Calc. Var. (2023) 62:64
https://doi.org/10.1007/s00526-022-02402-5

Calculus of Variations

Wolff potentials and measure data vectorial problems with Orlicz growth

Iwona Chlebicka[1] · Yeonghun Youn[2] · Anna Zatorska-Goldstein[1]

Received: 1 December 2021 / Accepted: 2 December 2022 / Published online: 7 January 2023

Abstract
We study solutions to measure data elliptic systems with Uhlenbeck-type structure that involve operator of divergence form, depending continuously on the spacial variable, and exposing doubling Orlicz growth with respect to the second variable. Pointwise estimates for the solutions that we provide are expressed in terms of a nonlinear potential of generalized Wolff type. Not only we retrieve the recent sharp results proven for p-Laplace systems, but additionally our study covers the natural scope of operators with similar structure and natural class of Orlicz growth.

Mathematics Subject Classification 35B45 · 35J47

Contents

Communicated by A. Mondino.

✉ Iwona Chlebicka
i.chlebicka@mimuw.edu.pl

Yeonghun Youn
yeonghunyoun@yu.ac.kr

Anna Zatorska-Goldstein
azator@mimuw.edu.pl

1 Institute of Applied Mathematics and Mechanics, University of Warsaw, ul. Banacha 2, 02-097 Warsaw, Poland

2 Department of Mathematics, Yeungnam University, Gyeongsan, Gyeongbuk 38541, Republic of Korea

1 Introduction

A broad and profuse branch of the theory of nonlinear partial differential systems stems from seminal ideas by Ural'tseva [85] and Uhlenbeck [84]. There is a solid stream of recent studies on regularity of solutions to general growth elliptic systems [4, 25, 33, 38, 67, 72, 81] and related minimizers of vectorial functionals [7, 15, 30, 32, 35, 36, 47, 73]. Our aim is to contribute to the field by providing pointwise estimates for very weak solutions to measure data problems in the terms of potentials of relevant Orlicz growth. As a consequence, we infer sharp description of fine properties of solutions being the exact analogues of the ones available in the classical linear potential theory or in the case of p-Laplace systems [65].

Let us stress that weak solutions do not have to exist for arbitrary measure datum. Thus we employ a notion of very weak solutions obtained by an approximation studied since [9]. Despite they can be unbounded, they can be controlled by a certain potential. There are known deep classical results for scalar problems [57, 58] settling the nonlinear potential theory for solutions to $-\Delta_p u = -\operatorname{div}(|Du|^{p-2}Du) = \mu$, $1 < p < \infty$, followed by their variable exponent version [68] and recently by the Orlicz one [22, 70], as well as a counterpart proven for systems involving p-Laplace operator [27, 65]. We study the nonstandard growth version of pointwise estimates involving suitably generalized potential of the Wolff type and infer their regularity consequences.

In fact, we investigate very weak solutions $\mathbf{u} : \Omega \to \mathbb{R}^m$ to measure data elliptic systems involving nonlinear operators

$$-\,\mathbf{div}\left(a(x)\frac{g(|D\mathbf{u}|)}{|D\mathbf{u}|}D\mathbf{u}\right) = \boldsymbol{\mu} \quad \text{in} \quad \Omega, \tag{1}$$

where $a \in C(\Omega)$ is bounded and separated from zero, $g = G'$, $\boldsymbol{\mu} \in \mathcal{M}(\Omega, \mathbb{R}^m)$ is a bounded measure, whereas **div** stands for the $\mathbb{R}^m$-valued divergence operator. Here we admit G to be a Young function satisfying both Δ_2- and ∇_2-conditions, which follows from the condition

$$2 \le i_G = \inf_{t>0} \frac{tg(t)}{G(t)} \le \sup_{t>0} \frac{tg(t)}{G(t)} = s_G < \infty. \tag{2}$$

See *Assumption* **(A-vect)** in Sect. 2 for all details. We stress that we impose the typical assumption of quasi-diagonal structure of $\mathcal{A}$ naturally covering the case of (possibly weighted) p-Laplacian when $G_p(s) = s^p$, for every $p \ge 2$, together with operators governed by the Zygmund-type functions $G_{p,\alpha}(s) = s^p \log^\alpha(1+s)$, $p \ge 2$, $\alpha \in \mathbb{R}$, as well as their multiplications and compositions with various parameters. In turn, we generalize the corresponding results of [27, 65] to embrace also p-Laplace systems with continuous coefficients, and – on the other hand – cover the natural scope of operators with similar structure and Orlicz growth. Let us stress that the operator we consider is *not* assumed to enjoy homogeneity of a form $\mathcal{A}(x, k\xi) = |k|^{p-2}k\mathcal{A}(x, \xi)$. Consequently, our class of solutions is *not* invariant with respect to scalar multiplication.

Obtaining sharp regularity results for solutions to nonlinear systems is particularly challenging. In the scalar case one can infer continuity or Hölder continuity of the solution to $-\mathrm{div}\mathcal{A}(x, Du) = \mu$ when the dependence of the operator on the spacial variable is merely bounded and measurable and the growth of $\mathcal{A}$ with respect to the second variable is governed by an arbitrary doubling Young function (cf. [22]). The same is not possible for systems even with far less complicated growth and null datum, cf. [29, 55, 82] and [75, Sect. 3]. To justify why we restrict our attention to operators having a specific form as in (1), let us point out that the typical assumption of a so-called Uhlenbeck structure is imposed in order to control energy of solutions. The continuity of the coefficients is a minimal assumption to get continuity of the solution in the view of counterexample of [29].

The studies on the potential theory to measure data problems dates back to [52, 74]. We refer to [62] for an overview of the nonlinear potential theory, to [54, 57, 58, 69] for cornerstones of the field, and [1, 53] for well-present background of the p-growth case. In the scalar case, in their seminal works [57, 58], Kilpeläinen and Malý provided optimal Wolff potential estimates for p-superharmonic functions u generating a nonnegative measure μ from above and below

$$\frac{1}{c}\mathcal{W}^{\mu}_{p}(x_0,R) \le u(x_0) \le c\Big(\inf_{B_R(x_0)} u + \mathcal{W}^{\mu}_{p}(x_0,R)\Big) \quad \text{for some } \ c = c(n,p) \tag{3}$$

with the so-called Wolff potential

$$\mathcal{W}^{\mu}_{p}(x_0,R) = \int_0^R \left(r^{p-n}\mu(B_r(x_0))\right)^{\frac{1}{p-1}}\,\frac{dr}{r} = \int_0^R \left(\frac{\mu(B_r(x_0))}{r^{n-1}}\right)^{\frac{1}{p-1}} dr,$$

see also [60, 83]. In the linear case ($p = 2$) the estimates of (3) become the classical Riesz potential bounds. The precise Orlicz counterpart of this result with nonnegative measure μ is proven with the nonstandard growth potential

$$\mathcal{W}^{\mu}_{G}(x_0,R) = \int_0^R g^{-1}\left(\frac{\mu(B_r(x_0))}{r^{n-1}}\right) dr,$$

see [22, 70]. To our best knowledge, the only reference one can find on the related results for systems are [27, 65] that involves problems with the p-Laplace operator. The estimate related to (3) provided therein establishes the upper bound only. Note however that no lower bound can be available in the vectorial case. Indeed, it origins in the lack of the possibility of proving maximum principle. Our method of proof relies on the ideas of [65]. We employ a properly adapted Orlicz version of $\mathcal{A}$-harmonic approximation relevant for measure data problems (Theorem 4.1 in Sect. 4) and careful estimates on concentric balls. The proof of Theorem 4.1 is based on the arguments that essentially apply the Uhlebeck structure of the system and the lower growth restriction (2). By the very nature of the result one cannot hope for $\mathcal{A}$-harmonic approximation of functions that do not have at least Sobolev regularity. In order to ensure that a distributional solution to a measure data problem belongs at least to $W^{1,1}_{loc}$, the operator need to be far from 1-Laplacian. Note that already in the case of scalar p-Laplace equation, the value $p = 2-\frac{1}{n}$ is the integrability threshold for a fundamental solution. Indeed, for $p \neq n$ a function $u(x) = c(|x|^{\frac{p-n}{n-1}}-1)$, being a solution to $-\Delta_p u = \delta_0$ on a ball, has locally integrable distributional gradient only for $p > 2-\frac{1}{n}$. Similarly, one can show the natural Orlicz counterpart of this threshold. If this condition is violated, a fundamental part of the proof of Theorem 4.1, i.e. the estimate on summability of $\mathbf{u}$ and $D\mathbf{u}$, is false. Note that Remark 4.2 explains that the proof of Theorem 4.1 works for functions G having i_G slightly below 2. Other available proofs of summability estimates for solutions to similar problems with Orlicz growth also do not cover the natural range reflecting $p > 2-\frac{1}{n}$, see [4]. We stress that rest of the proofs, including those of pointwise Wolff potential estimate (Theorem 2.1), continuity criterium (Theorem 2.3) and Hölder continuity criterium (Theorem 2.9), do not make use of other growth assumptions than $G \in \Delta_2 \cap \nabla_2$. Nonetheless, all of them substantially rely on measure data $\mathcal{A}$-harmonic approximation result.

Let us also mention that a lot of attention is attracted by potential estimates on gradients of solutions, generalized harmonic approximation, and their application in the theory of partial regularity [5, 6, 10–13, 16, 17, 19, 34, 37, 41–44, 62, 64, 71, 76] which is an open path from now on.

The paper is organized as follows. Our assumptions, main results and their regularity consequences are presented in Sect. 2. Sect. 3 is devoted to notation and information on the setting. In particular see Sect. 3.5 for the precise definition of the notion of very weak solutions we employ. Sect. 4 provides the most important tool of paper—a measure data $\mathcal{A}$-harmonic approximation. Sect. 5 contains the proofs of comparison estimates, the sufficient condition for $\mathbf{u}$ to be in VMO of Proposition 2.2, the potential estimates of Theorem 2.1, the continuity criterion of Theorem 2.3 and the Hölder continuity criterion of Theorem 2.9.

2 Main result and its consequences

2.1 The statement of the problem

Let us present an essential notation and details of the measure data problem we study.

Essential notation. By '$\cdot$' we denote the scalar product of two vectors, i.e. for $\boldsymbol{\xi} = (\xi_1, \dots, \xi_m) \in \mathbb{R}^m$ and $\boldsymbol{\eta} = (\eta_1, \dots, \eta_m) \in \mathbb{R}^m$ we have $\boldsymbol{\xi} \cdot \boldsymbol{\eta} = \sum_{i=1}^m \xi_i \eta_i$; by '$:$' – the Frobenius product of the second-order tensors, i.e. for $\xi = [\xi_j^\alpha]_{j=1,\dots,n,\, \alpha=1,\dots,m}$ and $\eta = [\eta_j^\alpha]_{j=1,\dots,n,\, \alpha=1,\dots,m}$ we have

$$\xi : \eta = \sum_{\alpha=1}^{m} \sum_{j=1}^{n} \xi_j^\alpha \eta_j^\alpha.$$

By '$\otimes$' we denote the tensor product of two vectors, i.e for $\boldsymbol{\xi} = (\xi_1, \dots, \xi_k) \in \mathbb{R}^k$ and $\boldsymbol{\eta} = (\eta_1, \dots, \eta_\ell) \in \mathbb{R}^\ell$, we have $\boldsymbol{\xi} \otimes \boldsymbol{\eta} := [\xi_i \eta_j]_{i=1,\dots,k,\, j=1,\dots,\ell}$, that is

$$\boldsymbol{\xi} \otimes \boldsymbol{\eta} := \begin{pmatrix} \xi_1\eta_1 & \xi_1\eta_2 & \cdots & \xi_1\eta_\ell \\ \xi_2\eta_1 & \xi_2\eta_2 & \cdots & \xi_2\eta_\ell \\ \vdots & \vdots & & \vdots \\ \xi_k\eta_1 & \xi_k\eta_2 & \cdots & \xi_k\eta_\ell \end{pmatrix} \in \mathbb{R}^{k\times\ell}.$$

Assumption **(A-vect)**. Given a bounded, open, Lipschitz set $\Omega \subset \mathbb{R}^n$, $n \geq 2$, we investigate solutions $\mathbf{u} : \Omega \to \mathbb{R}^m$ to the problem

$$\begin{cases} -\mathbf{div}\mathcal{A}(x, D\mathbf{u}) = \boldsymbol{\mu} & \text{in } \Omega, \\ \mathbf{u} = 0 & \text{on } \partial\Omega \end{cases} \tag{4}$$

with a datum $\boldsymbol{\mu}$ being a vector-valued bounded Radon measure and a function $\mathcal{A} : \Omega \times \mathbb{R}^{n\times m} \to \mathbb{R}^{n\times m}$ is a weighted operator of Orlicz growth expressed by the means of $g(t) := G'(t)$, where an N-function $G \in C^2((0,\infty)) \cap C([0,\infty))$ satisfies $i_G \geq 2$ with i_G given by (2). Let $g \in \Delta_2 \cap \nabla_2$. Namely, $\mathcal{A}$ is assumed to admit a form

$$\mathcal{A}(x, \xi) = a(x)\frac{g(|\xi|)}{|\xi|}\,\xi, \tag{5}$$

where $a : \Omega \to [c_a, C_a]$, $0 < c_a < C_a$ is a continuous function with a modulus of continuity ω_a. We define a potential

$$\mathcal{W}_G^{\mu}(x_0, R) = \int_0^R g^{-1}\left(\frac{|\boldsymbol{\mu}|(B_r(x_0))}{r^{n-1}}\right) dr. \tag{6}$$

For the case of referring to the dependence of some quantities on the parameters of the problem, we collect them as

$$\texttt{data} = \texttt{data}(i_G, s_G, c_a, C_a, \omega_a, n, m).$$

Having (5), one can infer the strong monotonicity of the vector field $\mathcal{A}$ of a form given by Lemma 3.2.

Let us define the notion of very weak solutions we employ. Let us define the notion of very weak solutions we employ, cf. [23].

Solutions Obtained as a Limit of Approximation. A map $\mathbf{u} \in W_0^{1,1}(\Omega, \mathbb{R}^m)$ such that $\int_\Omega g(|D\mathbf{u}|)\,dx < \infty$ is called a SOLA to (4) under the regime of Assumption **(A-vect)**, if there exists a sequence $(\mathbf{u}_h) \subset W^{1,G}(\Omega, \mathbb{R}^m)$ of local energy solutions to the systems

$$-\mathbf{div}\mathcal{A}(x, D\mathbf{u}_h) = \boldsymbol{\mu}_h$$

such that $\mathbf{u}_h \to \mathbf{u}$ locally in $W^{1,1}(\Omega, \mathbb{R}^m)$ and $(\boldsymbol{\mu}_h) \subset L^\infty(\Omega, \mathbb{R}^m)$ is a sequence of maps that converges to $\boldsymbol{\mu}$ weakly in the sense of measures and satisfies

$$\limsup_h |\boldsymbol{\mu}_h|(B) \le |\boldsymbol{\mu}|(B) \tag{7}$$

for every ball $B \subset \Omega$.

Observe that the above approximation property immediately implies that a SOLA $\mathbf{u}$ is a distributional solution to (4), that is,

$$\int_\Omega \mathcal{A}(x, D\mathbf{u}) : D\boldsymbol{\varphi}\,dx = \int \boldsymbol{\varphi}\,d\boldsymbol{\mu} \qquad \text{for every } \boldsymbol{\varphi} \in C^\infty(\Omega, \mathbb{R}^m).$$

2.2 Main results

Our main accomplishment reads as follows.

Theorem 2.1 (Pointwise Wolff potential estimates) *Suppose* $\mathbf{u} : \Omega \to \mathbb{R}^m$ *is a SOLA to* (4) *with* $\mathcal{A}$ *satisfying* Assumption **(A-vect)** *and* $\boldsymbol{\mu} \in \mathcal{M}(\Omega, \mathbb{R}^m)$. *Let* $B_r(x_0) \Subset \Omega$ *with* $r < R_0$ *for some* $R_0 = R_0(\texttt{data})$. *If* $\mathcal{W}_G^{\boldsymbol{\mu}}(x_0, r)$ *is finite, then* x_0 *is a Lebesgue's point of* $\mathbf{u}$ *and*

$$|\mathbf{u}(x_0) - (\mathbf{u})_{B_r(x_0)}| \le C_{\mathcal{W}} \left(\mathcal{W}_G^{\boldsymbol{\mu}}(x_0, r) + \fint_{B_r(x_0)} |\mathbf{u} - (\mathbf{u})_{B_r(x_0)}|\,dx \right) \tag{8}$$

holds for $C_{\mathcal{W}} > 0$ *depending only on* $\texttt{data}$. *In particular, we have the following pointwise estimate*

$$|\mathbf{u}(x_0)| \le C_{\mathcal{W}} \left(\mathcal{W}_G^{\boldsymbol{\mu}}(x_0, r) + \fint_{B_r(x_0)} |\mathbf{u}(x)|dx \right). \tag{9}$$

2.3 Local behaviour of very weak solutions

Potential estimates are known to be efficient tools to bring precise information on the local behaviour of solutions. We refer to [62] for clearly presented overview of consequences of estimates like (9) in studies on p-superharmonic functions together with a bunch of related references and to [22] for similar results for $\mathcal{A}$-harmonic functions where the operator $\mathcal{A}$ is exhibiting Orlicz type of growth. Notice however we referred to the scalar case. In the vectorial one, the only investigations on the potential estimates to solutions to measure data problem we are aware of are available in [27, 65] for p-Laplace systems. Let us present the regularity consequences of Wolff potential estimates to p-Laplace system with continuous coefficients and the natural scope of operators with similar structure and Orlicz growth.

We start with finding a density condition around a point x_0 implying that a solution has vanishing mean oscillations at x_0. The proposition below does not follow from Theorem 2.1, for the proof see Sect. 5.

Proposition 2.2 (VMO criterion) *Suppose $\mathcal{A}$ satisfies* Assumption **(A-vect)** *and $\boldsymbol{\mu} \in \mathcal{M}(\Omega, \mathbb{R}^m)$. Let $\mathbf{u}$ be a SOLA to (4) and let $B_r(x_0) \Subset \Omega$. If*

$$\lim_{\varrho \to 0} \varrho\, g^{-1}\left(\frac{|\boldsymbol{\mu}|(B_\varrho(x_0))}{\varrho^{n-1}}\right) = 0, \tag{10}$$

then $\mathbf{u}$ has vanishing mean oscillations at x_0, i.e.

$$\lim_{\varrho \to 0} \fint_{B_\varrho(x_0)} |\mathbf{u} - (\mathbf{u})_{B_\varrho(x_0)}|\, dx = 0. \tag{11}$$

An application of Theorem 2.1 is the following continuity criterion proven also in Sect. 5.

Theorem 2.3 (Continuity criterion) *Suppose $\mathbf{u}$ is a SOLA to (4) under the regime of* Assumption **(A-vect)** *and $B_r(x_0) \Subset \Omega$. If*

$$\lim_{\varrho \to 0} \sup_{x \in B_r(x_0)} \mathcal{W}_G^{\boldsymbol{\mu}}(x, \varrho) = 0, \tag{12}$$

then $\mathbf{u}$ is continuous in $B_r(x_0)$.

If $\boldsymbol{\mu} = 0$, then $\mathcal{W}_G^{\boldsymbol{\mu}}(x, \varrho) = 0$, thus trivially we have the following consequence.

Corollary 2.4 *Under* Assumption **(A-vect)** *if $\mathbf{u}$ is an $\mathcal{A}$-harmonic map in Ω, then $\mathbf{u}$ is continuous in every $\Omega' \Subset \Omega$.*

Condition (12) holds true provided the datum belongs to a Lorentz-type space. In order to define it we recall some definitions. We denote by $f^\star$ the decreasing rearrangement of a measurable function $f : \Omega \to \mathbb{R}$ by

$$f^\star(t) = \sup\{s \geq 0 \colon |\{x \in \mathbb{R}^n : f(x) > s\}| > t\},$$

the maximal rearrangement by

$$f^{\star\star}(t) = \frac{1}{t} \int_0^t f^\star(s)\, ds \quad \text{and} \quad f^{\star\star}(0) = f^\star(0).$$

Following [80] by Lorentz space $L(\alpha, \beta)(\Omega)$ for $\alpha, \beta > 0$ we mean the class of measurable functions such that

$$\int_0^\infty \left(t^{1/\alpha} f^{\star\star}(t)\right)^\beta \frac{dt}{t} < \infty.$$

The following fact is proven in Appendix.

Lemma 2.5 *Suppose $\boldsymbol{\mu} = \mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$ is a locally integrable map vanishing outside Ω, then there exists a constant $c = c(n, i_G, s_G) > 0$, such that*

$$\mathcal{W}_G^{\mathbf{F}}(x, R) \leq c \int_0^{|B_R|} t^{\frac{1}{n}} g^{-1}\left(t^{\frac{1}{n}} |\mathbf{F}|^{\star\star}(t)\right) \frac{dt}{t} =: \mathcal{I}_R.$$

Note that if $\mathcal{I}_R < \infty$, the datum $\mathbf{F}$ is in the dual space to $W^{1,G}(B_R)$ and, consequently, we deal with a weak solution, see [3]. Let us present a corollary of Theorem 2.3 holding due to Lemma 2.5.

Corollary 2.6 *If* $\mathbf{u}$ *is a weak solution to* $-\mathbf{div}\mathcal{A}(x, D\mathbf{u}) = \mathbf{F}$ *with* $\mathcal{A}$ *satisfying* Assumption **(A-vect)** *and* $\mathbf{F} : \Omega \to \mathbb{R}^m$ *such that*

$$\mathcal{I} := \int_0^{|\Omega|} t^{\frac{1}{n}} g^{-1}\left(t^{\frac{1}{n}} |\mathbf{F}|^{\star\star}(t)\right) \frac{dt}{t} < \infty \tag{13}$$

for $\Omega_0 \Subset \Omega$, *then* $\mathbf{u} \in C(\Omega_0, \mathbb{R}^m)$ *and* $\|\mathbf{u}\|_{L^\infty(\Omega_0,\mathbb{R}^m)} \le c(\mathtt{data})\mathcal{I}$. *This bound is optimal and attained by a radial solution on a ball, see* [3]. *Moreover, as a special case we get that* $\mathbf{u}$ *is continuous under the regularity restrictions on* $|\mathbf{F}| = f$ *of [3, Example 1 (A) and Example 2 (A)], still within our regime requiring* $g \in \Delta_2 \cap \nabla_2$ *and* $i_G \ge 2$.

The above corollary results in the following extension of [65, Theorem 10.6] to the weighted case.

Remark 2.7 If $\mathbf{u}$ is a weak solution to $-\mathbf{div}\left(a(x)|D\mathbf{u}|^{p-2}D\mathbf{u}\right) = \mathbf{F}$ for $p \ge 2$ and $0 < a \in C(\Omega)$ is separated from zero and $|\mathbf{F}|$ belongs locally to the Lorentz space $L(\frac{n}{p}, \frac{1}{p-1})(\Omega)$ then $\mathbf{u}$ is continuous in Ω.

As another application of Corollary 2.6 let us present its consequences for the Zygmund case.

Remark 2.8 Suppose that $2 \le p < n$, $\alpha \ge 0$, $0 < a \in C(\Omega)$ separated from zero is bounded, and $\mathbf{u}$ is a weak solution to

$$-\,\mathbf{div}\left(a(x)|D\mathbf{u}|^{p-2}\log^\alpha(\mathrm{e} + |D\mathbf{u}|)\, D\mathbf{u}\right) = \mathbf{F}. \tag{14}$$

Observe that in this case $g^{-1}(\lambda) \approx \lambda^{\frac{1}{p-1}} \log^{-\frac{\alpha}{p-1}}(\mathrm{e} + \lambda)$. If $\mathbf{F}$ satisfies (13), then $\mathbf{u}$ is continuous in Ω.

Wolff potential estimates can be used to find a relevant condition on a measure $\boldsymbol{\mu}$ to infer Hölder continuity of solutions. One of the natural ones is expressed in the Orlicz modification of the Morrey-type scale.

Theorem 2.9 (Hölder continuity criterion) *Suppose* $\mathbf{u}$ *is a SOLA to* (4) *under the regime of* Assumption **(A-vect)**. *Assume further that for* $\boldsymbol{\mu}$ *there exist positive constants* $c = c(\mathtt{data}) > 0$ *and* $\theta \in (0, 1)$ *such that*

$$|\boldsymbol{\mu}|(B_r(x)) \le c r^{n-1} g(r^{\theta-1}) \approx r^{n-\theta} G(r^{\theta-1}) \tag{15}$$

for each $B_r(x) \Subset \Omega$ *with sufficiently small radius. Then* $\mathbf{u}$ *is locally Hölder continuous in* Ω.

For p-growth problems condition (15) reads as $|\boldsymbol{\mu}|(B_r(x)) \le c r^{n-p+\theta(p-1)}$ well known since [14, 57, 59, 78]. Moreover, in the scalar case (15) is proven in [22] to be equivalent to Hölder continuity of solutions, while in [20] to characterize removable sets for Hölder continuous solutions. In the vectorial case we cannot get equivalence, because by Theorem 2.1 we are equipped with one-sided estimate only. Specializing Theorem 2.9, we have the following results.

Remark 2.10 Suppose $p \ge 2$, positive $a \in C(\Omega)$ is separated from zero, and $\mathbf{u}$ is a SOLA to

$$-\mathbf{div}\left(a(x)|D\mathbf{u}|^{p-2}D\mathbf{u}\right) = \boldsymbol{\mu}$$

with $|\boldsymbol{\mu}|(B_r(x)) \le c r^{n-p+\theta(p-1)}$ for some $c > 0$, $\theta \in (0, 1)$ and all sufficiently small $r > 0$. Then $\mathbf{u}$ is locally Hölder continuous in Ω.

Remark 2.11 Suppose $p \geq 2$, $\alpha \in \mathbb{R}$, positive $a \in C(\Omega)$ is separated from zero, and $\mathbf{u}$ is a SOLA to

$$-\,\mathbf{div}\,\big(a(x)|D\mathbf{u}|^{p-2}\log^{\alpha}(\mathrm{e}+|D\mathbf{u}|)D\mathbf{u}\big)=\boldsymbol{\mu} \tag{16}$$

with $|\boldsymbol{\mu}|(B_r(x)) \leq cr^{n-p+\theta(p-1)}\log^{\alpha}(\mathrm{e}+r^{\theta-1})$ for some $c>0$, $\theta \in (0,1)$ and all sufficiently small $r>0$. Then $\mathbf{u}$ is locally Hölder continuous in Ω.

The sufficient condition for (15) and, in turn, for the Hölder continuity of the solution is to assume that $|\boldsymbol{\mu}| = |\mathbf{F}|$ belongs to a relevant Marcinkiewicz-type space. Following [77] for a continuous increasing function $\psi : (0, |\Omega|) \to (0, \infty)$ we say that $f \in L(\psi, \infty)(\Omega)$ if the maximal rearrangement $f^{\star\star}$ of f satisfies

$$\sup_{s\in(0,|\Omega|)} \frac{f^{\star\star}(s)}{\psi^{-1}(1/s)} < \infty.$$

There holds the following consequence of Theorem 2.9.

Corollary 2.12 *Suppose* $\mathbf{u}$ *is a SOLA to* $-\mathbf{div}\mathcal{A}(x, D\mathbf{u}) = \mathbf{F}$ *under the regime of* Assumption **(A-vect)** *and* $|\mathbf{F}|$ *belongs locally to the Marcinkiewicz-type space* $L(\psi, \infty)(\Omega)$ *with* $\psi^{-1}(1/\lambda) = \lambda^{-\frac{1}{n}} g\big(\lambda^{\frac{\theta-1}{n}}\big)$ *for some* $\theta \in (0,1)$*, then* $\mathbf{u}$ *is locally Hölder continuous.*

For justification that indeed under the assumptions of Corollary 2.12 the condition (15) is satisfied see the calculations provided for the scalar case [22, Sect. 2]. The above fact has the best possible consequence in the p-Laplace case.

Remark 2.13 If $p \geq 2$, positive $a \in C(\Omega)$ is separated from zero, $\mathbf{u}$ is a SOLA to $-\mathbf{div}\,\big(a(x)|D\mathbf{u}|^{p-2}D\mathbf{u}\big) = \mathbf{F}$ and $|\mathbf{F}|$ belongs locally to the Marcinkiewicz space $L(\frac{n}{p+\theta(p-1)}, \infty)(\Omega)$ for some $\theta \in (0,1)$, i.e. $\sup_{\lambda>0}\Big(\lambda^{\frac{n}{p+\theta(p-1)}}\big|\{x \in \Omega_0 : |\mathbf{F}(x)| > \lambda\}\big|\Big) < \infty$ for any $\Omega_0 \Subset \Omega$, then $\mathbf{u}$ is locally Hölder continuous in Ω.

Remark 2.14 When $G(t) \approx t^p \log^{\alpha}(\mathrm{e}+t)$, $p \geq 2$, $\alpha \in \mathbb{R}$, $\mathbf{u}$ is a SOLA to (14) with $\mathbf{F}$ such that

$$\sup_{\lambda>0}\left(\lambda^{\frac{n}{p+\theta(p-1)}}\log^{-\frac{\alpha(1-\theta)}{p+\theta(p-1)}}(\mathrm{e}+\lambda^{\frac{n}{1-\theta}})\big|\{x \in \Omega_0 : |\mathbf{F}(x)| > \lambda\}\big|\right) < \infty,$$

for any $\Omega_0 \Subset \Omega$, then $\mathbf{u}$ is locally Hölder continuous in Ω.

3 Preliminaries

3.1 Notation

We shall adopt the customary convention of denoting by c a constant that may vary from line to line. To skip rewriting a constant, we use $\lesssim$. By $a \approx b$, we mean $a \lesssim b$ and $b \lesssim a$. To stress the dependence of the intrinsic constants on the parameters of the problem, we write $\lesssim_{\mathtt{data}}$ or $\approx_{\mathtt{data}}$. By B_R we denote a ball skipping prescribing its center, when it is not important. By $cB_R = B_{cR}$ we mean a ball with the same center as B_R, but with rescaled radius cR. We make use of the truncation operator, $T_k : \mathbb{R}^m \to \mathbb{R}^m$, defined as follows

$$T_k(\boldsymbol{\xi}) := \min\left\{1, \frac{k}{|\boldsymbol{\xi}|}\right\}\boldsymbol{\xi}. \tag{17}$$

Then, of course, $DT_k : \mathbb{R}^m \to \mathbb{R}^{m\times m}$ is given by

$$DT_k(\boldsymbol{\xi}) = \begin{cases} \mathsf{Id} & \text{if } |\boldsymbol{\xi}| \le k, \\ \frac{k}{|\boldsymbol{\xi}|}\left(\mathsf{Id} - \frac{\boldsymbol{\xi}\otimes\boldsymbol{\xi}}{|\boldsymbol{\xi}|^2}\right) & \text{if } |\boldsymbol{\xi}| > k. \end{cases} \tag{18}$$

For a measurable set $U \subset \mathbb{R}^n$ with finite and positive n-dimensional Lebesgue measure $|U| > 0$ and $\mathbf{f}\colon U \to \mathbb{R}^k$, $k \ge 1$ being a measurable map, we define

$$(\mathbf{f})_U = \fint_U \mathbf{f}(x)\,dx = \frac{1}{|U|}\int_U \mathbf{f}(x)\,dx.$$

By $C^{0,\gamma}(U)$, $\gamma \in (0,1]$, we mean the family of Hölder continuous functions, i.e. measurable functions $f\colon U \to \mathbb{R}^k$ for which

$$[\mathbf{f}]_{0,\gamma} := \sup_{\substack{x,y\in U,\\ x\neq y}} \frac{|\mathbf{f}(x)-\mathbf{f}(y)|}{|x-y|^\gamma} < \infty.$$

We describe the ellipticity of a vector field $\mathcal{A}$ using a function $\mathcal{V} : \mathbb{R}^{n\times m} \to \mathbb{R}^{n\times m}$ given by

$$\mathcal{V}(\xi) = \left(\frac{g(|\xi|)}{|\xi|}\right)^{1/2} \xi. \tag{19}$$

3.2 Basic definitions

References for this section are [61, 79].

We say that a function $G : [0,\infty) \to [0,\infty]$ is a Young function if it is convex, vanishes at 0, and is neither identically equal to 0, nor to infinity. A Young function G which is finite-valued, vanishes only at 0 and satisfies the additional growth conditions

$$\lim_{t\to 0} \frac{G(t)}{t} = 0 \quad \text{and} \quad \lim_{t\to\infty} \frac{G(t)}{t} = \infty$$

is called an N-function. The complementary function $\widetilde{G}$ (called also the Young conjugate, or the Legendre transform) to a nondecreasing function $G : [0,\infty) \to [0,\infty)$ is given by the following formula

$$\widetilde{G}(s) := \sup_{t>0}(s\cdot t - G(t)).$$

If G is a Young function, so is $\widetilde{G}$. If G is an N-function, so is $\widetilde{G}$.

Having Young functions $G, \widetilde{G}$, we are equipped with Young's inequality reading as follows

$$ts \le G(t) + \widetilde{G}(s) \quad \text{for all } s,t \ge 0. \tag{20}$$

We say that a function $G : [0,\infty) \to [0,\infty)$ satisfies Δ_2-condition if there exists $c_{\Delta_2} > 0$ such that $G(2t) \le c_{\Delta_2} G(t)$ for $t > 0$. We say that G satisfy ∇_2-condition if $\widetilde{G} \in \Delta_2$. Note that it is possible that G satisfies only one of the conditions Δ_2/∇_2. For instance, for $G(t) = ((1+|t|)\log(1+|t|) - |t|) \in \Delta_2$, its complementary function is $\widetilde{G}(s) = (\exp(|s|)-|s|-1) \notin \Delta_2$. See [79, Sect. 2.3, Theorem 3] for equivalence of various definitions of these conditions and [21, 31] for illustrating the subtleties. In particular, $G \in \Delta_2 \cap \nabla_2$ if and only if $1 < i_G \le s_G < \infty$, see (2). This assumption implies a comparison with power-type functions i.e. $\frac{G(t)}{t^{i_G}}$ is non-decreasing and $\frac{G(t)}{t^{s_G}}$ is non-increasing, but it is stronger than being sandwiched between power functions.

Lemma 3.1 *If an N-function $G \in \Delta_2 \cap \nabla_2$, then $g(t)t \approx G(t)$ and $\widetilde{G}(g(t)) \approx G(t)$ with the constants depending only on the growth indexes of G, that is i_G and s_G. Moreover, $g^{-1}(2t) \le cg^{-1}(t)$ with $c = c(i_G, s_G)$.*

Due to Lemma 3.1 and [33, Lemmas 3 and 21], we have the following relations.

Lemma 3.2 *If G is an N-function of class $C^2((0,\infty)) \cap C([0,\infty))$, $G, g \in \Delta_2 \cap \nabla_2$, $\mathcal{A}$ is given by (5), then for every $\xi, \eta \in \mathbb{R}^{n\times m}$ it holds*

$$(\mathcal{A}(x,\xi) - \mathcal{A}(x,\eta)) : (\xi - \eta) \gtrsim_{data} \frac{g(|\xi| + |\eta|)}{|\xi| + |\eta|}|\xi - \eta|^2 \approx_{data} |\mathcal{V}(\xi) - \mathcal{V}(\xi)|^2\,, \tag{21}$$

and

$$g(|\xi| + |\eta|)|\xi - \eta| \approx_{data} G^{\frac{1}{2}}(|\xi| + |\eta|)|\mathcal{V}(\xi) - \mathcal{V}(\eta)|. \tag{22}$$

3.3 Orlicz spaces

Basic reference for this section is [2], where the theory of Orlicz spaces is presented for scalar functions. The proofs for functions with values in $\mathbb{R}^m$ can be obtained by obvious modifications.

We study the solutions to PDEs in the Orlicz-Sobolev spaces equipped with a modular function $G \in C^1((0,\infty))$ - a strictly increasing and convex function such that $G(0) = 0$ and satisfying (2). Let us define a modular

$$\varrho_{G,U}(\boldsymbol{\xi}) = \int_U G(|\boldsymbol{\xi}|)\,dx. \tag{23}$$

For any bounded $\Omega \subset \mathbb{R}^n$, by Orlicz space $L^G(\Omega, \mathbb{R}^m)$ we understand the space of measurable functions endowed with the Luxemburg norm

$$||\mathbf{f}||_{L^G(\Omega,\mathbb{R}^m)} = \inf\left\{\lambda > 0 :\ \ \varrho_{G,\Omega}\left(\tfrac{1}{\lambda}|\mathbf{f}|\right) \le 1\right\}.$$

We define the Orlicz-Sobolev space $W^{1,G}(\Omega)$ as follows

$$W^{1,G}(\Omega, \mathbb{R}^m) = \left\{\mathbf{f} \in W^{1,1}(\Omega, \mathbb{R}^m) :\ \ |\mathbf{f}|, |D\mathbf{f}| \in L^G(\Omega, \mathbb{R}^m)\right\},$$

where the gradient is understood in the distributional sense, endowed with the norm

$$\|\mathbf{f}\|_{W^{1,G}(\Omega,\mathbb{R}^m)} = \inf\left\{\lambda > 0 :\ \ \varrho_{G,\Omega}\left(\tfrac{1}{\lambda}|\mathbf{f}|\right) + \varrho_{G,\Omega}\left(\tfrac{1}{\lambda}|D\mathbf{f}|\right) \le 1\right\}$$

and by $W_0^{1,G}(\Omega, \mathbb{R}^m)$ we denote the closure of $C_c^\infty(\Omega, \mathbb{R}^m)$ under the above norm. Since condition (2) imposed on G implies $G, \widetilde{G} \in \Delta_2$, the Orlicz-Sobolev space $W^{1,G}(\Omega, \mathbb{R}^m)$ we deal with is separable and reflexive. Moreover, one can apply arguments of [49] to infer density of smooth functions in $W^{1,G}(\Omega, \mathbb{R}^m)$.

The counterpart of the Hölder inequality in this setting reads

$$\|\mathbf{f}\mathbf{g}\|_{L^1(\Omega,\mathbb{R}^m)} \le 2\|\mathbf{f}\|_{L^G(\Omega,\mathbb{R}^m)}\|\mathbf{g}\|_{L^{\widetilde{G}}(\Omega,\mathbb{R}^m)} \tag{24}$$

for all $\mathbf{f} \in L^G(\Omega, \mathbb{R}^m)$ and $\mathbf{g} \in L^{\widetilde{G}}(\Omega, \mathbb{R}^m)$.

3.4 The operator

We notice that in such regime the operator A_G acting as

$$\langle \mathsf{A}_G \mathbf{u}, \boldsymbol{\phi} \rangle := \int_\Omega \mathcal{A}(x, D\mathbf{u}) : D\boldsymbol{\phi}\, dx \quad \text{for} \quad \boldsymbol{\phi} \in C_0^\infty(\Omega, \mathbb{R}^m)$$

is well defined on a reflexive and separable Banach space $W_0^{1,G}(\Omega, \mathbb{R}^m)$ and $\mathsf{A}_G (W_0^{1,G}(\Omega, \mathbb{R}^m)) \subset (W_0^{1,G}(\Omega, \mathbb{R}^m))'$. Indeed, when $\mathbf{u} \in W_0^{1,G}(\Omega, \mathbb{R}^m)$ and $\boldsymbol{\phi} \in C_c^\infty(\Omega, \mathbb{R}^m)$, structure condition (5), Hölder's inequality (24), and Lemma 3.1 justify that

$$\begin{aligned} |\langle \mathsf{A}_G \mathbf{u}, \boldsymbol{\phi} \rangle| &\le c \int_\Omega g(|D\mathbf{u}|)|D\boldsymbol{\phi}|\, dx \le c\, \|g(|D\mathbf{u}|)\|_{L^{\widetilde{G}(\cdot)}} \||D\boldsymbol{\phi}|\|_{L^G} \\ &\le c \||D\mathbf{u}|\|_{L^G} \||D\boldsymbol{\phi}|\|_{L^G} \le c \|\boldsymbol{\phi}\|_{W^{1,G}}. \end{aligned}$$

3.5 Definitions of solutions and comments on existence results

We stress that the problems are considered under the regime of Assumption **(A-vect)**.

A function $\mathbf{v} \in W_{loc}^{1,G}(\Omega, \mathbb{R}^m)$ is called an $\mathcal{A}$-harmonic map in $\Omega \subset \mathbb{R}^n$ provided

$$\int_\Omega \mathcal{A}(x, D\mathbf{v}) : D\boldsymbol{\varphi}\, dx = 0 \qquad \text{for all} \;\; \boldsymbol{\varphi} \in C_c^\infty(\Omega, \mathbb{R}^m). \tag{25}$$

As a consequence of our main result, in Corollary 2.4, we prove that $\mathcal{A}$-harmonic maps are continuous. In fact, by Campanato's characterization [48, Theorem 2.9] one can infer from Proposition 3.13 Hölder continuity $C^{0,\gamma}(\Omega, \mathbb{R}^m)$ of $\mathcal{A}$-harmonic maps with any exponent $\gamma \in (0, 1)$.

A function $\mathbf{u} \in W_{loc}^{1,G}(\Omega, \mathbb{R}^m)$ is called a weak solution to (4), if

$$\int_\Omega \mathcal{A}(x, D\mathbf{u}) : D\boldsymbol{\varphi}\, dx = \int_\Omega \boldsymbol{\varphi}\, d\boldsymbol{\mu}(x) \quad \text{for every} \;\; \boldsymbol{\varphi} \in C_c^\infty(\Omega, \mathbb{R}^m). \tag{26}$$

Recall that $W_0^{1,G}(\Omega, \mathbb{R}^m)$ is separable and by its definition $C_c^\infty(\Omega, \mathbb{R}^m)$ is dense there.

Remark 3.3 (Existence and uniqueness of weak solutions) For $\boldsymbol{\mu} \in (W_0^{1,G}(\Omega, \mathbb{R}^m))'$, due to the strict monotonicity of the operator, there exists a unique weak solution to (4), see [56, Sect. 3.1].

Recall that the notion of SOLA is defined in Sect. 2.2. The problem (4) admits a solution of this type for arbitrary bounded measure.

Proposition 3.4 *If a vector field $\mathcal{A}$ satisfies* Assumption **(A-vect)** *and $\boldsymbol{\mu} \in \mathcal{M}(\Omega, \mathbb{R}^m)$, then there exists a SOLA $\mathbf{u} \in W_0^{1,1}(\Omega, \mathbb{R}^m)$ to* (4).

The idea to prove it is to consider

$$\mathbf{f}_k(x) := \int_{\mathbb{R}^m} \varrho_k(x - y)\, d\boldsymbol{\mu}(y),$$

where ϱ_k stands for a standard mollifier i.e. for a nonnegative, smooth, and even function such that $\int_{\mathbb{R}^m} \varrho(s)\, \mathrm{d}s = 1$ we define $\varrho_k(s) = k^n \varrho(ks)$ for $k \in \mathbb{N}$. Of course

$$\mathbf{f}_k \xrightharpoonup{*} \boldsymbol{\mu} \qquad \text{and} \qquad \sup_k \|\mathbf{f}_k\|_{L^1(\Omega)} \le |\boldsymbol{\mu}|(\mathbb{R}^n) < \infty.$$

By Remark 3.3 one finds $\mathbf{u}_k \in W_0^{1,G}(\Omega, \mathbb{R}^m)$ such that for every $\boldsymbol{\varphi} \in W_0^{1,G}(\Omega, \mathbb{R}^m)$ it holds that

$$\int_\Omega \mathcal{A}(x, D\mathbf{u}_k) : D\boldsymbol{\varphi}\, dx = \int_\Omega \boldsymbol{\varphi}\, \mathbf{f}_k dx.$$

The existence of SOLA by passing to the limit can be justified by modification of arguments of [39].
See [3, 8, 18, 21, 24, 26] for related existence and regularity results in the scalar case and [28, 39, 40, 46, 66] for vectorial existence results in various regimes.

3.6 Auxiliary results

Lemma 3.5 *For* $\mathbf{g} : B \to \mathbb{R}^k$, $k \geq 1$ *and any* $\boldsymbol{\xi} \in \mathbb{R}^k$ *it holds that*

$$\fint_B |\mathbf{g} - (\mathbf{g})_B|\, dx \leq 2 \fint_B |\mathbf{g} - \boldsymbol{\xi}|\, dx.$$

We have the following corollary of the Cavalieri Principle.

Lemma 3.6 *If* $\nu \in \mathcal{M}(\Omega)$ *has a density* ω *(i.e.* $d\nu = \omega(x)\, dx$ *with* $\omega \in L^1(\Omega)$*) and* $(1 + |f|)^{-(\gamma+1)}\omega \in L^1(\mathbb{R}^n)$ *for some* $\gamma > 0$*, then*

$$\int_0^\infty \frac{\nu(\{|f| < t\})}{(1+t)^{2+\gamma}}\, dt = \frac{1}{1+\gamma} \int_{\mathbb{R}^n} \frac{d\nu}{(1+|f|)^{\gamma+1}}.$$

Lemma 3.7 ([48], Lemma 6.1) *Let* $\phi : [R/2, 3R/4] \to [0, \infty)$ *be a function such that*

$$\phi(r_1) \leq \frac{1}{2}\phi(r_2) + A + \frac{B}{(r_2 - r_1)^\beta} \qquad \textit{for every} \qquad R/2 \leq r_1 < r_2 \leq 3R/4$$

with $A, B \geq 0$ *and* $\beta > 0$*. Then there exists* $c = c(\beta)$*, such that*

$$\phi(R/2) \leq c\left(A + \frac{B}{R^\beta}\right).$$

Lemma 3.8 ([50], Lemma 3.4) *Let* $\phi(t)$ *be a nonnegative and nondecreasing function on* $[0, R]$*. Suppose that*

$$\phi(\rho) \leq A\left[\left(\frac{\rho}{r}\right)^\alpha + \epsilon\right]\phi(r) + Br^\beta$$

for any $0 < \rho \leq r \leq R$*, with* A, B, α, β *nonnegative constants and* $\beta < \alpha$*. Then for any* $\gamma \in (\beta, \alpha)$*, there exists a constant* $\epsilon_0 = \epsilon_0(A, \alpha, \beta, \gamma)$ *such that if* $\epsilon < \epsilon_0$*, then for all* $0 < \rho \leq r \leq R$ *we have*

$$\phi(\rho) \leq c\left\{\left(\frac{\rho}{r}\right)^\gamma \phi(r) + B\rho^\beta\right\}$$

where c *is a positive constant depending on* A, α, β, γ.

The next lemma is a self-improving property for the reverse Hölder inequalities.

Lemma 3.9 ([53], Lemma 3.38) *Let* $0 < r < q < p < \infty$, $0 < \rho < R \leq 1$ *and* $w \in L^p(B_1)$*. If the following reverse Hölder inequality holds*

$$\left(\int_{B_{\sigma'}} w^p\, dx\right)^{1/p} \leq \frac{c_0}{(\sigma - \sigma')^\varkappa}\left(\int_{B_\sigma} w^q\, dx\right)^{1/q} + c_1$$

for some constants c_0 *and* c_1*, whenever* $\rho \leq \sigma' \leq \sigma \leq R$*. Then there exists* $c = c(c_0, \xi, p, q, r)$ *such that*

$$\left(\int_{B_\rho} w^p\,dx\right)^{1/p} \leq \frac{c}{(R-\rho)^{\widetilde{\varkappa}}}\left(\int_{B_R} w^r\,dx\right)^{1/r} + c_1,$$

where

$$\widetilde{\varkappa} = \frac{\varkappa r(p-q)}{q(p-r)}.$$

The following modular version of the Sobolev-Poincaré inequality follows almost directly as in [5], but we present the proof for vector-valued functions for the sake of completeness.

Proposition 3.10 *Suppose* Ω *is a bounded Lipschitz domain in* $\mathbb{R}^n$*,* $m, n \geq 1$*, and* $G : [0,\infty) \to [0,\infty)$ *is an* N*-function such that* $G \in \Delta_2 \cap \nabla_2$*. Then there exist a constant* $C = C(n, m, |\Omega|, G) > 0$*, such that for every* $\mathbf{u} \in W_0^{1,G}(\Omega, \mathbb{R}^m)$

$$\int_\Omega G^{n'}(|\mathbf{u}|)\,dx \leq C\left(\int_\Omega G(|D\mathbf{u}|)\,dx\right)^{n'}.$$

Proof We provide the proof only in the case of continuously differentiable G. Otherwise every time one can find a sufficiently smooth function $G_\circ$ comparable to G, i.e. such that there exists $c > 0$, such that $G_\circ(t)/c \leq G(t) \leq cG_\circ(t)$. Moreover, we start with the proof for fixed $\mathbf{u} \in C_c^\infty(\Omega, \mathbb{R}^m)$ and then conclude by the density argument. The classical Sobolev inequality in $W^{1,1}$ gives

$$\left(\int_\Omega G^{n'}(|\mathbf{u}|)\,dx\right)^{\frac{1}{n'}} \leq c\int_\Omega |D(G(|\mathbf{u}|))|\,dx. \tag{27}$$

Since $G \in \Delta_2$, it satisfies $g(t) \leq c\,G(t)/t$ and $G^*\left(\frac{G(t)}{t}\right) \leq G(t)$. Thus by the Young inequality we arrive at

$$\begin{aligned} |D(G(|\mathbf{u}|))| = g(|\mathbf{u}|)\big|D|\mathbf{u}|\big| &\leq c\frac{G(|\mathbf{u}|)}{|\mathbf{u}|}|D\mathbf{u}| \\ &\leq \varepsilon G^*\left(\frac{G(|\mathbf{u}|)}{|\mathbf{u}|}\right) + c\,G(|D\mathbf{u}|) \leq \varepsilon G(|\mathbf{u}|) + c\,G(|D\mathbf{u}|). \end{aligned} \tag{28}$$

Summing up, we have

$$\left(\int_\Omega G^{n'}(|\mathbf{u}|)\,dx\right)^{\frac{1}{n'}} \leq C\varepsilon\int_\Omega G(|\mathbf{u}|)\,dx + Cc_\varepsilon\int_\Omega G(|D\mathbf{u}|)\,dx,$$

where according to the Hölder inequality we obtain

$$\left(\int_\Omega G^{n'}(|\mathbf{u}|)\,dx\right)^{\frac{1}{n'}} \leq \varepsilon C|\Omega|^{\frac{1}{n}}\left(\int_\Omega G^{n'}(|\mathbf{u}|)\,dx\right)^{\frac{1}{n'}} + Cc_\varepsilon\int_\Omega G(|D\mathbf{u}|)\,dx.$$

Now we can choose ε small enough to absorb it on the right-hand side and obtain the claim for $\mathbf{u} \in C_c^\infty(\Omega, \mathbb{R}^m)$. Since smooth function are dense in $W_0^{1,G}(\Omega, \mathbb{R}^m)$ by standard approximation argument we get the claim for all $\mathbf{u} \in W_0^{1,G}(\Omega, \mathbb{R}^m)$. □

3.7 Properties of $\mathcal{A}$-harmonic maps

Let us establish some fundamental properties of $\mathcal{A}$-harmonic maps.

Proposition 3.11 (Caccioppoli estimate) *If $\mathbf{v} \in W_0^{1,G}(\Omega, \mathbb{R}^m)$ is a nonnegative $\mathcal{A}$-harmonic map, $\boldsymbol{\lambda} \in \mathbb{R}^m$, and $7/8 \le \sigma' < \sigma \le 1$, then there exists $c = c(\mathtt{data}) > 0$, such that*

$$\fint_{B_{\sigma' r}} G(|D\mathbf{v}|)\, dx \le \frac{c}{(\sigma' - \sigma)^{s_G}} \fint_{B_{\sigma r}} G\left(\frac{|\mathbf{v} - \boldsymbol{\lambda}|}{r}\right) dx. \tag{29}$$

Proof Let us pick a cutoff function $\eta \in C_c^\infty(B_{\sigma r})$ such that $\mathbb{1}_{B_{\sigma' r}} \le \eta \le \mathbb{1}_{B_{\sigma r}}$ and $|D\eta| \le c_1/(\sigma' - \sigma)$. We use $\xi = \eta^q(\mathbf{v} - \boldsymbol{\lambda})$ as a test function to get

$$\int_\Omega \mathcal{A}(x, D\mathbf{v}) : D\mathbf{v}\, \eta^q\, dx = \int_\Omega \mathcal{A}(x, D\mathbf{v}) : (-q\eta^{q-1}(\mathbf{v} - \boldsymbol{\lambda}) \otimes D\eta)\; dx.$$

Therefore, due to the coercivity of $\mathcal{A}$ and the Cauchy–Schwarz inequality we have

$$\int_{B_{\sigma r}} G(|D\mathbf{v}|)\eta^q\, dx \le c \int_{B_{\sigma r}} \frac{g(|D\mathbf{v}|)}{|D\mathbf{v}|} \eta^{q-1} |D\mathbf{v} : ((\mathbf{v} - \boldsymbol{\lambda}) \otimes D\eta)|\, dx =: \mathcal{K}$$

Noting that q is large enough to satisfy $s'_G \ge q'$, we have in turn that $\widetilde{G}(\eta^{q-1}t) \le c\eta^q \widetilde{G}(t)$ and

$$\widetilde{G}(\eta^{q-1} g(t)) \le c\eta^q \widetilde{G}(g(t)) \le c\eta^q G(t).$$

Then, using Young inequality (20) applied to the integrand of $\mathcal{K}$ we get

$$\begin{aligned}\mathcal{K} &\le \varepsilon \int_{B_{\sigma r}} \widetilde{G}(\eta^{q-1}|D\mathbf{v}|)\, dx + c_\varepsilon \int_{B_{\sigma r}} G\left(|\mathbf{v} - \boldsymbol{\lambda}|\, |D\eta|/c_1\right)\, dx \\ &\le \varepsilon c \int_{B_{\sigma r}} \eta^q G(|D\mathbf{v}|)\, dx + c_\varepsilon \int_{B_{\sigma r}} G\left(|\mathbf{v} - \boldsymbol{\lambda}|\, |D\eta|/c_1\right)\, dx\end{aligned}$$

with arbitrary $\varepsilon < 1$. Choosing ε small enough to absorb the term, and finally by properties of η we obtain (29). □

An $\mathcal{A}$-harmonic function $\mathbf{v}$ is a minimizer of a functional

$$\mathbf{v} \mapsto \int_\Omega G(|D\mathbf{v}|)\, dx.$$

Therefore, taking the operator independent of the spacial variable

$$\mathcal{A}(x, \xi) = c_a \frac{g(|\xi|)}{|\xi|} \xi \tag{30}$$

$\mathcal{A}$-harmonic functions are Lipschitz regular by the following fact.

Lemma 3.12 ([35], Lemma 5.8) *Suppose G is an N-function of class $C^2((0,\infty)) \cap C([0,\infty))$, $G, g \in \Delta_2 \cap \nabla_2$, $\mathbf{w}$ is $\mathcal{A}$-harmonic in Ω for $\mathcal{A}$ given by (30). Let $B \subset 2B \Subset \Omega$. Then there exists $c = c(\mathtt{data}) > 0$ such that*

$$\sup_B G(|D\mathbf{w}|) \le c \fint_{2B} G(|D\mathbf{w}|)\, dx.$$

Proposition 3.13 *Suppose a vector field* $\mathcal{A}$ *satisfies* Assumption **(A-vect)**. *If* $\mathbf{v} \in W^{1,G}(\Omega, \mathbb{R}^m)$ *is* $\mathcal{A}$*-harmonic, then for any* $\varsigma \in (0,1)$ *there exists* $R_0 = R_0(\mathtt{data}, \varsigma) \in (0,1]$ *such that for any* $0 < R \leq R_0$ *and* $B_R \Subset \Omega$ *it holds that*

$$\fint_{B_{\delta R}} |\mathbf{v} - (\mathbf{v})_{B_{\delta R}}|\, dx \leq c_o \delta^{1+(\varsigma-1)K} \fint_{B_R} |\mathbf{v} - (\mathbf{v})_{B_R}|\, dx \tag{31}$$

whenever $\delta \in (0, 1/4]$, *where* $c_o \geq 1$ *is a constant depending only on* $\mathtt{data}$ *and* ς *and* $K = K(i_G, s_G)$.

Proof In this proof, we use a classical perturbation argument, see for instance [50, Theorem 3.8]. It suffices to prove (31) for $\mathbf{v}$ solving

$$-\mathbf{div}\left(a(x)\frac{g(|D\mathbf{v}|)}{|D\mathbf{v}|}D\mathbf{v}\right) = 0 \quad \text{in} \quad B_1.$$

Indeed, the general case can be deduced then by considering $\widetilde{\mathbf{v}}(x) = \mathbf{v}(x_0 + Rx)/R$ solving $-\mathbf{div}\bar{\mathcal{A}}(x, D\widetilde{\mathbf{v}}) = 0$ on $B_1(0)$ with

$$\bar{\mathcal{A}}(x,\xi) = \bar{a}(x)\frac{\bar{g}(|\xi|)}{|\xi|}\xi = \frac{1}{g(R)}\mathcal{A}(x_0 + Rx, R\xi) = a(x_0 + Rx)\frac{g(R|\xi|)}{g(R)|\xi|}\xi.$$

In this case, the modulus $\omega_{\bar{a}}$ of continuity of $\bar{a}$ satisfies $\omega_{\bar{a}}(r) = \omega_a(rR)$.

Note first Propositions 3.10 and 3.11 imply that for any $7/8 \leq \sigma' < \sigma \leq 1$ it holds

$$\left(\fint_{B_{\sigma'}} G^{n'}\left(\frac{|\mathbf{v} - (\mathbf{v})_{B_{\sigma'}}|}{\sigma'}\right) dx\right)^{1/n'} \leq \frac{c}{(\sigma - \sigma')^{s_G}} \fint_{B_\sigma} G\left(\frac{|\mathbf{v} - (\mathbf{v})_{B_\sigma}|}{\sigma}\right) dx.$$

From the doubling property of G and the upper and lower bound on σ' and σ, we have

$$\left(\fint_{B_{\sigma'}} G^{n'}\left(|\mathbf{v} - (\mathbf{v})_{B_{\sigma'}}|\right) dx\right)^{1/n'} \leq \frac{c}{(\sigma - \sigma')^{s_G}} \fint_{B_\sigma} G\left(|\mathbf{v} - (\mathbf{v})_{B_\sigma}|\right) dx.$$

Using the triangle inequality and Jensen's inequality

$$\begin{aligned}
&\left(\fint_{B_{\sigma'}} G^{n'}\left(|\mathbf{v} - (\mathbf{v})_{B_1}|\right) dx\right)^{1/n'} \\
&\leq c\left(\fint_{B_{\sigma'}} G^{n'}\left(|\mathbf{v} - (\mathbf{v})_{B_{\sigma'}}|\right) dx\right)^{1/n'} + c\, G\left(|(\mathbf{v})_{B_{\sigma'}} - (\mathbf{v})_{B_1}|\right) \\
&\leq \frac{c}{(\sigma - \sigma')^{s_G}} \fint_{B_\sigma} G\left(|\mathbf{v} - (\mathbf{v})_{B_\sigma}|\right) dx + c \fint_{B_{\sigma'}} G\left(|\mathbf{v} - (\mathbf{v})_{B_1}|\right) dx \\
&\leq c\left(\frac{1}{(\sigma - \sigma')^{s_G}} + 1\right) \fint_{B_\sigma} G\left(|\mathbf{v} - (\mathbf{v})_{B_1}|\right) dx \\
&\leq \frac{c}{(\sigma - \sigma')^{s_G}} \fint_{B_\sigma} G\left(|\mathbf{v} - (\mathbf{v})_{B_1}|\right) dx,
\end{aligned}$$

for $c = c(\mathtt{data})$. From Lemma 3.9, for any $t \in (0, 1/s_G)$, we have

$$\left(\fint_{B_{7/8}} G^{n'}\left(|\mathbf{v}-(\mathbf{v})_{B_{7/8}}|\right)\,dx\right)^{1/n'} \le c\left(\fint_{B_{7/8}} G^{n'}\left(|\mathbf{v}-(\mathbf{v})_{B_1}|\right)\,dx\right)^{1/n'}$$
$$\le c\left(\fint_{B_1} G^{t}\left(|\mathbf{v}-(\mathbf{v})_{B_1}|\right)\,dx\right)^{1/t}$$
$$\le c\,G\left(\fint_{B_1} |\mathbf{v}-(\mathbf{v})_{B_1}|\,dx\right). \tag{32}$$

Here, in the last line, as $t \mapsto G^t(t)$ is a concave function for every $t \in (0, 1/s_G)$ we have used Jensen's inequality.

Let $\mathbf{w} \in \mathbf{v} + W_0^{1,G}(B_\sigma, \mathbb{R}^m)$ for any $\sigma \in (1, 1/2)$ be the weak solution to

$$-\,\mathbf{div}\left(a(x_0)\frac{g(|D\mathbf{w}|)}{|D\mathbf{w}|}D\mathbf{w}\right) = 0 \qquad \text{in } B_\sigma. \tag{33}$$

Recall that the function $\mathcal{V}$ describing ellipticity of $\mathcal{A}$ is defined in (19) and a is a function bounded below by $c_a > 0$ and with a modulus of continuity ω_a. Testing (33) and (25) against $(\mathbf{w}-\mathbf{v})$, and applying Lemma 3.2, for any $\epsilon \in (0, 1]$ we see

$$\begin{aligned}
&\int_{B_\sigma} c_a\,|\mathcal{V}(D\mathbf{w})-\mathcal{V}(D\mathbf{v})|^2\,dx\\
&\le c\int_{B_\sigma} a(x_0)\left(\frac{g(|D\mathbf{w}|)}{|D\mathbf{w}|}D\mathbf{w}-\frac{g(|D\mathbf{v}|)}{|D\mathbf{v}|}D\mathbf{v}\right):(D\mathbf{w}-D\mathbf{v})\,dx\\
&= c\int_{B_\sigma} (a(x_0)-a(x))\frac{g(|D\mathbf{v}|)}{|D\mathbf{v}|}D\mathbf{v}:(D\mathbf{w}-D\mathbf{v})\,dx\\
&\le c\,\omega_a(\sigma)\int_{B_\sigma} g(|D\mathbf{w}|+|D\mathbf{v}|)\,|D\mathbf{w}-D\mathbf{v}|\,dx\\
&\le c\,\epsilon\int_{B_\sigma} |\mathcal{V}(D\mathbf{w})-\mathcal{V}(D\mathbf{v})|^2\,dx + c\,\frac{\omega_a(1/2)^2}{\epsilon}\int_{B_\sigma} G(|D\mathbf{w}|+|D\mathbf{v}|)\,dx,
\end{aligned}$$

where in the last line we used Young's inequality. As $c = c(\mathtt{data})$ and $c_a > 0$, we can take ϵ small enough to absorb the first term on the right-hand side. Then Jensen's inequality implies

$$\int_{B_\sigma} |\mathcal{V}(D\mathbf{w})-\mathcal{V}(D\mathbf{v})|^2\,dx \le c\,\omega_a(1/2)^2\int_{B_\sigma} [G(|D\mathbf{v}|)+G(|D\mathbf{w}|)]\,dx.$$

Since $\mathbf{w}$ is a minimizer of the integral functional

$$\mathbf{w} \mapsto \int_{B_\sigma} G(|D\mathbf{w}|)\,dx,$$

we have

$$\int_{B_\sigma} |\mathcal{V}(D\mathbf{w})-\mathcal{V}(D\mathbf{v})|^2\,dx \le c\,\omega_a(1/2)^2\int_{B_\sigma} G(|D\mathbf{v}|)\,dx. \tag{34}$$

Since we know the Lipschitz regularity of $\mathbf{w}$, provided by Lemma 3.12

$$\sup_{B_{\sigma/2}} G(|D\mathbf{w}|) \le c\fint_{B_\sigma} G(|D\mathbf{w}|)\,dx,$$

it follows from (34) that

$$\begin{aligned}\int_{B_{\delta\sigma}} G(|D\mathbf{v}|)\,dx &\le c\int_{B_{\delta\sigma}} |\mathcal{V}(D\mathbf{v})|^2\,dx\\ &\le c\int_{B_{\delta\sigma}} |\mathcal{V}(D\mathbf{v})-\mathcal{V}(D\mathbf{w})|^2\,dx+\int_{B_{\delta\sigma}} |\mathcal{V}(D\mathbf{w})|^2\,dx\\ &\le c\int_{B_{\sigma}} |\mathcal{V}(D\mathbf{v})-\mathcal{V}(D\mathbf{w})|^2\,dx+\delta^n\int_{B_{\sigma}} |\mathcal{V}(D\mathbf{w})|^2\,dx\\ &\le c\left(\delta^n+\omega_a(1/2)^2\right)\int_{B_\sigma} G(|D\mathbf{v}|)\,dx.\end{aligned}$$

We take $R_0 = R_0(\mathtt{data}, \varsigma)$ small enough to ensure that $\omega_a(R_0)^2 < \epsilon_0$, where ϵ_0 is a constant given in Lemma 3.8. Then Lemma 3.8, Proposition 3.11 and (32) give

$$⨍_{B_\delta} G(|D\mathbf{v}|)\,dx \le c\delta^{\varsigma-1}⨍_{B_{1/2}} G(|D\mathbf{v}|)\,dx \le c\delta^{\varsigma-1}G\left(\int_{B_1} |\mathbf{v}-(\mathbf{v})_{B_1}|\,dx\right) \tag{35}$$

where c depends only on $\mathtt{data}$ and ς.

Using the Sobolev-Poincaré inequality in $W^{1,1}$ and Jensen's inequality, we conclude that

$$\begin{aligned}⨍_{B_\delta} |\mathbf{v}-(\mathbf{v})_{B_\delta}|\,dx &\le c\,\delta ⨍_{B_\delta} |D\mathbf{v}|\,dx\\ &\le c\,\delta\, G^{-1}\left(⨍_{B_\delta} G(|D\mathbf{v}|)\,dx\right)\\ &\le c\,\delta^{1+(\varsigma-1)K}\int_{B_1} |\mathbf{v}-(\mathbf{v})_{B_1}|\,dx,\end{aligned}$$

what completes the proof. □

4 Measure data $\mathcal{A}$-harmonic approximation

In this section we provide the tool of crucial meaning for our further reasoning – the approximation of a $W^{1,G}$-function by an $\mathcal{A}$-harmonic map for weighted operator $\mathcal{A}$ of an Orlicz growth given by (5). Results in this spirit can be found in [37, 41, 45], but most preeminently for the approximation relevant for application to measure data problems we refer to [65, Theorem 4.1].

We define an auxiliary function

$$H_s(t) = \int_0^t \frac{g(r)^{1-s}G(r)^s}{r}\,dr \qquad \text{for } s\in[0,1/2). \tag{36}$$

It is readily checked that when $s > \max\{2-i_G, 0\}$, H_s is a Young function satisfying

$$H_s(t) \approx g(t)^{1-s}G(t)^s \tag{37}$$

with intrinsic constants depending only on i_G, s_G and s. In fact, $H_s \in \Delta_2\cap\nabla_2$ since

$$0 < s+i_G-2 \le \frac{tH_s''(t)}{H_s'(t)} = \frac{(1-s)tg'(t)}{g(t)} + \frac{stg(t)}{G(t)} \le s+s_G-2. \tag{38}$$

Furthermore, there exist $\epsilon, c, t_0 > 0$, such that $H_s(t) \ge ct^{1+\epsilon}$ and $H_s(t) \ge cg^{1+\epsilon}(t)$ for all $t \ge t_0$.

Theorem 4.1 *Under Assumption* **(A-vect)** *let* $\varepsilon > 0$, $\gamma \in (0, 1/(s_G n))$, *and*

$$\max\{2 - i_G, 0\} < s < s_{\mathrm{m}} := \frac{i_G - \gamma s_G n}{i_G + s_G n}. \tag{39}$$

Suppose that $\mathbf{u} \in W^{1,G}(B_r(x_0), \mathbb{R}^m)$ *satisfies*

$$\fint_{B_r(x_0)} |\mathbf{u}|\, dx \le Mr, \quad M \ge 1 \tag{40}$$

then there exists $\delta = \delta(\mathtt{data}, s, M, \varepsilon) \in (0, 1]$ *such that if* $\mathbf{u}$ *is almost* $\mathcal{A}$*-harmonic in a sense that for every* $\boldsymbol{\varphi} \in W_0^{1,G}(B_r(x_0), \mathbb{R}^m) \cap L^\infty(B_r(x_0), \mathbb{R}^m)$ *it holds*

$$\left| \fint_{B_r(x_0)} \mathcal{A}(x, D\mathbf{u}) : D\boldsymbol{\varphi}\, dx \right| \le \frac{\delta}{r} \|\boldsymbol{\varphi}\|_{L^\infty(B_r(x_0),\mathbb{R}^m)}, \tag{41}$$

then there exists an $\mathcal{A}$*-harmonic map* $\mathbf{v} \in W^{1,G}(B_{r/2}(x_0), \mathbb{R}^m)$ *satisfying*

$$\fint_{B_{r/2}(x_0)} H_s(|D\mathbf{u} - D\mathbf{v}|)\, dx \le \varepsilon \tag{42}$$

together with

$$\fint_{B_{r/2}(x_0)} |\mathbf{v}|\, dx \le 2^n Mr \quad \text{and} \quad \fint_{B_{r/2}(x_0)} H_s(|D\mathbf{v}|)\, dx \le c H_s(M), \tag{43}$$

where $c = c(\mathtt{data}) > 0$.

Remark 4.2 The limitation that G has to be superquadratic ($i_G \ge 2$) can be a little bit relaxed in Theorem 4.1 and later on the restriction is not needed. The key property is to ensure that the range of admissible s from (39) is nonempty. We need to assume that i_G is either bigger or equal to 2, or close to 2 in a sense that

$$2 - i_G < \frac{i_G - 1}{i_G + s_G n}.$$

Proof The plan is to first establish suitable a priori estimates for the rescaled problem and then proceed with the proof via contradiction. The proof is presented in 6 steps.

Step 1. Scaling. We fix arbitrary $\boldsymbol{\varphi} \in W_0^{1,G}(B_r(x_0), \mathbb{R}^m) \cap L^\infty(B_r(x_0), \mathbb{R}^m)$ satisfying (41). Let us change variables putting

$$\bar{\mathbf{u}}(x) := \frac{\mathbf{u}(x_0 + rx)}{Mr}, \ \bar{\mathcal{A}}(x_0 + rx, \boldsymbol{\xi}) = \mathcal{A}(x_0 + rx, M\boldsymbol{\xi}), \text{ and } \boldsymbol{\eta}(x) := \frac{\boldsymbol{\varphi}(x_0 + rx)}{r}. \tag{44}$$

Then $\bar{\mathcal{A}}$ satisfies the same conditions as $\mathcal{A}$ with the functions $\bar{g}(t) := g(Mt)$ and $\bar{G}(t) := G(Mt)/M$ (with $\bar{G}' = \bar{g}$), and the constants depending on $\mathtt{data}$. Of course in such a case $i_G = i_{\bar{G}}$ and $s_G = s_{\bar{G}}$.

Having (40) and (41), by denoting the unit ball by B_1, we get further that

$$\fint_{B_1} |\bar{\mathbf{u}}|\, dx \le 1, \tag{45}$$

$$\left| \fint_{B_1} \bar{\mathcal{A}}(x, D\bar{\mathbf{u}}) : D\boldsymbol{\eta}\, dx \right| \le \delta \|\boldsymbol{\eta}\|_{L^\infty(B_1,\mathbb{R}^m)}. \tag{46}$$

Step 2. A priori estimates. We choose $q \geq s_G$, pick

$$\boldsymbol{\eta} := \phi^q T_k(\bar{\mathbf{u}}) \quad \text{with some } \phi \in C_c^\infty(B_1),\ 0 \leq \phi \leq 1,\ k \geq 0, \tag{47}$$

and denote

$$\mathsf{P} := \frac{\bar{\mathbf{u}} \otimes \bar{\mathbf{u}}}{|\bar{\mathbf{u}}|^2}.$$

Then

$$\begin{aligned} D\boldsymbol{\eta} &= \mathbb{1}_{\{|\bar{\mathbf{u}}| \leq k\}}(\phi^q D\bar{\mathbf{u}} + q\phi^{q-1}\bar{\mathbf{u}} \otimes D\phi) \\ &\quad + \mathbb{1}_{\{|\bar{\mathbf{u}}| > k\}}(\phi^q(\mathsf{Id} - \mathsf{P})D\bar{\mathbf{u}} + q\phi^{q-1}\bar{\mathbf{u}} \otimes D\phi). \end{aligned}$$

We use (47) in (46) to get

$$\begin{aligned} &\Bigg| \int_{B_1 \cap \{|\bar{\mathbf{u}}| \leq k\}} \bar{\mathcal{A}}(x, D\bar{\mathbf{u}}) : (\phi^q D\bar{\mathbf{u}} + q\phi^{q-1}\bar{\mathbf{u}} \otimes D\phi)\, dx \\ &\quad + \int_{B_1 \cap \{|\bar{\mathbf{u}}| > k\}} \bar{\mathcal{A}}(x, D\bar{\mathbf{u}}) : (\phi^q(\mathsf{Id} - \mathsf{P})D\bar{\mathbf{u}} + q\phi^{q-1}\bar{\mathbf{u}} \otimes D\phi)\, dx \Bigg| \leq \delta|B_1|\|\boldsymbol{\eta}\|_{L^\infty(B_1,\mathbb{R}^m)}. \end{aligned} \tag{48}$$

Since $\bar{\mathcal{A}}$ has the quasi-diagonal structure resulting from (5) and

$$D\bar{\mathbf{u}} : \big((\mathsf{Id} - \mathsf{P})D\bar{\mathbf{u}}\big) = |D\bar{\mathbf{u}}|^2 - \frac{D_j\bar{\mathbf{u}}^\alpha \bar{\mathbf{u}}^\alpha D_j\bar{\mathbf{u}}^\beta \bar{\mathbf{u}}^\beta}{|\bar{\mathbf{u}}|^2} = |D\bar{\mathbf{u}}|^2 - \frac{\sum_{j=1}^m \langle D_j\bar{\mathbf{u}}, \bar{\mathbf{u}}\rangle^2}{|\bar{\mathbf{u}}|^2} \geq 0,$$

we infer that

$$\bar{\mathcal{A}}(x, D\bar{\mathbf{u}}) : \big((\mathsf{Id} - \mathsf{P})D\bar{\mathbf{u}}\big) \geq 0. \tag{49}$$

By rearranging terms in (48), applying Lemma 3.1, noting that $\|\boldsymbol{\eta}\|_{L^\infty(B_1,\mathbb{R}^m)} \leq k$, and dropping a nonnegative term due to (49), we get for some $c = c(\mathtt{data}, q)$

$$\begin{aligned} \int_{B_1 \cap \{|\bar{\mathbf{u}}| \leq k\}} \bar{G}(|D\bar{\mathbf{u}}|)\phi^q\, dx &\leq c \int_{B_1 \cap \{|\bar{\mathbf{u}}| \leq k\}} \frac{\bar{g}(|D\bar{\mathbf{u}}|)}{|D\bar{\mathbf{u}}|}\phi^{q-1}\big|D\bar{\mathbf{u}} : (\bar{\mathbf{u}} \otimes D\phi)\big|\, dx \\ &\quad + c \int_{B_1 \cap \{|\bar{\mathbf{u}}| > k\}} \frac{k}{|\bar{\mathbf{u}}|}\frac{\bar{g}(|D\bar{\mathbf{u}}|)}{|D\bar{\mathbf{u}}|}\phi^{q-1}\big|D\bar{\mathbf{u}} : (\bar{\mathbf{u}} \otimes D\phi)\big|\, dx + c|B_1|\delta k. \end{aligned}$$

We estimate the first term on the right-hand side of the last display by the use of Young inequality with a parameter, use Lemma 3.1, and we absorb one term. The second term can be estimated by the Cauchy–Schwarz inequality. Altogether we obtain

$$\begin{aligned} \int_{B_1 \cap \{|\bar{\mathbf{u}}| < k\}} \bar{G}(|D\bar{\mathbf{u}}|)\phi^q\, dx &\leq c \int_{B_1 \cap \{|\bar{\mathbf{u}}| < k\}} \bar{G}(|\bar{\mathbf{u}}|\,|D\phi|)\, dx + c|B_1|\delta k \\ &\quad + ck \int_{B_1 \cap \{|\bar{\mathbf{u}}| \geq k\}} \bar{g}(|D\bar{\mathbf{u}}|)\phi^{q-1}|D\phi|\, dx \end{aligned} \tag{50}$$

for some $c = c(\mathtt{data}, q)$.

Step 3. Summability of $\bar{\mathbf{u}}$ and $D\bar{\mathbf{u}}$. We choose $k = t$ in (50), then multiply this line by $(1+t)^{-(\gamma+2)}$, where $\gamma > 0$, integrate it from zero to infinity and apply Cavalieri's principle (Lemma 3.6) twice (with $\nu_1 = \bar{G}(|D\bar{\mathbf{u}}|)\phi^q$ and $\nu_2 = \bar{G}(|\bar{\mathbf{u}}|\,|D\phi|)$. Altogether we get

$$\frac{1}{1+\gamma}\int_{B_1}\frac{\bar G(|D\bar{\mathbf{u}}|)\phi^q}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\,dx \le \frac{c}{1+\gamma}\int_{B_1}\frac{\bar G(|\bar{\mathbf{u}}|\,|D\phi|)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\,dx+\frac{c}{\gamma}\delta$$
$$+\,c\int_0^\infty \frac{t}{(1+t)^{\gamma+2}}\int_{B_1\cap\{|\bar{\mathbf{u}}|>t\}}\bar g(|D\bar{\mathbf{u}}|)\phi^{q-1}|D\phi|\,dx\,dt. \tag{51}$$

The right-most term in the last display can be estimated as follows

$$\begin{aligned} &c\int_0^\infty \frac{t}{(1+t)^{\gamma+2}}\int_{B_1\cap\{|\bar{\mathbf{u}}|>t\}}\bar g(|D\bar{\mathbf{u}}|)\phi^{q-1}|D\phi|\,dx\,dt\\ &\le c\int_0^\infty\frac{1}{(1+t)^{\gamma+1}}dt\int_{B_1}\bar g(|D\bar{\mathbf{u}}|)\phi^{q-1}|D\phi|\,dx\\ &\le \frac{c}{\gamma}\int_{B_1}\bar g(|D\bar{\mathbf{u}}|)\phi^{q-1}|D\phi|\,dx. \end{aligned}\tag{52}$$

To estimate it further we note that q is large enough to satisfy $s'_G \ge q'$, there exist $c_0, c_1 > 0$ depending on i_G, s_G, such that we have $\bar G^*(c_0\phi^{q-1}\bar g(t)) \le c_1\phi^q\bar G^*(\bar g(t)) \le \phi^q\bar G(t)$. Then, using Young inequality (20) applied to the integrand in (52) and the above observation we get

$$\begin{aligned} &\int_{B_1}\bar g(|D\bar{\mathbf{u}}|)\phi^{q-1}|D\phi|\,dx\\ &\le \frac{1}{2(1+\gamma)}\int_{B_1}\frac{\bar G^*\left(c_0\phi^{q-1}\bar g(|D\bar{\mathbf{u}}|)\right)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\,dx+\widetilde c\int_{B_1}\frac{\bar G\left((1+|\bar{\mathbf{u}}|)^{1+\gamma}|D\phi|\right)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\,dx\\ &\le \frac{1}{2(1+\gamma)}\int_{B_1}\frac{\bar G\left(|D\bar{\mathbf{u}}|\right)\phi^q}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\,dx+\widetilde c\int_{B_1}\frac{\bar G\left((1+|\bar{\mathbf{u}}|)^{1+\gamma}|D\phi|\right)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\,dx \end{aligned}$$

with $\widetilde c = \widetilde c(\gamma, i_G, s_G)$. By applying this estimate in (51) and rearranging terms we obtain

$$\int_{B_1}\frac{\bar G(|D\bar{\mathbf{u}}|)\phi^q}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\,dx \le c\int_{B_1}\frac{\bar G\left((1+|\bar{\mathbf{u}}|)^{1+\gamma}|D\phi|\right)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\,dx + c\,\delta\frac{1+\gamma}{\gamma}. \tag{53}$$

Observe that $1/(s_G n) < i_G - 1$, as otherwise the condition required by Remark 4.2 is violated. Recall that since $\gamma < 1/(s_G n)$, we have $\gamma < i_G - 1$. Let us set

$$\vartheta(x) := \frac{\bar G((1+|\bar{\mathbf{u}}|)\phi^q)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}} \tag{54}$$

and notice that since $|D|\bar{\mathbf{u}}|| \le |D\bar{\mathbf{u}}|$, using Lemma 3.1, we can estimate

$$\begin{aligned} |D\vartheta| &= \frac{\left|D(\bar G((1+|\bar{\mathbf{u}}|)\phi^q)(1+|\bar{\mathbf{u}}|)^{1+\gamma} - D\left((1+|\bar{\mathbf{u}}|)^{1+\gamma}\right)\bar G((1+|\bar{\mathbf{u}}|)\phi^q)\right|}{(1+|\bar{\mathbf{u}}|)^{2+2\gamma}}\\ &\le \frac{\bar g((1+|\bar{\mathbf{u}}|)\phi^q)\,|D((1+|\bar{\mathbf{u}}|)\phi^q)|}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}} + (1+\gamma)\frac{|D\bar{\mathbf{u}}|\,\bar G((1+|\bar{\mathbf{u}}|)\phi^q)}{(1+|\bar{\mathbf{u}}|)^{2+\gamma}}\\ &\le \frac{\bar g((1+|\bar{\mathbf{u}}|)\phi^q)|D\bar{\mathbf{u}}|\phi^q}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\\ &\quad + q\,\frac{\bar g((1+|\bar{\mathbf{u}}|)\phi^q)(1+|\bar{\mathbf{u}}|)\phi^{q-1}\,|D\phi|}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}} + (1+\gamma)\frac{|D\bar{\mathbf{u}}|}{1+|\bar{\mathbf{u}}|}\vartheta\\ &\le c(\gamma)\frac{|D\bar{\mathbf{u}}|}{1+|\bar{\mathbf{u}}|}\vartheta + c\,\frac{|D\phi|}{\phi}\vartheta. \end{aligned}\tag{55}$$

For the later use in **Step 6**, we emphasize the dependence of constants on γ by denoting $c(\gamma)$. Note that every $c(\gamma)$ in (55)-(59) is an increasing function of γ.

Since $q \geq s_G$, for any $\kappa \in [1, i_G)$ we see

$$\left|D\left(\vartheta^{\frac{1}{\kappa}}\right)\right| \leq \frac{1}{\kappa}\vartheta^{\frac{1}{\kappa}-1}|D\vartheta| \leq c(\gamma)\,\vartheta^{\frac{1}{\kappa}}\frac{|D\bar{\mathbf{u}}|}{1+|\bar{\mathbf{u}}|} + c\,\vartheta^{\frac{1}{\kappa}}\frac{|D\phi|}{\phi}$$

and

$$\left|D\left(\vartheta^{\frac{1}{\kappa}}\right)\right|^{\kappa} \leq c(\gamma)\frac{G((1+|\bar{\mathbf{u}}|)\phi^q)}{[(1+|\bar{\mathbf{u}}|)\phi^q]^{\kappa}}\frac{(|D\bar{\mathbf{u}}|\phi^q)^{\kappa}}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}} + c\frac{G(1+|\bar{\mathbf{u}}|)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}|D\phi|^{\kappa}. \tag{56}$$

To proceed further, we define an auxiliary function h_κ by setting

$$h_\kappa^{-1}(t) := \int_0^t \frac{1}{\left[G^{-1}(\tau)\right]^{\kappa}}\,d\tau.$$

A straightforward calculation gives

$$-1 < -\frac{\kappa}{i_G} \leq \frac{t[h_\kappa^{-1}]''(t)}{[h_\kappa^{-1}]'(t)} = -\frac{\kappa t[G^{-1}]'(t)}{G^{-1}(t)} \leq -\frac{\kappa}{s_G} < 0, \tag{57}$$

which implies that h_κ^{-1} is an increasing concave function on $[0, \infty)$. Note that (57) also gives

$$h_\kappa\left(\frac{G(t)}{t^\kappa}\right) \approx G(t)$$

with intrinsic constants depending on i_G, s_G and κ only. Moreover, we have

$$\bar{g}(t)[h_\kappa^{-1}]'(\bar{G}(t)) = [h_\kappa^{-1}(\bar{G}(t))]' = \frac{d}{dt}\left(\int_0^{G(t)} \frac{1}{[G^{-1}(\tau)]^{\kappa}}\,d\tau\right) = \frac{\bar{g}(t)}{t^\kappa},$$

and so

$$[\widetilde{h_\kappa}]^{-1}(G(t)) \approx h_\kappa'(h_\kappa^{-1}(\bar{G}(t)) = \frac{1}{[h_\kappa^{-1}]'(h_\kappa(h_\kappa^{-1}(G(t))))} = \frac{1}{[h_\kappa^{-1}]'(G(t))} = t^\kappa.$$

Hence, $\widetilde{h_\kappa}(t) \approx \bar{G}(t^{1/\kappa})$.

Applying Young's inequality (20) with the pair of Young functions $(h_\kappa, \widetilde{h_\kappa})$ to (56), for any $\epsilon_0 \in (0, 1)$ we discover

$$\left|D\left(\vartheta^{\frac{1}{\kappa}}\right)\right|^{\kappa} \leq \epsilon_0\frac{G((1+|\bar{\mathbf{u}}|)\phi^q)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}} + c(\epsilon_0)c(\gamma)\frac{\bar{G}\left(|D\bar{\mathbf{u}}|\phi^q\right)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}} + c\frac{G(1+|\bar{\mathbf{u}}|)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}|D\phi|^{\kappa}.$$

By the classical Sobolev inequality we get that

$$\left(\int_{B_1} |\vartheta|^{\frac{n}{n-\kappa}}\,dx\right)^{\frac{n-\kappa}{n}} \leq c\int_{B_1}\left|D\left(\vartheta^{\frac{1}{\kappa}}\right)\right|^{\kappa}\,dx. \tag{58}$$

Merging (51), (53), (54), (55) and (58) and taking ϵ_0 small enough, we get

$$\left(\int_{B_1}\left(\frac{\bar{G}((1+|\bar{\mathbf{u}}|)\phi^q)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\right)^{\frac{n}{n-\kappa}}\,dx\right)^{\frac{n-\kappa}{n}} \leq \bar{c}\left(\int_{B_1}\frac{\bar{G}\left((1+|\bar{\mathbf{u}}|)^{1+\gamma}|D\phi|\right)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\,dx + \frac{1}{\gamma}\right). \tag{59}$$

It is worth to mention that $\bar{c} = \bar{c}(\mathtt{data}, \gamma, \delta) > 0$ depends on $\mathtt{data}$, γ and δ and it is an increasing function of γ and δ. Since $\gamma \in (0, 1/(s_G n))$ is fixed, we may choose α such that

$$\alpha \in \left(1, \frac{n}{n-\kappa}\right) \qquad \text{and} \qquad \frac{\alpha s_G \gamma}{\alpha - 1} \le 1. \tag{60}$$

Then

$$\int_{B_1} (1+|\bar{\mathbf{u}}|)^{\frac{\alpha s_G \gamma}{\alpha-1}}\, dx \le \int_{B_1} 1 + |\bar{\mathbf{u}}|\, dx \le 1 + |B_1|$$

and we can estimate

$$\begin{aligned}
\int_{B_1} \frac{\bar{G}\left((1+|\bar{\mathbf{u}}|)^{1+\gamma}|D\phi|\right)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\, dx &\le \int_{B_1} \frac{\bar{G}\left((1+|\bar{\mathbf{u}}|)|D\phi|\right)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma-\gamma s_G}}\, dx \\
&\le \left(\int_{B_1} (1+|\bar{\mathbf{u}}|)^{\frac{\alpha s_G \gamma}{\alpha-1}}\, dx\right)^{\frac{\alpha-1}{\alpha}} \left(\int_{B_1} \left(\frac{\bar{G}\left((1+|\bar{\mathbf{u}}|)|D\phi|\right)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\right)^{\alpha} dx\right)^{\frac{1}{\alpha}} \\
&\le c \left(\int_{B_1} \left(\frac{\bar{G}\left((1+|\bar{\mathbf{u}}|)|D\phi|\right)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\right)^{\alpha} dx\right)^{\frac{1}{\alpha}}.
\end{aligned} \tag{61}$$

Thus, from (59) and (61) we obtain

$$\left(\int_{B_1} \left(\frac{\bar{G}((1+|\bar{\mathbf{u}}|)\phi^q)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\right)^{\frac{n}{n-\kappa}} dx\right)^{\frac{n-\kappa}{n}} \le c \left(\int_{B_1} \left(\frac{\bar{G}\left((1+|\bar{\mathbf{u}}|)|D\phi|\right)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\right)^{\alpha} dx\right)^{\frac{1}{\alpha}} + c. \tag{62}$$

For $7/8 \le r_1 < r_2 \le 1$, we take a cut-off function ϕ satisfying

$$\phi \equiv 1 \quad \text{on} \quad B_{r_1} \qquad \text{and} \qquad |D\phi| \le \frac{100}{r_2 - r_1}.$$

It then follows from the doubling property of $\bar{G}$ and Lemma 3.9 that for any $\upsilon \in (0, 1/s_G)$ we have

$$\begin{aligned}
\left(\int_{B_{7/8}} \left(\frac{\bar{G}(1+|\bar{\mathbf{u}}|)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\right)^{\frac{n}{n-\kappa}} dx\right)^{\frac{n-\kappa}{n}} &\le c \left(\int_{B_1} \left(\frac{\bar{G}\,(1+|\bar{\mathbf{u}}|)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\right)^{\upsilon} dx\right)^{\frac{1}{\upsilon}} + c \\
&\le c \left(\int_{B_1} (\bar{G}\,(1+|\bar{\mathbf{u}}|))^{\upsilon}\, dx\right)^{\frac{1}{\upsilon}} + c \\
&\le \bar{G}\left(\int_{B_1} (1+|\bar{\mathbf{u}}|)\, dx\right) + c.
\end{aligned}$$

In the last line, we have used Jensen's inequality with the concave function $t \mapsto \bar{G}(t)^{\upsilon}$. Recalling (45), we obtain

$$\left(\int_{B_{7/8}} \left(\frac{\bar{G}(1+|\bar{\mathbf{u}}|)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\right)^{\frac{n}{n-\kappa}} dx\right)^{\frac{n-\kappa}{n}} \le c \tag{63}$$

for some $c = c(\mathtt{data}, \gamma) > 0$.

To proceed further, we recall H_s defined in (36) with s from (39). By (37) function H_s satisfies also $H_s(Mt) \approx_{\mathtt{data}} \bar{g}(t)^{1-s}\bar{G}(t)^s$. We apply Young's inequality, (53), and (63) with suitable choice of ϕ. In turn we see

$$
\begin{aligned}
\int_{B_{\rho_1}} H_s(M|D\bar{\mathbf{u}}|)\,dx &\lesssim_{\mathtt{data}} \int_{B_{\rho_1}} \frac{\bar{g}(|D\bar{\mathbf{u}}|)^{1-s}\bar{G}(|D\bar{\mathbf{u}}|)^s}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}(1+|\bar{\mathbf{u}}|)^{1+\gamma}\,dx\\
&\le \int_{B_{\rho_1}} \frac{\bar{G}(|D\bar{\mathbf{u}}|)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\,dx + \int_{B_{\rho_1}} \frac{\bar{g}(|D\bar{\mathbf{u}}|)(1+|\bar{\mathbf{u}}|)^{\frac{1+\gamma}{1-s}}}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\,dx\\
&\le \int_{B_{\rho_1}} \frac{\bar{G}(|D\bar{\mathbf{u}}|)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\,dx + \int_{B_{\rho_1}} \frac{\bar{G}(1+|\bar{\mathbf{u}}|)(1+|\bar{\mathbf{u}}|)^{\frac{(s+\gamma)s_G}{(1-s)}}}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\,dx\\
&\le \int_{B_{\rho_1}} \frac{\bar{G}(|D\bar{\mathbf{u}}|)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\,dx\\
&\quad + \left(\int_{B_{\rho_1}} \left(\frac{\bar{G}(1+|\bar{\mathbf{u}}|)}{(1+|\bar{\mathbf{u}}|)^{1+\gamma}}\right)^{\frac{n}{n-\kappa}}\,dx\right)^{\frac{n-\kappa}{n}} \left(\int_{B_{\rho_1}} (1+|\bar{\mathbf{u}}|)^{\frac{(s+\gamma)s_G n}{(1-s)\kappa}}\,dx\right)^{\frac{\kappa}{n}}
\end{aligned}
\tag{64}
$$

In order to use (45) we need to have

$$
\frac{(s+\gamma)s_G n}{(1-s)\kappa} \le 1. \tag{65}
$$

Observe that by Remark 4.2 we have $2-i_G < (i_G-\gamma s_G n)/(i_G+s_G n)$, and we can choose $\kappa \in [1, i_G)$ such that

$$
\max\{2-i_G, 0\} < s \le \frac{\kappa-\gamma s_G n}{\kappa+s_G n} < s_{\mathrm{m}}
$$

and the bound (65) follows. Then (64) combined with (53), (61) and (63) implies for $s < s_{\mathrm{m}}$ that

$$
\int_{B_{3/4}} H_s(M|D\bar{\mathbf{u}}|)\,dx \le c_{\mathrm{ap}} = c_{\mathrm{ap}}(\mathtt{data}, \gamma). \tag{66}
$$

Step 4. Contradiction argument. We state the counter–assumption.

Namely, we assume that there exists ε and sequences of balls $\{B_{r_j}(x_j)\}$ and almost $\mathcal{A}$-harmonic maps $\{\mathbf{u}_j\} \subset W^{1,G}(B_{r_j}(x_j), \mathbb{R}^m)$ such that

$$
\fint_{B_{r_j}(x_j)} |\mathbf{u}_j|\,dx \le Mr_j, \quad M \ge 1, \tag{67}
$$

$$
\left|\fint_{B_{r_j}(x_j)} \mathcal{A}(x, D\mathbf{u}_j) : D\boldsymbol{\varphi}\,dx\right| \le \frac{2^{-j}}{r_j}\|\boldsymbol{\varphi}\|_{L^\infty(B_{r_j}(x_j),\mathbb{R}^m)} \tag{68}
$$

for all $\boldsymbol{\varphi} \in W_0^{1,G}(B_{r_j}(x_j), \mathbb{R}^m) \cap L^\infty(B_{r_j}(x_j), \mathbb{R}^m)$, but such that

$$
\fint_{B_{r_j/2}(x_j)} H_s(|D\mathbf{u}_j - D\mathbf{v}|)\,dx > \varepsilon \tag{69}
$$

whenever $\mathbf{v} \in W^{1,G}(B_{r/2}(x_0), \mathbb{R}^m)$ is an $\mathcal{A}$-harmonic map in $B_{r/2}(x_0)$ satisfying

$$
\fint_{B_{r_j/2}(x_0)} |\mathbf{v}|\,dx \le 2^n Mr_j \quad \text{and} \quad \left(\fint_{B_{r_j/2}(x_j)} H_s(M|D\mathbf{v}|)\,dx\right) \le c, \tag{70}
$$

where $c = c(\mathtt{data}) > 0$.

Let $\bar{\mathbf{u}}$ be scaled as in (44), but with the use of x_j and r_j, that is we set

$$\bar{\mathbf{u}}_j(x) := \frac{\mathbf{u}(x_j + r_j x)}{M r_j} \quad \text{and} \quad \bar{\mathcal{A}}(x_j + r_j x, \boldsymbol{\xi}) := \mathcal{A}(x_j + r_j x, M\boldsymbol{\xi}).$$

In such a case by (67) we get that

$$\fint_{B_1} |\bar{\mathbf{u}}_j|\, dx \le 1, \tag{71}$$

so by (68) we infer that for all $\boldsymbol{\eta} = \boldsymbol{\varphi}(x_j + r_j x)/r_j \in W_0^{1,G}(B_1, \mathbb{R}^m) \cap L^\infty(B_1, \mathbb{R}^m)$ it holds

$$\left| \fint_{B_1} \bar{\mathcal{A}}(x, D\bar{\mathbf{u}}_j) : D\boldsymbol{\eta}\, dx \right| \le 2^{-j} \|\boldsymbol{\eta}\|_{L^\infty(B_1, \mathbb{R}^m)} \tag{72}$$

and

$$\fint_{B_1} H_s(M|D\bar{\mathbf{u}}_j - D\bar{\mathbf{v}}|)\, dx > \varepsilon \tag{73}$$

whenever $\bar{\mathbf{v}} \in W^{1,G}(B_{1/2}(x_0), \mathbb{R}^m)$ is an $\bar{\mathcal{A}}$-harmonic map in $B_{1/2}(x_0)$ satisfying

$$\fint_{B_{1/2}} |\bar{\mathbf{v}}|\, dx \le 2^n \quad \text{and} \quad \left(\fint_{B_{1/2}} H_s(M|D\bar{\mathbf{v}}|)\, dx \right) \le c, \tag{74}$$

where $c = c(\mathtt{data}) > 0$.

Since $\bar{\mathbf{u}}$ satisfies (71), we have (66) for $s_o < s_m$ from (39). Therefore

$$\int_{B_{3/4}} H_{s_o}(M|D\bar{\mathbf{u}}_j|)\, dx \le C \quad \text{for} \quad C = C(\mathtt{data}, \gamma).$$

We fix any $s < s_o$ from the range (39). Then we pick $\epsilon > 0$ for which there exist $c, t_0 > 0$, such that $H_s(t) \ge c t^{1+\epsilon}$ and $H_s(t) \ge c g^{1+\epsilon}(t)$ for all $t \ge t_0$. In turn, we conclude with the following estimates uniform in j

$$\int_{B_{3/4}} g^{1+\epsilon}(M|D\bar{\mathbf{u}}_j|)\, dx \le c_1 \quad \text{and} \quad \int_{B_{3/4}} (M|D\bar{\mathbf{u}}_j|)^{1+\epsilon}\, dx \le c_2 \tag{75}$$

with c_1, c_2 depending on $\mathtt{data}$ and γ only. Further we infer that there exist

$$\widetilde{\mathbf{u}} \in W^{1,H_s}(B_{3/4}, \mathbb{R}^m), \quad \mathfrak{A} \in L^{1+\epsilon}(B_{3/4}, \mathbb{R}^{n\times m}), \quad \text{and} \quad \mathfrak{h} \in L^{H_s}(B_{3/4})$$

such that up to a subsequence

$$\begin{aligned}
&D\bar{\mathbf{u}}_j - D\widetilde{\mathbf{u}} \rightharpoonup 0 \quad \text{in } L^{H_s}(B_{3/4}, \mathbb{R}^{n\times m}),\\
&|D\bar{\mathbf{u}}_j - D\widetilde{\mathbf{u}}| \rightharpoonup \mathfrak{h} \quad \text{in } L^{H_s}(B_{3/4}),\\
&\bar{\mathcal{A}}(x, D\bar{\mathbf{u}}_j) \rightharpoonup \mathfrak{A} \quad \text{in } L^{1+\epsilon}(B_{3/4}, \mathbb{R}^{n\times m}),\\
&\bar{\mathbf{u}}_j \to \widetilde{\mathbf{u}} \quad \text{strongly in } L^{H_s}(B_{3/4}, \mathbb{R}^m) \text{ and a.e. in } B_{3/4}.
\end{aligned} \tag{76}$$

By (71), lower semicontinuity of a functional $\boldsymbol{\varphi} \mapsto \fint_{B_{1/2}} H_s(M|D\boldsymbol{\varphi}|)\, dx$, and (66) we have

$$\fint_{B_{1/2}} |\widetilde{\mathbf{u}}|\, dx \le 2^n \quad \text{and} \quad \fint_{B_{1/2}} H_s(M|D\widetilde{\mathbf{u}}|)\, dx \le c, \tag{77}$$

Step 5. Strong convergence of gradients. Our aim is now to prove that

$$D\bar{\mathbf{u}}_j \to D\widetilde{\mathbf{u}} \quad \text{ in } L^{H_s}(B_{3/4}, \mathbb{R}^{n\times m}). \tag{78}$$

For this we need to show that $\mathfrak{h} \in L^{H_s}(B_{3/4})$ from (76) satisfies

$$\mathfrak{h} = 0 \tag{79}$$

a.e. in $B_{3/4}$. This almost everywhere and weak convergence in L^1 implies strong L^1-convergence of $D\bar{\mathbf{u}}_j \to D\widetilde{\mathbf{u}}$ in $B_{3/4}$. Using the monotonicity property of H_s and (74), for sufficiently small $\widetilde{\epsilon} \in (0, \frac{1}{s_{H_s}})$ we have

$$\begin{aligned}
&\int_{B_{3/4}} H_s(M|D\bar{\mathbf{u}}_j - D\widetilde{\mathbf{u}}|)\,dx \\
&\le \int_{B_{3/4}} H_s(M|D\bar{\mathbf{u}}_j - D\widetilde{\mathbf{u}}|)^{\widetilde{\epsilon}^2} H_s(M|D\bar{\mathbf{u}}_j| + M|D\widetilde{\mathbf{u}}|)^{1-\widetilde{\epsilon}^2}\,dx \\
&\le \left(\int_{B_{3/4}} H_s(M|D\bar{\mathbf{u}}_j - D\widetilde{\mathbf{u}}|)^{\widetilde{\epsilon}}\,dx\right)^{\widetilde{\epsilon}} \left(\int_{B_{3/4}} H_s(M|D\bar{\mathbf{u}}_j| + M|D\widetilde{\mathbf{u}}|)^{1+\widetilde{\epsilon}}\,dx\right)^{1-\widetilde{\epsilon}} \\
&\le c\left(\int_{B_{3/4}} H_s(M|D\bar{\mathbf{u}}_j - D\widetilde{\mathbf{u}}|)^{\widetilde{\epsilon}}\,dx\right)^{\widetilde{\epsilon}}.
\end{aligned} \tag{80}$$

Denoting

$$\Psi(t) = \int_0^t \frac{H_s^{-1}(\tau^{1/\widetilde{\epsilon}})}{\tau}\,d\tau,$$

one can immediately check

$$\frac{t\Psi''(t)}{\Psi'(t)} = \frac{t^{1/\widetilde{\epsilon}}}{\widetilde{\epsilon} H_s'(H_s^{-1}(t^{1/\widetilde{\epsilon}}))H_s^{-1}(t^{1/\widetilde{\epsilon}})} - 1 \ge \frac{1}{\widetilde{\epsilon} s_{H_s}} - 1 > 0,$$

and so Ψ is a Young function. We then apply Jensen's inequality to (80) to obtain

$$\int_{B_{3/4}} H_s(M|D\bar{\mathbf{u}}_j - D\widetilde{\mathbf{u}}|)\,dx \le c\left[H_s\left(M\int_{B_{3/4}} |D\bar{\mathbf{u}}_j - D\widetilde{\mathbf{u}}|\,dx\right)\right]^{\widetilde{\epsilon}^2} \xrightarrow{j\to\infty} 0.$$

Hence, it remains to show (79) to obtain (78).

We pick $\bar{x}$ being a Lebesgue's point simultaneously for $\widetilde{\mathbf{u}}$, $D\widetilde{\mathbf{u}}$, $\mathfrak{h}$, $\mathfrak{A}$, that is

$$\begin{aligned}
\lim_{\varrho\to 0} \fint_{B_\varrho(\bar{x})} & H_s(M|\widetilde{\mathbf{u}} - \widetilde{\mathbf{u}}(\bar{x})|) + H_s(M|D\widetilde{\mathbf{u}} - D\widetilde{\mathbf{u}}(\bar{x})|) \\
& + H_s(M|\mathfrak{h} - \mathfrak{h}(\bar{x})|) + |\mathfrak{A} - \mathfrak{A}(\bar{x})|^{1+\epsilon}\,dx = 0
\end{aligned} \tag{81}$$

and

$$|\widetilde{\mathbf{u}}(\bar{x})| + |D\widetilde{\mathbf{u}}(\bar{x})| + |\mathfrak{h}(\bar{x})| + |\mathfrak{A}(\bar{x})| < \infty. \tag{82}$$

Almost every point of $B_{3/4}$ satisfies this conditions. Thus it is enough to show that (79) holds for $\bar{x}$.

We restrict our attention to ϱ small enough for $B_\varrho(\bar{x}) \subset B_{3/4}$ and we set the linearization of $\widetilde{\mathbf{u}}$ at $\bar{x}$

$$\boldsymbol{\ell}_{\boldsymbol{\varrho}}(x) := (\widetilde{\mathbf{u}})_{B_\varrho(\bar{x})} + D\widetilde{\mathbf{u}}(\bar{x}) : (x - \bar{x}). \tag{83}$$

Having the classical Poincaré inequality and (81), we obtain that

$$\lim_{\varrho\to 0} \fint_{B_\varrho(\bar x)} \left|\frac{\widetilde{\mathbf{u}}-\boldsymbol{\ell}_\varrho}{\varrho}\right|^{1+\epsilon} dx \le c \lim_{\varrho\to 0} \fint_{B_\varrho(\bar x)} |D\widetilde{\mathbf{u}} - D\widetilde{\mathbf{u}}(\bar x)|^{1+\epsilon}\, dx = 0. \tag{84}$$

Let us set

$$\mathrm{I}^0_{j,\varrho} := \fint_{B_{\varrho/2}(\bar x)} |D\bar{\mathbf{u}}_j - D\widetilde{\mathbf{u}}|\, dx.$$

By (76) we have the weak convergence of $|D\bar{\mathbf{u}}_j - D\widetilde{\mathbf{u}}| \to \mathfrak{h}$ in $L^{H_s}(B_\varrho)$. Since $\bar x$ is a Lebesgue's point of $D\widetilde{\mathbf{u}}$ we infer that

$$\mathfrak{h}(\bar x) = \lim_{\varrho\to 0}\lim_{j\to\infty} \mathrm{I}^0_{j,\varrho}. \tag{85}$$

In order to prove that $\mathfrak{h}(\bar x) = 0$, let us write

$$\begin{aligned}\mathrm{I}^0_{j,\varrho} = \fint_{B_{\varrho/2}(\bar x)} |D\bar{\mathbf{u}}_j - D\widetilde{\mathbf{u}}|\, dx &= \fint_{B_{\varrho/2}(\bar x)} \mathbb{1}_{\{|\bar{\mathbf{u}}_j-\boldsymbol{\ell}_\varrho|\ge\varrho\}} |D\bar{\mathbf{u}}_j - D\widetilde{\mathbf{u}}|\, dx\\ &\quad+ \fint_{B_{\varrho/2}(\bar x)} \mathbb{1}_{\{|\bar{\mathbf{u}}_j-\boldsymbol{\ell}_\varrho|<\varrho\}} |D\bar{\mathbf{u}}_j - D\widetilde{\mathbf{u}}|\, dx := \mathrm{I}^1_{j,\varrho} + \mathrm{I}^2_{j,\varrho}\end{aligned} \tag{86}$$

and prove the convergence of both terms first when $j\to\infty$ and then $\varrho\to 0$.

We start with $\mathrm{I}^1_{j,\varrho}$. Let us observe that

$$\begin{aligned}\mathrm{I}^1_{j,\varrho} &\le \fint_{B_{\varrho/2}(\bar x)} \mathbb{1}_{\{|\bar{\mathbf{u}}_j-\widetilde{\mathbf{u}}|\ge\varrho\}} |D\bar{\mathbf{u}}_j - D\widetilde{\mathbf{u}}|\, dx\\ &\quad+ \fint_{B_{\varrho/2}(\bar x)} \mathbb{1}_{\{|\widetilde{\mathbf{u}}-\boldsymbol{\ell}_\varrho|\ge\varrho\}} |D\bar{\mathbf{u}}_j - D\widetilde{\mathbf{u}}|\, dx =: \mathrm{I}^{1,1}_{j,\varrho} + \mathrm{I}^{1,2}_{j,\varrho}.\end{aligned}$$

Notice that $\mathrm{I}^{1,1}_{j,\varrho}$ vanishes as $j\to\infty$. Indeed, $\mathrm{I}^{1,1}_{j,\varrho}\ge 0$ and by Hölder inequality we have

$$\mathrm{I}^{1,1}_{j,\varrho} \le \frac{1}{|B_{\varrho/2}(\bar x)|}\left(\int_{B_{\varrho/2}(\bar x)} |D\bar{\mathbf{u}}_j - D\widetilde{\mathbf{u}}|^{1+\epsilon}\, dx\right)^{\frac{1}{1+\epsilon}} \left(|\{x\in B_{3/4}:\ |\bar{\mathbf{u}}_j - \widetilde{\mathbf{u}}|\ge\varrho/2\}|\right)^{\frac{\epsilon}{1+\epsilon}}.$$

Since by (76) one has that $|\bar{\mathbf{u}}_j - \widetilde{\mathbf{u}}|\to 0$ strongly in $L^1(B_{3/4})$, so

$$\lim_{j\to\infty} |\{x\in B_{3/4}:\ |\bar{\mathbf{u}}_j - \widetilde{\mathbf{u}}|\ge\varrho/2\}| = 0.$$

The rest of the terms are bounded as ϵ is chosen such that (75) is true and $|D\widetilde{\mathbf{u}}|$ shares the same a priori estimates as $|D\bar{\mathbf{u}}_j|$. Therefore, we infer that $\lim_{j\to\infty}\mathrm{I}^{1,1}_{j,\varrho} = 0$. On the other hand, $\mathrm{I}^{1,2}_{j,\varrho}$ is convergent when $j\to\infty$, because of the weak convergence of $|D\bar{\mathbf{u}}_j - D\widetilde{\mathbf{u}}|\to\mathfrak{h}$ in $L^{H_s}(B_\varrho)$. Hence, we get

$$\lim_{j\to\infty}\mathrm{I}^{1,2}_{j,\varrho} = \fint_{B_{\varrho/2}(\bar x)} \mathbb{1}_{\{|\widetilde{\mathbf{u}}-\boldsymbol{\ell}_\varrho|\ge\varrho\}}\mathfrak{h}\, dx =: \mathrm{I}^{1,2}_{\varrho}.$$

We can estimate further

$$\begin{aligned}\mathrm{I}_{\varrho}^{1,2} &\le \left(\fint_{B_\varrho(\bar{x})} \mathfrak{h}^{1+\epsilon}\,dx\right)^{\frac{1}{1+\epsilon}} \left(\fint_{B_\varrho(\bar{x})} \mathbb{1}_{\{|\widetilde{\mathbf{u}}-\boldsymbol{\ell}_{\boldsymbol{\varrho}}|\ge\varrho/2\}}\,dx\right)^{\frac{\epsilon}{1+\epsilon}}\\
&\le c\left[\left(\fint_{B_\varrho(\bar{x})} |\mathfrak{h}-\mathfrak{h}(\bar{x})|^{1+\epsilon}\,dx\right)^{\frac{1}{1+\epsilon}} + \mathfrak{h}(\bar{x})\right]\left(\fint_{B_\varrho(\bar{x})} \left|\frac{\widetilde{\mathbf{u}}-\boldsymbol{\ell}_{\boldsymbol{\varrho}}}{\varrho}\right|^{1+\epsilon} dx\right)^{\frac{\epsilon}{1+\epsilon}}\end{aligned}$$

which tends to 0 as $\varrho \to 0$ as $\bar{x}$ is a Lebesgue's point of $\mathfrak{h}$ as in (81) and the last bracket converges to 0 due to (84). Altogether, we have that $\mathrm{I}_{j,\varrho}^{1}$ vanishes in the limit, so we will now concentrate on $\mathrm{I}_{j,\varrho}^{2}$ for which we have

$$\begin{aligned}\mathrm{I}_{j,\varrho}^{2} &\le \fint_{B_{\varrho/2}(\bar{x})} \mathbb{1}_{\{|\bar{\mathbf{u}}_j-\boldsymbol{\ell}_{\boldsymbol{\varrho}}|<\varrho\}}|D\bar{\mathbf{u}}_j - D\boldsymbol{\ell}_{\boldsymbol{\varrho}}|\,dx + 2^n\fint_{B_\varrho(\bar{x})} \mathbb{1}_{\{|\bar{\mathbf{u}}_j-\boldsymbol{\ell}_{\boldsymbol{\varrho}}|<\varrho\}}|D\boldsymbol{\ell}_{\boldsymbol{\varrho}} - D\widetilde{\mathbf{u}}|\,dx\\
&=: \mathrm{I}_{j,\varrho}^{2,1} + \mathrm{I}_{j,\varrho}^{2,2}.\end{aligned}$$

By (76), (81), and (83) we have that $\lim_{\varrho\to 0}\limsup_{j\to\infty} \mathrm{I}_{j,\varrho}^{2,2} = 0$. Proving the convergence

$$\limsup_{\varrho\to 0}\limsup_{j\to\infty} \mathrm{I}_{j,\varrho}^{2,1} = 0 \tag{87}$$

requires more arguments. We take

$$\phi \in C_c^\infty(B_\varrho(\bar{x})) \quad \text{with}\, 0\le\phi\le 1,\ \phi\equiv 1 \,\text{on}\, B_{\varrho/2}(\bar{x}) \ \text{and}\ |D\phi|\le 4/\varrho.$$

Let

$$\boldsymbol{\eta} = \phi T_\varrho(\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}}),$$

where the truncation is defined in (17). Let us denote

$$\mathsf{P_j} := \frac{(\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}})\otimes(\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}})}{|\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|^2} \qquad \text{and} \qquad \mathsf{P} := \frac{(\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}})\otimes(\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}})}{|\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|^2} \tag{88}$$

when $|\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}}| \ne 0$ and $|\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}| \ne 0$, respectively. Within this notation we have that

$$\begin{aligned}&\big(\bar{\mathcal{A}}(x, D\bar{\mathbf{u}}_j) - \bar{\mathcal{A}}(x, D\boldsymbol{\ell}_{\boldsymbol{\varrho}})\big) : D\boldsymbol{\eta}\\
&\quad= \mathbb{1}_{\{|\bar{\mathbf{u}}_j-\boldsymbol{\ell}_{\boldsymbol{\varrho}}|<\varrho\}}\left[\big(\bar{\mathcal{A}}(x, D\bar{\mathbf{u}}_j) - \bar{\mathcal{A}}(x, D\boldsymbol{\ell}_{\boldsymbol{\varrho}})\big) : D(\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}})\right]\phi\\
&\qquad+ \frac{\varrho\mathbb{1}_{\{|\bar{\mathbf{u}}_j-\boldsymbol{\ell}_{\boldsymbol{\varrho}}|\ge\varrho\}}}{|\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|}\left[\big(\bar{\mathcal{A}}(x, D\bar{\mathbf{u}}_j) - \bar{\mathcal{A}}(x, D\boldsymbol{\ell}_{\boldsymbol{\varrho}})\big) : (\mathsf{Id}-\mathsf{P_j})D(\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}})\right]\phi\\
&\qquad+ \big(\bar{\mathcal{A}}(x, D\bar{\mathbf{u}}_j) - \bar{\mathcal{A}}(x, D\boldsymbol{\ell}_{\boldsymbol{\varrho}})\big) : \left[T_\varrho(\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}})\otimes D\phi\right]\\
&\quad=: G_{j,\varrho}^1(x) + G_{j,\varrho}^2(x) + G_{j,\varrho}^3(x).\end{aligned}$$

Moreover, we recall that since $\boldsymbol{\ell}_{\boldsymbol{\varrho}}$ is affine, whenever $B \subset B_1$ it holds that

$$\int_B \bar{\mathcal{A}}(x, D\boldsymbol{\ell}_{\boldsymbol{\varrho}}) : D\boldsymbol{\varphi}\,dx = 0 \qquad \text{for every } \boldsymbol{\varphi}\in W_0^{1,1}(B,\mathbb{R}^m).$$

Therefore, by (72) it is justified to write that

$$0 \le \fint_{B_\varrho(\bar{x})} G_{j,\varrho}^1(x)\,dx \le 2^{-j}\varrho^{1-n} - \fint_{B_\varrho(\bar{x})} G_{j,\varrho}^2(x)\,dx - \fint_{B_\varrho(\bar{x})} G_{j,\varrho}^3(x)\,dx. \tag{89}$$

The first term in the above display is nonnegative because of the monotonicity of $\bar{\mathcal{A}}$. Instrumental for proving that $\mathrm{I}^{2,1}_{j,\varrho} \to 0$ is to establish that

$$\limsup_{\varrho\to 0}\limsup_{j\to\infty} \fint_{B_\varrho(\bar{x})} G^1_{j,\varrho}(x)\,dx = 0, \tag{90}$$

which will be proven provided one justifies that the last two terms of (89) vanish in the limit. We will show first that

$$\limsup_{\varrho\to 0}\limsup_{j\to\infty} \left(-\fint_{B_\varrho(\bar{x})} G^2_{j,\varrho}\,dx\right) \le 0. \tag{91}$$

The quasi-diagonal structure of $\bar{\mathcal{A}}$ ensures that $\bar{\mathcal{A}}(x, D\bar{\mathbf{u}}_j) : [(\mathsf{Id}-\mathsf{P_j})D\bar{\mathbf{u}}_j] \ge 0$, see (49). Therefore,

$$\begin{aligned}&\big(\bar{\mathcal{A}}(x, D\bar{\mathbf{u}}_j) - \bar{\mathcal{A}}(x, D\boldsymbol{\ell}_{\boldsymbol{\varrho}})\big) : (\mathsf{Id}-\mathsf{P_j})D(\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}})\\ &\quad\ge -\bar{\mathcal{A}}(x, D\bar{\mathbf{u}}_j) : (\mathsf{Id}-\mathsf{P_j})D\boldsymbol{\ell}_{\boldsymbol{\varrho}} - \bar{\mathcal{A}}(x, D\boldsymbol{\ell}_{\boldsymbol{\varrho}}) : (\mathsf{Id}-\mathsf{P_j})D(\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}}).\end{aligned} \tag{92}$$

Recall that $\mathsf{P_j}$ and P, defined in (88), are bounded. Notice that for $j \to \infty$ we have $\mathbb{1}_{\{|\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|\ge\varrho\}}\mathsf{P_j} \to \mathbb{1}_{\{|\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|\ge\varrho\}}\mathsf{P}$ almost everywhere and thus, by the Lebesgue's dominated convergence theorem, also strongly in $L^t(B_{3/4})$ for every $t \ge 1$. Moreover, $\mathbb{1}_{\{|\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|\ge\varrho\}}|\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|^{-1} \to \mathbb{1}_{\{|\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|\ge\varrho\}}|\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|^{-1}$ almost everywhere and, as a uniformly bounded sequence of functions, it converges also strongly in $L^t(B_{3/4})$ for every $t \ge 1$. Having this, (92), and (76), we obtain

$$\begin{aligned}\limsup_{j\to\infty}\left(-\fint_{B_\varrho(\bar{x})} G^2_{j,\varrho}\,dx\right) &\le \fint_{B_\varrho(\bar{x})} \mathfrak{A} : (\mathsf{Id}-\mathsf{P})D\boldsymbol{\ell}_{\boldsymbol{\varrho}} \frac{\varrho\mathbb{1}_{\{|\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|\ge\varrho\}}}{|\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|}\,dx\\ &\quad+\fint_{B_\varrho(\bar{x})} \bar{\mathcal{A}}(x, D\boldsymbol{\ell}_{\boldsymbol{\varrho}}) : (\mathsf{Id}-\mathsf{P})D(\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}})\frac{\varrho\mathbb{1}_{\{|\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|\ge\varrho\}}}{|\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|}\,dx =: \mathrm{II}^1_\varrho + \mathrm{II}^2_\varrho.\end{aligned} \tag{93}$$

We can estimate

$$|\mathrm{II}^1_\varrho| \le c\fint_{B_\varrho(\bar{x})} |\mathfrak{A}| \left|\frac{\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}}{\varrho}\right|^\epsilon dx \le c\left(\fint_{B_\varrho(\bar{x})} |\mathfrak{A}|^{1+\epsilon}\,dx\right)^{\frac{1}{1+\epsilon}} \left(\fint_{B_\varrho(\bar{x})} \left|\frac{\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}}{\varrho}\right|^{1+\epsilon} dx\right)^{\frac{\epsilon}{1+\epsilon}}$$

where the first term is bounded and the second term is convergent to zero by (84). On the other hand, by (83) and (82) we may estimate

$$\begin{aligned}|\mathrm{II}^2_\varrho| &\le c\fint_{B_\varrho(\bar{x})} |D(\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}})| \left|\frac{\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}}{\varrho}\right|^\epsilon dx\\ &\le c\left(\fint_{B_\varrho(\bar{x})} |D\widetilde{\mathbf{u}} - D\widetilde{\mathbf{u}}(\bar{x})|^{1+\epsilon}\,dx\right)^{\frac{1}{1+\epsilon}} \left(\fint_{B_\varrho(\bar{x})} \left|\frac{\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}}{\varrho}\right|^{1+\epsilon} dx\right)^{\frac{\epsilon}{1+\epsilon}}\end{aligned}$$

where, again, the first term is bounded and the second convergent to zero by (84). Summing up the information from the last three displays we get (91).

Now we concentrate on justifying that

$$\lim_{\varrho\to 0}\lim_{j\to\infty} \left|\fint_{B_\varrho(\bar{x})} G^3_{j,\varrho}(x)\,dx\right| = 0. \tag{94}$$

Let us observe that because (76) provides weak convergence of $\bar{A}(x, D\bar{\mathbf{u}}_j)$ to $\mathfrak{A}$ in $L^{1+\epsilon}$ and strong convergence $\bar{\mathbf{u}}_j$ to $\widetilde{\mathbf{u}}$ in $L^{1+\epsilon}$, we obtain that

$$\lim_{j\to\infty} \fint_{B_\varrho(\bar{x})} G^3_{j,\varrho}(x)\,dx = \fint_{B_\varrho(\bar{x})} \left(\mathfrak{A} - \bar{A}(x, D\boldsymbol{\ell}_{\boldsymbol{\varrho}})\right) : [T_\varrho(\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}) \otimes D\phi]\,dx.$$

By Hölder inequality and the choice of ϕ, we estimate further

$$\lim_{j\to\infty} \fint_{B_\varrho(\bar{x})} G^3_{j,\varrho}(x)\,dx \le \left(\fint_{B_\varrho(\bar{x})} |\mathfrak{A} - \mathfrak{A}(\bar{x})|^{1+\epsilon} + |\mathfrak{A}(\bar{x})|^{1+\epsilon} + g(|D\widetilde{\mathbf{u}}(\bar{x})|)^{1+\epsilon}\,dx\right)^{\frac{1}{1+\epsilon}}$$
$$\cdot \left(\fint_{B_\varrho(\bar{x})} \left(\frac{\min\{\varrho, |\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|\}}{\varrho}\right)^{\frac{1+\epsilon}{\epsilon}} dx\right)^{\frac{\epsilon}{1+\epsilon}},$$

where the first integral on the right-hand side is finite and the second term converges to zero. Indeed, since $0 \le \min\{\varrho, |\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|\} \le \varrho$, we have

$$\fint_{B_\varrho(\bar{x})} \left(\frac{\min\{\varrho, |\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|\}}{\varrho}\right)^{\frac{1+\epsilon}{\epsilon}} dx \le \fint_{B_\varrho(\bar{x})} \left(\frac{\min\{\varrho, |\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|\}}{\varrho}\right)^{1+\epsilon} dx$$
$$\le \fint_{B_\varrho(\bar{x})} \left(\frac{|\widetilde{\mathbf{u}} - \boldsymbol{\ell}_{\boldsymbol{\varrho}}|}{\varrho}\right)^{1+\epsilon} dx \xrightarrow[\varrho\to 0]{} 0,$$

where the last convergence results from (84). Therefore, we get (94).

We have shown (91) and (94), so because of (89) the limit (90) follows. Hence, we are in the position to prove (87). In the view of (21), (90) implies that

$$\limsup_{\varrho\to\infty} \limsup_{j\to\infty} \fint_{B_\varrho(\bar{x})} \mathbb{1}_{\{|\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}}| < \varrho\}} \frac{g(|D\bar{\mathbf{u}}_j| + |D\boldsymbol{\ell}_{\boldsymbol{\varrho}}|)}{|D\bar{\mathbf{u}}_j| + |D\boldsymbol{\ell}_{\boldsymbol{\varrho}}|} |D(\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}})|^2\,dx = 0. \tag{95}$$

At this stage, we calculate similarly to (80) in order to show (87). For any $\hat{\epsilon} < \frac{2}{s_G}$, it is readily checked that $t \mapsto tg(t)^{-\hat{\epsilon}/(2-\hat{\epsilon})}$ is a monotone increasing function. Then

$$\fint_{B_\varrho(\bar{x})} \mathbb{1}_{\{|\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}}| < \varrho\}} |D(\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}})|\,dx$$
$$\le \fint_{B_\varrho(\bar{x})} \mathbb{1}_{\{|\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}}| < \varrho\}} \left(|D(\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}})|^2 \frac{g(|D\bar{\mathbf{u}}_j| + |D\boldsymbol{\ell}_{\boldsymbol{\varrho}}|)}{|D\bar{\mathbf{u}}_j| + |D\boldsymbol{\ell}_{\boldsymbol{\varrho}}|}\right)^{\frac{\hat{\epsilon}}{2}} \frac{(|D\bar{\mathbf{u}}_j| + |D\boldsymbol{\ell}_{\boldsymbol{\varrho}}|)^{1-\frac{\hat{\epsilon}}{2}}}{g(|D\bar{\mathbf{u}}_j| + |D\boldsymbol{\ell}_{\boldsymbol{\varrho}}|)^{\frac{\hat{\epsilon}}{2}}}\,dx$$
$$\le \left(\fint_{B_\varrho(\bar{x})} \mathbb{1}_{\{|\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}}| < \varrho\}} |D(\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}})|^2 \frac{g(|D\bar{\mathbf{u}}_j| + |D\boldsymbol{\ell}_{\boldsymbol{\varrho}}|)}{|D\bar{\mathbf{u}}_j| + |D\boldsymbol{\ell}_{\boldsymbol{\varrho}}|}\,dx\right)^{\frac{\hat{\epsilon}}{2}}$$
$$\cdot \left(\fint_{B_\varrho(\bar{x})} \mathbb{1}_{\{|\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}}| < \varrho\}} \frac{|D\bar{\mathbf{u}}_j| + |D\boldsymbol{\ell}_{\boldsymbol{\varrho}}|}{g(|D\bar{\mathbf{u}}_j| + |D\boldsymbol{\ell}_{\boldsymbol{\varrho}}|)^{\frac{\hat{\epsilon}}{2-\hat{\epsilon}}}}\,dx\right)^{1-\frac{\hat{\epsilon}}{2}}$$
$$\le c \left(\fint_{B_\varrho(\bar{x})} \mathbb{1}_{\{|\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}}| < \varrho\}} |D(\bar{\mathbf{u}}_j - \boldsymbol{\ell}_{\boldsymbol{\varrho}})|^2 \frac{g(|D\bar{\mathbf{u}}_j| + |D\boldsymbol{\ell}_{\boldsymbol{\varrho}}|)}{|D\bar{\mathbf{u}}_j| + |D\boldsymbol{\ell}_{\boldsymbol{\varrho}}|}\,dx\right)^{\frac{\hat{\epsilon}}{2}}$$
$$\cdot \left(\fint_{B_\varrho(\bar{x})} (|D\bar{\mathbf{u}}_j| + |D\boldsymbol{\ell}_{\boldsymbol{\varrho}}| + 1)\,dx\right)^{1-\frac{\hat{\epsilon}}{2}},$$

where $c = c(g) > 0$. Noting that the very last term in the above display is bounded, by (95) we infer that (87) holds.

Summing up all the convergences of this step, we get in (85) that $\mathfrak{h}(\bar{x}) = 0$ and, consequently, (79) holds almost everywhere in $B_{3/4}$. As explained in the beginning of this step, this suffices to get the final aim of **Step 5**, that is strong convergence of gradients (78).

Step 6. $\bar{\mathcal{A}}$-harmonicity of the limit map and conclusion by contradiction. Having (78), we can pass to the limit in (72) with $j \to \infty$ getting that

$$\fint_{B_{1/2}} \bar{\mathcal{A}}(x, D\widetilde{\mathbf{u}}) : D\boldsymbol{\eta}\, dx = 0 \qquad \text{for} \quad \boldsymbol{\eta} \in C_c^\infty(B_{1/2}, \mathbb{R}^m). \tag{96}$$

Therefore, if $\widetilde{\mathbf{u}} \in W^{1,G}(B_{1/2}, \mathbb{R}^m)$, then it will be proven to be $\bar{\mathcal{A}}$-harmonic. Indeed, since we know (77), it is allowed to take $\bar{\mathbf{v}} = \widetilde{\mathbf{u}}$ in (73). Note that in such a case (77) is precisely the restriction on the test function from (74). Then, in the view of (78), taking j large enough, we will get the desired contradiction. Hence, it remains to prove that $|D\widetilde{\mathbf{u}}| \in L^G(B_{1/2})$.

We have (63) for each $\bar{\mathbf{u}}_j$ with the constant independent of j, so by the lower semicontinuity we can write that

$$\begin{aligned}\left(\fint_{B_{7/8}} \left(\frac{\bar{G}(1+|\widetilde{\mathbf{u}}|)}{(1+|\widetilde{\mathbf{u}}|)^{1+\widetilde{\gamma}}}\right)^{\frac{n}{n-\kappa}} dx\right)^{\frac{n-\kappa}{n}} &\le \liminf_{j\to\infty} \left(\fint_{B_{7/8}} \left(\frac{\bar{G}(1+|\bar{\mathbf{u}}_j|)}{(1+|\bar{\mathbf{u}}_j|)^{1+\widetilde{\gamma}}}\right)^{\frac{n}{n-\kappa}} dx\right)^{\frac{n-\kappa}{n}} \\ &\le c\end{aligned}$$

for some $c = c(\mathtt{data}, \widetilde{\gamma}) > 0$. Analogically, by (50) for $\bar{\mathbf{u}}_j$ and with $\delta = 2^{-j}$, by letting $j \to \infty$, Fatou's lemma on the left-hand side of the resultant inequality and (78) on its right-hand side, we get

$$\begin{aligned}\int_{B_{3/4}\cap\{|\widetilde{\mathbf{u}}|<t\}} \bar{G}(|D\widetilde{\mathbf{u}}|)\phi^q\, dx \le c_* &\int_{B_{3/4}\cap\{|\widetilde{\mathbf{u}}|<t\}} \bar{G}(|\widetilde{\mathbf{u}}|\,|D\phi|)\, dx \\ &+ c_*\, t \int_{B_{3/4}\cap\{|\widetilde{\mathbf{u}}|\ge t\}} \bar{g}(|D\widetilde{\mathbf{u}}|)\phi^{q-1}|D\phi|\, dx\end{aligned}$$

for every $t > 0$ and $\phi \in C_c^\infty(B_{3/4})$ with $\phi \ge 0$, and $c_* = c_*(\mathtt{data}, q)$. We proceed as in the beginning of **Step 3**. We multiply the above display by $(1+t)^{-(\widetilde{\gamma}+1)}$, $\widetilde{\gamma} > 0$ to be chosen sufficiently small in a few lines, integrate it from zero to infinity and apply Cavalieri's principle (Lemma 3.6) twice (with $\nu_1 = \bar{G}(|D\widetilde{\mathbf{u}}|)\phi^q$ and $\nu_2 = \bar{G}(|\widetilde{\mathbf{u}}|\,|D\phi|)$. Altogether we get

$$\begin{aligned}\frac{1}{\widetilde{\gamma}} \int_{B_{3/4}} \frac{\bar{G}(|D\widetilde{\mathbf{u}}|)\phi^q}{(1+|\widetilde{\mathbf{u}}|)^{\widetilde{\gamma}}}\, dx &\le \frac{c_*}{\widetilde{\gamma}} \int_{B_{3/4}} \frac{\bar{G}(|\widetilde{\mathbf{u}}|\,|D\phi|)}{(1+|\widetilde{\mathbf{u}}|)^{\widetilde{\gamma}}}\, dx \\ &+ c_* \int_0^\infty \frac{1}{(1+t)^{\widetilde{\gamma}}} \int_{B_{3/4}\cap\{|\widetilde{\mathbf{u}}|\ge t\}} \bar{g}(|D\widetilde{\mathbf{u}}|)\phi^{q-1}|D\phi|\, dx\, dt = \mathrm{III}_1 + \mathrm{III}_2.\end{aligned}$$

To estimate further the very last term we note that q is large enough to satisfy $s_G' \ge q'$, and so Lemma 3.1 implies that $\bar{G}^*(\phi^{q-1}\bar{g}(t)) \le c_G \phi^q \bar{G}(t)$. Then, using Young inequality (20) and by taking $\widetilde{\gamma} \in (0, 1/(2c_*c_G + 1)]$, get

$$
\begin{aligned}
\mathrm{III}_2 &\le \frac{c_*}{1-\widetilde{\gamma}} \int_{B_{3/4}} \bar{g}(|D\widetilde{\mathbf{u}}|)(1+|\widetilde{\mathbf{u}}|)^{1-\widetilde{\gamma}} \phi^{q-1} |D\phi| \, dx \\
&\le \frac{1}{2\widetilde{\gamma} c_G} \int_{B_{3/4}} \frac{\bar{G}^*\left(\phi^{q-1}\bar{g}(|D\widetilde{\mathbf{u}}|)\right)}{(1+|\widetilde{\mathbf{u}}|)^{\widetilde{\gamma}}} \, dx + \frac{1}{2\widetilde{\gamma}} \int_{B_{3/4}} \frac{\bar{G}\left(\frac{2\widetilde{\gamma} c_* c_G}{1-\widetilde{\gamma}}(1+|\widetilde{\mathbf{u}}|)|D\phi|\right)}{(1+|\widetilde{\mathbf{u}}|)^{\widetilde{\gamma}}} \, dx \\
&\le \frac{1}{2\widetilde{\gamma}} \int_{B_{3/4}} \frac{\bar{G}\left(|D\widetilde{\mathbf{u}}|\right)\phi^q}{(1+|\widetilde{\mathbf{u}}|)^{\widetilde{\gamma}}} \, dx + \frac{c_* c_G}{1-\widetilde{\gamma}} \int_{B_{3/4}} \frac{\bar{G}\left((1+|\widetilde{\mathbf{u}}|)|D\phi|\right)}{(1+|\widetilde{\mathbf{u}}|)^{\widetilde{\gamma}}} \, dx
\end{aligned}
$$

where in the last line we used that $\frac{2\widetilde{\gamma} c_* c_G}{1-\widetilde{\gamma}} < 1$ can be taken out of the integrand by Jensen's inequality. Summing up we obtain

$$
\int_{B_{3/4}} \frac{\bar{G}(|D\widetilde{\mathbf{u}}|)\phi^q}{(1+|\widetilde{\mathbf{u}}|)^{\widetilde{\gamma}}} \, dx \le C \int_{B_{3/4}} \frac{\bar{G}\left((1+|\widetilde{\mathbf{u}}|)|D\phi|\right)}{(1+|\widetilde{\mathbf{u}}|)^{\widetilde{\gamma}}} \, dx \tag{97}
$$

where $C = C(\mathtt{data}) > 0$. Then similar calculations to (55)-(59) yield that

$$
\left(\int_{B_{3/4}} \left(\frac{\bar{G}((1+|\bar{\mathbf{u}}|)\phi^q)}{(1+|\bar{\mathbf{u}}|)^{\widetilde{\gamma}}} \right)^{\frac{n}{n-1}} dx \right)^{\frac{n-1}{n}} \le c \int_{B_{3/4}} \frac{\bar{G}\left((1+|\bar{\mathbf{u}}|)|D\phi|\right)}{(1+|\bar{\mathbf{u}}|)^{\widetilde{\gamma}}} \, dx.
$$

holds with $c = c(\mathtt{data}) > 0$. Indeed, in **Step 4**, we have checked that the above $c = c(\mathtt{data}, \widetilde{\gamma})$ is an increasing function of $\widetilde{\gamma}$. As we consider small $\widetilde{\gamma}$, c in fact depends only on $\mathtt{data}$. For $5/8 \le r_1 < r_2 \le 3/4$ we take $\phi \in C_c^\infty(B_{r_2})$ to satisfy

$$
\phi \equiv 1 \quad \text{on} \quad B_{r_1} \qquad \text{and} \qquad |D\phi| \le \frac{100}{r_2 - r_1}.
$$

Then the doubling property of $\bar{G}$ and Lemma 3.9 gives

$$
\left(\int_{B_{5/8}} \left(\frac{\bar{G}(1+|\bar{\mathbf{u}}|)}{(1+|\bar{\mathbf{u}}|)^{\widetilde{\gamma}}} \right)^{\frac{n}{n-1}} dx \right)^{\frac{n-1}{n}} \le c \left(\int_{B_{3/4}} \left(\frac{\bar{G}(1+|\bar{\mathbf{u}}|)}{(1+|\bar{\mathbf{u}}|)^{\widetilde{\gamma}}} \right)^{\frac{1}{2s_G}} dx \right)^{2s_G}. \tag{98}
$$

We now restrict ourselves to $\widetilde{\gamma} \in \left(0, \min\{1/(2c_* c_G + 1), \frac{i_G}{2}\}\right)$ and define

$$
\Psi_{\widetilde{\gamma}}(t) = \int_0^t \frac{1}{\tau} \left(\frac{\bar{G}(\tau)}{\tau^{\widetilde{\gamma}}} \right)^{\frac{1}{2s_G}} d\tau,
$$

which is an increasing concave function on $[0, \infty)$ satisfying

$$
\frac{t\Psi''_{\widetilde{\gamma}}(t)}{\Psi'_{\widetilde{\gamma}}(t)} = \frac{tg(t)}{2s_G \bar{G}(t)} - 1 - \frac{\widetilde{\gamma}}{2s_G} \in (-1, -1/2) \quad \text{and} \quad \Psi_{\widetilde{\gamma}}(t) \approx \left(\frac{\bar{G}(t)}{t^{\widetilde{\gamma}}} \right)^{\frac{1}{2s_G}}.
$$

Then Jensen's inequality gives

$$
\begin{aligned}
\int_{B_{3/4}} \left(\frac{\bar{G}(1+|\bar{\mathbf{u}}|)}{(1+|\bar{\mathbf{u}}|)^{\widetilde{\gamma}}} \right)^{\frac{1}{2s_G}} dx &\le c \int_{B_{3/4}} \Psi_{\widetilde{\gamma}}(1+|\bar{\mathbf{u}}|) \, dx \le c\, \Psi_{\widetilde{\gamma}} \left(\int_{B_{3/4}} (1+|\bar{\mathbf{u}}|) \, dx \right) \\
&\le c\, \Psi_{\widetilde{\gamma}} \left(\int_{B_{3/4}} (1+|\bar{\mathbf{u}}|) \, dx + 1 \right) \le c\, \bar{G}^{\frac{1}{2s_G}} \left(\int_{B_{3/4}} (1+|\bar{\mathbf{u}}|) \, dx + 1 \right) \le c,
\end{aligned} \tag{99}
$$

with $c = c(\mathtt{data}) > 0$. We used that if $t > 1$ is arbitrarily fixed there exists $c > 0$ independent of $\widetilde{\gamma}$ such that for all $t > 1$ and all $\widetilde{\gamma}$, we have $\Psi_{\widetilde{\gamma}}(t) \le c\left(\bar{G}(t)\right)^{\frac{1}{2s_G}}$. By Hölder inequality, (98) and (99) we get that

$$\int_{B_{5/8}} \frac{\bar{G}(1+|\bar{\mathbf{u}}|)}{(1+|\bar{\mathbf{u}}|)^{\widetilde{\gamma}}}\,dx \le \left(\int_{B_{5/8}} \left(\frac{\bar{G}(1+|\bar{\mathbf{u}}|)}{(1+|\bar{\mathbf{u}}|)^{\widetilde{\gamma}}}\right)^{\frac{n}{n-1}} dx\right)^{\frac{n-1}{n}} \le c \tag{100}$$

with $c = c(\mathtt{data}) > 0$.

We now consider (97) with a cutoff function $\phi \in C_c^\infty(B_{5/8})$ satisfying

$$\phi \equiv 1 \quad \text{in} \quad B_{1/2} \qquad \text{and} \qquad |D\phi| \le 100$$

and combine it with (100), to obtain for some $c = c(\mathtt{data}) > 0$ that

$$\int_{B_{1/2}} \frac{\bar{G}(|D\widetilde{\mathbf{u}}|)}{(1+|\widetilde{\mathbf{u}}|)^{\widetilde{\gamma}}}\,dx \le c\int_{B_{3/4}} \frac{\bar{G}(1+|\widetilde{\mathbf{u}}|)}{(1+|\widetilde{\mathbf{u}}|)^{\widetilde{\gamma}}}\,dx \le c.$$

Therefore, using Fatou's lemma we justify that

$$\int_{B_{1/2}} \bar{G}(|D\widetilde{\mathbf{u}}|)\,dx \le \limsup_{\widetilde{\gamma}\to 0}\int_{B_{1/2}} \frac{\bar{G}(|D\widetilde{\mathbf{u}}|)}{(1+|\widetilde{\mathbf{u}}|)^{\widetilde{\gamma}}}\,dx \le c.$$

Consequently, we conclude that $|D\widetilde{\mathbf{u}}| \in L^G(B_{1/2})$. This completes the proof of Theorem 4.1. □

5 Proof of Wolff potential estimates

5.1 Comparison estimate

We need one more auxiliary estimate yielding comparison between energy of a weak solution and an $\mathcal{A}$-harmonic function.

Lemma 5.1 *Under Assumption* **(A-vect)** *suppose* $u \in W^{1,G}(B_r, \mathbb{R}^m)$ *is a weak solution to (4) in* $B_r = B_r(x_0)$, $r < 1$ *and let* $\varepsilon \in (0,1)$. *Then there exists a positive constant* $c_s = c_s(\mathtt{data}, \varepsilon)$ *and a map* $\mathbf{v}$ *being* $\mathcal{A}$*-harmonic in* $B_{r/2}$ *and such that*

$$\fint_{B_{r/2}} |D\mathbf{u} - D\mathbf{v}|\,dx \le \frac{\varepsilon}{r}\fint_{B_r} |\mathbf{u} - (\mathbf{u})_{B_r}|\,dx + c_s g^{-1}\left(\frac{|\boldsymbol{\mu}|(B_r)}{r^{n-1}}\right). \tag{101}$$

Proof Let us fix

$$\lambda := \frac{1}{r}\fint_{B_r} |\mathbf{u} - (\mathbf{u})_{B_r}|\,dx + g^{-1}\left(\delta\frac{|\boldsymbol{\mu}|(B_r)}{r^{n-1}}\right)$$

with $\delta = \delta(\mathtt{data}, \varepsilon)$ from Theorem 4.1 with $M = 1$. If $\lambda = 0$, then $\mathbf{u}$ is constant and $\mathbf{v} = \mathbf{u}$. Otherwise $\lambda > 0$ and we can argue by scaling

$$\bar{\mathbf{u}} := \frac{\mathbf{u} - (\mathbf{u})_{B_r}}{\lambda}, \qquad \bar{\boldsymbol{\mu}} := \frac{\boldsymbol{\mu}}{g(\lambda)}, \qquad \bar{\mathcal{A}}(x,\boldsymbol{\xi}) = \frac{\mathcal{A}(x,\lambda\boldsymbol{\xi})}{g(\lambda)}.$$

Then

$$\left|\fint_{B_r} \bar{\mathcal{A}}(x, D\bar{\mathbf{u}}) : D\boldsymbol{\varphi}\,dx\right| \le \frac{\|\boldsymbol{\varphi}\|_{L^\infty(B_r)}|\boldsymbol{\mu}|(B_r)}{g(\lambda)r^{n-1}} \le \frac{\delta}{r}\|\boldsymbol{\varphi}\|_{L^\infty(B_r)}.$$

By definition of $\bar{\mathbf{u}}$ and λ we notice that

$$\fint_{B_r} |\bar{\mathbf{u}}|\,dx \le r.$$

Therefore, by Theorem 4.1 applied to $\bar{\mathbf{u}}$ we get that there exists $\bar{\mathbf{v}}$ being $\bar{\mathcal{A}}$-harmonic in $B_{r/2}$ and such that

$$\fint_{B_{r/2}} |D\bar{\mathbf{u}} - D\bar{\mathbf{v}}|\,dx \le \varepsilon.$$

Then (101) follows by rescaling back with $\mathbf{v} = \lambda\bar{\mathbf{v}}$ which is $\mathcal{A}$-harmonic. □

5.2 Estimates on concentric balls

This subsection is devoted to prove some properties of weak solutions to (4) with $\boldsymbol{\mu} \in C^\infty(\Omega, \mathbb{R}^m)$ holding over a family of concentric balls $\{B^j\}$. Before we pass to this, let us fix some notation and parameters. Recall that we have chosen $R_0 = R_0(\texttt{data}, \varsigma)$ in Proposition 3.13. We take an arbitrary constant $\alpha_V \in (0, 1)$ and take $\varsigma = \varsigma(s_G, \alpha_V)$ to satisfy $\alpha_D := \frac{\alpha_V+1}{2} \le 1 + (\varsigma - 1)K$. To prove Theorem 2.1, it is enough to take $\alpha_V = \frac{1}{2}$, but for the later use in the proof of Theorem 2.9, we have taken α_V arbitrarily. We now choose

$$\sigma_0 := \min\left\{\left(\frac{1}{2^{n+6}c_o}\right)^{\frac{1-\alpha_V}{2}}, \frac{1}{4}\right\}. \tag{102}$$

If $r \in (0, R_0)$ is given, for every $j \in \mathbb{N} \cup \{0\}$ let us fix

$$r_j := \sigma^{j+1} r, \qquad B^j := \overline{B_{r_j}(x_0)},$$

so that $r_{-1} = r$. We denote

$$E_j := \fint_{B^j} |\mathbf{u} - (\mathbf{u})_{B^j}|\,dx. \tag{103}$$

Lemma 5.2 *Suppose Assumption* **(A-vect)** *is satisfied. If* $\mathbf{u} \in W^{1,G}(\Omega, \mathbb{R}^m)$ *is a weak solution to* (4) *with* $\boldsymbol{\mu} \in C^\infty(\Omega, \mathbb{R}^m)$*,* $j \in \mathbb{N}$ *is fixed,* E_j *is given by* (103)*, while* $0 < \sigma \le \sigma_0$ *is arbitrary, then we have that*

$$E_{j+1} \le c_D \sigma^{\alpha_D} E_j + c_E r_j g^{-1}\left(\frac{|\boldsymbol{\mu}|(B^j)}{r_j^{n-1}}\right) \tag{104}$$

for $c_D = c_D(\texttt{data}, \alpha_V) = 2^{n+4}c_o$ *and* $c_E = c_E(\texttt{data}, \alpha_V) = 2^{n+2}c_s c_P c_o \sigma^{-n}$*, where* c_P *is the constant from Poincaré inequality in* $W^{1,1}(\Omega, \mathbb{R}^m)$*.*

Proof We may apply Lemma 5.1 in $B_r = B_{r_j}(x_0)$ to get that there exists an $\mathcal{A}$-harmonic map $\mathbf{v}_j \in W^{1,G}(B_{r_{j/2}}, \mathbb{R}^m)$ in $B_{r_{j/2}}$ such that

$$\fint_{\frac{1}{2}B^j} |D\mathbf{u} - D\mathbf{v}_j|\,dx \le \frac{\varepsilon}{r_j} E_j + c_s g^{-1}\left(\frac{|\boldsymbol{\mu}|(B^j)}{r_j^{n-1}}\right). \tag{105}$$

By Poincaré inequality in $W^{1,1}(\Omega, \mathbb{R}^m)$ for $\mathbf{w}_j = \mathbf{u} - \mathbf{v}_j$ we have

$$\fint_{\frac{1}{2}B^j} |\mathbf{w}_j - (\mathbf{w}_j)_{\frac{1}{2}B^j}|\,dx \le c_P r_j \fint_{\frac{1}{2}B^j} |D\mathbf{u} - D\mathbf{v}_j|\,dx.$$

Then

$$\fint_{\frac{1}{2}B^j} |\mathbf{w}_j - (\mathbf{w}_j)_{\frac{1}{2}B^j}|\,dx \le \varepsilon c_{\mathrm{P}}\, E_j + c_{\mathrm{s}} c_{\mathrm{P}}\, r_j g^{-1}\left(\frac{|\boldsymbol{\mu}|(B_{r_j})}{r_j^{n-1}}\right). \tag{106}$$

Thus by Lemma 3.5, the triangle inequality, (106), and Proposition 3.13, we estimate

$$\begin{aligned}
E_{j+1} &= \fint_{B^{j+1}} |\mathbf{u} - (\mathbf{u})_{B^{j+1}}|\,dx \\
&\le \fint_{B^{j+1}} |\mathbf{w}_j - (\mathbf{w}_j)_{B^{j+1}}|\,dx + \fint_{B^{j+1}} |\mathbf{v}_j - (\mathbf{v}_j)_{B^{j+1}}|\,dx \\
&\le 2\fint_{B^{j+1}} |\mathbf{w}_j - (\mathbf{w}_j)_{\frac{1}{2}B^j}|\,dx + 2c_{\mathrm{o}}\sigma^{1+(\varsigma-1)/s_G}\fint_{\frac{1}{2}B^j} |\mathbf{v}_j - (\mathbf{v}_j)_{\frac{1}{2}B^j}|\,dx \\
&\le \left(\frac{2^{n+1}}{\sigma^n} + 2c_{\mathrm{o}}\sigma^{\alpha_V}\right)\fint_{\frac{1}{2}B^j} |\mathbf{w}_j - (\mathbf{w}_j)_{\frac{1}{2}B^j}|\,dx + 2c_{\mathrm{o}}\sigma^{\alpha_V}\fint_{\frac{1}{2}B^j} |\mathbf{u} - (\mathbf{u})_{\frac{1}{2}B^j}|\,dx \\
&\le 2^{n+2}c_{\mathrm{o}}\sigma^{\alpha_V} E_j + \left(\frac{2^{n+1}}{\sigma^n} + 2c_{\mathrm{o}}\sigma^{\alpha_V}\right)\fint_{\frac{1}{2}B^j} |\mathbf{w}_j - (\mathbf{w}_j)_{\frac{1}{2}B^j}|\,dx \\
&\le \left(2^{n+2}c_{\mathrm{o}}\sigma^{\alpha_V} + 2\varepsilon c_{\mathrm{o}} c_{\mathrm{P}}\sigma^{\alpha_V} + \varepsilon c_{\mathrm{P}}\frac{2^{n+1}}{\sigma^n}\right) E_j \\
&\quad + c_{\mathrm{s}} c_{\mathrm{P}}\left(\frac{2^{n+1}}{\sigma^n} + 2c_0\sigma^{\alpha_V}\right) r_j g^{-1}\left(\frac{|\boldsymbol{\mu}|(B_{r_j})}{r_j^{n-1}}\right).
\end{aligned}$$

By choosing $\varepsilon = \frac{\sigma^{n+\alpha_V}}{c_{\mathrm{P}}}$ we complete the proof. □

Proposition 5.3 *Suppose Assumption* **(A-vect)** *is satisfied. If* $\mathbf{u} \in W^{1,G}(\Omega, \mathbb{R}^m)$ *is a weak solution to* (4) *with* $\boldsymbol{\mu} \in C^\infty(\Omega, \mathbb{R}^m)$*, then there exists a constant* $c_V = c_V(\mathtt{data}, \alpha_V) \ge 1$ *such that for every* $\tau \in (0, 1]$ *we have*

$$\begin{aligned}
&\fint_{B_{\tau r}(x_0)} |\mathbf{u} - (\mathbf{u})_{B_{\tau r(x_0)}}|\,dx \\
&\qquad \le c_V \tau^{\alpha_V}\fint_{B_r(x_0)} |\mathbf{u} - (\mathbf{u})_{B_r(x_0)}|\,dx + c_V \sup_{0<\varrho<r} \varrho\, g^{-1}\left(\frac{|\boldsymbol{\mu}|(B_\varrho(x_0))}{\varrho^{n-1}}\right).
\end{aligned}$$

Proof Lemma 5.2 implies that for $j \in \mathbb{N} \cup \{0\}$ it holds that

$$E_{j+1} \le c_{\mathrm{D}}\sigma^{\alpha_D} E_j + c_{\mathrm{E}} r_j g^{-1}\left(\frac{|\boldsymbol{\mu}|(B^j)}{r_j^{n-1}}\right).$$

Iterating this estimate we get that for any $k \in \mathbb{N} \cup \{0\}$ we have

$$E_{k+1} \le (c_D\sigma^{\alpha_D})^{k+1} E_0 + c_{\mathrm{E}}\sum_{j=0}^{k} (c_{\mathrm{D}}\sigma^{\alpha_{\mathrm{D}}})^j r_j\, g^{-1}\left(\frac{|\boldsymbol{\mu}|(B^j)}{r_j^{n-1}}\right),$$

where $c = c(\mathtt{data})$. Recalling $\alpha_D = \frac{\alpha_V+1}{2}$, $c_{\mathrm{D}} = 2^{n+4}c_{\mathrm{o}}$ and (102), we see

$$c_{\mathrm{D}}\sigma^{\alpha_{\mathrm{D}}} \le \frac{\sigma^{\alpha_V}}{4}.$$

By Lemma 3.5 and direct computation we have for any $k \in \mathbb{N} \cup \{0\}$ that

$$E_k \le \sigma^{k\alpha_V} \fint_{B_r(x_0)} |\mathbf{u} - (\mathbf{u})_{B_r(x_0)}|\, dx + 2c_{\mathrm{E}} \sup_{0<\varrho<r} \varrho\, g^{-1}\left(\frac{|\boldsymbol{\mu}|(B_\varrho(x_0))}{\varrho^{n-1}}\right). \tag{107}$$

We take $\tau \in (0, \sigma)$ and $k \ge 1$ such that $\sigma^{k+1} < \tau \le \sigma^k$. Then by Lemma 3.5 and (107) we obtain

$$\begin{aligned}
&\fint_{B_{\tau r}(x_0)} |\mathbf{u} - (\mathbf{u})_{B_{\tau r}(x_0)}|\, dx \\
&\quad \le \frac{2\sigma^{kn}}{\tau^n} \fint_{B_{r\sigma^k}(x_0)} |\mathbf{u} - (\mathbf{u})_{B^{k-1}}|\, dx \le \frac{2E_{k-1}}{\sigma^n} \\
&\quad \le \frac{\sigma^{(k+1)\alpha_V}}{\sigma^{n+2\alpha_V}} \fint_{B_r(x_0)} |\mathbf{u} - (\mathbf{u})_{B_r(x_0)}|\, dx + \frac{2c_{\mathrm{E}}}{\sigma^n} \sup_{0<\varrho<r} \varrho\, g^{-1}\left(\frac{|\boldsymbol{\mu}|(B_\varrho(x_0))}{\varrho^{n-1}}\right) \\
&\quad \le \frac{\tau^{\alpha_V}}{\sigma^{n+2\alpha_V}} \fint_{B_r(x_0)} |\mathbf{u} - (\mathbf{u})_{B_r(x_0)}|\, dx + \frac{2c_{\mathrm{E}}}{\sigma^n} \sup_{0<\varrho<r} \varrho\, g^{-1}\left(\frac{|\boldsymbol{\mu}|(B_\varrho(x_0))}{\varrho^{n-1}}\right).
\end{aligned}$$

By taking $c_V = c_V(\mathtt{data}, \alpha_V) = 2c_{\mathrm{E}}\sigma^{-n-2\alpha_V}$ we conclude the claim for $\tau \in (0, \sigma)$. For completing the range of $\tau \in [\sigma, 1]$ it suffices to note that

$$\fint_{B_{\tau r}(x_0)} |\mathbf{u} - (\mathbf{u})_{B_{\tau r}(x_0)}|\, dx \le \frac{2}{\sigma^n} \fint_{B_r(x_0)} |\mathbf{u} - (\mathbf{u})_{B_r(x_0)}|\, dx.$$

□

5.3 SOLA u belongs to VMO

Proof of Proposition 2.2 Suppose $\boldsymbol{\mu} \in C^\infty(\Omega, \mathbb{R}^m)$ and $\mathbf{u} \in W^{1,G}(\Omega, \mathbb{R}^m)$ is a weak solution to (4). By Proposition 5.3 we can find constants $c_V = c_V(\mathtt{data}, \alpha_V)$ we have

$$\fint_{B_{\tau r}} |\mathbf{u} - (\mathbf{u})_{B_{\tau r}}|\, dx \le c_V \tau^{\alpha_V} \fint_{B_r} |\mathbf{u} - (\mathbf{u})_{B_r}|\, dx + c_V \sup_{0<\varrho<r} \varrho\, g^{-1}\left(\frac{|\boldsymbol{\mu}|(B_\varrho)}{\varrho^{n-1}}\right). \tag{108}$$

Let us consider a SOLA $\mathbf{u} \in W^{1,1}(\Omega, \mathbb{R}^m)$ existing due to Proposition 3.4. Suppose $(\mathbf{u}_h)$ and $(\boldsymbol{\mu}_h)$ are approximating sequences from definition of SOLA, see Sect. 2.2. Inequality (108) hold for each $\mathbf{u}_h$ and $\boldsymbol{\mu}_h$. We have to motivate passing to the limit with $h \to \infty$. Since (7) holds, we can write (108) for the original SOLA too. From now on this kind of solution is considered.

Our aim now is to show that $\mathbf{u}$ is VMO at x_0 provided (10) is assumed. Let $\delta \in (0, 1)$. By (10) we find a positive radius $r_{1,\delta} < r$ such that

$$c_V \sup_{0<\varrho<r_{1,\delta}} \varrho\, g^{-1}\left(\frac{|\boldsymbol{\mu}|(B_\varrho(x_0))}{\varrho^{n-1}}\right) \le \frac{\delta}{2}$$

and then τ_δ so small that

$$c_V \tau_\delta^{\alpha_V} \fint_{B_{r_{1,\delta}}} |\mathbf{u} - (\mathbf{u})_{B_{r_{1,\delta}}(x_0)}|\, dx \le \frac{\delta}{2}.$$

For $r_\delta := \tau_\delta r_{1,\delta}$ from estimate (108) (applied with $r = r_{1,\delta}$) it follows that

$$\sup_{0<\varrho<r_\delta} \fint_{B_\varrho(x_0)} |\mathbf{u} - (\mathbf{u})_{B_{\varrho(x_0)}}| dx \le \delta,$$

that is that $\mathbf{u}$ has vanishing mean oscillation at x_0. □

5.4 Proofs of Theorems 2.1, 2.3 and 2.9

We start with the proof of pointwise Wolff potential estimate, then pass to continuity and Hölder continuity criteria.

Proof of Theorem 2.1 We notice that having E_j defined in (103) with $r = r^j$ we can fix σ in Lemma 5.2 to get that

$$E_{j+1} \le \frac{1}{2} E_j + c r^j g^{-1}\left(\frac{|\boldsymbol{\mu}|(B^j)}{r_j^{n-1}}\right) \qquad \text{for every } \; j \in \mathbb{N} \cup \{0\}. \tag{109}$$

We sum up inequalities from (109) to obtain

$$\sum_{j=1}^{k+1} E_j \le \frac{1}{2} \sum_{j=0}^{k} E_j + c \sum_{j=0}^{k} r_j g^{-1}\left(\frac{|\boldsymbol{\mu}|(B^j)}{r_j^{n-1}}\right), \qquad k \in \mathbb{N} \cup \{0\}.$$

By rearranging terms we have

$$\sum_{j=1}^{k+1} E_j \le 2E_0 + c \sum_{j=0}^{k} r_j g^{-1}\left(\frac{|\boldsymbol{\mu}|(B^j)}{r_j^{n-1}}\right).$$

We notice that for some $c = c(\mathtt{data})$ we can estimate

$$\sum_{j=0}^{k} r_j g^{-1}\left(\frac{|\boldsymbol{\mu}|(B^j)}{r_j^{n-1}}\right) \le c \int_0^r g^{-1}\left(\frac{|\boldsymbol{\mu}|(B_\varrho)}{\varrho^{n-1}}\right) d\varrho = c\, \mathcal{W}_G^{\boldsymbol{\mu}}(x_0, r).$$

Last two displays imply that

$$\sum_{j=1}^{k+1} E_j \le 2E_0 + c\mathcal{W}_G^{\boldsymbol{\mu}}(x_0, r).$$

For every $m, k \in \mathbb{N}$ such that $m < k$ we have

$$\begin{aligned}
|(\mathbf{u})_{B^k} - (\mathbf{u})_{B^m}| &\le \sum_{j=m}^{k-1} |(\mathbf{u})_{B^{j+1}} - (\mathbf{u})_{B^j}| \le \sigma^{-n} \sum_{j=m}^{k+1} E_j \le \sigma^{-n} \sum_{j=0}^{k+1} E_j \\
&\le 2\sigma^{-n} E_0 + c\sigma^{-n} \mathcal{W}_G^{\boldsymbol{\mu}}(x_0, r) \\
&\le 2\sigma^{-n} \fint_{B_r(x_0)} |\mathbf{u} - (\mathbf{u})_{B_r(x_0)}|\, dx + c\sigma^{-n} \mathcal{W}_G^{\boldsymbol{\mu}}(x_0, r),
\end{aligned}$$

where $\sigma = \sigma(\mathtt{data})$ and $c = c(\mathtt{data})$. For $j \to \infty$, $((\mathbf{u})_{B^j})_j$ is a Cauchy sequence that converges to $\mathbf{u}(x_0)$, that is

$$\lim_{\varrho \to 0} (\mathbf{u})_{B_\varrho(x_0)} = \mathbf{u}(x_0)$$

and x_0 is a Lebesgue's point of $\mathbf{u}$. This completes the proof of (8), while (9) follows as a direct corollary. □

Let us concentrate on the continuity criterion.

Proof of Theorem 2.3 Our aim is to infer continuity of $\mathbf{u}$ in $B_r(x_0)$ knowing that (12) holds. We will show that for every $\delta > 0$ and $x_1 \in B_r(x_0)$ we can find $r_\delta \in (0, \operatorname{dist}(B_r(x_0), \partial\Omega))$ such that

$$\operatorname{osc}_{B_{r_\delta}(x_1)} \mathbf{u} < \delta. \tag{110}$$

Without loss of generality we assume that $\boldsymbol{\mu}$ is defined on whole $\mathbb{R}^n$, as we can extend it by zero outside Ω. By (12) we can take ϱ_1 small enough for

$$\sup_{x \in B_r(x_0)} \mathcal{W}_G^{\boldsymbol{\mu}}(x, \varrho_1) \leq \frac{\delta}{16}. \tag{111}$$

Let $r_\delta > 0$ to be chosen in a moment. We take an arbitrary point $x_2 \in B_{r_\delta}(x_1)$ and estimate

$$\begin{aligned} |\mathbf{u}(x_1) - \mathbf{u}(x_2)| &\leq |\mathbf{u}(x_1) - (\mathbf{u})_{B_{2r_\delta}(x_1)}| + |(\mathbf{u})_{B_{2r_\delta}(x_1)} - (\mathbf{u})_{B_{r_\delta}(x_2)}| \\ &\quad + |(\mathbf{u})_{B_{r_\delta}(x_2)} - \mathbf{u}(x_2)| =: A_1 + A_2 + A_3. \end{aligned} \tag{112}$$

We start with estimating A_2 by noting that

$$A_2 = |(\mathbf{u})_{B_{r_\delta}(x_2)} - (\mathbf{u})_{B_{2r_\delta}(x_1)}| \leq \fint_{B_{r_\delta}(x_2)} |\mathbf{u} - (\mathbf{u})_{B_{2r_\delta}(x_1)}|\, dx.$$

Since (12) implies (10), Proposition 2.2 implies that $\mathbf{u}$ has vanishing mean oscillations at x_1. Therefore there exists $\varrho_2 \in (0, \min\{\varrho_1, \operatorname{dist}(x_1, \partial B_r(x_0))/4\})$ such that for every $\varrho \leq \varrho_2$ it holds

$$\fint_{B_\varrho(x_1)} |\mathbf{u} - (\mathbf{u})_{B_\varrho(x_1)}|\, dx \leq \frac{\delta}{2^{n+4}}.$$

We choose $r_\delta = \varrho_2/2$ and observe that (111) imply that

$$\fint_{B_{r_\delta}(x_2)} |\mathbf{u} - (\mathbf{u})_{B_{2r_\delta}(x_1)}|\, dx \leq 2^n \fint_{B_{2r_\delta}(x_1)} |\mathbf{u} - (\mathbf{u})_{B_{2r_\delta}(x_1)}|\, dx \leq \frac{\delta}{16}.$$

In turn $A_2 \leq \frac{\delta}{16}$. By Theorem 2.1 and (111) we get that x_1 and x_2 are Lebesgue's points and

$$A_1 + A_3 = |\mathbf{u}(x_1) - (\mathbf{u})_{B_{2r_\delta}(x_1)}| + |(\mathbf{u})_{B_{r_\delta}(x_2)} - \mathbf{u}(x_2)| \leq \frac{\delta}{4}.$$

Applying these observation we get from (112) that

$$|\mathbf{u}(x_1) - \mathbf{u}(x_2)| \leq \frac{\delta}{2}.$$

Since x_2 was an arbitrary point of $B_{r_\delta}(x_1)$, we have (110) justified, which completes the proof. □

We are in the position to prove the Hölder continuity criterion.

Proof of Theorem 2.9 Notice that assumption (15) implies that there exists $c = c(\mathtt{data}) > 0$, such that for all sufficiently small r we have

$$\mathcal{W}_G^{\boldsymbol{\mu}}(x, r) \leq c r^\theta.$$

Applying assumption (15) to Proposition 5.3 with $\alpha_V = \frac{\theta+1}{2}$, we have

$$\fint_{B_\rho} |\mathbf{u} - (\mathbf{u})_{B_\rho}|\, dx \leq c \left(\frac{\rho}{r}\right)^{\frac{\theta+1}{2}} \fint_{B_r} |\mathbf{u} - (\mathbf{u})_{B_r}|\, dx + c r^\theta$$

for any $0 < \rho < r \le R_0$. Now we apply Lemma 3.8 to see

$$\fint_{B_\rho} |\mathbf{u} - (\mathbf{u})_{B_\rho}|\, dx \le c \left(\frac{\rho}{r}\right)^{\theta} \fint_{B_r} |\mathbf{u} - (\mathbf{u})_{B_r}|\, dx + c\rho^{\theta}$$

By Campanato's characterization [48, Theorem 2.9], we complete the proof. □

Acknowledgements I. Chlebicka is supported by NCN Grant No. 2019/34/E/ST1/00120. A. Zatorska-Goldstein is supported by NCN Grant No. 2019/33/B/ST1/00535. Y. Youn is supported by NRF Grant No. 2020R1C1C1A01009760.

Appendix

Proof of Lemma 2.5 Set $x \in \Omega_0$, $R_k = 2^{1-k}R$ and $B_k = B_{R_k}(x)$ for $k = 0, 1, \dots$. As $\mathbf{F}$ is taken in a place of a measure with a slight abuse of notation we write $|\mathbf{F}|(B_{R_k}(x)) = \int_{B_k} |\mathbf{F}(y)|\, dy$. We notice that we have

$$\mathcal{W}_G^{|\mathbf{F}|}(x, R) = \sum_{k=1}^{\infty} \int_{R_{k+1}}^{R_k} g^{-1}\left(\frac{|\mathbf{F}|(B_r(x))}{r^{n-1}}\right) dr \lesssim \sum_{k=1}^{\infty} R_k\, g^{-1}\left(\frac{|\mathbf{F}|(B_{R_k}(x))}{R_k^{n-1}}\right).$$

To estimate the series we employ the decreasing rearrangement $|\mathbf{F}|^\star$ of $|\mathbf{F}|$ and its maximal rearrangement $|\mathbf{F}|^{\star\star}$. When w_n is the volume of the unit ball, we have that

$$\begin{aligned}
\frac{|\mathbf{F}|(B_{R_k}(x))}{R_k^{n-1}} &= \frac{1}{R_k^{n-1}} \int_{B_{R_k}(x)} |\mathbf{F}(y)|\, dy \\
&\le w_n R_k \fint_0^{w_n R_k^n} |\mathbf{F}|^\star(t)\, dt = w_n R_k\, |\mathbf{F}|^{\star\star}(w_n R_k^n).
\end{aligned}$$

Then we have

$$\begin{aligned}
R_k\, g^{-1}\left(\frac{|\mathbf{F}|(B_{R_k}(x))}{R_k^{n-1}}\right) &\lesssim R_k\, g^{-1}\left(w_n R_k\, |\mathbf{F}|^{\star\star}(w_n R_k^n)\right) \\
&\lesssim \int_{w_n R_k^n}^{w_n R_{k-1}^n} \rho^{\frac{1}{n}} g^{-1}\left(\rho^{\frac{1}{n}}\, |\mathbf{F}|^{\star\star}(\rho)\right) \frac{d\rho}{\rho}
\end{aligned} \tag{113}$$

with implicit constants independent of k. Therefore

$$\begin{aligned}
\sup_{x \in \Omega_0} \mathcal{W}_G^{|\mathbf{F}|}(x, R) &\lesssim \sum_{k=1}^{\infty} \int_{w_n R_k^n}^{w_n R_{k-1}^n} \rho^{\frac{1}{n}-1} g^{-1}\left(\rho^{\frac{1}{n}}\, |\mathbf{F}|^{\star\star}(\rho)\right) \frac{d\rho}{\rho} \\
&= \int_0^{w_n R^n} \rho^{\frac{1}{n}} g^{-1}\left(\rho^{\frac{1}{n}}\, |\mathbf{F}|^{\star\star}(\rho)\right) \frac{d\rho}{\rho}.
\end{aligned}$$

□

References

1. Adams, D.R., Hedberg, L.I.: Function spaces and potential theory, volume 314 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin (1996)
2. Adams, R.A., Fournier, J.J.F.: Sobolev Spaces, Volume 140 of Pure and Applied Mathematics, 2nd edn. Elsevier/Academic Press, Amsterdam (2003)
3. Alberico, A., Chlebicka, I., Cianchi, A., Zatorska-Goldstein, A.: Fully anisotropic elliptic problems with minimally integrable data. Calc. Var. PDE **58**, 186 (2019)
4. Balci, AKh., Cianchi, A., Diening, L., Maz'ya, V.: A pointwise differential inequality and second-order regularity for nonlinear elliptic systems. Math. Ann. **383**(3–4), 1775–1824 (2022)
5. Baroni, P.: Riesz potential estimates for a general class of quasilinear equations. Calc. Var. Partial Differ. Equ. **53**(3–4), 803–846 (2015)
6. Baroni, P., Habermann, J.: Elliptic interpolation estimates for non-standard growth operators. Ann. Acad. Sci. Fenn. Math. **39**(1), 119–162 (2014)
7. Beck, L., Mingione, G.: Lipschitz bounds and non-uniform ellipticity. Comm. Pure Appl. Math. **73**(5), 944–1034 (2020)
8. Bénilan, P., Boccardo, L., Gallouët, T., Gariepy, R., Pierre, M., Vázquez, J.-L.: An L^1-theory of existence and uniqueness of solutions of nonlinear elliptic equations. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **22**(2), 241–273 (1995)
9. Boccardo, L., Gallouët, T.: Nonlinear elliptic and parabolic equations involving measure data. J. Funct. Anal. **87**(1), 149–169 (1989)
10. Breit, D., Cianchi, A., Diening, L., Kuusi, T., Schwarzacher, S.: Pointwise Calderón–Zygmund gradient estimates for the p-Laplace system. J. Math. Pures Appl. **114**, 146–190 (2018)
11. Byun, S.-S., Youn, Y.: Optimal gradient estimates via Riesz potentials for $p(\cdot)$-Laplacian type equations. Q. J. Math. **68**(4), 1071–1115 (2017)
12. Byun, S.-S., Youn, Y.: Potential estimates for elliptic systems with subquadratic growth. J. Math. Pures Appl. **131**(9), 193–224 (2019)
13. Byun, S.-S., Youn, Y.: Riesz potential estimates for a class of double phase problems. J. Differ. Equ. **264**(2), 1263–1316 (2018)
14. Carleson, L.: Selected problems on exceptional sets. Van Nostrand Mathematical Studies, No. 13 D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London, v+151 pp (1967)
15. Celada, P., Ok, J.: Partial regularity for non-autonomous degenerate quasi-convex functionals with general growth. Nonlinear Anal. **194**, 111473 (2020)
16. Chlebicka, I.: A pocket guide to nonlinear differential equations in Musielak–Orlicz spaces. Nonlinear Anal. **175**, 1–27 (2018)
17. Chlebicka, I.: Gradient estimates for problems with Orlicz. Nonlinear Anal. **194**, 111364 (2020)
18. Chlebicka, I.: Measure data elliptic problems with generalized Orlicz growth. In: Proceedings of the Royal Society of Edinburgh Section A: Mathematics, to appear
19. Chlebicka, I.: Regularizing effect of the lower-order terms in elliptic problems with Orlicz growth. Israel J. Math. Anal. **236**(2), 967–1000 (2020)
20. Chlebicka, I., Karppinen, A.: Removable sets in elliptic equations with Musielak–Orlicz growth. J. Math. Anal. Appl. **501**(1), 124073 (2020)
21. Chlebicka, I., Giannetti, F., Zatorska-Goldstein, A.: Elliptic problems in the Orlicz setting without growth restrictions with measure or L^1 data. J. Math. Anal. Appl. **479**(1), 185–213 (2019)
22. Chlebicka, I., Giannetti, F., Zatorska-Goldstein, A.: Wolff potentials and local behaviour of solutions to measure data elliptic problems with Orlicz growth arXiv:2006.02172 (2020)
23. Chlebicka, I., Youn, Y., Zatorska-Goldstein, A.: Existence for measure data elliptic systems with Orlicz growth, arXiv:2106.11639 (2021)
24. Chlebicka, I., Zatorska-Goldstein, A.: Generalized superharmonic functions with strongly nonlinear operator. Potential Anal. **57**(3), 379–400 (2022)
25. Cianchi, A., Maz'ya, V.: Global boundedness of the gradient for a class of nonlinear elliptic systems. Arch. Ration. Mech. Anal. **212**(1), 129–177 (2014)
26. Cianchi, A., Maz'ya, V.: Quasilinear elliptic problems with general growth and merely integrable, or measure, data. Nonlinear Anal. **164**, 189–215 (2017)
27. Cianchi, A., Schwarzacher, S.: Potential estimates for the p-Laplace system with data in divergence form. J. Differ. Equ. **265**(1), 478–499 (2018)
28. Cupini, G., Leonetti, F., Mascolo, E.: Existence of weak solutions for elliptic systems with p, q-growth. Ann. Acad. Sci. Fenn. Math. **40**(2), 645–658 (2015)

29. De Giorgi, E.: Un esempio di estremali discontinue per un problema variazionale di tipo ellittico. Boll. Unione Mat. Ital. IV **1**, 135–137 (1968)
30. De Filippis, C.: On the regularity of the ω-minima of ϕ-functionals. Nonlinear Anal. **194**, 111464 (2020)
31. De Filippis, C., Leonetti, F.: Uniform ellipticity and (p, q)-growth. J. Math. Anal. Appl. **501**(1), 124451 (2021)
32. De Filippis, C., Mingione, G.: Lipschitz bounds and nonautonomous integrals. Arch. Ration. Mech. Anal. **242**, 973–1057 (2021)
33. Diening, L., Ettwein, F.: Fractional estimates for non-differentiable elliptic systems with general growth. Forum Math. **20**(3), 523–556 (2008)
34. Diening, L., Lengeler, D., Stroffolini, B., Verde, A.: Partial regularity for minimizers of quasi-convex functionals with general growth. SIAM J. Math. Anal. **44**(5), 3594–3616 (2012)
35. Diening, L., Stroffolini, B., Verde, A.: Everywhere regularity of functionals with ϕ-growth. Manuscr. Math. **129**(4), 449–481 (2009)
36. Diening, L., Stroffolini, B., Verde, A.: Lipschitz regularity for some asymptotically convex problems. ESAIM Control Optim. Calc. Var. **17**(1), 178–189 (2011)
37. Diening, L., Stroffolini, B., Verde, A.: The φ-harmonic approximation and the regularity of φ-harmonic maps. J. Differ. Equ. **253**(7), 1943–1958 (2012)
38. Di Marco, T., Marcellini, P.: A priori gradient bound for elliptic systems under either slow or fast growth conditions. Calc. Var. Partial Differ. Equ. **59**(4), 120 (2020)
39. Dolzmann, G., Hungerbühler, N., Müller, S.: The p-harmonic system with measure-valued right hand side. Ann. Inst. H. Poincaré Anal. Non Linéaire **14**, 353–364 (1997)
40. Dong, G., Shi, Z.: An existence theorem for weak solutions for a class of elliptic partial differential systems in Orlicz spaces. Nonlinear Anal. **68**(4), 1037–1042 (2008)
41. Duzaar, F., Grotowski, J.F.: Optimal interior partial regularity for nonlinear elliptic systems: the method of A-harmonic approximation. Manuscr. Math. **103**, 267–298 (2000)
42. Duzaar, F., Mingione, G.: Gradient continuity estimates. Calc. Var. Partial Differ. Equ. **39**(3–4), 379–418 (2010)
43. Duzaar, F., Mingione, G.: Gradient estimates via linear and nonlinear potentials. J. Funct. Anal. **259**(11), 2961–2998 (2010)
44. Duzaar, F., Mingione, G.: Gradient estimates via non-linear potentials. Am. J. Math. **133**(4), 1093–1149 (2011)
45. Duzaar, F., Mingione, G.: The p-harmonic approximation and the regularity of p-harmonic maps. Calc. Var. Partial Differ. Equ. **20**, 235–256 (2004)
46. Fuchs, M., Reuling, J.: Nonlinear elliptic systems involving measure data. Rend. Mat. **7**(15), 311–319 (1995)
47. Giannetti, F., Passarelli di Napoli, A., Tachikawa, A.: Partial regularity results for non-autonomous functionals with Φ-growth conditions. Ann. Mat. Pura Appl. (4) **196**(6), 2147–2165 (2017)
48. Giusti, E.: Direct Methods in the Calculus of Variations. World Scientific Publishing Co., Inc., River Edge (2003)
49. Gossez, J.-P.: Some approximation properties in Orlicz–Sobolev spaces. Stud. Math. **74**(1), 17–24 (1982)
50. Han, Q., Lin, F.: Elliptic partial differential equations, 2nd edn. In: Courant Lecture Notes in Mathematics. American Mathematical Society, Providence, RI (2011)
51. Hara, T.: The Wolff potential estimate for solutions to elliptic equations with signed data. Manuscr. Math. **150**, 45–58 (2016)
52. Havin, M., Maz'ya, V.G.: Nonlinear potential theory. Russ. Math. Surv. **27**, 71–148 (1972)
53. Heinonen, J., Kilpeläinen, T., Martio, O.: Nonlinear Potential Theory of Degenerate Elliptic Equations. Unabridged Republication of the 1993 Original. Dover Publications Inc., Mineola (2006)
54. Hedberg, L.I., Wolff, Th.H.: Thin sets in nonlinear potential theory. Ann. Inst. Fourier (Grenoble) **33**(4), 161–187 (1983)
55. Jin, T., Maz'ya, V., Van Schaftingen, J.: Pathological solutions to elliptic problems in divergence form with continuous coefficients. C. R. Math. Acad. Sci. Paris **347**(13–14), 773–778 (2009)
56. Kinderlehrer, D., Stampacchia, G.: An Introduction to Variational Inequalities and Their Applications, Pure Appl. Math., vol. 88. Academic Press, New York, London (1980)
57. Kilpeläinen, T., Malý, J.: Degenerate elliptic equations with measure data and nonlinear potentials. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **19**(4), 591–613 (1992)
58. Kilpeläinen, T., Malý, J.: The Wiener test and potential estimates for quasilinear elliptic equations. Acta Math. **172**(1), 137–161 (1994)
59. Kilpeläinen, T., Zhong, X.: Removable set for continuous solutions of quasilinear elliptic equations. Proc. Am. Math. Soc. **130**(6), 1681–1688 (2000)

60. Korte, R., Kuusi, T.: A note on the Wolff potential estimate for solutions to elliptic equations involving measures. Adv. Calc. Var. **3**, 99–113 (2010)
61. Krasnosel'skiĭ, M.A., Rutickiĭ, Y.B.: Convex Functions and Orlicz Spaces. P. Noordhoff Ltd., Groningen (1961)
62. Kuusi, T., Mingione, G.: Guide to nonlinear potential estimates. Bull. Math. Sci. **4**(1), 1–82 (2014)
63. Kuusi, T., Mingione, G.: Nonlinear potential theory of elliptic systems. Nonlinear Anal. **138**, 277–299 (2016)
64. Kuusi, T., Mingione, G.: Partial regularity and potentials. J. Éc. Polytech. Math. **3**, 309–363 (2016)
65. Kuusi, T., Mingione, G.: Vectorial nonlinear potential theory. J. Eur. Math. Soc. (JEMS) **20**(4), 929–1004 (2018)
66. Leonetti, F., Petricca, P.V.: Anisotropic elliptic systems with measure data. Ricerche Mat. **54**(2), 591–595 (2005)
67. Lieberman, G.M.: The natural generalization of the natural conditions of Ladyzhenskaya and Ural'tseva for elliptic equations. Comm. Partial Differ. Equ. **16**(2–3), 311–361 (1991)
68. Lukkari, T., Maeda, F.-Y., Marola, N.: Wolff potential estimates for elliptic equations with nonstandard growth and applications. Forum Math. **22**(6), 1061–1087 (2010)
69. Lindqvist, P., Martio, O.: Two theorems of N. Wiener for solutions of quasilinear elliptic equations. Acta Math. **155**(3–4), 153–171 (1985)
70. Malý, J.: Wolff potential estimates of superminimizers of Orlicz type Dirichlet integrals. Manuscr. Math. **110**(4), 513–525 (2003)
71. Marcellini, P.: Anisotropic and p, q-nonlinear partial differential equations. Rend. Fis. Acc. Lincei **31**, 295–301 (2020)
72. Marcellini, P.: Everywhere regularity for a class of elliptic systems without growth conditions. Ann. Scuola Norm. Pisa **23**, 1–25 (1996)
73. Marcellini, P., Papi, G.: Nonlinear elliptic systems with general growth. J. Differ. Equ. **221**(2), 412–443 (2006)
74. Maz'ya, V.G.: The continuity at a boundary point of the solutions of quasi-linear elliptic equations (Russian) Vestnik Leningrad. Univ **25**, 42–55 (1970)
75. Mingione, G.: Regularity of minima: an invitation to the dark side of the calculus of variations. Appl. Math. **51**(4), 355–426 (2006)
76. Nguyen, Q.-H., Phuc, N.C.: Pointwise gradient estimates for a class of singular quasilinear equations with measure data. J. Funct. Anal. **278**(5), 108391 (2020)
77. O'Neil, R.: Fractional integration in Orlicz spaces. I. Trans. Am. Math. Soc. **115**, 300–328 (1965)
78. Rakotoson, J.-M., Ziemer, W.: Local behavior of solutions of quasilinear elliptic equations with general structure. Trans. Am. Math. Soc. **319**(2), 747–764 (1990)
79. Rao, M.M., Ren, Z.D.: Theory of Orlicz Spaces, Volume 146 of Monographs and Textbooks in Pure and Applied Mathematics. Marcel Dekker Inc., New York (1991)
80. Stein, E.M., Weiss, G.: Introduction to Fourier Analysis on Euclidean Spaces, vol. 32. Princeton Univ. Press, Princeton (1971)
81. Stroffolini, B.: Partial regularity results for quasimonotone elliptic systems with general growth. Z. Anal. Anwend. **39**(3), 315–347 (2020)
82. Sverák, V., Yan, X.: Non-Lipschitz minimizers of smooth uniformly convex functionals. Proc. Natl. Acad. Sci. U.S.A. **99**(24), 15269–15276 (2002)
83. Trudinger, N.S., Wang, X.-J.: On the weak continuity of elliptic operators and applications to potential theory. Am. J. Math. **124**, 369–410 (2002)
84. Uhlenbeck, K.: Regularity for a class of nonlinear elliptic systems. Acta Math. **138**, 219–240 (1977)
85. Ural'tseva, N.N.: Degenerate quasilinear elliptic systems. Zap. Na. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **7**, 184–222 (1968)