Partial regularity for weak solutions of elliptic and parabolic systems with respect to regularity of the coefficients

Taku Kanazawa
Graduate School of Mathematics
Nagoya University, JAPAN

Jan. 19 2015

Contents

1 Introduction

In this paper, we consider nonlinear second order elliptic systems in divergence form of the following type:

$$-\mathrm{div}A(x,u,Du) = f(x,u,Du) \qquad \text{in } \Omega \tag{1.1}$$

or its parabolic version:

$$u_t - \mathrm{div}A(x,t,u,Du) = f(x,t,u,Du) \qquad \text{in } \Omega_T := \Omega \times (-T,0),\ T>0, \tag{1.2}$$

where Ω is a bounded domain in $\mathbb{R}^n$, $n \geq 2$, u takes values in $\mathbb{R}^N$, $N \geq 1$.

Considering about partial differential equations and systems has long history and so that there are many results we have known. For example, harmonic functions, solutions of Laplace's equation (with Dirichlet boundary condition)

$$\begin{cases} \Delta u = \displaystyle\sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} u = 0 & \text{in } \Omega \subset \mathbb{R}^n, \\ u = 0 & \text{on } \partial\Omega \end{cases}$$

has several well-known properties such as maximum principle, Liouville's theorem, mean value equality, etc.

Further, we knew that there are several ways of proving the existence of harmonic functions. Dirichlet integral

$$\mathcal{D}(u) = \frac{1}{2}\int_\Omega |Du|^2 dx, \qquad Du = \left(\frac{\partial u}{\partial x_i}\right)_{i=1,\dots,n}$$

is one of the idea to find harmonic functions. Riemann claimed that the minimum point (minimizer) of $\mathcal{D}$ is harmonic function. In fact, if a minimizer u exists, then the first variation of the Dirichlet integral vanishes:

$$\left.\frac{d}{dt}\mathcal{D}(u+t\varphi)\right|_{t=0} = 0$$

for all smooth compactly supported functions φ in Ω; an integration by parts then yields

$$\begin{aligned} 0 &= \left.\frac{d}{dt}\mathcal{D}(u+t\varphi)\right|_{t=0} \\ &= \int_\Omega \nabla u \cdot \nabla\varphi dx \\ &= -\int_\Omega \Delta u \varphi dx, \qquad \varphi \in C_0^\infty(\Omega), \end{aligned}$$

and arbitrariness of φ we conclude $\Delta u = 0$.

As like above, the problem that considering about minimum points of functionals is called *Calculus of Variations*. More precisely, it is to consider about the existence and differentiability of minimum points for variational integrals of the type

$$\mathcal{F}[u] := \int_\Omega F(x,u,Du)dx \tag{1.3}$$

where $F(x,u,p)\colon \Omega\times\mathbb{R}^N\times\mathbb{R}^{nN}\to\mathbb{R}$ and $u\colon\Omega\to\mathbb{R}^N$. Let $\mathcal{F}$ is of class C^1 and assume that u_0 is a minimum point. Then we have

$$\frac{d}{dt}\mathcal{F}(u_0+t\varphi)\bigg|_{t=0}=0,$$

i.e.,

$$\int_\Omega [F_{p^i_\alpha}(x,u,Du)D_\alpha\varphi^i+F_{u^i}(x,u,Du)\varphi^i]dx=0 \tag{1.4}$$

for all $\varphi\in C_0^\infty(\Omega,\mathbb{R}^N)$. Here and in the following, we use the notation that repeated indices are summed: here α goes from 1 to n and i from 1 to N.

If moreover we assume $\mathcal{F}$, Ω and u is sufficiently smooth then we can integrate by part in (1.4) and get

$$-D_\alpha F_{p^i_\alpha}(x,u,Du)+F_{u^i}(x,u,Du)=0 \quad \text{in } \Omega \qquad i=1,\dots,N \tag{1.5}$$

which is a quasilinear system of partial differential equations. (1.4) and (1.5) are called *Euler-Lagrange equations*.

The idea of using calculus of variations looks well to find harmonic functions but it is not trivial that the minimizer of variational integral $\mathcal{F}$ exists in $C^2(\Omega)$ in general.

This fault was solved by using Hilbert space $W^{1,2}(\Omega)$ instead of $C^2(\Omega)$. Let us present the existence theorem of solution for general elliptic equations (and minimizer for variational integrals) in Hilbert space.

Theorem 1.1 ([19, Theorem 3.12])**.** *Let $A^{\alpha\beta}\in L^\infty(\Omega)$ be elliptic and bounded, that is for some $\lambda,\Lambda>0$*

$$\lambda|\xi|^2\le A^{\alpha\beta}(x)\xi_\alpha\xi_\beta\le\Lambda|\xi|^2, \qquad {}^\forall x\in\Omega. \tag{1.6}$$

Then, for each $g\in W^{1,2}(\Omega)$ and $f_0,f^\alpha\in L^2(\Omega)$, $\alpha=1,\dots,n$, there exists one and only one weak solution $u\in W^{1,2}(\Omega)$ to the Dirichlet problem

$$\begin{cases}-D_\beta(A^{\alpha\beta}D_\alpha u)=f_0-D_\alpha f^\alpha & \text{in } \Omega\\ u=g & \text{on } \partial\Omega\end{cases} \tag{1.7}$$

meaning $u-g\in W_0^{1,2}(\Omega)$ and

$$\int_\Omega A^{\alpha\beta}D_\alpha uD_\beta\varphi dx=\int_\Omega(f_0\varphi+f^\alpha D_\alpha\varphi)dx$$

for all $\varphi\in C_0^\infty(\Omega)$.

If in addition $A^{\alpha\beta}=A^{\beta\alpha}$, then the solution u is the unique minimizer of the functional

$$\mathcal{F}(v)=\frac{1}{2}\int_\Omega A^{\alpha\beta}D_\alpha vD_\beta vdx-\int_\Omega f_0vdx-\int_\Omega f^\alpha D_\alpha vdx$$

in the class

$$\{v\in W^{1,2}(\Omega)\colon v-g\in W_0^{1,2}(\Omega)\}.$$

Note that the case $A^{\alpha\beta} = \delta^{\alpha\beta}$ is just as the case of Dirichlet integral. See also [19, Theorem 3.22] for the vector-valued case $u\colon \Omega \to \mathbb{R}^N$, $N \geq 1$.

The previous theorem ensures the existence of harmonic functions, solutions of (1.7), but it is able to find in $W^{1,2}(\Omega)$ and not in $C^2(\Omega)$. Therefore there is a gap in the regularity scale.

The regularity problem for partial differential equations and systems (and also for calculus of variations) is exactly the problem of filling this gap.

As the mathematicians considered more than a century for regularity problem, there are many different results and that we could not quote all of them. Thus, we only present some important results and we refer to [18, Chapter II] for detailed history about regularity problem.

One of the most important result in regularity problem is the theorem which due by De Giorgi and independently by Nash. Let us consider variational integral defined on $W^{1,2}(\Omega, \mathbb{R}^N)$

$$\mathcal{F}(u) = \int_\Omega F(Du)dx, \tag{1.8}$$

where $|F(p)| \leq L|p|^2$ for some $L > 0$ and the derivatives of $F \in C^\infty(\mathbb{R}^{nN})$, which we write as $A_i^\alpha := D_{p_\alpha^i} F$, satisfy the *growth and ellipticity conditions*

$$\begin{cases} |A_i^\alpha(p)| \leq c|p|, \qquad |D_{p_\beta^j} A_i^\alpha(p)| \leq M \\ D_{p_\beta^j} A_i^\alpha(p)\xi_\alpha^i \xi_\beta^j \geq \lambda|\xi|^2, \quad {}^\forall \xi \in \mathbb{R}^{nN}, \end{cases} \tag{1.9}$$

for some $\lambda, M > 0$.

Then any critical point u of $\mathcal{F}$ satisfies the Euler-Lagrange equation

$$\int_\Omega A_i^\alpha(Du) D_\alpha \varphi^i dx = 0, \qquad {}^\forall \varphi \in W_0^{1,2}(\Omega, \mathbb{R}^N). \tag{1.10}$$

By the method of difference quotient, we differentiate (1.10).

Theorem 1.2 ([19, Proposition 8.6])**.** *Let $u \in W^{1,2}(\Omega, \mathbb{R}^N)$ be a weak solution to the elliptic system* (1.10) *where A_i^α satisfy* (1.9)*. Then $u \in W^{2,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^N)$ and, for $1 \leq s \leq n$, $D_s u$ satisfies the elliptic system*

$$\int_\Omega D_{p_\beta^j} A_i^\alpha(Du) D_\beta(D_s u^j) D_\alpha \varphi^i dx = 0, \qquad {}^\forall \varphi \in W_0^{1,2}(\Omega, \mathbb{R}^N). \tag{1.11}$$

The previous mathematician's work (see for example [19, Corollary 5.16]) and Theorem 1.2 then imply that a critical point of the variational integral (1.8) is smooth as soon as the first derivatives are continuous. Thus, we need to show that the solutions $D_s u$ to the elliptic system (1.11) are continuous. This is false in general but it is true in the scalar case ($N = 1$). Let us assume $N = 1$ and rewrite (1.11) as

$$\int_\Omega A^{\alpha\beta}(x) D_\alpha v D_\beta \varphi dx = 0$$

where $v := D_s u$ and $A^{\alpha\beta}(x) := D_{p_\beta^j} A_i^\alpha(Du(x))$. Now under the assumption (1.9), we can only say that $A^{\alpha\beta} \in L^\infty$ and $A^{\alpha\beta}\xi_\alpha\xi_\beta \geq \lambda\,|\xi|^2$, and in this case we would like to show that v is continuous, or has more regularity, say Hölder continuous. This is exactly the claim of De Giorgi-Nash's theorem.

Theorem 1.3 (De Giorgi-Nash, [18, Theorem 2.1])**.** *Let $u \in W^{1,2}(\Omega)$ be a weak solution to*

$$\int_\Omega A^{\alpha\beta}(x) D_\alpha u D_\beta \varphi dx = 0, \qquad {}^\forall\varphi \in C_0^\infty(\Omega)$$

where $A^{\alpha\beta} \in L^\infty(\Omega)$ satisfies (1.6)*. Then $u \in C^{0,\alpha}_{\mathrm{loc}}(\Omega)$ for some positive α, and, for $\Omega' \Subset \Omega$, $\|u\|_{C^{0,\alpha}(\Omega')} \leq C(\Omega,\Omega') \|u\|_{L^2(\Omega)}$.*

Although regularity problem for elliptic systems in scalar case ($N = 1$) was solved by De Giorgi-Nash's theorem, we could not obtain the generalization of De Giorgi-Nash's theorem for general systems ($N > 1$). Nowadays, we know several counterexamples [18, Chapter II Section 3] to the generalization of De Giorgi-Nash's theorem for vector case. Therefore the goal is to obtain partial regularity result, that is, to prove the existence of a *regular set* $\Omega_u \subset \Omega$ such that

$$\Omega_u := \{x \in \Omega \colon u \text{ is continuous on a neighbourhood of } x\}.$$

Partial regularity result also involves obtaining estimates on the size of *singular set* $\Omega \backslash \Omega_u$ (proving that singular set has zero n-dimensional Lebesgue measure or better, controlling the Hausdorff dimension of the singular set), and higher regularity on the regular set Ω_u.

In the next section, we will present some previous partial regularity results and our main theorems.

2 Main theorems

The aim of this paper is to obtain a partial regularity of weak solutions to the elliptic system (1.1) or the parabolic system (1.2) under p-growth condition, $p \geq 2$, and some suitable conditions to the coefficients $A(x,u,w)$ or $A(z,u,w)$ where $z = (x,t)$. Here we assume the p-growth condition, that is, the estimate

$$|A(x,u,w)| + (1+|w|)\,|D_w A(x,u,w)| \leq L(1+|w|)^{p-1} \tag{2.1}$$

and the ellipticity condition

$$\Big\langle D_w A(x,u,w)\tilde{w}, \tilde{w} \Big\rangle := \sum_{\substack{1\leq i,\beta\leq N\\ 1\leq j,\alpha\leq n}} D_{w^j_\beta} A^i_\alpha(x,u,w)\tilde{w}^\alpha_i \tilde{w}^\beta_j \geq \lambda\,|\tilde{w}|^2\,(1+|w|^2)^{(p-2)/2} \tag{2.2}$$

holds to the elliptic system (1.1) for all $x \in \Omega$, $u \in \mathbb{R}^N$, $w \in \mathbb{R}^{nN}$ or

$$|A(z,u,w)| + (1+|w|)\,|D_w A(z,u,w)| \leq L(1+|w|)^{p-1} \tag{2.3}$$

and

$$\Big\langle D_w A(z,u,w)\tilde{w}, \tilde{w} \Big\rangle := \sum_{\substack{1\leq i,\beta\leq N\\ 1\leq j,\alpha\leq n}} D_{w^j_\beta} A^i_\alpha(z,u,w)\tilde{w}^\alpha_i \tilde{w}^\beta_j \geq \lambda\,|\tilde{w}|^2\,(1+|w|^2)^{(p-2)/2} \tag{2.4}$$

holds to the parabolic system (1.2) for all $z \in \Omega_T$, $u \in \mathbb{R}^N$, $w \in \mathbb{R}^{nN}$.

p-growth condition is originally comes from the growth order of integrand for integral variations. For example, Dirichlet integral,

$$\int_\Omega |Du|^2\, dx,$$

has *quadratic growth* ($p = 2$). Starting from Dirichlet integral, it is natural to consider the variational integral of the form

$$\int_\Omega F(Du) dx$$

where $|F(w)| \le L\,|w|^2$ or more developed version $|F(w)| \le L\,|w|^p$, $p > 1$. The case $1 < p < 2$ is called *subquadratic growth* and the case $p > 2$ is called *superquadratic growth.* If the integrand has p-growth, then by taking Euler-Lagrange equation, it is natural to assume that leading part of the equation has a growth order of $(p-1)$ as like in (1.9).

The regularity result for general nonlinear elliptic systems under p-growth condition (2.1) was first proved by Giaquinta-Modica [20] in case of $p = 2$. They proved that the weak solutions of (1.1) has Hölder continuous first derivatives for some Hölder exponent outside of a singular set of Lebesgue measure zero if

$$|D_w A(x, u, w)| \le L$$

holds for all (x, u, w) and $(1 + |w|)^{-1} A(x, u, w)$ is Hölder continuous in variables (x, u) uniformly with respect to w. Duzaar-Grotowski [12] used a new method, namely *$\mathcal{A}$-harmonic approximation technique* (see Lemma 3.2), and obtained the optimal partial regularity result with more simple proof in [20], i.e., weak solutions to (1.1) belong to $C^{1,\alpha}$ when the coefficients $(1+|w|)^{-1} A(x, u, w)$ are α-Hölder continuous. Note that the Hölder exponent for weak solutions α is same exponent as for coefficients.

The superquadratic version is proved by Chen-Tan [8] and subquadratic case is by Beck [2]. In both cases, we have the optimal result in the sense of quadratic case [12].

On the other hand, the coefficients of systems are considered. The previous results are all considered under Hölder continuous coefficients. The regularity results under more mild assumptions are first proved by Duzaar-Gastel [11]. They assume Dini-type condition instead of Hölder continuity to the coefficients $A(x, u, w)$ and proved C^1-regularity. More precisely, they assume that the continuity of $A(x, u, w)$ with respect to the variables (x, u) that

$$|A(x, u, w) - A(x_0, u_0, w)| \le \kappa(|u|)\eta(|x - x_0| + |u - u_0|)(1 + |w|) \tag{2.5}$$

for all $x, x_0 \in \Omega$, $u, u_0 \in \mathbb{R}^N$, $p \in \mathbb{R}^{nN}$, where $\kappa\colon [0, \infty) \to [1, \infty)$ is nondecreasing, and $\eta\colon (0, \infty) \to [0, \infty)$ is nondecreasing and concave with $\eta(+0) = 0$. They also require that $r \mapsto r^{-\alpha}\eta(r)$ is nonincreasing for some $0 < \alpha < 1$, and the Dini-type condition:

$$\int_0^r \frac{\eta(\rho)}{\rho} d\rho < +\infty \qquad \text{for some } r > 0. \tag{2.6}$$

The subquadratic case, $1 < p < 2$ was proved by Qiu [25].

Our first main theorem is an extension of the result due to Duzaar-Gastel [11] to superquadratic case, $p \geq 2$. Let us assume that the coefficients $A(x,u,w)$ has a modulus of continuity $\eta\colon [0,\infty) \to [0,\infty)$ and nondecreasing function $\kappa\colon [0,\infty) \to [1,\infty)$ such that

$$|A(x,u,w) - A(x_0,u_0,w)| \leq \kappa(|u|)\eta(|x-x_0| + |u-u_0|)(1+|w|)^{p-1} \tag{2.7}$$

for all $x, x_0 \in \Omega$, $u, u_0 \in \mathbb{R}^N$, $p \in \mathbb{R}^{nN}$. Further more we assume that

(η1) η is nondecreasing function with $\eta(0) = 0$.

(η2) η is concave; to prove the regularity theorem we require that $r \mapsto r^{-\alpha}\eta(r)$ is nonincreasing for some exponent $\alpha \in (0,1)$.

We also assume modified Dini condition:

(η3) $F(r) := \displaystyle\int_0^r \frac{\eta^\beta(\rho)}{\rho} d\rho < +\infty$ for some $r > 0$ and $\beta \in (0,1]$.

We assume that the inhomogeneous term f has p-growth, i.e., there exist constants a and b, with a possibly depending on M, such that

$$|f(x,u,w)| \leq a(M)\,|w|^p + b \tag{2.8}$$

for all $x \in \Omega$, $u \in \mathbb{R}^N$ with $|u| \leq M$ and $p \in \mathbb{R}^{nN}$.

Let us define that $u \in W^{1,p}(\Omega,\mathbb{R}^N)$ is a weak solution of (1.1) if u satisfies

$$\int_\Omega \langle A(x,u,Du), D\varphi\rangle dx = \int_\Omega \langle f, \varphi\rangle dx \tag{2.9}$$

for all $\varphi \in C_0^\infty(\Omega,\mathbb{R}^N)$, where $\langle\cdot,\cdot\rangle$ is the standard Euclidean inner product on $\mathbb{R}^N$ or $\mathbb{R}^{nN}$.

Now we are ready to state our first main theorem.

Theorem 2.1 (cf. [23], and Section 4)**.** *Let $u \in W^{1,p}(\Omega,\mathbb{R}^N) \cap L^\infty(\Omega,\mathbb{R}^N)$ be a bounded weak solution of the elliptic system* (1.1) *satisfying* (2.1), (2.2), (2.7), (2.8), (η1), (η2) *and* (η3) *with satisfying $\|u\|_\infty \leq M$ and $2^{(10-9p)/2}\lambda > a(M)M$. Then there exists an open set $\Omega_u \subset \Omega$ such that $u \in C^1(\Omega_u,\mathbb{R}^N)$ with $\mathscr{L}^n(\Omega \setminus \Omega_u) = 0$. Moreover, $\Omega \setminus \Omega_u \subset \Sigma^1 \cup \Sigma^2$ and*

$$\Sigma^1 = \left\{ x_0 \in \Omega\colon \liminf_{\rho\searrow 0} \fint_{B_\rho(x_0)} |Du - (Du)_{x_0,\rho}|^p\, dx > 0 \right\},$$

$$\Sigma^2 = \left\{ x_0 \in \Omega\colon \limsup_{\rho\searrow 0} |(Du)_{x_0,\rho}| = +\infty \right\}.$$

In addition, for $\sigma \in [\alpha,1)$ and $x_0 \in \Omega \setminus \Omega_u$ the derivatives of u has modulus of continuity $r \mapsto r^\sigma + F(r)$ in a neighborhood of x_0.

Note that our result is optimal in the sense that in the case $\eta(r) = r^\alpha$, $0 < \alpha < 1$, we have $F(r) = \alpha^{-1} r^\alpha$ and $C^{1,\alpha}$-regularity is known to be optimal in that case.

As we knew already, if the coefficients $A(x, u, p)$ is just continuous with respect to (x, u) then we could not expect the continuity (and not even boundedness) of the gradients Du. But Foss-Mingione [17] proved that we could still expect the Hölder continuity of the weak solution u itself under the superquadratic growth condition, $p \geq 2$. Subquadratic case was also proved by Beck [3].

Bögelein-Duzaar-Habermann-Scheven [4] proved that continuity of the coefficients with respect to (x, u) is not necessary to guarantee the regularity result, and VMO-condition is sufficient to prove Hölder continuity of the weak solutions in case of homogeneous systems ($f \equiv 0$). More precisely they assume that the partial mapping $x \mapsto A(x, u, w)/(1 + |w|)^{p-1}$ has vanishing mean oscillation (VMO), uniformly in (u, w), i.e., the coefficients $A(x, u, w)$ satisfies an estimate

$$|A(x, u, w) - (A(\cdot, u, w))_{x_0,\rho}| \leq V_{x_0}(x, \rho)(1 + |w|)^{p-1}, \quad \text{for all } x \in B_\rho(x_0) \tag{2.10}$$

where $V_{x_0} \colon \mathbb{R}^n \times [0, \rho_0] \to [0, 2L]$ are bounded functions with

$$\lim_{\rho \searrow 0} V(\rho) = 0, \qquad V(\rho) := \sup_{x_0 \in \Omega} \sup_{0 < r \leq \rho} \fint_{B_r(x_0) \cap \Omega} V_{x_0}(x, r) dx. \tag{2.11}$$

They also assume that $u \mapsto A(x, u, w)/(1 + |w|)^{p-1}$ is continuous, i.e., there exists a modulus of continuity $\omega \colon [0, \infty) \to [0, \infty)$ such that an estimate

$$|A(x, u, w) - A(x, u_0, w)| \leq L\omega(|u - u_0|^2)(1 + |w|)^{p-1} \tag{2.12}$$

holds for all $x \in \Omega$, $u, u_0 \in \mathbb{R}^N$, $w \in \mathbb{R}^{nN}$.

Here we extend the result in [4] and gives Hölder continuity of the weak solutions to inhomogeneous systems (1.1) with inhomogeneous term satisfying p-growth condition (2.8).

Theorem 2.2 (cf. [22], and Section 5)**.** *Let $u \in W^{1,p}(\Omega, \mathbb{R}^N) \cap L^\infty(\Omega, \mathbb{R}^N)$ be a bounded weak solution of the elliptic system* (1.1) *satisfying* (2.1)*,* (2.2)*,* (2.12)*,* (2.10)*,* (2.11) *and* (2.8) *with satisfying $\|u\|_\infty \leq M$ and $2^{(10-9p)/2}\lambda > a(M)M$. Then there exists an open set $\Omega_u \subset \Omega$ such that $u \in C^{0,\alpha}(\Omega_u, \mathbb{R}^N)$ with $\mathscr{L}^n(\Omega \setminus \Omega_u) = 0$ for every $\alpha \in (0, 1)$. Moreover, $\Omega \setminus \Omega_u \subset \Sigma^1 \cup \Sigma^2$ and*

$$\Sigma^1 = \left\{ x_0 \in \Omega \colon \liminf_{\rho \searrow 0} \fint_{B_\rho(x_0)} |Du - (Du)_{x_0,\rho}|^p \, dx > 0 \right\},$$

$$\Sigma^2 = \left\{ x_0 \in \Omega \colon \limsup_{\rho \searrow 0} |(Du)_{x_0,\rho}| = +\infty \right\}.$$

Regularity result for the parabolic systems with Hölder continuous coefficients are first proved by Duzaar-Mingione [13] in case of quadratic growth ($p = 2$). They proved an analogous of $\mathcal{A}$-harmonic approximation lemma, so called *$\mathcal{A}$-caloric approximation lemma* (see Lemma 3.3), and obtained the partial Hölder regularity to Du, where Du denotes the gradient with respect to the spacial variables x, i.e., $Du(x, t) \equiv D_x u(x, t)$. Then superquadratic case was proved by

Duzaar-Mingione-Steffen [15] and subquadratic case by Scheven [26]. Similarly as in the elliptic systems, Dini-type condition and continuous coefficients are considered. Baroni [1] proved that the special derivatives Du are continuous outside of singular set under Dini-type condition with quadratic growth ($p = 2$). Conditions under continuous coefficients are considered by Bögelein-Foss-Mingione [7] in case of $p \geq 2$ and by Foss-Geisbauer [16] in case of $1 < p < 2$.

Our last result is an regularity theorem of parabolic systems (1.2) under VMO-condition with superquadratic growth. As like in elliptic systems, let us assume that the partial mapping $z \mapsto A(z, u, w)/(1 + |w|)^{p-1}$ has VMO, uniformly in (u, w), i.e., the coefficients A satisfies the estimate

$$|A(z,u,w) - (A(\cdot,u,w))_{z_0,\rho}| \leq V_{z_0}(z,\rho)(1+|w|)^{p-1}, \quad \text{for all } z \in Q_\rho(z_0) \tag{2.13}$$

where $V_{z_0} \colon \mathbb{R}^{n+1} \times [0, \rho_0] \to [0, 2L]$ are bounded functions with

$$\lim_{\rho \searrow 0} V(\rho) = 0, \qquad V(\rho) := \sup_{z_0 \in \Omega_T} \sup_{0<r\leq\rho} ⨍_{Q_r(z_0)\cap\Omega_T} V_{z_0}(z,r)dz. \tag{2.14}$$

Moreover we assume that $u \mapsto A(z, u, w)/(1+|w|)^{p-1}$ are continuous, i.e., there exists a bounded, concave and non-decreasing function $\omega \colon [0, \infty) \to [0, \infty)$ satisfying

$$|A(z,u,w) - A(z,u_0,w)| \leq L\omega(|u-u_0|^2)(1+|w|)^{p-1} \tag{2.15}$$

for all $z \in \Omega_T$, $u, u_0 \in \mathbb{R}^N$, $w \in \mathbb{R}^{nN}$. The inhomogeneous term f also satisfies p-growth condition, i.e., there exist constants $a, b > 0$, with a possibly depending on M, such that

$$|f(z,u,w)| \leq a(M)\,|w|^p + b \tag{2.16}$$

for all $z \in \Omega_T$, $u \in \mathbb{R}^N$ with $|u| \leq M$ and $w \in \mathbb{R}^{nN}$.

We will prove the following theorem concerning weak solutions of (1.2), i.e., $u \in C^0(-T, 0; L^2(\Omega, \mathbb{R}^N)) \cap L^p(-T, 0; W^{1,p}(\Omega, \mathbb{R}^N))$, $p \geq 2$ satisfying

$$\int_{\Omega_T} \Big(\langle u, \varphi_t \rangle - \langle A(z,u,Du), D\varphi \rangle \Big) dz = \int_{\Omega_T} \langle f, \varphi \rangle dz \tag{2.17}$$

for all $\varphi \in C_0^\infty(\Omega_T, \mathbb{R}^N)$.

Theorem 2.3 (cf. [24], and Section 6)**.** *Let $u \in C_b^0(-T, 0; L^2(\Omega, \mathbb{R}^N)) \cap L^p(-T, 0; W^{1,p}(\Omega, \mathbb{R}^N))$ be a bounded weak solution of the parabolic system satisfying* (1.2) *satisfying* (2.3)*,* (2.4)*,* (2.13)*,* (2.14)*,* (2.15) *and* (2.16) *with $\|u\|_\infty \leq M$ and $2^{(10-9p)/2}\lambda > a(M)M$. Then there exists an open set $\Omega_u \subset \Omega_T$ such that $u \in C^{\alpha,\alpha/2}(\Omega_u, \mathbb{R}^N)$ with $\mathcal{H}^{n+2}_{\mathrm{par}}(\Omega_T \setminus \Omega_u) = 0$ for every $\alpha \in (0, 1)$. Moreover, $\Omega_T \setminus \Omega_u \subset \Sigma^1_{\mathrm{par}} \cup \Sigma^2_{\mathrm{par}}$ and*

$$\Sigma^1_{\mathrm{par}} = \left\{ z_0 \in \Omega_T \colon \liminf_{\rho \searrow 0} ⨍_{Q_\rho(z_0)} |Du - (Du)_{z_0,\rho}|^p \, dz > 0 \right\},$$

$$\Sigma^2_{\mathrm{par}} = \left\{ z_0 \in \Omega_T \colon \limsup_{\rho \searrow 0} |(Du)_{z_0,\rho}| = +\infty \right\}.$$

The previous result means that the weak solution u is Hölder continuous in Ω_u with exponent α with respect to the parabolic metric $d_{\mathrm{par}}(\cdot,\cdot)$ given by following:

$$d_{\mathrm{par}}(z,z_0) := \max\left\{|x-x_0|, \sqrt{|t-t_0|}\right\} \qquad \text{for } z=(x,t), z_0=(x_0,t_0)\in\Omega_T. \tag{2.18}$$

In other word, u is Hölder continuous in Ω_u with exponent α with respect to space variable x and with exponent $\alpha/2$ with respect to time variable t. Moreover, $\mathcal{H}^{n+2}_{\mathrm{par}}$ denotes $(n+2)$-dimensional parabolic Hausdorff measure which is defined by

$$\mathcal{H}^{n+2}_{\mathrm{par}}(X) := \sup_{\delta>0}\mathcal{H}^{n+2,\delta}_{\mathrm{par}}(X) \tag{2.19}$$

where

$$\mathcal{H}^{n+2,\delta}_{\mathrm{par}}(X) := \inf\left\{\sum_{i=1}^{\infty} R_i^{n+2} \colon X\subset\bigcup_{i=0}^{\infty} Q_{z_i,R_i},\ R_i\le\delta\right\}. \tag{2.20}$$

Note that $\mathcal{H}^{n+2}_{\mathrm{par}}$ is equivalent to $(n+1)$-dimensional Lebesgue measure.

We close this section by briefly summarizing the notation used in this paper. As mentioned above, we consider a bounded domain $\Omega\subset\mathbb{R}^n$ and a cylindrical domain $\Omega_T=\Omega\times(-T,0)\subset\mathbb{R}^{n+1}$ where $n\ge 2$ and $T>0$. u maps from Ω to $\mathbb{R}^N$, $N\ge 1$, in the elliptic setting and from Ω to $\mathbb{R}^N$ in the parabolic setting. Du denotes the gradient of u, especially with respect to the special variables x in the parabolic setting, i.e., $Du(x,t)\equiv D_x u(x,t)$. We write $B_\rho(x_0) := \{x\in\mathbb{R}^n\colon |x-x_0|<\rho\}$ and $Q_\rho(z_0) := B_\rho(x_0)\times(t_0-\rho^2,t_0)$ where $z_0=(x_0,t_0)\in\mathbb{R}^{n+1}$. For a given function g, we denote the average of g on ball $B_\rho(x_0)$ by $g_{x_0,\rho} = \fint_{B_\rho(x_0)\cap\Omega} g\,dx = \frac{1}{|B_\rho(x_0)\cap\Omega|}\int_{B_\rho(x_0)\cap\Omega} g\,dx$ or average on cylinder $Q_\rho(z_0)$ by $g_{z_0,\rho} = \fint_{Q_\rho(z_0)\cap\Omega} g\,dz = \frac{1}{|Q_\rho(z_0)\cap\Omega|}\int_{Q_\rho(z_0)\cap\Omega} g\,dz$. We denote c a positive constants, possibly varying from line by line and special occurrences will be denoted by capital letters C, C_1, C_2 or the likes.

3 Preliminaries

In this section we present $\mathcal{A}$-harmonic approximation lemma, $\mathcal{A}$-caloric approximation lemma and other lemmas which we use to prove regularity theorems.

$\mathcal{A}$-harmonic approximation technique (and also $\mathcal{A}$-caloric approximation technique) has its origin in De Giorgi's harmonic approximation lemma [10] and Simon's proof of the regularity theorem of Allard [27]. The idea of this technique is simple. If we prove that the original solution u is "close enough" to a solution v to an elliptic system with constant coefficients like

$$\operatorname{div}[D_w A(x_0,u_{x_0,\rho},Du_{x_0,\rho})Dv]=0 \quad \text{in } B_\rho(x_0),$$

then we knew that v is smooth in the interior of $B_\rho(x_0)$ and it satisfies good a priori estimates by classical regularity theory. Thus we may hope that the good regularity estimates available for v are in some sense inherited by u, and we may conclude the partial regularity of u. This technique allowed us to obtain the regularity result without heavy tools such as L^p-L^2-estimates for the gradient Du (which was known as Gehring's lemma). For further detail about $\mathcal{A}$-harmonic approximation techniques we may refer to the survey paper [14].

Before we introduce the $\mathcal{A}$-harmonic approximation lemma (and $\mathcal{A}$-caloric approximation lemma), let us define the $\mathcal{A}$-harmonic function and $\mathcal{A}$-caloric function.

Definition 3.1. Let $0 < \lambda < L$ be given and let $\mathcal{A}$ be a bilinear form with constant coefficients satisfying

$$\lambda\,|w|^2 \leq \mathcal{A}(w,w), \quad \mathcal{A}(w,\tilde{w}) \leq L\,|w|\,|\tilde{w}| \qquad \text{for all } w,\tilde{w} \in \mathbb{R}^{nN}. \tag{3.1}$$

A function h is called $\mathcal{A}$-harmonic in the ball $B_\rho(x_0)$ if and only if it satisfies

$$\int_{B_\rho(x_0)} \mathcal{A}(Dh, D\varphi)dx = 0 \qquad \text{for all } \varphi \in C_0^\infty(B_\rho(x_0),\mathbb{R}^N),$$

and a function g is called $\mathcal{A}$-caloric in the cylinder $Q_\rho(z_0)$ if and only if it satisfies

$$\int_{Q_\rho(z_0)} \Big(\langle g, \varphi_t\rangle - \mathcal{A}(Dg, D\varphi)\Big)dz = 0 \qquad \text{for all } \varphi \in C_0^\infty(Q_\rho(z_0),\mathbb{R}^N).$$

Lemma 3.2 ($\mathcal{A}$-harmonic approximation lemma, [4, Lemma 2.3])**.** *Given $\varepsilon > 0$, $0 < \lambda < L$ and $p \geq 2$ there exists $\delta = \delta(n,N,\lambda,L,\varepsilon) \leq 1$ with the following property: Whenever $\mathcal{A}$ is a bilinear form on $\mathbb{R}^{nN}$ satisfying* (3.1)*, $\gamma \in (0,1]$, and whenever*

$$w \in W^{1,2}(B_{\rho/2}(x_0),\mathbb{R}^N)$$

is a function satisfying

$$\fint_{B_{\rho/2}(x_0)} \Big(|Dw|^2 + \gamma^{p-2}\,|Dw|^p\Big)dx \leq 1 \tag{3.2}$$

and

$$\left|\int_{B_{\rho/2}(x_0)} \Big(\langle w, \varphi_t\rangle - \mathcal{A}(Dw, D\varphi)\Big)dx\right| \leq \delta \sup_{B_{\rho/2}(x_0)} |D\varphi| \tag{3.3}$$

for every $\varphi \in C_0^\infty(B_{\rho/2}(x_0),\mathbb{R}^N)$ then there exists a function

$$h \in W^{1,2}(B_{\rho/4}(x_0),\mathbb{R}^N)$$

which is $\mathcal{A}$-harmonic on $B_{\rho/4}(x_0)$ such that

$$\fint_{B_{\rho/4}(x_0)} \Big(|Dh|^2 + \gamma^{p-2}\,|Dh|^p\Big)dx \leq C(n,p) \tag{3.4}$$

and

$$\fint_{B_{\rho/4}(x_0)} \left(\left|\frac{w-h}{\rho/4}\right|^2 + \gamma^{p-2}\left|\frac{w-h}{\rho/4}\right|^p\right)dx \leq \varepsilon. \tag{3.5}$$

Lemma 3.3 ($\mathcal{A}$-caloric approximation lemma, [15, Lemma 3.2])**.** *Given $\varepsilon > 0$, $0 < \lambda < L$ and $p \geq 2$ there exists $\delta = \delta(n,N,p,\lambda,L,\varepsilon) \leq 1$ with the following property: Whenever $\mathcal{A}$ is a bilinear form on $\mathbb{R}^{nN}$ satisfying* (3.1)*, $\gamma \in (0,1]$, and whenever*

$$w \in L^p(t_0 - (\rho/2)^2, t_0; W^{1,2}(B_{\rho/2}(x_0),\mathbb{R}^N))$$

is a function satisfying

$$\fint_{Q_{\rho/2}(z_0)} \left(\left|\frac{w}{\rho/2}\right|^2 + \gamma^{p-2} \left|\frac{w}{\rho/2}\right|^p \right) dz + \fint_{Q_{\rho/2}(z_0)} \Big(|Dw|^2 + \gamma^{p-2} |Dw|^p \Big) dz \le 1 \tag{3.6}$$

and

$$\left| \int_{Q_{\rho/2}(z_0)} \Big(\langle w, \varphi_t \rangle - \mathcal{A}(Dw, D\varphi) \Big) dz \right| \le \delta \sup_{Q_{\rho/2}(z_0)} |D\varphi| \tag{3.7}$$

for every $\varphi \in C_0^\infty(Q_{\rho/2}(z_0), \mathbb{R}^N)$ *then there exists a function*

$$h \in L^p(t_0 - (\rho/4)^2, t_0; W^{1,2}(B_{\rho/4}(x_0), \mathbb{R}^N))$$

which is $\mathcal{A}$*-caloric on* $Q_{\rho/4}(z_0)$ *such that*

$$\fint_{Q_{\rho/4}(z_0)} \left(\left|\frac{h}{\rho/4}\right|^2 + \gamma^{p-2} \left|\frac{h}{\rho/4}\right|^p \right) dz + \fint_{Q_{\rho/4}(z_0)} \Big(|Dh|^2 + \gamma^{p-2} |Dh|^p \Big) dz \le \tilde{C}(n,p) \tag{3.8}$$

and

$$\fint_{Q_{\rho/4}(z_0)} \left(\left|\frac{w-h}{\rho/4}\right|^2 + \gamma^{p-2} \left|\frac{w-h}{\rho/4}\right|^p \right) dz \le \varepsilon. \tag{3.9}$$

The next two lemmas features standard estimates for $\mathcal{A}$-harmonic functions and $\mathcal{A}$-caloric functions.

Lemma 3.4 ([12, Theorem 2.3])**.** *Consider* $\mathcal{A}$*,* λ *and* L *as in Lemma 3.2. Then there exists* $C_0 \ge 1$ *depending only on* n*,* N*,* λ *and* L *such that any* $\mathcal{A}$*-harmonic function* h *on* $B_{\rho/2}(x_0)$ *satisfies*

$$\left(\frac{\rho}{2}\right)^2 \sup_{B_{\rho/4}(x_0)} |Dh|^2 + \left(\frac{\rho}{2}\right)^4 \sup_{B_{\rho/4}(x_0)} \left|D^2 h\right|^2 \le C_0 \left(\frac{\rho}{2}\right)^2 \fint_{B_{\rho/2}(x_0)} |Dh|^2 \, dx. \tag{3.10}$$

Lemma 3.5 ([15, Lemma 4.7])**.** *Let* $h \in L^2(t_0 - (\rho/4)^2, t_0; W^{1,2}(B_{\rho/4}(x_0), \mathbb{R}^N))$ *be* $\mathcal{A}$*-caloric function in* $Q_{\rho/4}(z_0)$ *with* $\mathcal{A}$ *satisfying* (3.1)*. Then* h *is smooth in* $B_{\rho/4}(x_0) \times (t_0 - (\rho/4)^2, t_0]$ *and for any* $s \ge 1$ *there exists a constant* $C = C(n, N, L, \lambda, s) \ge 1$ *such that for any affine function* $\ell \colon \mathbb{R}^n \to \mathbb{R}^N$ *there holds*

$$\fint_{Q_{\theta\rho}(z_0)} \left|\frac{h-\ell}{\theta\rho}\right|^s dz \le \tilde{C}_0 \theta^2 \fint_{Q_\rho(z_0)} \left|\frac{h-\ell}{\rho/4}\right|^s dz \quad \textit{for every } 0 < \theta \ge 1/4.$$

Here we state the Poincaré inequality in a convenient form. Its proof can be find in several literatures, for example [19, Proposition 3.10].

Lemma 3.6. *There exists* $C_P \ge 1$ *depending only on* n *such that every* $u \in W^{1,p}(B_\rho(x_0), \mathbb{R}^N)$ *satisfies*

$$\int_{B_\rho(x_0)} |u - u_{x_0,\rho}|^p \, dx \le C_P \rho^p \int_{B_\rho(x_0)} |Du|^p \, dx. \tag{3.11}$$

For given function $u \in L^2(B_\rho(x_0), \mathbb{R}^N)$ we denote by $\ell_{x_0,\rho}$ unique affine function minimizing

$$\ell \mapsto \fint_{B_\rho(x_0)} |u - \ell|^2 \, dx \tag{3.12}$$

among all affine functions. An elementary calculation yield that $\ell_{x_0,\rho}$ takes the form

$$\ell_{x_0,\rho}(x) = \ell_{x_0,\rho}(x_0) + D\ell_{x_0\rho}(x - x_0)$$

where

$$\ell_{x_0,\rho}(x_0) = u_{x_0,\rho} \quad \text{and} \quad D\ell_{x_0,\rho} = \frac{n+2}{\rho^2} \fint_{B_\rho(x_0)} u \otimes (x - x_0) dx.$$

Using the Cauchy-Schwarz inequality we have the following lemma.

Lemma 3.7 ([4, Lemma 2])**.** *Assume $u \in L^2(B_\rho(x_0), \mathbb{R}^N)$, $x_0 \in \mathbb{R}^n$, $\rho > 0$ and $0 < \theta \le 1$. With $\ell_{x_0,\rho}$ and $\ell_{x_0,\theta\rho}$ we denote the affine functions from $\mathbb{R}^n$ to $\mathbb{R}^N$ defined as above for the radii ρ and $\theta\rho$ respectively. Then we have*

$$|D\ell_{x_0,\rho} - D\ell_{x_0,\theta\rho}|^2 \le \frac{n(n+2)}{(\theta\rho)^2} \fint_{B_{\theta\rho}(x_0)} |u - \ell_{x_0,\rho}|^2 \, dx \tag{3.13}$$

and more generally,

$$|D\ell_{x_0,\rho} - D\ell|^2 \le \frac{n(n+2)}{\rho^2} \fint_{B_\rho(x_0)} |u - \ell|^2 \, dx \tag{3.14}$$

for all affine functions $\ell\colon \mathbb{R}^n \to \mathbb{R}^N$.

Lemma 3.7 implies that $\ell_{x_0,\rho}$ has the following quasi-minimizing property for the L^p-norm.

Lemma 3.8 ([4, Section 2])**.** *Consider the minimizer of* (3.12)*, that is, $\ell_{x_0,\rho}$. For any affine functions $\ell\colon \mathbb{R}^n \to \mathbb{R}^N$ and $p \ge 2$ we have*

$$\fint_{B_\rho(x_0)} |u - \ell_{x_0,\rho}|^p \, dx \le c(n,p) \fint_{B_\rho(x_0)} |u - \ell|^p \, dx.$$

Similarly as above, for given $u \in L^2(Q_\rho(z_0), \mathbb{R}^N)$, $z_0 \in \Omega_T$, $\rho > 0$, we denote by $\ell_{z_0,\rho}$ the unique affine function minimizing

$$\ell \mapsto \fint_{Q_\rho(z_0)} |u - \ell|^2 \, dz \tag{3.15}$$

among all affine functions $\ell(z) = \ell(x)$ which are independent of the time variable t. Note that $\ell_{z_0,\rho}$ takes form

$$\ell_{z_0,\rho}(x) = \ell_{z_0,\rho}(x_0) + D\ell_{z_0\rho}(x - x_0)$$

where

$$\ell_{z_0,\rho}(x_0) = u_{z_0,\rho} \quad \text{and} \quad D\ell_{z_0,\rho} = \frac{n+2}{\rho^2} \fint_{Q_\rho(z_0)} u \otimes (x - x_0) dz,$$

and same argument yields the parallel result of Lemma 3.7:

Lemma 3.9 ([15, Lemma 2.1]). *Assume $u \in L^2(Q_\rho(z_0), \mathbb{R}^N)$, $z_0 \in \mathbb{R}^{n+1}$, $\rho > 0$ and $0 < \theta \leq 1$. With $\ell_{z_0,\rho}$ and $\ell_{z_0,\theta\rho}$ we denote the affine functions from $\mathbb{R}^n$ to $\mathbb{R}^N$ defined in* (3.15) *for the radii ρ and $\theta\rho$ respectively. Then we have*

$$|D\ell_{z_0,\rho} - D\ell_{z_0,\theta\rho}|^2 \leq \frac{n(n+2)}{(\theta\rho)^2} \fint_{Q_{\theta\rho}(z_0)} |u - \ell_{z_0,\rho}|^2 \, dz \tag{3.16}$$

and more generally,

$$|D\ell_{z_0,\rho} - D\ell|^2 \leq \frac{n(n+2)}{\rho^2} \fint_{Q_\rho(z_0)} |u - \ell|^2 \, dz \tag{3.17}$$

for all affine functions $\ell(z) = \ell(x)$ which defined on $\mathbb{R}^n$ to $\mathbb{R}^N$.

The elementary calculation yields next two lemmas.

Lemma 3.10 ([22, Lemma 3.7]). *Consider fixed $a, b \geq 0$, $p \geq 1$. Then for any $\varepsilon > 0$, there exists $K = K(p, \varepsilon) \geq 0$ satisfying*

$$(a + b)^p \leq (1 + \varepsilon)a^p + Kb^p.$$

Lemma 3.11 ([20, Lemma 2.1]). *For $\delta > 0$, and for all $a, b \in \mathbb{R}^k$ we have*

$$4^{-(1+2\delta)}(1 + |a|^2 + |b - a|^2)^{\delta/2} \leq \int_0^1 (1 + |sa + (1 - s)b|^2)^{\delta/2} ds. \tag{3.18}$$

4 Elliptic system: Dini-continuous coefficients

Before we start proving the regularity theorem of the elliptic systems (1.1) with Dini continuous coefficients, i.e., Theorem 2.1, let us make a few remarks.

From (2.1), we infer the existence of a modulus of continuity $\mu\colon [0, \infty) \times [0, \infty) \to [0, 1]$ such that $\mu(s, 0) = 0$ for all s, $t \mapsto \mu(s, t)$ is nondecreasing for fixed s, $s \mapsto \mu(s, t)$ is concave and nondecreasing for fixed t, and μ also satisfies

$$\begin{aligned} &|D_w A(x, u, w) - D_w A(x_0, u_0, w_0)| \\ \leq &L\tilde{\mu} \left(|u| + |w| \, , |x - x_0|^2 + |u - u_0|^2 + |w - w_0|^2 \right) (1 + |w| + |w_0|)^{p-2} \end{aligned} \tag{4.1}$$

for all $x, x_0 \in \Omega$, $u, u_0 \in \mathbb{R}^N$, $w, w_0 \in \mathbb{R}^{nN}$ with $|u| + |w| \leq M$.

For technical reasons, we rewrite the modulus of continuity η by

$$\tilde{\eta}(t) := \eta^2 \left(\sqrt{t} \right).$$

Therefore we have

($\tilde{\eta}$1) $\tilde{\eta}$ is continuous, nondecreasing and $\tilde{\eta}(+0) = 0$.

($\tilde{\eta}$2) $\tilde{\eta}$ is concave and $t \mapsto t^{-\alpha}\tilde{\eta}(t)$ is nonincreasing for the same exponent α as in (η2).

($\tilde{\eta}$3) $\tilde{F}(t) := \left[2F(\sqrt{t}) \right]^2 = \left[\int_0^t \frac{\sqrt{\tilde{\eta}^\beta(\tau)}}{\tau} d\tau \right] < +\infty$ for some $t > 0$.

Changing κ by a constant, but keeping $\kappa \geq 1$, we can also assume that

($\tilde{\eta}4$) $\tilde{\eta}(1) = 1$, which implies $t \leq \tilde{\eta}(t) \leq 1$ for all $t \in (0, 1]$.

Fix $\sigma \leq 1/\alpha$. For $t \leq s$, we deduce $t\tilde{\eta}^\sigma(t) \leq s\tilde{\eta}^\sigma(s)$. For $s \leq t$, we use nonincreasing property of $t^{-\alpha}\tilde{\eta}(t)$ and $\tilde{\eta}(s) \leq 1$, we obtain $s\tilde{\eta}^\sigma(t) \leq t$. Combining both cases we have

$$s\tilde{\eta}^\sigma(t) \leq s\tilde{\eta}^\sigma(s) + t \quad \text{for } s \in [0, 1],\ t > 0,\ \sigma \leq \frac{1}{\alpha}.$$

In particular, we have

($\tilde{\eta}5$) $s\tilde{\eta}(t) \leq s\tilde{\eta}(s) + t$ for $s \in [0, 1]$, $t > 0$,

($\tilde{\eta}6$) $s\sqrt{\tilde{\eta}(t)} \leq s\sqrt{\tilde{\eta}(s)} + t$ for $s \in [0, 1]$, $t > 0$.

From ($\tilde{\eta}2$) we infer for $i \in \mathbb{N} \cup \{0\}$, $\theta \in (0, 1/8]$, $t > 0$

$$\int_{\theta^{2(i+1)}t}^{\theta^{2i}t} \frac{\sqrt{\tilde{\eta}^\beta(\tau)}}{\tau} d\tau \geq \sqrt{\frac{\tilde{\eta}^\beta(\theta^{2i}t)}{(\theta^{2i}t)^{\alpha\beta}}} \int_{\theta^{2(i+1)}t}^{\theta^{2i}t} \tau^{(\alpha\beta-2)/2} d\tau = \frac{2}{\alpha\beta}(1 - \theta^{\alpha\beta})\sqrt{\tilde{\eta}^\beta(\theta^{2i}t)},$$

which implies

$$\sum_{i=0}^{k-1} \sqrt{\tilde{\eta}^\beta(\theta^{2i}t)} \leq \frac{\alpha\beta}{2(1 - \theta^{\alpha\beta})} \sqrt{\tilde{F}(t)} \tag{4.2}$$

for $k \in \mathbb{N}$. This yields that

$$\tilde{\eta}(t) \leq \frac{\alpha^2\beta^2}{4(1 - \theta^{\alpha\beta})^2} \tilde{F}(t) \tag{4.3}$$

for all $t \in [0, 1]$. Moreover we have

$$\begin{aligned}
t^{-\alpha}\tilde{F}(t) &= t^{-\alpha}\left[\sqrt{\tilde{F}(\theta t)} + \int_{\theta t}^{t} \sqrt{\tau^{-\alpha}\tilde{\eta}(\tau)}\tau^{(\alpha-2)/2} d\tau\right]^2 \\
&\leq t^{-\alpha}\left[\sqrt{\tilde{F}(\theta t)} + \frac{2}{\alpha}\sqrt{(\theta t)^{-\alpha}\tilde{\eta}(\theta t)}\left\{\sqrt{t^\alpha} - \sqrt{(\theta t)^\alpha}\right\}\right]^2 \\
&\leq \left[\sqrt{t^{-\alpha}\tilde{F}(\theta t)} + \sqrt{(\theta t)^{-\alpha}\tilde{F}(\theta t)}\frac{1 - \theta^{\alpha/2}}{1 - \theta^{\alpha\beta}}\right]^2 \\
&\leq 4(\theta t)^{-\alpha}\tilde{F}(\theta t).
\end{aligned} \tag{4.4}$$

The first step of proving the regularity theorem is to establish a Caccioppoli-type inequality which able to control the derivatives Du by the solution u itself with increasing support.

For $s, t \geq 0$ let

$$\rho_1(s, t) := (1 + t)^{-1}\kappa^{-1}(s + t), \qquad G(s, t) := (1 + t)^2\kappa^{2p}(s + t).$$

Note that $\rho_1 \leq 1$ and $G \geq 1$.

Lemma 4.1. *Consider* $\nu \in \mathbb{R}^{nN}$ *and* $\xi \in \mathbb{R}^N$ *with* $|\xi| \le M$ *fixed. Let* $u \in W^{1,p}(\Omega, \mathbb{R}^N) \cap L^\infty(\Omega, \mathbb{R}^N)$ *be a weak solution of the elliptic system* (1.1) *satisfying* (2.1), (2.2), (2.7), (2.8), (η1), (η2), (η3) and (η4) *with* $\|u\|_\infty \le M$ *and* $2^{(10-9p)/2}\lambda > a(M)M$. *Then for any* $x_0 \in \Omega$ *and* $\rho \le \rho_1(|\xi|, |\nu|)$ *with* $B_\rho(x_0) \Subset \Omega$ *there holds*

$$\begin{aligned}
&\fint_{B_{\rho/2}(x_0)} \left\{ \frac{|Du-\nu|^2}{(1+|\nu|)^2} + \frac{|Du-\nu|^p}{(1+|\nu|)^p} \right\} dx \\
&\le C_1 \left[\fint_{B_\rho(x_0)} \left\{ \frac{|u-\xi-\nu(x-x_0)|^2}{\rho^2(1+|\nu|)^2} + \frac{|u-\xi-\nu(x-x_0)|^p}{\rho^p(1+|\nu|)^p} \right\} dx + G(|\xi|, |\nu|)\tilde{\eta}(\rho^2) + (a|\nu|+b)^2\rho^2 \right]
\end{aligned} \tag{4.5}$$

for some $C_1 = C_1(p, \lambda, L, a(M), M) \ge 1$.

Proof. Assume $x_0 \in \Omega$ and $\rho \le 1$ satisfying $B_\rho(x_0) \Subset \Omega$ and $\rho \le \rho_1(|\xi|, |\nu|)$. We denote $\xi + \nu(x - x_0)$ by $\ell(x)$ and let take a cut-off function $\psi \in C_0^\infty(B_\rho(x_0))$ satisfying $0 \le \psi \le 1$, $|D\psi| \le 4/\rho$ and $\psi \equiv 1$ on $B_{\rho/2}(x_0)$. Then $\varphi := \psi^p(u-\ell)$ is admissible as a test function in (2.9), and obtain

$$\begin{aligned}
&\fint_{B_\rho(x_0)} \psi^p \langle A(x,u,Du), Du-\nu\rangle dx \\
&\qquad = -\fint_{B_\rho(x_0)} \langle A(x,u,Du), p\psi^{p-1} D\psi \otimes (u-\ell)\rangle dx + \fint_{B_\rho(x_0)} \langle f, \varphi\rangle dx,
\end{aligned} \tag{4.6}$$

where $\xi \otimes \zeta := \xi_i \zeta^\alpha$. From equations

$$\begin{aligned}
&-\fint_{B_\rho(x_0)} \psi^p \langle A(x,u,\nu), Du-\nu\rangle dx \\
&\qquad = \fint_{B_\rho(x_0)} \langle A(x,u,\nu), p\psi^{p-1} \otimes (u-\ell)\rangle dx - \fint_{B_\rho(x_0)} \langle A(x,u,\nu), D\varphi\rangle dx,
\end{aligned} \tag{4.7}$$

and

$$\fint_{B_\rho(x_0)} \langle A(x_0,\xi,\nu), D\varphi\rangle dx = 0 \tag{4.8}$$

we have

$$
\begin{aligned}
&\fint_{B_\rho(x_0)} \psi^p \langle A(x,u,Du) - A(x,u,\nu), Du-\nu\rangle dx\\
=&-\fint_{B_\rho(x_0)} \langle A(x,u,Du) - A(x,u,\nu), p\psi^{p-1} D\psi \otimes (u-\ell)\rangle dx\\
&-\fint_{B_\rho(x_0)} \langle A(x,u,\nu) - A(x,\ell,\nu), D\varphi\rangle dx\\
&-\fint_{B_\rho(x_0)} \langle A(x,\ell,\nu) - A(x_0,\xi,\nu), D\varphi\rangle dx\\
&+\fint_{B_\rho(x_0)} \langle f,\varphi\rangle dx\\
=:&\mathrm{I}+\mathrm{II}+\mathrm{III}+\mathrm{IV}. \qquad (4.9)
\end{aligned}
$$

The terms I, II, III and IV are defined above. Using the ellipticity condition (2.2) and Lemma 3.11 to the left-hand side of (4.9), we have

$$
\begin{aligned}
&\fint_{B_\rho(x_0)} \psi^p \langle A(x,u,Du) - A(x,u,\nu), Du-\nu\rangle dx\\
=&\fint_{B_\rho(x_0)} \psi^p \int_0^1 \big\langle D_w A(x,u,sDu+(1-s)\nu)(Du-\nu), Du-\nu\big\rangle ds dx\\
\geq&\fint_{B_\rho(x_0)} \psi^p \lambda\,|Du-\nu|^2 \int_0^1 (1+|sDu+(1-s)\nu|)^{p-2} ds dx\\
\geq&2^{(12-9p)/2}\lambda\fint_{B_\rho(x_0)} \psi^p \big\{(1+|\nu|)^{p-2}\,|Du-\nu|^2 + |Du-\nu|^p\big\} dx. \qquad (4.10)
\end{aligned}
$$

For $\varepsilon > 0$ to be fixed later, using (2.1) and Young's inequality, we obtain

$$
\begin{aligned}
|\mathrm{I}| \leq& \varepsilon\fint_{B_\rho(x_0)} \psi^p \big\{(1+|\nu|)^{p-2}\,|Du-\nu|^2 + |Du-\nu|^p\big\} dx\\
&+ c(p,L,\varepsilon)\fint_{B_\rho(x_0)} \left\{(1+|\nu|)^{p-2}\left|\frac{u-\ell}{\rho}\right|^2 + \left|\frac{u-\ell}{\rho}\right|^p\right\} dx. \qquad (4.11)
\end{aligned}
$$

We use (2.7) and $D\varphi = \psi^p(Du-\nu) + p\psi^{p-1} D\psi \otimes (u-\ell)$, and split the term II as follows:

$$
\begin{aligned}
|\mathrm{II}| \leq& \fint_{B_\rho(x_0)} \kappa(|\xi| + |\nu|\,\rho)\eta(|u-\ell|)(1+|\nu|)^{p-1}\psi^p\,|Du-\nu|\,dx\\
&+\fint_{B_\rho(x_0)} \kappa(|\xi| + |\nu|\,\rho)\eta(|u-\ell|)(1+|\nu|)^{p-1} p\psi^{p-1}\,|D\psi|\,|u-\ell|\,dx\\
=:&\mathrm{II}_1 + \mathrm{II}_2. \qquad (4.12)
\end{aligned}
$$

Using Young's inequality we estimate the term II_1 as

$$
\mathrm{II}_1 \leq \varepsilon\fint_{B_\rho(x_0)} \psi^p (1+|\nu|)^{p-2}\,|Du-\nu|^2\,dx + \frac{1}{\varepsilon}\fint_{B_\rho(x_0)} (1+|\nu|)^p \kappa^2(|\xi|+|\nu|)\tilde{\eta}\left(|u-\ell|^2\right) dx.
$$

Note that our choice $\rho \le \rho_1(|\xi|, |\nu|)$ allow us to apply ($\tilde{\eta}5$), so that we get

$$\begin{aligned}\mathrm{II}_1 \le& \varepsilon \fint_{B_\rho(x_0)} \psi^p (1+|\nu|)^{p-2} |Du - \nu|^2 \, dx + \frac{1}{\varepsilon}\fint_{B_\rho(x_0)} (1+|\nu|)^{p-2} \left|\frac{u-\ell}{\rho}\right|^2 dx \\ &+ \frac{1}{\varepsilon}\fint_{B_\rho(x_0)} (1+|\nu|)^p \kappa^2(|\xi|+|\nu|)\tilde{\eta}\left(\rho^2(1+|\nu|)^2\kappa^2(|\xi|+|\nu|)\right) dx.\end{aligned}$$

Using the definition of $G(\cdot,\cdot)$ and the fact that $\tilde{\eta}(ct) \le c\tilde{\eta}(t)$ for $c \ge 1$, we deduce

$$\begin{aligned}\mathrm{II}_1 \le& \varepsilon \fint_{B_\rho(x_0)} \psi^p (1+|\nu|)^{p-2} |Du - \nu|^2 \, dx + \frac{1}{\varepsilon}\fint_{B_\rho(x_0)} (1+|\nu|)^{p-2} \left|\frac{u-\ell}{\rho}\right|^2 dx \\ &+ \frac{1}{\varepsilon}(1+|\nu|)^p G(|\xi|, |\nu|)\tilde{\eta}(\rho^2).\end{aligned}$$

Similarly we see

$$\mathrm{II}_2 \le c(p,\varepsilon)\fint_{B_\rho(x_0)} (1+|\nu|)^{p-2} \left|\frac{u-\ell}{\rho}\right|^2 dx + c(p,\varepsilon)(1+|\nu|)^p G(|\xi|, |\nu|)\tilde{\eta}(\rho^2).$$

Combining these two estimates and get

$$\begin{aligned}|\mathrm{II}| \le& \varepsilon \fint_{B_\rho(x_0)} \psi^p (1+|\nu|)^{p-2} |Du - \nu|^2 \, dx + c(p,\varepsilon)\fint_{B_\rho(x_0)} (1+|\nu|)^{p-2} \left|\frac{u-\ell}{\rho}\right|^2 dx \\ &+ c(p,\varepsilon)(1+|\nu|)^p G(|\xi|, |\nu|)\tilde{\eta}(\rho^2).\end{aligned} \tag{4.13}$$

In the same way we derive

$$\begin{aligned}|\mathrm{III}| \le& \fint_{B_\rho(x_0)} (1+|\nu|)^{p-1}\kappa(|\xi|+|\nu|)\eta\left((1+|\nu|)\rho\right)\psi^p |Du-\nu| \, dx \\ &+ \fint_{B_\rho(x_0)} (1+|\nu|)^{p-1}\kappa(|\xi|+|\nu|\,\rho)\eta((1+|\nu|)\rho)4p\left|\frac{u-\ell}{\rho}\right| dx \\ \le& \varepsilon \fint_{B\rho(x_0)} \psi^p (1+|\nu|)^{p-2} |Du - \nu|^2 \, dx + \varepsilon \fint_{B_\rho(x_0)} (1+|\nu|)^{p-2} \left|\frac{u-\ell}{\rho}\right|^2 dx \\ &+ c(p,\varepsilon)(1+|\nu|)^p G(|\xi|, |\nu|)\tilde{\eta}(\rho^2).\end{aligned} \tag{4.14}$$

For $\varepsilon' > 0$ to be fixed later, using (2.8), Lemma 3.10 and Young's inequality, we have

$$\begin{aligned}|\mathrm{IV}| \le& \fint_{B_\rho(x_0)} (a|Du|^p + b)\psi^p |u-\ell| \, dx \\ \le& a(1+\varepsilon')\fint_{B_\rho(x_0)} \psi^p |Du-\nu|^p |u-\ell| \, dx + \varepsilon b^2\rho^2 + \frac{1}{\varepsilon}\fint_{B_\rho(x_0)} \left|\frac{u-\ell}{\rho}\right|^2 dx \\ &+ \fint_{B_\rho(x_0)} \left\{aK(p,\varepsilon')\rho|\nu|^{(p+2)/2}\right\}(1+|\nu|)^{(p-2)/2}\left|\frac{u-\ell}{\rho}\right| dx \\ \le& a(1+\varepsilon')(2M+|\nu|\,\rho)\fint_{B_\rho(x_0)} \psi^p |Du-\nu|^p \, dx + \frac{2}{\varepsilon}\fint_{B_\rho(x_0)} (1+|\nu|)^{p-2}\left|\frac{u-\ell}{\rho}\right|^2 dx \\ &+ \varepsilon(1+|\nu|)^p\rho^2\{aK|\nu|+b\}^2.\end{aligned} \tag{4.15}$$

Combining above estimates, from (4.9) to (4.15), and set $\Lambda := 2^{(12-9p)/2}\lambda - 3\varepsilon - a(1+\varepsilon')(2M + |\nu|\rho)$, this gives

$$\begin{aligned}
&\Lambda \fint_{B_\rho(x_0)} \psi^p\{(1+|\nu|)^p |Du-\nu|^2 + |Du-\ell|^p\}dx \\
\leq & c(p,L,\varepsilon)\left[\fint_{B_\rho(x_0)}\left\{(1+|\nu|)^{p-2}\left|\frac{u-\ell}{\rho}\right|^2 + \left|\frac{u-\ell}{\rho}\right|^p\right\}dx + (1+|\nu|)^p G(|\xi|,|\nu|)\tilde{\eta}(\rho^2)\right] \\
&+ \varepsilon(1+|\nu|)^p\{aK|\nu|+b\}^2\rho^2.
\end{aligned}$$

Now choose $\varepsilon = \varepsilon(p,\lambda,a(M),M) > 0$ and $\varepsilon' = \varepsilon'(p,\lambda,a(M),M) > 0$ in a right way, we obtain the claim. □

Lemma 4.2. *Under the same assumptions in Lemma 4.1, take especially $\xi = u_{x_0,\rho}$. Then for any $x_0 \in \Omega$ and $\rho \leq \rho_1(|\xi|,|\nu|)$ satisfy $B_\rho(x_0) \Subset \Omega$, the inequality*

$$\begin{aligned}
\fint_{B_\rho(x_0)} \mathcal{A}(Dv, D\varphi)dx \leq C_2(1+|\nu|)\Big[&\tilde{\mu}^{1/2}(|\xi|+|\nu|, \Phi(x_0,\rho,\nu))\Phi^{1/2}(x_0,\rho,\nu) \\
&+ \Phi(x_0,\rho,\nu) + G(|\xi|,|\nu|)\sqrt{\tilde{\eta}(\rho^2)} + \rho(a|\nu|+b)\Big]\sup_{B_\rho(x_0)}|D\varphi|
\end{aligned} \tag{4.16}$$

holds for all $\varphi \in C_0^\infty(B_\rho(x_0),\mathbb{R}^N)$, where

$$\mathcal{A}(Dv, D\varphi) := \frac{1}{(1+|\nu|)^{p-1}}\langle D_w A(x_0,\xi,\nu)Dv, D\varphi\rangle,$$

$$\Phi(x_0,\rho,\nu) := \fint_{B_\rho(x_0)}\left\{\frac{|Du-\nu|^2}{(1+|\nu|)^2} + \frac{|Du-\nu|^p}{(1+|\nu|)^p}\right\}dx$$

and $C_2 = C_2(n,p,L,a(M)) \geq 1$.

Proof. Assume $x_0 \in \Omega$ and $\rho \leq 1$ which satisfy $B_\rho(x_0) \Subset \Omega$ and $\rho \leq \rho_1(|\xi|,|\nu|)$. Without loss of generality we may assume $\sup_{B_\rho(x_0)}|Du| \leq 1$. Note that this implies $\sup_{B_\rho(x_0)}|u| \leq \rho \leq 1$. Using the fact that $\int_{B_\rho(x_0)} A(x_0,\xi,\nu)D\varphi dx = 0$ for all $\varphi \in C_0^\infty(B_\rho(x_0),\mathbb{R}^N)$ we deduce

$$\begin{aligned}
&(1+|\nu|)^{p-1}\fint_{B_\rho(x_0)} \mathcal{A}(Dv, D\varphi)dx \\
=&\fint_{B_\rho(x_0)}\int_0^1 \langle [D_wA(x_0,\xi,\nu) - D_wA(x_0,\xi,\nu+s(Du-\nu))](Du-\nu), D\varphi\rangle ds dx \\
&+\fint_{B_\rho(x_0)}\langle A(x_0,\xi,Du) - A(x,\ell,Du), D\varphi\rangle dx \\
&+\fint_{B_\rho(x_0)}\langle A(x,\ell,Du) - A(x,u,Du), D\varphi\rangle dx \\
&+\fint_{B_\rho(x_0)}\langle f,\varphi\rangle dx \\
=:&\mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV}
\end{aligned} \tag{4.17}$$

where terms I, II, III, IV are defined above.

We estimate the term I using the modulus of continuity $\tilde{\mu}(\cdot,\cdot)$ from (4.1), Jensen's inequality and Hölder's inequality, and we get

$$\begin{aligned}
|\mathrm{I}| \le& c(p,L) \fint_{B_\rho(x_0)} \int_0^1 \tilde{\mu}(|\xi|+|\nu|\,, |Du-\nu|^2)(1+|\nu|+|Du-\nu|)^{p-2}\,|Du-\nu|\,dsdx \\
\le& c(1+|\nu|)^{p-1} \fint_{B_\rho(x_0)} \tilde{\mu}(|\xi|+|\nu|\,, |Du-D\ell|^2)\left\{\frac{|Du-\nu|}{1+|\nu|}+\frac{|Du-\nu|^{p-1}}{(1+|\nu|)^{p-1}}\right\}dx \\
\le& c(1+|\nu|)^{p-1}\Big[\tilde{\mu}^{1/2}\big(|\xi|+|\nu|\,,(1+|\nu|)^2\Phi(x_0,\rho,\nu)\big)\Phi^{1/2}(x_0,\rho,\nu) \\
&\qquad + \tilde{\mu}^{1/p}\big(|\xi|+|\nu|\,,(1+|\nu|)^2\Phi(x_0,\rho,\nu)\big)\Phi^{1/q}(x_0,\rho,\nu)\Big] \\
\le& c(1+|\nu|)^{p}\Big[\tilde{\mu}^{1/2}(|\xi|+|\nu|\,,\Phi(x_0,\rho,\nu))\Phi^{1/2}(x_0,\rho,\nu)+\Phi(x_0,\rho,\nu)\Big],
\end{aligned} \tag{4.18}$$

where $q>0$ is a dual exponent of $p\ge 2$, i.e., $q=p/(p-1)$. The last inequality follows from the fact that $a^{1/p}b^{1/q}=a^{1/p}b^{1/p}b^{(p-2)/2}\le a^{1/2}b^{1/2}+b$ holds by Young's inequality and the fact that $\tilde{\mu}(s,ct)\le c\tilde{\mu}(s,t)$ for $c\ge 1$ which deduce from the concavity of $t\mapsto\tilde{\mu}(s,t)$.

Similarly, using the modulus of continuity $\tilde{\eta}(\cdot)$ from (2.7), Young's inequality, and we deduce

$$\begin{aligned}
|\mathrm{II}| \le& 2^{p-2}\kappa(|\xi|+|\nu|)(1+|\nu|)^p\sqrt{\tilde{\eta}(\rho^2)} \\
&+ 2^{p-2}\fint_{B_\rho(x_0)} \kappa(|\xi|+|\nu|)\sqrt{\tilde{\eta}(\rho^2(1+|\nu|)^2)}\,|Du-\nu|^{p-1}\,dx \\
\le& 2^{p-1}(1+|\nu|)^p G(|\xi|\,,|\nu|)\sqrt{\tilde{\eta}(\rho^2)}+2^{p-2}(1+|\nu|)^p\Phi(x_0,\rho,\nu).
\end{aligned} \tag{4.19}$$

Here we have used $\tilde{\eta}^{p/2}(\rho^2(1+|\nu|)^2)\le\sqrt{\tilde{\eta}(\rho^2(1+|\nu|)^2)}$ which follows from the nondecreasing property of $t\mapsto\tilde{\eta}(t)$, ($\tilde{\eta}$4) and our assumption $\rho\le\rho_1\le 1$.

We derive, using again the modulus of continuity $\tilde{\eta}(\cdot)$ from (2.7),

$$\begin{aligned}
|\mathrm{III}| \le& c(p)\fint_{B_\rho(x_0)} \kappa(|\xi|+|\nu|)\sqrt{\tilde{\eta}(|u-\ell|^2)}(1+|\nu|)^{p-1}dx \\
&+ c(p)\fint_{B_\rho(x_0)} \kappa(|\xi|+|\nu|)\sqrt{\tilde{\eta}(|u-\ell|^2)}\,|Du-\nu|^{p-1}\,dx \\
=:& \mathrm{III}_1+\mathrm{III}_2.
\end{aligned}$$

Using Hölder's inequality, Jensen's inequality, ($\tilde{\eta}$6) and the Poincaré inequality, we have

$$\begin{aligned}
\mathrm{III}_1 \le& c(p)(1+|\nu|)^{p-1}\kappa(|\xi|+|\nu|)\tilde{\eta}^{1/2}\left(\fint_{B_\rho(x_0)}|u-\ell|^2\,dx\right) \\
\le& c\rho^{-2}(1+|\nu|)^{p-2}\left\{\rho^2(1+|\nu|)^2\kappa^2(|\xi|+|\nu|)\tilde{\eta}^{1/2}\left(\rho^2(1+|\nu|)^2\kappa^2(|\xi|+|\nu|)\right)+\fint_{B_\rho(x_0)}|u-\ell|^2\,dx\right\} \\
\le& c(p)(1+|\nu|)^p G(|\xi|\,,|\nu|)\sqrt{\tilde{\eta}(\rho^2)}+c(p,n)(1+|\nu|)^p\Phi(x_0,\rho,\nu).
\end{aligned}$$

Similarly, we have, using Young's inequality, ($\tilde{\eta}5$) and the Poincaré inequality,

$$\begin{aligned}
\mathrm{III}_2 &\le c(p) \fint_{B_\rho(x_0)} \kappa^p(|\xi| + |\nu|)\tilde{\eta}^{p/2}\left(|u-\ell|^2\right) dx + c(p)\fint_{B_\rho(x_0)} |Du-\nu|^p\, dx \\
&\le c\fint_{B_\rho(x_0)} \left[\rho^{-2}\{\kappa^2(|\xi|+|\nu|)\rho\tilde{\eta}\left(\kappa^2(|\xi|+|\nu|)\rho^2\right) + |u-\ell|^2\}\right]^{p/2} dx + c(1+|\nu|)^p\Phi(x_0,\rho,\nu) \\
&\le c(1+|\nu|)^p G(|\xi|\,,|\nu|)\sqrt{\tilde{\eta}(\rho^2)} + c(n,p)(1+|\nu|)^p\Phi(x_0,\rho,\nu).
\end{aligned}$$

Thus we obtain

$$|\mathrm{III}| \le c(p)(1+|\nu|)^p G(|\xi|\,,|\nu|)\sqrt{\tilde{\eta}(\rho^2)} + c(n,p)(1+|\nu|)^p\Phi(x_0,\rho,\nu). \tag{4.20}$$

Using (2.8) and recall our assumption $\sup_{B_\rho(x_0)}|\varphi| \le \rho$, we have

$$\begin{aligned}
|\mathrm{IV}| &\le \fint_{B_\rho(x_0)} \rho a(|Du-\nu|+|\nu|)^p dx + b\rho \\
&\le 2^{p-1}a(1+|\nu|)^p\Phi(x_0,\rho,\nu) + 2^{p-1}\rho(1+|\nu|)^p(a\,|\nu|+b).
\end{aligned} \tag{4.21}$$

Combining these estimates, from (4.17) to (4.21), we obtain the conclusion. □

For fixed $x_0 \in \Omega$ and $\rho \le 1$, let write $\Phi(\rho) = \Phi(x_0,\rho,(Du)_{x_0,\rho})$ from now on. Now we are ready to establish the excess improvement.

Lemma 4.3. *Assume the same assumptions with Lemma 4.1. Let $\theta \in (0,1/8]$ be arbitrary and impose the following smallness conditions on the excess:*

(i) $\mu^{1/2}(|u_{x_0,\rho}| + |(Du)_{x_0,\rho}|\,,\Phi(\rho)) + \sqrt{\Phi(\rho)} \le \frac{\delta}{2}$ *with the constant* $\delta = \delta(n,N,p,\lambda,L,\theta^{n+p+2})$ *from Lemma 3.2;*

(ii) $(1+|(Du)_{x_0,\rho}|)\gamma(\rho) \le \theta^n\left(2\sqrt{C_0C}\right)^{-1}$,
where C_0 and C are constants from Theorem 3.4 and Lemma 3.2, and
$\gamma(\rho) := C_2\left[\sqrt{\Phi(\rho)} + 2\delta^{-1}\left\{G(|u_{x_0,\rho}|\,,|(Du)_{x_0,\rho}|)\sqrt{\tilde{\eta}(\rho^2)} + \rho(a(1+|(Du)_{x_0,\rho}|)+b)\right\}\right]$,

(iii) $\rho \le \rho_1(|u_{x_0,\rho}|\,,|(Du)_{x_0,\rho}|)$.

Then there holds the excess improvement estimate

$$\Phi(\theta\rho) \le C_3\theta^2\Phi(\rho) + H(|u_{x_0,\rho}|\,,|(Du)_{x_0,\rho}|)\tilde{\eta}(\rho^2), \tag{4.22}$$

with a constant $C_3 = C_3(n,N,p,\lambda,L,a(M),M,\theta) \ge 1$, where

$$H(s,t) := 8\delta^{-2}C_3\{G^2(1+s,1+t) + a(1+t) + b\}.$$

Proof. We consider $B_\rho(x_0) \Subset \Omega$ and set $\xi = u_{x_0,\rho}$, $\nu = (Du)_{x_0,\rho}$, $\ell = \xi + \nu(x-x_0)$. Assume (i), (ii) and (iii) are satisfied. We rescale the solution u as

$$w := \frac{u-\ell}{(1+|\nu|)\gamma}.$$

Applying Lemma 4.2 on $B_\rho(x_0)$ to w and combining the assumption (i), we obtain

$$\begin{aligned}&\fint_{B_\rho(x_0)} \mathcal{A}(Dw, D\varphi)dx\\ \leq &\left[\mu^{1/2}\left(|\xi|+|\nu|\,,\sqrt{\Phi(\rho)}\right)+\sqrt{\Phi(\rho)}+\frac{\delta}{2}\right]\sup_{B_\rho(x_0)}|D\varphi|\\ \leq&\delta\sup_{B_\rho(x_0)}|D\varphi|\end{aligned}$$

for all $\varphi \in C_0^\infty(B_\rho(x_0), \mathbb{R}^N)$. Moreover, we have, note that $\gamma \geq C_2\sqrt{\Phi(\rho)}$ holds from the definition of γ,

$$\begin{aligned}\fint_{B_\rho(x_0)}\{|Dw|^2+\gamma^{p-2}|Dw|^p\}dx=&\fint_{B_\rho(x_0)}\left\{\frac{|Du-\nu|^2}{\gamma^2(1+|\nu|)^2}+\gamma^{p-2}\frac{|Du-\nu|^p}{\gamma^p(1+|\nu|)^p}\right\}dx\\ \leq&\frac{\Phi}{\gamma}\leq\frac{1}{{C_2}^2}\leq 1.\end{aligned}$$

Thus, these two inequalities allow us to apply $\mathcal{A}$-harmonic approximation lemma (Lemma 3.2), to conclude the existence of an $\mathcal{A}$-harmonic function h satisfying

$$\fint_{B_{\rho/2}(x_0)}\left\{\left|\frac{w-h}{\rho/2}\right|^2+\gamma^{p-2}\left|\frac{w-h}{\rho/2}\right|^p\right\}dx\leq\theta^{n+p+2}, \tag{4.23}$$

and

$$\fint_{B_{\rho/2}(x_0)}\{|Dh|+\gamma^{p-2}|Dh|^p\}dx\leq C(n,p), \tag{4.24}$$

where we taken $\varepsilon = \theta^{n+p+2}$. From Theorem 3.4 and (4.24) we have

$$\sup_{B_{\rho/4}(x_0)}\left|D^2h\right|^2\leq 4C_0C\rho^2.$$

From this we infer the following estimate for $s = 2$ respectively for $s = p$,

$$\sup_{B_{\rho/4}(x_0)}\left|D^h\right|^s\leq c(n,N,\lambda,L,p,s)\rho^{-s}.$$

For $\theta \in (0, 1/8]$, Taylor's theorem applied to h at x_0 yields

$$\sup_{x\in B_{2\theta\rho}(x_0)}|h(x)-h(x_0)-Dh(x_0)(x-x_0)|^s\leq c(n,N,\lambda,L,p,s)\theta^{2s}\rho^s.$$

We have then

$$\begin{aligned}&\gamma^{s-2}(2\theta\rho)^{-s}\fint_{B_{2\theta\rho}(x_0)}|w-h(x_0)-Dh(x_0)(x-x_0)|^s\,dx\\ \leq&c(s)\gamma^{s-2}(2\theta\rho)^{-s}\left[\fint_{B_{2\theta\rho}(x_0)}|w-h|^s\,dx+\fint_{B_{2\theta\rho}(x_0)}|h-h(x_0)-Dh(x_0)(x-x_0)|^s\,dx\right]\\ \leq&c(n,N,\lambda,L,p,s)\theta^2.\end{aligned}$$

Set $P_0 = \nu + \gamma(1+|nu|)Dh(x_0)$. Recall that the mean-value of $u - P_0(x-x_0)$ on $B_{2\theta\rho}(x_0)$ is $u_{x_0,2\theta\rho}$, we have

$$(2\theta\rho)^{-s}\fint_{B_{2\theta\rho}(x_0)} |u - u_{x_0,2\theta\rho} - P_0(x-x_0)|^s\,dx \tag{4.25}$$

$$\le c(s)(2\theta\rho)^{-s}\gamma^s(1+|\nu|)^s\fint_{B_{2\theta\rho}(x_0)} |w - h(x_0) - Dh(x_0)(x-x_0)|^s\,dx \tag{4.26}$$

$$\le c(n,N,\lambda,L,p,s)(1+|\nu|)^s\theta^2\gamma^2. \tag{4.27}$$

By assumption (ii), we infer $\sqrt{\Phi(\rho)} \le \theta^n/2$. This yields

$$|(Du)_{x_0,\theta\rho} - \nu| \le \theta^{-n}\fint_{B_\rho(x_0)} |Du - \nu|\,dx \le \theta^{-n}(1+|\nu|)\sqrt{\Phi(\rho)} \le \frac{1}{2}(1+|\nu|).$$

Thus, combining with the estimate $1+|\nu| \le 1 + |(Du)_{x_0,\theta\rho}| + |(Du)_{x_0,\theta\rho} - \nu|$, we obtain

$$1 + |\nu| \le 2(1 + |(Du)_{x_0,\theta\rho}|). \tag{4.28}$$

Then Theorem 3.4, (4.24) and assumption (iii) imply

$$|P_0| \le |\nu| + |\gamma(1+|\nu|)Dh(x_0)| \le |\nu| + \gamma(1+|\nu|)\sqrt{C_0C(n,p)} \le \frac{1}{2} + |\nu|\,. \tag{4.29}$$

Therefore, combining with (4.28), we have

$$1 + |P_0| \le 3(1 + |(Du)_{x_0,\theta\rho}|).$$

Applying the Caccioppoli-type inequality (Lemma 4.1) on $B_{2\theta\rho}(x_0)$ with $\xi = u_{x_0,2\theta\rho}$ and $\nu = P_0$ yields

$$\begin{aligned}\Phi(\theta\rho) \le& 6^p\Phi(x_0,\theta\rho,P_0)\\ \le& 6^pC_1\Bigg[\fint_{B_{2\theta\rho}(x_0)}\left\{\left|\frac{u - u_{x_0,2\theta\rho} - P_0(x-x_0)}{2\theta\rho(1+|P_0|)}\right|^2 + \left|\frac{u - u_{x_0,2\theta\rho} - P_0(x-x_0)}{2\theta\rho(1+|P_0|)}\right|^p\right\}dx\\ &+ G(|u_{x_0,2\theta\rho}|\,,|P_0|)\tilde{\eta}\big((2\theta\rho)^2\big) + (a\,|P_0| + b)^2(2\theta\rho)^2\Bigg].\end{aligned} \tag{4.30}$$

Using Hölder's inequality, the Poincaré inequality and assumption (ii) we have

$$\begin{aligned}
|u_{x_0,2\theta\rho}| &\leq |u_{x_0,\rho}| + \left| \fint_{B_{2\theta\rho}(x_0)} (u - u_{x_0,\rho} - \nu(x-x_0))dx \right| \\
&\leq |u_{x_0,\rho}| + \left(\fint_{B_{2\theta\rho}(x_0)} |u - u_{x_0,\rho} - \nu(x-x_0)|^2 \, dx \right)^{1/2} \\
&\leq |u_{x_0,\rho}| + (2\theta)^{-n/2} \left(\fint_{B_{\rho}(x_0)} |u - u_{x_0,\rho} - \nu(x-x_0)|^2 \, dx \right)^{1/2} \\
&\leq |u_{x_0,\rho}| + \theta^{-n/2} \sqrt{C_P}(1+|\nu|)\sqrt{\Phi(\rho)} \\
&\leq |u_{x_0,\rho}| + \theta^{-n/2} \frac{\sqrt{C_P}}{C_2}(1+|\nu|)\gamma \\
&\leq |u_{x_0,\rho}| + 1. \qquad (4.31)
\end{aligned}$$

Set $H_0(s,t) = G^2(1+s,1+t) + \{a(1+t)+b\}^q$ and using (4.29) we obtain

$$G(|u_{x_0,2\theta\rho}|, |P_0|)\tilde{\eta}((2\theta\rho)^2) + (a|P_0|+b)^2(2\theta\rho)^2 \leq H_0(|\xi|, |\nu|)\tilde{\eta}(\rho^2). \qquad (4.32)$$

Definitions of γ and H_0 imply

$$\begin{aligned}
\gamma^2 &\leq 2{C_2}^2 \left[\Phi(\rho) + 4\delta^{-2} \left\{ G(|\xi|, |\nu|)\sqrt{\tilde{\eta}(\rho^2)} + \rho(a(1+|\nu|)+b) \right\}^2 \right] \\
&\leq 2{C_2}^2 [\Phi(\rho) + 8\delta^{-2} H_0(|\xi|, |\nu|)\tilde{\eta}(\rho^2)]. \qquad (4.33)
\end{aligned}$$

Plugging (4.27), (4.32) and (4.33) into (4.30), we deduce

$$\begin{aligned}
\Phi(\theta\rho) &\leq 6^p C_1 \left[c(n,N,\lambda,L,p)\theta^2\gamma^2 + G(|u_{x_0,2\theta\rho}|, |P_0|)\tilde{\eta}((2\theta\rho)^2) + (a|P_0|+b)^2(2\theta\rho)^2 \right] \\
&\leq 6^p C_1 \left[c\theta^2 {C_2}^2 \left\{ \Phi(\rho) + \delta^{-2} H_0(|\xi|, |\nu|)\tilde{\eta}(\rho^2) \right\} + H_0(|\xi|, |\nu|)\tilde{\eta}(\rho^q) \right] \\
&\leq C_3 \left[\theta^2\Phi(\rho) + 8\delta^{-2} H_0(|\xi|, |\nu|)\tilde{\eta}(\rho^2) \right],
\end{aligned}$$

and this complete the proof. □

For $\sigma \in [\alpha, 1)$ we find $\theta \in (0, 1/8]$ such that $C_3\theta^2 \leq \theta^{2\sigma}/2$. For $T_0 \geq 1$ there exists $\Phi_0 > 0$ such that

$$\mu^{1/2}\left(2T_0, \sqrt{2\Phi_0}\right) + \sqrt{2\Phi_0} \leq \frac{\delta}{2}, \qquad (4.34)$$

$$2C_4(1+2T_0)\sqrt{2\Phi_0} \leq \theta^n, \qquad (4.35)$$

where $C_4 := C_3(1+\sqrt{C_P})$. Note that $\Phi_0 < 1$. Then we choose $0 < \rho_0 \leq 1$ such that

$$C_5\sqrt{\tilde{\eta}(\rho_0)} \leq \Phi_0, \qquad (4.36)$$

$$\frac{(1+2T_0)(1+\sqrt{C_P})}{\theta^{n/2}} \sqrt{\frac{C_5\alpha^2\beta^2\tilde{F}(\rho_0^2)}{4(1-\theta^{\alpha\beta})^2}} \leq \frac{1}{2}T_0, \qquad (4.37)$$

where

$$C_5 = C_5(n, N, \lambda, L, p, a(M), M, \alpha, \sigma, T_0) = \frac{2H(2T_0, 2T_0)}{2\theta^{2\alpha} - \theta^{2\sigma}}.$$

Lemma 4.4. *Assume that for some $T_0 \geq 1$ and $B_\rho(x_0) \Subset \Omega$ we have*

(a) $|u_{x_0,\rho}| + |(Du)_{x_0,\rho}| \leq T_0$,

(b) $\Phi(\rho) \leq \Phi_0$,

(c) $\rho \leq \rho_0$.

Then the smallness conditions (i), (ii) *and* (iii) *are satisfied on $B_{\theta^k\rho}(x_0)$ for $k \in \mathbb{N} \cup \{0\}$ in Lemma 4.3. Moreover, the limit*

$$\Lambda_{x_0} := \lim_{k\to\infty} (Du)_{x_0,\theta^k\rho}$$

exists, and the inequality

$$\fint_{B_r(x_0)} |Du - \Lambda_{x_0}|^s \, dx \leq C_6 \left[\left(\frac{r}{\rho}\right)^{2\sigma} \Phi(\rho) + \tilde{F}(r^2) \right] \tag{4.38}$$

is valid for $0 < r \leq \rho$ with a constant $C_6 = C_6(n, N, \lambda, L, p, a(M), M, \alpha, \beta, \sigma, T_0)$.

Proof. Inductively we shall derive for $k \in \mathbb{N} \cup \{0\}$ the following three assertions:

(I_k) $\Phi(\theta^k\rho) \leq 2\Phi_0$,

(II_k) $\left|u_{x_0,\theta^k\rho}\right| + \left|(Du)_{x_0,\theta^k\rho}\right| \leq 2T_0$,

(III_k) $\theta^k\rho \leq \rho_1(\left|u_{x_0,\theta^k\rho}\right|, \left|(Du)_{x_0,\theta^k\rho}\right|)$.

We first note that (I_k), (II_k) and (4.34) imply the smallness condition (i_k), i.e., (i) with $\theta^k\rho$ instead of ρ. Next we observe that (I_k), (II_k), (4.35) and (4.36) yield

$$\begin{aligned}
&(1 + \left|(Du)_{x_0,\theta^k\rho}\right|) \left(2\sqrt{C_0 C}\right) \gamma(\theta^k\rho) \\
\leq &(1 + \left|(Du)_{x_0,\theta^k\rho}\right|) \left[C_3\sqrt{2\Phi_0} + H(\left|u_{x_0,\theta^k\rho}\right|, \left|(Du)_{x_0,\theta^k\rho}\right|)\sqrt{\tilde{\eta}(\rho_0{}^2)}\right] \\
\leq &(1 + 2T_0) \left[C_3\sqrt{2\Phi_0} + H(2T_0, 2T_0)\sqrt{\tilde{\eta}(\rho_0{}^2)}\right] \\
\leq &(1 + 2T_0) \left[C_3\sqrt{2\Phi_0} + \frac{2\theta^{2\alpha} - \theta^{2\sigma}}{2}\Phi_0\right] \\
\leq &2C_3(1 + 2T_0)\sqrt{2\Phi_0} \\
\leq &1.
\end{aligned}$$

Thus we have (ii_k). Note that $C_2\left(2\sqrt{C_0 C}\right) \leq C_3$ and $\Phi_0 < 1$ are hold from there definitions. Finally (iii_k) is just (III_k).

By assumption (a), (b) and (c), there hold (I_0), (II_0) and (III_0). Now suppose that we have (I_l), (II_l) and (III_l) for $l = 0, 1, \dots, k-1$ with some $k \in \mathbb{N}$. Then we can use Lemma 4.3 with ρ, $\theta\rho$, …, $\theta^{k-1}\rho$, and yield

$$\begin{aligned}\Phi(\theta^k\rho) &\le \left(\frac{1}{2}\theta^{2\sigma}\right)^k \Phi(\rho) + \sum_{l=0}^{k-1}\left(\frac{1}{2}\theta^{2\sigma}\right)^l H(\left|u_{x_0,\theta^{k-1-l}\rho}\right|, \left|(Du)_{x_0,\theta^{k-1-l}\rho}\right|)\tilde{\eta}((\theta^{k-1-l}\rho)^2)\\ &\le \left(\frac{1}{2}\theta^{2\sigma}\right)^k \Phi(\rho) + H(2T_0, 2T_0)\sum_{l=0}^{k-1}\left(\frac{1}{2}\theta^{2\sigma}\right)^l \tilde{\eta}((\theta^{k-1-l}\rho)^2).\end{aligned}$$

The nondecreasing property of $t \mapsto t^{-\alpha}\tilde{\eta}(t)$ and the choice of σ imply

$$\begin{aligned}\sum_{l=0}^{k-1}\left(\frac{1}{2}\theta^{2\sigma}\right)^l \tilde{\eta}((\theta^{k-1-l}\rho)^2) &\le \theta^{-2\alpha}\tilde{\eta}((\theta^k\rho)^2)\sum_{l=0}^{k-1}\left(\frac{1}{2}\theta^{2\alpha-2\sigma}\right)^l\\ &\le \frac{2\tilde{\eta}((\theta^k\rho)^2)}{2\theta^{2\alpha}-\theta^{2\sigma}}.\end{aligned}$$

Therefore we have

$$\Phi(\theta^k\rho) \le \left(\frac{1}{2}\theta^{2\sigma}\right)^k \Phi(\rho) + C_5\tilde{\eta}((\theta^k\rho)^2). \tag{4.39}$$

Keeping in mind of (b), (c) and the choice of ρ, we prove (I_k). We next want to show (II_k). Using the fact that $⨍_{B_\rho(x_0)} \nu(x - x_0)dx = 0$ holds for all $\nu \in \mathbb{R}^{nN}$, Hölder's inequality and the Poincaré inequality, we obtain

$$\begin{aligned}\left|u_{x_0,\theta^k\rho}\right| &\le \left|u_{x_0,\theta^{k-1}\rho}\right| + \left|⨍_{B_{\theta^k\rho}(x_0)}(u - u_{x_0,\theta^{k-1}\rho} - (Du)_{x_0,\theta^{k-1}\rho}(x - x_0))dx\right|\\ &\le \left|u_{x_0,\theta^{k-1}\rho}\right| + \theta^{-n/2}\sqrt{C_P}(1 + \left|(Du)_{x_0,\theta^{k-1}\rho}\right|)\sqrt{\Phi(\theta^{k-1}\rho)}\\ &\le |u_{x_0,\rho}| + \theta^{-n/2}\sqrt{C_P}\sum_{l=0}^{k-1}(1 + \left|(Du)_{x_0,\theta^l\rho}\right|)\sqrt{\Phi(\theta^l\rho)}.\end{aligned}$$

Similarly we see

$$\begin{aligned}\left|(Du)_{x_0,\theta^k\rho}\right| &\le \left|(Du)_{x_0,\theta^{k-1}\rho}\right| + \left|⨍_{B_{\theta^k\rho}(x_0)}(Du - (Du)_{x_0,\theta^{k-1}\rho})dx\right|\\ &\le |(Du)_{x_0,\rho}| + \theta^{-n/2}\sum_{l=0}^{k-1}(1 + \left|(Du)_{x_0,\theta^l\rho}\right|)\sqrt{\Phi(\theta^l\rho)}.\end{aligned}$$

Combining above two estimates and using (4.39) and (4.2) we infer

$$
\begin{aligned}
&\left|u_{x_0,\theta^k\rho}\right| + \left|(Du)_{x_0,\theta^k\rho}\right| \\
\le& |u_{x_0,\rho}| + |(Du)_{x_0,\rho}| + \frac{(1+\sqrt{C_P})(1+2T_0)}{\theta^{n/2}} \sum_{l=0}^{k-1} \sqrt{\Phi(\theta^l\rho)} \\
\le& T_0 + \frac{(1+\sqrt{C_P})(1+2T_0)}{\theta^{n/2}} \sum_{l=0}^{k-1} \left\{ \left(\frac{1}{\sqrt{2}}\theta^\sigma\right)^l \sqrt{\Phi(\rho)} + \sqrt{C_5\tilde{\eta}(\theta^{2l}\rho^2)} \right\} \\
\le& T_0 + \frac{(1+\sqrt{C_P})(1+2T_0)}{\theta^{n/2}} \left\{ \frac{\sqrt{2\Phi(\rho)}}{\sqrt{2}-\theta^\sigma} + \sqrt{\frac{C_5\alpha^2\beta^2\tilde{F}(\rho^2)}{4(1-\theta^{\alpha\beta})^2}} \right\} \\
\le& T_0 + \frac{1}{\sqrt{2}-\theta^\sigma}\frac{\theta^{n/2}}{2} + \frac{1}{2}T_0 \\
\le& 2T_0.
\end{aligned}
$$

This proves (II$_k$). By (c), (II$_k$), ($\tilde{\eta}$4), the definition of H and (4.36), we easily derive

$$
\begin{aligned}
&(1+\left|(Du)_{x_0,\theta^k\rho}\right|)\kappa(\left|u_{x_0,\theta^k\rho}\right|, \left|(Du)_{x_0,\theta^k\rho}\right|)\theta^k\rho \\
\le& H(2T_0,2T_0)\sqrt{\tilde{\eta}(\rho_0)} \\
\le& 1.
\end{aligned}
$$

Thus, we prove (III$_k$).

We next want to prove that $(Du)_{x_0,\theta^k\rho}$ converges to some limit Λ_{x_0} in $\mathbb{R}^{nN}$. Arguing as in the proof of (II$_k$) we deduce for $k > j$

$$
\begin{aligned}
\left|(Du)_{x_0,\theta^k\rho} - (Du)_{x_0,\theta^j\rho}\right| &\le \sum_{l=j+1}^{k} \left|(Du)_{x_0,\theta^l\rho} - (Du)_{x_0,\theta^{l-1}\rho}\right| \\
&\le \sum_{l=j+1}^{k} \theta^{-n/2}(1+\left|(Du)_{x_0,\theta^{l-1}\rho}\right|)\sqrt{\Phi(\theta^{l-1}\rho)} \\
&\le \frac{(1+2T_0)\sqrt{\theta^{2\sigma j}\Phi(\rho}}{\theta^{n/2}(\sqrt{2}-\theta^\sigma)} + \frac{1+2T_0}{\theta^{n/2}}\sqrt{\frac{C_5\alpha^2\beta^2\tilde{F}(\theta^{2j}\theta^2)}{4(1-\theta^{\alpha\beta})^2}}. \qquad (4.40)
\end{aligned}
$$

Taking into account our assumption ($\tilde{\eta}$3) we see that $\{(Du)_{x_0,\theta^k\rho}\}_k$ is a Cauchy sequence in $\mathbb{R}^{nN}$. Therefore the limit

$$\Lambda_{x_0} := \lim_{k\to\infty}(Du)_{x_0,\theta^k\rho}$$

exists and from (4.40) we infer for $j \in \mathbb{N}\cup\{0\}$

$$
\begin{aligned}
\left|(Du)_{x_0,\theta^j\rho} - \Lambda_{x_0}\right| &\le \left|(Du)_{x_0,\theta^k\rho} - (Du)_{x_0,\theta^j\rho}\right| + \left|(Du)_{x_0,\theta^k\rho} - \Lambda_{x_0}\right| \\
&\to C_7\sqrt{\theta^{2\sigma j}\Phi(\rho) + \tilde{F}(\theta^{2j}\rho^2)} \qquad (\text{as } k\to\infty)
\end{aligned}
$$

where

$$C_7 := \frac{\sqrt{2}(1+2T_0)}{\theta^{n/2}}\sqrt{\frac{1}{(\sqrt{2}-\theta^\sigma)^2}+\frac{C_5\alpha^2\beta^2}{4(1-\theta^{\alpha\beta})^2}}.$$

Combining this with (4.39), and recalling the estimate (4.3) we arrive at

$$\begin{aligned}\fint_{B_{\theta^j\rho}(x_0)} |Du-\Lambda_{x_0}|^2\,dx &\le 2(1+2T_0)\Phi(\theta^j\rho)+2\left|(Du)_{x_0,\theta^j\rho}-\Lambda_{x_0}\right|^2\\ &\le C_8\left\{\theta^{2\sigma j}\Phi(\rho)+\tilde{F}(\theta^{2\sigma j}\rho^2)\right\}\end{aligned}$$

with

$$C_8 := 2\left\{1+2T_0+{C_7}^2+\frac{C_5\alpha^2\beta^2(1+2T_0)}{4(1-\theta^{\alpha\beta})^2}\right\}.$$

For $0<r\le\rho$ we find $j\in\mathbb{N}\cup\{0\}$ such that $\theta^{j+1}\rho\le r<\theta^j\rho$. Then using the above estimate with (4.4) imply

$$\begin{aligned}\fint_{B_r(x_0)} |Du-\Lambda_{x_0}|^2\,dx &\le \theta^{-n}\fint_{B_\theta^j\rho(x_0)} |Du-\Lambda_{x_0}|^2\,dx\\ &\le C_8\theta^{-n}\{\theta^{2\sigma j}\Phi(\rho)+\tilde{F}(\theta^{2\sigma}\rho^2)\\ &\le 4C_8\theta^{-n-2\sigma}\left\{\left(\frac{r}{\rho}\right)^{2\sigma}\Phi(\rho)+\tilde{F}(r^2)\right\}.\end{aligned}$$

This proves (4.38) with $C_6 := 4C_8\theta^{-n-2\sigma}$. □

The regularity theorem (Theorem 2.1) is obtained from Lemma 4.4 by using standard arguments.

5 Elliptic system: VMO-coefficients

From p-growth condition (2.1), we may infer the modulus of continuity function $\mu\colon[0,\infty)\to[0,\infty)$ such that μ is bounded, concave, non-decreasing and we have

$$|D_wA(x,u,w)-D_wA(x,u,w_0)|\le L\mu\left(\frac{|w-w_0|}{1+|w|+|w_0|}\right)(1+|w|+|w_0|)^{p-2} \tag{5.1}$$

for all $x\in\Omega$, $u\in\mathbb{R}^N$, $w,w_0\in\mathbb{R}^{nN}$. Without loss of generality, we may assume $\mu\le 1$. Note that μ is differ from the one we take in the case of Dini-continuous coefficients, (4.1).

Lemma 5.1. *Let $u\in W^{1,p}(\Omega,\mathbb{R}^N)\cap L^\infty(\Omega,\mathbb{R}^N)$ be a weak solution of the elliptic system* (1.1) *satisfying* (2.1)*,* (2.2)*,* (2.12)*,* (2.10)*,* (2.11) *and* (2.8) *with $\|u\|_\infty\le M$ and $2^{(10-9p)/2}\lambda>a(M)M$. For any $x_0\in\Omega$ and $\rho\le 1$ with $B_\rho(x_0)\Subset\Omega$ and any affine function $\ell\colon\mathbb{R}^n\to\mathbb{R}^N$ with $|\ell(x_0)|\le M$,* *we have the estimate*

$$\fint_{B_{\rho/2}(x_0)} \left\{ \frac{|Du - D\ell|^2}{(1+|D\ell|)^2} + \frac{|Du - D\ell|^p}{(1+|D\ell|)^p} \right\} dx$$
$$\leq C_9 \Bigg[\fint_{B_\rho(x_0)} \left\{ \frac{|u-\ell|^2}{\rho^2(1+|D\ell|)^2} + \frac{|u-\ell|^p}{\rho^p(1+|D\ell|)^p} \right\} dx$$
$$+ \omega\left(\fint_{B_\rho(x_0)} |u - \ell(x_0)|^2 \, dx \right) + V(\rho) + (a^q |D\ell|^q + b^q)\rho^q \Bigg] \tag{5.2}$$

with the constant $C_9 = C_9(p, \lambda, L, a(M), M) \leq 1$.

Proof. Similar as in the proof of Lemma 4.1, we take a standard cut-off function $\psi \in C_0^\infty(\Omega, \mathbb{R}^N)$ and insert the admissible test function $\varphi := \psi^p(u - \ell)$ to (2.9) we obtain

$$\fint_{B_\rho(x_0)} \psi^p \langle A(x,u,Du) - A(x,u,D\ell), Du - D\ell \rangle dx$$
$$= -\fint_{B_\rho(x_0)} \langle A(x,u,Du) - A(x,u,D\ell), p\psi^{p-1} D\psi \otimes (u-\ell) \rangle dx$$
$$- \fint_{B_\rho(x_0)} \langle A(x,u,D\ell) - A(x,\ell(x_0),D\ell), D\varphi \rangle dx$$
$$- \fint_{B_\rho(x_0)} \langle A(x,\ell(x_0),D\ell) - (A(\cdot,\ell(x_0),D\ell))_{x_0,\rho}, D\varphi \rangle dx$$
$$+ \fint_{B_\rho(x_0)} \langle f, \varphi \rangle dx$$
$$=: \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV}. \tag{5.3}$$

Note that the term III is differ from (4.9). The left-hand side of (5.3), the term I and IV are estimated as similar as in Lemma 4.1, and we have

$$\fint_{B_\rho(x_0)} \psi^p \langle A(x,u,Du) - A(x,u,D\ell), Du - D\ell \rangle dx$$
$$\geq 2^{(12-9p)/2} \lambda \fint_{B_\rho(x_0)} \psi^p \left\{ (1+|D\ell|)^{p-2} |Du - D\ell|^2 \, |Du - D\ell|^p \right\} dx \tag{5.4}$$

$$|\mathrm{I}| \leq \varepsilon \fint_{B_\rho(x_0)} \psi^p \left\{ (1+|D\ell|)^{p-2} |Du - D\ell|^2 \, |Du - D\ell|^p \right\} dx$$
$$+ c(p, L, \varepsilon) \fint_{B_\rho(x_0)} \left\{ (1+|D\ell|)^{p-2} \left| \frac{u-\ell}{\rho} \right|^2 + \left| \frac{u-\ell^p}{\rho} \right| \right\} dx \tag{5.5}$$

$$|\mathrm{IV}| \leq a(1+\varepsilon')(2M + |D\ell|\,\rho) \fint_{B_\rho(x_0)} \psi^p \, |Du - D\ell|^p \, dx$$
$$+ c(p, \varepsilon) \fint_{B_\rho(x_0)} \left| \frac{u-\ell}{\rho} \right|^p dx + \varepsilon (1+|D\ell|)^p \rho^q \{ a^q K^q |D\ell|^q + b^q \} \tag{5.6}$$

where $\varepsilon, \varepsilon' > 0$ to be fixed later. In order to estimate the term II, we use (2.12), $D\varphi = \psi^p(Du - D\ell) + p\psi^{p-1} \otimes (u - \ell)$, and Young's inequality, we get

$$
\begin{aligned}
|\mathrm{II}| \leq & \varepsilon \fint_{B_\rho(x_0)} |Du - D\ell|^p \, dx + \varepsilon^{-q/p} \fint_{B_\rho(x_0)} L^q \omega^q (|u - \ell(x_0)|^2)(1 + |D\ell|)^p dx \\
& + \varepsilon \fint_{B_\rho(x_0)} \left| \frac{u - \ell}{\rho} \right|^p dx + \varepsilon^{-q/p} \fint_{B_\rho(x_0)} (4Lp)^q \omega^q (|u - \ell(x_0)|^2)(1 + |D\ell|)^p dx \\
\leq & \varepsilon \fint_{B_\rho(x_0)} \psi^p |Du - D\ell|^p \, dx + \varepsilon \fint_{B_\rho(x_0)} \left| \frac{u - \ell}{\rho} \right|^p dx \\
& + c(p, L, \varepsilon)(1 + |D\ell|)^p \omega \left(\fint_{B_\rho(x_0)} |u - \ell(x_0)|^2 \right) dx, \qquad (5.7)
\end{aligned}
$$

where we use Jensen's inequality in the last step. We next estimate the term III by using VMO-condition (2.10) and again Young's inequality, we have

$$
|\mathrm{III}| \leq \frac{\varepsilon}{2^{p-1}} \fint_{B_\rho(x_0)} \left\{ \psi^p |Du - D\ell| + \frac{4p |u - \ell|}{\rho} \right\}^p dx + \left(\frac{2^{p-1}}{\varepsilon} \right)^{q/p} \fint_{B_\rho(x_0)} {V_{x_0}}^q(x, \rho)(1 + |D\ell|)^p dx.
$$

Then using the fact that ${V_{x_0}}^q = {V_{x_0}}^{q-1} \cdot V_{x_0} \leq (2L)^{q-1} V_{x_0} \leq 2L V_{x_0}$, and (2.11) we infer

$$
|\mathrm{III}| \leq \varepsilon \fint_{B_\rho(x_0)} \psi^p |Du - D\ell|^p \, dx + c(p, \varepsilon) \fint_{B_\rho(x_0)} \left| \frac{u - \ell}{\rho} \right|^p dx + c(p, L, \varepsilon)(1 + |D\ell|)^p V(\rho). \qquad (5.8)
$$

Combining (5.3) through (5.8), and set $\Lambda := 2^{(12-9p)/2} \lambda - 3\varepsilon - a(1 + \varepsilon')(2M + |D\ell| \, \rho)$, this gives

$$
\begin{aligned}
& \Lambda \fint_{B_\rho(x_0)} \psi^p \left\{ \frac{|Du - D\ell|^2}{(1 + |D\ell|)^2} + \frac{|Du - D\ell|^p}{(1 + |D\ell|)^p} \right\} dx \\
\leq & c(p, L, \varepsilon) \left[\fint_{B_\rho(x_0)} \left\{ \left| \frac{u - \ell}{\rho(1 + |D\ell|)} \right|^2 + \left| \frac{u - \ell}{\rho(1 + |D\ell|)} \right|^p \right\} dx + \omega \left(\fint_{B_\rho(x_0)} |u - \ell(x_0)|^2 \, dx \right) + V(\rho) \right] \\
& + \varepsilon \{ a^q (1 + K(q, \varepsilon'))^q |D\ell|^q + b^q \} \rho^q.
\end{aligned}
$$

Now choose $\varepsilon = \varepsilon(\lambda, p, a(M), M) > 0$ and $\varepsilon' = \varepsilon'(\lambda, p, a(M), M) > 0$ in a right way, we obtain (5.2). $\square$

Here let us write

$$
\begin{aligned}
v &:= u - \ell = u - \ell(x_0) - D\ell(x - x_0), \\
\mathcal{A}(Dv, D\varphi) &:= \frac{1}{(1+|D\ell|)^{p-1}} \Big\langle \big(D_\xi A(\cdot, \ell(x_0), D\ell)\big)_{x_0,\rho} Dv, D\varphi \Big\rangle, \\
\Phi(x_0, \rho, \ell) &:= \fint_{B_\rho(x_0)} \left\{ \frac{|Du - D\ell|^2}{(1+|D\ell|)^2} + \frac{|Du - D\ell|^p}{(1+|D\ell|)^p} \right\} dx \\
\Psi(x_0, \rho, \ell) &:= \fint_{B_\rho(x_0)} \left\{ \frac{|u-\ell|^2}{\rho^2(1+|D\ell|)^2} + \frac{|u-\ell|^p}{\rho^p(1+|D\ell|)^p} \right\} dx \\
\Psi_*(x_0, \rho, \ell) &:= \Psi(x_0, \rho, \ell) + \omega \left(\fint_{B_\rho(x_0)} |u - \ell(x_0)|^2 \, dx \right) + V(\rho) + (a^q \, |D\ell|^q + b^q)\rho^q
\end{aligned}
$$

for shorten.

Lemma 5.2. *Assume the same assumptions in Lemma 5.1. Then for any $x_0 \in \Omega$ and $\rho \le \rho_0$ with satisfy $B_{2\rho}(x_0) \Subset \Omega$, and any affine function $\ell\colon \mathbb{R}^n \to \mathbb{R}^N$ with $|\ell(x_0)| \le M$, the inequality*

$$
\begin{aligned}
\fint_{B_\rho(x_0)} \mathcal{A}(Dv, D\varphi) dx \le C_{10}(1+|D\ell|) \Big[&\mu^{1/2} \left(\sqrt{\Psi_*(x_0, 2\rho, \ell)} \right) \sqrt{\Psi_*(x_0, 2\rho, \ell)} \\
&+ \Psi_*(x_0, 2\rho, \ell) + \rho(a|D\ell|^p + b) \Big] \sup_{B_\rho(x_0)} |D\varphi| \qquad (5.9)
\end{aligned}
$$

holds for all $\varphi \in C_0^\infty(B_\rho(x_0), \mathbb{R}^N)$ and a constant $C_{10} = C_{10}(n, p, \lambda, L, a(M)) \ge 1$.

Proof. Assume $x_0 \in \Omega$ and $\rho \le 1$ satisfy $B_\rho(x_0) \Subset \Omega$.

$$
\begin{aligned}
&(1+|D\ell|)^{p-1} \fint_{B_\rho(x_0)} \mathcal{A}(Dv, D\varphi) dx \\
&\quad = \fint_{B_\rho(x_0)} \int_0^1 \Big\langle \Big[\big(D_w A(\cdot, \ell(x_0), D\ell)\big)_{x_0,\rho} - \big(D_w A(\cdot, \ell(x_0), D\ell + sDv)\big)_{x_0,\rho} \Big] Dv, D\varphi \Big\rangle ds dx \\
&\qquad + \fint_{B_\rho(x_0)} \Big\langle \big(A(\cdot, \ell(x_0), Du)\big)_{x_0,\rho} - A(x, \ell(x_0), Du), D\varphi \Big\rangle dx \\
&\qquad + \fint_{B_\rho(x_0)} \langle A(x, \ell(x_0), Du) - A(x, u, Du), D\varphi \rangle dx \\
&\qquad + \fint_{B_\rho(x_0)} \langle f, \varphi \rangle dx \\
&\quad =: \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} \qquad (5.10)
\end{aligned}
$$

Using the modulus of continuity $\mu(\cdot)$ from (5.1), Jensen's inequality and Hölder's inequality, we estimate

$$
\begin{aligned}
|\,\mathrm{I}\,| &\le c(p,L)(1+|D\ell|)^{p-1} \fint_{B_\rho(x_0)} \mu\left(\frac{|Du-D\ell|}{1+|D\ell|}\right)\left\{\frac{|Du-D\ell|}{1+|D\ell|}+\frac{|Du-D\ell|^{p-1}}{(1+|D\ell|)^{p-1}}\right\}dx \\
&\le c(1+|D\ell|)^{p-1}\left[\mu^{1/2}\left(\sqrt{\Phi(x_0,\rho,\ell)}\right)\sqrt{\Phi(x_0,\rho,\ell)}+\mu^{1/p}\left(\Phi^{1/2}(x_0,\rho,\ell)\right)\Phi^{1/q}(x_0,\rho,\ell)\right] \\
&\le c(1+|D\ell|)^{p-1}\left[\mu^{1/2}\left(\sqrt{\Phi(x_0,\rho,\ell)}\right)\sqrt{\Phi(x_0,\rho,\ell)}+\Phi(x_0,\rho,\ell)\right]. \qquad (5.11)
\end{aligned}
$$

The last inequality follows from the fact that $a^{1/p}b1/q = a^{1/p}b^{1/p}b^{(p-2)/p} \le a^{1/2}b^{1/2}+b$ holds by Young's inequality.

By using the VMO-conditions (2.10), (2.11), Young's inequality and the bound $V_{x_0}(x,\rho)\le 2L$, the term II can be estimated as

$$
\begin{aligned}
|\mathrm{II}| &\le c(p)(1+|D\ell|)^{p-1}\fint_{B_\rho(x_0)}\left\{V_{x_0}(x,\rho)+V_{x_0}(x,\rho)\frac{|Du-D\ell|^{p-1}}{(1+|D\ell|)^{p-1}}\right\}dx \\
&\le c(1+|D\ell|)^{p-1}[(1+(2L)^{p-1})V(\rho)+\Phi(x_0,\rho,\ell)] \qquad (5.12)
\end{aligned}
$$

Similarly, we estimate the term III by using the continuity condition (2.12), Young's inequality, the bound $\omega\le 1$ and Jensen's inequality. This leads us to

$$
\begin{aligned}
|\mathrm{III}| &\le L\fint_{B_\rho(x_0)}(1+|D\ell|+|Du-D\ell|)^{p-1}\omega(|u-\ell(x_0)|^2)dx \\
&\le c(p,L)(1+|D\ell|)^{p-1}\left[\omega\left(\fint_{B_\rho(x_0)}|u-\ell(x_0)|^2\,dx\right)+\Phi(x_0,\rho,\ell)\right] \qquad (5.13)
\end{aligned}
$$

The term IV is estimated as similar as in Dini-continuous coefficients case, Lemma 4.2, therefore we have

$$
\begin{aligned}
|\mathrm{IV}| &\le \fint_{B_\rho(x_0)}\rho(a\,|Du|^p+b)dx \\
&\le 2^{p-1}a(1+|D\ell|)^p\Phi(x_0,\rho,\ell)+2^{p-1}\rho(1+|D\ell|)^{p-1}(a\,|D\ell|^p+b) \qquad (5.14)
\end{aligned}
$$

Combining estimates (5.10) through (5.14), we arrive at

$$
\begin{aligned}
&\fint_{B_\rho(x_0)}\mathcal{A}(Dv,D\varphi)dx \\
&\quad\le c(p,L,a(M))(1+|D\ell|) \\
&\qquad\times\left[\mu^{1/2}\left(\sqrt{\Phi(x_0,\rho,\ell)}\right)\sqrt{\Phi(x_0,\rho,\ell)}+\Phi(x_0,\rho,\ell)+\Psi_*(x_0,\rho,\ell)+\rho(a\,|D\ell|^p+b)\right] \\
&\quad\le C_{10}(1+|D\ell|)\left[\mu^{1/2}\left(\sqrt{\Psi_*(x_0,2\rho,\ell)}\right)\sqrt{\Psi_*(x_0,2\rho,\ell)}+\Psi_*(x_0,2\rho,\ell)+\rho(a\,|D\ell|^p+b)\right],
\end{aligned}
$$

where we use the Caccioppoli-type inequality (Lemma 5.2), $\Phi(x_0,\rho,\ell)\le C_1\Psi_*(x_0,2\rho,\ell)$ and the concavity of μ to have $\mu(cs)\le c\mu(s)$ for $c\ge 1$ at the last step and this complete the proof. □

From now on, for fixed $x_0 \in \Omega$ and $\rho \leq 1$ let us write $\Phi(\rho) = \Phi(x_0, \rho, \ell_{x_0,\rho})$, $\Psi(\rho) = \Psi(x_0, \rho, \ell_{x_0,\rho})$ and $\Psi_*(\rho) = \Psi_*(x_0, \rho, \ell_{x_0,\rho})$ for shorten. Here $\ell_{x_0,\rho}$ is a minimizer which we introduce in (3.12).

Now we are ready to establish the excess improvement estimate for VMO-coefficients system. Note that the excess improvement estimate (5.15) is established by functional Ψ and not by Φ as like in (4.22), Lemma 4.3, because our purpose is to prove $u \in C^{0,\alpha}$ and not $Du \in C^{0,\alpha}$ in this section.

Lemma 5.3. *Assume the same assumptions in Lemma 5.1. Let $\theta \in (0, 1/4]$ be arbitrary and impose the following smallness conditions on the excess:*

(i) $\mu^{1/2}(\sqrt{\Psi(\rho)}) + \sqrt{\Psi(\rho)} \leq \frac{\delta}{2}$ *with the constant* $\delta = \delta(n, N, p, \lambda, L, \theta^{n+p+2})$ *from Lemma 3.2.*

(ii) $\Psi(\rho) \leq \frac{\theta^{n+2}}{4n(n+2)}$,

(iii) $\gamma(\rho) := [\Psi_*^{q/2}(\rho) + \delta^{-q}\rho^q(a|D\ell_{x_0,\rho}| + b)^q]^{1/q} \leq 1.$

Then there holds the excess improvement estimate

$$\Psi(\theta\rho) \leq C_{11}\theta^2\Psi_*(\rho) \tag{5.15}$$

with a constant $C_{11} = C_{11}(n, N, p, \lambda, L, a(M), M, \theta) \geq 1$.

Proof. We first rescale u and set

$$w := \frac{u - \ell_{x_0,\rho}}{C(1 + |D\ell_{x_0,\rho}|)\gamma}.$$

Similarly as Dini-coefficient case, by Lemma 5.2 and assumption (i), the map w is approximately $\mathcal{A}$-harmonic in the sense that

$$\begin{aligned} \fint_{B_{\rho/2}(x_0)} \mathcal{A}(Dw, D\varphi)dx &\leq \left[\mu^{1/2}\left(\sqrt{\Psi_*(\rho)}\right) + \sqrt{\Psi_*(\rho)} + \frac{\delta}{2}\right] \sup_{B_{\rho/2}(x_0)} |D\varphi| \\ &\leq \delta \sup_{B_{\rho/2}(x_0)} |D\varphi|, \end{aligned}$$

for all $\varphi \in C_0^\infty(B_{\rho/2}(x_0), \mathbb{R}^N)$. Moreover, we have

$$\fint_{B_{\rho/2}(x_0)} \left\{|Dw|^2 + \gamma^{p-2}|Dw|^p\right\} dx \leq \frac{C_9\Psi_*(\rho)}{{C_{10}}^2\gamma^2} \leq \frac{C_9}{{C_{10}}^2} \leq 1,$$

and thus Lemma 3.2 ensures the existence of an $\mathcal{A}$-harmonic function h with the properties

$$\fint_{B_{\rho/2}(x_0)} \left\{\left|\frac{w-h}{\rho/2}\right|^2 + \gamma^{p-2}\left|\frac{w-h}{\rho/2}\right|^p\right\} dx \leq \theta^{n+p+2} \tag{5.16}$$

$$\fint_{B_{\rho/2}(x_0)} \left\{|Dh|^2 + \gamma^{p-2}|Dh|^p\right\} dx \leq C(n,p). \tag{5.17}$$

Using Lemma 3.4 and Taylor's theorem, we have the decay estimate for $s = 2$ as well as for $s = p$, where $\theta \in (0, 1/8]$ can be chosen arbitrarily:

$$\begin{aligned}
&\gamma^{s-2}(\theta\rho)^{-s} \fint_{B_{\theta\rho}(x_0)} |w - h(x_0) - Dh(x_0)(x - x_0)|^s \, dx \\
\leq& 2^{s-1}\gamma^{s-2}(\theta\rho)^{-s} \left[\fint_{B_{\theta\rho}(x_0)} |w - h|^s \, dx + \fint_{B_{\theta\rho}(x_0)} |h - h(x_0) - Dh(x_0)(x - x_0)|^s \, dx \right] \\
\leq& c(s, n, N, p, \lambda, L)\theta^2
\end{aligned}$$

Scaling back to u and using Lemma 3.8 we get

$$\begin{aligned}
&(\theta\rho)^{-s} \fint_{B_{\theta\rho}(x_0)} |u - \ell_{x_0,\theta\rho}|^s \, dx \\
\leq& c(s, n)(\theta\rho)^{-s} \fint_{B_{\theta\rho}(x_0)} |u - \ell_{x_0,\theta\rho} - C_{10}\gamma(1 + |D\ell_{x_0,\rho}|)(h(x_0) - Dh(x_0)(x - x_0))|^s \, dx \\
=& c(s, n, N, p, \lambda, L, a(M))(\theta\rho)^{-s}\gamma^s(1 + |D\ell_{x_0,\rho}|)^s \fint_{B_{\theta\rho}(x_0)} |w - h(x_0) - Dh(x_0)(x - x_0)|^s \, dx \\
\leq& c\gamma^2(1 + |D\ell_{x_0,\rho}|)^s\theta^2 \\
\leq& c(1 + |D\ell_{x_0,\rho}|)^s\theta^2 \left[\Psi_*{}^{q/2}(\rho) + 2^{q/p}\delta^{-q}\Psi_*(\rho) \right]^{2/q} \\
\leq& c(1 + |D\ell_{x_0,\rho}|)^s\theta^2\Psi_*(\rho) \qquad (5.18)
\end{aligned}$$

Here we would like to replace the term $|D\ell_{x_0,\rho}|$ on the right-hand side by $|D\ell_{x_0,\theta\rho}|$. For this, we use Lemma 3.7 and the assumption (ii) in order to estimate

$$\begin{aligned}
|D\ell_{x_0,\rho} - D\ell_{x_0,\theta\rho}| &\leq \frac{n(n+2)}{(\theta\rho)^2} \fint_{B_{\theta\rho}(x_0)} |u - \ell_{x_0,\rho}|^2 \, dx \\
&\leq \frac{n(n+2)}{\theta^{n+2}\rho^2} \fint_{B_\rho(x_0)} |u - \ell_{x_0,\rho}|^2 \, dx \\
&\leq \frac{n(n+2)}{\theta^{n+2}}(1 + |D\ell_{x_0,\rho}|)^2\Psi(\rho) \leq \frac{1}{4}(1 + |D\ell_{x_0,\rho}|)^2.
\end{aligned}$$

This yields

$$1 + |D\ell_{x_0,\rho}| \leq 1 + |D\ell_{x_0,\theta\rho}| + |D\ell_{x_0,\rho} - D\ell_{x_0,\theta\rho}| \leq 1 + |D\ell_{x_0,\theta\rho}| + \frac{1}{2}(1 + |D\ell_{x_0,\rho}|),$$

and after reabsorbing the last term from the right-hand side on the left, we obtain

$$1 + |D\ell_{x_0,\rho}| \leq 2(1 + |D\ell_{x_0,\theta\rho}|).$$

Plugging this into (5.18)

$$(\theta\rho)^{-s} \fint_{B_{\theta\rho}(x_0)} |u - \ell_{x_0,\theta\rho}|^s \, dx \leq c(s, n, N, p, \lambda, L, a(M))(1 + |D\ell_{x_0,\theta\rho}|)^s\Psi_*(\rho)$$

for $s = 2$ and $s = p$. Dividing by $(1 + |D\ell_{x_0,\theta\rho}|)^s$, then adding the corresponding terms for $s = 2$ and $s = p$, we deduce the claim. □

We fix an arbitrarily Hölder exponent $\alpha \in (0,1)$ and define the Campanato-type excess

$$C_\alpha(\rho) = C_\alpha(x_0,\rho) = \rho^{-2\alpha} \fint_{B_\rho(x_0)} |u - u_{x_0,\rho}|^2 \, dx.$$

In the following lemma, we iterate the excess improvement estimate (5.15) from Lemma 5.3 and obtain the boundedness of the two excess functionals, C_α and Ψ.

Lemma 5.4. *Under the same assumption with Lemma 5.1, for every $\alpha \in (0,1)$, there exist constants $\varepsilon_*, \kappa_*, \rho_* > 0$ and $\theta_* \in (0,1/8]$, all depending at most on n, N, p, λ, L, $a(M)$, b, α, $\mu(\cdot)$, ρ_0, $\omega(\cdot)$, $V(\cdot)$ and M, such that the conditions*

$$\Psi(\rho) < \varepsilon_*, \quad \text{and} \quad C_\alpha(x_0,\rho) < \kappa_* \tag{(A_0)}$$

for all $\rho \in (0,\rho_*)$ with $B_\rho(x_0) \Subset \Omega$, imply

$$\Psi(\theta *^k \rho) < \varepsilon_*, \quad \text{and} \quad C_\alpha(x_0,\theta_*^k\rho) < \kappa_* \tag{(A_k)}$$

respectively, for every $k \in \mathbb{N}$.

Proof. We begin by choosing the constants. First, let

$$\theta_* := \min\left\{ \left(\frac{1}{16n(n+2)}\right)^{1/(2-2\alpha)}, \frac{1}{\sqrt{4C_{11}}} \right\} \le \frac{1}{8},$$

with the constant C_{11} determined in Lemma 5.3. In particular, the choice of $\theta_* > 0$ fixes the constant $\delta > 0$ from Lemma 3.2. Next, we fix an $\varepsilon_* > 0$ sufficiently small to ensure

$$\varepsilon_* \le \frac{\theta_*^{n+2}}{16n(n+2)} \quad \text{and} \quad \mu^{1/2}\left(\sqrt{\varepsilon_*}\right) + \sqrt{\varepsilon_*} \le \frac{\delta}{2}.$$

Then, we choose $\kappa_* > 0$ so small that

$$\omega(\kappa_*) < \varepsilon_*.$$

Finally, we fix $\rho_* > 0$ small enough to guarantee

$$\rho_* \le \min\{\rho_0, \kappa_*^{1/(2-2\alpha)}, 1\}, \quad V(\rho_*) < \varepsilon_* \quad \text{and} \quad \left\{ \left(a\sqrt{n(n+2)\kappa_*}\right)^q + b^q \right\} \rho_*^{q\alpha} < \varepsilon_*.$$

Recall that $q = p/(p-1)$. Now we prove the assertion (A_k) by induction. We assume that we have already established (A_k) up to some $k \in \mathbb{N} \cup \{0\}$. We begin with proving the first part of the assertion (A_{k+1}), that is, the one concerning $\Psi(\theta_*^{k+1}\rho)$. First, using Lemma 3.7 with $\ell \equiv u_{x_0,\theta_*^k\rho}$, we obtain

$$\begin{aligned}
\left| D\ell_{x_0,\theta_*^k\rho} \right| &\le \frac{n(n+2)}{(\theta_*^k\rho)^2} \fint_{B_{\theta_*^k\rho}(x_0)} \left| u - u_{x_0,\theta_*^k\rho} \right|^2 dx \\
&= n(n+2)(\theta_*^k\rho)^{2\alpha-2} C_\alpha(\theta_*^k\rho) \\
&\le n(n+2)\rho_*^{2\alpha-2}\kappa_*.
\end{aligned} \tag{5.19}$$

Thus, the assumption (A_k), the choice of κ_* and ρ_* and the above estimate infer

$$\begin{aligned}\Psi_*(\theta_*^k\rho) &\le \Psi(\theta_*^k\rho) + \omega\left(C_\alpha(\theta_*^k\rho)\right) + V(\theta_*^k\rho) + \left(a\left|D\ell_{x_0,\theta_*^k\rho}\right|^q + b\right)(\theta_*^k\rho)^q \\ &\le \varepsilon_* + \omega(\kappa_*) + V(\rho_*) + \left(\left(a\sqrt{n(n+2)\kappa_*}\right)^q + b^q\right)\rho_*^{q\alpha} < 4\varepsilon_*. \end{aligned} \tag{5.20}$$

Now it is easy to check that our choice of ε_* implies that the smallness condition assumptions (i) and (ii) in Lemma 5.3 are satisfied on the level $\theta_*^k\rho$, that is, we have

$$\mu^{1/2}\left(\sqrt{\Psi_*(\theta_*^k\rho)}\right) + \sqrt{\Psi_*(\theta_*^k\rho)} < \mu\left(\sqrt{4\varepsilon_*}\right) + \sqrt{4\varepsilon_*} \le \frac{\delta}{2},$$

and

$$\Psi(\theta_*^k\rho) < \varepsilon_* < \frac{\theta_*^{n+2}}{4n(n+2)}.$$

Furthermore, we have the smallness condition assumption (iii), that is,

$$\gamma(\theta_*^k\rho) = \left[\Psi_*^{q/2}(\theta_*^k\rho) + \delta^{-q}(\theta_*^k\rho)^q\left(a\left|D\ell_{x_0,\theta_*^k\rho}\right| + b\right)^q\right]^{1/q} \le 1. \tag{5.21}$$

To check (5.21), first, note that $\Psi_*(\theta_*^k\rho) < 1$ holds by the estimate (5.20) and the choice of ε_*. This implies

$$\Psi_*^{q/2}(\theta_*^k\rho) \le \Psi_*^{1/2}(\theta_*^k\rho) < \sqrt{4\varepsilon_*} \le \frac{\delta}{4}. \tag{5.22}$$

Next, using (5.19) and $\rho_*^{\alpha-1} \ge 1$, we obtain

$$\begin{aligned}\delta^{-q}(\theta_*^k\rho)^q\left(a\left|D\ell_{x_0,\theta_*^k\rho}\right| + b\right)^q &\le \delta^{-q}\rho_*^q\left(a\sqrt{n(n+2)\kappa_*}\rho_*^{\alpha-1} + b\right)^q \\ &\le \delta^{-q}\rho_*^{q\alpha}\left(a\sqrt{n(n+2)\kappa_*} + b\right)^q \\ &\le \delta^{-q}\rho_*^{q\alpha}2^{q/p}\left\{\left(a\sqrt{n(n+2)\kappa_*}\right)^q + b^q\right\}.\end{aligned}$$

Then the choice of ρ_* and ε_* imply

$$\delta^{-q}(\theta_*^k\rho)^q\left(a\left|D\ell_{x_0,\theta_*^k\rho}\right| + B\right)^q \le \delta^{-q}2^{-4+q/p}\delta^{2-q} \le \frac{\delta}{8}. \tag{5.23}$$

Therefore combining (5.22) and (5.23), we have (5.21). We may thus apply Lemma 5.3 with the radius $\theta_*^k\rho$ instead of ρ, which yields

$$\Psi(\theta_*^{k+1}\rho) \le C_{11}\theta_*^2\Psi_*(\theta^k\rho) < 4C_{11}\theta^2\varepsilon_* \le \varepsilon_*,$$

by the choice of $\theta_* > 0$. We have thus established the first part of the assertion (A_{k+1}) and it remains to prove the second one, that is, the one concerning $C_\alpha(x_0, \theta_*^{k+1}\rho)$. For this aim, we first compute

$$\frac{1}{(\theta_*^k\rho)^2}⨍_{B_{\theta_*^k\rho}(x_0)}\left|u - \ell_{x_0,\theta_*^k}\right|^2 dx \le (1 + \left|D\ell_{x_0,\theta_*^k\rho}\right|)^2\Psi(\theta_*^k\rho) \le 2\varepsilon_* + 2\varepsilon_*\left|D\ell_{x_0,\theta_*^k\rho}\right|^2$$

where we used the assumption (A_k) in the last step. Since $\ell_{x_0,\theta_*^k\rho}(x) = u_{x_0,\theta_*^k} + D\ell_{x_0,\theta_*^k\rho}(x - x_0)$, we can estimate

$$
\begin{aligned}
C_\alpha(\theta_*^{k+1}\rho) &\le (\theta_*^{k+1}\rho)^{-2\alpha} \fint_{B_{\theta_*^{k+1}\rho}(x_0)} \left| u - u_{x_0,\theta_*^k\rho} \right|^2 dx \\
&\le 2(\theta_*^{k+1}\rho)^{-2\alpha} \left[\fint_{B_{\theta_*^{k+1}\rho}(x_0)} \left| u - \ell_{x_0,\theta_*^k\rho} \right|^2 dx + \left| D\ell_{x_0,\theta_*^k\rho} \right|^2 (\theta_*^{k+1}\rho)^2 \right] \\
&\le 2(\theta_*^{k+1}\rho)^{-2\alpha} \left[\theta^{-n} \fint_{B_{\theta_*^k\rho}(x_0)} \left| u - \ell_{x_0,\theta_*^k\rho} \right|^2 dx + \left| D\ell_{x_0,\theta_*^k\rho} \right|^2 (\theta_*^{k+1}\rho)^2 \right] \\
&\le 4(\theta_*^k\rho)^{2-2\alpha} \left[\varepsilon_* \theta_*^{-n-2\alpha} + \left| D\ell_{x_0,\theta_*^k\rho} \right|^2 (\varepsilon_* \theta_*^{-n-2\alpha} + \theta_*^{2-2\alpha}) \right].
\end{aligned}
$$

Using (5.19) and recalling the choice of ρ_*, ε_* and θ_*, we deduce

$$
\begin{aligned}
C_\alpha(\theta_*^{k+1}\rho) &\le 4\rho_*^{2-2\alpha} \left[\varepsilon_* \theta_*^{-n-2\alpha} + n(n+2)\kappa_* \rho_*^{2-2\alpha} (\varepsilon_* \theta_*^{-n-2\alpha} + \theta_*^{2-2\alpha}) \right] \\
&\le \frac{1}{4} \rho_*^{2-2\alpha} \theta_*^{2-2\alpha} + 8n(n+2)\kappa_* \theta_*^{2-2\alpha} \\
&\le \frac{1}{4}\kappa_* + \frac{1}{2}\kappa_* < \kappa_*.
\end{aligned}
$$

This proves the second part of the assertion (A_{k+1}) and finally we conclude the proof of the lemma. □

Now, to obtain the regularity theorem (Theorem 2.2), it is similar argument as in [4, Section 3.5] by using Lemma 5.4 and so we omit it.

6 Parabolic system: VMO-coefficients

Similarly as in the elliptic case, from (2.3), we may infer a modulus of continuity function $\mu\colon [0, \infty) \to [0, \infty)$ such that μ is bounded, concave, non-decreasing and we have

$$
|D_w A(z, u, w) - D_w A(z, u, w_0)| \le L\mu \left(\frac{|w - w_0|}{1 + |w| + |w_0|} \right) (1 + |w| + |w_0|)^{p-2} \tag{6.1}
$$

for all $z \in \Omega_T$, $u \in \mathbb{R}^N$, $w, w_0 \in \mathbb{R}^{nN}$. Without loss of generality, we may assume $\mu \le 1$.

Lemma 6.1. *Let $u \in C_b^0(-T, 0; L^2(\Omega, \mathbb{R}^N)) \cap L^p(-T, 0; W^{1,p}(\Omega, \mathbb{R}^N))$ be a bounded weak solution to the parabolic system (1.2) under the structure conditions (2.3), (2.4), (2.15) and (2.16) with $\|u\|_\infty \le M$ and $2^{(10-9p)/2}\lambda > a(M)M$. For any $z_0 = (x_0, t_0) \in \Omega_T$ and $\rho \le 1$ with $Q_\rho(z_0) \Subset \Omega_T$,* *and any affine functions* $\ell\colon \mathbb{R}^n \to \mathbb{R}^N$ *with* $|\ell(x_0)| \le M$*, we have the estimate*

$$\sup_{t_0-(\rho/2)^2<t<t_0} \fint_{B_{\rho/2}(x_0)} \left|\frac{u-\ell}{\rho(1+|D\ell|)}\right|^2 dx + \fint_{Q_{\rho/2}(z_0)} \left\{\left|\frac{Du-D\ell}{(1+|D\ell|)}\right|^2 + \left|\frac{Du-D\ell}{(1+|D\ell|)}\right|^p\right\} dz$$
$$\le C_1 \Bigg[\fint_{Q_\rho(z_0)} \left\{\left|\frac{u-\ell}{\rho(1+|D\ell|)}\right|^2 + \left|\frac{u-\ell}{\rho(1+|D\ell|)}\right|^p\right\} dz$$
$$+ \omega\left(\fint_{Q_\rho(z_0)} |u-\ell(x_0)|dz\right) + V(\rho) + (a^q|D\ell|^q + b^q)\rho^q\Bigg], \tag{6.2}$$

with the constant $C_1 = C_1(\lambda, p, L, a(M), M) \ge 1$.

Proof. The following calculations are not rigorous enough. To proceed in a rigorous way, we should use a smoothing procedure in time via a family of non-negative mollifying functions or via Steklov averages, but this is standard argument and yields only technical minor changes we shall proceed formally.

Assume $z_0 \in \Omega_T$ and $\rho \le 1$ satisfy $Q_\rho(z_0) \Subset \Omega_T$. We take a cut-off functions $\chi \in C_0^\infty(B_\rho(x_0))$ and $\zeta \in C^1(\mathbb{R})$. More precisely, let us take $\tilde{t} \in (t_0 - \rho^2/4, t_0)$ and $\vartheta \in (0, \rho^2/4 - \tilde{t})$ and then $\zeta \in C^1(\mathbb{R})$ satisfying

$$\begin{cases} \zeta \equiv 1, & \text{on } (-\rho^2/4, \tilde{t} - \vartheta), \\ \zeta \equiv 0, & \text{on } (-\infty, -\rho^2) \cup (\tilde{t}, \infty), \\ 0 \le \zeta \le 1 & \text{on } \mathbb{R}, \\ \zeta_t = -1/\vartheta, & \text{on } (\tilde{t} - \vartheta, \tilde{t}), \\ |\zeta_t| \le 1/\rho^2, & \text{on } (-\rho^2, -\rho^2/4). \end{cases} \tag{6.3}$$

Moreover, $\chi \in C_0^\infty(B_\rho(x_0))$ satisfy $0 \le \chi \le 1$ on $B_\rho(x_0)$, $\chi \equiv 1$ on $B_{\rho/2}(x_0)$ and $|D\chi| \le 4/\rho$. Then $\varphi(x,t) := \zeta(t)\chi^p(x)(u(x,t) - \ell(x))$ is admissible as a test function in (2.17), and we obtain

$$\fint_{Q_\rho(z_0)} \zeta\chi^p \langle A(z,u,Du), Du - D\ell\rangle dz$$
$$= -\fint_{Q_\rho(z_0)} \langle A(z,u,Du), p\zeta\chi^{p-1} D\chi \otimes (u-\ell)\rangle dz$$
$$+ \fint_{Q_\rho(z_0)} \langle u, \partial_t \varphi\rangle dz + \fint_{Q_\rho(z_0)} \langle f, \varphi\rangle dz. \tag{6.4}$$

Furthermore, we have

$$-\fint_{Q_\rho(z_0)} \zeta\chi^p \langle A(z,u,D\ell), Du - D\ell\rangle dz$$
$$= \fint_{Q_\rho(z_0)} \langle A(z,u,D\ell), p\zeta\chi^{p-1} D\chi \otimes (u-\ell)\rangle dz - \fint_{Q_\rho(z_0)} \langle A(z,u,D\ell), D\varphi\rangle dz, \tag{6.5}$$

and

$$\fint_{Q_\rho(z_0)} \langle (A(\cdot, \ell(x_0), D\ell))_{x_0,\rho}, D\varphi\rangle dz = 0. \tag{6.6}$$

Adding these three equations and we obtain

$$
\begin{aligned}
&\fint_{Q_\rho(z_0)} \zeta\chi^p \langle A(z,u,Du) - A(z,u,D\ell), Du - D\ell\rangle dz \\
=& -\fint_{Q_\rho(z_0)} \langle A(z,u,Du) - A(z,u,D\ell), p\zeta\chi^{p-1} D\chi \otimes (u-\ell)\rangle dz \\
&-\fint_{Q_\rho(z_0)} \langle A(z,u,D\ell) - A(z,\ell(x_0),D\ell), D\varphi\rangle dz \\
&-\fint_{Q_\rho(z_0)} \langle A(z,\ell(x_0),D\ell) - (A(\cdot,\ell(x_0),D\ell))_{z_0,\rho}, D\varphi\rangle dz \\
&+\fint_{Q_\rho(z_0)} \langle u-\ell, \partial_t\varphi\rangle dz \\
&+\fint_{Q_\rho(z_0)} \langle f, \varphi\rangle dz \\
=:& \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} + \mathrm{V}.
\end{aligned} \tag{6.7}
$$

As like in elliptic case, using the ellipticity condition (2.4) and Lemma 3.11 to the left-hand side of (6.7), we get

$$
\begin{aligned}
&\langle A(z,u,Du) - A(z,u,D\ell), Du - D\ell\rangle \\
\geq& 2^{(12-9p)/2}\lambda\{(1+|D\ell|)^{p-2}|Du-D\ell|^2 + |Du-D\ell|^p\}.
\end{aligned} \tag{6.8}
$$

The terms I, II, III and V are estimated as similar as in elliptic case:

$$
\begin{aligned}
|\mathrm{I}| \leq& \varepsilon\fint_{Q_\rho(z_0)} \zeta\chi^p\{(1+|D\ell|)^{p-2}|Du-D\ell|^2 + |Du-D\ell|^p\} dz \\
&+ c(p,L,\varepsilon)\fint_{Q_\rho(z_0)} \left\{(1+|D\ell|)^{p-2}\left|\frac{u-\ell}{\rho}\right|^2 + \left|\frac{u-\ell}{\rho}\right|\right\} dz
\end{aligned} \tag{6.9}
$$

$$
\begin{aligned}
|\mathrm{II}| \leq& \varepsilon\fint_{Q_\rho(z_0)} \zeta\chi^p(1+|D\ell|)^{p-2}|Du-D\ell|^2\, dz + \varepsilon\fint_{Q_\rho(z_0)} (1+|D\ell|)^{p-2}\left|\frac{u-\ell}{\rho}\right|^2 dz \\
&+ c(p,L,\varepsilon)(1+|D\ell|)^p\omega^2\left(\fint_{Q_\rho(z_0)} |u-\ell(x_0)|^2\, dz\right),
\end{aligned} \tag{6.10}
$$

$$
|\mathrm{III}| \leq \varepsilon\fint_{Q_\rho(z_0)} \zeta\chi^p|Du-D\ell|^p\, dz + c(p,\varepsilon)\fint_{Q_\rho(z_0)} \left|\frac{u-\ell}{\rho}\right|^p dz + c(p,L,\varepsilon)(1+|D\ell|)^p V(\rho), \tag{6.11}
$$

$$
\begin{aligned}
|\mathrm{V}| \leq& a(1+\varepsilon')(2M + |D\ell|\,\rho)\fint_{Q_\rho(z_0)} \zeta\chi^p|Du-D\ell|^p\, dz + c(p,\varepsilon)\fint_{Q_\rho(z_0)} \left|\frac{u-\ell}{\rho}\right|^p dz \\
&+ \varepsilon(1+|D\ell|)^p\rho^q\left\{a^q K^q(p,\varepsilon')\,|D\ell|^q + b^q\right\}.
\end{aligned} \tag{6.12}
$$

where $\varepsilon, \varepsilon' > 0$ are fixed later and $K(p,\varepsilon') \geq 0$ is a constant from Lemma 3.10.

To estimate the term IV, recall that ζ_t satisfies $\zeta = -1/\vartheta$ on $(\tilde{t} - \vartheta, \tilde{t})$ and $|\zeta_t| \le 1/\rho^2$ on $(-\rho^2, -\rho^2/4)$. This implies

$$\begin{aligned}
\mathrm{IV} &= \fint_{Q_\rho(z_0)} \zeta_t \chi^p |u-\ell|^2 \, dz + \fint_{Q_\rho(z_0)} \zeta \chi^p \cdot \partial_t \frac{1}{2} |u-\ell|^2 \, dz \\
&= \frac{1}{2} \fint_{Q_\rho(z_0)} \zeta_t \chi^p |u-\ell|^2 \, dz \\
&= \frac{1}{2|Q_\rho(z_0)|} \int_{t_0-\rho^2}^{t_0-\rho^2/4} \int_{B_\rho(x_0)} \chi^p \left| \frac{u-\ell}{\rho} \right|^2 dxdt - \frac{1}{2|Q_\rho(z_0)|} \int_{\tilde{t}-\vartheta}^{\tilde{t}} \int_{B_\rho(x_0)} \chi^p |u-\ell|^2 \, dxdt \\
&\le \frac{1}{2} \fint_{Q_\rho(z_0)} (1+|D\ell|)^{p-2} \left| \frac{u-\ell}{\rho} \right|^2 dz - \frac{1}{2|Q_\rho(z_0)|} \int_{\tilde{t}-\vartheta}^{\tilde{t}} \int_{B_\rho(x_0)} \chi^p |u-\ell|^2 \, dz. \qquad (6.13)
\end{aligned}$$

Combining (6.7) through (6.13) then choose suitable $\varepsilon = \varepsilon(p, \lambda, a(M), M) > 0$ and $\varepsilon' = \varepsilon'(p, \lambda, a(M), M) > 0$. Further, taking the limit $\vartheta \to 0$, we obtain (6.2). $\square$

Lemma 6.2. *Assume the same assumptions in Lemma 6.1. Then for any $z_0 = (x_0, t_0) \in \Omega_T$ and $\rho \le \rho_0$ satisfying $Q_{2\rho}(z_0) \Subset \Omega_T$, and any affine functions $\ell \colon \mathbb{R}^n \to \mathbb{R}^N$ with $|\ell(x_0)| \le M$, the inequality*

$$\begin{aligned}
&\fint_{Q_\rho(z_0)} \big(\langle v, \varphi_t \rangle - \mathcal{A}(Dv, D\varphi) \big) dz \\
&\le C_2 (1+|D\ell|) \left[\mu^{1/2} \Big(\sqrt{\Psi_*(z_0, 2\rho, \ell)} \Big) \sqrt{\Psi_*(z_0, 2\rho, \ell)} + \Psi_*(z_0, 2\rho, \ell) + \rho(a|D\ell|^p + b) \right] \sup_{Q_\rho(z_0)} |D\varphi|
\end{aligned} \qquad (6.14)$$

holds for all $\varphi \in C_0^\infty(Q_\rho(x_0), \mathbb{R}^N)$ and a constant $C_2 = C_2(n, \lambda, L, p, a(M)) \ge 1$, where

$$\begin{aligned}
v &:= u - \ell = u - \ell(x_0) - D\ell(x - x_0), \\
\mathcal{A}(Dv, D\varphi) &:= \frac{1}{(1+|D\ell)^{p-1}} \Big\langle \big(\partial_w A(\cdot, \ell(x_0), D\ell) \big)_{z_0,\rho} Dv, D\varphi \Big\rangle, \\
\Psi(z_0, \rho, \ell) &:= \fint_{Q_\rho(z_0)} \left\{ \left| \frac{u-\ell}{\rho(1+|D\ell|)} \right|^2 + \left| \frac{u-\ell}{\rho(1+|D\ell|)} \right|^p \right\} dz, \\
\Psi_*(z_0, \rho, \ell) &:= \Psi(z_0, \rho, \ell) + \omega \left(\fint_{Q_\rho(z_0)} |u - \ell(x_0)|^2 dz \right) + V(\rho) + (a^q |D\ell|^q + b^q) \rho^q.
\end{aligned}$$

To obtain the estimate (6.14), calculation goes parallel as in Lemma 5.2 and so that we omit the proof of Lemma 6.2.

From now on, let us write $\Phi(\rho) = \Phi(z_0, \rho, \ell_{z_0,\rho})$, $\Psi(\rho) = \Psi(z_0, \rho, \ell_{z_0,\rho})$, $\Psi_*(\rho) = \Psi_*(z_0, \rho, \ell_{z_0,\rho})$ for fixed $z_0 \in \Omega_T$ and $0 < \rho \le 1$ as like in previous sections. Here $\ell_{z_0,\rho}$ is a minimizer which we introduce in Section 3.

Lemma 6.3. *Assume the same assumptions in Lemma 6.1. Let $\theta \in (0, 1/4]$ be an arbitrary and impose the following smallness conditions on the excess:*

(i) $\mu^{1/2}\Big(\sqrt{\Psi_*(\rho)}\Big) + \sqrt{\Psi_*(\rho)} \le \frac{\delta}{2}$ *with the constant* $\delta = \delta(n, N, p, \lambda, L, \theta^{n+p+4})$ *from Lemma 3.3,*

(ii) $\Psi(\rho) \le \frac{\theta^{n+4}}{4n(n+2)}$,

(iii) $\gamma(\rho) := [\Psi_*^{q/2}(\rho) + \delta^{-q}\rho^q(a|D\ell| + b)^q]^{1/q} \le 1$.

Then there holds the excess improvement estimate

$$\Psi(\theta\rho) \le C_3\theta^2\Psi_*(\rho) \tag{6.15}$$

with a constant $C_3 = C_3(n, \lambda, L, p, a(M)) \ge 1$.

Proof. Set

$$w := \frac{u-\ell}{C(1+|D\ell|)\gamma(\rho)}.$$

As like in Lemma 5.3, the assumptions (i), (ii), (iii) and the claim of Lemma 5.2 enable us to use $\mathcal{A}$-caloric approximation lemma (Lemma 3.3) so that there exists a function

$$h \in L^p(t_0 - (\rho/4)^2, t_0; W^{1,2}(B_{\rho/4}(x_0)), \mathbb{R}^N)$$

which is $\mathcal{A}$-caloric on $Q_{\rho/4}(z_0)$ and satisfies

$$⨍_{Q_\rho(z_0)} \left(\left|\frac{h}{\rho/4}\right|^2 + \gamma^{p-2}\left|\frac{h}{\rho/4}\right|^p\right) dz + ⨍_{Q_\rho(z_0)} \left(|Dh|^2 + \gamma^{p-2}|Dh|^p\right) dz \le 2\cdot 2^{n+2+2p}$$

and

$$⨍_{Q_\rho(z_0)} \left(\left|\frac{w-h}{\rho/4}\right|^2 + \gamma^{p-2}\left|\frac{w-h}{\rho/4}\right|^p\right) dz \le \theta^{n+p+4}. \tag{6.16}$$

Then from Lemma 3.5, we have for $s=2$ respectively for $s=p$

$$\begin{aligned}
&\gamma^{s-2}(\theta\rho)^{-s}⨍_{Q_{\theta\rho}(z_0)} \left|h - h_{z_0,\rho/4} - (Dh)_{z_0,\rho/4}(x-x_0)\right|^s dz\\
\le& c(s)\gamma^{s-2}\theta^s\left(\frac{\rho}{4}\right)^{-s}⨍_{Q_{\rho/4}(z_0)} \left|h - h_{z_0,\rho/4} - (Dh)_{z_0,\rho/4}(x-x_0)\right|^s dz\\
\le& 3^{s-1}c\gamma^{s-2}\theta^s\left(\frac{\rho}{4}\right)^{-s}\left[⨍_{Q_{\rho/4}(z_0)} |h|^s\,dz + \left|h_{z_0,\rho/4}\right|^s + \left|(Dh)_{z_0,\rho/4}\right|^s\left(\frac{\rho}{4}\right)^s\right]\\
\le& 2\cdot 3^{s-1}c\gamma^{s-2}\theta^s\left[\left(\frac{\rho}{4}\right)^{-s}⨍_{Q_{\rho/4}(z_0)} |h|^s\,dz + ⨍_{Q_{\rho/4}(z_0)} |Dh|^s\,dz\right]\\
\le& 2^{n+4+p}\cdot 3^{s-1}c\theta^s.
\end{aligned}$$

Thus, using (6.16) we obtain

$$\begin{aligned}
&\gamma^{s-2}(\theta\rho)^{-s} \fint_{Q_{\theta\rho}(z_0)} \left|w - h_{z_0,\rho/4} - (Dh)_{z_0,\rho/4}(x-x_0)\right|^s dz \\
\le& 2^{s-1}(\theta\rho)^{-s}\left[\fint_{Q_{\theta\rho}(z_0)} \gamma^{s-2}|w-h|^s\, dz + \gamma^{s-2}\fint_{Q_{\theta\rho}(z_0)} \left|h - h_{z_0,\rho/4} - (Dh)_{z_0,\rho/4}(x-x_0)\right|^s dz\right] \\
\le& 2^{s-1}\left[4^{n+2-s}\theta^{-n-2-s}\fint_{Q_{\rho/4}(z_0)} \gamma^{s-2}\left|\frac{w-h}{\rho/4}\right|^s dz + 3^{s-1}\cdot 2^{n+4+p}c\theta^s\right] \\
\le& 2^{s-1}\Big(4^{n+2-s} + 3^{s-1}\cdot 2^{n+4+p}c(s)\Big)\theta^2.
\end{aligned}$$

Scaling back to u, we have

$$\begin{aligned}
&(\theta\rho)^{-s}\fint_{Q_{\theta\rho}(z_0)} |u - \ell_{z_0,\theta\rho}|^s\, dz \\
\le& c(n,s)(\theta\rho)^{-s}\fint_{Q_{\theta\rho}(z_0)} \left|u - \ell_{z_0,\rho} - C\gamma(1+|D\ell_{z_0,\rho}|)(h_{z_0,\rho/4} - (Dh)_{z_0,\rho/4}(x-x_0))\right|^s dz \\
=& cC^s\gamma^s(1+|D\ell_{z_0,\rho}|)^s(\theta\rho)^{-s}\fint_{Q_{\theta\rho}(z_0)} \left|w - h_{z_0,\rho/4} - (Dh)_{z_0,\rho/4}(x-x_0)\right|^s dz \\
\le& c(n,s,p,C)\gamma^2(1+|D\ell_{z_0,\rho}|)^s\theta^s \\
\le& c(1+|D\ell_{z_0,\rho}|)^s\theta^2[\Psi_*{}^{q/2}(\rho) + 2^{q/p}\delta^{-q}\Psi_*(\rho)]^{2/q} \\
\le& c(1+|D\ell_{z_0,\rho}|)^s\theta^s\Psi_*(\rho).
\end{aligned}$$

By the similar arguments in Lemma 5.2, we can replace the term $(1+|D\ell_{z_0,\rho}|)$ by $(1+|D\ell_{z_0,\theta\rho}|)$ and this immediately yields the claim. □

Let fix an arbitrarily Hölder exponent $\alpha \in (0,1)$ and define the Campanato-type excess

$$C_\alpha(z_0,\rho) := C_\alpha(\rho) = \rho^{-2\alpha}\fint_{Q_\rho(z_0)} |u - u_{z_0,\rho}|^2\, dz.$$

The similar arguments as in Lemma 5.4 yields the following lemma.

Lemma 6.4. *Assume the same assumption in Lemma 6.1. For every $\alpha \in (0,1)$ there exist constants $\varepsilon_*, \kappa_*, \rho_* > 0$ and $\theta_* \in (0,1/8]$ such that the conditions*

$$\Psi(\rho) < \varepsilon_* \qquad \text{and} \qquad C_\alpha(\rho) < \kappa_* \tag{A_0}$$

for all $0 < \rho < \rho_*$ with $Q_\rho(z_0) \Subset \Omega_T$, imply

$$\Psi(\theta_*^k\rho) < \varepsilon_* \qquad \text{and} \qquad C_\alpha(\theta_*^k\rho) < \kappa_* \tag{A_k}$$

respectively, for every $k \in \mathbb{N}\cup\{0\}$.

Now it is easy to obtain Theorem 2.3 from Lemma 6.4 by using the integral characterization of Hölder continuous functions with respect to the parabolic metric of Campanato-Da Prato [9].

Acknowledgments
The author thanks Professor Hisashi Naito for helpful discussions through out the paper. The author also thanks Professor Atsushi Tachikawa from Tokyo University of Science for variable comments to our second result, the case of elliptic systems with VMO-coefficients, which is obtained as the Master thesis in Tokyo University of Science under advising by Professor Atsushi Tachikawa.

References

[1] P. Baroni, *Regularity in parabolic Dini continuous systems*, Forum Math. **23** (2011), no. 6, 1281–1322.

[2] L. Beck, *Partial regularity for weak solutions of nonlinear elliptic systems: the subquadratic case*, Manuscripta Math. **123** (2007), no. 4, 453–491.

[3] L. Beck, *Partial Hölder continuity for solutions of subquadratic elliptic systems in low dimensions*, J. Math. Anal. Appl. **354** (2009), no. 1, 301–318.

[4] V. Bögelein, F. Duzaar, J. Habermann, and C. Scheven, *Partial Hölder continuity for discontinuous elliptic problems with VMO-coefficients*, Proc. Lond. Math. Soc. (3) **103** (2011), 371–404.

[5] V. Bögelein, F. Duzaar, and G. Mingione, *The boundary regularity of non-linear parabolic systems. I*, Ann. Inst. H. Poincaré Anal. Non Linéaire **27** (2010), no. 1, 201–255.

[6] V. Bögelein, F. Duzaar, and G. Mingione, *The boundary regularity of non-linear parabolic systems. II*, Ann. Inst. H. Poincaré Anal. Non Linéaire **27** (2010), no. 1, 145–200.

[7] V. Bögelein, M. Foss, and G. Mingione, *Regularity in parabolic systems with continuous coefficients*, Math. Z. **270** (2012), no. 3-4, 903–938.

[8] S. Chen and Z. Tan, *Optimal interior partial regularity for nonlinear elliptic systems under the natural growth condition: the method of A-harmonic approximation*, Acta Math. Sci. Ser. B Engl. Ed. **27** (2007), no. 3, 491–508.

[9] G. Da Prato, *Spazi* $\mathcal{L}^{(p,\theta)}(\Omega,\delta)$ *e loro proprietà*, Ann. Mat. Pura Appl. (4) **69** (1965), 383–392.

[10] E. De Giorgi, *Frontiere orientate di misura minima*, Seminario di Matematica della Scuola Normale Superiore di Pisa, 1960-61, Editrice Tecnico Scientifica, Pisa, 1961.

[11] F. Duzaar and A. Gastel, *Nonlinear elliptic systems with Dini continuous coefficients*, Arch. Math. (Basel) **78** (2002), no. 1, 58–73.

[12] F. Duzaar and J. F. Grotowski, *Optimal interior partial regularity for nonlinear elliptic systems: the method of A-harmonic approximation*, Manuscripta Math. **103** (2000), 267–298.

[13] F. Duzaar and G. Mingione, *Second order parabolic systems, optimal regularity, and singular sets of solutions*, Ann. Inst. H. Poincaré Anal. Non Linéaire **22** (2005), no. 6, 705–751.

[14] F. Duzaar and G. Mingione, *Harmonic type approximation lemmas*, J. Math. Anal. Appl. **352** (2009), no. 1, 301–335.

[15] F. Duzaar, G. Mingione, and K. Steffen, *Parabolic systems with polynomial growth and regularity*, Mem. Amer. Math. Soc. **214** (2011), no. 1005, x+118.

[16] M. Foss and J. Geisbauer, *Partial regularity for subquadratic parabolic systems with continuous coefficients*, Manuscripta Math. **139** (2012), no. 1-2, 1–47.

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

[18] M. Giaquinta, *Multiple integrals in the calculus of variations and nonlinear elliptic systems*, Annals of Mathematics Studies, vol. 105, Princeton University Press, Princeton, NJ, 1983.

[19] M. Giaquinta and L. Martinazzi, *An introduction to the regularity theory for elliptic systems, harmonic maps and minimal graphs*, Appunti. Scuola Normale Superiore di Pisa (Nuova Serie) [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)], vol. 2, Edizioni della Normale, Pisa, 2005.

[20] M. Giaquinta and G. Modica, *Almost-everywhere regularity results for solutions of nonlinear elliptic systems*, Manuscripta Math. **28** (1979), 109–158.

[21] M. Giaquinta and G. Modica, *Partial regularity of minimizers of quasiconvex integrals*, Ann. Inst. H. Poincaré Anal. Non Linéaire **3** (1986), 185–208.

[22] T. Kanazawa, *Partial regularity for elliptic systems with VMO-coefficients*, Riv. Math. Univ. Parma (N.S.) **5** (2014), no. 2, 311-333, arXiv:1302.4148 [math.AP].

[23] T. Kanazawa, *Partial regularity result for elliptic systems with Dini continuous coefficients and q-growth*, arXiv:1307.1946 [math.AP]

[24] T. Kanazawa, *Partial regularity for parabolic systems with VMO-coefficients*, arXiv:1312.5044 [math.AP]

[25] Y. Qiu, *Interior partial regularity for nonlinear elliptic systems with Dini continuous coefficients for the case:* $1 < m < 2$, J. Math. Anal. Appl. **387** (2012), 885–908.

[26] C. Scheven, *Partial regularity for subquadratic parabolic systems by $\mathcal{A}$-caloric approximation*, Rev. Mat. Iberoam. **27** (2011), no. 3, 751–801.

[27] L. Simon, *Theorems on regularity and singularity of energy minimizing maps*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1996, Based on lecture notes by Norbert Hungerbühler.

Taku Kanazawa
Graduate School of Mathematics
Nagoya University
Chikusa-ku, Nagoya, 464-8602, JAPAN
taku.kanazawa@math.nagoya-u.ac.jp