Partial regularity for minimizers of discontinuous quasiconvex integrals with general growth

Christopher Goodrich [ID]

School of Mathematics and Statistics, UNSW Sydney, Sydney, NSW 2052, Australia (c.goodrich@unsw.edu.au)

Giovanni Scilla [ID]

Dipartimento di Scienze di Base ed Applicate per l'Ingegneria (SBAI), Sapienza Università di Roma, Via A. Scarpa 16, 00169 Roma, Italy (giovanni.scilla@uniroma1.it)

Bianca Stroffolini [ID]

Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione, Università di Napoli Federico II, Via Claudio, 80125 Napoli, Italy (bstroffo@unina.it)

We prove the partial Hölder continuity for minimizers of quasiconvex functionals

$$\mathcal{F}(\mathbf{u}) := \int_\Omega f(x, \mathbf{u}, D\mathbf{u})\, \mathrm{d}x,$$

where f satisfies a uniform VMO condition with respect to the x-variable and is continuous with respect to $\mathbf{u}$. The growth condition with respect to the gradient variable is assumed a general one.

1. Introduction

In this paper we study the partial regularity of minimizers of the integral functional

$$\mathcal{F}(\mathbf{u}) := \int_\Omega f(x, \mathbf{u}, D\mathbf{u})\, \mathrm{d}x, \tag{1.1}$$

where $\Omega \subseteq \mathbb{R}^n$ is an open bounded set and $\mathbf{u} : \Omega \to \mathbb{R}^N$, with n, $N \geqslant 2$; i.e., we consider vectorial minimizers of $\mathcal{F}$. The growth conditions we impose on $f = f(x, \mathbf{u}, \mathbf{P})$ are quite general, being as they permit 'general growth condition' with respect to the gradient variable. This allows us to treat in a unified way the degenerate (when $p > 2$) or singular (when $p < 2$) behaviour. We assume with respect to x a

weak VMO condition, uniformly in $(\mathbf{u}, \mathbf{P})$, and continuity with respect to $\mathbf{u}$. Our main result, theorem 1.1, proves that a minimizer of (1.1) is locally Hölder continuous for any Hölder exponent $0 < \alpha < 1$; i.e., if $\mathbf{u}$ is a minimizer of (1.1), then $\mathbf{u} \in C^{0,\alpha}_{\mathrm{loc}}(\Omega_0, \mathbb{R}^N)$, where $\Omega_0 \subset \Omega$ is an open set of full measure specified in the statement of theorem 1.1 later in this section.

1.1. Literature review

We begin by explaining how the study of functional (1.1) fits into the broader regularity theory research over the past many years. Before proceeding further, we point out that Mingione [**35**] has provided a comprehensive account of the various areas of study within regularity theory for integral functionals and PDEs; it is an excellent reference for those wishing to read a broad overview of the various areas of interest within the larger realm of regularity theory.

As already mentioned we allow f to satisfy a VMO-type condition with respect to x. More precisely the partial map $x \mapsto (f(x, \mathbf{u}, \mathbf{P}))/(\varphi(|\mathbf{P}|))$ satisfies a uniform VMO condition; here φ is an N-function—see condition (F4) later in this section for the precise formulation. As a consequence we allow a certain controlled discontinuous behaviour with respect to the spatial variable in the integrand of (1.1). We prove partial Hölder continuity for the local minimizers. The first paper who considered low order regularity (for variational integrals) was the one by Foss and Mingione [**23**], where they were assuming continuity with respect to x and $\mathbf{u}$. Thereafter Kristensen and Mingione [**29**] proved Hölder continuity for convex integral functionals with continuous coefficients for a fixed Hölder exponent depending on the dimension and the growth exponent. Stronger assumptions as Dini-type conditions [**20**] lead to partial C^1-regularity. It is worth mentioning the uniform porosity of the singular set for Lipschitzian minimizers of quasiconvex functionals [**30**].

The space of functions with vanishing mean oscillation (VMO) has been introduced by Sarason in the realm of harmonic analysis, see [**37**]. It has had several applications in connection with Hardy spaces, Riesz transforms or nonlinear commutators, see [**27**, **39**] and references therein. In the early 90s, Chiarenza, Frasca and Longo [**8**] studied non-divergence form equations with VMO coefficients by means of singular integrals operators, see also [**18**, **19**].

The study of functionals satisfying a VMO-type condition has been broadened considerably over the past couple decades, see [**9**, **38**]. Recently, Bögelein, Duzaar, Habermann and Scheven [**5**] considered a functional of the form (1.1) under the assumption that $(x, \mathbf{u}, \mathbf{P}) \mapsto f(x, \mathbf{u}, \mathbf{P})$ satisfies a type of VMO assumption in x, uniformly with respect to $\mathbf{u}$ and $\mathbf{P}$; they further considered an analogous elliptic system of the form $\nabla \cdot a(x, \mathbf{u}, D\mathbf{u}) = 0$, in which, again, the coefficient a was assumed to satisfy a VMO-type condition with respect to its spatial coordinate. Moreover, the integral functional they studied was assumed to be quasi-convex. However, unlike our study, they assumed that the growth of f with respect to $\mathbf{P}$ was standard p-growth, $p \geqslant 2$.

Similarly, Bögelein [**4**] studied quasi-convex integral functionals in the vectorial case. But the assumed growth of the integrand with respect to the gradient was standard p-growth. It was also assumed that the map $x \mapsto (f(x, \mathbf{u}, \mathbf{P}))/((1 + |\mathbf{P}|)^p)$ was VMO, uniformly with respect to $\mathbf{u}$ and $\mathbf{P}$. Bögelein, Duzaar, Habermann and

Scheven [**6**] made some similar assumptions when considering a system of PDEs involving the symmetric part of the gradient $D\mathbf{u}$, wherein the coefficients on the symmetric part were VMO.

Goodrich [**26**] then further generalized, in part, the results of [**5**] by considering (1.1) in the case where $x \mapsto f(x, \mathbf{u}, \mathbf{P})$ was VMO, uniformly with respect to $\mathbf{u}$ and $\mathbf{P}$, and, furthermore, in which f was only asymptotically convex.

Next, the study of problems with general growth conditions has been initiated by Marcellini in a list of papers [**32**–**34**] and it is now very rich—see, e.g., [**7**, **10**, **15**–**17**, **40**]. In particular, Marcellini and Papi proved the Lipschitz bound for a solution of an elliptic system of Uhlenbeck type with general growth. In view of comparison estimates, it is worth mentioning the paper [**15**], where the $C^{1,\alpha}$ regularity is proven via an excess decay estimate. Very recently, De Filippis and Mingione have relaxed the hypotheses by considering also growth of exponential type (no Δ_2-condition) [**11**].

So, we see that many papers in recent years have treated *either* VMO-type coefficient problems *or* general growth problems. To our knowledge, it seems that the combination of these two generalities has not been considered as we do in this paper. Thus, the results of this paper significantly generalize many of the previously mentioned papers.

1.2. Strategy of the proof

We briefly explain the strategy of the proof of the main result. As a major difficulty with respect to the proof by Bögelein or Duzaar *et al.* in the p-setting, we can't rely on homogeneity of the function φ. In particular, an analogue of the Campanato excess

$$\Psi_\alpha(x_0, \varrho) := \varrho^{-\alpha p} \fint_{B_\varrho(x_0)} |\mathbf{u} - (\mathbf{u})_{x_0,\varrho}|^p \, \mathrm{d}x$$

defined there and playing a key role in the iteration process could not be easily handled in the Orlicz setting.

Our strategy is to find carefully the two quantities which play the role both in the non-degenerate and in the degenerate cases. The first leading quantity is the excess functional

$$\Phi(x_0, \varrho) := \fint_{B_\varrho(x_0)} \varphi_{|(D\mathbf{u})_{x_0,\varrho}|}(|D\mathbf{u} - (D\mathbf{u})_{x_0,\varrho}|) \, \mathrm{d}x$$

(see (3.14)). In the non-degenerate case, when

$$\Phi(x_0, \varrho) \leqslant \varphi(|(D\mathbf{u})_{x_0,\varrho}|), \tag{1.2}$$

we linearize the problem, via the $\mathcal{A}$-harmonic approximation [**17**]. This procedure, exploiting assumptions (F4)–(F5) and a freezing technique (with respect to the variables x and $\mathbf{u}$) based on the Ekeland variational principle, provides a comparison map which is an almost minimizer of the frozen functional and whose gradient is L^1-close to that of the original minimizer (see lemma 3.8). Such comparison map

is shown to be approximately $\mathcal{A}$-harmonic, and this property is inherited by the minimizer itself via the comparison estimate. This allows to prove an excess-decay estimate, which, in turn, permits the iteration of the rescaled excess $\frac{\Phi(x_0, \varrho)}{\varphi(|(D\mathbf{u})_{x_0,\varrho}|)}$ and of a 'Morrey-type' excess

$$\Theta(x_0, \varrho) := \varrho \varphi^{-1} \left(\fint_{B_\varrho(x_0)} \varphi(|D\mathbf{u}|) \, \mathrm{d}x \right)$$

at each scale. Namely, there exists $\vartheta \in (0,1)$ such that, if the boundedness conditions

$$\frac{\Phi(x_0, \varrho)}{\varphi(|(D\mathbf{u})_{x_0,\varrho}|)} \leqslant \varepsilon_* \quad \text{and} \quad \Theta(x_0, \varrho) \leqslant \delta_*$$

hold on some ball $B_\varrho(x_0)$, then

$$\frac{\Phi(x_0, \vartheta^m \varrho)}{\varphi(|(D\mathbf{u})_{x_0,\vartheta^m \varrho}|)} \leqslant \varepsilon_* \quad \text{and} \quad \Theta(x_0, \vartheta^m \varrho) \leqslant \delta_*$$

hold for every $m = 0, 1, \ldots$. Therefore, $\Theta(x_0, \varrho)$ is the adequate excess playing the role of Ψ_α in our setting.

In the degenerate case, when

$$\Phi(x_0, \varrho) \geqslant \kappa \varphi(|(D\mathbf{u})_{x_0,\varrho}|) \tag{1.3}$$

for some $\kappa < 1$, we perform a different linearization procedure: the assumption (F7) coupled with an analogous freezing argument as before provides, now, the almost φ-harmonicity of the minimizer via the application of the φ-harmonic approximation [16] to the comparison map. The corresponding excess improvement implies that if the excess is small at radius ϱ it is also small at some smaller radius $\theta\varrho$, for $\theta < 1$. The key point in this iteration process is that the boundedness of both the excess Φ and the Morrey excess Θ at some scale $\vartheta\theta^{k_0}\varrho$ ('switching radius') under assumption (1.2) is satisfied exactly when the degenerate bound (1.3) fails and therefore we can proceed the iteration in the non-degenerate regime. Notice that, if on the one hand $|(D\mathbf{u})_{x_0,\varrho}|$ might blow up in the iteration since we cannot expect C^1-regularity, on the other hand the Morrey excess $\Theta(x_0, \theta^k\varrho)$ stays bounded, exactly as it should be for a $C^{0,\alpha}$-regularity result. In addition, if at level k_0 the regime is non-degenerate, the behaviour stays non-degenerate at any subsequent level $k \geqslant k_0$, and the iteration can proceed. The smallness of Θ at any level ensures Hölder continuity of $\mathbf{u}$ in x_0 provided the excess functionals Φ and Θ are small at some initial radius ϱ (actually, this holds in a neighbourhood of x_0, since these smallness conditions are open). Finally, it is then proven that such a smallness condition on the excesses is indeed satisfied on the complement of the set $\Sigma_1 \cup \Sigma_2$ of theorem 1.1.

1.3. Assumptions and statement of the main result

We list here the main assumptions on the integral functional that we are going to study throughout the paper. We assume that $\varphi : [0, \infty) \to [0, \infty)$ is an N-function such that

(φ1) $\varphi \in C^1([0, \infty)) \cap C^2((0, \infty))$;

(φ2) $0 < \mu_1 - 1 \leqslant \inf_{t>0} \frac{t\varphi''(t)}{\varphi'(t)} \leqslant \sup_{t>0} \frac{t\varphi''(t)}{\varphi'(t)} \leqslant \mu_2 - 1$, for suitable constants $1 < \mu_1 \leqslant \mu_2$.

We may assume, without loss of generality, that $1 < \mu_1 < 2 < \mu_2$.

For the precise notation and definitions, as well as the additional assumptions we will require on φ, we refer to § 2.

We assume the integrand $f : \Omega \times \mathbb{R}^N \times \mathbb{R}^{N \times n} \to \mathbb{R}$, $f = f(x, \mathbf{u}, \mathbf{P})$ to be Borel-measurable, such that the partial map $\mathbf{P} \to f(\cdot, \cdot, \mathbf{P}) \in C^1(\mathbb{R}^{N \times n}) \cap C^2(\mathbb{R}^{N \times n} \setminus \{\mathbf{0}\})$. We will denote by Df and $D^2 f$ the corresponding first and second gradients, respectively, for fixed x and $\mathbf{u}$. We require f to comply with the following assumptions:

(F1) *coercivity:* there exists $\nu > 0$ such that

$$\nu\varphi(|\mathbf{P}|) \leqslant f(x, \mathbf{u}, \mathbf{P}) - f(x, \mathbf{u}, \mathbf{0})$$

uniformly in $x \in \Omega$ and $\mathbf{u} \in \mathbb{R}^N$, for every $\mathbf{P} \in \mathbb{R}^{N \times n}$;

(F2) *φ-growth conditions with respect to the $\mathbf{P}$ variable:* there exists a constant $L > 0$ such that

$$|Df(x, \mathbf{u}, \mathbf{P})| \leqslant L\varphi'(|\mathbf{P}|), \quad |D^2 f(x, \mathbf{u}, \mathbf{P})| \leqslant L\varphi''(|\mathbf{P}|),$$

uniformly in $x \in \Omega$ and $\mathbf{u} \in \mathbb{R}^N$, for every $\mathbf{P} \in \mathbb{R}^{N \times n}$ with $|\mathbf{P}| \neq 0$;

(F3) *f is degenerate quasiconvex;* i.e.,

$$\int_B f(x, \mathbf{u}, \mathbf{P} + D\boldsymbol{\eta}(y)) - f(x, \mathbf{u}, \mathbf{P}) \, dy$$

$$\geqslant \nu \int_B \varphi''(\mu + |\mathbf{P}| + |D\boldsymbol{\eta}(y)|) |D\boldsymbol{\eta}(y)|^2 \, dy,$$

for every $x \in \Omega$, $\mathbf{u} \in \mathbb{R}^N$, every ball $B \subset \Omega$, $\mathbf{P} \in \mathbb{R}^{N \times n}$ and $\boldsymbol{\eta} \in C_0^\infty(B, \mathbb{R}^N)$, $\mu \geqslant 0$;

(F4) *the function $x \mapsto f(x, \mathbf{u}, \mathbf{P})/\varphi(|\mathbf{P}|)$ satisfies a VMO-condition, uniformly with respect to $(\mathbf{u}, \mathbf{P})$:*

$$|f(x, \mathbf{u}, \mathbf{P}) - (f(\cdot, \mathbf{u}, \mathbf{P}))_{x_0, r}| \leqslant v_{x_0}(x, r)\varphi(|\mathbf{P}|), \quad \text{for all} x \in B_r(x_0)$$

where $x_0 \in \Omega$, $r \in (0, 1]$ and $\mathbf{P} \in \mathbb{R}^{N \times n}$ and $v_{x_0} : \mathbb{R}^n \times [0, 1] \to [0, 2L]$ are bounded functions such that

$$\lim_{\varrho \to 0} \mathcal{V}(\varrho) = 0, \quad \text{where } \mathcal{V}(\varrho) := \sup_{x_0 \in \Omega} \sup_{0 < r \leqslant \varrho} \fint_{B_r(x_0)} v_{x_0}(x, r) \, dx,$$

and

$$(f(\cdot, \mathbf{u}, \mathbf{P}))_{x_0, r} := \frac{1}{|B_r(x_0)|} \int_{B_r(x_0)} f(x, \mathbf{u}, \mathbf{P}) \, dx \, ;$$

(F5) *f is uniformly continuous with respect to the* **u** *variable; i.e.,*

$$|f(x, \mathbf{u}, \mathbf{P}) - f(x, \mathbf{u}_0, \mathbf{P})| \leqslant L\omega(|\mathbf{u} - \mathbf{u}_0|)\varphi(|\mathbf{P}|),$$

where $\omega : [0, \infty) \to [0, 1]$ is a nondecreasing, concave modulus of continuity; i.e., $\lim_{t\downarrow 0} \omega(t) = \omega(0) = 0$.

(F6) *the second derivatives $D^2 f$ are Hölder continuous away from* **0** *with some exponent $\beta_0 \in (0, 1)$ such that uniformly in $(x, \mathbf{u})$ and for $0 < |\mathbf{P}| \leqslant \frac{1}{2}|\mathbf{Q}|$*

$$|D^2 f(x, \mathbf{u}, \mathbf{P}) - D^2 f(x, \mathbf{u}, \mathbf{P} + \mathbf{Q})| \leqslant c_0\, \varphi'(|\mathbf{Q}|)\, |\mathbf{Q}|^{-\beta_0} |\mathbf{P}|^{\beta_0}\, ;$$

(F7) *the function $\mathbf{P} \to Df(x, \mathbf{u}, \mathbf{P})$ behaves asymptotically at 0 as the φ-Laplacian; i.e.,*

$$\lim_{t\to 0} \frac{Df(x, \mathbf{u}, t\mathbf{P})}{\varphi'(t)} = \mathbf{P},$$

uniformly in $\{\mathbf{P} \in \mathbb{R}^{N\times n} : |\mathbf{P}| = 1\}$ and uniformly for all $x \in \Omega$ and $\mathbf{u} \in \mathbb{R}^N$.

Our main regularity result can be stated as follows. Note that the definition of **V** appearing in Σ_1 can be found in (2.3).

THEOREM 1.1. *Let $\Omega \subset \mathbb{R}^n$ be an open bounded domain, φ a convex function satisfying assumptions $(\varphi 1)$–$(\varphi 3)$ and consider a minimizer $\mathbf{u} \in W^{1,\varphi}(\Omega, \mathbb{R}^N)$ to the functional (1.1) under assumptions (F1)–(F7). Then there exists an open subset $\Omega_0 \subset \Omega$ such that*

$$\mathbf{u} \in C_{loc}^{0,\alpha}\left(\Omega_0, \mathbb{R}^N\right) \qquad and \qquad |\Omega \setminus \Omega_0| = 0$$

for every $\alpha \in (0, 1)$. Moreover, $\Omega \setminus \Omega_0 \subset \Sigma_1 \cup \Sigma_2$ where

$$\Sigma_1 := \left\{ x_0 \in \Omega : \liminf_{\varrho \searrow 0} \fint_{B_\varrho(x_0)} |\mathbf{V}_{|(D\mathbf{u})_{x_0,\varrho}}(D\mathbf{u} - (D\mathbf{u})_{x_0,\varrho})|^2 \, \mathrm{d}x > 0 \right\},$$

$$\Sigma_2 := \left\{ x_0 \in \Omega : \limsup_{\varrho \searrow 0} |(D\mathbf{u})_{x_0,\varrho}| = +\infty \right\}.$$

2. Preliminaries and basic results

2.1. Some basic facts on N–functions

We recall here some elementary definitions and basic results about Orlicz functions. The following definitions and results can be found, e.g., in [**1**, **3**, **28**, **31**].

A real-valued function $\varphi : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ is said to be an *N-function* if it is convex and satisfies the following conditions: $\varphi(0) = 0$, φ admits the derivative φ' and this derivative is right continuous, non-decreasing and satisfies $\varphi'(0) = 0$, $\varphi'(t) > 0$ for $t > 0$, and $\lim_{t\to\infty} \varphi'(t) = \infty$.

We say that φ satisfies the Δ_2-*condition* if there exists $c > 0$ such that for all $t \geqslant 0$ holds $\varphi(2t) \leqslant c\,\varphi(t)$. We denote the smallest possible such constant by $\Delta_2(\varphi)$. Since

$\varphi(t) \leqslant \varphi(2t)$, the Δ_2-condition is equivalent to $\varphi(2t) \sim \varphi(t)$, where '$\sim$' indicates the equivalence between N-functions.

By L^φ and $W^{1,\varphi}$ we denote the classical Orlicz and Orlicz-Sobolev spaces, i. e. $f \in L^\varphi$ if and only if $\int \varphi(|f|)\,\mathrm{d}x < \infty$ and $f \in W^{1,\varphi}$ if and only if $f, Df \in L^\varphi$. The space $W_0^{1,\varphi}(\Omega)$ will denote the closure of $C_0^\infty(\Omega)$ in $W^{1,\varphi}(\Omega)$.

We define the function $(\varphi')^{-1} \colon \mathbb{R}_0^+ \to \mathbb{R}_0^+$ as

$$(\varphi')^{-1}(t) := \sup\{s \in \mathbb{R}_0^+ \ : \ \varphi'(s) \leqslant t\}.$$

If φ' is strictly increasing, then $(\varphi')^{-1}$ is the inverse function of φ'. Then $\varphi^* \colon \mathbb{R}_0^+ \to \mathbb{R}_0^+$ with

$$\varphi^*(t) := \int_0^t (\varphi')^{-1}(s)\,\mathrm{d}s$$

is again an N-function and $(\varphi^*)'(t) = (\varphi')^{-1}(t)$ for $t > 0$. φ^* is the Young-Fenchel-Yosida conjugate function of φ. Note that $\varphi^*(t) = \sup_{a \geqslant 0}(at - \varphi(a))$ and $(\varphi^*)^* = \varphi$. When both φ and φ^* satisfy Δ_2-condition, by elementary convex analysis it is easy to see that for all $\delta > 0$ there exists c_δ (only depending on $\Delta_2(\varphi)$ and $\Delta_2(\varphi^*)$) such that for all $t, a \geqslant 0$ it holds that

$$at \leqslant \delta\,\varphi(t) + c_\delta\,\varphi^*(a).$$

PROPOSITION 2.1. *Let φ be an N-function complying with $(\varphi1)$ and $(\varphi2)$. Then*

(i) *it holds that*

$$\varphi'(t) \sim t\,\varphi''(t) \tag{2.1}$$

uniformly in $t > 0$. The constants in (2.1) are called the characteristics of φ;

(ii) *it holds that*

$$\mu_1 \leqslant \inf_{t>0} \frac{t\varphi'(t)}{\varphi(t)} \leqslant \sup_{t>0} \frac{t\varphi'(t)}{\varphi(t)} \leqslant \mu_2\,;$$

(iii) *the mappings*

$$t \in (0, +\infty) \to \frac{\varphi'(t)}{t^{\mu_1 - 1}}, \frac{\varphi(t)}{t^{\mu_1}} \quad and \quad t \in (0, +\infty) \to \frac{\varphi'(t)}{t^{\mu_2 - 1}}, \frac{\varphi(t)}{t^{\mu_2}}$$

are increasing and decreasing, respectively;

(iv) *as for the functions φ and φ' applied to multiples of given arguments, the following inequalities hold for every $t \geqslant 0$:*

$$a^{\mu_2}\varphi(t) \leqslant \varphi(at) \leqslant a^{\mu_1}\varphi(t) \quad and$$
$$a^{\mu_2 - 1}\varphi'(t) \leqslant \varphi'(at) \leqslant a^{\mu_1 - 1}\varphi'(t) \quad if \ \ 0 < a \leqslant 1\,;$$
$$a^{\mu_1}\varphi(t) \leqslant \varphi(at) \leqslant a^{\mu_2}\varphi(t) \quad and$$
$$a^{\mu_1 - 1}\varphi'(t) \leqslant \varphi'(at) \leqslant a^{\mu_2 - 1}\varphi'(t) \ if \ a \geqslant 1.$$

In particular, from (iv) it follows that both φ and φ^* satisfy the Δ_2-condition with constants $\Delta_2(\varphi)$ and $\Delta_2(\varphi^*)$ determined by μ_1 and μ_2. We will denote by $\Delta_2(\varphi, \varphi^*)$ constants depending on $\Delta_2(\varphi)$ and $\Delta_2(\varphi^*)$. Moreover, for $t > 0$ we have

$$\varphi(t) \sim \varphi'(t)\, t, \qquad \varphi(t) \sim \varphi''(t)\, t^2, \qquad \varphi^*\big(\varphi'(t)\big) \sim \varphi^*\big(\varphi(t)/t\big) \sim \varphi(t).$$

We recall also that the following inequalities hold for the inverse function φ^{-1}:

$$a^{\frac{1}{\mu_1}} \varphi^{-1}(t) \leqslant \varphi^{-1}(at) \leqslant a^{\frac{1}{\mu_2}} \varphi^{-1}(t) \tag{2.2}$$

for every $t \geqslant 0$ with $0 < a \leqslant 1$. The same result holds also for $a \geqslant 1$ by exchanging the role of μ_1 and μ_2.

For given φ we define the associated N-function ψ by

$$\psi'(t) := \sqrt{\varphi'(t)\, t}\,.$$

Notice that if φ satisfies assumption (2.1), then also φ^*, ψ and ψ^* satisfy this assumption.

Define $\mathbf{V} : \mathbb{R}^{N \times n} \to \mathbb{R}^{N \times n}$ in the following way:

$$\mathbf{V}(\mathbf{Q}) = \psi'(|\mathbf{Q}|) \frac{\mathbf{Q}}{|\mathbf{Q}|}. \tag{2.3}$$

It is easy to check that

$$|\mathbf{V}(\mathbf{Q})|^2 \sim \varphi(|\mathbf{Q}|),$$

uniformly in $\mathbf{Q} \in \mathbb{R}^{N \times n}$.

Another important set of tools are the shifted N-functions $\{\varphi_a\}_{a \geqslant 0}$ (see [**12**]). We define for $t \geqslant 0$

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

We have the following relations:

$$\varphi_a(t) \sim \varphi'_a(t)\, t\,;$$

$$\varphi_a(t) \sim \varphi''(a+t)t^2 \sim \frac{\varphi(a+t)}{(a+t)^2} t^2 \sim \frac{\varphi'(a+t)}{a+t} t^2, \tag{2.4}$$

$$\varphi(a+t) \sim [\varphi_a(t) + \varphi(a)]. \tag{2.5}$$

The families $\{\varphi_a\}_{a \geqslant 0}$ and $\{(\varphi_a)^*\}_{a \geqslant 0}$ satisfy the Δ_2-condition uniformly in $a \geqslant 0$. The connection between $\mathbf{V}$ and φ_a (see [**12**]) is the following:

$$|\mathbf{V}(\mathbf{P}) - \mathbf{V}(\mathbf{Q})|^2 \sim \varphi_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|), \tag{2.6}$$

uniformly in $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{N \times n}$. The following lemma (see [**14**, corollary 26]) deals with the *change of shift* for N-functions.

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

$$\varphi_{|\mathbf{a}|}(t) \leqslant c_\eta \varphi_{|\mathbf{b}|}(t) + \eta \varphi_{|\mathbf{a}|}(|\mathbf{a} - \mathbf{b}|). \tag{2.7}$$

We define the function $\mathbf{V}_a : \mathbb{R}^{N \times n} \to \mathbb{R}^{N \times n}$ for $a \geqslant 0$ by

$$\mathbf{V}_a(\mathbf{Q}) := \sqrt{\varphi_a'(|\mathbf{Q}|)|\mathbf{Q}|} \frac{\mathbf{Q}}{|\mathbf{Q}|}, \tag{2.8}$$

where φ_a is the shifted N-function of φ. Since $\varphi_0 = \varphi$, we retrieve in (2.8) the function $\mathbf{V}$ for $a = 0$. With the following lemma, we list some properties of functions $\mathbf{V}_a$ which will be useful in the sequel.

LEMMA 2.3. *Let $a \geqslant 0$ and $\mathbf{V}_a$ be as above. Then for any $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{N \times n}$ a Young-type inequality holds:*

$$\varphi_a'(|\mathbf{Q}|)|\mathbf{P}| \leqslant c(|\mathbf{V}_a(\mathbf{Q})|^2 + |\mathbf{V}_a(\mathbf{P})|^2), \tag{2.9}$$

where the constant c depends only on $\Delta_2(\varphi)$.

Let $\mathbf{P}_0, \mathbf{P}_1 \in \mathbb{R}^{N \times n}$, $\theta \in [0, 1]$ and define $\mathbf{P}_\theta := (1 - \theta)\mathbf{P}_0 + \theta\mathbf{P}_1$. Then the following result holds (see [**12**, lemma 20]).

LEMMA 2.4. *Let φ be a N-function with $\Delta_2(\varphi, \varphi^*) < \infty$. Then uniformly for all $\mathbf{P}_0, \mathbf{P}_1 \in \mathbb{R}^{N \times n}$ with $|\mathbf{P}_0| + |\mathbf{P}_1| > 0$ holds*

$$\int_0^1 \frac{\varphi'(|\mathbf{P}_\theta|)}{|\mathbf{P}_\theta|} \, \mathrm{d}\theta \sim \frac{\varphi'(|\mathbf{P}_0| + |\mathbf{P}_1|)}{|\mathbf{P}_0| + |\mathbf{P}_1|}$$

where the constants only depend on $\Delta_2(\varphi, \varphi^*)$.

In view of the previous considerations, the same proposition holds true for the shifted functions, uniformly in $a \geqslant 0$.

From assumption (F2) we can easily infer an upper bound for $f(x, \mathbf{u}, \mathbf{P}) - f(x, \mathbf{u}, \mathbf{Q})$, uniformly in $x \in \Omega$ and $\mathbf{u} \in \mathbb{R}^N$, for every $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{N \times n}$; namely,

$$\begin{aligned}
|f(x, \mathbf{u}, \mathbf{P}) - f(x, \mathbf{u}, \mathbf{Q})| &\leqslant |\mathbf{P} - \mathbf{Q}| \int_0^1 |Df(x, \mathbf{u}, \mathbf{P} + t(\mathbf{Q} - \mathbf{P}))| \, \mathrm{d}t \\
&\leqslant L|\mathbf{P} - \mathbf{Q}| \int_0^1 \varphi'(|\mathbf{P} + t(\mathbf{Q} - \mathbf{P})|) \, \mathrm{d}t \\
&\leqslant cL\varphi(|\mathbf{P}| + |\mathbf{Q}|).
\end{aligned} \tag{2.10}$$

The following estimate is a consequence of (F2) and lemma 2.4 (see [**17**, eq. (2.14)]):

$$\begin{aligned}
|Df(x, \mathbf{u}, \mathbf{P}) - Df(x, \mathbf{u}, \mathbf{Q})| &\leqslant c(\varphi, L)\varphi''(|\mathbf{P}| + |\mathbf{Q}|)|\mathbf{P} - \mathbf{Q}| \\
&\leqslant c(\varphi, L)\varphi'_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|) \\
&= c(\varphi, L)\frac{\varphi'(|\mathbf{P}| + |\mathbf{P} - \mathbf{Q}|)}{|\mathbf{P}| + |\mathbf{P} - \mathbf{Q}|}|\mathbf{P} - \mathbf{Q}|,
\end{aligned} \tag{2.11}$$

for every $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{N \times n}$.

The following version of Sobolev-Poincaré inequality can be found in [**12**, lemma 7].

THEOREM 2.5. *Let φ be an N-function with $\Delta_2(\varphi, \varphi^*) < +\infty$. Then there exist numbers $\alpha = \alpha(n, \Delta_2(\varphi, \varphi^*)) \in (0,1)$ and $K = K(n, N, \Delta_2(\varphi, \varphi^*)) > 0$ such that the following holds. If $B \subset \mathbb{R}^n$ is any ball with radius R and $\mathbf{w} \in W^{1,\varphi}(B, \mathbb{R}^N)$, then*

$$
\fint_B \varphi\left(\frac{|\mathbf{w} - (\mathbf{w})_B|}{R}\right) \mathrm{d}x \leqslant K \left(\fint_B \varphi^\alpha\left(|D\mathbf{w}|\right) \mathrm{d}x\right)^{\frac{1}{\alpha}}, \qquad (2.12)
$$

where $(\mathbf{w})_B := \fint_B \mathbf{w}(x)\, \mathrm{d}x$. Moreover, if $\mathbf{w} \in W_0^{1,\varphi}(B, \mathbb{R}^N)$, then

$$
\fint_B \varphi\left(\frac{|\mathbf{w}|}{R}\right) \mathrm{d}x \leqslant K \left(\fint_B \varphi^\alpha\left(|D\mathbf{w}|\right) \mathrm{d}x\right)^{\frac{1}{\alpha}},
$$

where K and α have the same dependencies as before.

2.2. Some useful lemmas

The following lemma, useful in order to re-absorb certain terms, is a variant of the classical [**25**, lemma 6.1] (see [**17**, lemma 3.1]).

LEMMA 2.6. *Let ψ be an N-function with $\psi \in \Delta_2$, let $\varrho > 0$ and $h \in L^\psi(B_\varrho(x_0))$. Let $g : [r, \varrho] \to \mathbb{R}$ be nonnegative and bounded such that for all $r \leqslant s < t \leqslant \varrho$*

$$
g(s) \leqslant \theta g(t) + A \int_{B_t(x_0)} \psi\left(\frac{|h(y)|}{t-s}\right) \mathrm{d}y + \frac{B}{(t-s)^\beta} + C,
$$

where $A, B, C \geqslant 0$, $\beta > 0$ and $\theta \in [0,1)$. Then

$$
g(r) \leqslant c(\theta, \Delta_2(\psi), \beta) \left[A \int_{B_\varrho(x_0)} \psi\left(\frac{|h(y)|}{\varrho - r}\right) \mathrm{d}y + \frac{B}{(\varrho - r)^\beta} + C\right].
$$

The following lemma is useful to derive reverse Hölder estimates. It is a variant of the results by Gehring [**24**] and Giaquinta-Modica [**25**, theorem 6.6].

LEMMA 2.7. *Let $B_0 \subset \mathbb{R}^n$ be a ball, $f \in L^1(B_0)$, and $g \in L^{\sigma_0}(B_0)$ for some $\sigma_0 > 1$. Assume that for some $\theta \in (0,1)$, $c_1 > 0$ and all balls B with $2B \subset B_0$*

$$
\fint_B |f|\, \mathrm{d}x \leqslant c_1 \left(\fint_{2B} |f|^\theta\, \mathrm{d}x\right)^{1/\theta} + \fint_{2B} |g|\, \mathrm{d}x.
$$

Then there exist $\sigma_1 > 1$ and $c_2 > 1$ such that $g \in L^{\sigma_1}_{loc}(B)$ and for all $\sigma_2 \in [1, \sigma_1]$

$$\left(\fint_B |f|^{\sigma_2} \, dx \right)^{1/\sigma_2} \leqslant c_2 \fint_{2B} |f| \, dx + c_2 \left(\fint_{2B} |g|^{\sigma_2} \, dx \right)^{1/\sigma_2}.$$

2.3. $\mathcal{A}$-harmonic and φ-harmonic functions

Let $\mathcal{A}$ be a bilinear form on $\mathbb{R}^{N \times n}$. We say that $\mathcal{A}$ is *strongly elliptic in the sense of Legendre-Hadamard* if for all $\boldsymbol{\xi} \in \mathbb{R}^N$, $\boldsymbol{\zeta} \in \mathbb{R}^n$ it holds that

$$\kappa_{\mathcal{A}} |\boldsymbol{\xi}|^2 |\boldsymbol{\zeta}|^2 \leqslant \langle \mathcal{A}(\boldsymbol{\xi} \otimes \boldsymbol{\zeta}) | (\boldsymbol{\xi} \otimes \boldsymbol{\zeta}) \rangle \leqslant L_{\mathcal{A}} |\boldsymbol{\xi}|^2 |\boldsymbol{\zeta}|^2 \tag{2.13}$$

for some $L_{\mathcal{A}} \geqslant \kappa_{\mathcal{A}} > 0$. We say that a Sobolev function $\mathbf{w}$ on a ball $B_\varrho(x_0)$ is *$\mathcal{A}$-harmonic* on $B_\varrho(x_0)$ if it satisfies $-\mathrm{div}(\mathcal{A}D\mathbf{w}) = 0$ in the sense of distributions; i.e.,

$$\int_{B_\varrho(x_0)} \langle \mathcal{A}D\mathbf{w} | D\boldsymbol{\psi} \rangle \, dx = 0, \quad \text{for all } \boldsymbol{\psi} \in C_0^\infty(B_\varrho(x_0), \mathbb{R}^N).$$

It is well known from the classical theory (see, e.g. [**25**, chapter 10]) that $\mathbf{w}$ is smooth in the interior of $B_\varrho(x_0)$, and it satisfies the estimate

$$\sup_{B_{\varrho/2}(x_0)} |D\mathbf{w}|^2 + \varrho^2 \sup_{B_{\varrho/2}(x_0)} |D^2\mathbf{w}|^2 \leqslant c(n, N, \nu, L) \fint_{B_\varrho(x_0)} |D\mathbf{w}|^2 \, dx. \tag{2.14}$$

Let φ be an Orlicz function. We say that a map $\mathbf{w} \in W^{1,\varphi}(B_\varrho(x_0), \mathbb{R}^N)$ is *φ-harmonic* on $B_\varrho(x_0)$ (see [**16**]) if and only if

$$\int_{B_\varrho(x_0)} \left\langle \frac{\varphi'(|D\mathbf{w}|)}{|D\mathbf{w}|} D\mathbf{w} \middle| D\boldsymbol{\psi} \right\rangle \, dx = 0, \quad \text{for all } \boldsymbol{\psi} \in C_0^\infty(B_\varrho(x_0), \mathbb{R}^N).$$

More precisely, $D\mathbf{w}$ and $\mathbf{V}(D\mathbf{w})$ are Hölder continuous due to the following decay estimate, see [**15**].

PROPOSITION 2.8. *Let φ be a convex function complying with $(\varphi 1)$, $(\varphi 2)$ and*

$(\varphi 3)$

> *φ' is Hölder continuous off the diagonal:*
>
> $$|\varphi''(s + t) - \varphi''(t)| \leqslant c_0 \, \varphi''(t) \left(\frac{|s|}{t} \right)^{\beta_0}, \quad \beta_0 > 0,$$
>
> *for all $t > 0$ and $s \in \mathbb{R}$ with $|s| < \frac{1}{2}t$.*

Then there exist a constant $c \geqslant 1$ and an exponent $\gamma_0 \in (0, 1)$ depending only on n, N and the characteristics of φ, such that the following statement holds true:

whenever $\mathbf{w} \in W^{1,\varphi}(B_R(x_0), \mathbb{R}^N)$ *is a weak solution of the system*

$$\mathrm{div}\left(\frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|}\,D\mathbf{u}\right) = 0 \qquad in\ B_R(x_0),$$

then for every $\tau \in (0,1)$ *there hold*

$$\sup_{B_{\tau R/2}(x_0)} \varphi(|D\mathbf{w}|) \leqslant c \fint_{B_{\tau R}(x_0)} \varphi(|D\mathbf{w}|)\,\mathrm{d}x,$$

$$\fint_{B_{\tau R}(x_0)} |\mathbf{V}(D\mathbf{w}) - (\mathbf{V}(D\mathbf{w}))_{x_0,\tau R}|^2\,\mathrm{d}x$$

$$\leqslant c\tau^{2\gamma_0} \fint_{B_R(x_0)} |\mathbf{V}(D\mathbf{w}) - (\mathbf{V}(D\mathbf{w}))_{x_0,R}|^2\,\mathrm{d}x.$$

This result can be viewed as the Orlicz version of the milestone theorem of Uhlenbeck [**41**] for differential forms solving a p-harmonic system, see also [**2**].

2.4. Harmonic type approximation results

We recall here two different harmonic type approximation results. The first one is the $\mathcal{A}$-*harmonic approximation*: given a Sobolev function $\mathbf{u}$ on a ball B, we want to find an $\mathcal{A}$-harmonic function $\mathbf{w}$ which is 'close' to the function $\mathbf{u}$. It will be the $\mathcal{A}$-harmonic function with the same boundary values as $\mathbf{u}$; i.e., a Sobolev function $\mathbf{w}$ which satisfies

$$\begin{cases} -\mathrm{div}(\mathcal{A}D\mathbf{w}) = 0 & \text{on } B \\ \mathbf{w} = \mathbf{u} & \text{on } \partial B \end{cases} \tag{2.15}$$

in the sense of distributions.

Setting $\mathbf{z} := \mathbf{w} - \mathbf{u}$, then (2.15) is equivalent to finding a Sobolev function $\mathbf{z}$ which satisfies

$$\begin{cases} -\mathrm{div}(\mathcal{A}D\mathbf{z}) = -\mathrm{div}(\mathcal{A}D\mathbf{u}) & \text{on } B \\ \mathbf{z} = \mathbf{0} & \text{on } \partial B \end{cases} \tag{2.16}$$

in the sense of distributions.

The following $\mathcal{A}$-harmonic approximation result in the setting of Orlicz spaces has been proved in [**17**, theorem 14].

THEOREM 2.9. *Let* $B \subset\subset \Omega$ *be a ball with radius* r_B *and let* $\widetilde{B} \subset \Omega$ *denote either* B *or* $2B$. *Let* $\mathcal{A}$ *be a strongly elliptic (in the sense of Legendre-Hadamard) bilinear form on* $\mathbb{R}^{N \times n}$. *Let* ψ *be an* N-*function with* $\psi \in \Delta_2(\psi, \psi^*)$ *and let* $s > 1$. *Then for every* $\varepsilon > 0$, *there exists* $\delta > 0$ *only depending on* n, N, κ_A, $|\mathcal{A}|$, $\Delta_2(\psi, \psi^*)$ *and* $s > 1$ *such that the following holds. Let* $\mathbf{u} \in W^{1,\psi}(\widetilde{B}, \mathbb{R}^N)$ *be almost* $\mathcal{A}$-*harmonic*

on B in the sense that

$$\left| \fint_B \langle \mathcal{A}D\mathbf{u}|D\boldsymbol{\eta}\rangle \, \mathrm{d}x \right| \leqslant \delta \fint_{\tilde{B}} |D\mathbf{u}| \, \mathrm{d}x \|D\boldsymbol{\eta}\|_{L^\infty(B)} \tag{2.17}$$

for all $\boldsymbol{\eta} \in C_0^\infty(B, \mathbb{R}^N)$. Then the unique solution $\mathbf{z} \in W_0^{1,\psi}(B, \mathbb{R}^N)$ of (2.16) satisfies

$$\fint_B \psi\left(\frac{|\mathbf{z}|}{r_B}\right) \mathrm{d}x + \fint_B \psi(|D\mathbf{z}|) \, \mathrm{d}x \leqslant \varepsilon \left(\left(\fint_{\tilde{B}} (\psi(|D\mathbf{u}|))^s \, \mathrm{d}x \right)^{\frac{1}{s}} + \fint_{\tilde{B}} \psi(|D\mathbf{u}|) \, \mathrm{d}x \right).$$

REMARK 2.10. We will exploit the previous approximation result in a slightly modified version. Indeed, following [**7**, lemma 2.7], under the additional assumption

$$\fint_{\tilde{B}} \psi(|D\mathbf{u}|) \, \mathrm{d}x \leqslant \left(\fint_{\tilde{B}} [\psi(|D\mathbf{u}|)]^s \, \mathrm{d}x \right)^{\frac{1}{s}} \leqslant \psi(\mu)$$

for some exponent $s > 1$ and for a constant $\mu > 0$, and (2.17) replaced by

$$\left| \fint_B \langle \mathcal{A}D\mathbf{u}|D\boldsymbol{\eta}\rangle \, \mathrm{d}x \right| \leqslant \delta\mu \|D\boldsymbol{\eta}\|_{L^\infty(B)},$$

it can be seen with minor changes in the proof that the unique solution $\mathbf{z} \in W_0^{1,\psi}(B, \mathbb{R}^N)$ of (2.16) satisfies

$$\fint_B \psi\left(\frac{|\mathbf{z}|}{r_B}\right) dx + \fint_B \psi(|D\mathbf{z}|) \, dx \leqslant \varepsilon\psi(\mu).$$

Now, moving on to φ-harmonic functions, the following *φ-harmonic approximation* lemma ([**16**, lemma 1.1]) is the extension to general convex functions of the p-harmonic approximation lemma [**21**], [**22**, lemma 1], and allows to approximate 'almost φ-harmonic' functions by φ-harmonic functions.

LEMMA 2.11. *Let φ satisfy assumption (2.1). For every $\varepsilon > 0$ and $\theta \in (0,1)$ there exists $\delta > 0$ which only depends on ε, θ, and the characteristics of φ such that the following holds. Let $B \subset \mathbb{R}^n$ be a ball and let $\tilde{B}$ denote either B or $2B$. If $\mathbf{u} \in W^{1,\varphi}(\tilde{B}, \mathbb{R}^N)$ is almost φ-harmonic on a ball $B \subset \mathbb{R}^n$ in the sense that*

$$\fint_B \left\langle \frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|} D\mathbf{u} \middle| D\boldsymbol{\eta} \right\rangle \, \mathrm{d}x \leqslant \delta \left(\fint_{\tilde{B}} \varphi(|D\mathbf{u}|) \, \mathrm{d}x + \varphi(\|D\boldsymbol{\eta}\|_\infty) \right) \tag{2.18}$$

for all $\boldsymbol{\eta} \in C_0^\infty(B, \mathbb{R}^N)$, then the unique φ-harmonic $\mathbf{w} \in W^{1,\varphi}(B, \mathbb{R}^N)$ with $\mathbf{w} = \mathbf{u}$ on ∂B satisfies

$$\left(\fint_B |\mathbf{V}(D\mathbf{u}) - \mathbf{V}(D\mathbf{w})|^{2\theta} \, \mathrm{d}x \right)^{\frac{1}{\theta}} \leqslant \varepsilon \fint_{\tilde{B}} \varphi(|D\mathbf{u}|) \, \mathrm{d}x, \qquad (2.19)$$

where $\mathbf{V}$ is as in (2.3).

The estimate (2.19) can be improved when $\varphi(|D\mathbf{u}|)$ satisfies a reverse Hölder inequality as follows (see [**7**, corollary 2.10]).

LEMMA 2.12. *Let $B \subset \mathbb{R}^n$ be a ball. Let $\mathbf{u} \in W^{1,\varphi}(2B, \mathbb{R}^N)$ be such that*

$$\left(\fint_B \varphi^{s_1}(|D\mathbf{u}|) \, \mathrm{d}x \right)^{\frac{1}{s_1}} \leqslant \tilde{c}_0 \fint_{2B} \varphi(|D\mathbf{u}|) \, \mathrm{d}x$$

for $s_1 > 1$ and $\tilde{c}_0 > 0$. Then for every $\varepsilon \in (0, 1)$ there exists $\delta_0 = \delta_0(n, N, \mu_1, \mu_2, s_1, \tilde{c}_0, \varepsilon) > 0$ such that the following holds: if $\mathbf{u}$ is almost φ-harmonic as in (2.18) *with δ_0 in place of δ, then the unique φ-harmonic function $\mathbf{w} \in W^{1,\varphi}(B, \mathbb{R}^N)$ such that $\mathbf{w} = \mathbf{u}$ on ∂B satisfies*

$$\fint_B |\mathbf{V}(D\mathbf{u}) - \mathbf{V}(D\mathbf{w})|^2 \, \mathrm{d}x \leqslant \varepsilon \fint_{2B} \varphi(|D\mathbf{u}|) \, \mathrm{d}x.$$

3. Partial regularity for functionals

3.1. Caccioppoli inequalities and higher integrability results

As usual, the first step in proving a regularity theorem for the minimizers of integral functionals is to establish suitable Caccioppoli-type inequalities.

First, we state a 'zero order' Caccioppoli inequality. The proof is an adaptation to the φ-setting of [**4**, lemma 3.1], we then omit the details (see also [**7**, theorem 2.4]).

LEMMA 3.1. *Let $\mathbf{u} \in W^{1,\varphi}(\Omega, \mathbb{R}^N)$ be a minimizer of the functional* (1.1), *under assumptions* (F1)–(F2). *Then, for every $\mathbf{u}_0 \in \mathbb{R}^N$ and $x_0 \in \Omega$ and all $0 < \varrho < \mathrm{dist}(x_0, \partial\Omega)$ and $r \in [\varrho/2, \varrho)$ there holds*

$$\fint_{B_r(x_0)} \varphi(|D\mathbf{u}|) \, \mathrm{d}x \leqslant c \fint_{B_\varrho(x_0)} \varphi\left(\frac{|\mathbf{u} - \mathbf{u}_0|}{\varrho - r} \right) \mathrm{d}x$$

for some constant $c = c(\varphi, L, \nu) > 0$.

From lemma 3.1 together with the Sobolev-Poincaré inequality (theorem 2.5) and Gehring's Lemma (lemma 2.7), one can infer in a standard way the following higher integrability result (see, e.g., [**7**, theorem 2.5]).

LEMMA 3.2. *There exist an exponent $s_0 = s_0(n, N, \varphi, L, \nu) > 1$ and a constant c depending only on n, N, φ, L, ν such that, if $\mathbf{u} \in W^{1,\varphi}(\Omega; \mathbb{R}^N)$ is a minimizer of the functional (1.1), complying with (F1)–(F2), then the following holds: for every $s \in (1, s_0]$, for any $x_0 \in \Omega$, any radius $0 < \varrho < \mathrm{dist}(x_0, \partial\Omega)$ and $r \in [\varrho/2, \varrho)$, one has*

$$\fint_{B_r(x_0)} \varphi^s(|D\mathbf{u}|) \, \mathrm{d}x \leqslant c \left(\frac{\varrho}{\varrho - r} \right)^{n(s-1)} \left(\fint_{B_\varrho(x_0)} \varphi(|D\mathbf{u}|) \, \mathrm{d}x \right)^s.$$

Another useful tool will be the following global higher integrability result on balls for minimizers of (1.1), which has been proven in the Orlicz setting for more general integrands in [**10**, lemma 4.3].

LEMMA 3.3. *Let $\mathbf{u} \in W^{1,\varphi}(B_r(x_0), \mathbb{R}^N)$ be such that $\varphi(|D\mathbf{u}|) \in L^{s_0}(B_r(x_0), \mathbb{R}^N)$ for some $s_0 > 1$. Then there exists an exponent $s = s(n, N, \varphi, L, \nu, s_0) \in (1, s_0]$ and a constant $c = c(n, N, \varphi, L, \nu)$ such that, if $\mathbf{v} \in \mathbf{u} + W_0^{1,\varphi}(B_r(x_0), \mathbb{R}^N)$ is a minimizer of the functional $\mathcal{G}[\mathbf{v}] := \int_{B_r(x_0)} g(D\mathbf{v}) \, \mathrm{d}x$ with a C^1-integrand $g : \mathbb{R}^{N \times n} \to \mathbb{R}$ complying with the growth assumptions*

$$\nu\varphi(|\boldsymbol{\xi}|) \leqslant g(\boldsymbol{\xi}) \leqslant L\varphi(1 + |\boldsymbol{\xi}|) \quad and \quad |Dg(\boldsymbol{\xi})| \leqslant L\varphi'(|\boldsymbol{\xi}|)$$

for all $\boldsymbol{\xi} \in \mathbb{R}^{nN}$, then we have $\varphi(|D\mathbf{v}|) \in L^s(B_r(x_0), \mathbb{R}^N)$ and

$$\left(\fint_{B_r(x_0)} \varphi^s(|D\mathbf{v}|) \, \mathrm{d}x \right)^{\frac{1}{s}} \leqslant c \left(\fint_{B_r(x_0)} \varphi^{s_0}(|D\mathbf{u}|) \, \mathrm{d}x \right)^{\frac{1}{s_0}}.$$

We have the following Caccioppoli inequality of second type for local minimizers of (1.1), involving affine functions.

LEMMA 3.4. *There exists a constant $c = c(n, N, \Delta_2(\varphi), \nu, L) > 0$ such that, if $\mathbf{u} \in W^{1,\varphi}(\Omega; \mathbb{R}^N)$ is a minimizer of the functional (1.1) under assumptions (F1)–(F7), and $\boldsymbol{\ell} : \mathbb{R}^n \to \mathbb{R}^N$ is an affine function, say $\boldsymbol{\ell}(x) := \mathbf{u}_0 + \mathbf{Q}(x - x_0)$ for some $\mathbf{u}_0 \in \mathbb{R}^N$ and $\mathbf{Q} \in \mathbb{R}^{N \times n}$, then for any ball $B_\varrho(x_0) \subseteq \Omega$ with $\varrho \leqslant \varrho_0$ there holds*

$$\fint_{B_{\varrho/2}(x_0)} \varphi_{|\mathbf{Q}|}(|D\mathbf{u} - \mathbf{Q}|) \, \mathrm{d}x$$

$$\leqslant c \fint_{B_\varrho(x_0)} \varphi_{|\mathbf{Q}|} \left(\frac{|\mathbf{u} - \boldsymbol{\ell}|}{\varrho} \right) \, \mathrm{d}x$$

$$+ c\varphi(|\mathbf{Q}|) \left[\omega \left(\fint_{B_\varrho(x_0)} |\mathbf{u} - \mathbf{u}_0| + |\mathbf{u} - \boldsymbol{\ell}| \, \mathrm{d}x \right)^{1 - \frac{1}{s}} + [\mathcal{V}(\varrho)]^{1 - \frac{1}{s}} \right]$$

for every $s \in (1, s_0]$ where s_0 is that of lemma 3.2.

Proof. We follow the argument of [**4**, lemma 3.5] for functionals with p-growth, just mentioning how to obtain the analogous main estimates therein. We assume, without loss of generality, that $x_0 = 0$. For radii $\frac{\varrho}{2} \leqslant r < \tau < t \leqslant \frac{3\varrho}{4}$ with $\tau := \frac{r+t}{2}$ we consider a cut-off function $\eta \in C_0^\infty(B_\tau; [0,1])$ such that $\eta \equiv 1$ on B_r and $|D\eta| \leqslant \frac{4}{t-r}$ on B_τ. Correspondingly, we define the functions $\boldsymbol{\xi} := \eta(\mathbf{u} - \boldsymbol{\ell}) \in W^{1,\varphi}(B_\tau; \mathbb{R}^N)$ and $\boldsymbol{\psi} := (1 - \eta)(\mathbf{u} - \boldsymbol{\ell}) \in W^{1,\varphi}(B_\tau; \mathbb{R}^N)$. Note that $\boldsymbol{\ell} + \boldsymbol{\xi} = \mathbf{u} - \boldsymbol{\psi}$. From the quasi-convexity assumption (F3), (2.4) and simple manipulations we obtain

$$
\begin{aligned}
\int_{B_\tau} \varphi_{|\mathbf{Q}|}(|D\boldsymbol{\xi}|) \, \mathrm{d}x &\leqslant c(\nu, \varphi) \int_{B_\tau} \varphi''(|\mathbf{Q}| + |D\boldsymbol{\xi}|)|D\boldsymbol{\xi}|^2 \, \mathrm{d}x \\
&\leqslant c \int_{B_\tau} [(f(\cdot, \mathbf{u}_0, \mathbf{Q} + D\boldsymbol{\xi}(x)))_\tau - (f(\cdot, \mathbf{u}_0, \mathbf{Q}))_\tau] \, \mathrm{d}x \\
&= c(J_1 + J_2 + J_3 + J_4 + J_5 + J_6 + J_7),
\end{aligned}
\tag{3.1}
$$

where

$$
J_1 := \int_{B_\tau} [(f(\cdot, \mathbf{u}_0, D\mathbf{u}(x) - D\boldsymbol{\psi}(x)))_\tau - (f(\cdot, \mathbf{u}_0, D\mathbf{u}(x)))_\tau] \, \mathrm{d}x,
$$

$$
J_2 := \int_{B_\tau} [(f(\cdot, \mathbf{u}_0, D\mathbf{u}(x)))_\tau - (f(\cdot, \mathbf{u}(x), D\mathbf{u}(x)))_\tau] \, \mathrm{d}x,
$$

$$
J_3 := \int_{B_\tau} [(f(\cdot, \mathbf{u}(x), D\mathbf{u}(x)))_\tau - f(x, \mathbf{u}(x), D\mathbf{u}(x))] \, \mathrm{d}x,
$$

$$
J_4 := \int_{B_\tau} [f(x, \mathbf{u}(x), D\mathbf{u}(x)) - f(x, \mathbf{u}(x) - \boldsymbol{\xi}(x), D\mathbf{u}(x) - D\boldsymbol{\xi}(x))] \, \mathrm{d}x,
$$

$$
J_5 := \int_{B_\tau} [f(x, \mathbf{u}(x) - \boldsymbol{\xi}(x), \mathbf{Q} + D\boldsymbol{\psi}(x)) - f(x, \mathbf{u}_0, \mathbf{Q} + D\boldsymbol{\psi}(x))] \, \mathrm{d}x,
$$

$$
J_6 := \int_{B_\tau} [f(x, \mathbf{u}_0, \mathbf{Q} + D\boldsymbol{\psi}(x)) - (f(\cdot, \mathbf{u}_0, \mathbf{Q} + D\boldsymbol{\psi}(x)))_\tau] \, \mathrm{d}x,
$$

$$
J_7 := \int_{B_\tau} [(f(\cdot, \mathbf{u}_0, \mathbf{Q} + D\boldsymbol{\psi}(x)))_\tau - (f(\cdot, \mathbf{u}_0, \mathbf{Q}))_\tau] \, \mathrm{d}x.
$$

Now, we proceed to estimate each term above separately. From the minimizing property of $\mathbf{u}$ we infer that $J_4 \leqslant 0$, and by assumptions (F5) and (F4) we obtain the estimates

$$
J_2 \leqslant \int_{B_\tau} \omega(|\mathbf{u} - \mathbf{u}_0|)\varphi(|D\mathbf{u}|) \, \mathrm{d}x,
$$

$$
J_3 \leqslant \int_{B_\tau} v_0(\cdot, \tau)\varphi(|D\mathbf{u}|) \, \mathrm{d}x,
$$

respectively. Again by exploiting property (F5), the monotonicity of ω and φ, and the fact that

$$
|\boldsymbol{\xi}| \leqslant |\mathbf{u} - \boldsymbol{\ell}|, \quad \text{and} \quad |D\boldsymbol{\psi}| \leqslant |D\mathbf{u} - \mathbf{Q}| + 4\left|\frac{\mathbf{u} - \boldsymbol{\ell}}{t - \tau}\right|,
$$

we can estimate J_5 as

$$J_5 \leqslant c(\varphi) \int_{B_\tau} \omega(|\mathbf{u} - \mathbf{u}_0| + |\mathbf{u} - \boldsymbol{\ell}|)\varphi\left(|\mathbf{Q}| + |D\mathbf{u}| + \left|\frac{\mathbf{u} - \boldsymbol{\ell}}{t - \tau}\right|\right) \, \mathrm{d}x$$

$$\leqslant c(\varphi) \int_{B_\tau} \omega(|\mathbf{u} - \mathbf{u}_0| + |\mathbf{u} - \boldsymbol{\ell}|)\left[\varphi(|\mathbf{Q}| + |D\mathbf{u}|) + \varphi\left(\left|\frac{\mathbf{u} - \boldsymbol{\ell}}{t - \tau}\right|\right)\right] \, \mathrm{d}x,$$

whence, taking into account that by virtue of (2.5),

$$\varphi\left(\left|\frac{\mathbf{u} - \boldsymbol{\ell}}{t - \tau}\right|\right) \leqslant c\varphi_{|\mathbf{Q}|}\left(\left|\frac{\mathbf{u} - \boldsymbol{\ell}}{t - \tau}\right|\right) + c\varphi(|\mathbf{Q}|)$$

$$\leqslant c(\varphi)\left|\mathbf{V}_{|\mathbf{Q}|}\left(\left|\frac{\mathbf{u} - \boldsymbol{\ell}}{t - \tau}\right|\right)\right|^2 + c\varphi(|\mathbf{Q}|)$$

(3.2)

and recalling that $\omega \leqslant 1$, we get

$$J_5 \leqslant c(\varphi)\left(\int_{B_\tau}\left|\mathbf{V}_{|\mathbf{Q}|}\left(\left|\frac{\mathbf{u} - \boldsymbol{\ell}}{t - \tau}\right|\right)\right|^2 \, \mathrm{d}x\right.$$

$$\left. + \int_{B_\tau} \omega(|\mathbf{u} - \mathbf{u}_0| + |\mathbf{u} - \boldsymbol{\ell}|)\varphi\left(|\mathbf{Q}| + |D\mathbf{u}|\right) \, \mathrm{d}x\right).$$

For what concerns J_6, an analogous computation as for the estimate of J_5 based on (3.2) and the VMO assumption (F4) gives

$$J_6 \leqslant \int_{B_\tau} v_0(\cdot, \tau)\varphi(|\mathbf{Q} + D\boldsymbol{\psi}|) \, \mathrm{d}x$$

$$\leqslant c(\Delta_2(\varphi))\left(\int_{B_\tau}\left|\mathbf{V}_{|\mathbf{Q}|}\left(\left|\frac{\mathbf{u} - \boldsymbol{\ell}}{t - \tau}\right|\right)\right|^2 \, \mathrm{d}x + \int_{B_\tau} v_0(\cdot, \tau)\varphi\left(|\mathbf{Q}| + |D\mathbf{u}|\right) \, \mathrm{d}x\right).$$

The terms J_1 and J_7 can be combined together as

$$J_7 + J_1 = \int_{B_\tau} \fint_{B_\tau} \int_0^1 \langle Df(y, \mathbf{u}_0, \mathbf{Q} + \theta D\boldsymbol{\psi}(x)) - Df(y, \mathbf{u}_0, \mathbf{Q})|D\boldsymbol{\psi}(x)\rangle \, \mathrm{d}\theta \, \mathrm{d}y \, \mathrm{d}x$$

$$+ \int_{B_\tau} \fint_{B_\tau} \int_0^1 \langle Df(y, \mathbf{u}_0, \mathbf{Q}) - Df(y, \mathbf{u}_0, D\mathbf{u}(x)$$

$$- (1 - \theta)D\boldsymbol{\psi}(x))|D\boldsymbol{\psi}(x)\rangle \, \mathrm{d}\theta \, \mathrm{d}y \, \mathrm{d}x$$

$$=: J_7' + J_1'.$$

From the Cauchy-Schwarz inequality, (2.11) and the fact that $D\psi = \mathbf{0}$ on B_r we infer

$$
\begin{aligned}
J_7' &\leqslant \int_{B_\tau} \fint_{B_\tau} \int_0^1 |Df(y, \mathbf{u}_0, \mathbf{Q} + \theta D\psi(x)) - Df(y, \mathbf{u}_0, \mathbf{Q})||D\psi(x)| \, \mathrm{d}\theta \, \mathrm{d}y \mathrm{d}x \\
&\leqslant c \int_{B_\tau} \int_0^1 \varphi'_{|\mathbf{Q}|}(\theta|D\psi(x)|)|D\psi(x)| \, \mathrm{d}\theta \, \mathrm{d}x \\
&\leqslant c \int_{B_\tau} \varphi_{|\mathbf{Q}|}(|D\psi(x)|) \, \mathrm{d}x \leqslant c \int_{B_\tau \setminus B_r} |\mathbf{V}_{|\mathbf{Q}|}(D\psi(x))|^2 \, \mathrm{d}x.
\end{aligned}
$$

We can estimate J_1' analogously, by recalling that $D\mathbf{u} - (1 - \theta)D\psi = \mathbf{Q} + D\boldsymbol{\xi} + \theta D\psi$, $D\psi = \mathbf{0}$ on B_r and applying the triangle inequality for $\varphi'_{|\mathbf{Q}|}$, (2.11) and the Young's inequality (2.9). In this way we get

$$
\begin{aligned}
J_1' &\leqslant \int_{B_\tau} \fint_{B_\tau} \int_0^1 |Df(y, \mathbf{u}_0, \mathbf{Q}) - Df(y, \mathbf{u}_0, D\mathbf{u}(x) \\
&\qquad - (1 - \theta)D\psi(x))||D\psi(x)| \, \mathrm{d}\theta \, \mathrm{d}y \, \mathrm{d}x \\
&\leqslant c \int_{B_\tau} \int_0^1 \varphi'_{|\mathbf{Q}|}(|D\boldsymbol{\xi} + \theta D\psi|)|D\psi| \, \mathrm{d}\theta \, \mathrm{d}x \\
&\leqslant c \int_{B_\tau} \varphi'_{|\mathbf{Q}|}(|D\psi|)|D\psi| \, \mathrm{d}x + c \int_{B_\tau} \varphi'_{|\mathbf{Q}|}(|D\boldsymbol{\xi}|)|D\psi| \, \mathrm{d}x \\
&\leqslant c \int_{B_\tau \setminus B_r} (|\mathbf{V}_{|\mathbf{Q}|}(D\psi)|^2 + |\mathbf{V}_{|\mathbf{Q}|}(D\boldsymbol{\xi})|^2) \, \mathrm{d}x.
\end{aligned}
$$

Recalling the definitions of $\boldsymbol{\xi}$ and ψ, by a simple computation we find that

$$
\begin{aligned}
D\psi &= (1 - \eta)(D\mathbf{u} - \mathbf{Q}) - D\eta \otimes (\mathbf{u} - \boldsymbol{\ell}), \\
D\boldsymbol{\xi} &= \eta(D\mathbf{u} - \mathbf{Q}) + D\eta \otimes (\mathbf{u} - \boldsymbol{\ell}),
\end{aligned}
$$

whence

$$
\begin{aligned}
\int_{B_\tau \setminus B_r} &(\varphi_{|\mathbf{Q}|}(|D\psi|) + \varphi_{|\mathbf{Q}|}(|D\boldsymbol{\xi}|)) \, \mathrm{d}x \\
&\leqslant c \int_{B_\tau \setminus B_r} \varphi_{|\mathbf{Q}|}(|D\mathbf{u} - \mathbf{Q}|) \, \mathrm{d}x + c \int_{B_\tau} \varphi_{|\mathbf{Q}|}\left(\frac{\mathbf{u} - \boldsymbol{\ell}}{t - r}\right) \, \mathrm{d}x,
\end{aligned}
$$

so that combining with the previous estimates we get

$$
J_1 + J_7 \leqslant c \left(\int_{B_\tau \setminus B_r} |\mathbf{V}_{|\mathbf{Q}|}(D\mathbf{u} - \mathbf{Q})|^2 \, \mathrm{d}x + \int_{B_\tau} \left| \mathbf{V}_{|\mathbf{Q}|}\left(\frac{\mathbf{u} - \boldsymbol{\ell}}{t - r}\right) \right|^2 \, \mathrm{d}x \right).
$$

Since $\boldsymbol{\xi} = \mathbf{u} - \boldsymbol{\ell}$ on B_r and $\tau \leqslant \varrho$, from (3.1) and the estimates for $J_1 - J_7$ we obtain

$$
\fint_{B_r} |\mathbf{V}_{|\mathbf{Q}|}(D\mathbf{u} - \mathbf{Q})|^2 \, \mathrm{d}x
$$
$$
\leqslant \tilde{c} \left(\fint_{B_\tau \backslash B_r} |\mathbf{V}_{|\mathbf{Q}|}(D\mathbf{u} - \mathbf{Q})|^2 \, \mathrm{d}x + \fint_{B_\varrho} \left| \mathbf{V}_{|\mathbf{Q}|} \left(\frac{\mathbf{u} - \boldsymbol{\ell}}{t - r} \right) \right|^2 \, \mathrm{d}x \right)
$$
$$
+ \tilde{c} \fint_{B_\tau} \left(\omega(|\mathbf{u} - \mathbf{u}_0| + |\mathbf{u} - \boldsymbol{\ell}|) + v_0(\cdot, \tau) \right) \varphi(|\mathbf{Q}| + |D\mathbf{u}|) \, \mathrm{d}x.
$$

Now, in a standard way we 'fill the hole' thus obtaining

$$
\fint_{B_r} |\mathbf{V}_{|\mathbf{Q}|}(D\mathbf{u} - \mathbf{Q})|^2 \, \mathrm{d}x
$$
$$
\leqslant \sigma \fint_{B_\tau} |\mathbf{V}_{|\mathbf{Q}|}(D\mathbf{u} - \mathbf{Q})|^2 \, \mathrm{d}x + \fint_{B_\varrho} \left| \mathbf{V}_{|\mathbf{Q}|} \left(\frac{\mathbf{u} - \boldsymbol{\ell}}{t - r} \right) \right|^2 \, \mathrm{d}x
$$
$$
+ \fint_{B_\tau} \left(\omega(|\mathbf{u} - \mathbf{u}_0| + |\mathbf{u} - \boldsymbol{\ell}|) + v_0(\cdot, \tau) \right) \varphi(|\mathbf{Q}| + |D\mathbf{u}|) \, \mathrm{d}x, \tag{3.3}
$$

where $\sigma := \frac{\tilde{c}}{\tilde{c}+1} < 1$. In order to bound the latter term further, we exploit the higher integrability result of lemma 3.2. Thus, with fixed $s \in (1, s_0]$, as a consequence of Hölder's inequality, the concavity of ω, the bounds $\omega \leqslant 1$ and $v_0 \leqslant 2L$, and Jensen's inequality also we obtain

$$
\fint_{B_\tau} \left(\omega(|\mathbf{u} - \mathbf{u}_0| + |\mathbf{u} - \boldsymbol{\ell}|) + v_0(\cdot, \tau) \right) \varphi(|\mathbf{Q}| + |D\mathbf{u}|) \, \mathrm{d}x
$$
$$
\leqslant c|B_\tau| \left(\fint_{B_\varrho} \omega(|\mathbf{u} - \mathbf{u}_0| + |\mathbf{u} - \boldsymbol{\ell}|)^{\frac{s}{s-1}} \, \mathrm{d}x + \fint_{B_\tau} v_0(\cdot, \tau)^{\frac{s}{s-1}} \, \mathrm{d}x \right)^{1 - \frac{1}{s}}
$$
$$
\times \left(\fint_{B_\tau} \varphi^s(|\mathbf{Q}|) + \varphi^s(|D\mathbf{u}|) \, \mathrm{d}x \right)^{\frac{1}{s}}
$$
$$
\leqslant c\tau^n \left(\frac{t}{t - r} \right)^{n(s-1)} \left[\omega \left(\fint_{B_\varrho} |\mathbf{u} - \mathbf{u}_0| + |\mathbf{u} - \boldsymbol{\ell}| \, \mathrm{d}x \right)^{1 - \frac{1}{s}} + \mathcal{V}(\tau)^{1 - \frac{1}{s}} \right]
$$
$$
\times \fint_{B_t} \varphi(|\mathbf{Q}|) + \varphi(|D\mathbf{u}|) \, \mathrm{d}x
$$

$$\leqslant c \left(\frac{\varrho}{t-r} \right)^{n(s-1)} \left[\omega \left(\fint_{B_\varrho} |\mathbf{u} - \mathbf{u}_0| + |\mathbf{u} - \boldsymbol{\ell}| \, \mathrm{d}x \right)^{1-\frac{1}{s}} + \mathcal{V}(\tau)^{1-\frac{1}{s}} \right]$$

$$\times \fint_{B_{3\varrho/4}} \varphi(|\mathbf{Q}|) + \varphi(|D\mathbf{u}|) \, \mathrm{d}x,$$

where $c = c(n, N, \Delta_2(\varphi), \nu, L)$. This estimate, combined with (3.3) gives

$$\fint_{B_r} |\mathbf{V}_{|\mathbf{Q}|}(D\mathbf{u} - \mathbf{Q})|^2 \, \mathrm{d}x$$

$$\leqslant \sigma \fint_{B_t} |\mathbf{V}_{|\mathbf{Q}|}(D\mathbf{u} - \mathbf{Q})|^2 \, \mathrm{d}x + c \fint_{B_\varrho} \left| \mathbf{V}_{|\mathbf{Q}|} \left(\frac{\mathbf{u} - \boldsymbol{\ell}}{t-r} \right) \right|^2 \, \mathrm{d}x$$

$$+ c \left(\frac{\varrho}{t-r} \right)^{n(s-1)} \left[\omega \left(\fint_{B_\varrho} |\mathbf{u} - \mathbf{u}_0| + |\mathbf{u} - \boldsymbol{\ell}| \, \mathrm{d}x \right)^{1-\frac{1}{s}} + \mathcal{V}(\tau)^{1-\frac{1}{s}} \right]$$

$$\times \fint_{B_{3\varrho/4}} \varphi(|\mathbf{Q}|) + \varphi(|D\mathbf{u}|) \, \mathrm{d}x,$$

$$=: \sigma \fint_{B_t} |\mathbf{V}_{|\mathbf{Q}|}(D\mathbf{u} - \mathbf{Q})|^2 \, \mathrm{d}x + c \fint_{B_\varrho} \left| \mathbf{V}_{|\mathbf{Q}|} \left(\frac{\mathbf{u} - \boldsymbol{\ell}}{t-r} \right) \right|^2 \, \mathrm{d}x + c \left(\frac{\varrho}{t-r} \right)^{n(s-1)} \mathcal{U}.$$

Now, since the previous estimate holds for arbitrary radii r, t such that $\varrho/2 \leqslant r < t \leqslant 3\varrho/4$, the constant c depends only on $n, N, \Delta_2(\varphi), \nu, L$ and $\sigma < 1$, as a consequence of lemma 2.6 applied with $\beta := n(s-1)$ we obtain

$$\fint_{B_{\varrho/2}} |\mathbf{V}_{|\mathbf{Q}|}(D\mathbf{u} - \mathbf{Q})|^2 \, \mathrm{d}x \leqslant c \fint_{B_\varrho} \left| \mathbf{V}_{|\mathbf{Q}|} \left(\frac{\mathbf{u} - \boldsymbol{\ell}}{\varrho} \right) \right|^2 \, \mathrm{d}x + c\mathcal{U}. \qquad (3.4)$$

In view of lemma 3.1 applied with ϱ in place of $t - s$ and from (3.2) we get

$$\fint_{B_{3\varrho/4}} \varphi(|D\mathbf{u}|) \, \mathrm{d}x \leqslant c \fint_{B_\varrho} \varphi \left(\left| \frac{\mathbf{u} - \mathbf{u}_0}{\varrho} \right| \right) \, \mathrm{d}x$$

$$\leqslant c \fint_{B_\varrho} \varphi \left(\left| \frac{\mathbf{u} - \boldsymbol{\ell}}{\varrho} \right| \right) \, \mathrm{d}x + c\varphi(|\mathbf{Q}|)$$

$$\leqslant c \left[\fint_{B_\varrho} \left| \mathbf{V}_{|\mathbf{Q}|} \left(\frac{\mathbf{u} - \boldsymbol{\ell}}{\varrho} \right) \right|^2 \, \mathrm{d}x + \varphi(|\mathbf{Q}|) \right],$$

which combined with (3.4) and using the fact that $\omega \leqslant 1$ as well as $\mathcal{V}(\varrho) \leqslant 2L$ gives

$$
\begin{aligned}
\fint_{B_{\varrho/2}} & |\mathbf{V}_{|\mathbf{Q}|}(D\mathbf{u} - \mathbf{Q})|^2 \, dx \\
& \leqslant c \fint_{B_\varrho} \left| \mathbf{V}_{|\mathbf{Q}|} \left(\frac{\mathbf{u} - \boldsymbol{\ell}}{\varrho} \right) \right|^2 \, dx \\
& \quad + c\varrho^n \varphi(|\mathbf{Q}|) \left[\omega \left(\fint_{B_\varrho} |\mathbf{u} - \mathbf{u}_0| + |\mathbf{u} - \boldsymbol{\ell}| \, dx \right)^{1-\frac{1}{s}} + \mathcal{V}(\tau)^{1-\frac{1}{s}} \right],
\end{aligned}
$$

where $c = c(n, N, \Delta_2(\varphi), \nu, L)$. The Caccioppoli inequality then follows by taking means on both sides of the latter inequality. $\qquad\square$

We can apply lemma 3.4 to affine functions $\boldsymbol{\ell}_{x_0,r}(x) := (\mathbf{u})_{x_0,\varrho} + \mathbf{Q}(x - x_0)$ for some $\mathbf{Q} \in \mathbb{R}^{N \times n}$, and the resulting Caccioppoli inequality can be compared with that of [**7**, theorem 3.1]. We notice that, apart from an extra VMO term due to assumption (F4), the dependence of the integrand f also on $\mathbf{u}$ implies that the remainder term inside ω; i.e.,

$$
R(x_0, \varrho, \mathbf{u}, \mathbf{Q}) := \fint_{B_\varrho(x_0)} |\mathbf{u} - (\mathbf{u})_{x_0,\varrho}| + |\mathbf{u} - \boldsymbol{\ell}_{x_0,\varrho}| \, dx \tag{3.5}
$$

is, in general, non-monotone in the radius ϱ. Indeed, it can be estimated from above by the *Morrey-type excess*

$$
\Theta(x_0, \varrho) := \varrho \varphi^{-1} \left(\fint_{B_\varrho(x_0)} \varphi(|D\mathbf{u}|) \, dx \right), \tag{3.6}
$$

which fails to be monotone for small ϱ (lemma 3.5(i)). This does not allow, in general, for an application of Gehring's lemma in order to infer an higher integrability result: for this purpose, a suitable 'smallness' regime (3.9) has to be imposed (lemma 3.5(ii)).

LEMMA 3.5. *Let $\boldsymbol{\ell}_{x_0,\varrho}$ be an affine function as above, and $R(x_0, \varrho, \mathbf{u}, \boldsymbol{\ell}_{x_0,\varrho})$ be defined as in (3.5). Then*

(i)

$$
R(x_0, \varrho, \mathbf{u}, \mathbf{Q}) \leqslant c\Theta(x_0, \varrho) + \varrho|\mathbf{Q}|. \tag{3.7}
$$

In particular, if $\mathbf{Q} = (D\mathbf{u})_{x_0,\varrho}$, we have

$$
R(x_0, \varrho, \mathbf{u}, \mathbf{Q}) \leqslant c\Theta(x_0, \varrho). \tag{3.8}
$$

(ii) *if the smallness assumption*

$$\fint_{B_\varrho(x_0)} \varphi_{|\mathbf{Q}|}(|D\mathbf{u} - \mathbf{Q}|)\,\mathrm{d}x \leqslant \Lambda\varphi(|\mathbf{Q}|) \tag{3.9}$$

holds for some $\Lambda \in (0, 1]$, *then there exists a constant* $c = c(\varphi) > 0$ *such that*

$$\left(\fint_{B_\varrho(x_0)} |\mathbf{u} - (\mathbf{u})_{x_0,\varrho}|^{\mu_1}\,\mathrm{d}x\right)^{\frac{1}{\mu_1}} \leqslant c\Theta(x_0, \varrho) \leqslant c\varrho(|\mathbf{Q}|)\,; \tag{3.10}$$

hence

$$R(x_0, \varrho, \mathbf{u}, \mathbf{Q}) \leqslant c\varrho(|\mathbf{Q}|).$$

Proof. (i) First, from Poincaré inequality and Jensen's inequality we obtain

$$\varphi\left(\fint_{B_\varrho(x_0)} \frac{|\mathbf{u} - (\mathbf{u})_{x_0,\varrho}|}{\varrho}\,\mathrm{d}x\right) \leqslant \fint_{B_\varrho(x_0)} \varphi\left(\frac{|\mathbf{u} - (\mathbf{u})_{x_0,\varrho}|}{\varrho}\right)\,\mathrm{d}x \leqslant c \fint_{B_\varrho(x_0)} \varphi(|D\mathbf{u}|)\,\mathrm{d}x,$$

whence

$$\fint_{B_\varrho(x_0)} |\mathbf{u} - (\mathbf{u})_{x_0,\varrho}|\,\mathrm{d}x \leqslant c\Theta(x_0, \varrho).$$

Then, recalling the definition of $\ell_{x_0,r}$, it is immediate to infer the estimate (3.7). As for (3.8), it follows from (3.7) since $\varrho|(D\mathbf{u})_{x_0,\varrho}| \leqslant c\Theta(x_0, \varrho)$.

(ii) We note from (φ2) that $\varphi(t^{1/\mu_1})$ is convex for $t \geqslant 0$. Applying Jensen's inequality, the Poincaré type estimate in theorem 2.5 and the change-shift formula (2.7) with $\mathbf{a} = \mathbf{0}$, and using assumption (3.9), we obtain

$$\varphi\left(\left(\fint_{B_\varrho(x_0)} \left[\frac{|\mathbf{u} - (\mathbf{u})_{x_0,\varrho}|}{\varrho}\right]^{\mu_1}\,\mathrm{d}x\right)^{\frac{1}{\mu_1}}\right)$$

$$\leqslant \fint_{B_\varrho(x_0)} \varphi\left(\frac{|\mathbf{u} - (\mathbf{u})_{x_0,\varrho}|}{\varrho}\right)\,\mathrm{d}x \leqslant c \fint_{B_\varrho(x_0)} \varphi(|D\mathbf{u}|)\,\mathrm{d}x$$

$$\leqslant c \fint_{B_\varrho(x_0)} \varphi(|D\mathbf{u} - \mathbf{Q}|)\,\mathrm{d}x + c\varphi(|\mathbf{Q}|)$$

$$\leqslant c \fint_{B_\varrho(x_0)} \varphi_{|\mathbf{Q}|}(|D\mathbf{u} - \mathbf{Q}|)\,\mathrm{d}x + c\varphi(|\mathbf{Q}|) \leqslant \varphi(c(|\mathbf{Q}|)),$$

which yields (3.10) up to applying φ^{-1} to both sides. $\qquad\square$

Now, we are in position to establish a 'conditioned' higher integrability result for $\varphi_{|\mathbf{Q}|}(|D\mathbf{u} - \mathbf{Q}|)$, under the smallness assumption (3.9). The result follows as a consequence of Gehring's lemma with increasing supports (lemma 2.7):

COROLLARY 3.6. *If $\mathbf{u} \in W^{1,\varphi}(\Omega; \mathbb{R}^N)$ is a minimizer of the functional (1.1) under assumptions* (F1)–(F7), *and $\mathbf{Q} \in \mathbb{R}^{N \times n}$ is such that (3.9) holds for some $\Lambda \in (0,1]$, then there exist a constant $c = c(n, N, \Delta_2(\varphi), \nu, L) > 0$ and $\sigma > 1$ such that*

$$
\left(\fint_{B_{\varrho/2}(x_0)} \varphi_{|\mathbf{Q}|}^\sigma(|D\mathbf{u} - \mathbf{Q}|) \, dx \right)^{\frac{1}{\sigma}}
$$

$$
\leqslant c \fint_{B_\varrho(x_0)} \varphi_{|\mathbf{Q}|}(|D\mathbf{u} - \mathbf{Q}|) \, dx + c\varphi(|\mathbf{Q}|) \left[\omega\left(\varrho|\mathbf{Q}|\right)^{1-\frac{1}{s}} + [\mathcal{V}(\varrho)]^{1-\frac{1}{s}} \right] \quad (3.11)
$$

holds for every $s \in (1, s_0]$ where s_0 is that of lemma 3.2.

Proof. Let $y \in \Omega$ and $r > 0$ be such that $B_{2r}(y) \subset\subset B_\varrho(x_0)$. In view of lemma 3.4 applied with $\varrho = 2r$, $x_0 = y$, $\mathbf{u}_0 = (\mathbf{u})_{y,2r}$ and an arbitrary $\mathbf{Q}$, we obtain

$$
\fint_{B_r(y)} \varphi_{|\mathbf{Q}|}(|D\mathbf{u} - \mathbf{Q}|) \, dx \leqslant c \fint_{B_{2r}(y)} \varphi_{|\mathbf{Q}|}\left(\frac{|\mathbf{u} - (\mathbf{u})_{y,2r} - \mathbf{Q}(x-y)|}{2r} \right) dx
$$

$$
+ c\varphi(|\mathbf{Q}|) \left[\omega\left(\fint_{B_{2r}(y)} |\mathbf{u} - (\mathbf{u})_{y,2r}| + |\mathbf{Q}||x - y| \, dx \right)^{1-\frac{1}{s}} + \mathcal{V}(2r)^{1-\frac{1}{s}} \right]. \quad (3.12)
$$

Here, we observe that

$$
\fint_{B_{2r}(y)} |\mathbf{u} - (\mathbf{u})_{y,2r}| + |\mathbf{Q}||x - y| \, dx \leqslant c \fint_{B_{2r}(y)} |\mathbf{u} - (\mathbf{u})_{y,2r}| \, dx + c|\mathbf{Q}|r
$$

$$
\leqslant c \left(\frac{1}{\varrho|\mathbf{Q}|} \fint_{B_{2r}(y)} |\mathbf{u} - (\mathbf{u})_{y,2r}| \, dx + 1 \right) \varrho|\mathbf{Q}|,
$$

which, recalling that $\omega(ct) \leqslant c\omega(t)$ when $c \geqslant 1$ since ω is concave and $\omega(0) = 0$, yields

$$
\omega\left(\fint_{B_{2r}(y)} |\mathbf{u} - (\mathbf{u})_{y,2r}| + |\mathbf{Q}||x - y| \, dx \right)^{1-\frac{1}{s}}
$$

$$
\leqslant c \left(\frac{1}{\varrho|\mathbf{Q}|} \fint_{B_{2r}(y)} |\mathbf{u} - (\mathbf{u})_{y,2r}| \, dx + 1 \right) \omega(\varrho|\mathbf{Q}|)^{1-\frac{1}{s}}.
$$

Moreover, as $\fint_{B_{2r}(y)} \mathbf{u} - (\mathbf{u})_{y,2r} - \mathbf{Q}(x - y)\, \mathrm{d}x = \mathbf{0}$, by the Sobolev–Poincaré type inequality (2.12),

$$
\fint_{B_{2r}(y)} \varphi_{|\mathbf{Q}|}\left(\frac{|\mathbf{u} - (\mathbf{u})_{y,2r} - \mathbf{Q}(x-y)|}{r}\right)\, \mathrm{d}x \leqslant c \left(\fint_{B_{2r}(y)} \varphi^{\alpha}_{|\mathbf{Q}|}(|D\mathbf{u} - \mathbf{Q}|)\, \mathrm{d}x\right)^{\frac{1}{\alpha}}
$$

for some $\alpha \in (0,1)$. Therefore, plugging the preceding two estimates into (3.12) and taking into account that

$$
\fint_{B_{2r}(y)} |\mathbf{u} - (\mathbf{u})_{y,2r}|\, \mathrm{d}x \leqslant 2 \fint_{B_{2r}(y)} |\mathbf{u} - (\mathbf{u})_{x_0,\varrho}|\, \mathrm{d}x,
$$

we obtain

$$
\fint_{B_r(y)} \varphi_{|\mathbf{Q}|}(|D\mathbf{u} - \mathbf{Q}|)\, \mathrm{d}x
$$

$$
\leqslant c \left(\fint_{B_{2r}(y)} \varphi^{\alpha}_{|\mathbf{Q}|}(|D\mathbf{u} - \mathbf{Q}|)\, \mathrm{d}x\right)^{\frac{1}{\alpha}}
$$

$$
+ c\frac{\varphi(|\mathbf{Q}|)\omega(\varrho|\mathbf{Q}|)^{1-\frac{1}{s}}}{\varrho|\mathbf{Q}|} \fint_{B_{2r}(y)} |\mathbf{u} - (\mathbf{u})_{x_0,\varrho}|\, \mathrm{d}x + c\varphi(|\mathbf{Q}|)
$$

$$
\times \left[\omega\left(\varrho|\mathbf{Q}|\right)^{1-\frac{1}{s}} + \mathcal{V}(\varrho)^{1-\frac{1}{s}}\right].
$$

Now, since $\mathbf{u} - (\mathbf{u})_{x_0,\varrho} \in L^{\mu_1}(B_\varrho(x_0))$, as a consequence of Gehring's lemma there exists $\sigma = \sigma(n, N, \mu_1, \mu_2, \nu, L) \in (1, \mu_1)$ such that

$$
\left(\fint_{B_{\rho/2}(x_0)} \varphi^{\sigma}_{|\mathbf{Q}|}(|D\mathbf{u} - \mathbf{Q}|)\, \mathrm{d}x\right)^{\frac{1}{\sigma}} \leqslant c \fint_{B_\varrho(x_0)} \varphi_{|\mathbf{Q}|}(|D\mathbf{u} - \mathbf{Q}|)\, \mathrm{d}x
$$

$$
+ c\frac{\varphi(|\mathbf{Q}|)\omega(\varrho|\mathbf{Q}|)^{1-\frac{1}{s}}}{\varrho|\mathbf{Q}|} \left(\fint_{B_\varrho(x_0)} |\mathbf{u} - (\mathbf{u})_{x_0,\varrho}|^{\sigma}\, \mathrm{d}x\right)^{\frac{1}{\sigma}}
$$

$$
+ c\varphi(|\mathbf{Q}|) \left[\omega\left(\varrho|\mathbf{Q}|\right)^{1-\frac{1}{s}} + \mathcal{V}(\varrho)^{1-\frac{1}{s}}\right].
$$

Finally, applying lemma 3.5 (ii), we obtain (3.11). □

We conclude this section by introducing the excess functional and other tools useful in the sequel. Let $\boldsymbol{L}_{x_0,\varrho} : \mathbb{R}^n \to \mathbb{R}^N$ be the affine function associated to $\mathbf{u}$

defined as

$$\boldsymbol{L}_{x_0,\varrho}(x) := (\mathbf{u})_{x_0,\varrho} + \mathbf{Q}_{x_0,\varrho}(x - x_0), \tag{3.13}$$

where $\mathbf{Q}_{x_0,\varrho} := (D\mathbf{u})_{x_0,\varrho}$. For $x_0 \in \Omega$ and $\varrho \in (0, \mathrm{dist}(x_0, \partial\Omega))$, $\varrho \leqslant 1$, we define the *excess functional* as

$$\Phi(x_0, \varrho) \equiv \Phi(x_0, \varrho, \boldsymbol{L}_{x_0,\varrho}) := \fint_{B_\varrho(x_0)} \varphi_{|(D\mathbf{u})_{x_0,\varrho}|}(|D\mathbf{u} - (D\mathbf{u})_{x_0,\varrho}|) \, \mathrm{d}x \tag{3.14}$$

and

$$\Psi(x_0, \varrho) := \fint_{B_\varrho(x_0)} \varphi\left(\frac{|\mathbf{u} - (\mathbf{u})_{x_0,\varrho}|}{\varrho}\right) \, \mathrm{d}x. \tag{3.15}$$

Moreover, we define also

$$H(x_0, \varrho) := \frac{1}{1 + (2L)^{1-\frac{1}{s}}} \left([\omega(\varrho|(D\mathbf{u})_{x_0,\varrho}|)]^{1-\frac{1}{s}} + [\mathcal{V}(\varrho)]^{1-\frac{1}{s}}\right), \tag{3.16}$$

and

$$\widetilde{H}(x_0, \varrho) := \frac{1}{1 + (2L)^{1-\frac{1}{s}}} \left([\omega(\Theta(x_0, \varrho))]^{1-\frac{1}{s}} + [\mathcal{V}(\varrho)]^{1-\frac{1}{s}}\right), \tag{3.17}$$

where $s \in (1, s_0]$ is the exponent of lemma 3.3 and $\Theta(x_0, \varrho)$ is the excess defined in (3.6). Since $\omega \leqslant 1$ and $\mathcal{V}(\varrho) \leqslant 2L$, we have that $H(x_0, \varrho), \widetilde{H}(x_0, \varrho) \leqslant 1$, and

$$H(x_0, \varrho) \leqslant c\widetilde{H}(x_0, \varrho)$$

as a consequence of lemma 3.5(i). Under the smallness assumption $\Phi(x_0, \varrho) \leqslant \Lambda\varphi(|(D\mathbf{u})_{x_0,\varrho}|)$, by virtue of lemma 3.5(ii) there exists a constant $\tilde{c} = \tilde{c}(\varphi)$ such that

$$\frac{1}{\tilde{c}}\widetilde{H}(x_0, \varrho) \leqslant H(x_0, \varrho) \leqslant c\widetilde{H}(x_0, \varrho).$$

We can rewrite the Caccioppoli inequality (3.11) as

$$\Phi(x_0, \varrho/2) \leqslant c\Phi(x_0, \varrho) + c\varphi(|(D\mathbf{u})_{x_0,\varrho}|)H(x_0, \varrho). \tag{3.18}$$

Note also that by (2.6) and, e.g., [13, lemma A.2] we have the following equivalence:

$$\Phi(x_0, \varrho) \sim \fint_{B_\varrho(x_0)} |\mathbf{V}(D\mathbf{u}) - \mathbf{V}((D\mathbf{u})_{x_0,\varrho})|^2 \, \mathrm{d}x \sim \fint_{B_\varrho(x_0)} |\mathbf{V}(D\mathbf{u}) - (\mathbf{V}(D\mathbf{u}))_{x_0,\varrho}|^2 \, \mathrm{d}x.$$

In the case $x_0 = 0$, we will use the shorthands $\Phi(\varrho)$, $\Psi(\varrho)$, $\Theta(\varrho)$, $H(\varrho)$ and $\widetilde{H}(\varrho)$ in place of $\Phi(0, \varrho)$, $\Psi(0, \varrho)$, $\Theta(0, \varrho)$, $H(0, \varrho)$ and $\widetilde{H}(0, \varrho)$, respectively.

3.2. Comparison maps via Ekeland's variational principle

The proof of the main results will require suitable comparison functions, which will be constructed with a freezing argument in the variables $(x, \mathbf{u})$ based on Ekeland's variational principle. We recall below a version of this classical tool, whose proof can be found, e.g., in [**25**, theorem 5.6].

LEMMA 3.7 (Ekeland's principle). *Let (X, d) be a complete metric space, and assume that $F : X \to [0, \infty]$ be not identically ∞ and lower semicontinuous with respect to the metric topology on X. If for some $u \in X$ and some $\kappa > 0$, there holds*

$$F(u) \leqslant \inf_X F + \kappa,$$

then there exists $v \in X$ with the properties

$$d(u, v) \leqslant 1 \ \text{ and } \ F(v) \leqslant F(w) + \kappa d(v, w) \quad \forall w \in X.$$

Although a similar analysis in the Orlicz setting, for integrands $f = f(x, \boldsymbol{\xi})$, has been performed in [**7**, theorem 3.3], we will follow a quite different argument, which refers to the case of p-growth as in [**4**, lemma 3.7]. We will also specify the appropriate complete metric space X, which is not explicitly mentioned in [**7**, theorem 3.3].

To this aim, let $B_\varrho(x_0) \subseteq \Omega$ with $\varrho \leqslant \varrho_0$ and set

$$g(\boldsymbol{\xi}) \equiv g_{x_0, \varrho}(\boldsymbol{\xi}) := (f(\cdot, (\mathbf{u})_{x_0, \varrho}, \boldsymbol{\xi}))_{x_0, \varrho} \quad \text{ for all } \boldsymbol{\xi} \in \mathbb{R}^{N \times n}, \tag{3.19}$$

and

$$K(x_0, \varrho) := \widetilde{H}(x_0, \varrho) \Psi(x_0, \varrho) \tag{3.20}$$

where $\widetilde{H}(x_0, \varrho)$ and $\Psi(x_0, \varrho)$ are defined as in (3.17) and (3.15), respectively.

As for the complete metric space (X, d), following [**36**, lemma 4.4] we consider

$$X := \left\{ \mathbf{w} \in \mathbf{u} + W_0^{1,1}(B_{\varrho/2}(x_0)) : \fint_{B_{\varrho/2}(x_0)} \varphi(|D\mathbf{w}|) \, \mathrm{d}x \leqslant \fint_{B_{\varrho/2}(x_0)} \varphi(|D\mathbf{u}|) \, \mathrm{d}x \right\}$$

with the metric

$$d(\mathbf{w}_1, \mathbf{w}_2) := \frac{1}{c_* \varphi^{-1}(K(\varrho))} \fint_{B_{\varrho/2}(x_0)} |D\mathbf{w}_1 - D\mathbf{w}_2| \, \mathrm{d}x,$$

for $\mathbf{w}_1, \mathbf{w}_2 \in \mathbf{u} + W_0^{1,1}(B_{\varrho/2(x_0)}, \mathbb{R}^N)$, and note that the functional

$$\mathcal{G}[\mathbf{w}] := \fint_{B_{\varrho/2}(x_0)} g(D\mathbf{w}) \, \mathrm{d}x \quad \text{ in } \mathbf{u} + W_0^{1,1}(B_{\varrho/2}(x_0), \mathbb{R}^N), \tag{3.21}$$

is lower semicontinuous in the metric topology. We would get a comparison map $\mathbf{v} \in \mathbf{u} + W_0^{1,1}(B_{\varrho/2}(x_0), \mathbb{R}^N)$ by proving the following lemma.

LEMMA 3.8. *Assume that* $\mathbf{u} \in W^{1,\varphi}(\Omega, \mathbb{R}^N)$ *is a minimizer of the functional* (1.1)*, under assumptions* (F1)–(F6)*. Then there exists a minimizer* $\mathbf{v} \in \mathbf{u} + W_0^{1,1}(B_{\varrho/2}(x_0), \mathbb{R}^N)$ *of the functional*

$$\widetilde{\mathcal{G}}[\mathbf{w}] := \fint_{B_{\varrho/2}(x_0)} g(D\mathbf{w}) \, \mathrm{d}x + \frac{K(x_0, \varrho)}{\varphi^{-1}(K(x_0, \varrho))} \fint_{B_{\varrho/2}(x_0)} |D\mathbf{v} - D\mathbf{w}| \, \mathrm{d}x,$$

that satisfies

$$\fint_{B_{\varrho/2}(x_0)} |D\mathbf{v} - D\mathbf{u}| \, \mathrm{d}x \leqslant c_* \varphi^{-1}(K(x_0, \varrho)) \tag{3.22}$$

for some constant $c_* = c_*(n, N, \Delta_2(\varphi), \nu, L)$*. Moreover,* $\mathbf{v}$ *fulfils the following Euler-Lagrange variational inequality:*

$$\left| \fint_{B_{\varrho/2}(x_0)} \langle Dg(D\mathbf{v}) | D\boldsymbol{\eta} \rangle \, \mathrm{d}x \right| \leqslant \frac{K(x_0, \varrho)}{\varphi^{-1}(K(x_0, \varrho))} \fint_{B_{\varrho/2}(x_0)} |D\boldsymbol{\eta}| \, \mathrm{d}x \tag{3.23}$$

for every $\boldsymbol{\eta} \in C_0^\infty(B_{\varrho/2}(x_0), \mathbb{R}^N)$*.*

Proof. We may assume, without loss of generality, that $x_0 = 0$ and, correspondingly, we use the shorthand $K(\varrho)$ for $K(0, \varrho)$. As a first remark, we recall that from lemma 3.1 with $r = \frac{3}{4}\varrho$ we have

$$\fint_{B_{3\varrho/4}} \varphi(|D\mathbf{u}|) \, \mathrm{d}x \leqslant c \fint_{B_\varrho} \varphi\left(\frac{|\mathbf{u} - (\mathbf{u})_\varrho|}{\varrho} \right) \, \mathrm{d}x = c\Psi(\varrho), \tag{3.24}$$

where $c = c(\Delta_2(\varphi), L, \nu)$. We then denote by $\tilde{\boldsymbol{v}} \in X$ a minimizer of the functional (3.21) whose existence is ensured by the direct method under assumptions (F1)–(F2). From the minimality of $\tilde{\boldsymbol{v}}$, assumption (F1) and (2.10) we get

$$\begin{aligned}
\fint_{B_{\varrho/2}} \varphi(|D\tilde{\boldsymbol{v}}|) \, \mathrm{d}x &\leqslant \frac{1}{\nu} \fint_{B_{\varrho/2}} g(D\tilde{\boldsymbol{v}}) - g(\mathbf{0}) \, \mathrm{d}x \\
&\leqslant \frac{1}{\nu} \fint_{B_{\varrho/2}} g(D\mathbf{u}) - g(\mathbf{0}) \, \mathrm{d}x \leqslant \frac{c(\varphi)L}{\nu} \fint_{B_{\varrho/2}} \varphi(|D\mathbf{u}|) \, \mathrm{d}x.
\end{aligned} \tag{3.25}$$

By the sublinearity of φ, the Poincaré inequality (theorem 2.5), Jensen's inequality and (3.24) this gives

$$\varphi\left(\fint_{B_{\varrho/2}} \frac{|\tilde{\boldsymbol{v}} - (\mathbf{u})_\varrho|}{\varrho}\,\mathrm{d}x\right)$$

$$\leqslant c\left(\fint_{B_{\varrho/2}} \varphi\left(\frac{|\tilde{\boldsymbol{v}} - \mathbf{u}|}{\varrho}\right)\mathrm{d}x + \fint_{B_{\varrho/2}} \varphi\left(\frac{|\mathbf{u} - (\mathbf{u})_\varrho|}{\varrho}\right)\mathrm{d}x\right)$$

$$\leqslant c\left[\left(\fint_{B_{\varrho/2}} \varphi^\alpha(|D\tilde{\boldsymbol{v}} - D\mathbf{u}|)\,\mathrm{d}x\right)^{\frac{1}{\alpha}} + \left(\fint_{B_{\varrho/2}} \varphi^\alpha(|D\mathbf{u}|)\,\mathrm{d}x\right)^{\frac{1}{\alpha}}\right]$$

$$\leqslant c\fint_{B_{\varrho/2}} \varphi(|D\tilde{\boldsymbol{v}}|) + \varphi(|D\mathbf{u}|)\,\mathrm{d}x$$

$$\leqslant c\fint_{B_{\varrho/2}} \varphi(|D\mathbf{u}|)\,\mathrm{d}x,$$

whence

$$\fint_{B_{\varrho/2}} |\tilde{\boldsymbol{v}} - (\mathbf{u})_\varrho|\,\mathrm{d}x \leqslant c\Theta(\varrho) \tag{3.26}$$

where $c = c(\varphi, n, L, \nu)$. Moreover, as a consequence of the higher integrability results of both lemmas 3.2 and 3.3, together with (3.25) and (3.24), we infer the higher integrability result

$$\left(\fint_{B_{\varrho/2}} \varphi^s(|D\tilde{\boldsymbol{v}}|)\,\mathrm{d}x\right)^{\frac{1}{s}} \leqslant c\fint_{B_{3\varrho/4}} \varphi(|D\mathbf{u}|)\,\mathrm{d}x \leqslant c\Psi(\varrho), \tag{3.27}$$

where $c = c(n, N, \varphi, \nu, L)$ and $s = s(n, N, \varphi, \nu, L) \in (1, s_0]$.

Now we prove that $\mathbf{u}$ is an almost minimizer of the functional $\mathcal{G}$. Indeed, from the minimality of $\mathbf{u}$ and assumptions (F4), (F3) we get

$$\fint_{B_{\varrho/2}} f(x, \mathbf{u}, D\mathbf{u})\,\mathrm{d}x - \mathcal{G}[\tilde{\boldsymbol{v}}] \leqslant \fint_{B_{\varrho/2}} f(x, \tilde{\boldsymbol{v}}, D\tilde{\boldsymbol{v}})\,\mathrm{d}x - \mathcal{G}[\tilde{\boldsymbol{v}}]$$

$$= \fint_{B_{\varrho/2}} f(x, \tilde{\boldsymbol{v}}, D\tilde{\boldsymbol{v}}) - (f(\cdot, \tilde{\boldsymbol{v}}, D\tilde{\boldsymbol{v}}))_\varrho\,\mathrm{d}x$$

$$+ \fint_{B_{\varrho/2}} (f(\cdot, \tilde{\boldsymbol{v}}, D\tilde{\boldsymbol{v}}))_\varrho - (f(\cdot, (\mathbf{u})_\varrho, D\tilde{\boldsymbol{v}}))_\varrho \, \mathrm{d}x$$

$$\leqslant c(L) \fint_{B_{\varrho/2}} [v_0(\cdot, \varrho) + \omega(|\tilde{\boldsymbol{v}} - (\mathbf{u})_\varrho|)]\varphi(|D\tilde{\boldsymbol{v}}|) \, \mathrm{d}x.$$

Then, by using Jensen's inequality, the concavity and sub-linearity of ω, (3.26) and (3.27), from the previous estimate we obtain

$$\fint_{B_{\varrho/2}} f(x, \mathbf{u}, D\mathbf{u}) \, \mathrm{d}x - \mathcal{G}[\tilde{\boldsymbol{v}}]$$

$$\leqslant c \left[\omega \left(\fint_{B_{\varrho/2}} |\tilde{\boldsymbol{v}} - (\mathbf{u})_\varrho| \, \mathrm{d}x \right)^{1-\frac{1}{s}} + [\mathcal{V}(\varrho)]^{1-\frac{1}{s}} \right] \left(\fint_{B_{\varrho/2}} \varphi^s(|D\tilde{\boldsymbol{v}}|) \, \mathrm{d}x \right)^{\frac{1}{s}}$$

$$\leqslant c \left[\omega(\Theta(\varrho))^{1-\frac{1}{s}} + [\mathcal{V}(\varrho)]^{1-\frac{1}{s}} \right] \Psi(\varrho) = cK(\varrho), \tag{3.28}$$

where $c = c(n, N, \Delta_2(\varrho), \nu, L)$. Arguing similarly, we can estimate

$$\mathcal{G}[\mathbf{u}] - \fint_{B_{\varrho/2}} f(x, \mathbf{u}, D\mathbf{u}) \, \mathrm{d}x = \fint_{B_{\varrho/2}} [(f(\cdot, (\mathbf{u})_\varrho, D\mathbf{u}))_\varrho - f(x, (\mathbf{u})_\varrho, D\mathbf{u})] \, \mathrm{d}x$$

$$+ \fint_{B_{\varrho/2}} [f(x, (\mathbf{u})_\varrho, D\mathbf{u}) - f(x, \mathbf{u}, D\mathbf{u})] \, \mathrm{d}x$$

$$\leqslant c \left[\omega(\Theta(\varrho))^{1-\frac{1}{s}} + [\mathcal{V}(\varrho)]^{1-\frac{1}{s}} \right] \Psi(\varrho) = cK(\varrho), \tag{3.29}$$

where the constant c has the same dependencies as before. Adding term by term (3.28)–(3.29) and taking into account the minimality of $\tilde{\boldsymbol{v}}$, we infer

$$\mathcal{G}[\mathbf{u}] \leqslant \mathcal{G}[\tilde{\boldsymbol{v}}] + c_* K(\varrho) = \min_{\mathbf{u}+W_0^{1,1}(B_{\varrho/2}, \mathbb{R}^N)} \mathcal{G} + c_* K(\varrho),$$

for a constant $c_* = c_*(n, N, \Delta_2(\varphi), \nu, L)$. Finally, Ekeland's variational principle (lemma 3.7) with the choice $\kappa = c_* K(\varrho)$ provides the existence of a function $\mathbf{v} \in X$ with the desired property of minimality for the functional $\widetilde{\mathcal{G}}$ and such that $d(\mathbf{u}, \mathbf{v}) \leqslant 1$, which corresponds to (3.22). The inequality (3.23) follows from the validity of the associated Euler-Lagrange variational inequality for $\mathbf{v}$ in a standard way. $\qquad \square$

3.3. Approximate $\mathcal{A}$-harmonicity and φ-harmonicity

In this section, we provide two different linearization strategies for the minimization problem, along the lines of [**4**, section 3.2], where an analogous analysis has been performed for functionals with p-growth. On the one hand, with lemma 3.9 we will show that the minimizer $\mathbf{u}$ of $\mathcal{F}$ is an almost $\mathcal{A}$-harmonic function for a

suitable elliptic bilinear form $\mathcal{A}$. On the other hand, this **u** turns out to be an almost φ-harmonic function (see lemma 3.10). These results will allow us to apply the $\mathcal{A}$-harmonic approximation lemma, respectively the φ-harmonic approximation lemma. The proof will require, in both cases, the comparison maps obtained with lemma 3.8.

We start by proving the approximate $\mathcal{A}$-harmonicity of a minimizer to (1.1). To this aim, only assumptions (F1)–(F6) are required on f.

Let $\boldsymbol{L}_{x_0,\varrho}$ be the affine function associated to **u** as in (3.13), which complies with $\boldsymbol{L}_{x_0,\varrho}(x_0) = (\mathbf{u})_{x_0,\varrho}$ and $D\boldsymbol{L}_{x_0,\varrho} = (D\mathbf{u})_{x_0,\varrho} =: \mathbf{Q}_{x_0,\varrho}$. We set

$$\mathcal{A} := \frac{D^2 g((D\mathbf{u})_{x_0,\varrho})}{\varphi''(|(D\mathbf{u})_{x_0,\varrho}|)} \equiv \frac{\left(D^2 f(\cdot, (\mathbf{u})_{x_0,\varrho}, (D\mathbf{u})_{x_0,\varrho})\right)_{x_0,\varrho}}{\varphi''(|(D\mathbf{u})_{x_0,\varrho}|)}.$$

We point out that $\mathcal{A}$ defined above is a bilinear form on $\mathbb{R}^{N \times n}$, satisfying the ellipticity assumption (2.13) by virtue of (F2) and (F3).

LEMMA 3.9. *Let* $\mathbf{u} \in W^{1,\varphi}(\Omega, \mathbb{R}^N)$ *be a minimizer of the functional* (1.1), *under assumptions* (F1)–(F6), *and assume that for a ball* $B_\varrho(x_0) \subseteq \Omega$ *the non-degeneracy assumptions*

$$\Phi(x_0, \varrho) \leqslant \varphi(|(D\mathbf{u})_{x_0,\varrho}|) \quad and \quad \varrho \leqslant 1,$$

are satisfied. Then, **u** *is approximately* $\mathcal{A}$-*harmonic on the ball* $B_{\varrho/2}(x_0)$, *in the sense that there exists* $\beta_1 = \beta_1(n, N, \mu_1, \mu_2, \nu, L, \beta_0) \in (0, \frac{1}{2})$ *such that*

$$\left| \fint_{B_{\varrho/2}(x_0)} \langle \mathcal{A}(D\mathbf{u} - (D\mathbf{u})_{x_0,\varrho}) | D\boldsymbol{\eta} \rangle \, \mathrm{d}x \right|$$

$$\leqslant c |(D\mathbf{u})_{x_0,\varrho}| \|D\boldsymbol{\eta}\|_\infty \left\{ [H(x_0, \varrho)]^{\beta_1} + \frac{\Phi(x_0, \varrho)}{\varphi(|(D\mathbf{u})_{x_0,\varrho}|)} + \left(\frac{\Phi(x_0, \varrho)}{\varphi(|(D\mathbf{u})_{x_0,\varrho}|)} \right)^{\frac{1+\beta_0}{2}} \right\}$$

$$\tag{3.30}$$

holds for every $\boldsymbol{\eta} \in C_c^\infty(B_{\varrho/2}(x_0), \mathbb{R}^N)$ *for some constant* $c = c(n, N, \mu_1, \mu_2, \nu, c_0, L) > 0$, *where* μ_1, μ_2 *are the characteristics of* φ *and* c_0, β_0 *are the constants of assumption* (F6).

Proof. See [**7**, lemma 4.1]. □

If, in addition, f complies also with (F7), we can show that each local minimizer of the functional $\mathcal{F}(\mathbf{u})$ (eq. (1.1)) is almost φ-harmonic.

For this, we preliminarily note (see [**7**, eq. (4.19)–(4.20)]) that assumption (F7) implies the following:

for all $\delta > 0$, exists $\sigma = \sigma(\delta) > 0$ such that $\left| Dg(\mathbf{P}) - \frac{\mathbf{P}}{|\mathbf{P}|} \varphi'(|\mathbf{P}|) \right| \leqslant \delta \varphi'(|\mathbf{P}|),$

$$\tag{3.31}$$

for every $\mathbf{P} \in \mathbb{R}^{N \times n}$ with $0 < |\mathbf{P}| \leqslant \sigma$, where the function g has been introduced in (3.19).

We then have the following result.

LEMMA 3.10. *Let $\mathbf{u} \in W_{loc}^{1,\varphi}(\Omega, \mathbb{R}^N)$ be a local minimizer of the functional (1.1), and assume that f complies also with* (F7). *Then there exists $\beta_2 = \beta_2(n, N, \mu_1, \mu_2, c_0, L) \in (0, \frac{1}{2})$ such that, for every $\delta > 0$ and for $\sigma = \sigma(\delta) > 0$ given by (3.31), the inequality*

$$
\left| \fint_{B_{\varrho/2}(x_0)} \left\langle \frac{\varphi'(|D\mathbf{u}|)}{|D\mathbf{u}|} D\mathbf{u} \middle| D\boldsymbol{\eta} \right\rangle \, \mathrm{d}x \right|
$$

$$
\leqslant c \left(\delta + [\widetilde{H}(x_0, \varrho)]^{\beta_2} + \frac{\varphi^{-1}(\Psi(x_0, \varrho))}{\sigma} \right) \left(\fint_{B_{\varrho}(x_0)} \varphi(|D\mathbf{u}|) \, \mathrm{d}x + \varphi(\|D\boldsymbol{\eta}\|_\infty) \right)
$$

holds for every $\boldsymbol{\eta} \in C_c^\infty(B_{\varrho/2}(x_0), \mathbb{R}^N)$ for some constant $c = c(n, N, \mu_1, \mu_2, c_0, \nu, L) > 0$.

Proof. See [**7**, lemma 4.3]. $\qquad \square$

3.4. Excess decay estimates: the non-degenerate regime

We start by establishing excess improvement estimates in the *non-degenerate regime* characterized by (3.33) below, i.e. the fact that $\Phi(x_0, \varrho) \leqslant c\varphi(|(D\mathbf{u})_{x_0,\varrho}|)$. The strategy of the proof is to exploit lemma 3.9 to approximate the given minimizer by $\mathcal{A}$-harmonic functions, for which suitable decay estimates are available from theorem 2.9.

We introduce the *hybrid excess functional*

$$
\Phi_*(x_0, \varrho) := \Phi(x_0, \varrho) + \varphi(|(D\mathbf{u})_{x_0,\varrho}|)[H(x_0, \varrho)]^{\beta_1}, \tag{3.32}
$$

where β_1 is the exponent of lemma 3.9. Since $\beta_1 < 1/2$ and $H(x_0, \varrho) \leqslant 1$, we deduce, in particular, that $H(x_0, \varrho) \leqslant [H(x_0, \varrho)]^{\beta_1}$. Thus, the Caccioppoli inequality (3.18) can be re-read as

$$
\Phi(x_0, \varrho/2) \leqslant c\Phi_*(x_0, \varrho),
$$

where $c = c(n, N, \mu_1, \mu_2, \nu, L)$.

LEMMA 3.11. *For every $\varepsilon \in (0, 1)$ there exist $\delta_1, \delta_2 \in (0, 1]$, where $\delta_i = \delta_i(n, N, \mu_1, \mu_2, \beta_0, \nu, L, \varepsilon)$, $i = 1, 2$, with the following property: if*

$$
\frac{\Phi(x_0, \varrho)}{\varphi(|(D\mathbf{u})_{x_0,\varrho}|)} \leqslant \delta_1 \tag{3.33}
$$

$$
[H(x_0, \varrho)]^{\beta_1} \leqslant \delta_2 \tag{3.34}
$$

then the excess improvement estimate

$$\Phi(x_0, \vartheta\varrho) \leqslant c_{dec}\vartheta^2 \left[1 + \frac{\varepsilon}{\vartheta^{n+2}}\right] \Phi_*(x_0, \varrho) \tag{3.35}$$

holds for every $\vartheta \in (0,1)$ for some constant $c_{dec} = c_{dec}(n, N, \mu_1, \mu_2, \nu, L, c_1) > 0$, where Φ_* is defined in (3.32).

Proof. The proof follows the argument of [**7**, lemma 4.2]. We emphasize that corollary 3.6 is crucial in order to obtain the estimate

$$\left(\fint_{B_{\varrho/2}} \left[\frac{\varphi_{|\mathbf{Q}_\varrho|}(|D\mathbf{u} - \mathbf{Q}_\varrho|)}{\varphi(|\mathbf{Q}_\varrho|)}\right]^{s_0} \mathrm{d}x\right)^{\frac{1}{s_0}}$$

$$\leqslant \frac{c}{\varphi(|\mathbf{Q}_\varrho|)} \fint_{B_\varrho} \varphi_{|\mathbf{Q}_\varrho|}(|D\mathbf{u} - \mathbf{Q}_\varrho|)\, \mathrm{d}x + c[H(\varrho)]^{\beta_1}$$

$$\leqslant c\frac{\Phi_*(\varrho)}{\varphi(|\mathbf{Q}_\varrho|)},$$

which comes into play in applying the $\mathcal{A}$-harmonic approximation theorem in the modified version of remark 2.10. □

LEMMA 3.12. *Let $\vartheta \in (0,1)$, and assume that*

$$\frac{\Phi(x_0, \varrho)}{\varphi(|(D\mathbf{u})_{x_0,\varrho}|)} \leqslant \frac{\vartheta^n}{2^{\mu_2+1}c_{\mu_2}}, \tag{3.36}$$

where c_{μ_2} is the constant of the change of shift formula (2.7) with $\eta = \frac{1}{2^{\mu_2+1}}$. Then it holds that

$$|(D\mathbf{u})_{x_0,\varrho}| \leqslant 2|(D\mathbf{u})_{x_0,\vartheta\varrho}|. \tag{3.37}$$

Proof. As a consequence of (2.7) for $\eta = \frac{1}{2^{\mu_2+1}}$ and with (3.36) we get

$$\varphi(|(D\mathbf{u})_{x_0,\varrho} - (D\mathbf{u})_{x_0,\vartheta\varrho}|) \leqslant \fint_{B_{\vartheta\varrho}(x_0)} \varphi(|D\mathbf{u} - (D\mathbf{u})_{x_0,\varrho}|)\, \mathrm{d}x$$

$$\leqslant c_{\mu_2}\vartheta^{-n}\Phi(x_0, \varrho) + \frac{1}{2^{\mu_2+1}}\varphi(|(D\mathbf{u})_{x_0,\varrho}|)$$

$$\leqslant \frac{1}{2^{\mu_2}}\varphi(|(D\mathbf{u})_{x_0,\varrho}|),$$

whence, passing to φ^{-1} and taking into account (2.2), we obtain

$$|(D\mathbf{u})_{x_0,\varrho} - (D\mathbf{u})_{x_0,\vartheta\varrho}| \leqslant \frac{1}{2}|(D\mathbf{u})_{x_0,\varrho}|.$$

Now,

$$|(D\mathbf{u})_{x_0,\varrho}| \leqslant |(D\mathbf{u})_{x_0,\varrho} - (D\mathbf{u})_{x_0,\vartheta\varrho}| + |(D\mathbf{u})_{x_0,\vartheta\varrho}| \leqslant \frac{1}{2}|(D\mathbf{u})_{x_0,\varrho}| + |(D\mathbf{u})_{x_0,\vartheta\varrho}|,$$

whence (3.37) follows by re-absorbing the first term of the right-hand side into the left. □

The excess-decay estimate (3.35) can be iterated, as the non-degeneracy conditions (3.33)–(3.34) are also satisfied on any smaller ball $B_{\vartheta^m \varrho}(x_0)$, $m \in \mathbb{N}$, $\vartheta < 1$.

LEMMA 3.13. *Let $\Phi(x_0, \varrho)$ and $\Theta(x_0, \varrho)$ be defined as in (3.14) and (3.6), respectively. Then there exist constants δ_*, ε_*, $\varrho_* \in (0, 1]$ and ϑ such that the following holds: if the conditions*

$$\frac{\Phi(x_0, \varrho)}{\varphi(|(D\mathbf{u})_{x_0, \varrho}|)} \leqslant \varepsilon_* \quad and \quad \Theta(x_0, \varrho) \leqslant \delta_*. \tag{3.38}$$

hold on $B_\varrho(x_0) \subseteq \Omega$ for $\varrho \in (0, \varrho_*]$, then

$$\frac{\Phi(x_0, \vartheta^m \varrho)}{\varphi(|(D\mathbf{u})_{x_0, \vartheta^m \varrho}|)} \leqslant \varepsilon_* \quad and \quad \Theta(x_0, \vartheta^m \varrho) \leqslant \delta_* \tag{3.39}$$

for every $m = 0, 1, \ldots$. As a consequence, for any $\alpha \in (0, 1)$ the following Morrey-type estimate holds:

$$\Theta(y, r) \leqslant c \delta_* \left(\frac{r}{\varrho} \right)^\alpha \tag{3.40}$$

for all $y \in B_{\varrho/2}(x_0)$ and $r \in (0, \varrho/2]$.

Proof. As usual, we omit the explicit dependence on x_0. Let $\vartheta \in (0, 1)$ be such that

$$\vartheta \leqslant \min \left\{ (6 c_{\mathrm{dec}} 2^{\mu_2})^{-\frac{1}{2}}, \frac{1}{2^{\mu_2}}, \frac{1}{2^{\frac{\mu_2}{\mu_1(1-\alpha)}}} \right\}, \tag{3.41}$$

where c_{dec} is the constant of lemma 3.11 depending only on $n, N, \mu_1, \mu_2, \nu, L, c_0$. Correspondingly, let $\delta_i = \delta_i(n, N, \mu_1, \mu_2, \beta_0, \nu, L, \vartheta)$, $i = 1, 2$ be the constants of lemma 3.11, applied with the choice $\varepsilon = \vartheta^{n+2}$. We choose $\varepsilon_* > 0$ such that

$$\varepsilon_* \leqslant \min \left\{ \frac{\delta_1}{3}, \frac{\delta_2}{2}, \frac{\vartheta^n}{\max\{2 c_{\frac{1}{2}}, 2^{\mu_2+1} c_{\mu_2}\}} \right\}, \tag{3.42}$$

where $c_{\frac{1}{2}}$ is the constant in the change-shift formula (2.7) with $\eta = \frac{1}{2}$, and we fix the constant $\delta_* > 0$ so small that

$$\left(\frac{\omega(\delta_*)^{1-\frac{1}{s}}}{1 + (2L)^{1-\frac{1}{s}}} \right)^{\beta_1} < \varepsilon_*. \tag{3.43}$$

Moreover, we choose a radius $\varrho_* > 0$ such that

$$\varrho_* \leqslant 1 \quad \text{and} \quad \left(\frac{\mathcal{V}(\varrho_*)^{1-\frac{1}{s}}}{1 + (2L)^{1-\frac{1}{s}}} \right)^{\beta_1} < \varepsilon_*. \tag{3.44}$$

As a consequence, ε_*, δ_* and ϱ_* have the same dependencies as δ_1, δ_2. In addition, δ_* depends also on ω, while ϱ_* also on ω and $\mathcal{V}$.

We argue by induction on m. Since (3.39) are trivially true for $m = 0$ by assumption (3.38), our aim is to show that if (3.39) holds for some $m \geqslant 1$, then the corresponding inequalities hold with $m + 1$ in place of m. Setting

$$E(B_{\vartheta^m \varrho}) := \fint_{B_{\vartheta^m \varrho}} \varphi(|D\mathbf{u}|)\, \mathrm{d}x,$$

in order to prove the second inequalities in (3.39) it will suffice to show that

$$E(B_{\vartheta^m \varrho}) \leqslant \varphi\left(\frac{\delta_*}{\vartheta^m \varrho}\right). \tag{3.45}$$

We have, with (3.39) at step m, the shift-change formula (2.7) with $\eta = \frac{1}{2}$ and (3.42), the estimate

$$
\begin{aligned}
E(B_{\vartheta^{m+1} \varrho}) &\leqslant 2^{\mu_2 - 1}\left(c_{\frac{1}{2}}\vartheta^{-n}\Phi(\vartheta^m \varrho) + \frac{1}{2}\varphi(|(D\mathbf{u})_{\vartheta^m \varrho}|) + \varphi(|(D\mathbf{u})_{\vartheta^m \varrho}|)\right)\\
&\leqslant 2^{\mu_2 - 1}\left(c_{\frac{1}{2}}\vartheta^{-n}\Phi(\vartheta^m \varrho) + \frac{3}{2}E(B_{\vartheta^m \varrho})\right)\\
&\leqslant 2^{\mu_2 - 1}\left(c_{\frac{1}{2}}\vartheta^{-n}\varepsilon_* + \frac{3}{2}\right)E(B_{\vartheta^m \varrho}) \tag{3.46}\\
&\leqslant 2^{\mu_2 - 1}\left(c_{\frac{1}{2}}\vartheta^{-n}\varepsilon_* + \frac{3}{2}\right)\vartheta\varphi\left(\frac{\delta_*}{\vartheta^{m+1}\varrho}\right)\\
&\leqslant \varphi\left(\frac{\delta_*}{\vartheta^{m+1}\varrho}\right).
\end{aligned}
$$

Now, we prove by induction the first inequality in (3.39) for $m + 1$. From (3.39) at step k and the choices of δ_* and ϱ_* as in (3.43)–(3.44), we have

$$\frac{\Phi(\vartheta^m \varrho)}{\varphi(|(D\mathbf{u})_{\vartheta^m \varrho}|)} \leqslant \varepsilon_* < 3\varepsilon_* \leqslant \delta_1,$$

$$[H(\vartheta^m \varrho)]^{\beta_1} < 2\varepsilon_* \leqslant \delta_2,$$

and

$$\Phi_*(\vartheta^m \varrho) = \Phi(\vartheta^m \varrho) + \varphi(|(D\mathbf{u})_{\vartheta^m \varrho}|)[H(\vartheta^m \varrho)]^{\beta_1} \leqslant 3\varepsilon_*\varphi(|(D\mathbf{u})_{\vartheta^m \varrho}|).$$

Then, by virtue of lemma 3.11 and lemma 3.12 applied with radius $\vartheta^m \varrho$ in place of ϱ, and recalling the choice of ϑ (3.41), we get

$$
\begin{aligned}
\Phi(\vartheta^{m+1}\varrho) \leqslant 2c_{\mathrm{dec}}\vartheta^2 \Phi_*(\vartheta^m \varrho) &\leqslant 6c_{\mathrm{dec}}\varepsilon_*\vartheta^2\varphi(|(D\mathbf{u})_{\vartheta^m \varrho}|)\\
&\leqslant \varepsilon_*\varphi(|(D\mathbf{u})_{\vartheta^{m+1}\varrho}|).
\end{aligned}
$$

Finally, since the iteration starting from $m = 0$ of the estimate $\varphi^{-1}(E(B_{\vartheta^{m+1}\varrho})) \leqslant 2^{\frac{\mu_2}{\mu_1}}\varphi^{-1}(E(B_{\vartheta^m \varrho}))$, obtained by (3.46) and (2.2), with (3.41) yields

$$(\vartheta^m \varrho)^{1-\alpha}\varphi^{-1}(E(B_{\vartheta^m \varrho})) \leqslant \varrho^{1-\alpha}\varphi^{-1}(E(B_\varrho)) \leqslant \delta_*\varrho^{-\alpha},$$

and this estimate *a fortiori* holds if we consider $E(B_{\vartheta^m \varrho}(y))$ for $y \in B_{\varrho/2}$ in place of $E(B_{\vartheta^m \varrho})$, we deduce the Morrey-type estimate

$$r^{1-\alpha}\varphi^{-1}(E(B_r(y))) \leqslant c\delta_* \varrho^{-\alpha}$$

for all $y \in B_{\varrho/2}$ and $r \leqslant \varrho/2$, which is equivalent to (3.40). The proof is now concluded. $\qquad\square$

3.5. Excess decay estimate: the degenerate regime

In this section, with lemma 3.14 we will establish an excess improvement estimate for the degenerate case which is characterized by the fact that $\Phi(x_0, \varrho)$ is 'large' compared to $\varphi(|(D\mathbf{u})_{x_0,\varrho}|)$.

In view of lemma 3.10, this will be achieved via the φ-harmonic approximation lemma (lemma 2.11) which allows to approximate the original minimizer by a φ-harmonic function. In this way, one can transfer the a priori estimates for φ-harmonic functions (proposition 2.8) to the minimizer.

LEMMA 3.14. *Let $\gamma_0 > 0$ be the exponent of proposition 2.8. Then, for every $0 < \gamma < \gamma_0$ and every $\kappa, \mu \in (0,1)$ there exist $\varepsilon_{\#}, \tau \in (0,1)$ and $\varrho_{\#} \in (0,1]$ depending on $n, N, \mu_1, \mu_2, c_0, \beta_0, L, \nu, \gamma, \gamma_0, \mu$ and κ ($\varepsilon_{\#}$ also depends on τ and $\sigma(\delta)$, where δ satisfies (3.51) below, and $\varrho_{\#}$ also depends on ω and $\mathcal{V}$) with the following property: if*

$$\kappa\varphi(|(D\mathbf{u})_{x_0,\varrho}|) \leqslant \Phi(x_0,\varrho) \leqslant \varepsilon_{\#} \qquad (3.47)$$

for $B_\varrho(x_0) \subseteq \Omega$ with $\varrho \in (0, \varrho_{\#}]$, then

$$\Phi(x_0, \tau\varrho) \leqslant \tau^{2\gamma}\Phi(x_0, \varrho) \quad and \quad \Theta(x_0, \tau\varrho) < \mu. \qquad (3.48)$$

Proof. Without loss of generality, we assume that $x_0 = 0$ and, correspondingly, we use the abbreviations $\Phi(\varrho) = \Phi(0, \varrho)$, $\Psi(\varrho) = \Psi(0, \varrho)$, $\Theta(\varrho) = \Theta(0, \varrho)$. Let $0 < \gamma < \gamma_0$ be fixed, $\tau \in (0, \frac{1}{2^{\mu_2}}]$ to be specified later, and we set $\varepsilon := \tau^{2\gamma_0 + n}$. Furthermore, let $\delta_0 = \delta_0(n, N, \varphi, \nu, L, \varepsilon) \in (0,1]$ be the constant according to the φ-harmonic approximation (lemma 2.11) with θ the exponent of higher integrability as in (2.19).

Now, we have to check that $\mathbf{u}$ complies with (2.18), in order to apply lemma 2.11. From the Poincaré inequality, the shift change formula (2.7) with $\eta = \kappa$ and (3.47) we have

$$\Psi(\varrho) \leqslant c_P \fint_{B_\varrho} \varphi(|D\mathbf{u}|)\,dx \leqslant 2^{\mu_2 - 1}c_P(c_\kappa + 1 + \kappa^{-1})\Phi(\varrho) \leqslant c_P c(\kappa, \mu_2)\varepsilon_{\#}, \quad (3.49)$$

where $c(\kappa, \mu_2) := 2^{\mu_2 - 1}(c_\kappa + 1 + \kappa^{-1}) > 1$. An analogous computation and the concavity of φ^{-1} give

$$E(B_\varrho) := \fint_{B_\varrho} \varphi(|D\mathbf{u}|)\,dx \leqslant \varphi\left(\varphi^{-1}\left(\frac{c(\kappa, \mu_2)\varepsilon_{\#}}{\varrho}\right)\right)$$

$$\leqslant \varphi\left(\frac{(c(\kappa, \mu_2))^{\frac{1}{\mu_1}}\varphi^{-1}(\varepsilon_{\#})}{\varrho}\right),$$

whence

$$\Theta(\varrho) \leqslant (c(\kappa,\mu_2))^{\frac{1}{\mu_1}} \varphi^{-1}(\varepsilon_\#) =: \tilde{c}(\kappa,\mu_1,\mu_2)\varphi^{-1}(\varepsilon_\#). \qquad (3.50)$$

In applying lemma 3.10 we choose $\delta > 0$ (which, in turn, determines $\sigma = \sigma(\delta) > 0$ such that (3.31) holds) in such a way that

$$c_* \delta \leqslant \frac{\delta_0}{2}, \qquad (3.51)$$

where c_* is the constant of lemma 3.10. Then, we choose $\varepsilon_\# < 1$ such that

$$\varepsilon_\# \leqslant \min\left\{ \frac{1}{c_P c(\kappa,\mu_2)} \varphi\left(\frac{\delta_0 \sigma c_*}{4}\right), \varphi\left(\frac{\mu}{\tilde{c}(\kappa,\mu_1,\mu_2)}\right), \frac{\tau^n}{2c_{\frac{1}{2}}} \varphi(\mu) \right\},$$

so that, with (3.49)–(3.50), we have

$$\frac{\varphi^{-1}(\Psi(\varrho))}{\sigma} \leqslant \frac{\delta_0}{4} \quad \text{and} \quad \Theta(x_0,\varrho) \leqslant \mu.$$

We also determine a radius $\varrho_\# \in (0,1]$ according to

$$\left[\omega(\mu)^{1-\frac{1}{s}} + \mathcal{V}(\varrho_\#)^{1-\frac{1}{s}}\right]^{\beta_2} \leqslant \frac{\delta_0}{4},$$

where β_2 is defined in lemma 3.10. Recalling the definition of $\widetilde{H}(\varrho)$, for $\varrho \leqslant \varrho_\#$ we then have

$$[\widetilde{H}(\varrho)]^{\beta_2} \leqslant \frac{\delta_0}{4}.$$

For such choice of $\varepsilon_\#$ and $\varrho_\#$ it holds that

$$c_*\left(\delta + [\widetilde{H}(\varrho)]^{\beta_2} + \frac{\varphi^{-1}(\Psi(\varrho))}{\sigma}\right) \leqslant \delta_0,$$

so that, by lemma 2.11, there exists a unique φ-harmonic $\mathbf{w} \in W^{1,\varphi}(B_{\varrho/2}, \mathbb{R}^N)$ with $\mathbf{w} = \mathbf{u}$ on $\partial B_{\varrho/2}$ that satisfies

$$\left(\fint_{B_{\varrho/2}} |\mathbf{V}(D\mathbf{u}) - \mathbf{V}(D\mathbf{w})|^{2\theta} \, dx \right)^{\frac{1}{\theta}} \leqslant \tau^{2\gamma_0+n} \fint_{B_\varrho} \varphi(|D\mathbf{u}|) \, dx,$$

and

$$\fint_{B_{\varrho/2}} \varphi(|D\mathbf{w}|) \, dx \leqslant c \fint_{B_{\varrho/2}} \varphi(|D\mathbf{u}|) \, dx.$$

Taking into account the higher integrability result of lemma 3.2, lemma 2.12 implies that

$$\fint_{B_{\varrho/2}} |\mathbf{V}(D\mathbf{u}) - \mathbf{V}(D\mathbf{w})|^2 \, dx \leqslant \tau^{2\gamma_0+n} \fint_{B_\varrho} \varphi(|D\mathbf{u}|) \, dx, \qquad (3.52)$$

and since $\tau < \frac{1}{2}$, from proposition 2.8 we also have

$$\fint_{B_{\tau\varrho}} |\mathbf{V}(D\mathbf{w}) - (\mathbf{V}(D\mathbf{w}))_{\tau\varrho}|^2 \, \mathrm{d}x \leqslant c\tau^{2\gamma_0} \fint_{B_{\varrho/2}} |\mathbf{V}(D\mathbf{w}) - (\mathbf{V}(D\mathbf{w}))_{\varrho/2}|^2 \, \mathrm{d}x. \tag{3.53}$$

Thus, with (3.52) and (3.53) we infer

$$\begin{aligned}
\Phi(\tau\varrho) &\leqslant 4 \fint_{B_{\tau\varrho}} |\mathbf{V}(D\mathbf{u}) - (\mathbf{V}(D\mathbf{w}))_{\tau\varrho}|^2 \, \mathrm{d}x \\
&\leqslant 8 \fint_{B_{\tau\varrho}} |\mathbf{V}(D\mathbf{u}) - \mathbf{V}(D\mathbf{w})|^2 \, \mathrm{d}x + 8 \fint_{B_{\tau\varrho}} |\mathbf{V}(D\mathbf{w}) - (\mathbf{V}(D\mathbf{w}))_{\tau\varrho}|^2 \, \mathrm{d}x \\
&\leqslant c\tau^{-n}(\tau^{2\gamma_0+n}) \fint_{B_\varrho} \varphi(|D\mathbf{u}|) \, \mathrm{d}x + c\tau^{2\gamma_0} \fint_{B_{\varrho/2}} |\mathbf{V}(D\mathbf{w}) - (\mathbf{V}(D\mathbf{w}))_{\varrho/2}|^2 \, \mathrm{d}x \\
&\leqslant \tilde{c}_1 \tau^{2\gamma_0} \fint_{B_\varrho} \varphi(|D\mathbf{u}|) \, \mathrm{d}x
\end{aligned}$$

for some constant $\tilde{c}_1 = \tilde{c}_1(n, N, \mu_1, \mu_2, c_1) > 0$. Now, by virtue of the computation in (3.49) we conclude that

$$\Phi(\tau\varrho) \leqslant \tilde{c}_1 c(\kappa, \mu_2) \tau^{2\gamma_0} \Phi(\varrho),$$

whence (3.48) follows if we choose τ such that

$$\tau \leqslant \left(\frac{1}{\tilde{c}_1 c(\kappa, \mu_2)} \right)^{\frac{1}{2(\gamma_0 - \gamma)}}.$$

As for the second assertion in (3.48), the shift change formula (2.7) for $\eta = \frac{1}{2}$, with (3.47), (3.50) and the choice of $\varepsilon_\#$ shows that

$$\begin{aligned}
E(B_{\tau\varrho}) &:= \fint_{B_{\tau\varrho}} \varphi(|D\mathbf{u}|) \, \mathrm{d}x \\
&\leqslant 2^{\mu_2-1} \left(c_{\frac{1}{2}} \tau^{-n} \fint_{B_\varrho} \varphi_{|(D\mathbf{u})_\varrho|}(|D\mathbf{u} - (D\mathbf{u})_\varrho|) \, \mathrm{d}x + \frac{3}{2} \varphi(|(D\mathbf{u})_\varrho|) \right) \\
&\leqslant 2^{\mu_2-1} \left(c_{\frac{1}{2}} \tau^{-n} \Phi(\varrho) + \frac{3}{2} E(B_\varrho) \right) \\
&\leqslant 2^{\mu_2-1} \left(c_{\frac{1}{2}} \tau^{-n} \varepsilon_\# + \frac{3}{2} E(B_\varrho) \right) \\
&\leqslant 2^{\mu_2} \tau \varphi\left(\frac{\mu}{\tau\varrho} \right)
\end{aligned}$$

whence the assertion follows since $\tau \leqslant \frac{1}{2^{\mu_2}}$. $\qquad\square$

3.6. Proof of theorem 1.1

LEMMA 3.15. *Under the assumptions of theorem* 1.1, *let* $\alpha \in (0,1)$. *Then there exist constants* $\varepsilon_\#$, δ_* *and* $\tilde{\varrho}$ *such that the conditions*

$$\Phi(x_0, \varrho) < \varepsilon_\# \quad and \quad \Theta(x_0, \varrho) < \delta_*, \tag{3.54}$$

for $B_\varrho(x_0) \subseteq \Omega$ *with* $\varrho \in (0, \tilde{\varrho}]$ *imply*

$$\mathbf{u} \in C^{0,\alpha}(\overline{B_{\varrho/2}(x_0)}). \tag{3.55}$$

Proof. Without loss of generality, we assume that $x_0 = 0$ and, correspondingly, we omit the dependence on it. Let $\varepsilon_*, \delta_*, \vartheta \in (0,1)$ and $\varrho_* \in (0,1]$ be the constants of lemma 3.13. We then choose $\mu = \delta_*$ in lemma 3.14 leaving κ unchanged. This fixes the constants $\varepsilon_\#, \tau$ and $\varrho_\#$. We set $\tilde{\varrho} := \min\{\varrho_*, \varrho_\#\}$.

We introduce the set of integers

$$\mathbb{S} := \left\{ k \in \mathbb{N}_0 : \kappa\varphi(|(D\mathbf{u})_\varrho|) \leqslant \Phi(\tau^k \varrho) \right\},$$

and we distinguish between the cases $\mathbb{S} = \mathbb{N}_0$ and $\mathbb{S} \neq \mathbb{N}_0$.

The case $\mathbb{S} = \mathbb{N}_0$. We prove by induction that the bounds

$$\Phi(\tau^k \varrho) < \varepsilon_\# \quad and \quad \Theta(\tau^k \varrho) < \delta_* \tag{3.56}$$

hold for every $k \in \mathbb{N}_0$. The case $k = 0$ is trivial from the assumption (3.54). Now, since $k \in \mathbb{S} = \mathbb{N}_0$, the assumption (3.47) of lemma 3.14 hold with $\tau^k \varrho$ in place of ϱ. Then, an application of lemma 3.14 gives (3.56) for $k + 1$ (recall that $\tau < 1$). The validity of (3.56) implies, as in lemma 3.13, that the Morrey-type estimate

$$\Theta(y, r) \leqslant c\delta_* \left(\frac{r}{\varrho} \right)^\alpha \tag{3.57}$$

holds for every $\alpha \in (0,1)$, for all $y \in B_{\varrho/2}(x_0)$ and $r \in (0, \varrho/2]$. For $y, z \in B_{\varrho/2}$, with $|y - z| \leqslant \varrho/4$ we estimate the telescopic sum

$$\frac{|\mathbf{u}(y) - \mathbf{u}(z)|}{|y - z|} \leqslant \sum_{j \in \mathbb{Z}} \frac{1}{2^j} \fint_{B_{r_j}} \frac{|\mathbf{u}(x) - (\mathbf{u})_j|}{|y - z|} \, \mathrm{d}x \leqslant \sum_{j \in \mathbb{Z}} \frac{1}{2^j} \fint_{B_{r_j}} \frac{|\mathbf{u}(x) - (\mathbf{u})_j|}{r_j} \, \mathrm{d}x,$$

where $B_{r_j} := B_{2^{1-j}|y-z|}(y)$ for $j \geqslant 0$ and $B_{r_j} := B_{2^{1+j}|y-z|}(z)$ for $j < 0$. Now, with the Poincaré inequality we get

$$\frac{|\mathbf{u}(y) - \mathbf{u}(z)|}{|y - z|} \leqslant c \sum_{j \in \mathbb{Z}} \fint_{B_{r_j}} |D\mathbf{u}| \, \mathrm{d}x$$

$$\leqslant c \sum_{j \in \mathbb{Z}} \varphi^{-1} \left(\fint_{B_{r_j}} \varphi(|D\mathbf{u}|) \, \mathrm{d}x \right)$$

$$\leqslant c \sum_{j \geqslant 0} \frac{\Theta(y, 2^{1-j}|y-z|)}{2^{1-j}|y-z|} + c \sum_{j < 0} \frac{\Theta(z, 2^{1+j}|y-z|)}{2^{1+j}|y-z|}.$$

Finally, with the estimate (3.57) we infer

$$
|\mathbf{u}(y) - \mathbf{u}(z)| \leqslant c\delta_* |y - z|^\alpha \varrho^{-\alpha} \left(\sum_{j \geqslant 0} 2^{(\alpha-1)(1-j)} + \sum_{j < 0} 2^{(\alpha-1)(1+j)} \right) \tag{3.58}
$$
$$
\leqslant c\delta_* |y - z|^\alpha \varrho^{-\alpha},
$$

which, in particular, implies that $\mathbf{u} \in C^{0,\alpha}(\overline{B_{\varrho/2}})$.

The case $\mathbb{S} \neq \mathbb{N}_0$. In this case, there exists $k_0 := \min \mathbb{N} \backslash \mathbb{S}$. Since $k \in \mathbb{S}$ for any integer $k < k_0$ we can iterate as in the case $\mathbb{S} = \mathbb{N}_0$ for $k = 0, 1, \ldots, k_0 - 1$ to infer that (3.56) holds for any $k \leqslant k_0$. By the definition of $\mathbb{S}$ we have

$$
\Phi(\tau^{k_0}\varrho) < \kappa \varphi(|(D\mathbf{u})_\varrho|),
$$

which together with the second inequality in (3.56) with $k = k_0$ ensures that the assumptions (3.38) of lemma 3.13 are satisfied for ϱ replaced by $\tau^{k_0}\varrho$. Then, by virtue of this lemma, we have

$$
\frac{\Phi(\vartheta^m \tau^{k_0}\varrho)}{\varphi(|(D\mathbf{u})_{\vartheta^m \tau^{k_0}\varrho}|)} \leqslant \kappa \quad \text{and} \quad \Theta(\vartheta^m \tau^{k_0}\varrho) \leqslant \delta_* \tag{3.59}
$$

for every $m \in \mathbb{N}_0$.

Now, we consider an arbitrary radius $r \in (0, \varrho]$. If $r \in (\tau^{k_0}\varrho/2, \varrho]$ we find $0 \leqslant k \leqslant k_0$ such that $\tau^{k+1} \leqslant r \leqslant \theta^k$ and then we can argue as in the case $\mathbb{S} = \mathbb{N}_0$. In the case $r \in (0, \tau^{k_0}\varrho/2]$, instead, we find $m \in \mathbb{N}_0$ such that $\vartheta^{m+1}\tau^{k_0}\varrho < r \leqslant \vartheta^m \tau^{k_0}\varrho$. Then arguing as in the proof of (3.40) and taking into account the second estimate in (3.59), we have

$$
r^{1-\alpha}\varphi^{-1}(E(B_r(y))) \leqslant c(\vartheta^m \tau^{k_0}\varrho)^{1-\alpha}\varphi^{-1}(E(B_{\vartheta^m \tau^{k_0}\varrho})) \leqslant \frac{c\delta_*}{(\vartheta^m \tau^{k_0}\varrho)^\alpha}
$$

for every $y \in B_{\vartheta^m \tau^{k_0}\varrho/2} \subseteq B_{\varrho/2}$ whence

$$
\Theta(y, r) \leqslant c\frac{\delta_*}{(\vartheta^m \tau^{k_0})^\alpha} \left(\frac{r}{\varrho} \right)^\alpha.
$$

At this point, we can argue as in the case $\mathbb{S} = \mathbb{N}_0$ for the proof of (3.58), whence (3.55) follows thus concluding the proof. $\qquad\square$

Proof of theorem 1.1. Let $\varepsilon_\#$, δ_* and $\tilde{\varrho}$ be the constants of lemma 3.15. Let Σ_1 and Σ_2 be defined as in the statement of theorem 1.1. Note that, by Lebesgue's differentiation theorem, it holds that $|\Sigma_1 \cup \Sigma_2| = 0$. Thus, we are reduced to show

that each $x_0 \in \Omega \backslash (\Sigma_1 \cup \Sigma_2)$ belongs to the set

$$\Omega_0 := \left\{ z_0 \in \Omega : \mathbf{u} \in C^{0,\alpha}(U_{z_0}, \mathbb{R}^N) \text{ for every } \alpha \in (0,1) \text{ and for some } U_{z_0} \subset \Omega \right\},$$

where U_{z_0} is an open neighbourhood of z_0. For this, let $x_0 \in \Omega$ be such that both the conditions

$$\liminf_{\varrho \searrow 0} \fint_{B_\varrho(x_0)} |\mathbf{V}_{|(D\mathbf{u})_{x_0,\varrho}}(D\mathbf{u} - (D\mathbf{u})_{x_0,\varrho})|^2 \, dx = 0 \text{ and}$$

$$m_{x_0} := \limsup_{\varrho \searrow 0} |(D\mathbf{u})_{x_0,\varrho}| < +\infty \tag{3.60}$$

hold.

We set

$$\sigma := \min \left\{ \frac{1}{c_\varphi c_{\frac{1}{2}} 2^{\mu_2}} \varphi\left(\delta_*\right), \frac{\varepsilon_\#}{c_\varphi} \right\}, \tag{3.61}$$

where the constants c_φ, $c_{\frac{1}{2}}$ are specified later and, correspondingly, with (3.60) we can find a radius $\bar{\varrho}$ such that

$$\bar{\varrho} \leqslant \frac{\varphi(\delta_*)}{3 \cdot 2^{\mu_2+1} \varphi(m_{x_0} + 1)} \tag{3.62}$$

and

$$\fint_{B_{\bar{\varrho}}(x_0)} |\mathbf{V}_{|(D\mathbf{u})_{x_0,\bar{\varrho}}}(D\mathbf{u} - (D\mathbf{u})_{x_0,\bar{\varrho}})|^2 \, dx \leqslant \sigma \quad \text{and} \quad |(D\mathbf{u})_{x_0,\bar{\varrho}}| \leqslant m_{x_0} + 1. \tag{3.63}$$

Now, recalling that

$$\Phi(x_0, \bar{\varrho}) \leqslant c_\varphi \fint_{B_{\bar{\varrho}}(x_0)} |\mathbf{V}_{|(D\mathbf{u})_{x_0,\bar{\varrho}}}(D\mathbf{u} - (D\mathbf{u})_{x_0,\bar{\varrho}})|^2 \, dx$$

and observing that, as a consequence of the shift-change formula (2.7) with $\eta = \frac{1}{2}$, conditions (3.63) imply

$$\fint_{B_{\bar{\varrho}}(x_0)} \varphi(|D\mathbf{u}|) \, dx$$

$$\leqslant 2^{\mu_2-1} \left(c_{\frac{1}{2}} c_\varphi \fint_{B_{\bar{\varrho}}(x_0)} |\mathbf{V}_{|(D\mathbf{u})_{x_0,\bar{\varrho}}}(D\mathbf{u} - (D\mathbf{u})_{x_0,\bar{\varrho}})|^2 \, dx + \frac{3}{2} \varphi(|(D\mathbf{u})_{x_0,\bar{\varrho}}|) \right)$$

$$\leqslant 2^{\mu_2-1} \left(c_{\frac{1}{2}} c_\varphi \sigma + \frac{3}{2} \varphi(m_{x_0} + 1) \right) \leqslant \frac{1}{2} \varphi(\delta_*) + \frac{1}{2\bar{\varrho}} \varphi(\delta_*) \leqslant \varphi\left(\frac{\delta_*}{\bar{\varrho}} \right),$$

with the choice of σ (3.61), corresponding to the radius $\bar{\varrho} \in (0, \tilde{\varrho}]$ as in (3.62) there holds

$$\Phi(x_0, \bar{\varrho}) < \varepsilon_{\#} \quad \text{and} \quad \Theta(x_0, \bar{\varrho}) < \delta_*.$$

By the absolute continuity of the integral, we can find an open neighbourhood U_{x_0} of x_0 such that

$$\Phi(x, \bar{\varrho}) < \varepsilon_{\#} \quad \text{and} \quad \Theta(x, \bar{\varrho}) < \delta_*$$

for every $x \in U_{x_0}$. We can apply lemma 3.15 at each point of U_{x_0}, proving that $\mathbf{u} \in C^{0,\alpha}(U_{x_0}, \mathbb{R}^N)$ for every $\alpha \in (0, 1)$. Thus, $x_0 \in \Omega_0$ and the proof is concluded. $\square$

Acknowledgments

References

1   R. A. Adams. *Sobolev Spaces* (New York: Academic Press, 1975).

2   L. Beck and B. Stroffolini. Regularity results for differential forms solving degenerate elliptic systems. *Calc. Var. Partial Differential Equations* **46** (2013), 769–808.

3   C. Bennett and R. Sharpley. *Interpolation of Operators. Academic Press Inc.* vol. 129 of Pure and Applied Mathematics. (MA: Boston, 1988).

4   V. Bögelein. Partial regularity for minimizers of discontinuous quasi-convex integrals with degeneracy. *J. Differential Equations* **252** (2012), 1052–1100.

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

6   V. Bögelein, F. Duzaar, J. Habermann and C. Scheven. Stationary electro-rheological fluids: low order regularity for systems with discontinuous coefficients. *Adv. Calc. Var.* **5** (2012), 1–57.

7   P. Celada and J. Ok. Partial regularity for non-autonomous degenerate quasi-convex functionals with general growth. *Nonliear Anal.* **194** (2020), 111473.

8   F. Chiarenza, M. Frasca and P. Longo. W^2, p-solvability of the Dirichlet problem for non-divergence elliptic equations with VMO coefficients. *Trans. Amer. Math. Soc.* **336** (1993), 841–853.

9   J. Danéček and E. Viszus. C^0, γ-regularity for vector-valued minimizers of quasilinear functionals with VMO-coefficients. *Mediterr J. Math.* **12** (2015), 1287–1305.

10   C. De Filippis. On the regularity of the ω-minima of φ-functionals. *Nonlinear Anal.* **194** (2020), 111464.

11   C. De Filippis and G. Mingione. On the regularity of minima of non-autonomous functionals. *J. Geom. Anal.* **30** (2020), 1584–1626.

12   L. Diening and F. Ettwein. Fractional estimates for non-differentiable elliptic systems with general growth. *Forum Math.* **20** (2008), 523–556.

13   L. Diening, P. Kaplický and S. Schwarzacher. BMO estimates for the p-Laplacian. *Nonlinear Anal.* **75** (2012), 637–650.

14   L. Diening and C. Kreuzer. Linear convergence of an adaptive finite element method for the p-laplacian equation. *SIAM J. Numer. Anal.* **46** (2008), 614–638.

15   L. Diening, B. Stroffolini and A. Verde. Everywhere regularity of functionals with φ-growth. *Manuscr. Math.* **129** (2009), 449–481.

16   L. Diening, B. Stroffolini and A. Verde. The φ-harmonic approximation and the regularity of φ-harmonic maps. *J. Differential Equations* **253** (2012), 1943–1958.

17   L. Diening, D. Lengeler, B. Stroffolini and A. Verde. Partial regularity for minimizers of quasiconvex functionals with general growth. *SIAM J. Math. Anal.* **44** (2012), 3594–3616.

18   G. Di Fazio and M. A. Ragusa. Interior estimates in Morrey spaces for strong solutions to nondivergence form equations with discontinuous coefficients. *J. Funct. Anal.* **112** (1993), 241–256.

19   G. Di Fazio, D. K. Palagachev and M. A. Ragusa. Global Morrey regularity of strong solutions to the Dirichlet problem for elliptic equations with discontinuous coefficients. *J. Funct. Anal.* **166** (1999), 179–196.

20   F. Duzaar, A. Gastel and G. Mingione. Elliptic systems, singular sets and Dini continuity. *Comm. Partial Differential Equations* **29** (2004), 1215–1240.

21   F. Duzaar and G. Mingione. Regularity for degenerate elliptic problems via p-harmonic approximation. *Ann. Inst. H. Poincaré Anal. Non. Linéaire* **21** (2004), 735–766.

22   F. Duzaar and G. Mingione. The p-harmonic approximation and the regularity of p-harmonic maps. *Calc. Var. Partial Differential Equations* **20** (2004), 235–256.

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

24   F. W. Gehring. The L^p-integrability of the partial derivatives of a quasiconformal mapping. *Acta Math.* **130** (1973), 265–277.

25   E. Giusti. *Direct Methods in the Calculus of Variations*, First Edition (World Scientific, 2003).

26   C. S. Goodrich. Partial Hölder continuity of minimizers of functionals satisfying a VMO condition. *Adv. Calc. Var.* **10** (2017), 83–110.

27   T. Iwaniec and G. Martin. *Geometric Function Theory and Nonlinear Analysis*, Oxford Mathematical Monographs (Oxford: Clarendon Press).

28   M. A. Krasnosel'skií and Ya. B. Rutickií. *Convex Functions and Orlicz Spaces* (The Netherlands: P. Noordhoff LTD. - Groningen, 1961).

29   J. Kristensen and G. Mingione. The singular set of minima of integral functionals. *Arch. Rational Mech. Anal.* **180** (2006), 331–398.

30   J. Kristensen and G. Mingione. The singular set of Lipschitzian minima of multiple integrals. *Arch. Rational Mech. Anal.* **184** (2007), 341–369.

31   A. Kufner, O. John and S. Fucik. *Function Spaces* (The Netherlands: Springer, 1977).

32   P. Marcellini. Regularity for elliptic equations with general growth conditions. *J. Differ. Equations* **105** (1993), 296–333.

33   P. Marcellini. Everywhere regularity for a class of elliptic systems without growth conditions. *Ann. Sc. Norm. Super. Pisa, Cl. Sci., IV. Ser.* **23** (1996), 1–25.

34   P. Marcellini and G. Papi. Nonlinear elliptic systems with general growth. *J. Differ. Equations* **221** (2006), 412–443.

35   G. Mingione. Regularity of minima: an invitation to the dark side of the calculus of variations. *Appl. Math.* **51** (2006), 355–425.

36   J. Ok. Regularity of ω-minimizers for a class of functionals with non-standard growth. *Calc. Var.* **56** (2017), 48.

37   D. Sarason. Functions of vanishing mean oscillation. *T.A.M.S.* **207** (1975), 391–405.

38   M. Ragusa and A. Tachikawa. Partial regularity of the minimizers of quadratic functionals with VMO coefficients. *J. Lond. Math. Soc.* **2** (2005), 609–620.

39   E. Stein. *Harmonic Analysis: Real-Variable Methods and Oscillatory Integrals* (Princeton University Press, Volume 43 Orthogonality, 1993).

40   B. Stroffolini. Partial regularity results for quasimonotone elliptic systems with general growth. *Z. Anal. Anwend.* **39** (2020), 315–347.

41   K. Uhlenbeck. Regularity for a class of nonlinear elliptic systems. *Acta. Math.* **138** (1977), 219–240.