok, Regularity theory for non-autonomous problems with a priori assumptions, *Calc. Var. Partial Differential Equations* **62** (2023), no. 9, Paper No. 251, 28 pp.

45. P. Hästö and J. Ok, Regularity theory for non-autonomous partial differential equations without Uhlenbeck structure, *Arch. Rational. Mech. Anal.* **245** (2022), 1401–1436.

46. P. Hästö and J. Ok, Higher integrability for parabolic PDEs with generalized Orlicz growth, Preprint.

47. W. Kim, Calderón-Zygmund type estimate for the parabolic double-phase system, arXiv:2304.10615. (to appear in *Ann. Sc. Norm. Super. Pisa*).

48. W. Kim, Calderón-Zygmund type estimate for the singular parabolic double-phase system, *J. Math. Anal. Appl.* **551** (2025), no. 1, Paper No. 129593, 33 pp.

49. W. Kim, J. Kinnunen and K. Moring, Gradient higher integrability for degenerate parabolic double-phase systems, *Arch. Ration. Mech. Anal.* **247** (2023), no. 5, art. 79.

50. W. Kim, J. Kinnunen and L. Särkiö, Lipschitz truncation method for parabolic double-phase systems and applications, *J. Funct. Anal.* **288** (2025), no. 3, Paper No. 110738, 60 pp.

51. W. Kim, K. Moring and L. Särkiö, Hölder regularity for degenerate parabolic double-phase equations, *J. Differential Equations* **434** (2025), Paper No. 113231, 34 pp.

52. W. Kim and L. Särkiö, Gradient higher integrability for singular parabolic double-phase systems, *NoDEA Nonlinear Differential Equations Appl.* **31** (2024), no. 3, Paper No. 40, 38 pp.

53. M. A. Krasnosel'skiĭ and Ya. B. Rutickiĭ, *Convex Functions and Orlicz Spaces*, P. Noordhoff LTD. - Groningen - The Netherlands, (1961).

54. A. Kufner, O. John and S. Fucĭk, *Function Spaces*, Springer Netherlands (1977).

55. G. Mingione, Regularity of minima: an invitation to the dark side of the calculus of variations. *Appl. Math.* **51** (2006), no. 4, 355–426.

56. J. Ok, Regularity of $\omega$-minimizers for a class of functionals with non-standard growth, *Calc. Var. Partial Differential Equations* **56** (2017), no. 2, Paper No. 48, 31 pp.

57. J. Ok, Partial Hölder regularity for elliptic systems with non-standard growth, *J. Funct. Anal.* **274** (3) (2018), 723–768.

58. J. Ok, Partial regularity for general systems of double phase type with continuous coefficients. *Nonlinear Anal.* **177** (2018), 673–698.

59. J. Ok, Regularity for double phase problems under additional integrability assumptions, *Nonlinear Anal.* **194** (2020), 111408, 13 pp.

60. J. Ok, G. Scilla and B. Stroffolini, Partial regularity for degenerate parabolic systems with general growth via caloric approximations, *Calc. Var. Partial Differential Equations* **63** (2024), no. 4, Paper No. 105, 49 pp.

61. J. Ok, G. Scilla and B. Stroffolini, Partial regularity for degenerate systems of double phase type, *J. Differential Equations* **432** (2025), 113207.

62. G. Scilla and B. Stroffolini, Partial regularity for steady double phase fluids, *Math. Eng.* **5**(5) (2023), 1–47.

63. T. Singer, Existence of weak solutions of parabolic systems with $p, q$-growth. *Manuscr. Math.* 151(1–2) (2016), 87–112.

64. V. V. Zhikov, Averaging of functionals of the calculus of variations and elasticity theory, *Izv. Akad. Nauk SSSR Ser. Mat.* **50** (1986), no. 4, 675–710, 877.

65. V. V. Zhikov, On Lavrentiev's phenomenon, *Russian J. Math. Phys.* **3** (1995), no. 2, 249–269.

66. V. V. Zhikov, On some variational problems, *Russian J. Math. Phys.* **5** (1997), no. 1, 105–116 (1998).

67. V. V. Zhikov, S. M. Kozlov and O. A. Oleinik, Homogenization of differential operators and integral functionals, *Springer-Verlag*, Berlin, 1994. xii+570 pp. Scuol

(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: `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*, Bianca Stroffolini: `bstroffo@unina.it`