

Manifold constrained non-uniformly elliptic problems

CRISTIANA DE FILIPPIS* AND GIUSEPPE MINGIONE[†]

September 21, 2019

Abstract

We consider the problem of minimizing variational integrals defined on nonlinear Sobolev spaces of competitors taking values into the sphere. The main novelty is that the underlying energy features a non-uniformly elliptic integrand exhibiting different polynomial growth conditions and no homogeneity. We develop a few intrinsic methods aimed at proving partial regularity of minima and providing techniques for treating larger classes of similar constrained non-uniformly elliptic variational problems. In order to give estimates for the singular sets we use a general family of Hausdorff type measures following the local geometry of the integrand. A suitable comparison is provided with respect to the naturally associated capacities.

Contents

1	Introduction	2
1.1	Partial regularity	4
1.2	Weighted Hausdorff measures, intrinsic capacities and singular sets	6
1.3	Overview of the paper	9
2	Notation	10
3	Basic material	11
3.1	Caccioppoli's and higher integrability inequalities	11
3.2	On the Euler-Lagrange equation under non-standard growth conditions	15
3.3	A Morrey type decay estimate	17
4	Harmonic type approximation	19
5	Proof of Theorem 1	27
6	Weighted Hausdorff measures and singular sets	43
6.1	Proof of Proposition 1	43
6.2	Proof of Theorem 2	44
6.3	Proof of Theorem 3	45
	References	46

*Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX26GG, Oxford, United Kingdom. E-mail: Cristiana.DeFilippis@maths.ox.ac.uk

[†]Dipartimento SMFI, Università di Parma, Viale delle Scienze 53/a, Campus, 43124 Parma, Italy E-mail: giuseppe.mingione@unipr.it

1 Introduction

In this paper we want to treat, from the regularity theory viewpoint, a special but yet significant class of non-uniformly variational problems characterized by the fact that minimizers and competitors take their values into the sphere. At the same time, we want to introduce a few intrinsic methods and viewpoints that should be useful in order to prove regularity theorems for more general classes of non-uniformly elliptic equations and functionals with geometric constraints. For this reason we shall combine and present both old techniques from different perspectives and new ones. Specifically, we shall consider a class of variational integrals of the type

$$W^{1,1}(\Omega, \mathbb{R}^N) \ni w \mapsto \mathcal{F}(w, \Omega) := \int_{\Omega} F(x, w, Dw) dx, \quad (1.1)$$

where the main model is provided by the so-called double phase functional

$$\left\{ \begin{array}{l} W^{1,1}(\Omega, \mathbb{R}^N) \ni w \mapsto \mathcal{P}(w, \Omega) := \int_{\Omega} (|Dw|^p + a(x)|Dw|^q) dx \\ 1 < p < q < N, \quad 0 \leq a(\cdot) \in C^{0,\alpha}(\Omega). \end{array} \right. \quad (1.2)$$

Here, as in the rest of the paper, $\Omega \subset \mathbb{R}^n$ denotes (unless otherwise specified) a bounded open domain, $n \geq 2$, and, again unless otherwise stated, we consider $N > 1$ and $a(\cdot)$ satisfies the condition in (1.2)₃ for some $\alpha \in (0, 1]$. When $a(\cdot) \equiv 0$, the integral in (1.2) reduces to the familiar p -Dirichlet integral

$$w \mapsto \int_{\Omega} |Dw|^p dx, \quad (1.3)$$

whose Euler-Lagrange equation is given by the p -Laplacean system $-\operatorname{div}(|Du|^{p-2}Du) = 0$. The regularity theory for minimizers of the functional in (1.3) has been treated at length starting from the seminal papers of Uraltseva [63] and Uhlenbeck [62], in the scalar and vectorial case, respectively. For results concerning more general functionals as in (1.1) with p -growth, that is, modelled on the one in (1.2) and therefore satisfying

$$|z|^p \lesssim F(x, w, z) \lesssim |z|^p + 1, \quad (1.4)$$

see for instance [41, 42, 45] and related references. Under suitable assumptions, the final outcome is that minima are locally of class $C^{1,\beta}$, for some $\beta \in (0, 1)$, on a subset of full n -dimensional Lebesgue measure. The regularity theory in the case when both minimizers and competitors take values into a manifold $\mathcal{M} \subset \mathbb{R}^N$ poses additional difficulties. In particular, the case $\mathcal{M} = \mathbb{S}^{N-1}$ is the $(N-1)$ -dimensional sphere in $\mathbb{R}^N$ has been treated extensively. The theory started with the fundamental papers of Eells & Sampson [24] and Schoen & Uhlenbeck [58, 59], analyzing harmonic maps, i.e., constrained minimizers of the functional in (1.3) for $p = 2$. The extension of such basic results to the case $p \neq 2$ has been done in the by now classical papers of Fuchs [29, 30], Hardt & Lin [32] and Luckhaus [44]. Moreover, several results have been extended to more general functionals with p -growth, that is, functionals as in (1.1) with $F(\cdot)$ satisfying (1.4); see for instance [33]. On the other hand, we notice that energies of the type in (1.2) do not satisfy conditions as in (1.4), but rather, the more general and flexible ones

$$|z|^p \lesssim F(x, w, z) \lesssim |z|^q + 1, \quad 1 < p < q. \quad (1.5)$$

These are known in the literature as (p, q) -growth conditions or non-standard growth conditions. They have pioneered by Uraltseva & Urdaletova [64] and Zhikov [65, 66, 67] in the context of Homogenization (see also the recent paper [22]). In the setting of the Calculus of Variations they have been systematically studied by Marcellini [46, 47]. We refer to [50] for a reasonable survey on the subject. Growth conditions of the

type in (1.5) often occur when considering variational models for physical phenomena. For instance, in the setting of Homogenization, a model as the double phase functional can be used to describe a composite of two materials with hardening exponents p and q respectively, whose geometry is dictated by the zero set $\{a(x) = 0\}$ of the coefficient $a(\cdot)$. Obviously, both in the case $a(\cdot) \equiv 0$ and in the one when $\inf a(\cdot) > 0$, we have a functional with standard polynomial growth of the type in (1.4) (with p replaced by q in the second case). In the remaining one, the nature of ellipticity of the functional $\mathcal{P}$ switches between the p and q rates accordingly to the value of $a(\cdot)$. For this reason models as those in (1.2) are particularly useful to describe strongly anisotropic media. We refer to the papers [2, 7, 9, 17, 18, 54, 65] for more results, different directions and related topics. Another, softer instance of functional with non-standard growth used to describe anisotropic models [1, 65, 66, 67] is the variable exponent one

$$w \mapsto \int_{\Omega} |Dw|^{p(x)} dx, \quad (1.6)$$

that has attracted a lot of attention in the last years [55]; a match between the two cases has been recently proposed in [15, 19, 25, 52, 55, 56, 57]. The growth conditions in (1.5) are typically linked to the non-uniform ellipticity of the Euler-Lagrange equations associated to the functionals in question. In case of (1.2), such equation is

$$-\operatorname{div} A(x, Du) = 0 \quad \text{with} \quad A(x, z) = |z|^{p-2}z + (q/p)a(x)|z|^{q-2}z, \quad (1.7)$$

and therefore, the ellipticity ratio $\mathcal{R}(z, B)$ on any ball $B \subset \Omega$ touching the transition set $\{a(x) = 0\}$, which is defined by

$$\mathcal{R}(z, B) := \frac{\sup_{x \in B} \text{of the highest eigenvalue of } \partial_z A(x, z)}{\inf_{x \in B} \text{of the lowest eigenvalue of } \partial_z A(x, z)} \approx 1 + \|a\|_{L^\infty(B)} |z|^{q-p},$$

becomes unbounded as $|z| \rightarrow \infty$. This means that the equation in (1.7) is non-uniformly elliptic. More specifically, the asymptotics of the ratio $\mathcal{R}(z, B)$ exhibit a delicate interplay between the size of $|z|^{q-p}$ and the one of the coefficient $a(\cdot)$ which is crucially close to the zero set $\{a(x) = 0\}$. As a matter of fact, the rate $a(\cdot)$ approaches to zero rebalances the rate of potential blow-up. This is displayed in the sharp condition $q \leq p + \alpha$, which in [4, 10, 11] is found to be necessary and sufficient for unconstrained, *bounded*, scalar minimizers of the model functional (1.2) to be regular; otherwise discontinuous minimizers of $\mathcal{P}$ may exist, [25, 27]. Let us remark that both the model functional in (1.2) and the one in (1.6) fall in the realm of non-autonomous functionals defined in Musielak-Orlicz spaces. These are functionals of the type

$$w \mapsto \int_{\Omega} \Phi(x, |Dw|) dx, \quad (1.8)$$

where, $\Phi: \Omega \times [0, \infty) \rightarrow [0, \infty)$ is a Caratheodory function such that for each choice of $x \in \Omega$, the partial map $t \mapsto \Phi(x, t)$ is a Young function and thereby generates an Orlicz space that changes with x . Such functionals are naturally defined on Musielak-Orlicz-Sobolev spaces/classes $W^{1,\Phi}$, i.e.,

$$W^{1,\Phi}(\Omega, \mathbb{R}^N) = \left\{ f \in W^{1,1}(\Omega, \mathbb{R}^N) : \Phi(\cdot, |f|) + \Phi(\cdot, |Df|) \in L^1(\Omega) \right\}. \quad (1.9)$$

For such spaces we refer to the recent interesting monograph [34]. A main problem here is the one of finding general conditions ensuring the regularity of minimizers. This appears as a non-trivial and challenging issue. The main idea is that the regularity of minima of (1.8) is governed by a delicate interplay between the regularity of the function $x \mapsto \Phi(x, \cdot)$ and the growth conditions of $t \mapsto \Phi(\cdot, t)$. For instance, in the double phase case, relations as $q \leq p + \alpha$ or $q/p \leq 1 + \alpha/n$ define sharp conditions for regularity [4, 6, 9, 10]. In the variable exponent case the log-modulus of continuity of $p(\cdot)$ is another instance of such conditions. See

[3] for a picture concerning the similarities between (1.2) and (1.6) as particular cases of the one in (1.8). More general conditions unifying those for (1.2) and (1.6) have been formulated in the interesting paper [35] and lead to a full De Giorgi-Nash-Moser theory. A general approach to the gradient regularity has been devised and suggested in [4]. The regularity problem in the case of constrained minimizers for functionals with non-standard growth conditions has recently received some attention, see the higher integrability result recently obtained in [13] and the singular set estimates proved for the variable exponent case in [15]. The main difficulties essentially rely in the lack of a certain number of properties, that are typically linked to uniform ellipticity and that are essential in order to treat constrained minimizers. In this paper we undertake this issue in the model case of double phase energies, i.e., functionals of the type in (1.1) controlled by the one in (1.2) in the sense of (1.12) below. Moreover, we consider the case when the manifold is the $(N - 1)$ -dimensional sphere $\mathbb{S}^{N-1}$ in $\mathbb{R}^N$, that already incorporates several of the new difficulties. Our aim here is also to propose an intrinsic approach which departs from the usual estimates, and it is designed for treating the quantity $\Phi(\cdot, |Dw|)$ which in this case is $|Dw|^p + a(\cdot)|Dw|^q$, as a sort of replacement of $|Dw|^p$. We shall therefore formulate and use a certain number of tools (harmonic approximation lemmas, a priori estimates and so on) in terms of the quantity $\Phi(\cdot, |Dw|)$. Accordingly to this viewpoint, in order to characterize the singular sets, we shall use an intrinsic Hausdorff type measure aimed at catching the local geometry of the integrand $\Phi(\cdot, t)$. Such measures give back the standard Hausdorff measure in the case $\Phi(\cdot, t) = t^p$ as well as other examples of measures available in the literature. We then compare these measures to the natural capacities generated by functionals of the type in (1.8) and relate the corresponding outcomes to the size of the singular sets, that are indeed found to have zero capacity.

1.1 Partial regularity

It is convenient to introduce some notation (see also Section 2 below). We shall denote

$$H(x, z) := |z|^p + a(x)|z|^q, \quad \text{for all } z \in \mathbb{R}^{N \times n}, \quad x \in \Omega, \quad (1.10)$$

where $a(\cdot)$ is as in (1.2) and recall that in the following it will always be $n \geq 2$ and $N > 1$ (in fact, sometimes we shall consider also the case $z \in \mathbb{R}^{(N-1) \times n}$; see Section 2 and Step 4 from the proof of Theorem 1 below). Moreover, with $B \Subset \Omega$ being a ball, we introduce the auxiliary Young functions

$$\begin{cases} H_B^-(z) := |z|^p + a_i(B)|z|^q, & \text{where } a_i(B) := \inf_{x \in B} a(x); \\ H_B^+(z) := |z|^p + a_s(B)|z|^q, & \text{where } a_s(B) := \sup_{x \in B} a(x). \end{cases} \quad (1.11)$$

With abuse of notation, we shall keep on denoting $H(x, t) = t^p + a(x)t^q$ for $t \geq 0$ (and the like for the functions in (1.11)), that is when in (1.10) z is a non-negative number. In the rest of the paper, with $B \subset \mathbb{R}^n$ being a ball, we shall denote by $r(B)$ its radius. Following [4], we then consider variational integrals of the type in (1.1), where $F: \Omega \times \mathbb{R}^N \times \mathbb{R}^{N \times n} \rightarrow \mathbb{R}$ is a Carathéodory integrand, such that $z \mapsto F(\cdot, z)$ is $C^1(\mathbb{R}^{N \times n}) \cap C^2(\mathbb{R}^{N \times n} \setminus \{0\})$ and satisfies the following assumptions:

$$\begin{cases} \nu H(x, z) \leq F(x, v, z) \leq LH(x, z) \\ |\partial F(x, v, z)||z| + |\partial^2 F(x, v, z)||z|^2 \leq LH(x, z) \\ \nu(|z|^{p-2} + a(x)|z|^{q-2})|\xi|^2 \leq \langle \partial^2 F(x, v, z)\xi, \xi \rangle \\ |\partial F(x_1, v, z) - \partial F(x_2, v, z)||z| \leq L\omega(|x_1 - x_2|)[H(x_1, z) + H(x_2, z)] + L|a(x_1) - a(x_2)||z|^q \\ |F(x, v_1, z) - F(x, v_2, z)| \leq L\omega(|v_1 - v_2|)H(x, z). \end{cases} \quad (1.12)$$

These are assumed to hold whenever $x, x_1, x_2 \in \Omega$, $v, v_1, v_2 \in \mathbb{R}^N$, $z, z_1, z_2 \in \mathbb{R}^{N \times n} \setminus \{0\}$, $\xi \in \mathbb{R}^{N \times n}$, where $0 < \nu \leq 1 \leq L$ and $\vartheta \in (0, 1)$ are fixed constants and, for every non-negative number t ,

$$\omega(t) := \min\{t^\beta, 1\}, \quad \beta \in (0, 1], \quad (1.13)$$

is defined as the standard concave β -Hölder modulus of continuity. Note that, here as in the following, by "∂" we always mean the partial derivative with respect to the gradient variable z . We finally consider the necessary structure assumption to deal with the vectorial case, that is, we assume that for every choice of $(x, v) \in \Omega \times \mathbb{R}^N$ there exists a function $\tilde{F}_{x,v}(\cdot) \equiv \tilde{F}(x, v, \cdot): [0, \infty) \rightarrow [0, \infty)$ of class $C^1[0, \infty) \cap C^2(0, \infty)$, such that

$$\begin{cases} F(x, v, z) = \tilde{F}(x, v, |z|) \text{ holds for every } z \in \mathbb{R}^{N \times n} \text{ with } t \mapsto \tilde{F}(x, v, t) \text{ being non-decreasing,} \\ |\tilde{F}''(x, v, t+s) - \tilde{F}''(x, v, t)| \leq \frac{LH(x, t)}{t^2} \left(\frac{|s|}{t} \right)^{\beta_1}. \end{cases} \quad (1.14)$$

The inequality in the last line is assumed to hold whenever $s, t \in \mathbb{R}$ are such that $t > 0$ and $2|s| < t$, and with a fixed constant $\beta_1 \in (0, 1]$ (which is independent of the considered (x, v)). As for the exponents p, q , we assume

$$p < q < p + \alpha, \quad q < N. \quad (1.15)$$

We remark that the inequality $q < p + \alpha$ in the last display, apart from the missing equality case, is a sharp condition for regularity, as shown in [25, 27]. The second inequality $q < N$ relates the growth conditions of the problem and the topological properties of the target manifold, which is in this case $\mathbb{S}^{N-1}$. This is necessary in order to use certain projection operators (see Lemma 1 below). Assumptions of this kind are considered by Hardt & Lin [32, 33] in the convex case and by Hopper [38] in the quasiconvex one. The related definition of local minimizer we are going to consider is the following:

Definition 1 *A function $u \in W_{\text{loc}}^{1,1}(\Omega, \mathbb{S}^{N-1})$ is a local minimizer of the functional $\mathcal{F}$ defined in (1.1) under assumptions (1.12)₁, if and only if $H(\cdot, Du) \in L_{\text{loc}}^1(\Omega)$ and the minimality condition $\mathcal{F}(u, \text{supp}(u - v)) \leq \mathcal{F}(v, \text{supp}(u - v))$ is satisfied whenever $v \in W_{\text{loc}}^{1,1}(\Omega, \mathbb{S}^{N-1})$ is such that $\text{supp}(u - v) \subset \Omega$.*

By definition a local minimizer belongs to $W_{\text{loc}}^{1,p}(\Omega, \mathbb{S}^{N-1})$; in the rest of the paper we shall appeal such local minimizers sometimes as constrained local minimizers to emphasize that the presence of the constraint $|u| = 1$. We notice when $a(\cdot) \equiv 0$ assumptions (1.12) reduce to the standard ones considered for functionals with p -growth when considering partial regularity problems (see for instance [39, 40, 48, 50] and related references). In particular, assumptions (1.12)-(1.14) are devised to cover functionals of the type

$$w \mapsto \int_{\Omega} [F_1(x, w, Dw) + a(x)F_2(x, w, Dw)] dx,$$

where $F_1(\cdot)$ and $F_2(\cdot)$ have p - and q -growth, respectively, accordingly to the standard assumptions described for instance in [39]. Another functional covered by our set of assumptions is

$$W^{1,1}(\Omega) \ni w \mapsto \int_{\Omega} b(x, w)H(x, Dw) dx,$$

where $0 < v_1 \leq b(x, v) \leq L_1$, for some constants v_1, L_1 . Here, $b(\cdot)$ is a Hölder-continuous function. For later convenience, we shall denote

$$\text{data} \equiv \text{data}(n, N, v, L, p, q, \alpha, [a]_{0,\alpha}), \quad (1.16)$$

as the set of basic parameters intervening in the problem. Our first main result is the following:

Theorem 1 *Let $u \in W_{\text{loc}}^{1,1}(\Omega, \mathbb{S}^{N-1})$ be a local minimizer of the functional $\mathcal{F}$ in (1.1) under the assumptions (1.12)-(1.15). There exists $\delta_g \equiv \delta_g(\text{data}) > 0$ such that*

$$H(\cdot, Du) \in L_{\text{loc}}^{1+\delta_g}(\Omega). \quad (1.17)$$

Moreover, there exist $\beta_0 \equiv \beta_0(\text{data}, \beta, \beta_1) > 0$ and an open subset $\Omega_u \subset \Omega$, called the regular set, such that

$$Du \in C_{\text{loc}}^{0, \beta_0}(\Omega_u, \mathbb{R}^{N \times n}) \quad \text{and} \quad |\Omega \setminus \Omega_u| = 0. \quad (1.18)$$

In the case $p(1 + \delta_g) > n$ we have $\Omega = \Omega_u$. When $p(1 + \delta_g) \leq n$, there exists a number $\varepsilon \equiv \varepsilon(\text{data}, \beta)$, such that a point $x_0 \in \Omega$ belongs to Ω_u iff

$$\left[H_{B_r(x_0)}^- \left(\frac{\varepsilon}{r} \right) \right]^{-1} \int_{B_r(x_0)} H(x, Du) dx < 1 \quad (1.19)$$

holds for some ball $B_r(x_0) \Subset \Omega$ with $r \leq 1$. Finally, as for the so-called singular set $\Sigma_u := \Omega \setminus \Omega_u$, it follows that

$$\Sigma_u = \left\{ x_0 \in \Omega : \limsup_{\varrho \rightarrow 0} \left[H_{B_\varrho(x_0)}^- \left(\frac{1}{\varrho} \right) \right]^{-1} \int_{B_\varrho(x_0)} H(x, Du) dx > 0 \right\}. \quad (1.20)$$

The ε -regularity condition (1.19) differs from the usual ones given in the case of functionals with p -growth as it gives an intrinsic quantified version of the amount of energy needed for regularity; see also the interesting paper [21] for the case of autonomous functionals. The shape of (1.19) suggests an intrinsic path to estimate the size of the so-called singular set $\Omega \setminus \Omega_u$. Indeed, this can be done via a general definition of certain Hausdorff type measures that can be useful in general contexts too; for this we refer to the next section. It is worth remarking that the results of Theorem 1 continue to hold in the case of unconstrained *bounded* minimizers and it is new in the vectorial case (it extends the scalar one in [4]). In the unconstrained case the condition $q < N$ in (1.5) can be dropped (see Remark 5 below).

Remark 1 It is still possible to get a partial regularity result by weakening the assumptions on the function $\tilde{F}_{x,v}(t) \equiv F(x, v, t)$ considered in (1.14). Specifically, we can drop (1.14)₂, thereby replacing (1.18) with the weaker outcome

$$u \in C_{\text{loc}}^{0, \beta_2}(\Omega_u, \mathbb{R}^{N \times n}) \quad \text{and} \quad |\Omega \setminus \Omega_u| = 0, \quad (1.21)$$

for every $\beta_2 < 1$.

1.2 Weighted Hausdorff measures, intrinsic capacities and singular sets

Here we shall be slightly more general than what is needed in the present setting, as we wish to settle down a general approach valid also for other contexts. We shall produce a family of Hausdorff type measures that are naturally linked to general functionals of the type in (1.8). In the following we consider a Carathéodory function $\Phi: \Omega \times [0, \infty) \rightarrow [0, \infty)$, i.e., such that $x \mapsto \Phi(x, t)$ is measurable for every $t \geq 0$ and $t \mapsto \Phi(x, t)$ is continuous and non-decreasing for almost every $x \in \Omega$. Here, $\Omega \subset \mathbb{R}^n$ denotes an open subset. Moreover, we assume that $\Phi(x, 0) = 0$ and that $\lim_{t \rightarrow \infty} \Phi(x, t) = \infty$ for every $x \in \Omega$. We also assume that

$$\Phi(x, t) \lesssim m(x)t^n, \quad \text{for all } t \geq 1, \text{ a.e. } x \in \Omega, \quad \text{where } 0 \leq m(\cdot) \in L_{\text{loc}}^1(\Omega) \quad (1.22)$$

and that

$$\text{there exists } \beta_3 \in (0, 1) \text{ such that } \Phi(x, \beta_3) \leq 1 \text{ and } \Phi(x, 1/\beta_3) \geq 1 \text{ for every } x \in \Omega, \quad (1.23)$$

$$\frac{\Phi(x, s)}{s} \lesssim \frac{\Phi(x, t)}{t} \quad \text{whenever } 0 < s \leq t, \text{ for all } x \in \Omega. \quad (1.24)$$

These assumptions, also considered in [5], are trivially verified by all the relevant model examples motivating us; see Remark 2 below. To proceed, for any n -dimensional open ball $B \subset \Omega$ (there is no difference in the following in taking closed balls in this respect) of radius $r(B) \in (0, \infty)$, we define

$$h_\Phi(B) = \int_B \Phi(x, 1/r(B)) dx. \quad (1.25)$$

Notice that this function is always finite and that this is guaranteed by (1.22). It results:

$$h_\Phi(B) \lesssim \int_B m(x) dx .$$

We then use the standard Carathéodory's construction to obtain an outer measure. For this, let $E \subset \Omega$ be any subset. We define the weighted κ -approximating Hausdorff measure of E , $\mathcal{H}_{\Phi,\kappa}(E)$ with $0 < \kappa \leq 1$, by

$$\mathcal{H}_{\Phi,\kappa}(E) = \inf_{C_E^\kappa} \sum_j h_\Phi(B_j) , \quad (1.26)$$

$$C_E^\kappa = \left\{ \{B_j\}_{j \in \mathbb{N}} \text{ is a countable collection of balls } B_j \subset \Omega \text{ covering } E \text{ such that } r(B_j) \leq \kappa \right\} .$$

As $0 < \kappa_1 < \kappa_2 < \infty$ implies $C_E^{\kappa_1} \subset C_E^{\kappa_2}$, we have that $\mathcal{H}_{\Phi,\kappa_1}(E) \geq \mathcal{H}_{\Phi,\kappa_2}(E)$ and there exists the limit

$$\mathcal{H}_\Phi(E) := \lim_{\kappa \rightarrow 0} \mathcal{H}_{\Phi,\kappa}(E) = \sup_{\kappa > 0} \mathcal{H}_{\Phi,\kappa}(E) . \quad (1.27)$$

When considering functionals of the type in (1.8), it is convenient to localize the x -dependence and locally compare the starting integrand $\Phi(\cdot)$ with similar maps that are independent of x . This means that, with a ball $B \subset \Omega$ being fixed, we consider the functions $t \mapsto \text{ess inf}_{x \in B} \Phi(x, t)$ and $t \mapsto \text{ess sup}_{x \in B} \Phi(x, t)$, and define

$$h_\Phi^+(B) = |B| \text{ess sup}_{x \in B} \Phi(x, 1/r(B)) \quad \text{and} \quad h_\Phi^-(B) = |B| \text{ess inf}_{x \in B} \Phi(x, 1/r(B)) \quad (1.28)$$

so that $h_\Phi^-(B) \leq h_\Phi(B) \leq h_\Phi^+(B)$. Accordingly, keeping (1.26)-(1.27), we finally set

$$\mathcal{H}_{\Phi,\kappa}^\pm(E) = \inf_{C_E^\kappa} \sum_j h_\Phi^\pm(B_j) \quad \text{and} \quad \mathcal{H}_\Phi^\pm(E) = \lim_{\kappa \rightarrow 0} \mathcal{H}_{\Phi,\kappa}^\pm(E) . \quad (1.29)$$

The above definitions obviously imply that $\mathcal{H}_{\Phi,\kappa}^-(E) \leq \mathcal{H}_{\Phi,\kappa}(E) \leq \mathcal{H}_{\Phi,\kappa}^+(E)$ holds for every $\kappa \in (0, 1]$ and therefore, upon letting $\kappa \rightarrow 0$, it follows that

$$\mathcal{H}_\Phi^-(E) \leq \mathcal{H}_\Phi(E) \leq \mathcal{H}_\Phi^+(E) . \quad (1.30)$$

Remark 2 Definition (1.27) is aimed at catching and unifying several instances of similar objects. Furthermore, let us notice that

- In the case $\Phi(x, t) \equiv t^p$ for $p \leq n$, then $\mathcal{H}_\Phi$ is equivalent (up to constants) to the usual $(n - p)$ -dimensional spherical Hausdorff measure.
- In the case $\Phi(x, t) \equiv t^{p(x)}$ for $p(x) \leq n$ being a continuous function defined on an open subset Ω , then $\mathcal{H}_\Phi$ falls in the class of the variable exponent Hausdorff measures studied in [51, 61].
- In the case $\Phi(x, t) \equiv w(x)t^p$ for $p \leq n$ and $w(\cdot)$ being a non-negative and measurable function, $\mathcal{H}_\Phi$ is equivalent to the weighted Hausdorff measures introduced in [51, 61], with particular emphasis on the situations when $w(\cdot)$ is a Muckenhoupt weight.
- The case we are mostly interested in is when $\Phi(x, t) = [H(x, t)]^{1+\delta} \equiv [t^p + a(x)t^q]^{1+\delta}$ for some $\delta \geq 0$, with $H(\cdot)$ as in (1.10) and under the condition that $q(1 + \delta) \leq n$. In this case we shall use the notation $\mathcal{H}_\Phi \equiv \mathcal{H}_{H^{1+\delta}}$ and $\mathcal{H}_\Phi^\pm \equiv \mathcal{H}_{H^{1+\delta}}^\pm$. These measures reveal to be an essential tool to prove sharp theorems of removable singularities in double phase problems, as shown in the forthcoming paper [8].

Remark 3 By standard arguments, i.e., those of the type needed in the case of the usual Hausdorff measures, the set function $\mathcal{H}_\Phi$ turns out to be a Borel-regular measure (here we adopt the standard terminology from [26]). We notice that the definition of $\mathcal{H}_{\Phi,x}$ is invariant when using open or closed balls in (1.25). As for the set functions $\mathcal{H}_\Phi^\pm$ in (1.29), these turn out to be Borel regular measures too. We mention an alternative way to describe measures as $\mathcal{H}_\Phi^\pm$. This occurs upon replacing (1.28) by

$$h_\Phi^+(B) = |B| \sup_{x \in B} \Phi(x, 1/r(B)) \quad \text{and} \quad h_\Phi^-(B) = |B| \inf_{x \in B} \Phi(x, 1/r(B)) . \quad (1.31)$$

In this case the corresponding set functions $\mathcal{H}_\Phi^\pm$ are again Borel measures and are Borel regular too if $\Phi(\cdot)$ is continuous. Alternatively, one can use in the definition (1.31) closed balls instead of open ones, thereby always getting automatically a Borel regular measure.

In order to place the above measures in the setting of regularity of minimizers and to connect the three measures appearing in (1.30), we next consider the following assumption:

$$\operatorname{ess\,sup}_{x \in B} \Phi(x, \beta_4 t) \leq c_d \operatorname{ess\,inf}_{x \in B} \Phi(x, t) \quad (1.32)$$

to hold whenever $1 \leq t \leq 1/r(B)$ for all balls $B \subset \Omega$ with $r(B) \leq 1$ and for some constants $\beta_4 \in (0, 1]$, $c_d \geq 1$. This assumption is known to be crucial to prove the local Hölder continuity of *bounded* minimizers of functionals of the type in (1.8) and in certain Harmonic Analysis questions related to Musielak-Orlicz spaces; see [35, 36]. In this respect, assumption (1.32) is sharp by the examples in [25, 27]. When applied to the choice $\Phi(x, t) = t^p + a(x)t^q$ and $a(\cdot) \in C^{0,\alpha}(\Omega)$, (1.32) amounts to require that $q \leq p + \alpha$ as first considered in [10]; see Proposition 1 and again (6.1) below. An immediate consequence of (1.32) is the following fact, whose proof is reported in Section 6.

Proposition 1 *Assume that (1.32) holds. Then, for any subset $E \subset \Omega$ it follows that*

$$\mathcal{H}_\Phi^+(E) \leq \frac{c_d}{\beta_4^n} \mathcal{H}_\Phi^-(E) . \quad (1.33)$$

As a consequence, if $a(\cdot) \in C^{0,\alpha}(\Omega)$, $q \leq p + \alpha$ and $\delta \geq 0$, then there exists a constant $c \equiv c([a]_{0,\alpha}, \delta) \geq 1$ such that the following inequality holds for every subset $E \subset \mathbb{R}^n$:

$$\mathcal{H}_{H^{1+\delta}}^-(E) \leq \mathcal{H}_{H^{1+\delta}}(E) \leq \mathcal{H}_{H^{1+\delta}}^+(E) \leq c \mathcal{H}_{H^{1+\delta}}^-(E) . \quad (1.34)$$

Following [5], we now introduce a notion of (relative) capacity generated by the function $\Phi(\cdot)$. For a compact subset $K \subset \mathbb{R}^n$, we denote

$$\operatorname{Cap}_\Phi^*(K) \equiv \operatorname{Cap}_\Phi^*(K, \Omega) := \inf_{f \in \mathcal{R}(K)} \int_\Omega \Phi(x, |Df|) dx \quad (1.35)$$

where

$$\mathcal{R}(K) := \left\{ f \in W^{1,\Phi}(\Omega) \cap C_0(\Omega) : f \geq 1 \text{ in } K, f \geq 0 \right\} .$$

As usual, for open subsets $U \subset \Omega$ we set

$$\operatorname{Cap}_\Phi(U) := \sup_{K \subset U, K \text{ is compact}} \operatorname{Cap}_\Phi^*(K)$$

and then, for general sets $E \subset \Omega$ we finally define

$$\operatorname{Cap}_\Phi(E) := \inf_{E \subset \tilde{U} \subset \Omega, \tilde{U} \text{ is open}} \operatorname{Cap}_\Phi(\tilde{U}) .$$

It turns out that, under the present assumptions on $\Phi(\cdot)$, we have $\text{Cap}_\Phi^*(K) = \text{Cap}_\Phi(K)$, whenever $K \subset \Omega$ is a compact subset and therefore the symbol Cap_Φ^* will not be used anymore, see [5, Proposition 6.3]. Anisotropic capacities of this kind have been studied at length in the literature. Classical reference in this respect are [12, 28, 37, 49, 51]. Here we refer to the recent paper [5], where such capacities have been studied in detail under the assumptions in (1.23)-(1.24) considered here. These ensure that Cap_Φ enjoys the standard properties of Sobolev capacities; in particular Cap_Φ is a Choquet capacity in the sense that

$$\text{Cap}_\Phi(E) = \sup \left\{ \text{Cap}_\Phi(K) : K \subset E \text{ and } K \text{ is compact} \right\} \quad (1.36)$$

holds for every Borel set $E \subset \Omega$, [5, Remark 3.6]. Needless to say, in the case $\Phi(x, t) \equiv t^p$, Cap_Φ coincides with the usual relative $W^{1,p}$ -capacity. In the following we shall denote $\text{Cap}_H = \text{Cap}_\Phi$ when $\Phi(x, t) = t^p + a(x)t^q$. Exactly as in the case of the $W^{1,p}$ -capacity, we can prove a relation between capacity and Haudorff measures. For this, we need some more assumptions. Specifically, we assume that there exist $1 < p < q < \infty$ such that

$$\frac{\Phi(x, s)}{s^p} \leq c_g \frac{\Phi(x, t)}{t^p} \quad \text{and} \quad \frac{\Phi(x, t)}{t^q} \leq c_g \frac{\Phi(x, s)}{s^q} \quad \text{whenever } 0 < s \leq t, \quad (1.37)$$

for some $c_g \geq 1$. We then have

Theorem 2 *Assume that (1.32) and (1.37) are in force. Let $E \subset \mathbb{R}^n$ be such that $\mathcal{H}_\Phi(E) < \infty$, then $\text{Cap}_\Phi(E) = 0$.*

It is now possible to improve the estimates of the Hausdorff measure of the singular set $\Sigma_u := \Omega \setminus \Omega_u$ from Theorem 1. This is in the following:

Theorem 3 *Let $u \in W_{\text{loc}}^{1,1}(\Omega, \mathbb{S}^{N-1})$ be a local minimizer of the functional $\mathcal{F}$ in (1.1) under the assumptions (1.12)-(1.15), and let $\Omega_u \subset \Omega$ be its regular set in the sense of Theorem 1. Assume that $q(1 + \delta_g) \leq n$, where δ_g is the number appearing in (1.17). Then*

$$\mathcal{H}_{H^{1+\delta_g}}(\Omega \setminus \Omega_u) = 0 \quad \text{and therefore} \quad \text{Cap}_{H^{1+\delta_g}}(\Omega \setminus \Omega_u) = 0. \quad (1.38)$$

In particular, we have

$$\mathcal{H}^{n-p-p\delta_g}(\Sigma_u) = 0, \quad (1.39)$$

and

$$\mathcal{H}^{n-q-q\delta_g}(\Sigma_u \cap \{a(x) > 0\}) = 0. \quad (1.40)$$

1.3 Overview of the paper

As mentioned at the beginning of the Introduction, the aim of this paper is not only to prove regularity results for constrained local minimizers of double phase functionals, but also to expose intrinsic techniques bound to cover general functionals of the type in (1.8). In this sense, this paper further develops the ideas introduced in [4] to get general regularity methods for non-autonomous functionals and also simplifies some of the arguments presented there. Moreover, the techniques considered here provide new results also in the unconstrained case. For instance, a partial regularity theory which is analogous to the classical one for standard p -functionals can be derived in the double phase case too (see Remark 5). The paper is structured as follows. In Section 2 we fix some notation. In Section 3 we establish some basic energy and higher integrability inequalities adapting the path developed in [9, 10] to the manifold constrained case. This is based on a projection argument exposed in Lemma 1. Moreover, we derive the precise form of the Euler-Lagrange equation of functionals of the type in (1.1), under assumptions (1.12)-(1.14). We finally readapt a

Morrey type decay estimate originally proved in [10] (see Theorem 4). In Section 4 we develop an intrinsic harmonic type approximation result (compactness lemma), which is Lemma 7. The main novelty is that the energy bounds and the approximation are given directly in the intrinsic terms of a Musielak-Orlicz energy, rather than a more typical Orlicz one, as usually done in the literature [21, 23]. The lemma is quantitative, in the sense that it reveals a power type dependence of the constants. It therefore extends a similar result previously obtained in [4], which was there considered in a more classical Orlicz setting. It is interesting to note that the proof of Lemma 7 involves the use of an a priori smallness assumption (see (4.5) below) which is exactly the one which is needed to prove partial regularity in the subsequent Section 4. The conceptual advantage of using such an approach becomes clear in Section 4, where partial regularity and Theorem 1 are proved. The intrinsic approach adopted in Lemma 7 allows avoiding to readapt the elaborate arguments of [4, 9, 10] as at this point we can directly use the intrinsic Morrey decay estimate of Theorem 4 as a natural reference estimate. This incorporates the regularity information on the solutions indeed developed in [4, 9, 10]. The final outcome is a treatment which is close to the classical one proposed by Simon [60] in the case of harmonic maps. Finally, in Section 6 we develop the arguments concerning the Hausdorff type measures presented in Section 1.2.

2 Notation

In this paper, following a usual custom, we denote by c a general constant larger than one. Different occurrences from line to line will be still denoted by c , while special occurrences will be denoted by $c_1, c_2, \tilde{c}$ or the like. Relevant dependencies on parameters will be emphasized using parentheses, i.e., $c_1 \equiv c_1(n, p, \nu, L)$ means that c_1 depends on n, p, ν, L . We denote by $B_r(x_0) := \{x \in \mathbb{R}^n : |x - x_0| < r\}$ the open ball with center x_0 and radius $r \equiv r(B) > 0$; when not important, or clear from the context, we shall omit denoting the center as follows: $B_r \equiv B_r(x_0)$. Very often, different balls in the same context will share the same center. We shall also denote $B_1 = B_1(0)$ if not differently specified. Moreover, with B being a given ball with radius $r \equiv r(B)$ and γ being a positive number, we denote by γB the concentric ball with radius γr . Finally, when referring to balls in $\mathbb{R}^N$, we will stress it with the apex N , i.e.: $B_r^N(a_0)$ is the open ball with center $a_0 \in \mathbb{R}^N$ and positive radius r . With $\mathcal{B} \subset \mathbb{R}^n$ being a measurable subset having finite and positive measure $|\mathcal{B}| > 0$, and with $g: \mathcal{B} \rightarrow \mathbb{R}^k, k \geq 1$, being a measurable map, we shall denote by

$$(g)_{\mathcal{B}} \equiv \int_{\mathcal{B}} g(x) dx := \frac{1}{|\mathcal{B}|} \int_{\mathcal{B}} g(x) dx$$

its integral average. Similarly, with $\gamma \in (0, 1)$ we denote

$$[g]_{0,\gamma;\mathcal{B}} := \sup_{x,y \in \mathcal{B}, x \neq y} \frac{|g(x) - g(y)|}{|x - y|^\gamma}.$$

In the case of the reference coefficient function $a(\cdot)$ in (1.2), we shall denote $[a]_{0,\alpha} \equiv [a]_{0,\alpha;\Omega}$. As usual when dealing with p -Laplacean type problems, we shall often use the auxiliary vector fields $V_p, V_q: \mathbb{R}^{k \times n} \rightarrow \mathbb{R}^{k \times n}$, with $k \geq 1$ being an integer, defined by

$$V_p(z) := |z|^{(p-2)/2} z \quad \text{and} \quad V_q(z) := |z|^{(q-2)/2} z \quad (2.1)$$

whenever $z \in \mathbb{R}^{k \times n}$; we notice that

$$H(x, z) = |V_p(z)|^2 + a(x)|V_q(z)|^2, \quad \text{for all } x \in \Omega, z \in \mathbb{R}^{k \times n}. \quad (2.2)$$

A useful related inequality is contained in the following

$$|V_t(z_1) - V_t(z_2)| \approx (|z_1| + |z_2|)^{(t-2)/2} |z_1 - z_2|, \quad t \in \{p, q\}, \quad (2.3)$$

where the equivalence holds up to constants depending only on n, k, t .

Remark 4 In this paper, when considering the maps in (2.1) and also $H(\cdot)$, we shall typically choose $k \in \{N-1, N\}$, where $N > 1$ is the number considered in Theorem 1. Actually, we shall almost always consider the case it is $z \in \mathbb{R}^{N \times n}$. When projecting on the sphere (in Step 4 of the proof of Theorem 1, for example), we shall also consider $z \in \mathbb{R}^{(N-1) \times n}$. Usually, we shall not specify the case occurring, as it will be clear from the context.

As a consequence of (1.12)₃, it can be proved that

$$|V_p(z_1) - V_p(z_2)|^2 + a(x)|V_q(z_1) - V_q(z_2)|^2 \leq c [\partial F(x, v, z_1) - \partial F(x, v, z_2)] \cdot (z_1 - z_2) \quad (2.4)$$

holds whenever $z_1, z_2 \in \mathbb{R}^{N \times n}$, $x \in \Omega$, $v \in \mathbb{R}^N$ and with $c \equiv c(n, N, v, p, q)$. For this see for instance [4, 39]. We similarly have, again from (1.12)₃

$$\begin{aligned} & |V_p(z_2) - V_p(z_1)|^2 + a(x)|V_q(z_2) - V_q(z_1)|^2 + \partial F(x, v, z_1) \cdot (z_2 - z_1) \\ & \leq c [F(x, v, z_2) - F(x, v, z_1)] , \end{aligned} \quad (2.5)$$

whenever $z_1, z_2 \in \mathbb{R}^{N \times n}$, again for $c \equiv c(n, N, v, p, q)$. We next recall some basic terminology about Musielak-Orlicz-Sobolev spaces. The space $W^{1,H}(\Omega, \mathbb{R}^N)$ is defined as in (1.9) with the choice $\Phi(\cdot) \equiv H(\cdot)$, with the local variants being defined in the obvious way and $W_0^{1,H}(\Omega) = W^{1,H}(\Omega) \cap W_0^{1,p}(\Omega)$. In the same way, we set

$$W^{1,H}(\Omega, \mathbb{S}^{N-1}) := \left\{ w \in W^{1,H}(\Omega, \mathbb{R}^N) : |w| = 1 \text{ holds a.e.} \right\} ,$$

with the local variants defined in a similar fashion. Finally, with $u \in W^{1,H}(\Omega, \mathbb{S}^{N-1})$ we denote the Dirichlet class

$$W_u^{1,H}(\Omega, \mathbb{S}^{N-1}) := \left\{ w \in W^{1,H}(\Omega, \mathbb{S}^{N-1}) : u - w \in W_0^{1,1}(\Omega, \mathbb{R}^N) \right\} .$$

We similarly define the Dirichlet class of unconstrained maps $W_u^{1,H}(\Omega, \mathbb{R}^N)$. Moreover, with $w \in W^{1,H}(\tilde{\Omega}, \mathbb{R}^N)$ and $\tilde{\Omega}$ being a domain that allows for a trace operator (for instance, this happens when $\partial\tilde{\Omega}$ is Lipschitz), we denote by $\text{tr}(w, \partial\tilde{\Omega})$ the trace of w on $\partial\tilde{\Omega}$.

3 Basic material

3.1 Caccioppoli's and higher integrability inequalities

Following the path established in [9, 10], in this section we gather a few technical inequalities for minimizers of functionals with double phase. The main difference is that now the setting is the one of constrained variational problems. Therefore, in several cases, we shall confine ourselves to give the necessary modifications to the proofs proposed in [9, 10]. We start with the following lemma; this provides an extension result in the spirit of [33].

Lemma 1 *Let $\tilde{\Omega} \subset \Omega$ be a bounded, Lipschitz domain in $\mathbb{R}^n$ and $v \in W^{1,H}(\tilde{\Omega}, \mathbb{R}^N)$ be such that $v(\partial\tilde{\Omega}) \subset \mathbb{S}^{N-1}$. Then there exists $c \equiv c(n, N, p, q)$ and $\tilde{v} \in W^{1,H}(\tilde{\Omega}, \mathbb{S}^{N-1})$ satisfying*

$$\int_{\tilde{\Omega}} H(x, D\tilde{v}) dx \leq c \int_{\tilde{\Omega}} H(x, Dv) dx \quad \text{and} \quad v - \tilde{v} \in W_0^{1,1}(\tilde{\Omega}, \mathbb{R}^N) . \quad (3.1)$$

Proof. For $a \in B_{1/2}^N(0) \subset \mathbb{R}^N$ and v as in the statement of the lemma, define the map

$$v^a(x) := \frac{v(x) - a}{|v(x) - a|} , \quad x \in \tilde{\Omega} .$$

Clearly, it is

$$|Dv^a(x)| \leq \frac{2|Dv(x)|}{|v(x) - a|},$$

so that we can estimate

$$\begin{aligned} \int_{B_{1/2}^N(0)} H(x, Dv^a) da &\leq c \int_{B_{1/2}^N(0)} H\left(x, \frac{Dv}{|v-a|}\right) da \\ &= c \left(|Dv|^p \int_{B_{1/2}^N(0)} \frac{da}{|v-a|^p} + a(x)|Dv|^q \int_{B_{1/2}^N(0)} \frac{da}{|v-a|^q} \right) \leq cH(x, Dv). \end{aligned}$$

Here $c \equiv c(N, p, q)$ and we have used the assumption (1.15)₂; this makes the integrals in the above line finite. Integrating over $\tilde{\Omega}$, using Fubini's theorem and the content of the last display, we obtain

$$\int_{B_{1/2}^N(0)} \int_{\tilde{\Omega}} H(x, Dv^a) dx da \leq c \int_{\tilde{\Omega}} H(x, Dv) dx.$$

By Chebyshev inequality, this yields the existence of $a_0 \in B_{1/2}^N(0)$ such that

$$\int_{\tilde{\Omega}} H(x, Dv^{a_0}) dx \leq c \int_{\tilde{\Omega}} H(x, Dv) dx, \quad (3.2)$$

with $c \equiv c(n, N, p, q)$. Let us consider the projector

$$\Pi_a(y) := \frac{y-a}{|y-a|}, \quad \text{for } y \in \mathbb{S}^{N-1} \text{ and } a \in B_{1/2}^N(0).$$

Such a projector is a bilipschitz map $\mathbb{S}^{N-1}$ into itself, and it is such that

$$\left[\nabla(\Pi_a^{-1}) \right]_{0,1} \leq c = c(N), \quad (3.3)$$

an estimate which is independent of $a \in B_{1/2}^N(0)$. Since $v^a(x) \in \mathbb{S}^{N-1}$ for a.e. $x \in \Omega$ and all $a \in B_{1/2}^N(0)$, we may define $\tilde{v} := \Pi_{a_0}^{-1} \circ v^{a_0}$ which has the requested features. In fact, since $v(\partial\tilde{\Omega}) \subset \mathbb{S}^{N-1}$, we have

$$\text{tr}(\tilde{v}, \partial\tilde{\Omega}) = \Pi_{a_0}^{-1} \left(\text{tr}(v^{a_0}, \partial\tilde{\Omega}) \right) = \Pi_{a_0}^{-1} \left(\Pi_{a_0} \left(\text{tr}(v, \partial\tilde{\Omega}) \right) \right) = \text{tr}(v, \partial\tilde{\Omega})$$

and, by (3.3) and (3.2),

$$\int_{\tilde{\Omega}} H(x, D\tilde{v}) dx \leq c \int_{\tilde{\Omega}} H(x, Dv^{a_0}) dx \leq c \int_{\tilde{\Omega}} H(x, Dv) dx,$$

where $c \equiv c(n, N, p, q)$, so that (3.1) is proved in view of the last two displays. $\square$

Remark 5 The condition $q < N$ in (1.5) enters only in the proof of the above lemma and therefore can be dropped when adapting the proofs given here to the unconstrained case.

Lemma 1 allows to derive in the new constrained setting a number of preliminary tools that have been already obtained and used in the unconstrained one [9, 10]. We shortly report them, with some additional modification and informations.

Lemma 2 (Caccioppoli's Inequality) *Let $u \in W^{1,H}(\Omega, \mathbb{S}^{N-1})$ be a constrained local minimizer of the functional $\mathcal{F}$ in (1.1) under (only) assumptions (1.12)₁ and $q \leq p + \alpha$. Then there exists $c \equiv c(n, N, \nu, L, p, q) > 0$ such that for any choice of concentric balls $B_r \subset B_R \Subset \Omega$ there holds*

$$\int_{B_r} H(x, Du) dx \leq c \int_{B_R} H\left(x, \frac{u - (u)_{B_R}}{R - r}\right) dx \quad (3.4)$$

and, if $R \leq 1$, it also holds that

$$\int_{B_{R/2}} H(x, Du) dx \leq c \int_{B_R} H_{B_R}^-\left(\frac{u - (u)_{B_R}}{R}\right) dx, \quad (3.5)$$

for $c \equiv c(n, N, \nu, L, p, q, [a]_{0,\alpha}) > 0$. Moreover, if

$$\inf_{B_R} a(x) \leq 4[a]_{0,\alpha} R^\alpha \quad (3.6)$$

holds and again it is $R \leq 1$, then (3.5) reduces to

$$\int_{B_{R/2}} H(x, Du) dx \leq c \int_{B_R} \left| \frac{u - (u)_{B_R}}{R} \right|^p dx, \quad (3.7)$$

with $c \equiv c(n, N, \nu, L, p, q, [a]_{0,\alpha})$. Finally, these facts still hold for an unconstrained local minimizer $u \in (W^{1,H} \cap L^\infty)(\Omega, \mathbb{R}^N)$, with all the constants depending in addition on $\|u\|_{L^\infty}$.

Proof. The proof is a modification of the one originally given in [9, Theorem 1.1, (1.8)]; we furthermore specialize to the case of constrained minimizers, the unconstrained one being totally analogous. In the following all the balls will be concentric to the ones mentioned in the statement of the lemma. With $r \leq t < s \leq R$, we determine a cut-off function $\eta \in C_c^\infty(B_s)$ such that $\chi_{B_t} \leq \eta \leq \chi_{B_s}$ and $|D\eta| \leq 4/(s - t)$. Consider now the function $w(x) = u(x) - \eta(u - (u)_{B_R})$. Since η is smooth and $u \in W^{1,H}(B_s, \mathbb{S}^{N-1})$, then obviously $w \in W^{1,H}(B_s, \mathbb{R}^N)$ and $u - w \in W_0^{1,1}(B_s, \mathbb{R}^N)$. Lemma 1 yields the existence of $\tilde{w} \in W^{1,H}(B_s, \mathbb{S}^{N-1})$ such that (3.1) holds with $\tilde{\Omega} \equiv B_s$, where $c \equiv c(n, N, p, q)$. The minimality of u , (1.12)₁ and (3.1) (with $\Omega \equiv B_s$) yield

$$\begin{aligned} \nu \int_{B_s} H(x, Du) dx &\leq \int_{B_s} F(x, u, Du) dx \leq \int_{B_s} F(x, \tilde{w}, D\tilde{w}) dx \leq L \int_{B_s} H(x, D\tilde{w}) dx \\ &\leq c \int_{B_s} H(x, Dw) dx \leq c \int_{B_s \setminus B_t} \left[H(x, (1 - \eta)Du) + H(x, (u - (u)_{B_R}) \otimes D\eta) \right] dx \\ &\leq c \int_{B_s \setminus B_t} H(x, Du) dx + c \int_{B_s \setminus B_t} H\left(x, \frac{u - (u)_{B_R}}{s - t}\right) dx, \end{aligned} \quad (3.8)$$

with $c \equiv c(N, \nu, L, p, q)$. The proof of (3.4) can be now concluded by filling the hole and iteration, as in [9, Theorem 1.1, (1.8)], see also [10]. As for the proof of (3.5) we simply estimate (as it is $q < p + \alpha$ and $R \leq 1$), for $x \in B_R$

$$\begin{aligned} H\left(x, \frac{u - (u)_{B_R}}{R}\right) &\leq H_{B_R}^-\left(\frac{u - (u)_{B_R}}{R}\right) + \sup_{B_R} [a(x) - a_1(B_R)] \left| \frac{u - (u)_{B_R}}{R} \right|^q \\ &\leq H_{B_R}^-\left(\frac{u - (u)_{B_R}}{R}\right) + 2^{\alpha+q-p} [a]_{0,\alpha} R^{\alpha+p-q} \left| \frac{u - (u)_{B_R}}{R} \right|^p \leq c H_{B_R}^-\left(\frac{u - (u)_{B_R}}{R}\right), \end{aligned}$$

for $c = c(p, q, \alpha)$, and (3.5) follows from (3.4) with $r = R/2$. Finally, for (3.7), we similarly observe that (still $x \in B_R$)

$$a_1(B_R) \left| \frac{u - (u)_{B_R}}{R} \right|^q \stackrel{(3.6)}{\leq} 8[a]_{0,\alpha} R^{\alpha+p-q} \left| \frac{u - (u)_{B_R}}{R} \right|^p \leq c \left| \frac{u - (u)_{B_R}}{R} \right|^p,$$

so that (3.7) follows from (3.5) and the proof is complete. $\square$

We proceed with

Lemma 3 (Intrinsic Sobolev-Poincaré inequality) *Let $v \in (W^{1,H} \cap L^\infty)(\Omega, \mathbb{R}^N)$, $N \geq 1$, and $B_r \Subset \Omega$ be a ball with radius $r \leq 1$, and assume that $q \leq p + \alpha$. Then the following inequality holds*

$$\int_{B_r} H\left(x, \frac{v - (v)_{B_r}}{r}\right) dx \leq c \left(\int_{B_r} [H(x, Dv)]^d dx \right)^{1/d}, \quad (3.9)$$

where $c \equiv c(n, N, p, q, [a]_{0,\alpha}, \|v\|_{L^\infty(B_r)}) \geq 1$ and $d \equiv d(n, p, q) < 1$. In (3.9) we can replace $v - (v)_{B_r}$ by v in case we also have that $\text{tr}(v, \partial B_r) \equiv 0$. Finally, we can still replace $v - (v)_{B_r}$ by v , provided $v \equiv 0$ on $A \subset B_r$ and $|A|/|B_r| > \gamma > 0$; in this last case the constant c depends also on γ .

Proof. The proof is implicit in the one of [10, Theorem 1.2], with minor modifications that are left to the reader. See also [53]. $\square$

Lemma 4 (Inner higher integrability) *Let $u \in W^{1,H}(\Omega, \mathbb{S}^{N-1})$ be a constrained local minimizer of the functional $\mathcal{F}$ in (1.1) under (only) assumptions (1.12)₁ and (1.15). There exists a positive integrability exponent $\delta_g \equiv \delta_g(\text{data})$, such that the following reverse inequality holds for every $B_{2R} \subset \Omega$ such that $R \leq 1$:*

$$\left(\int_{B_R} [H(x, Du)]^{1+\delta_g} dx \right)^{1/(1+\delta_g)} \leq c \int_{B_{2R}} H(x, Du) dx, \quad (3.10)$$

where $c \equiv c(\text{data})$.

Proof. Also in this case, the proof follows the one for [10, Theorem 1.2], which in turn only uses the assumed bound $q < p + \alpha$ and the validity of (3.4). $\square$

Lemma 5 (Higher integrability up to the boundary) *Let $u \in W^{1,H}(\Omega, \mathbb{S}^{N-1})$ be such that $H(\cdot, Du) \in L_{\text{loc}}^{1+\delta}(\Omega)$, for some $\delta > 0$, and, for $B_R \Subset \Omega$, $R \leq 1$, let $v \in W_u^{1,H}(B_R, \mathbb{S}^{N-1})$ be a solution of*

$$v \mapsto \min_{w \in W_u^{1,H}(B_R, \mathbb{S}^{N-1})} \int_{B_R} F(x, w, Dw) dx,$$

where the Carathéodory integrand $F(\cdot)$ satisfies (only) (1.12)₁ and (1.15). Then there exists a positive exponent $\sigma_g \in (0, \delta)$ and a constant $c \geq 1$, both depending on $n, N, \nu, L, p, q, \alpha, [a]_{0,\alpha}$, such that

$$\left(\int_{B_R} [H(x, Dv)]^{1+\sigma_g} dx \right)^{1/(1+\sigma_g)} \leq c \left(\int_{B_R} H(x, Du)^{1+\sigma_g} dx \right)^{1/(1+\sigma_g)}.$$

Moreover, in the above display, σ_g can be replaced by any smaller and positive number.

Proof. With $x_0 \in B_R$, let us fix a ball $B_r(x_0) \subset \mathbb{R}^n$ such that it is $|B_r(x_0) \setminus B_R| > |B_r(x_0)|/10$. Let us fix $r/2 < t < s < r$ and take a cut-off function $\eta \in C_c^1(B_s(x_0))$ such that $\chi_{B_t(x_0)} \leq \eta \leq \chi_{B_s(x_0)}$ and $|D\eta| \leq 4/(s-t)$. The function $v - \eta(v - u)$ coincides with v in ∂B_R (in the sense of traces) and therefore we can apply Lemma 1. This provides us with a map $w \in W_v^{1,H}(B_s(x_0) \cap B_R, \mathbb{S}^{N-1})$ such that

$$\int_{B_s(x_0) \cap B_R} H(x, Dw) dx \leq c(n, N, \nu, L, p, q) \int_{B_s(x_0) \cap B_R} H(x, D(v - \eta(v - u))) dx.$$

The minimality of v and (1.12)₁, together with the above inequality, yield

$$\begin{aligned} \int_{B_s(x_0) \cap B_R} H(x, Dv) dx &\leq \frac{L}{v} \int_{B_s(x_0) \cap B_R} H(x, Dw) dx \\ &\leq c \int_{B_s(x_0) \setminus B_t(x_0) \cap B_R} H(x, Dv) dx + c \int_{B_s(x_0) \cap B_R} H(x, Du) dx + c \int_{B_s(x_0) \cap B_R} H\left(x, \frac{v-u}{s-t}\right) dx, \end{aligned}$$

with $c = c(n, N, v, L, p, q)$. By filling the hole and iterating as for instance done in [9, Proof of Theorem 1.8], we arrive at

$$\int_{B_{r/2}(x_0) \cap B_R} H(x, Dv) dx \leq c \int_{B_r(x_0) \cap B_R} H\left(x, \frac{v-u}{r}\right) dx + c \int_{B_r(x_0) \cap B_R} H(x, Du) dx,$$

for $c \equiv c(n, N, v, L, p, q)$. From this point on, we can follow the proof of [16, Lemma 5] but using the method of [10, Theorem 1.2], see also [15, Lemma 10]. $\square$

Remark 6 The assertion of Lemma 4 continues to hold in the case of unconstrained local minimizers $u \in W^{1,H}(\Omega, \mathbb{R}^N)$, $N > 1$, such that $u \in L^\infty(\Omega)$; in this case c and δ_g also depend on $\|u\|_{L^\infty}$; for this see the original proof in [10] and the extensions made in [53]. Moreover, Lemma 5 still holds when $u, v \in L^\infty(\Omega, \mathbb{R}^N)$ and, also in this case, c and σ_g again depend on $\|u\|_{L^\infty(B_R)}$ and $\|v\|_{L^\infty(B_R)}$. The proof follows again the one proposed in [16, Lemma 5] and [10, Theorem 1.2]. For later convenience we discuss a case when assumptions in (1.12) are relaxed. Instead of (1.12)₁, we consider

$$\tilde{v}(M)H(x, z) \leq F(x, v, z) \leq \tilde{L}(M)H(x, z) \quad (3.11)$$

to be satisfied as in (1.12)₁, whenever $|v| \leq M$, where $0 < v(M) \leq 1 \leq L(M)$ are, respectively, non-increasing and non-decreasing functions of $M \geq 3N$. Both Lemma 4 and of Lemma 5 hold assuming (3.11) instead of (1.12)₁, with exponents δ_g, σ_g depending again on $\|u\|_{L^\infty(B_R)}$ and $\|v\|_{L^\infty(B_R)}$. This can be easily seen (for instance in the proof of Lemma 4) by observing that $\|\tilde{w}\|_{L^\infty(B_R)} \leq 3\|u\|_{L^\infty(B_R)}$ and therefore (3.11) can be used with M depending only on $\|u\|_{L^\infty(B_R)}$ in (3.8). In the same way, the content of Lemma 2 still holds under assumptions (3.11).

Remark 7 The content of Lemma 5 applies in particular to the case when the function $a(x) \equiv a_0 \geq 0$ is constant and $H(x, z) \equiv H_0(z) := |z|^p + a_0|z|^q$. In this case assumption (1.15)₁ is not necessary and the statement continues to hold whenever $p \leq q$ are arbitrary.

3.2 On the Euler-Lagrange equation under non-standard growth conditions

Let us consider a ball $B_r \Subset \Omega$ and $v \in W^{1,H}(B_r, \mathbb{S}^{N-1})$ being a solution of the frozen Dirichlet problem

$$v \mapsto \min_{w \in W_u^{1,H}(B_r, \mathbb{S}^{N-1})} \int_{B_r} g(x, Dw) dx, \quad (3.12)$$

where, with $\bar{u} \in \mathbb{R}^N$ being fixed, we have set

$$g(x, z) := F(x, \bar{u}, z) = \tilde{F}(x, \bar{u}, |z|), \quad (3.13)$$

for $x \in \Omega$ and $z \in \mathbb{R}^{N \times n}$ and of course $F(\cdot)$ is the Carathéodory integrand considered in (1.1); this time we assume that $F(\cdot)$ satisfies only (1.12)_{1,2}. By definition, $g(\cdot)$ matches (1.12)₁-(1.12)₂. Because of the non-standard growth conditions considered here, we cannot derive the Euler-Lagrange equation for (3.12) in the

usual way, adopting variations defined through smooth φ and then concluding via a density argument. We shall rather use a direct argument, eventually leading to establish that

$$\begin{aligned} & \int_{B_r} \tilde{F}'(x, \bar{u}, |Dv|) \frac{Dv}{|Dv|} \cdot D\varphi - \tilde{F}'(x, \bar{u}, |Dv|) |Dv| (v \cdot \varphi) dx \\ &= \int_{B_r} \partial F(x, \bar{u}, |Dv|) \cdot D\varphi - \tilde{F}'(x, \bar{u}, |Dv|) |Dv| (v \cdot \varphi) dx = 0 \end{aligned} \quad (3.14)$$

holds whenever $\varphi \in (W_0^{1,H} \cap L^\infty)(B_r, \mathbb{R}^N)$; needless to say, the symbol $\tilde{F}'$ denotes the derivative of $\tilde{F}$ with respect to the last variable. For the sake of completeness we report all the details. To proceed, for $s \in (0, 1)$, define the variation $v_s := \Pi(v + s\varphi)$, where $\Pi(y) = y/|y|$ for $y \in \mathbb{R}^N \setminus \{0\}$. Clearly, for s sufficiently small, $v_s \in W_u^{1,H}(B_r, \mathbb{S}^{N-1})$. The minimality of v and the very definition of v_s tell us that

$$0 \leq \int_{B_r} \frac{g(x, Dv_s) - g(x, Dv)}{s} dx = \frac{1}{s} \int_{B_r} \left(\int_0^1 \partial g(x, \lambda Dv_s + (1-\lambda)Dv) d\lambda \right) \cdot (Dv_s - Dv) dx. \quad (3.15)$$

We aim to pass to the limit with $s \rightarrow 0$ in (3.15) via the dominated convergence theorem. A direct computation and the fact that $\nabla \Pi(v)Dv = Dv$ show that

$$\begin{aligned} |Dv_s - Dv| &= |\nabla \Pi(v + s\varphi)(Dv + sD\varphi) - \nabla \Pi(v)Dv| \\ &\leq |\nabla \Pi(v + s\varphi) - \nabla \Pi(v)| |Dv| + |s \nabla \Pi(v + s\varphi) D\varphi| \leq cs (|\varphi| |Dv| + |D\varphi|), \end{aligned} \quad (3.16)$$

where $c \equiv c(\|\nabla \Pi\|_\infty, \|\nabla^2 \Pi\|_\infty) \equiv c(N)$. Plugging (3.16) into the last term in the right-hand side of (3.15) we obtain

$$\begin{aligned} \frac{1}{s} \left| \left(\int_0^1 \partial g(x, \lambda Dv_s + (1-\lambda)Dv) d\lambda \right) \cdot (Dv_s - Dv) \right| &\leq c(|\varphi| |Dv| + |D\varphi|) \int_0^1 |\partial g(x, \lambda Dv_s + (1-\lambda)Dv)| d\lambda \\ &\leq c(\text{I}) + c(\text{II}), \end{aligned}$$

with $c \equiv c(N)$. From (1.12)₂, Young's inequality and (3.16), we estimate

$$\begin{aligned} |\text{I}| &\leq c \|\varphi\|_{L^\infty(B_r)} \int_0^1 \frac{H(x, \lambda Dv_s + (1-\lambda)Dv)}{|\lambda Dv_s + (1-\lambda)Dv|} |Dv| d\lambda \\ &\leq c \|\varphi\|_{L^\infty(B_r)} \int_0^1 [H(x, \lambda Dv_s) + H(x, (1-\lambda)Dv) + H(x, Dv)] d\lambda \leq cH(x, D\varphi) + cH(x, Dv), \end{aligned}$$

where $c \equiv c(N, L, p, q, \|\varphi\|_{L^\infty(B_r)})$. In a similar way we also have

$$\begin{aligned} |\text{II}| &\leq c \int_0^1 \frac{H(x, \lambda Dv_s + (1-\lambda)Dv)}{|\lambda Dv_s + (1-\lambda)Dv|} |D\varphi| d\lambda \\ &\leq c \int_0^1 [H(x, \lambda Dv_s) + H(x, (1-\lambda)Dv) + H(x, D\varphi)] d\lambda \leq cH(x, D\varphi) + cH(x, Dv), \end{aligned}$$

with $c \equiv c(N, L, p, q)$. Merging the content of the last three displays yields

$$\frac{1}{s} \left| \left(\int_0^1 \partial g(x, \lambda Dv_s + (1-\lambda)Dv) d\lambda \right) \cdot (Dv_s - Dv) \right| \leq cH(x, Dv) + cH(x, D\varphi), \quad (3.17)$$

again for $c \equiv c(N, L, p, q, \|\varphi\|_{L^\infty(B_r)})$. Finally, by the regularity of $z \mapsto g(\cdot, z)$, (3.16) and (1.12)₂ we notice that

$$\begin{cases} |\partial g(x, \lambda Dv_s + (1 - \lambda)Dv)| d\lambda \leq c \left(\frac{H(x, Dv)}{|Dv|} + \frac{H(x, D\varphi)}{|D\varphi|} \right) \\ \lim_{s \rightarrow 0} \partial g(x, \lambda Dv_s + (1 - \lambda)Dv) = \partial g(x, Dv) , \end{cases}$$

thus, by the dominated convergence theorem,

$$\lim_{s \rightarrow 0} \int_0^1 \partial g(x, \lambda Dv_s + (1 - \lambda)Dv) d\lambda = \partial g(x, Dv) . \quad (3.18)$$

Using (3.16) and (3.18) we can compute

$$\begin{aligned} \lim_{s \rightarrow 0} \frac{1}{s} \left(\int_0^1 \partial g(x, \lambda Dv_s + (1 - \lambda)Dv) d\lambda \right) \cdot (Dv_s - Dv) &= \partial g(x, Dv) \cdot \nabla \Pi(v) D\varphi \\ &+ \lim_{s \rightarrow 0} \left(\int_0^1 \partial g(x, \lambda Dv_s + (1 - \lambda)Dv) d\lambda \right) \cdot \frac{\nabla \Pi(v + s\varphi) - \nabla \Pi(v)}{s} Dv \\ &= \partial g(x, Dv) \cdot (\nabla \Pi(v) D\varphi + \nabla^2 \Pi(v) \varphi Dv) . \end{aligned} \quad (3.19)$$

Now, (3.15), (3.17), (3.19) and the dominated convergence theorem render

$$0 \leq \int_{B_r} \partial g(x, Dv) \cdot (\nabla \Pi(v) D\varphi + \nabla^2 \Pi(v) \varphi Dv) dx .$$

The same argument with $s \in (-1, 0)$ finally yields

$$\int_{B_r} \partial g(x, Dv) \cdot (\nabla \Pi(v) D\varphi + \nabla^2 \Pi(v) \varphi Dv) dx = 0 .$$

Taking into account the symmetry of the Jacobian of the projector, we can conclude that

$$\begin{aligned} 0 &= \int_{B_r} \tilde{F}'(x, \bar{u}, |Dv|) \frac{Dv}{|Dv|} \cdot (\nabla \Pi(v) D\varphi + \nabla^2 \Pi(v) \varphi Dv) dx \\ &= \int_{B_r} \tilde{F}'(x, \bar{u}, |Dv|) \frac{Dv}{|Dv|} \cdot D\varphi - \frac{\tilde{F}'(x, \bar{u}, |Dv|)}{|Dv|} A_v(Dv, Dv) \varphi dx \\ &= \int_{B_r} \tilde{F}'(x, \bar{u}, |Dv|) \frac{Dv}{|Dv|} \cdot D\varphi - \tilde{F}'(x, \bar{u}, |Dv|) |Dv| (v \cdot \varphi) dx , \end{aligned}$$

where in the last line we have used the explicit expression of the second fundamental form $A_v(\cdot, \cdot)$ of $\mathbb{S}^{N-1}$; see also [60, Section 2.2]. We have therefore proved the validity of (3.14).

3.3 A Morrey type decay estimate

In this section we briefly revisit some scalar regularity results reported in [4, 10], adapting them to the vectorial case. We consider unconstrained local minimizers of functionals of the type

$$w \mapsto \int_{B_r} g(x, Dw) dx \quad (3.20)$$

under the structure condition (3.13). We then have the following:

Theorem 4 ([4, 10]) Let $h \in (W^{1,H} \cap L^\infty)(\Omega, \mathbb{R}^N)$ be a local minimizer of the functional in (3.20) under assumptions (1.12)_{1,2,3,4}, (1.14)₁, (3.13) and (1.15)₁. Then for every $\sigma \in (0, n]$ there exists a constant $c \equiv c(\text{data}, \|u\|_{L^\infty(\Omega)}, \sigma)$ such that

$$\int_{B_t} H(x, Dh) dx \leq c \left(\frac{t}{s} \right)^{n-\sigma} \int_{B_s} H(x, Dh) dx \quad (3.21)$$

holds whenever $B_t \subset B_s \subset \Omega$ are concentric balls such that $t \leq 1$.

Proof. The proof can be obtained tracking the ones given for [10, Proposition 3.4] and [4, Theorem 2], according to the remarks made in the proof of [10, Theorem 5.2], that describes the modifications to make with respect to the scalar case. Reference [4] is actually more suitable as the regularity results are proved for general functionals, without assuming the splitting structure considered in [10]. The remarks given in [10, Section 5.2] will be also useful here. As an outcome of the proofs of [10, Proposition 3.4] and [4, Theorem 2], estimate (3.21) follows provided $B_t \subset B_s \subset \Omega_0 \Subset \Omega$ are concentric balls and with an additional dependence of the constant c on $\text{dist}(\Omega_0, \Omega)$, but under the full bound $q \leq p + \alpha$. As remarked in [4, 10], it is possible to reach the borderline case $q \leq p + \alpha$ in the scalar case by using the preliminary local Hölder continuity of h for some exponent $\gamma \in (0, 1)$ (see [10, Proposition 3.1]); the same happens in [4, Theorem 6]. This comes along with an a priori estimate of the type $[h]_{0,\gamma;\Omega_0} < c$, where, amongst the other things, c depends also on $\text{dist}(\Omega_0, \Omega)$. This is exactly the point where the dependence on $\text{dist}(\Omega_0, \Omega)$ comes from in the final statement of (3.21) from [10, Proposition 3.4] and [4, Theorem 2]. On the other hand, as already remarked in the proof of [10, Theorem 5.2], when considering the bound $q < p + \alpha$ we can avoid using that $u \in C_{\text{loc}}^{0,\gamma}(\Omega)$ and in this way, taking into account the proofs in [4], we arrive at (3.21) with the dependence of the constant c as described in the statement of Theorem 4. Notice that, in order to prove (3.21), in [4, 10] it is also necessary to replace the a priori Lipschitz estimate for minima of frozen functionals in [4, (132)] with an analogous one for the vectorial case. This is discussed in Remark 8 below. Notice that here we are not assuming that the function $\tilde{F}(\cdot)$ satisfies (1.14)₂. $\square$

Remark 8 Let us consider a local minimizer $v \in W^{1,H_0}(B_r, \mathbb{R}^{N \times n})$ of the functional

$$w \mapsto \int_{B_r} g(x_0, Dw) dx \quad x_0 \in \Omega, \quad (3.22)$$

where $H_0(z) \equiv |z|^p + a(x_0)|z|^q$. The following estimate holds:

$$\sup_{B_{r/2}} H_0(Dw) \leq c \int_{B_r} H_0(Dw) dx, \quad (3.23)$$

where $c \equiv c(n, N, \nu, L, p, q)$. This estimate plays a crucial role in the proofs given in [4, 10], and these are concerned with the scalar case. To get that this result holds in our vectorial case too it is sufficient to prove that

$$\tilde{F}''(x, \bar{u}, t)t \approx \tilde{F}'(x, \bar{u}, t) \quad \text{for every } t > 0 \quad (3.24)$$

(for implied constants depending only on n, N, ν, L, p, q) and then appeal for instance to [20, Lemma 5.8]. Indeed, following [14, Lemma 3.4], we see that

$$\partial^2 g(x_0, z) = \tilde{F}''(x_0, \bar{u}, |z|) \frac{z \otimes z}{|z|^2} + \tilde{F}'(x_0, \bar{u}, |z|) \left[\frac{\mathbb{I}_{N \times n}}{|z|} - \frac{z \otimes z}{|z|^3} \right], \quad (3.25)$$

holds for every $z \in \mathbb{R}^{N \times n}$ such that $|z| \neq 0$; here it is $\mathbb{I}_{N \times n} = \delta_{ij} \delta_{\alpha\beta}$. Testing the above inequality for $\xi \perp z$ and for $\xi = z$ and using (1.12)_{2,3} yields

$$\frac{\nu H(x_0, t)}{t} \leq \tilde{F}'(x_0, \bar{u}, t) \leq \frac{LH(x_0, t)}{t} \quad \text{and} \quad \frac{\nu H(x_0, t)}{t^2} \leq \tilde{F}''(x_0, \bar{u}, t) \leq \frac{LH(x_0, t)}{t^2}, \quad (3.26)$$

respectively, for every $t > 0$, so that (3.24) follows. Notice that here we are only assuming that $\tilde{F}(\cdot)$ satisfies only (1.14)₁.

Remark 9 This is a side remark of later use. Assuming that the function $\tilde{F}(\cdot)$ satisfies (1.14)_{1,2}, as in [14, Lemma 3.4], by using (3.26) we get that (1.14)₂ can be reformulated as

$$|\tilde{F}''(x, v, t + s) - \tilde{F}''(x, v, t)| \leq c(n, N, \nu, L, p, q) \tilde{F}''(x, v, t) \left(\frac{|s|}{t} \right)^{\beta_1}. \quad (3.27)$$

4 Harmonic type approximation

In this section we revisit the arguments of [4, 21], to give two kinds of harmonic type approximation lemmas. The most peculiar one is the first, which is given in terms of a generalized Young functions (specifically, $H(\cdot)$), rather than a usual Young function. Therefore all the arguments used there will be of intrinsic type. This perfectly combines with the type of intrinsic estimates already proved in [4, 9, 10], as we shall see in the next section when showing regularity theorems. Accordingly to the notation already established in (1.11), with $B_\varrho \Subset \Omega$ being a ball, we shall denote

$$H_{B_\varrho}^-(t) = t^p + a_i(B_\varrho)t^q \quad \text{and} \quad H_{B_\varrho}^+(t) = t^p + a_s(B_\varrho)t^q. \quad (4.1)$$

We shall again denote, with abuse of notation, $H_{B_\varrho}^-(z) \equiv H_{B_\varrho}^-(|z|)$ and so forth, also in the case $z \in \mathbb{R}^{N \times n}$.

Remark 10 We collect some features of the functions in (4.1). We first notice that $H_{B_\varrho}^\pm(\cdot)$ is a Young function in the sense of [20, Section 2], and satisfies the Δ_2 -condition. Since $t \mapsto H_{B_\varrho}^\pm(t)$ is strictly increasing and strictly convex, its inverse $(H_{B_\varrho}^\pm)^{-1}$ is strictly increasing and strictly concave and $(H_{B_\varrho}^\pm)^{-1}(0) = 0$, thus $(H_{B_\varrho}^\pm)^{-1}$ is subadditive. Therefore, for all $\lambda \geq 0$, the subadditivity and the monotonicity of $(H_{B_\varrho}^\pm)^{-1}(t)$ yield

$$(H_{B_\varrho}^\pm)^{-1}(\lambda t) \leq (\lambda + 1)(H_{B_\varrho}^\pm)^{-1}(t). \quad (4.2)$$

In particular, if $\lambda \geq 1$, $(H_{B_\varrho}^\pm)^{-1}(\lambda t) \leq 2\lambda(H_{B_\varrho}^\pm)^{-1}(t)$. Next, notice that if $B_\varrho = B_\varrho(x_0) \Subset \Omega$, then the function $(x_0, \varrho, t) \mapsto H_{B_\varrho(x_0)}^\pm(t)$ is continuous on $\Omega \times [0, \infty) \times [0, \infty)$. This easily follows from the Hölder continuity of $a(\cdot)$. Finally, for x_0 fixed, if $\varrho_1 \leq \varrho_2$, then $H_{B_{\varrho_1}(x_0)}^-(t) \geq H_{B_{\varrho_2}(x_0)}^-(t)$, $(H_{B_{\varrho_1}(x_0)}^+)^{-1}(t) \leq (H_{B_{\varrho_2}(x_0)}^+)^{-1}(t)$ holds uniformly in $t \geq 0$, and, as a consequence, $(H_{B_{\varrho_1}(x_0)}^-)^{-1}(t) \leq (H_{B_{\varrho_2}(x_0)}^-)^{-1}(t)$, (resp. $(H_{B_{\varrho_1}(x_0)}^+)^{-1}(t) \geq (H_{B_{\varrho_2}(x_0)}^+)^{-1}(t)$), holds too for all $t \geq 0$.

We start with a classical lemma (see [2] for a description and references), which is concerned with some properties of Maximal operators with respect to the so called gradient truncation. We recall that the Hardy-Littlewood maximal operator is defined as follows

$$\mathcal{M}(f)(x) := \sup_{B_\varrho(x) \subset \mathbb{R}^n} \int_{B_\varrho(x)} |f(y)| dy, \quad x \in \mathbb{R}^n,$$

whenever $f \in L_{\text{loc}}^1(\mathbb{R}^n)$.

Lemma 6 Let $B_\varrho \subset \mathbb{R}^n$ be a ball and $w \in W_0^{1,1}(B_\varrho, \mathbb{R}^N)$ (trivially extended by zero outside B_ϱ). Then for any $\lambda > 0$ there exists $w_\lambda \in W_0^{1,\infty}(B_\varrho, \mathbb{R}^N)$ such that

$$\|Dw_\lambda\|_{L^\infty(B_\varrho, \mathbb{R}^{N \times n})} \leq c\lambda, \quad (4.3)$$

for some positive constant $c \equiv c(n, N)$. Moreover, it holds that

$$B_\varrho \cap \{w \neq w_\lambda\} \subset (B_\varrho \cap \{\mathcal{M}(|Dw|) > \lambda\}) \cup \text{negligible set}. \quad (4.4)$$

We have a first quantitative harmonic approximation type lemma.

Lemma 7 (Intrinsic and quantitative $g(x, \cdot)$ -harmonic approximation) Let $B_r \subset \mathbb{R}^n$ be a ball with radius $r \leq 1$ and such that $B_{2r} \Subset \Omega$, $\varepsilon \in (0, 1)$ and $v \in (W^{1,H} \cap L^\infty)(B_r, \mathbb{R}^N)$, $N \geq 1$, be a function satisfying

$$\int_{B_r} H(x, Dv) dx \leq c_1 H_{B_r}^-\left(\frac{\varepsilon}{r}\right), \quad (4.5)$$

$$\left(\int_{B_{r/2}} [H(x, Dv)]^{1+\delta} dx \right)^{1/(1+\delta)} \leq \tilde{c}_1 \int_{B_r} H(x, Dv) dx, \quad (4.6)$$

and

$$\left| \int_{B_{r/2}} \partial g(x, Dv) \cdot D\varphi dx \right| \leq c_2 \varepsilon^t \int_{B_r} \left[H(x, Dv) + H\left(x, \|D\varphi\|_{L^\infty(B_{r/2})}\right) \right] dx, \quad (4.7)$$

for some $t \in (0, 1]$, $\delta \in (0, 1)$ and all $\varphi \in C_c^\infty(B_{r/2}, \mathbb{R}^N)$, where $g: \Omega \times \mathbb{R}^{N \times n} \rightarrow [0, \infty)$ is of the type in (3.13) under assumptions (1.12)_{1,2,3,4}, and where $c_1, \tilde{c}_1$ and c_2 are fixed constants larger than one. Then there exists $h \in W_v^{1,H}(B_{r/2}, \mathbb{R}^N)$ such that

$$\int_{B_{r/2}} \partial g(x, Dh) \cdot D\varphi dx = 0 \quad \text{for all } \varphi \in W_0^{1,H}(B_{r/2}, \mathbb{R}^N), \quad (4.8)$$

$$\|h\|_{L^\infty(B_{r/2})} \leq \sqrt{N} \|v\|_{L^\infty(B_{r/2})} \quad (4.9)$$

and

$$\int_{B_{r/2}} \left(|V_p(Dv) - V_p(Dh)|^2 + a(x) |V_q(Dv) - V_q(Dh)|^2 \right) dx \leq c \varepsilon^m \int_{B_r} H(x, Dv) dx, \quad (4.10)$$

with $c \equiv c(\text{data}_0)$ and $m = m(\text{data}, \|v\|_{L^\infty(B_r)}, t, \delta)$ (see (4.12) below for the meaning of data_0). Finally, the function $h \in W_v^{1,H}(B_{r/2}, \mathbb{R}^N)$ is the unique solution of the Dirichlet problem

$$h \mapsto \min_{w \in W_v^{1,H}(B_{r/2}, \mathbb{R}^N)} \int_{B_{r/2}} g(x, Dw) dx. \quad (4.11)$$

Remark 11 The assumptions considered in Lemma 7 are tailored to the situations where the Lemma will be applied. In the typical applications, v is a minimizer of a constrained problem as considered in Theorem 1. This means that the condition $v \in (W^{1,H} \cap L^\infty)$ is automatically satisfied. For the same reason, assumption (4.6) is satisfied by Lemma 4. Finally, the smallness condition (4.5) typically occurs when proving partial regularity theorems (see next section). We also wish to point out that the proof we are going to give here allows for further generalizations to cases where instead of the function $H(\cdot)$ one considers more general instances, as for example those in Section 1.2.

Proof. In the following we shall abbreviate, as in (1.16), as follows

$$\text{data}_0 \equiv \text{data}_0(n, N, \nu, L, p, q, \alpha, [a]_{0,\alpha}, \|v\|_{L^\infty(B_r)}, c_1, \tilde{c}_1, c_2) . \quad (4.12)$$

By a standard approximation argument we notice that, if (4.7) holds for every $\varphi \in C_c^\infty(B_{r/2}, \mathbb{R}^N)$, then it also holds for every $\varphi \in W_0^{1,\infty}(B_{r/2}, \mathbb{R}^N)$. Now, let $h \in W_v^{1,H}(B_{r/2}, \mathbb{R}^N)$ be the unique solution to the Dirichlet problem (4.11). This can be obtained as follows. First, notice that solutions are always unique as a consequence of the strict convexity of $z \mapsto g(\cdot, z)$. Then, existence of $h \in W_v^{1,p}(B_{r/2}, \mathbb{R}^N)$ results by Direct Methods of the Calculus of Variations. By minimality and (1.12)₁, there holds

$$\int_{B_{r/2}} H(x, Dh) dx \leq \frac{L}{\nu} \int_{B_{r/2}} H(x, Dv) dx , \quad (4.13)$$

so h satisfies (4.8). Moreover, thanks to the assumptions in (3.13) and (1.14)₁, we can apply the maximum principle in [43, Theorem 2.3] and this yields (4.9). In particular, we conclude with $h \in W_v^{1,H}(B_{r/2}, \mathbb{R}^N)$. Thanks to Remark 6 and (4.6), we can also apply Lemma 5 (with v replaced by h and u replaced by v in the present situation); this, together with (4.13), yield that

$$\begin{aligned} \left(\int_{B_{r/2}} [H(x, Dh)]^{1+\sigma_g} dx \right)^{\frac{1}{1+\sigma_g}} &\leq c \left(\int_{B_{r/2}} [H(x, Dv)]^{1+\sigma_g} dx \right)^{\frac{1}{1+\sigma_g}} \\ &\leq c \left(\int_{B_{r/2}} [H(x, Dv)]^{1+\delta} dx \right)^{\frac{1}{1+\delta}} \stackrel{(4.6)}{\leq} c \int_{B_r} H(x, Dv) dx , \end{aligned} \quad (4.14)$$

for positive constants $c \equiv c(\text{data}, c_1)$ and $\sigma_g \equiv \sigma_g(\text{data}, \|v\|_{L^\infty(B_r)})$, with $c \geq 1$ and $\sigma_g \in (0, \delta)$. This peculiar dependence of the constants is also a consequence of (4.9) (see again Remark 6). In the application of Lemma 5 we are indeed getting rid of the dependence on $\|h\|_{L^\infty}$ by means of (4.9). Now, notice that there is no loss of generality in assuming that $\int_{B_r} H(x, Dv) dx > 0$, otherwise $v \equiv \text{const}$ on B_r and the thesis trivially holds for $h \equiv \text{const}$. From Remark 10, we have that $t \mapsto H_{B_r}^-(t)$ is a bijection, so there is a unique $\lambda > 0$ such that

$$H_{B_r}^-(\lambda) = M \int_{B_r} H(x, Dv) dx , \quad (4.15)$$

holds for some $M \geq 1$ whose size will be fixed later. Set $w = v - h \in W_0^{1,H}(B_{r/2}, \mathbb{R}^N)$ and consider $w_\lambda \in W_0^{1,\infty}(B_{r/2}, \mathbb{R}^N)$ given by Lemma 6, which satisfies (4.3) and (4.4). We deduce that

$$\begin{aligned} \frac{|B_{r/2} \cap \{w \neq w_\lambda\}|}{|B_{r/2}|} &\stackrel{(4.4)}{\leq} \frac{|B_{r/2} \cap \{\mathcal{M}(|Dw|) > \lambda\}|}{|B_{r/2}|} \\ &\stackrel{\text{Chebyshev}}{\leq} \frac{c}{[H_{B_r}^-(\lambda)]^{1+\sigma_g}} \int_{B_{r/2}} [H_{B_r}^-(\mathcal{M}(|Dw|))]^{1+\sigma_g} dx \\ &\stackrel{\text{maximal}}{\leq} \frac{c}{[H_{B_r}^-(\lambda)]^{1+\sigma_g}} \int_{B_{r/2}} [H_{B_r}^-(Dw)]^{1+\sigma_g} dx \\ &\leq \frac{c}{[H_{B_r}^-(\lambda)]^{1+\sigma_g}} \int_{B_{r/2}} [H_{B_r}^-(Dh)]^{1+\sigma_g} + [H_{B_r}^-(Dv)]^{1+\sigma_g} dx \\ &\stackrel{(4.14)}{\leq} \frac{c}{[H_{B_r}^-(\lambda)]^{1+\sigma_g}} \left(\int_{B_r} H(x, Dv) dx \right)^{1+\sigma_g} \stackrel{(4.15)}{\leq} \frac{c}{M^{1+\sigma_g}} , \end{aligned} \quad (4.16)$$

where $c \equiv c(\text{data}_0)$. Now we test the weak formulation of (4.11) against w_λ to get

$$\begin{aligned}\mathcal{T}_1 &:= \int_{B_{r/2}} (\partial g(x, Dv) - \partial g(x, Dh)) \cdot Dw_\lambda \chi_{\{w=w_\lambda\}} dx \\ &= \int_{B_{r/2}} \partial g(x, Dv) \cdot Dw_\lambda dx - \int_{B_{r/2}} (\partial g(x, Dv) - \partial g(x, Dh)) \cdot Dw_\lambda \chi_{\{w \neq w_\lambda\}} dx =: \mathcal{T}_2 + \mathcal{T}_3 .\end{aligned}\quad (4.17)$$

Upon setting (recall the definition in (2.1))

$$\mathcal{V}^2 := |V_p(Dv) - V_p(Dh)|^2 + a(x)|V_q(Dv) - V_q(Dh)|^2 , \quad (4.18)$$

the strict monotonicity (5.4) implies there exists a constant $c \equiv c(n, N, \nu, p, q)$ such that

$$\mathcal{T}_1 \geq \frac{1}{c} \int_{B_{r/2}} \mathcal{V}^2 \chi_{\{w=w_\lambda\}} dx .$$

Let us consider term $\mathcal{T}_2$; for this we start observing

$$\lambda \stackrel{(4.15)}{=} (H_{B_r}^-)^{-1} \left(M \int_{B_r} H(x, Dv) dx \right) \stackrel{(4.2), (4.5)}{\leq} 2c_1 M (H_{B_r}^-)^{-1} \left(H_{B_r}^- \left(\frac{\varepsilon}{r} \right) \right) \leq \frac{2c_1 \varepsilon M}{r} .$$

From this last inequality and (1.15)₁ we can estimate

$$\begin{aligned}H_{B_r}^+(\lambda) &= H_{B_r}^-(\lambda) + [a_s(B_r) - a_i(B_r)] \lambda^q \\ &\leq H_{B_r}^-(\lambda) + c r^{\alpha-(q-p)} (\varepsilon M)^{q-p} \lambda^p \leq c [1 + (\varepsilon M)^{q-p}] H_{B_r}^-(\lambda) ,\end{aligned}\quad (4.19)$$

with $c \equiv c(p, q, \alpha, [a]_{0,\alpha}, c_1)$. Now we have

$$\int_{B_r} H(x, \|Dw_\lambda\|_{L^\infty(B_{r/2})}) dx \stackrel{(4.3)}{\leq} c \int_{B_r} H(x, \lambda) dx \leq c H_{B_r}^+(\lambda) \stackrel{(4.19)}{\leq} c [1 + (\varepsilon M)^{q-p}] H_{B_r}^-(\lambda) , \quad (4.20)$$

where $c \equiv c(n, N, p, q, \alpha, [a]_{0,\alpha}, c_1)$, so that

$$\begin{aligned}|\mathcal{T}_2| &\stackrel{(4.7)}{\leq} c_2 \varepsilon^t \int_{B_r} \left[H(x, Dv) + H(x, \|Dw_\lambda\|_{L^\infty(B_{r/2})}) \right] dx \\ &\stackrel{(4.20)}{\leq} c \varepsilon^t \int_{B_r} H(x, Dv) dx + c \varepsilon^t [1 + (\varepsilon M)^{q-p}] H_{B_r}^-(\lambda) \\ &\stackrel{(4.15)}{\leq} c \varepsilon^t M [1 + (\varepsilon M)^{q-p}] \int_{B_r} H(x, Dv) dx ,\end{aligned}$$

for $c \equiv c(n, N, p, q, \alpha, [a]_{0,\alpha}, c_1, c_2)$. Finally, for $\mathcal{T}_3$, we fix $\kappa \in (0, 1)$ to be chosen later on and estimate as follows:

$$\begin{aligned}|\mathcal{T}_3| &\stackrel{(4.17)}{\leq} \int_{B_{r/2}} (|\partial g(x, Dh)| + |\partial g(x, Dv)|) |Dw_\lambda| \chi_{\{w \neq w_\lambda\}} dx \\ &\stackrel{(1.12)}{\leq} c \int_{B_{r/2}} \left(\frac{H(x, Dh)}{|Dh|} + \frac{H(x, Dv)}{|Dv|} \right) |Dw_\lambda| \chi_{\{w \neq w_\lambda\}} dx \\ &\stackrel{\text{Young}}{\leq} \kappa \int_{B_{r/2}} [H(x, Dh) + H(x, Dv)] dx + \frac{c}{\kappa^{q-1}} \int_{B_{r/2}} H(x, Dw_\lambda) \chi_{\{w \neq w_\lambda\}} dx \\ &\stackrel{(4.13)}{\leq} c \kappa \int_{B_{r/2}} H(x, Dv) dx + \frac{c}{\kappa^{q-1}} \int_{B_{r/2}} H(x, Dw_\lambda) \chi_{\{w \neq w_\lambda\}} dx\end{aligned}$$

$$\begin{aligned}
& \stackrel{(4.3)}{\leq} c\kappa \int_{B_r} H(x, Dv) dx + \frac{c}{\kappa^{q-1}} \frac{|B_{r/2} \cap \{w \neq w_\lambda\}|}{|B_{r/2}|} H_{B_r}^+(\lambda) \\
& \stackrel{(4.16)}{\leq} c\kappa \int_{B_r} H(x, Dv) dx + \frac{c}{\kappa^{q-1} M^{1+\sigma_g}} H_{B_r}^+(\lambda) \\
& \stackrel{(4.19)}{\leq} c\kappa \int_{B_r} H(x, Dv) dx + \frac{c}{\kappa^{q-1} M^{1+\sigma_g}} [1 + (\varepsilon M)^{q-p}] H_{B_r}^-(\lambda) \\
& \stackrel{(4.15)}{\leq} c \left\{ \kappa + \frac{1}{\kappa^{q-1} M^{\sigma_g}} [1 + (\varepsilon M)^{q-p}] \right\} \int_{B_r} H(x, Dv) dx,
\end{aligned}$$

with $c \equiv c(\text{data}_0)$. Collecting the estimates found for $\mathcal{T}_1, \mathcal{T}_2$ and $\mathcal{T}_3$ to (4.17), we get

$$\begin{aligned}
\int_{B_{r/2}} \mathcal{V}^2 \chi_{\{w=w_\lambda\}} dx & \leq c \left\{ \kappa + \varepsilon^t M + \varepsilon^{t+q-p} M^{q-p+1} + \frac{1}{\kappa^{q-1} M^{\sigma_g}} + \frac{(\varepsilon M)^{q-p}}{\kappa^{q-1} M^{\sigma_g}} \right\} \int_{B_r} H(x, Dv) dx \\
& =: cS(\kappa, \varepsilon, M) \int_{B_r} H(x, Dv) dx,
\end{aligned} \tag{4.21}$$

again with $c \equiv c(\text{data}_0)$. Now let $\theta \in (0, 1)$ be a number to be fixed in some lines. From Hölder's inequality, (4.16) and (4.13) we obtain

$$\left(\int_{B_{r/2}} \mathcal{V}^{2\theta} \chi_{\{w \neq w_\lambda\}} dx \right)^{1/\theta} \leq \left(\frac{|B_{r/2} \cap \{w \neq w_\lambda\}|}{|B_{r/2}|} \right)^{\frac{1-\theta}{\theta}} \int_{B_{r/2}} \mathcal{V}^2 dx \leq \frac{c}{M^{(1+\sigma_g)\frac{1-\theta}{\theta}}} \int_{B_r} H(x, Dv) dx$$

for $c \equiv c(\text{data}_0)$ and, again by Hölder's inequality and (4.21),

$$\left(\int_{B_{r/2}} \mathcal{V}^{2\theta} \chi_{\{w=w_\lambda\}} dx \right)^{1/\theta} \leq cS(\kappa, \varepsilon, M) \int_{B_r} H(x, Dv) dx,$$

with $c \equiv c(\text{data}_0)$. Merging the content of the last two displays now gives

$$\left(\int_{B_{r/2}} \mathcal{V}^{2\theta} dx \right)^{1/\theta} \leq c \left\{ S(\kappa, \varepsilon, M) + \frac{1}{M^{(1+\sigma_g)\frac{1-\theta}{\theta}}} \right\} \int_{B_r} H(x, Dv) dx. \tag{4.22}$$

In the above inequality ε is fixed in the statement of the theorem, while $\kappa \in (0, 1)$ and $M \geq 1$ are still free parameters to be chosen arbitrarily. We take

$$M = \frac{1}{\varepsilon^{\frac{t}{2}}} > 1 \quad \text{and} \quad \kappa = \varepsilon^{\frac{\sigma_g t}{4(q-1)}} \in (0, 1)$$

and set

$$\tilde{m} := \frac{t\sigma_g}{4} \min \left\{ 1, \frac{1}{q-1} \right\},$$

so that, recalling the expression of $S(\kappa, \varepsilon, M)$ in (4.21), we find

$$S(\kappa, \varepsilon, M) + \frac{1}{M^{(1+\sigma_g)\frac{1-\theta}{\theta}}} \leq 5\varepsilon^{\tilde{m}} + \varepsilon^{\frac{t(1+\sigma_g)(1-\theta)}{2\theta}} \leq 6\varepsilon^{\tilde{m}(\theta)}$$

where

$$\tilde{m}(\theta) := \min \left\{ \tilde{m}, \frac{t(1+\sigma_g)(1-\theta)}{2\theta} \right\} \equiv \tilde{m}(\sigma_g, t, q, \theta) \equiv \tilde{m}(\text{data}, \|v\|_{L^\infty(B_r)}, t, \delta, \theta), \tag{4.23}$$

and therefore (4.22) reads as

$$\left(\int_{B_{r/2}} \mathcal{V}^{2\theta} dx \right)^{\frac{1}{2\theta}} \leq c \varepsilon^{\frac{\tilde{m}(\theta)}{2}} \left(\int_{B_r} H(x, Dv) dx \right)^{\frac{1}{2}} \quad (4.24)$$

with $c \equiv c(\text{data}_0)$. The final dependence on the various constants of m in (4.23) has been obtained recalling that $\sigma_g \equiv \sigma_g(\text{data}, \|v\|_{L^\infty(B_r)})$; notice also that the dependence upon the initial higher integrability exponent δ appearing in (4.6) comes from the restriction $\sigma_g < \delta$. Next, notice that from the very definition of $\mathcal{V}$ in (4.18), and using (4.14), we readily infer

$$\left(\int_{B_{r/2}} \mathcal{V}^{2(1+\sigma_g)} dx \right)^{\frac{1}{2(1+\sigma_g)}} \leq c \left(\int_{B_r} H(x, Dv) dx \right)^{\frac{1}{2}}, \quad (4.25)$$

again for $c \equiv c(\text{data}_0)$. Next, we choose

$$\theta := \frac{1 + \sigma_g}{(1 + 2\sigma_g)} \equiv \theta(\text{data}, \|v\|_{L^\infty(B_r)}, \delta) \quad (4.26)$$

and apply Hölder's inequality with conjugate exponents $\frac{2(1+\sigma_g)}{1+2\sigma_g}$ and $2(1+\sigma_g)$, to get

$$\int_{B_{r/2}} \mathcal{V}^2 dx \leq \left(\int_{B_{r/2}} \mathcal{V}^{\frac{2(1+\sigma_g)}{1+2\sigma_g}} dx \right)^{\frac{1+2\sigma_g}{2(1+\sigma_g)}} \left(\int_{B_{r/2}} \mathcal{V}^{2(1+\sigma_g)} dx \right)^{\frac{1}{2(1+\sigma_g)}} \stackrel{(4.24), (4.25)}{\leq} c \varepsilon^{\frac{\tilde{m}(\theta)}{2}} \int_{B_r} H(x, Dv) dx,$$

with where $c \equiv c(\text{data}_0)$. This concludes the proof of (4.10), and of Lemma 7, by fixing $m := \tilde{m}(\theta)/2$ and the dependence $m = m(\text{data}, \|v\|_{L^\infty(B_r)}, t, \delta)$ claimed in the statement follows by looking at (4.23) and (4.26). $\square$

We next report another harmonic type approximation lemma of the type already considered in [4]. On the contrary to Lemma 7, this one involves a classical Young function $H_0(\cdot)$, i.e., no dependence on x is considered

$$H_0(t) = t^p + a_0 t^q, \quad a_0 \geq 0. \quad (4.27)$$

This time we shall consider a $C^1(\mathbb{R}^{N \times n}) \cap C^2(\mathbb{R}^{N \times n} \setminus \{0\})$ -regular integrand $g_0: \mathbb{R}^{N \times n} \rightarrow [0, \infty)$ such that

$$g_0(z) = \tilde{g}_0(|z|) \text{ holds for every } z \in \mathbb{R}^{N \times n} \text{ with } t \mapsto \tilde{g}_0(t) \text{ non-decreasing} \quad (4.28)$$

where $\tilde{g}_0: [0, \infty) \rightarrow [0, \infty)$ of class $C^1[0, \infty) \cap C^2(0, \infty)$. We shall consider the following set of assumptions:

$$\begin{cases} \nu H_0(z) \leq g_0(z) \leq L H_0(z) \\ |\partial g_0(z)| |z| + |\partial^2 g_0(z)| |z|^2 \leq L H_0(z) \\ \nu(|z|^{p-2} + a_0 |z|^{q-2}) |\xi|^2 \leq \langle \partial^2 g_0(z) \xi, \xi \rangle, \end{cases} \quad (4.29)$$

considered with the same notation as in (1.12), for suitable numbers $0 < \nu \leq 1 \leq L < +\infty$ (not necessarily the same as appearing in (1.12)). We then have the following approximation lemma, which is a different version of [4, Lemma 1]:

Lemma 8 (Quantitative g_0 -harmonic approximation) *Let $B_r \subset \mathbb{R}^n$ be a ball with radius $r \leq 1$, $\varepsilon \in (0, 1]$ and $v \in W^{1, H_0}(B_r, \mathbb{R}^N)$, $N \geq 1$, be a function satisfying*

$$\left(\int_{B_{r/2}} [H_0(Dv)]^{1+\delta} dx \right)^{1/(1+\delta)} \leq \tilde{c}_1 \int_{B_r} H_0(Dv) dx, \quad (4.30)$$

for some $\delta \in (0, 1)$ and

$$\left| \int_{B_{r/2}} \partial g_0(Dv) \cdot D\varphi \, dx \right| \leq c_2 \varepsilon^t \int_{B_r} \left[H_0(Dv) + H_0(\|D\varphi\|_{L^\infty(B_{r/2})}) \right] dx, \quad (4.31)$$

where $t \in (0, 1]$ and all $\varphi \in C_c^\infty(B_{r/2}, \mathbb{R}^N)$, where $\tilde{c}_1$ and c_2 are absolute constants and under assumptions (4.28)-(4.29). Then there exists $h_0 \in W_v^{1,H_0}(B_{r/2}, \mathbb{R}^N)$ such that

$$\int_{B_{r/2}} \partial g_0(Dh_0) \cdot D\varphi \, dx = 0 \quad \text{for all } \varphi \in W_0^{1,H_0}(B_{r/2}, \mathbb{R}^N), \quad (4.32)$$

$$\|h_0\|_{L^\infty(B_{r/2})} \leq \sqrt{N} \|v\|_{L^\infty(B_{r/2})} \quad (4.33)$$

and

$$\int_{B_{r/2}} \left(|V_p(Dv) - V_p(Dh_0)|^2 + a_0 |V_q(Dv) - V_q(Dh_0)|^2 \right) dx \leq c \varepsilon^m \int_{B_r} H_0(Dv) \, dx,$$

with $c \equiv c(n, N, v, L, p, q, \tilde{c}_1, c_2)$ and $m = m(n, N, v, L, p, q, t, \delta)$. Finally, the function $h_0 \in W_v^{1,H_0}(B_{r/2}, \mathbb{R}^N)$ is the unique solution of the Dirichlet problem

$$h \mapsto \min_{w \in W_v^{1,H_0}(B_{r/2}, \mathbb{R}^N)} \int_{B_{r/2}} g_0(Dh_0) \, dx. \quad (4.34)$$

Proof. The proof is rather close to that of Lemma 8. For this reason we only give a sketch of it. Again, in (4.31) we can consider $\varphi \in W_0^{1,\infty}(B_{r/2}, \mathbb{R}^N)$. This time we define $h_0 \in W^{1,H_0}(B_{r/2}, \mathbb{R}^N)$ as the unique solution of the Dirichlet problem (4.34), so that (4.32)-(4.33) hold. Moreover, (4.29)₁ and minimality yield

$$\int_{B_{r/2}} H_0(Dh_0) \, dx \leq \frac{L}{v} \int_{B_{r/2}} H_0(Dv) \, dx. \quad (4.35)$$

By Lemma 5 (with constant coefficients, see Remark 7), we get, as for (4.14) and using (4.35), that

$$\left(\int_{B_{r/2}} [H_0(Dh_0)]^{1+\sigma_g} \, dx \right)^{\frac{1}{1+\sigma_g}} \leq c \left(\int_{B_{r/2}} [H_0(Dv)]^{1+\sigma_g} \, dx \right)^{\frac{1}{1+\sigma_g}} \stackrel{(4.30)}{\leq} c \int_{B_r} H_0(Dv) \, dx, \quad (4.36)$$

holds for $\sigma_g \equiv \sigma_g(n, N, v, L, p, q) \in (0, \delta)$ and with $c \equiv c(n, N, v, L, p, q, \tilde{c}_1)$. Proceeding as for the proof of Lemma 7, we find $\lambda > 0$ such that

$$H_0(\lambda) = M \int_{B_r} H_0(Dv) \, dx, \quad (4.37)$$

for some $M \geq 1$ to be specified later on. Set $w = v - h_0 \in W_0^{1,H_0}(B_{r/2}, \mathbb{R}^N)$ and consider w_λ given by Lemma 6 matching (4.3)-(4.4). As for the proof of (4.16), but using (4.36) and (4.37), we deduce that

$$\frac{|B_{r/2} \cap \{w \neq w_\lambda\}|}{|B_{r/2}|} \leq \frac{c}{[H_0(\lambda)]^{1+\sigma_g}} \left(\int_{B_r} H_0(Dv) \, dx \right)^{1+\sigma_g} \stackrel{(4.37)}{\leq} \frac{c}{M^{1+\sigma_g}}, \quad (4.38)$$

with $c \equiv c(n, N, v, L, p, q, \tilde{c}_1)$. Now we test (4.32) against w_λ and set

$$\mathcal{T}_1 := \int_{B_{r/2}} (\partial g_0(Dv) - \partial g_0(Dh_0)) \cdot Dw_\lambda \chi_{\{w=w_\lambda\}} \, dx$$

$$= \int_{B_{r/2}} \partial g_0(Dv) \cdot Dw_\lambda dx - \int_{B_{r/2}} (\partial g_0(Dv) - \partial g_0(Dh_0)) \cdot Dw_\lambda \chi_{\{w \neq w_\lambda\}} dx =: \mathcal{T}_2 + \mathcal{T}_3 .$$

This time, as in (4.18), we set $\mathcal{V}_0^2 := |V_p(Dv) - V_p(Dh_0)|^2 + a_0 |V_q(Dv) - V_q(Dh_0)|^2$. By monotonicity of $\partial g_0(\cdot)$ (which is similar to (2.4) for $a(x) \equiv a_0$), there is $c \equiv c(n, N, \nu, p, q)$ such that

$$\mathcal{T}_1 \geq \frac{1}{c} \int_{B_{r/2}} \mathcal{V}_0^2 \chi_{\{w = w_\lambda\}} dx . \quad (4.39)$$

As for $\mathcal{T}_2$, from (4.31), (4.3) and (4.37), we obtain

$$|\mathcal{T}_2| \leq c \varepsilon^t M \int_{B_r} H_0(Dv) dx , \quad (4.40)$$

where $c \equiv c(n, N, L, p, q, c_1)$. Finally, for $\mathcal{T}_3$, we fix $\kappa \in (0, 1)$ to be chosen. Then, by using (4.29)₂, Young's inequality, (4.35), (4.38) and (4.3), and proceeding as in the proof of the analogous term $\mathcal{T}_3$ from Lemma 7, we have

$$\begin{aligned} |\mathcal{T}_3| &\leq c \int_{B_{r/2}} \left(\frac{H_0(Dh_0)}{|Dh_0|} + \frac{H_0(Dv)}{|Dv|} \right) |Dw_\lambda| \chi_{\{w \neq w_\lambda\}} dx \\ &\leq c \kappa \int_{B_{r/2}} H_0(Dv) dx + \frac{c}{\kappa^{q-1}} \int_{B_{r/2}} H_0(Dw_\lambda) \chi_{\{w \neq w_\lambda\}} dx \\ &\leq c \left(\kappa + \frac{1}{\kappa^{q-1} M^{\sigma_g}} \right) \int_{B_r} H_0(Dv) dx , \end{aligned} \quad (4.41)$$

with $c \equiv c(n, N, L, p, q, \tilde{c}_1)$. From estimates (4.39)-(4.41) we get

$$\int_{B_{r/2}} \mathcal{V}_0^2(x) \chi_{\{w = w_\lambda\}} \leq c \left\{ \kappa + \varepsilon^t M + \frac{1}{\kappa^{q-1} M^{\sigma_g}} \right\} \int_{B_r} H_0(Dv) dx ,$$

with $c \equiv c(n, N, \nu, L, p, q, \tilde{c}_1, c_2)$. Starting from the last inequality, the rest of the proof goes exactly as the one for Lemma 7, after (4.21). $\square$

Finally, an elementary Young type inequality.

Lemma 9 *Let $H_0(\cdot)$ be the function defined in (4.27). Then, whenever $\kappa \in (0, 1)$ it holds that*

$$st \leq \kappa H_0(t) + \kappa^{-1/(p-1)} H_0^*(t) , \quad \text{for all } s, t \geq 0 . \quad (4.42)$$

where

$$H_0^*(t) := \sup_{s>0} (st - H_0(s)) , \quad \text{for all } t \geq 0$$

denotes the convex conjugate function to $H_0(\cdot)$.

Proof. Notice that, for $A \geq 1$, as it is $q \geq p$, we have

$$\begin{aligned} H_0^*(At) &= \sup_{s>0} (sAt - H_0(s)) = A^{\frac{p}{p-1}} \sup_{s>0} \left(sA^{-\frac{1}{p-1}} t - A^{-\frac{p}{p-1}} H_0(s) \right) \\ &\leq A^{\frac{p}{p-1}} \sup_{s>0} \left(sA^{-\frac{1}{p-1}} t - H_0(sA^{-\frac{1}{p-1}}) \right) = A^{\frac{p}{p-1}} H_0^*(t) . \end{aligned}$$

Therefore, we find, for $\kappa \in (0, 1)$

$$st \leq H_0(\kappa^{1/p} t) + H_0^*(t/\kappa^{1/p}) \leq \kappa H_0(t) + \kappa^{-1/(p-1)} H_0^*(t)$$

that is, (4.42). $\square$

5 Proof of Theorem 1

In the following, $u \in W_{\text{loc}}^{1,H}(\Omega, \mathbb{S}^{N-1})$ is as in the statement of Theorem 1. We start recalling Lemma 4, according to which there exists $\delta_g \equiv \delta_g(\text{data}) > 0$ such that $H(\cdot, Du) \in L_{\text{loc}}^{1+\delta_g}(\Omega)$ holds, i.e., (1.17) is proved. For the proof of Theorem 1, we first treat the case when $p(1 + \delta_g) \leq n$, and then we describe how to get the result in the remaining one $p(1 + \delta_g) > n$. The proof now goes in six steps. The first three are devoted to the proof of the partial Hölder continuity of a constrained local minimizer of (1.1); in particular, in the third step we describe the regular and the singular sets. In the fourth step we exploit this continuity to move to a single chart. Step five is devoted to show partial Hölder continuity for the gradient in the regular set. In the final step we briefly mention how to treat the case $p(1 + \delta_g) > n$.

Step 1: Freezing.

Let $B_r = B_r(x_0)$ be any ball such that $B_{2r} \Subset \Omega$ and $r \leq 1/2$; more in general, every ball B considered in the rest of the proof will have radius $r(B) \leq 1/2$. We assume that the smallness condition

$$\int_{B_{2r}(x_0)} H(x, Du) dx < H_{B_{2r}(x_0)}^-\left(\frac{\varepsilon}{2r}\right), \quad (5.1)$$

holds for some $\varepsilon \in (0, 1)$ which is going to be chosen in due course of the proof. Let $v \in W_u^{1,H}(B_r, \mathbb{S}^{N-1})$ be a solution to the frozen Dirichlet problem

$$v \mapsto \min_{w \in W_u^{1,H}(B_r, \mathbb{S}^{N-1})} \int_{B_r} F(x, (u)_{B_r}, Dw) dx.$$

This functional satisfies the same growth assumptions (in particular (1.12)₁) of the original one minimized by u and therefore Lemma 4 applies, giving

$$\left(\int_{B_{r/2}} [H(x, Dv)]^{1+\delta_g} dx \right)^{1/(1+\delta_g)} \leq \tilde{c}_1 \int_{B_r} H(x, Dv) dx, \quad (5.2)$$

where the exponent $\delta_g \equiv \delta_g(\text{data}) > 0$ is the same one appearing in (3.10) and $\tilde{c}_1 \equiv \tilde{c}_1(\text{data})$. Taking into account the content of Section 3.2, and in particular (3.14), v solves the Euler-Lagrange equation

$$\int_{B_r} \partial F(x, (u)_{B_r}, Dv) \cdot D\varphi dx = \int_{B_r} \tilde{F}'(x, (u)_{B_r}, Dv) |Dv| (v \cdot \varphi) dx, \quad (5.3)$$

which is valid for any $\varphi \in (W_0^{1,H} \cap L^\infty)(B_r, \mathbb{R}^N)$. Moreover, (2.5) becomes

$$\begin{aligned} & |V_p(z_2) - V_p(z_1)|^2 + a(x)|V_q(z_2) - V_q(z_1)|^2 + \partial F(x, (u)_{B_r}, z_1) \cdot (z_2 - z_1) \\ & \leq c \left[F(x, (u)_{B_r}, z_2) - F(x, (u)_{B_r}, z_1) \right] \end{aligned} \quad (5.4)$$

which holds for any choice of $z_1, z_2 \in \mathbb{R}^{N \times n}$ and $x \in \Omega$, for a constant $c \equiv c(n, N, \nu, p, q)$, see for instance [4, (90)]. The map $w = u - v \in (W_0^{1,H} \cap L^\infty)(B_r, \mathbb{R}^N)$ is an admissible test function in (5.3), therefore we have

$$\begin{aligned} & \int_{B_r} (|V_p(Du) - V_p(Dv)|^2 + a(x)|V_q(Du) - V_q(Dv)|^2) dx \\ & \stackrel{(5.4)}{\leq} c \int_{B_r} \left[F(x, (u)_{B_r}, Du) - F(x, (u)_{B_r}, Dv) \right] dx - \int_{B_r} \partial F(x, (u)_{B_r}, Dv) \cdot (Du - Dv) dx \\ & \stackrel{(5.3)}{=} c \int_{B_r} \left[F(x, (u)_{B_r}, Du) - F(x, (u)_{B_r}, Dv) \right] dx - \int_{B_r} \tilde{F}'(x, (u)_{B_r}, Dv) |Dv| v \cdot (u - v) dx \end{aligned}$$

$$\begin{aligned}
&= c \int_{B_r} \left[F(x, (u)_{B_r}, Du) - F(x, u, Du) \right] dx + c \int_{B_r} \left[F(x, u, Du) - F(x, v, Dv) \right] dx \\
&\quad + c \int_{B_r} \left[F(x, v, Dv) - F(x, (v)_{B_r}, Dv) \right] dx + c \int_{B_r} \left[F(x, (v)_{B_r}, Dv) - F(x, (u)_{B_r}, Dv) \right] dx \\
&\quad - \int_{B_r} \tilde{F}'(x, (u)_{B_r}, Dv) |Dv|v \cdot (u - v) dx \\
&=: \text{(I)} + \text{(II)} + \text{(III)} + \text{(IV)} + \text{(V)}, \tag{5.5}
\end{aligned}$$

where $c \equiv c(n, N, \nu, p, q)$. Before starting working on terms (I)-(V) in (5.5), let us estimate some quantities which will be recurrent in the forthcoming computations. First, notice that the minimality of ν and (1.12)₁ yield

$$\nu \int_{B_r} H(x, Dv) dx \leq \int_{B_r} F(x, (u)_{B_r}, Dv) dx \leq \int_{B_r} F(x, (u)_{B_r}, Du) dx \leq L \int_{B_r} H(x, Du) dx. \tag{5.6}$$

Lemma 5 gives

$$\left(\int_{B_r} [H(x, Dv)]^{1+\sigma_g} dx \right)^{1/(1+\sigma_g)} \leq c \left(\int_{B_r} [H(x, Du)]^{1+\sigma_g} dx \right)^{1/(1+\sigma_g)} \stackrel{(4.6)}{\leq} c \int_{B_{2r}} H(x, Du) dx, \tag{5.7}$$

where $0 < \sigma_g \equiv \sigma_g(\text{data}) < \delta_g$. We are next going to use the function $H_{B_{2r}}^-(t) := t^p + a_1(B_{2r})t^q$. By Jensen's inequality (recall that $\omega(\cdot)$ in (1.13) is a concave function, while $H_{B_{2r}}^-(\cdot)$ is convex), Remark 10 and the smallness condition (5.1), we get

$$\begin{aligned}
\int_{B_r} \omega(|u - (u)_{B_r}|) dx &\leq \omega \left(r \int_{B_r} \left| \frac{u - (u)_{B_r}}{r} \right| dx \right) \\
&= \omega \left[r (H_{B_{2r}}^-)^{-1} \circ (H_{B_{2r}}^-) \left(\int_{B_r} \left| \frac{u - (u)_{B_r}}{r} \right| dx \right) \right] \\
&\leq \omega \left[r (H_{B_{2r}}^-)^{-1} \left(\int_{B_r} H_{B_{2r}}^- \left(\frac{u - (u)_{B_r}}{r} \right) dx \right) \right] \\
&\leq \omega \left[r (H_{B_{2r}}^-)^{-1} \left(\int_{B_r} H \left(x, \frac{u - (u)_{B_r}}{r} \right) dx \right) \right] \\
&\stackrel{(3.9)}{\leq} c \omega \left[r (H_{B_{2r}}^-)^{-1} \left(c \int_{B_{2r}} H(x, Du) dx \right) \right] \\
&\stackrel{(4.2), (5.1)}{\leq} c \omega \left[r (H_{B_{2r}}^-)^{-1} \circ H_{B_{2r}}^- \left(\frac{\varepsilon}{2r} \right) \right] \leq c \varepsilon^\beta, \tag{5.8}
\end{aligned}$$

with $c \equiv c(\text{data}, \beta)$. Similarly, we have

$$\begin{aligned}
\int_{B_r} \omega(|v - (v)_{B_r}|) dx &\leq c \omega \left[r (H_{B_{2r}}^-)^{-1} \left(\int_{B_r} H(x, Dv) dx \right) \right] \\
&\stackrel{(5.6)}{\leq} c \omega \left[r (H_{B_{2r}}^-)^{-1} \left(\int_{B_{2r}} H(x, Du) dx \right) \right] \stackrel{(5.8)}{\leq} c \varepsilon^\beta, \tag{5.9}
\end{aligned}$$

with $c \equiv c(\text{data}, \beta)$. In a totally similar way, in particular again using Lemma 3 and repeatedly the content of Remark 10, we get

$$\omega(|(u)_{B_r} - (v)_{B_r}|) \leq \omega \left(\int_{B_r} |u - v| dx \right)$$

$$\begin{aligned}
&\leq \omega \left[r \left(H_{B_{2r}}^- \right)^{-1} \circ \left(H_{B_{2r}}^- \right) \left(\int_{B_r} \left| \frac{u-v}{r} \right| dx \right) \right] \\
&\leq c\omega \left[r \left(H_{B_{2r}}^- \right)^{-1} \left(\int_{B_r} H(x, Du - Dv) dx \right) \right] \\
&\leq c\omega \left[r \left(H_{B_{2r}}^- \right)^{-1} \left(\int_{B_r} [H(x, Du) + H(x, Dv)] dx \right) \right] \\
&\stackrel{(5.6)}{\leq} c\omega \left[r \left(H_{B_{2r}}^- \right)^{-1} \left(\int_{B_{2r}} H(x, Du) dx \right) \right] \stackrel{(5.8)}{\leq} c\varepsilon^\beta, \tag{5.10}
\end{aligned}$$

again with $c \equiv c(\text{data}, \beta)$. We can now start estimating the terms (I)-(V) in (5.5); we have

$$\begin{aligned}
|\text{(I)}| &\leq c \int_{B_r} \omega(|u - (u)_{B_r}|) H(x, Du) dx \\
&\leq c \left(\int_{B_r} \omega(|u - (u)_{B_r}|)^{\frac{1+\delta_g}{\delta_g}} dx \right)^{\frac{\delta_g}{1+\delta_g}} \left(\int_{B_r} [H(x, Du)]^{1+\delta_g} dx \right)^{\frac{1}{1+\delta_g}} \\
&\stackrel{(3.10)}{\leq} c \left(\int_{B_r} \omega(|u - (u)_{B_r}|) dx \right)^{\frac{\delta_g}{1+\delta_g}} \left(\int_{B_{2r}} H(x, Du) dx \right) \\
&\stackrel{(5.8)}{\leq} c(\text{data}, \beta) \varepsilon^{\frac{\beta\delta_g}{1+\delta_g}} \int_{B_{2r}} H(x, Du) dx. \tag{5.11}
\end{aligned}$$

By minimality we see that

$$\int_{B_r} F(x, u, Du) dx \leq \int_{B_r} F(x, v, Dv) dx \implies |\text{(II)}| \leq 0. \tag{5.12}$$

As for (III), we have

$$\begin{aligned}
|\text{(III)}| &\leq c \int_{B_r} \omega(|v - (v)_{B_r}|) H(x, Dv) dx \\
&\leq c \left(\int_{B_r} \omega(|v - (v)_{B_r}|)^{\frac{1+\sigma_g}{\sigma_g}} dx \right)^{\frac{\sigma_g}{1+\sigma_g}} \left(\int_{B_r} [H(x, Dv)]^{1+\sigma_g} dx \right)^{\frac{1}{1+\sigma_g}} \\
&\stackrel{(5.7)}{\leq} c \left(\int_{B_r} \omega(|v - (v)_{B_r}|) dx \right)^{\frac{\sigma_g}{1+\sigma_g}} \int_{B_{2r}} H(x, Du) dx \\
&\stackrel{(5.9)}{\leq} c(\text{data}, \beta) \varepsilon^{\frac{\beta\sigma_g}{1+\sigma_g}} \int_{B_{2r}} H(x, Du) dx. \tag{5.13}
\end{aligned}$$

The estimation of (IV) is analogous to that of (III), the only difference being that in this case we must use (5.10); we end up with

$$|\text{(IV)}| \leq c(\text{data}, \beta) \varepsilon^{\frac{\beta\sigma_g}{1+\sigma_g}} \int_{B_{2r}} H(x, Du) dx. \tag{5.14}$$

Finally we look at term (V). Proceeding as for the previous terms, and in particular using the smallness condition (5.1) as done in the last line of display (5.8), we have

$$|\text{(V)}| \leq c \int_{B_r} H(x, Dv) |u - v| dx$$

$$\begin{aligned}
&\leq c \left(\int_{B_r} |u - v|^{\frac{1+\sigma_g}{\sigma_g}} dx \right)^{\frac{\sigma_g}{1+\sigma_g}} \left(\int_{B_r} [H(x, Dv)]^{1+\sigma_g} dx \right)^{\frac{1}{1+\sigma_g}} \\
&\leq c \left(\int_{B_r} |u - v| dx \right)^{\frac{\sigma_g}{1+\sigma_g}} \left(\int_{B_r} [H(x, Dv)]^{1+\sigma_g} dx \right)^{\frac{1}{1+\sigma_g}} \\
&\leq c \left[r (H_{B_{2r}}^-)^{-1} \left(\int_{B_r} H_{B_{2r}} \left(\frac{u - v}{r} \right) dx \right) \right]^{\frac{\sigma_g}{1+\sigma_g}} \int_{B_{2r}} H(x, Du) dx \\
&\leq c \left[r (H_{B_{2r}}^-)^{-1} \left(\int_{B_r} H \left(x, \frac{u - v}{r} \right) dx \right) \right]^{\frac{\sigma_g}{1+\sigma_g}} \int_{B_{2r}} H(x, Du) dx \\
&\leq c \left[r (H_{B_{2r}}^-)^{-1} \left(\int_{B_r} H(x, Du - Dv) dx \right) \right]^{\frac{\sigma_g}{1+\sigma_g}} \int_{B_{2r}} H(x, Du) dx \\
&\leq c \left[r (H_{B_{2r}}^-)^{-1} \left(\int_{B_{2r}} H(x, Du) dx \right) \right]^{\frac{\sigma_g}{1+\sigma_g}} \int_{B_{2r}} H(x, Du) dx \\
&\leq c(\text{data}, \beta) \varepsilon^{\frac{\sigma_g}{1+\sigma_g}} \int_{B_{2r}} H(x, Du) dx .
\end{aligned} \tag{5.15}$$

Connecting estimates (5.11)-(5.15) to (5.5), and recalling that $\varepsilon < 1, \beta \leq 1$ and $\sigma_g < \delta_g$, we conclude with

$$\int_{B_r} (|V_p(Du) - V_p(Dv)|^2 + a(x)|V_q(Du) - V_q(Dv)|^2) dx \leq c \varepsilon^{\frac{\beta \sigma_g}{1+\sigma_g}} \int_{B_{2r}} H(x, Du) dx , \tag{5.16}$$

holds for $c \equiv c(\text{data}, \beta)$.

Step 2: $\partial F(\cdot, (u)_{B_r}, \cdot)$ -harmonic approximation.

We aim to show that v matches the assumptions of Lemma 7, with the choice $g(x, z) \equiv \partial F(x, (u)_{B_r}, z)$; obviously, (1.12)_{1,2,3,4} are satisfied, as well as (1.14)₁, by the very definition of $g(\cdot)$. As for (4.5), we have

$$\begin{aligned}
E(v; B_r) &:= \int_{B_r} H(x, Dv) dx \stackrel{(5.6)}{\leq} \frac{L}{\nu} \int_{B_r} H(x, Du) dx \\
&=: \frac{L}{\nu} E(u; B_r) \stackrel{(5.1)}{<} 2^n \frac{L}{\nu} H_{B_{2r}}^- \left(\frac{\varepsilon}{2r} \right) < c_1 H_{B_r}^- \left(\frac{\varepsilon}{r} \right) ,
\end{aligned} \tag{5.17}$$

which is in fact (4.5) with $c_1 := 2^n L/\nu$. On the other hand, the validity of (4.6) is stated in (5.2). To verify (4.7), we look at the Euler-Lagrange equation (3.14) solved by v on $B_{r/2}$. By (1.12)₁, for any $\varphi \in W_0^{1,\infty}(B_{r/2}, \mathbb{R}^N)$ we have

$$\begin{aligned}
\left| \int_{B_{r/2}} \partial F(x, (u)_{B_r}, Dv) \cdot D\varphi dx \right| &= \left| \int_{B_{r/2}} \tilde{F}'(x, (u)_{B_r}, Dv) |Dv| (v \cdot \varphi) dx \right| \\
&\leq c \int_{B_{r/2}} H(x, Dv) |\varphi| dx \\
&\leq cr \|D\varphi\|_{L^\infty(B_{r/2})} \int_{B_{r/2}} H(x, Dv) dx \\
&\leq cr \|D\varphi\|_{L^\infty(B_{r/2})} E(v; B_r)
\end{aligned} \tag{5.18}$$

with $c \equiv c(n, L, p, q)$. The last term in display (5.18) can be estimated via (4.42)

$$cr \|D\varphi\|_{L^\infty(B_{r/2})} E(v; B_r) \leq \delta_1 H_{B_r}^- \left(\|D\varphi\|_{L^\infty(B_{r/2})} \right) + \frac{c}{\delta_1^{1/(p-1)}} \left(H_{B_r}^- \right)^* (r E(v; B_r)) , \tag{5.19}$$

with $c \equiv c(n, L, p, q)$ and $\delta_1 \in (0, 1)$, where $(H_{B_r}^-)^*$ denotes the convex conjugate of $H_{B_r}^-$. Since

$$\left((H_{B_r}^-)^*\right)' = \left((H_{B_r}^-)'\right)^{-1}, \quad (5.20)$$

then, for

$$\varepsilon < \frac{p}{c_1} = \frac{pv}{2^n L}, \quad (5.21)$$

(recall that ε in (5.1) is chosen in due course of the proof via various size restrictions), we find

$$\begin{aligned} \left((H_{B_r}^-)^*\right)'(rE(v; B_r)) &\stackrel{(5.17)}{\leq} \left((H_{B_r}^-)^*\right)' \left(c_1 r H_{B_r}^- \left(\frac{\varepsilon}{r}\right)\right) \leq \left((H_{B_r}^-)^*\right)' \left(\frac{c_1}{p} \varepsilon (H_{B_r}^-)' \left(\frac{\varepsilon}{r}\right)\right) \\ &\stackrel{(5.21)}{\leq} \left((H_{B_r}^-)^*\right)' \left((H_{B_r}^-)' \left(\frac{\varepsilon}{r}\right)\right) \stackrel{(5.20)}{=} \left((H_{B_r}^-)'\right)^{-1} \left((H_{B_r}^-)' \left(\frac{\varepsilon}{r}\right)\right) = \frac{\varepsilon}{r}, \end{aligned}$$

so we can conclude with

$$(H_{B_r}^-)^*(rE(v; B_r)) \leq rE(v; B_r) \left((H_{B_r}^-)^*\right)'(rE(v; B_r)) \leq \varepsilon E(v; B_r).$$

In this way (5.19) becomes

$$cr \|D\varphi\|_{L^\infty(B_{r/2})} E(v; B_r) \leq \delta_1 H_{B_r}^- \left(\|D\varphi\|_{L^\infty(B_{r/2})}\right) + \frac{c\varepsilon}{\delta_1^{1/(p-1)}} \int_{B_r} H(x, Dv) dx.$$

Now select $\delta_1 = \varepsilon^{(p-1)/2}$, and define $2t := \min\{p-1, 1\}$. The last inequality used in (5.18) gives

$$\left| \int_{B_{r/2}} \partial F(x, (u)_{B_r}, Dv) \cdot D\varphi dx \right| \leq c_2 \varepsilon^t \int_{B_r} \left[H(x, Dv) + H(x, \|D\varphi\|_{L^\infty(B_{r/2})}) \right] dx,$$

for some $c_2 = c(n, L, p, q)$ which is in fact (4.7). So Lemma 7 applies and yields a $\partial F(\cdot, (u)_{B_r}, \cdot)$ -harmonic map $h \in W_v^{1,H}(B_{r/2}, \mathbb{R}^N)$, specifically, a solution to

$$h \mapsto \min_{w \in W_v^{1,H}(B_{r/2}, \mathbb{R}^N)} \int_{B_{r/2}} F(x, (u)_{B_r}, Dw) dx, \quad (5.22)$$

such that (4.10) holds; this, together with (5.6), allows to get

$$\int_{B_{r/2}} \left(|V_p(Dv) - V_p(Dh)|^2 + a(x) |V_q(Dv) - V_q(Dh)|^2 \right) dx \leq c\varepsilon^m \int_{B_r} H(x, Du) dx, \quad (5.23)$$

where $c \equiv c(\text{data}, \beta)$ and $m = m(\text{data})$. Moreover, there holds that

$$\|h\|_{L^\infty(B_{r/2})} \leq \sqrt{N}. \quad (5.24)$$

By virtue of (5.22) and of the previous inequality, we are then able to apply Theorem 4. For every $\sigma \in (0, n]$, estimate (3.21) reads as

$$\int_{B_t} H(x, Dh) dx \leq c \left(\frac{t}{s}\right)^{-\sigma} \int_{B_s} H(x, Dh) dx, \quad (5.25)$$

that holds whenever $B_t \subset B_s \subset B_{r/2}$ are concentric balls, and where $c \equiv c(\text{data}, \beta, \sigma)$, again by virtue of (5.24); in the following we take $\sigma < 1/4$. With $\tau \in (0, 1/2)$, recalling (2.2) and using (5.16) and (5.23), we can then estimate

$$\int_{B_{2\tau r}} H(x, Du) dx \leq c \int_{B_{2\tau r}} \left(|V_p(Du) - V_p(Dv)|^2 + a(x) |V_q(Du) - V_q(Dv)|^2 \right) dx$$

$$\begin{aligned}
& +c \int_{B_{2\tau r}} (|V_p(Dv) - V_p(Dh)|^2 + a(x)|V_q(Dv) - V_q(Dh)|^2) dx \\
& +c \int_{B_{2\tau r}} H(x, Dh) dx \\
\leq & c\tau^{-n} \int_{B_r} (|V_p(Du) - V_p(Dv)|^2 + a(x)|V_q(Du) - V_q(Dv)|^2) dx \\
& +c\tau^{-n} \int_{B_{r/2}} (|V_p(Dv) - V_p(Dh)|^2 + a(x)|V_q(Dv) - V_q(Dh)|^2) dx \\
& +c \int_{B_{2\tau r}} H(x, Dh) dx \\
\leq & c \left(\tau^{-n} \varepsilon^{\frac{\beta\sigma_g}{1+\sigma_g}} + \tau^{-n} \varepsilon^m + \tau^{-\sigma} \right) \int_{B_{2r}} H(x, Du) dx ,
\end{aligned}$$

where $c \equiv c(\text{data}, \beta, \sigma)$. Recalling the notation adopted in (5.17), the conclusion of the last display reads as

$$\begin{aligned}
E(u; B_{2\tau r}) & \leq c(\text{data}, \beta, \sigma) \left(\tau^{-n} \varepsilon^{\frac{\beta\sigma_g}{1+\sigma_g}} + \tau^{-n} \varepsilon^m + \tau^{-\sigma} \right) E(u; B_{2r}) \\
& = c(\text{data}, \beta, \sigma) \tau^\sigma \left(\tau^{-n-\sigma} \varepsilon^{\frac{\beta\sigma_g}{1+\sigma_g}} + \tau^{-n-\sigma} \varepsilon^m + \tau^{-2\sigma} \right) E(u; B_{2r}) .
\end{aligned} \tag{5.26}$$

We can now determine $\tau \equiv \tau(\text{data}, \beta, \sigma) \in (0, 1)$ such that

$$c(\text{data}, \sigma) \tau^\sigma \leq \frac{1}{6} \tag{5.27}$$

It is now time to choose the number ε coming from (5.1). Recalling (5.21), we now further reduce ε to have

$$\varepsilon < \min \left\{ \frac{p\nu}{2^n L}, \tau^{\frac{(n-\sigma)(1+\sigma_g)}{\beta\sigma_g}}, \tau^{\frac{n-\sigma}{m}} \right\} \tag{5.28}$$

and notice that this fixes the dependence $\varepsilon \equiv \varepsilon(\text{data}, \beta, \sigma)$. By using (5.27) and (5.28) in (5.26), this last inequality reads as

$$E(u; B_{2\tau r}) \leq \tau^{-2\sigma} E(u; B_{2r}) \tag{5.29}$$

that is, recalling the definition in (5.17)

$$\int_{B_{2\tau r}} H(x, Du) dx \leq \tau^{n-2\sigma} \int_{B_{2r}} H(x, Du) dx . \tag{5.30}$$

Next, we observe that

$$E(u; B_{2\tau r}) \stackrel{(5.29)}{\leq} \tau^{-2\sigma} E(u; B_{2r}) \stackrel{(5.1)}{<} \tau^{1-2\sigma} \tau^{-1} H_{B_{2r}}^- \left(\frac{\varepsilon}{2r} \right) \leq \tau^{-1} H_{B_{2\tau r}}^- \left(\frac{\varepsilon}{2r} \right) \leq H_{B_{2\tau r}}^- \left(\frac{\varepsilon}{2\tau r} \right) ,$$

and we conclude with

$$E(u; B_{2\tau r}) < H_{B_{2\tau r}}^- \left(\frac{\varepsilon}{2\tau r} \right) .$$

We have therefore proved that, for the choice of $\tau \equiv \tau(\text{data}, \beta, \sigma)$ and $\varepsilon \equiv \varepsilon(\text{data}, \beta, \sigma)$ made in (5.27) and (5.28), respectively, if the smallness condition (5.1) is satisfied on the ball B_{2r} it is also satisfied on the ball $B_{2\tau r}$. We can therefore repeat the whole argument developed after (5.1) starting from the ball $B_{2\tau r}$ instead of B_{2r} , thereby arriving at the analog of (5.29), that is $E(u; B_{2\tau^2 r}) \leq \tau^{-2\sigma} E(u; B_{2\tau r})$. This argument can

obviously be iterated on the family of shrinking balls $\{B_{\tau^j r}\}$, thereby concluding that, for every $j \in \mathbb{N}$, it holds that

$$E(u; B_{2\tau^j r}) < H_{B_{2\tau^j r}}^-\left(\frac{\varepsilon}{2\tau^j r}\right)$$

and

$$\int_{B_{2\tau^j r}} H(x, Du) dx < \tau^{(n-2\sigma)j} \int_{B_{2r}} H(x, Du) dx .$$

In turn, a standard interpolation argument leads to conclude that

$$\int_{B_t(x_0)} H(x, Du) dx \leq c \left(\frac{t}{r}\right)^{n-2\sigma} \int_{B_{2r}(x_0)} H(x, Du) dx , \quad \forall t \leq 2r , \quad (5.31)$$

where $c \equiv c(\text{data}, \beta, \sigma)$. Notice that the above inequality has been derived for $4\sigma < 1$ but it is then easily seen to hold whenever $\sigma \in (0, 1)$. Going back to (5.1), we observe that the two functions

$$x_0 \mapsto \int_{B_{2r}(x_0)} H(x, Du) dx \quad \text{and} \quad x_0 \mapsto H_{B_{2r}(x_0)}^-\left(\frac{\varepsilon}{2r}\right)$$

are continuous. This is a consequence of the absolute continuity of the integral for the former, and of Remark 10 for the latter. We conclude that, with σ being fixed, if (5.1) is satisfied at a point $x_0 \in \Omega$, then there exists ball $B_{4r_{x_0}}(x_0)$ such that

$$y \in B_{4r_{x_0}}(x_0) \implies \int_{B_{2r}(y)} H(x, Du) dx < H_{B_{2r}(y)}^-\left(\frac{\varepsilon}{2r}\right) . \quad (5.32)$$

We then conclude that (5.31) holds (with y replacing x_0), and with the same constant $c \equiv c(\text{data}, \beta, \sigma)$, whenever $y \in B_{4r_{x_0}}(x_0)$. By a standard characterization of Hölder continuity it then follows that $u \in C^{0,\gamma}(B_{r_{x_0}}(x_0))$ with $\gamma = 1 - 2\sigma/p$ (see Remark 12 below). As we can choose $\sigma \in (0, 1/4)$ arbitrarily, we have finally proved the following (we can switch from $2r$ to r now):

Proposition 2 *Let $u \in W_{\text{loc}}^{1,1}(\Omega, \mathbb{S}^{N-1})$ be a constrained local minimizer of the functional $\mathcal{F}$ in (1.1), under the assumptions (1.12)-(1.15). Assume that $p(1 + \delta_g) \leq n$. Then, for every positive exponent $\gamma < 1$, there exists another positive number $\varepsilon_\gamma \equiv \varepsilon_\gamma(\text{data}, \beta, \gamma)$ such that if $B_r(x_0) \Subset \Omega$, $r \leq 1$, and*

$$\left[H_{B_r(x_0)}^-\left(\frac{\varepsilon_\gamma}{r}\right) \right]^{-1} \int_{B_r(x_0)} H(x, Du) dx < 1 , \quad (5.33)$$

then u is of class $C^{0,\gamma}$ in a neighbourhood of x_0 .

Remark 12 Let us make the last argument somehow more quantitative. With $\gamma \in (0, 1)$ being fixed, let $\tilde{r}(x_0)$ be the largest radius, such that the smallness condition (5.33) is satisfied with $r \equiv \tilde{r}(x_0)$ (together with $B_{2\tilde{r}(x_0)} \Subset \Omega$). Then we have that

$$\begin{aligned} \int_{B_t(x_0)} H(x, Du) dx &\leq c \left(\frac{t}{\tilde{r}(x_0)}\right)^{n-p+p\gamma} \int_{B_{2\tilde{r}(x_0)}(x_0)} H(x, Du) dx \\ &\stackrel{(3.7)}{\leq} ct^{n-p+p\gamma} \left[[\tilde{r}(x_0)]^{p-p\gamma} \int_{B_{4\tilde{r}(x_0)}(x_0)} \left(\frac{|u|^p}{[\tilde{r}(x_0)]^p} + \|a\|_{L^\infty} \frac{|u|^q}{[\tilde{r}(x_0)]^q} \right) dx \right] . \end{aligned}$$

By using Poincaré's inequality, we get

$$\int_{B_r(x_0)} |u - (u)_{B_r}|^p dx \leq \frac{c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma)}{[\tilde{r}(x_0)]^{q-p+p\gamma}} t^{p\gamma}.$$

This estimate remains stable whenever x_0 is replaced by $y \in B_{4r_{x_0}}(x_0)$ as in (5.32) and therefore, by a standard integral characterization of Hölder continuity, for all $\gamma \in (0, 1)$ it follows that

$$[u]_{0,\gamma;B_{r_{x_0}}(x_0)} \leq \frac{c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma)}{[\tilde{r}(x_0)]^{q/p-1+\gamma}}. \quad (5.34)$$

Step 3: Dimension of the singular set; the first estimate and proof of (1.20).

Following a standard terminology, we denote by

$$\Omega_u := \left\{ x_0 \in \Omega : \text{there exists a ball } B_{r_{x_0}}(x_0) \subset \Omega \text{ such that } u \in C^{0,\gamma}(B_{r_{x_0}}(x_0)) \text{ for some } \gamma \in (0, 1) \right\}.$$

This set is open by definition and we denote $\Sigma_u := \Omega \setminus \Omega_u$, the so-called singular set. Indeed, we shall later on prove that Du is locally Hölder continuous on Ω_u and this justifies the terminology used here with respect to the one in the statement of Theorem 1. Let us now prove (1.20); call Σ the set in the right-hand side of (1.20). The inclusion $\Sigma_u \subset \Sigma$ is obvious in view of Proposition 2. On the other hand, take, by contradiction $x_0 \in \Sigma \setminus \Sigma_u$. Then there exists $\gamma \in (0, 1)$ such that u is of class $C^{0,\gamma}(B_{r_{x_0}}(x_0))$ for some ball $B_{r_{x_0}}(x_0)$. We then look at the Caccioppoli type inequality in (3.5); this implies that, for $\varrho \leq r_{x_0}/2$

$$\int_{B_\varrho(x_0)} H(x, Du) dx \leq c \int_{B_{2\varrho}(x_0)} H_{B_{2\varrho}(x_0)}^- \left(\frac{u - (u)_{B_{2\varrho}(x_0)}}{2\varrho} \right) dx \leq c H_{B_\varrho(x_0)}^- (\varrho^{\gamma-1})$$

and

$$\left[H_{B_\varrho(x_0)}^- \left(\frac{1}{\varrho} \right) \right]^{-1} \int_{B_\varrho(x_0)} H(x, Du) dx \leq c \varrho^{\gamma p}, \quad (5.35)$$

for $c \equiv c(\text{data}, \beta, [u]_{0,\gamma})$ (here we have used that $H_{B_{2\varrho}}^- \leq H_{B_\varrho}^-$). Letting $\varrho \rightarrow 0$ in the above display implies $x_0 \notin \Sigma$, a contradiction. This proves that $\Sigma \subset \Sigma_u$ and therefore (1.20). Next, observe that by definition we have $H_{B_\varrho(x_0)}^-(1/\varrho) \geq \varrho^{-p}$. Combining this with Lemma 4 (which is used to assert the integrability of $[H(x, Du)]^{1+\delta_g}$) we conclude with

$$\Sigma_u \subset \left\{ x_0 \in \Omega : \limsup_{\varrho \rightarrow 0} \varrho^{p(1+\delta_g)-n} \int_{B_\varrho(x_0)} [H(x, Du)]^{1+\delta_g} dx > 0 \right\}. \quad (5.36)$$

Giusti's Lemma ([31, Proposition 2.7]) then implies

$$\dim_{\mathcal{H}}(\Sigma_u) \leq n - p - p\delta_g \implies \mathcal{H}^{n-p}(\Sigma_u) = 0. \quad (5.37)$$

Recall we are treating the case when $p(1 + \delta_g) \leq n$. In particular, we have $|\Sigma_u| = 0$. Notice also that, once proved that Du is locally Hölder continuous in Ω_u , we shall have proved the validity of (1.39). In a totally similar way, assume also that $q(1 + \delta_g) \leq n$; we observe that if $x_0 \in \Sigma_u$ is such that $a(x_0) > 0$, then $H_{B_\varrho(x_0)}^-(1/\varrho) \geq a_i(B_\varrho(x_0))\varrho^{-q} > 0$ for ϱ sufficiently small, and we have

$$\Sigma_u \cap \{a(x) > 0\} \subset \left\{ x_0 \in \Omega : \limsup_{\varrho \rightarrow 0} \varrho^{q(1+\delta_g)-n} \int_{B_\varrho(x_0)} [H(x, Du)]^{1+\delta_g} dx > 0 \right\}.$$

Therefore, again by Giusti's Lemma, we also have

$$\dim_{\mathcal{H}}(\Sigma_u) \leq n - q - q\delta_g \implies \mathcal{H}^{n-q}(\Sigma_u \cap \{a(x) > 0\}) = 0. \quad (5.38)$$

Finally, we observe that in fact we have

$$\Omega_u = \left\{ x_0 \in \Omega : \text{there exists a ball } B_{r_{x_0}}(x_0) \subset \Omega \text{ such that } u \in C^{0,\gamma}(B_{r_{x_0}}(x_0)) \text{ for every } \gamma \in (0, 1) \right\}. \quad (5.39)$$

Indeed, call $\tilde{\Omega}_u$ the set in the right-hand side of the previous display; $\tilde{\Omega}_u \subset \Omega_u$, again by Proposition 2. On the other hand, the us take $x_0 \in \Omega_u$; it follows that there exists $B_{r_{x_0}}(x_0) \subset \Omega$ with $u \in C^{0,\tilde{\gamma}}(B_{r_{x_0}}(x_0))$ for some $\tilde{\gamma} < 1$; then, fix $\gamma < 1$ and determine the corresponding $\varepsilon_\gamma \equiv \varepsilon_\gamma(\text{data}, \beta, \gamma)$ according to Proposition 2. We can take ϱ small enough in (5.35) (this time with γ replaced by $\tilde{\gamma}$) in such a way that the smallness condition in (5.33) is satisfied. This implies that u is γ -Hölder continuous in a neighbourhood of x_0 . As $\gamma \in (0, 1)$ has been chosen arbitrarily, we deduce that $x_0 \in \tilde{\Omega}_u$ and therefore $\Omega_u = \tilde{\Omega}_u$, that is, (5.39) is completely proved.

By (5.39) and matching the content of Remark 12 (in particular, see (5.34)) with a standard covering argument, we conclude with

Proposition 3 *Let $u \in W_{\text{loc}}^{1,1}(\Omega, \mathbb{S}^{N-1})$ be a constrained local minimizer of the functional $\mathcal{F}$ in (1.1), under the assumptions (1.12)-(1.15). Assume that $p(1 + \delta_g) \leq n$. For every open subset $\Omega_0 \Subset \Omega_u$ and every $\gamma \in (0, 1)$, there exist constants $c, \tilde{c} \equiv c, \tilde{c}(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \Omega_0)$ such that*

$$\int_{B_\varrho} H(x, Du) dx \leq c\varrho^{(\gamma-1)q} \quad \text{and} \quad [u]_{0,\gamma;\Omega_0} \leq \tilde{c}, \quad (5.40)$$

with the first one that holds whenever $B_{2\varrho} \subset \Omega_0$ with $2\varrho \leq 1$.

In (5.40) we notice that the second inequality actually implies the first one via (3.5). As for the rest of the proof, as mentioned above, we only need to show that Du is locally Hölder continuous in Ω_u .

Step 4: Passage to coordinates.

After the proof of the local partial Hölder continuity of u , we can now pass to coordinates using stereographic projections. The procedure is standard in the case of functionals with p -growth but, since we are dealing with non-standard growth conditions, we need to check extra regularity conditions and therefore we shall repeat it in some detail. Having (5.40) in mind, we fix a certain initial γ , say $\gamma = 1/2$. We then consider bounded open subsets $\tilde{\Omega} \Subset \Omega_0 \Subset \Omega_u$ and cover $\tilde{\Omega}$ with finitely many balls $B \subset \Omega_0$ with sufficiently small radius (size and number here only depend on $n, N, [u]_{0,1/2;\Omega_0}$ and $\text{diam}(\Omega_0)$) such that $u(B)$ lies in single coordinate neighbourhood of $\mathbb{S}^{N-1}$. More precisely, if $B \equiv B_{r(B)}(x_0)$ is one of such balls, up to rotations we can assume that $u(x_0) = (-1, 0, \dots, 0)$ and that $u^1(x) \leq -1/2$ for every $x \in B$. Given this, with no loss of generality we can reduce to the case in which we are working on an open subset $\tilde{\Omega} \Subset \Omega$ such that $u^1(x) \leq -1/2$ for every $x \in \tilde{\Omega}$. This is the setting we shall use in the rest of the proof and our next goal is now to prove that Du is locally β_0 -Hölder continuous in $\tilde{\Omega}$, with β_0 depending only on $(\text{data}, \beta, \beta_1)$, where β_1 is the exponent appearing in (1.14)₂. The full statement of Theorem 1 then follows again via a standard covering argument. To proceed, denoting by $P(\cdot)$ the usual stereographic projection $P: \mathbb{S}^{N-1} \setminus \{(1, \dots, 0)\} \rightarrow \mathbb{R}^{N-1}$ and by $S := P^{-1}: \mathbb{R}^{N-1} \rightarrow \mathbb{S}^{N-1} \setminus \{(1, \dots, 0)\}$ its inverse, i.e.,

$$S(y) = \left(\frac{|y|^2 - 1}{|y|^2 + 1}, \frac{2y}{|y|^2 + 1} \right), \quad S^{-1}(v) = \left(\frac{v^i}{1 - v^1} \right)_{i=2}^N, \quad (5.41)$$

we then define $\tilde{u} := S^{-1}(u)$. We note that

$$\|\nabla S\|_{L^\infty} \leq c(N), \quad \|\nabla^2 S\|_{L^\infty} \leq c(N) \quad \text{and} \quad |\nabla(S^{-1})(u(x))| \leq c(N), \quad (5.42)$$

the last inequality being valid for all $x \in \tilde{\Omega}$ (as $u^1(x) \leq -1/2$ whenever $x \in \tilde{\Omega}$). Recalling (5.41), again that $u^1(x) \leq -1/2$ whenever $x \in \tilde{\Omega}$, and that $\tilde{u} = S^{-1}(u)$, we get

$$\|\tilde{u}\|_{L^\infty(\tilde{\Omega})} \leq \frac{2}{3} \leq 1. \quad (5.43)$$

Again (5.42) implies that if $\tilde{w} \in W^{1,H}(\tilde{\Omega}, \mathbb{R}^{N-1})$, then $S(\tilde{w}) \in W^{1,H}(\tilde{\Omega}, \mathbb{S}^{N-1})$. Therefore, by the minimality of u it follows that the map $\tilde{u} \in W^{1,H}(\tilde{\Omega}, \mathbb{R}^{N-1})$ is a local minimizer of the functional

$$W^{1,H}(\tilde{\Omega}, \mathbb{R}^{N-1}) \ni w \mapsto \int_{\tilde{\Omega}} G(x, w, Dw) dx, \quad (5.44)$$

where the integrand $G(\cdot)$ is defined by

$$G(x, y, z) := F(x, S(y), \nabla S(y)z) \quad \text{for } x \in \tilde{\Omega}, y \in \mathbb{R}^{N-1}, z \in \mathbb{R}^{(N-1) \times n}.$$

As it is

$$|\nabla S(y)z| = \frac{2}{(1 + |y|^2)} |z| \quad (5.45)$$

(as it follows from an elementary but lengthy computation), recalling (1.14)₁ we conclude with

$$G(x, y, z) = \tilde{G}(x, y, |z|) \equiv \tilde{G}_{x,y}(|z|) := \tilde{F}(x, S(y), 2(1 + |y|^2)^{-1}|z|). \quad (5.46)$$

By using the starting assumptions (1.12), it is now not difficult to show that for every $M \geq 3N$ there exist new constants $0 < \tilde{\nu} \equiv \tilde{\nu}(\text{data}, M) \leq 1 \leq \tilde{L} \equiv \tilde{L}(\text{data})$, such that

$$\left\{ \begin{array}{l} \tilde{\nu}H(x, z) \leq G(x, y, z) \leq \tilde{L}H(x, z) \\ |\partial G(x, y, z)||z| + |\partial^2 G(x, y, z)||z|^2 \leq \tilde{L}H(x, z) \\ \tilde{\nu}(|z|^{p-2} + a(x)|z|^{q-2})|\xi|^2 \leq \langle \partial^2 G(x, y, z)\xi, \xi \rangle \\ |\partial G(x_1, y, z) - \partial G(x_2, y, z)||z| \leq \tilde{L}\omega(|x_1 - x_2|)[H(x_1, z) + H(x_2, z)] + \tilde{L}|a(x_1) - a(x_2)||z|^q \\ |G(x, y_1, z) - G(x, y_2, z)| \leq \tilde{L}\omega(|y_1 - y_2|)H(x, z), \end{array} \right. \quad (5.47)$$

hold whenever $x, x_1, x_2 \in \Omega, z \in \mathbb{R}^{(N-1) \times n} \setminus \{0\}, \xi \in \mathbb{R}^{(N-1) \times n}$ and $y, y_1, y_2 \in \mathbb{R}^{N-1}$ are such that $|y| + |y_1| + |y_2| \leq M$. In the lines above, $\tilde{\nu}$ is a non-increasing function of M . All the inequalities in (5.47) are consequences of the definition in (5.46) and of (5.45) and we leave the details of the verification to the reader. We just spend a few words on the verification of (5.47)₃. By using (3.26), from the explicit representation in (5.46) we get

$$\partial^2 G(x, y, z) = 4 \frac{\tilde{F}''(x, S(y), |\tilde{z}|)}{(1 + |y|^2)^2} \frac{\tilde{z} \otimes \tilde{z}}{|\tilde{z}|^2} + 4 \frac{\tilde{F}'(x, S(y), |\tilde{z}|)}{(1 + |y|^2)^2} \left[\frac{\mathbb{I}_{N \times n}}{|\tilde{z}|} - \frac{\tilde{z} \otimes \tilde{z}}{|\tilde{z}|^3} \right], \quad (5.48)$$

where we have denoted $\tilde{z} := 2(1 + |y|^2)^{-1}z$. Taking $\xi \in \mathbb{R}^{(N-1) \times n}$ and adding one more null component to both $\tilde{z}$ and ξ (thereby making then $\mathbb{R}^{N \times n}$ matrices), and using (3.25) and (1.12)₃, yields

$$\langle \partial^2 G(x, y, z)\xi, \xi \rangle \geq \frac{\nu}{2^q} \left(\frac{|z|^{p-2}}{(1 + |y|^2)^p} + a(x) \frac{|z|^{q-2}}{(1 + |y|^2)^q} \right) |\xi|^2 \geq \frac{\nu(|z|^{p-2} + a(x)|z|^{q-2})}{(2 + 4M^2)^q} |\xi|^2,$$

that is, (5.47)₃. Finally, fix $x \in \tilde{\Omega}$ and $y \in \mathbb{R}^{N-1}$. By (1.14)₁ and (5.46) it follows that

$$t \mapsto \tilde{G}_{x,y}(\cdot, t) \text{ is non-decreasing} \quad (5.49)$$

and this means that the structure assumption (1.14)₁ is verified also by $G(\cdot)$. As for the analog of (1.14)₂, observe that (5.46) implies that $\tilde{G}_{x,y}''(t) = [2(1 + |y|^2)^{-1}]^2 \tilde{F}_{x,S(y)}''(2(1 + |y|^2)^{-1}t)$. Then, with $|s| < t/2$ and $t > 0$, we define $\tilde{s} := 2(1 + |y|^2)^{-1}s$ and $\tilde{t} := 2(1 + |y|^2)^{-1}t$; taking into account (3.27) we find

$$\begin{aligned} |\tilde{G}_{x,y}''(t+s) - \tilde{G}_{x,y}''(t)| &= \frac{4}{(1 + |y|^2)^2} |\tilde{F}_{x,S(y)}''(\tilde{t} + \tilde{s}) - \tilde{F}_{x,S(y)}''(\tilde{t})| \\ &\leq \frac{c}{(1 + |y|^2)^2} \tilde{F}_{x,S(y)}''(\tilde{t}) \left(\frac{|\tilde{s}|}{\tilde{t}} \right)^{\beta_1} = c \tilde{G}_{x,y}''(t) \left(\frac{|s|}{t} \right)^{\beta_1}, \end{aligned} \quad (5.50)$$

where it is again $c \equiv c(n, N, \nu, L, p, q)$. We moreover remark that, by the growth conditions in (5.47), exactly as done for (3.24) in Remark 8 we infer that

$$\tilde{G}_{x,y}''(t)t \approx \tilde{G}_{x,y}'(t) \quad \text{holds for every } t > 0 \quad (5.51)$$

and here the implied constants depend on n, N, ν, L, p, q, M and they are independent of (x, y) .

Step 5: Partial Hölder continuity of the gradient.

First of all, let us observe that for reasons that will be clear in a few lines, and in view of (5.43), in the following, when considering (5.47), we shall permanently use the choice $M = 10N$. Recall from Step 2 that $u \in C_{\text{loc}}^{0,\gamma}(\Omega_u, \mathbb{S}^{N-1})$ for every $\gamma \in (0, 1)$, and so, by (5.42) and (5.40), we have that $\tilde{u} \in C^{0,\gamma}(\tilde{\Omega}, \mathbb{R}^{N-1})$ for every $\gamma \in (0, 1)$, with

$$[\tilde{u}]_{0,\gamma;\tilde{\Omega}} \leq c(N)[u]_{0,\gamma;\tilde{\Omega}} \leq c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega}). \quad (5.52)$$

For any ball $B_{4r} \Subset \tilde{\Omega}$ with $r \leq 1/8$, as $\tilde{u}$ minimizes the functional in (5.44) and (5.43) holds, also taking Remark 6 into account, Lemma 2 provides

$$\int_{B_{2r}} H(x, D\tilde{u}) \, dx \leq c \int_{B_{4r}} H_{B_{4r}}^- \left(\frac{\tilde{u} - (\tilde{u})_{B_{4r}}}{4r} \right) \, dx \stackrel{(5.52)}{\leq} c H_{B_r}^- (r^{\gamma-1}), \quad (5.53)$$

holds with $c \equiv c(\text{data}, \beta, \gamma, \tilde{\Omega})$, for every $\gamma \in (0, 1)$. In particular, it follows that

$$\int_{B_{2r}} H(x, D\tilde{u}) \, dx \leq c r^{(\gamma-1)q}, \quad (5.54)$$

for all $\gamma \in (0, 1)$, where $c \equiv c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega})$. We start fixing $\gamma \geq 1/2$; we shall further increase the value of γ in due course of the proof (and the constants involved will increase accordingly). Fix $B_r \equiv B_r(x_0)$ such that $B_{4r} \Subset \tilde{\Omega}$, $r \leq 1/64$, and let $\tilde{v} \in W^{1,H}(B_r, \mathbb{R}^{N-1})$ be a solution to the frozen Dirichlet problem

$$\tilde{v} \mapsto \min_{w \in W_{\tilde{u}}^{1,H}(B_r, \mathbb{R}^{N-1})} \int_{B_r} G(x, (\tilde{u})_{B_r}, Dw) \, dx. \quad (5.55)$$

By (5.43) the integrand $G(\cdot, (\tilde{u})_{B_r}, \cdot)$ satisfies assumptions (5.47) with $\tilde{v}$ and $\tilde{L}$ only depending on data as we have fixed $M = 10N$. In particular, there exist positive numbers $\tilde{v}, \tilde{L}$ such that

$$\tilde{v}(\text{data})H(x, z) \leq G(x, (\tilde{u})_{B_r}, z) \leq \tilde{L}(\text{data})H(x, z) \quad (5.56)$$

holds whenever $x \in \Omega$ and $z \in \mathbb{R}^{(N-1) \times n}$. By (5.49) and the maximum principle [43, Theorem 2.3] we then have

$$\|\tilde{v}\|_{L^\infty(B_r)} \leq \sqrt{N} \|\tilde{u}\|_{L^\infty(B_r)} \stackrel{(5.43)}{\leq} \frac{2\sqrt{N}}{3}. \quad (5.57)$$

The validity of the Euler-Lagrange equation for (5.55) can be checked as done in [10] (see also Section 3.2 and apply the same arguments exposed there without using projections, that is, when no constraints are involved). Specifically, $\tilde{v}$ solves

$$\int_{B_r} \partial G(x, (\tilde{u})_{B_r}, D\tilde{v}) \cdot D\varphi \, dx = 0, \quad (5.58)$$

for all $\varphi \in W_0^{1,H}(B_r, \mathbb{R}^{N-1})$.

Remark 13 From now on, we adopt the following convention. Also taking the content of Remarks 6-7 and (5.56)-(5.57) into account, the results of Lemma 4 and 5 apply to $\tilde{u}, \tilde{v}$ and lead to new higher integrability exponents δ_g and σ_g . The values of δ_g and σ_g are different from those used in the previous steps for u , but essentially equivalent to them. Indeed, they still depend on the same set of parameters, that is data. Therefore, with some abuse of notation, we shall keep on denoting by $\sigma_g < \delta_g$ the higher integrability exponents provided by the application of Lemmas 4-5 in the present setting (notice that what is denoted by N in Lemmas 4-5 is actually $N - 1$ here; this can be made rigorous eventually taking the smallest amongst all the exponents considered when the values of ν and L attain their minimum and maximum, respectively). All in all, the following inequalities hold as in (5.6) and (5.7):

$$\int_{B_r} H(x, D\tilde{v}) \, dx \leq c \int_{B_r} H(x, D\tilde{u}) \, dx, \quad \left(\int_{B_{r/2}} [H(x, D\tilde{v})]^{1+\delta_g} \, dx \right)^{1/(1+\delta_g)} \leq \tilde{c}_1 \int_{B_r} H(x, D\tilde{v}) \, dx \quad (5.59)$$

and

$$\left(\int_{B_r} [H(x, D\tilde{v})]^{1+\sigma_g} \, dx \right)^{1/(1+\sigma_g)} \leq c \left(\int_{B_r} [H(x, D\tilde{u})]^{1+\sigma_g} \, dx \right)^{1/(1+\sigma_g)} \leq c \int_{B_{2r}} H(x, D\tilde{u}) \, dx, \quad (5.60)$$

for $c, \tilde{c}_1 \equiv c, \tilde{c}_1(\text{data})$. This last dependence on the constants is a consequence of (5.56)-(5.57).

As $\varphi \equiv \tilde{u} - \tilde{v}$ is a legal choice in (5.58), using (5.4) as in (5.5), we end up with

$$\begin{aligned} & c \int_{B_r} (|V_p(D\tilde{u}) - V_p(D\tilde{v})|^2 + a(x)|V_q(D\tilde{u}) - V_q(D\tilde{v})|^2) \, dx \\ & \leq \int_{B_r} \left[G(x, (\tilde{u})_{B_r}, D\tilde{u}) - G(x, (\tilde{u})_{B_r}, D\tilde{v}) \right] \, dx \\ & = \int_{B_r} \left[G(x, (\tilde{u})_{B_r}, D\tilde{u}) - G(x, \tilde{u}, D\tilde{u}) \right] \, dx + \int_{B_r} [G(x, \tilde{u}, D\tilde{u}) - G(x, \tilde{v}, D\tilde{v})] \, dx \\ & \quad + \int_{B_r} \left[G(x, \tilde{v}, D\tilde{v}) - G(x, (\tilde{v})_{B_r}, D\tilde{v}) \right] \, dx \\ & \quad + \int_{B_r} \left[G(x, (\tilde{v})_{B_r}, D\tilde{v}) - G(x, (\tilde{u})_{B_r}, D\tilde{v}) \right] \, dx =: \sum_{j=1}^4 (I)_j, \end{aligned} \quad (5.61)$$

with $c \equiv c(\text{data})$. Before taking care of terms (I)₁-(I)₄, we derive a few preliminary inequalities. The first one is an obvious consequence of (5.52) and is

$$\sup_{x \in B_r} \omega(|\tilde{u}(x) - (\tilde{u})_{B_r}|) \leq cr^{\beta\gamma} \leq cr^{\beta/2}, \quad (5.62)$$

for $c \equiv c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega})$ (here recall that $\gamma \geq 1/2$ and $r \leq 1$). Proceeding as for (5.9), and recalling (5.59), we instead have

$$\int_{B_r} \omega(|\tilde{v} - (\tilde{v})_{B_r}|) \, dx \leq c\omega \left[r(H_{B_r}^-)^{-1} \left(\int_{B_r} H(x, D\tilde{v}) \, dx \right) \right]$$

$$\begin{aligned}
&\leq c\omega \left[r \left(H_{B_r}^- \right)^{-1} \left(\int_{B_r} H(x, D\tilde{u}) dx \right) \right] \\
&\stackrel{(5.53)}{\leq} c\omega \left[r \left(H_{B_r}^- \right)^{-1} \left(c H_{B_r}^- (r^{\gamma-1}) \right) \right] \leq cr^{\beta\gamma} \leq cr^{\beta/2}, \tag{5.63}
\end{aligned}$$

with $c \equiv c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega})$. In a totally similar fashion, as done in (5.10), we have

$$\omega \left(|(\tilde{u})_{B_r} - (\tilde{v})_{B_r}| \right) \leq cr^{\beta\gamma} \leq cr^{\beta/2}, \tag{5.64}$$

where $c \equiv c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega})$. We are now ready to estimate terms (I)₁-(I)₄. We have

$$(I)_1 \leq \int_{B_r} \omega \left(|\tilde{u} - (\tilde{u})_{B_r}| \right) H(x, D\tilde{u}) dx \stackrel{(5.62)}{\leq} cr^{\beta/2} \int_{B_{2r}} H(x, D\tilde{u}) dx,$$

with $c \equiv c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega})$. The minimality of $\tilde{u}$ gives (I)₂ ≤ 0 . As for term (I)₃, using (5.60) and (5.63), we get

$$\begin{aligned}
(I)_3 &\leq c \left(\int_{B_r} \omega \left(|\tilde{v} - (\tilde{v})_{B_r}| \right) dx \right)^{\frac{\sigma_g}{1+\sigma_g}} \left(\int_{B_r} [H(x, D\tilde{v})]^{1+\sigma_g} dx \right)^{\frac{1}{1+\sigma_g}} \\
&\stackrel{(5.60)}{\leq} c \left(\int_{B_r} \omega \left(|\tilde{v} - (\tilde{v})_{B_r}| \right) dx \right)^{\frac{\sigma_g}{1+\sigma_g}} \int_{B_{2r}} H(x, D\tilde{u}) dx \\
&\stackrel{(5.63)}{\leq} cr^{\frac{\beta\sigma_g}{2(1+\sigma_g)}} \int_{B_{2r}} H(x, D\tilde{u}) dx,
\end{aligned}$$

for $c \equiv c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega})$ and, in a totally similar fashion, arguing as for (5.15), but using this time (5.64), we find

$$(I)_4 \stackrel{(5.62)}{\leq} cr^{\frac{\beta\sigma_g}{2(1+\sigma_g)}} \int_{B_{2r}} H(x, D\tilde{u}) dx,$$

where, again it is $c \equiv c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega})$. Collecting the estimates found above for the terms (I)₁-(I)₄ to (5.61) yields

$$\int_{B_r} \left(|V_p(D\tilde{u}) - V_p(D\tilde{v})|^2 + a(x) |V_q(D\tilde{u}) - V_q(D\tilde{v})|^2 \right) dx \leq cr^{\frac{\beta\sigma_g}{2(1+\sigma_g)}} \int_{B_{2r}} H(x, D\tilde{u}) dx,$$

where $c \equiv c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega})$. In particular, there holds

$$\int_{B_r} \left(|V_p(D\tilde{u}) - V_p(D\tilde{v})|^2 + a_i(B_r) |V_q(D\tilde{u}) - V_q(D\tilde{v})|^2 \right) dx \leq cr^{\frac{\beta\sigma_g}{2(1+\sigma_g)}} \int_{B_{2r}} H(x, D\tilde{u}) dx. \tag{5.65}$$

We now select $x_m \in \bar{B}_r$ such that

$$a(x_m) = a_i(B_r) \tag{5.66}$$

and define (keep in mind the notation in (5.46))

$$G_m(z) := G \left(x_m, (\tilde{u})_{B_r}, z \right) \equiv \tilde{G}_{x_m, (\tilde{u})_{B_r}}(|z|) \quad \text{for every } z \in \mathbb{R}^{(N-1) \times n}. \tag{5.67}$$

The newly defined integrand $G_m(\cdot)$ is of the type $g_0(\cdot)$ considered in Lemma 8 and satisfies assumptions (4.29) with the choice $H_0(\cdot) \equiv H_{B_r}^-(\cdot)$ and for suitable constants ν, L depending on data (see the discussion in Step 4 and, in particular, (5.47) and (5.56)). We now proceed applying Lemma 8 to $\tilde{v}$; notice that (4.30)

is automatically satisfied by the second inequality in (5.59) and therefore verify (4.31). According to the terminology adopted in [4, 9, 10] the p -phase occurs if

$$a_i(B_r) \leq 4[a]_{0,\alpha} r^{\alpha-s}, \quad s := \alpha + (\gamma - 1)(q - p) \stackrel{(1.15)_1}{>} 0, \quad (5.68)$$

while the (p, q) -phase is defined by the complementary condition

$$a_i(B_r) > 4[a]_{0,\alpha} r^{\alpha-s}. \quad (5.69)$$

It is then easy to see that

$$\begin{cases} a_s(B_r) \leq 6[a]_{0,\alpha} r^{\alpha-s} & \text{holds in the } p\text{-phase} \\ a_s(B_r) \leq \frac{3}{2}a_i(B_r) & \text{holds in the } (p, q)\text{-phase.} \end{cases} \quad (5.70)$$

Notice that the two phases described above depend on the number $\gamma \geq 1/2$, which is going to be chosen later as an absolute function of data. We start estimating, for every $\varphi \in C_c^\infty(B_{r/2}, \mathbb{R}^N)$

$$\begin{aligned} \left| \int_{B_{r/2}} \partial G_m(D\tilde{v}) \cdot D\varphi \, dx \right| &\stackrel{(5.58)}{=} \left| \int_{B_{r/2}} \left[\partial G_m(D\tilde{v}) - \partial G(x, (\tilde{u})_{B_r}, D\tilde{v}) \right] \cdot D\varphi \, dx \right| \\ &\stackrel{(5.47)_4}{\leq} c \|D\varphi\|_{L^\infty(B_{r/2})} \int_{B_{r/2}} \omega(|x_m - x|) \left[\frac{H(x_m, D\tilde{v})}{|D\tilde{v}|} + \frac{H(x, D\tilde{v})}{|D\tilde{v}|} \right] dx \\ &\quad + c \|D\varphi\|_{L^\infty(B_{r/2})} \int_{B_{r/2}} [a(x) - a(x_m)] |D\tilde{v}|^{q-1} dx \\ &\stackrel{(1.13)}{\leq} cr^\beta \|D\varphi\|_{L^\infty(B_{r/2})} \int_{B_{r/2}} |D\tilde{v}|^{p-1} dx \\ &\quad + cr^\beta \|D\varphi\|_{L^\infty(B_{r/2})} \int_{B_{r/2}} (a(x) + a_i(B_r)) |D\tilde{v}|^{q-1} dx \\ &\quad + c[a]_{0,\alpha} r^\alpha \|D\varphi\|_{L^\infty(B_{r/2})} \int_{B_{r/2}} |D\tilde{v}|^{q-1} dx \\ &=: \quad (\text{II})_1 + (\text{II})_2 + (\text{II})_3, \end{aligned} \quad (5.71)$$

and proceed estimating the terms appearing in the last three lines. In any case we have, Young's inequality gives

$$(\text{II})_1 \leq c(\text{data}) r^\beta \int_{B_r} \left(|D\tilde{v}|^p + \|D\varphi\|_{L^\infty(B_{r/2})}^p \right) dx. \quad (5.72)$$

In order to bound the remaining two terms we distinguish between the p -phase (5.68) and the (p, q) -phase (5.69). We start noticing that in the p -phase we have

$$\int_{B_{r/2}} H(x, D\tilde{v}) \, dx \stackrel{(5.59)}{\leq} c \int_{B_r} H(x, D\tilde{u}) \, dx \stackrel{(5.53)}{\leq} c H_{B_r}^-(r^{\gamma-1}) \stackrel{(5.68)}{\leq} cr^{(\gamma-1)p}, \quad (5.73)$$

for $c \equiv c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega})$. As $q - 1 < p$, using Hölder's inequality we have

$$\begin{aligned} (\text{II})_2 &\stackrel{(5.70)_1}{\leq} cr^{\beta+\alpha-s} \|D\varphi\|_{L^\infty(B_{r/2})} \left(\int_{B_{r/2}} |D\tilde{v}|^p \, dx \right)^{\frac{q-p}{p}} \left(\int_{B_{r/2}} |D\tilde{v}|^p \, dx \right)^{\frac{p-1}{p}} \\ &\leq cr^{\beta+\alpha-s} \|D\varphi\|_{L^\infty(B_{r/2})} \left(\int_{B_{r/2}} H(x, D\tilde{v}) \, dx \right)^{\frac{q-p}{p}} \left(\int_{B_{r/2}} |D\tilde{v}|^p \, dx \right)^{\frac{p-1}{p}} \end{aligned}$$

$$\begin{aligned}
& \stackrel{(5.73)}{\leq} c r^{\beta+\alpha-s+(\gamma-1)(q-p)} \|D\varphi\|_{L^\infty(B_{r/2})} \left(\int_{B_{r/2}} |D\tilde{v}|^p dx \right)^{\frac{p-1}{p}} \\
& \stackrel{(5.68)}{\leq} c r^\beta \int_{B_r} \left(|D\tilde{v}|^p + \|D\varphi\|_{L^\infty(B_{r/2})}^p \right) dx,
\end{aligned}$$

for $c \equiv c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega})$. Finally, we similarly have

$$\begin{aligned}
(\text{II})_3 & \leq c r^s r^{\alpha-s} \|D\varphi\|_{L^\infty(B_{r/2})} \left(\int_{B_{r/2}} |D\tilde{v}|^p dx \right)^{\frac{q-p}{p}} \left(\int_{B_{r/2}} |D\tilde{v}|^p dx \right)^{\frac{p-1}{p}} \\
& \stackrel{(5.68), (5.73)}{\leq} c r^s \int_{B_r} \left(|D\tilde{v}|^p + \|D\varphi\|_{L^\infty(B_{r/2})}^p \right) dx,
\end{aligned}$$

with $c \equiv c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega})$. We now consider the occurrence of the (p, q) -phase (5.69). We have, again by Hölder's inequality

$$\begin{aligned}
(\text{II})_2 & \stackrel{(5.70)_2}{\leq} c r^\beta \|D\varphi\|_{L^\infty(B_{r/2})} [a_i(B_r)]^{\frac{1}{q}} \int_{B_r} [a_i(B_r)]^{\frac{q-1}{q}} |D\tilde{v}|^{q-1} dx \\
& \leq c r^\beta \|D\varphi\|_{L^\infty(B_{r/2})} [a_i(B_r)]^{\frac{1}{q}} \left(\int_{B_r} a_i(B_r) |D\tilde{v}|^q dx \right)^{\frac{q-1}{q}} \\
& \leq c r^\beta \int_{B_r} \left(a_i(B_r) |D\tilde{v}|^q + a_i(B_r) \|D\varphi\|_{L^\infty(B_{r/2})}^q \right) dx
\end{aligned}$$

and

$$\begin{aligned}
(\text{II})_3 & = c[a]_{0,\alpha} r^{\alpha-s} r^s \|D\varphi\|_{L^\infty(B_{r/2})} \int_{B_{r/2}} |D\tilde{v}|^{q-1} dx \\
& \stackrel{(5.69)}{\leq} c r^s \|D\varphi\|_{L^\infty(B_{r/2})} [a_i(B_r)]^{\frac{1}{q}} \int_{B_r} [a_i(B_r)]^{\frac{q-1}{q}} |D\tilde{v}|^{q-1} dx \\
& \leq c r^s \int_{B_r} \left(a_i(B_r) |D\tilde{v}|^q + a_i(B_r) \|D\varphi\|_{L^\infty(B_{r/2})}^q \right) dx,
\end{aligned}$$

where $c \equiv c(\text{data}, \beta)$. Collecting the estimates founds for the terms $(\text{II})_1, (\text{II})_2, (\text{II})_3$ to (5.71) and recalling (5.68), in any case we conclude with

$$\left| \int_{B_{r/2}} \partial G_m(D\tilde{v}) \cdot D\varphi dx \right| \leq c r^{\min\{\beta, \alpha-(q-p)/2\}} \int_{B_r} \left[H_{B_r}^-(D\tilde{v}) + H_{B_r}^-(\|D\varphi\|_{L^\infty(B_{r/2})}) \right] dx,$$

with $c \equiv c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega})$. Here we have used $\gamma \geq 1/2$, so that it is $s \geq \alpha - (q-p)/2 > 0$. Lemma 8 (notice that the number N used there is actually $N-1$ in this context; recall again that here it is $N > 1$) yields the existence of $\tilde{h} \in W_{\tilde{v}}^{1, H_{B_r}^-}(B_{r/2}, \mathbb{R}^{N-1})$ satisfying

$$\int_{B_{r/2}} \partial G_m(D\tilde{h}) \cdot D\varphi dx = 0 \quad \text{for every } \varphi \in W_0^{1, H_{B_r}^-}(B_{r/2}, \mathbb{R}^{N-1}),$$

and, such that

$$\int_{B_{r/2}} \left(|V_p(D\tilde{v}) - V_p(D\tilde{h})|^2 + a_i(B_r) |V_q(D\tilde{v}) - V_q(D\tilde{h})|^2 \right) dx \leq c r^m \int_{B_r} H_{B_r}^-(D\tilde{v}) dx, \quad (5.74)$$

where $c \equiv c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega})$ and $m = m(n, N, \nu, L, p, q, \alpha)$. Needless to say, $\tilde{h}$ solves

$$\tilde{h} \mapsto \min_{w \in W_{\tilde{\nu}}^{1, H^-}_{B_r}(B_{r/2}, \mathbb{R}^{N-1})} \int_{B_{r/2}} G_m(Dw) \, dx \quad (5.75)$$

so that, recalling again (5.66), we find

$$\int_{B_{r/2}} H_{B_r}^-(D\tilde{h}) \, dx \leq c \int_{B_r} H_{B_r}^-(D\tilde{v}) \, dx \leq \int_{B_{r/2}} H(x, D\tilde{v}) \, dx \stackrel{(5.59)}{\leq} c \int_{B_r} H(x, D\tilde{u}) \, dx. \quad (5.76)$$

Hence, from (5.65), (5.74), and the last inequality in the above display, we obtain

$$\int_{B_{r/2}} (|V_p(D\tilde{u}) - V_p(D\tilde{h})|^2 + a_i(B_r)|V_q(D\tilde{u}) - V_q(D\tilde{h})|^2) \, dx \leq cr^\kappa \int_{B_{2r}} H(x, D\tilde{u}) \, dx, \quad (5.77)$$

for

$$\kappa := \min \left\{ m, \frac{\beta \sigma_g}{2(1 + \sigma_g)} \right\} < 1$$

and $c \equiv c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega})$. After some standard manipulations on (5.77), see e.g. [4, Section 10] and [9, pp. 483], we have

$$\int_{B_{r/2}} |D\tilde{u} - D\tilde{h}|^p \, dx \leq \int_{B_{r/2}} H_{B_r}^-(D\tilde{u} - D\tilde{h}) \, dx \leq cr^{\kappa/2} \int_{B_{2r}} H(x, D\tilde{u}) \, dx, \quad (5.78)$$

for $c \equiv c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega})$. Now we make a further restriction on the size of γ imposing that

$$\gamma \geq 1 - \frac{\kappa}{4q} \quad (5.79)$$

(which is still larger than $1/2$), and apply (5.53); we use the resulting inequality in (5.78) to obtain

$$\int_{B_{r/2}} |D\tilde{u} - D\tilde{h}|^p \, dx \leq cr^{\kappa/4}, \quad (5.80)$$

where $c \equiv c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega})$. By using the content of Remark 14 below we have that

$$\int_{B_\varrho} H_{B_r}^-(D\tilde{h} - (D\tilde{h})_{B_\varrho}) \, dx \leq c \left(\frac{\varrho}{r} \right)^\mu \int_{B_r} H(x, D\tilde{u}) \, dx, \quad (5.81)$$

holds for concentric balls $B_\varrho \subset B_{r/2}$; here we take $\varrho \leq r/8$. Here it is $c \equiv c(n, N, \nu, L, p, q)$ and $\mu \equiv \mu(n, N, \nu, L, p, q, \beta_1)$. We estimate

$$\begin{aligned} \int_{B_\varrho} |D\tilde{u} - (D\tilde{u})_{B_\varrho}|^p \, dx &\leq c \int_{B_\varrho} |D\tilde{u} - (D\tilde{h})_{B_\varrho}|^p \, dx \\ &\stackrel{(5.80)}{\leq} c \left\{ \left(\frac{r}{\varrho} \right)^n r^{\kappa/4} + \int_{B_\varrho} H_{B_r}^-(D\tilde{h} - (D\tilde{h})_{B_\varrho}) \, dx \right\} \\ &\stackrel{(5.81)}{\leq} c \left\{ \left(\frac{r}{\varrho} \right)^n r^{\kappa/4} + \left(\frac{\varrho}{r} \right)^\mu \int_{B_r} H(x, D\tilde{u}) \, dx \right\} \\ &\stackrel{(5.53)}{\leq} c \left\{ \left(\frac{r}{\varrho} \right)^n r^{\kappa/4} + \left(\frac{\varrho}{r} \right)^\mu r^{(\gamma-1)q} \right\}, \end{aligned} \quad (5.82)$$

where $c \equiv c(\text{data}, \|a\|_{L^\infty}, \beta, \gamma, \tilde{\Omega})$. In (5.82) we pick

$$\varrho = \frac{r^{1+a}}{8} \quad \text{with} \quad a := \frac{(1-\gamma)q + \kappa/4}{\mu + n} \quad \text{and} \quad \gamma = 1 - \frac{\kappa\mu}{8nq},$$

(γ also meets the condition in (5.79)) so that (5.82) yields

$$\int_{B_\varrho} |D\tilde{u} - (D\tilde{u})_{B_\varrho}|^p dx \leq c\varrho^{\frac{\kappa/4-na}{1+a}} \leq c\varrho^{\beta_0 p}, \quad \beta_0 := \frac{\kappa\mu}{16p(n+\mu)}. \quad (5.83)$$

The classical Hölder continuity characterization of Campanato and Meyers, a standard covering argument, and the fact that $\tilde{\Omega} \Subset \Omega_u$ is arbitrary, allow to conclude that $D\tilde{u} \in C_{\text{loc}}^{0,\beta_0}(\Omega_u, \mathbb{R}^{(N-1) \times n})$, with β_0 as in (5.83) and therefore depends on $(\text{data}, \beta, \beta_1)$. Finally, using (5.42)_{1,2} we get $Du = D(S(\tilde{u})) \in C_{\text{loc}}^{0,\beta_0}(\Omega_u, \mathbb{R}^{N \times n})$ and the proof of the partial local Hölder continuity of the gradient as stated in (1.18) is complete in the case it is $p(1 + \delta_g) \leq n$.

Step 6: The case $p(1 + \delta_g) > n$.

In this case the singular set is empty $\Sigma_u = \Omega_u = \Omega$ as the right-hand side in (5.36) is empty. Therefore we see from Step 3 that $u \in C_{\text{loc}}^{0,\gamma}(\Omega, \mathbb{S}^{N-1})$ for every $\gamma < 1$ and the rest of the proof, i.e., the local Hölder continuity of Du follows as for the case when $p(1 + \delta_g) \leq n$.

Remark 14 We briefly explain how to get estimate (5.81) from the results of [20]. Recalling the notation in (5.67), by (5.51) we can argue exactly as in Remark 8 to get that

$$\sup_{B_{r/4}} H_{B_r}^-(D\tilde{h}) \leq c \int_{B_{r/2}} H_{B_r}^-(D\tilde{h}) dx \quad (5.84)$$

holds for a constant $c \equiv c(n, N, \nu, L, p, q)$. Moreover, by (5.50) we are able to satisfy [21, Assumption 2.2], where we can take $\varphi(\cdot) \equiv \tilde{G}_{x_m, (\tilde{u})_{B_r}}(\cdot)$ and therefore (also taking into account the definitions in (2.1) and in [21, (1.3)]) we can apply [21, Theorem 6.4] that in the present setting gives

$$\begin{aligned} & \int_{B_\varrho} (|V_p(D\tilde{h}) - (V_p(D\tilde{h}))_{B_\varrho}|^2 + a_i(B_r)|V_q(D\tilde{h}) - (V_q(D\tilde{h}))_{B_\varrho}|^2) dx \\ & \leq c \left(\frac{\varrho}{r}\right)^{2\mu} \int_{B_{r/2}} (|V_p(D\tilde{h}) - (V_p(D\tilde{h}))_{B_{r/2}}|^2 + a_i(B_r)|V_q(D\tilde{h}) - (V_q(D\tilde{h}))_{B_{r/2}}|^2) dx, \end{aligned} \quad (5.85)$$

whenever $B_\varrho \subset B_{r/2}$ is concentric to $B_{r/2}$, where $c \geq 1$ and $\mu \in (0, 1/2)$ are both depending on n, N, p, q, ν, L and β_1 . Following the method explained in [3, Theorem 3.1], see also [14, Proposition 3.3], and combining (5.84)-(5.85) with (2.3), finally yields

$$\int_{B_\varrho} H_{B_r}^-(D\tilde{h} - (D\tilde{h})_{B_\varrho}) dx \leq c \left(\frac{\varrho}{r}\right)^\mu \int_{B_{r/2}} H_{B_r}^-(D\tilde{h}) dx,$$

from which (5.81) obviously follows.

6 Weighted Hausdorff measures and singular sets

6.1 Proof of Proposition 1

Observe that, for a ball B such that $\beta_4^{-1}B \subset \Omega$ and $r(B) \leq 1$ we then have

$$|\beta_4^{-1}B| \operatorname{ess\,sup}_{x \in \beta_4^{-1}B} \Phi(x, 1/r(\beta_4^{-1}B)) = \frac{|B|}{\beta_4^n} \operatorname{ess\,sup}_{x \in \beta_4^{-1}B} \Phi(x, \beta_4/r(B))$$

$$\stackrel{(1.32)}{\leq} \frac{c_d}{\beta_4^n} |B| \operatorname{ess\,inf}_{x \in \beta_4^{-1} B} \Phi(x, 1/r(B)) \leq \frac{c_d}{\beta_4^n} |B| \operatorname{ess\,inf}_{x \in B} \Phi(x, 1/r(B)) .$$

By taking balls B such that $r(B) \leq \kappa$ and $\kappa \leq 1$, we find $\mathcal{H}_{\Phi, \kappa/\beta_4}^+ \leq (c_d/\beta_4^n) \mathcal{H}_{\Phi, \kappa}^-$, from which (1.33) follows by letting $\kappa \rightarrow 0$. As for the proof of (1.34), we observe that in this case it is $\Phi(x, t) = [t^p + a(x)t^q]^{1+\sigma} \approx t^{p(1+\sigma)} + [a(x)]^{1+\sigma} t^{q(1+\sigma)}$, with constants implicit in " $\approx$ " depending on σ . Then, for $B \subset \Omega$ and $1 \leq t \leq 1/r(B)$, we have

$$\begin{aligned} \operatorname{ess\,sup}_{x \in B} \Phi(x, t) &\leq \operatorname{ess\,inf}_{x \in B} \Phi(x, t) + \left\{ [a_s(B)]^{1+\sigma} - [a_i(B)]^{1+\sigma} \right\} t^{q(1+\sigma)} \\ &\leq \operatorname{ess\,inf}_{x \in B} \Phi(x, t) + c[r(B)]^{\alpha(1+\sigma)} t^{q(1+\sigma)} \\ &\leq \operatorname{ess\,inf}_{x \in B} \Phi(x, t) + c[r(B)]^{(\alpha-q+p)(1+\sigma)} t^{p(1+\sigma)} \leq c \operatorname{ess\,inf}_{x \in B} \Phi(x, t) \end{aligned} \quad (6.1)$$

where $c \equiv c([a]_{0, \alpha}, \sigma)$. Therefore (1.32) is satisfied for $\beta_4 = 1$ and assertion (1.34) follows by (1.33).

6.2 Proof of Theorem 2

The proof is a suitable modification of the one which is valid for the standard $W^{1,p}$ -capacity. Thanks to the Choquet property (1.36), we can reduce to the case when E is a compact subset. Therefore, recalling that $\operatorname{Cap}_{\Phi}^*(K) = \operatorname{Cap}_{\Phi}(K)$ whenever $K \subset \Omega$ is a compact subset, we can then compute $\operatorname{Cap}_{\Phi}(E)$ via (1.35). We now claim that there exists a positive constant c , essentially depending on E , such that if V is a bounded open set such that $E \Subset V \subset \Omega$, then there exists an open set W and a function $f \in \mathcal{R}(E)$ with the following features:

$$\begin{cases} E \subset W \subset \{x \in \Omega : f(x) = 1\} , & \operatorname{supp} f \subset V , f \in C_0(\Omega); \\ \int_{\Omega} \Phi(x, |Df|) dx < c . \end{cases} \quad (6.2)$$

Let $V \subset \Omega$ be an open set as above and fix $\kappa = \frac{1}{4} \min \{\operatorname{dist}(E, \mathbb{R}^n \setminus V), 1\}$. Since $\mathcal{H}_{\Phi}(E) < \infty$ and E is compact, there exists a positive integer $m = m(E)$ and a finite collection of open balls $\{B_{r_j}(x_j)\}_{j \leq m} \in C_E^{\kappa/2}$ such that $\{x_j\}_{j \leq m} \subset E$, $B_{2r_j}(x_j) \Subset \Omega$ for all $j \in \{1, \dots, m\}$ and

$$B_{r_j}(x_j) \cap E \neq \emptyset \text{ for all } j \in \{1, \dots, m\} \quad \text{and} \quad \sum_{j=1}^m \int_{B_{r_j}(x_j)} \Phi(x, 1/r_j) dx \leq 2 [\mathcal{H}_{\Phi}(E) + 1] . \quad (6.3)$$

We introduce $W := \bigcup_{j=1}^m B_{r_j}(x_j)$ and the maps f_j , defined on the whole $\mathbb{R}^n$, such that $f_j(x) := 1$ if $|x - x_j| \leq r_j$, $f_j(x) := 2 - |x - x_j|/r_j$ if $r_j < |x - x_j| \leq 2r_j$ and $f_j(x) := 0$ if $2r_j < |x - x_j|$. With $\beta_4 \in (0, 1)$ being the constant appearing in (1.32), we have

$$\begin{aligned} \int_{\Omega} \Phi(x, |Df_j|) dx &\leq \int_{B_{2r_j}(x_j)} \Phi(x, 1/r_j) dx \\ &\stackrel{(1.37)}{\leq} \frac{c_g}{\beta_4^q} \int_{B_{2r_j}(x_j)} \Phi(x, \beta_4/r_j) dx \\ &\leq \frac{2^n c_g}{\beta_4^q} |B_{r_j}(x_j)| \operatorname{ess\,sup}_{x \in B_{2r_j}(x_j)} \Phi(x, \beta_4/r_j) \\ &\stackrel{(1.32)}{\leq} \frac{2^n c_g c_d}{\beta_4^q} |B_{r_j}(x_j)| \operatorname{ess\,inf}_{x \in B_{r_j}(x_j)} \Phi(x, 1/r_j) \leq \frac{2^n c_g c_d}{\beta_4^q} \int_{B_{r_j}(x_j)} \Phi(x, 1/r_j) dx . \end{aligned}$$

In particular, it follows that $f_j \in W^{1,\Phi}(\Omega)$. We then set

$$f := \max_{j \in \{1, \dots, m\}} f_j ,$$

which is continuous, as every f_j is. Moreover, if $x \in W$, then $x \in B_{r_j}(x_j)$ for some $j \in \{1, \dots, m\}$ and, as a consequence, $f_j(x) = 1$; thus $f(x) = 1$, so $f(W) \equiv \{1\}$ and using the content of the last display, the lattice property of Sobolev functions, see [37, Theorem 1.20] and (6.3), we obtain

$$\begin{aligned} \int_{\mathbb{R}^n} \Phi(x, |Df|) dx &\leq \sum_{j=1}^m \int_{\Omega} \Phi(x, |Df_j|) dx \\ &\leq c \sum_{j=1}^m \int_{B_{r_j}(x_j)} \Phi(x, 1/r_j) dx \stackrel{(6.3)}{\leq} c [\mathcal{H}_{\Phi}(E) + 1] , \end{aligned}$$

with $c = c(n, \tilde{c}_g, c_d, \beta_4, q)$, hence $f \in W^{1,\Phi}(\Omega) \cap C_0(\Omega)$. Finally, observe that if $x \in \mathbb{R}^n \setminus V$, then $|x - x_j| \geq \text{dist}(E, \mathbb{R}^n \setminus V) \geq 4\kappa$, so $f(x) = 0$ and $\text{supp } f \subset V$. This completes the proof of (6.2). Using the above construction inductively, for any $k \in \mathbb{N}$ we find a collection of open sets $\{V_k\}_k$, with $V_0 = \emptyset$ and a sequence of functions $\{\tilde{f}_k\}_k$ such that

$$\begin{cases} E \subset V_{k+1} \subset V_k , \quad \bar{V}_{k+1} \subset \{x \in \Omega : \tilde{f}_k(x) = 1\} , \quad \text{supp } \tilde{f}_k \subset V_k \\ \int_{\Omega} \Phi(x, |D\tilde{f}_k|) dx \leq c_* , \end{cases} \quad (6.4)$$

with c_* being independent of $k \in \mathbb{N}$. For $j \in \mathbb{N}$, define

$$S_j = \sum_{k=1}^j \frac{1}{k} \quad \text{and} \quad g_j = \frac{1}{S_j} \sum_{k=1}^j \frac{\tilde{f}_k}{k} .$$

From the above discussion, $\tilde{f}_j$ belongs to $\mathcal{R}(E)$ for every j , so, by construction, $g_j \in \mathcal{R}(E)$, and, given that $\text{supp } |D\tilde{f}_k| \subset V_k \setminus \bar{V}_{k+1}$, we find

$$\text{Cap}_{\Phi}(E) \leq \int_{\mathbb{R}^n} \Phi(x, |Dg_j|) dx = \sum_{k=1}^j \int_{V_k \setminus V_{k+1}} \Phi(x, S_j^{-1} k^{-1} |D\tilde{f}_k|) dx \stackrel{(1.37), (6.4)}{\leq} \frac{c_* \tilde{c}_g}{S_j^p} \sum_{k=1}^j \frac{1}{k^p} \rightarrow 0$$

as $j \rightarrow \infty$, because $p > 1$. The proof is complete.

6.3 Proof of Theorem 3

The assertions (1.39)-(1.40) have already been proved in Step 3 from the proof of Theorem 1; see (5.37)-(5.38). We therefore proceed with the proof of (1.38). We abbreviate $\Sigma_u = \Sigma_u^p \cup \Sigma_u^q$ where $\Sigma_u^p := \Sigma_u \cap \{x_0 \in \Omega : a(x_0) = 0\}$ and $\Sigma_u^q := \Sigma_u \cap \{x_0 \in \Omega : a(x_0) > 0\}$. It is therefore sufficient to show

$$\mathcal{H}_{H^{1+\delta_g}}(\Sigma_u^p) = 0 \quad \text{and} \quad \mathcal{H}_{H^{1+\delta_g}}(\Sigma_u^q) = 0 . \quad (6.5)$$

The implication concerning the capacity in (1.38) will then be a consequence of Theorem 2. To prove (6.5), we use (1.34) from Proposition 1, that gives, in particular, that $\mathcal{H}_{H^{1+\delta_g}} \lesssim \mathcal{H}_{H^{1+\delta_g}}^-$. On the other hand, by the very definition of Σ_u^p , we have that

$$\mathcal{H}_{H^{1+\delta_g}}^-(\Sigma_u^p) \lesssim \mathcal{H}^{n-p-\delta_g}(\Sigma_u^p) . \quad (6.6)$$

Indeed, taking a covering from $C_{\Sigma_u^p}^\kappa$ for any $\kappa \in (0, 1)$ as in (1.26), we see that every ball B of the covering (that for obvious reasons can be assumed to touch Σ_u^p) is such that $a_i(B) = 0$. Therefore $\mathcal{H}_{H^{1+\delta_g}}^-(\Sigma_u^p)$ is equivalent to the $(n - p - p\delta_g)$ -dimensional spherical Hausdorff measure of Σ_u^p and (6.6) is proved. We conclude that the first equality in (6.5) follows from the already proved fact that $\mathcal{H}^{n-p-p\delta_g}(\Sigma_u^p) = 0$. We now prove the second equality in (6.5). For this we show that $\mathcal{H}_{H^{1+\delta_g}}(\Sigma_{u,m}^q) = 0$ for every integer m , where $\Sigma_{u,m}^q := \Sigma_u^q \cap \{a(x) > 1/m\}$. The key observation is that there exists a positive number $\tilde{\kappa} \equiv \tilde{\kappa}(\alpha, [a]_{0,\alpha}, m) \in (0, 1)$ such that, if $r(B) < \tilde{\kappa}$ and B touches $\Sigma_{u,m}^q$, then $2ma_i(B) > 1$. Therefore, we take a covering $\{B_j\} \in C_{\Sigma_{u,m}^q}^\kappa$ (again we assume that each of the balls from the covering is touching $\Sigma_{u,m}^q$) with $\kappa < \tilde{\kappa}$, and estimate as follows

$$\begin{aligned} & \sum_{j \in \mathbb{N}} [r(B_j)]^n \left\{ [r(B_j)]^{-p} + a_i(B_j)[r(B_j)]^{-q} \right\}^{1+\delta_g} \\ & \leq cm^{1+\delta_g} \sum_{j \in \mathbb{N}} [r(B_j)]^n \left\{ a_i(B_j)[r(B_j)]^{-p} + a_i(B_j)[r(B_j)]^{-q} \right\}^{1+\delta_g} \\ & \leq cm^{1+\delta_g} \|a\|_{L^\infty}^{1+\delta_g} \sum_{j \in \mathbb{N}} [r(B_j)]^{n-q(1+\delta_g)}. \end{aligned}$$

The above relation implies then that

$$\mathcal{H}_{H^{1+\delta_g}}^-(\Sigma_{u,m}^q) \lesssim \mathcal{H}^{n-q-q\delta_g}(\Sigma_{u,m}^q) \leq \mathcal{H}^{n-q-q\delta_g}(\Sigma_u^q) \implies \mathcal{H}_{H^{1+\delta_g}}^-(\Sigma_{u,m}^q) = 0$$

by (1.40) for every positive integer m . By again appealing to (1.34) we deduce that $\mathcal{H}_{H^{1+\delta_g}}(\Sigma_{u,m}^q) = 0$ for every positive integer m and the proof is complete.

Acknowledgements. This work is supported by the Engineering and Physical Sciences Research Council [EP/L015811/1]. The authors thank I. Chlebicka and J. Ok for comments on a preliminary version of the manuscript.

References

- [1] Acerbi E. & Mingione G.: Regularity results for electrorheological fluids: the stationary case. *C. R. Math. Acad. Sci. Paris* 334 (2002), 817–822.
- [2] Baroni P. & Colombo M. & Mingione G.: Harnack inequalities for double phase functionals. *Nonlinear Anal.* 121 (2015), 206–222.
- [3] Baroni P. & Colombo M. & Mingione G.: Non-autonomous functionals, borderline cases and related function classes. *St. Petersburg Math. J.* 27 (2016), 347–379.
- [4] Baroni P. & Colombo M. & Mingione G.: Regularity for general functionals with double phase. *Calc. Var. & PDE* 57 (2018), Paper No. 62, 48 pages.
- [5] Baruah D. & Harjulehto P. & Hästö P.: Capacities in generalized Orlicz spaces. *J. Funct. Spaces* 2018, Article ID 8459874.
- [6] Byun S.S. & Oh J.: Global gradient estimates for non-uniformly elliptic equations. *Calc. Var. & PDE* 56 (2017), no. 2, Paper No. 46, 36 pp.
- [7] Cencelj M. & Radulescu V. & Repovš D.D.: Double phase problems with variable growth. *Nonlinear Anal.* 177 (2018), 270–287.

- [8] Chlebicka I. & De Filippis C.: Removable sets in non-uniformly elliptic problems. *Ann. Mat. Pura Appl.* (IV), to appear.
- [9] Colombo M. & Mingione G.: Regularity for double phase variational problems. *Arch. Rat. Mech. Anal.* 215 (2015), 443–496.
- [10] Colombo M. & Mingione G.: Bounded minimisers of double phase variational integrals. *Arch. Rat. Mech. Anal.* 218 (2015), 219–273.
- [11] Colombo M. & Mingione G.: Calderón-Zygmund estimates and non-uniformly elliptic operators. *J. Funct. Anal.* 270 (2016), 1416–1478.
- [12] Choquet G.: Theory of capacities. *Ann. Inst. Fourier, Grenoble* 5 (1955), 131–295.
- [13] De Filippis C.: Higher integrability for constrained minimizers of integral functionals with (p, q) -growth in low dimension. *Nonlinear Analysis* 170 (2018), 1–20.
- [14] De Filippis C.: On the regularity of the ω -minima of φ -functionals. *Nonlinear Anal.* <https://doi.org/10.1016/j.na.2019.02.017>.
- [15] De Filippis C.: Partial regularity for manifold constrained $p(x)$ -harmonic maps. *Calc. Var. & PDE*, 58:47, (2019) <https://doi.org/10.1007/s00526-019-1483-6>.
- [16] De Filippis C. & Mingione G.: A borderline case of Calderón-Zygmund estimates for non-uniformly elliptic problems. *St. Petersburg Math. J.*, to appear
- [17] De Filippis C. & Oh J.: Regularity for multi-phase variational problems. *J. Diff. Equ.*, 267 (2019), 1631–1670.
- [18] De Filippis C. & Palatucci G.: Hölder regularity for nonlocal double phase equations. *J. Diff. Equ.*, 267 (2019), 547–586.
- [19] Diening L. & Harjulehto P. & Hästö P. & Růžička M.: *Lebesgue and Sobolev spaces with a variable growth exponent*. Springer Lecture notes Math. 2017 (2011).
- [20] Diening L. & Stroffolini B. & Verde A.: Everywhere regularity of functionals with φ -growth. *Manuscripta Math.* 129 (2012), 449–481.
- [21] Diening L. & Stroffolini B. & Verde A.: The φ -harmonic approximation and the regularity of φ -harmonic maps. *J. Diff. Equ.* 253 (2012), 1943–1958.
- [22] Duerinckx M. & Gloria A.: Stochastic homogenization of nonconvex unbounded integral functionals with convex growth. *Arch. Ration. Mech. Anal.* 221 (2016), 1511–1584.
- [23] Duzaar F. & Mingione G.: The p -harmonic approximation and the regularity of p -harmonic maps. *Calc. Var. & PDE* 20 (2004), 235–256.
- [24] Eells J. & Sampson J. H.: Harmonic mappings of Riemannian manifolds. *Amer. J. Math.* 86 (1964), 109–160.
- [25] Esposito L. & Leonetti F. & Mingione G.: Sharp regularity for functionals with (p, q) growth. *J. Diff. Equ.* 204 (2004), 5–55.
- [26] Evans L. C. & Gariepy R. F.: *Measure theory and fine properties of functions*. CRC Press, (1991).

- [27] Fonseca I. & Malý J. & Mingione G.: Scalar minimizers with fractal singular sets. *Arch. Rat. Mech. Anal.* 172 (2004), 295–307.
- [28] Frehse J.: Capacity methods in the theory of partial differential equations. *Jahresber. Deutsch. Math.-Verein.* 84 (1982), 1–44.
- [29] Fuchs M.: p -harmonic obstacle problems. I: Partial regularity theory. *Ann. Mat. Pura Appl. (IV)* 156 (1990), 127–158.
- [30] Fuchs M.: p -harmonic obstacle problems. III. Boundary regularity. *Ann. Mat. Pura Appl. (IV)* 156 (1990), 159–180.
- [31] Giusti E.: *Direct methods in the calculus of variations*. World Scientific Publishing Co., Inc., River Edge, NJ, 2003.
- [32] Hardt R.M. & Lin F. G.: Mappings minimizing the L^p norm of the gradient. *Comm. Pure Appl. Math.* 40 (1987), no. 5, 555–588.
- [33] Hardt R.M. & Kinderlehrer D. & Lin F. G.: Stable defects of minimizers of constrained variational principles. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 5 (1988), 297–322.
- [34] Harjulehto P. & Hästö P.: *Orlicz Spaces and Generalized Orlicz Spaces*. Book, 2018.
- [35] Harjulehto P. & Hästö P. & Toivanen O.: Hölder regularity of quasiminimizers under generalized growth conditions. *Calc. Var. & PDE* 56 (2017), Paper No. 22, 26 pp.
- [36] Hästö P.: The maximal operator on generalized Orlicz spaces. *J. Funct. Anal.* 269 (2015), 4038–4048.
- [37] Heinonen J. & Kilpeläinen T. & Martio O.: *Nonlinear potential theory of degenerate elliptic equations*. Dover Publications Inc., Mineola, NY, 2006. Unabridged republication of the 1993 original.
- [38] Hopper C.: Partial regularity for holonomic minimizers of quasiconvex functionals. *Arch. Rational Mech. Anal.* 222 (2016), 91–141.
- [39] Kristensen J. & Mingione G.: The singular set of minima of integral functionals. *Arch. Rat. Mech. Anal.* 180 (2006), 331–398.
- [40] Kristensen J. & Mingione G.: The singular set of Lipschitzian minima of multiple integrals. *Arch. Rat. Mech. Anal.* 184 (2007), 341–369.
- [41] Kuusi T. & Mingione G.: Guide to nonlinear potential estimates. *Bull. Math. Sci.* 4 (2014), 1–82.
- [42] Kuusi T. & Mingione G.: Vectorial nonlinear potential theory. *J. Europ. Math. Soc. (JEMS)* 20 (2018), 929–1004.
- [43] Leonetti F. & Siepe F.: Maximum Principle for Vector Valued Minimizers. *J. Convex Anal.* 12 (2005), 267–278.
- [44] Luckhaus S.: Partial Hölder continuity for minima of certain energies among maps into Riemannian manifold. *Indiana Univ. Math. J.* 37 (1988), 349–367.
- [45] Manfredi J.J.: Regularity of the gradient for a class of nonlinear possibly degenerate elliptic equations. *Ph.D. Thesis*. University of Washington, St. Louis, 1986.

- [46] Marcellini P.: Regularity of minimizers of integrals of the calculus of variations with non standard growth conditions. *Arch. Rat. Mech. Anal.* 105 (1989), 267–284.
- [47] Marcellini P.: Regularity and existence of solutions of elliptic equations with p, q -growth conditions. *J. Diff. Equ.* 90 (1991), 1–30.
- [48] Mazowiecka M. & Miśkiewicz M. & Schikorra A.: On the size of the singular set of minimizing harmonic maps into the 2-sphere in dimension four and higher. *Arxiv preprint*, 2019
- [49] Mazya V.: *Sobolev spaces with applications to elliptic partial differential equations. Second, revised and augmented edition.* Grundlehren der Mathematischen Wissenschaften, 342. Springer, Heidelberg, 2011. xxviii+866 pp
- [50] Mingione G.: Regularity of minima: an invitation to the Dark Side of the Calculus of Variations. *Appl. Math.* 51 (2006), 355–425.
- [51] Nieminen E.: Hausdorff measures, capacities, and Sobolev spaces with weights. *Ann. Acad. Sci. Fenn. Ser. A I Math. Dissertationes* No. 81 (1991), 39 pages.
- [52] Ok J.: Gradient estimates for elliptic equations with $L^{p(\cdot)} \log L$ growth. *Calc. Var. & PDE* 55:26 (2016).
- [53] Ok J.: Regularity of ω -minimizers for a class of functionals with non-standard growth. *Calc. Var. & PDE* 56:48 (2017).
- [54] Papageorgiou N. & Radulescu V. & Repovš D.: Double-phase problems with reaction of arbitrary growth. *Z. Angew. Math. Phys.* 69 (2018), Art. 108, 21 pp.
- [55] Radulescu V. and Repovš D.: *Partial differential equations with variable exponents: variational methods and qualitative analysis.* Chapman & Hall/CRC Monographs and Research Notes in Mathematics. 2015
- [56] Ragusa M.A. & Tachikawa A.: Boundary regularity of minimizers of $p(x)$ -energy functionals. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 33 (2017), 451–476.
- [57] Ragusa M.A. & Tachikawa A.: Partial regularity of $p(x)$ -harmonic maps. *Trans. Amer. Math. Soc.* 365 (2013), 3329–3353.
- [58] Schoen R. & Uhlenbeck K.: A regularity theory for harmonic maps. *J. Diff. Geom.* 17 (1983), 307–336.
- [59] Schoen R. & Uhlenbeck K.: Boundary regularity and the Dirichlet problem for harmonic maps. *J. Diff. Geom.* 18 (1983), 253–268.
- [60] Simon L.: *Lectures on regularity and singularities of harmonic maps.* Birkhuser-Verlag, Basel-Boston-Berlin 1996.
- [61] Turesson B.O.: *Nonlinear potential theory and weighted Sobolev spaces.* Lecture Notes in Math., 1736. Springer-Verlag, Berlin, 2000. xiv+173.
- [62] Uhlenbeck K.: Regularity for a class of non-linear elliptic systems. *Acta Math.* 138 (1977), 219–240.
- [63] Ural'tseva N.N.: Degenerate quasilinear elliptic systems. *Zap. Na. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)* 7 (1968), 184–222.

- [64] Ural'tseva N.N. & Urdaletova A.B.: The boundedness of the gradients of generalized solutions of degenerate quasilinear non-uniformly elliptic equations. *Vestnik Leningrad Univ. Math.* 19 (1983) (russian) english. tran.: 16 (1984), 263-270.
- [65] Zhikov V.V.: Averaging of functionals of the calculus of variations and elasticity theory. *Izv. Akad. Nauk SSSR Ser. Mat.* 50 (1986), 675–710.
- [66] Zhikov V.V.: On Lavrentiev's Phenomenon. *Russian J. Math. Phys.* 3 (1995), 249–269.
- [67] Zhikov V.V.: On some variational problems. *Russian J. Math. Phys.* 5 (1997), 105–116.