

# REGULARITY FOR GENERAL FUNCTIONALS WITH DOUBLE PHASE

PAOLO BARONI, MARIA COLOMBO, AND GIUSEPPE MINGIONE

ABSTRACT. We prove sharp regularity results for a general class of functionals of the type

$$w \mapsto \int F(x, w, Dw) dx ,$$

featuring non-standard growth conditions and non-uniform ellipticity properties. The model case is given by the double phase integral

$$w \mapsto \int b(x, w)(|Dw|^p + a(x)|Dw|^q) dx , \quad 1 < p < q, \quad a(x) \geq 0 ,$$

with  $0 < \nu \leq b(\cdot) \leq L$ . This changes its ellipticity rate according to the geometry of the level set  $\{a(x) = 0\}$  of the modulating coefficient  $a(\cdot)$ . We also present new methods and proofs, that are suitable to build regularity theorems for larger classes of non-autonomous functionals. Finally, we disclose some new interpolation type effects that, as we conjecture, should draw a general phenomenon in the setting of non-uniformly elliptic problems. Such effects naturally connect with the Lavrentiev phenomenon.

*To Paolo Marcellini on his 70th birthday, with admiration for  
his pioneering work in the Calculus of Variations*

## CONTENTS

1.	Introduction and results	1
2.	Notation and preliminaries	7
3.	First regularity results	9
4.	Initial setting for Theorems 1.1-1.2	11
5.	Quantitative approximation	11
6.	Two phases	17
7.	$p$ -phase	18
8.	$(p, q)$ -phase	30
9.	Morrey decay, exit times and Theorem 1.2	32
10.	Gradient continuity and Theorem 1.1	35
11.	Interpolative effects and Theorem 1.3	37
12.	Lavrentiev phenomenon and Theorem 1.4	40
	References	41

## 1. INTRODUCTION AND RESULTS

The aim of this paper is to provide sharp and comprehensive regularity theorems for minimizers of a class of integral functionals of the Calculus of Variations exhibiting a degeneracy of double type and a strong non-uniform ellipticity. The main model case in question here is provided by the double phase functional

$$(1.1) \quad W^{1,1}(\Omega) \ni w \mapsto \mathcal{P}(w, \Omega) := \int_{\Omega} (|Dw|^p + a(x)|Dw|^q) dx ,$$

which is naturally defined for  $w \in W^{1,1}(\Omega)$ , where  $\Omega \subset \mathbb{R}^n$  is a bounded open domain,  $n \geq 2$  and it is initially assumed that  $1 < p \leq q$  and  $a(\cdot) \in L^\infty(\Omega)$ . In the recent papers [5, 16, 17] the authors have provided a rather comprehensive regularity theory for local minimizers of  $\mathcal{P}$ . This culminates in establishing the local gradient Hölder continuity of minima provided a sharp balancing condition between the closeness of  $p$  and  $q$  and the regularity of  $a(\cdot)$  is satisfied. Notice that the Hölder gradient regularity is optimal already in the classical  $p$ -Laplacean case, i.e., when  $a(\cdot) \equiv 0$  [53, 54]. More specifically, it turns out that the condition

$$(1.2) \quad a(\cdot) \in C^{0,\alpha}(\Omega), \quad \alpha \in (0, 1] \quad \text{and} \quad \frac{q}{p} \leq 1 + \frac{\alpha}{n}$$

is sufficient to get such maximal regularity. Actually, this has been proved in [16] apart from the controversial borderline case  $q/p = 1 + \alpha/n$ , which is instead handled here for the first time; see also Remark 1.1 below. Moreover, as proved in [17], assuming that minimizers are bounded allows to relax the relation between  $p$  and  $q$  in (1.2) as follows:

$$(1.3) \quad u \in L^\infty(\Omega), \quad a(\cdot) \in C^{0,\alpha}(\Omega), \quad \alpha \in (0, 1] \quad \text{and} \quad q \leq p + \alpha.$$

The counterexamples given in [26, 27] show that both (1.2) and (1.3) are sharp conditions for regularity. They do coincide for  $p = n$ , while (1.2) naturally gets better than (1.3) for  $p > n$ . Indeed, in this last case the boundedness assumption on  $u$  becomes irrelevant due to Sobolev-Morrey embedding theorem. Notice that such conditions essentially serve to contain the rate of non-uniform ellipticity of the functional  $\mathcal{P}$ , which can be measured by the distance  $q - p$ . We also remark that the first regularity results for minima of  $\mathcal{P}$  under assumptions (1.2), namely the higher integrability and fractional differentiability of gradient, have been obtained in [26] via the analysis of the related Lavrentiev phenomenon the functional  $\mathcal{P}$  might exhibit. The functional appearing in (1.1) has been first considered by Zhikov [56, 57, 58] in the setting of Homogenization theory and to give new instances of the Lavrentiev phenomenon. In this respect, the main feature of the integrand

$$(1.4) \quad H(x, z) := |z|^p + a(x)|z|^q =: K(x, |z|) \quad z \in \mathbb{R}^n$$

is that it changes its rate of ellipticity according to the positivity of  $a(x)$ . It shows  $q$ -growth with respect to the gradient variable  $z$  on the set  $\{a(x) > 0\}$ , and  $p$ -growth on  $\{a(x) = 0\}$ . The delicate transitions between the  $p$ - and the  $q$ -ellipticity explains the dependence of the bounds in (1.2)-(1.3) linking  $p$  and  $q$  on the Hölder exponent  $\alpha$  of  $a(\cdot)$ . We refer to [6, 16] for a more accurate description of the additional features of the integral in (1.1) and its occurrence in applications.

In this paper we shall deal with general functionals of the type

$$(1.5) \quad W^{1,1}(\Omega) \ni w \mapsto \mathcal{F}(w, \Omega) := \int_{\Omega} F(x, w, Dw) \, dx,$$

where  $F: \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$  is a Carathéodory integrand, initially satisfying the double-sided bound

$$(1.6) \quad \nu H(x, z) \leq F(x, v, z) \leq LH(x, z),$$

for constants  $0 < \nu \leq L$ . This aims at giving an intrinsic approach to regularity of double phase functional, drawing a parallel with the standard growth conditions of  $p$ -type, that is when  $a(x) \equiv 0$  and therefore  $H(x, z) \equiv |z|^p$ . Under growth conditions (1.6) the natural notion of minimality is given as follows:

**Definition 1.** A function  $u \in W_{\text{loc}}^{1,1}(\Omega)$  is a local minimiser of the functional  $\mathcal{F}$  defined in (1.5) if and only if  $H(\cdot, Du) \in L^1(\Omega)$  and the minimality condition  $\mathcal{F}(u, \text{supp}(u - v)) \leq \mathcal{F}(v, \text{supp}(u - v))$  is satisfied whenever  $v \in W_{\text{loc}}^{1,1}(\Omega)$  is such that  $\text{supp}(u - v) \subset \Omega$ . In particular, a local minimizer belongs to  $W^{1,p}(\Omega)$ .

Integrals of the type in (1.5) under the growth assumptions (1.6) belong to the family of functionals satisfying non-standard growth conditions of  $(p, q)$ -type, i.e., those whose integrands satisfy growth and coercivity conditions at different polynomial rates

$$(1.7) \quad \nu|z|^p \leq F(x, v, z) \leq L(|z|^q + 1), \quad p < q.$$

These functionals have been the object of intensive investigation over the last years, starting with the seminal papers of Marcellini [43, 44, 45]; see for instance [9, 11, 15, 26, 55, 57, 58]. See also the survey [46]. The regularity theory in the case of functionals with  $(p, q)$ -growth in the non-autonomous case, that is when the energy density  $F(\cdot)$  is allowed to depend on  $x$ , is still a non-trivial open issue, especially when considering sharp conditions for regularity. In this respect, the growth conditions considered in (1.6) represent a very significant case since the one in (1.1) provides one of the hardest model examples of non-autonomous functionals with  $(p, q)$ -growth available in the literature. See for instance the lists of examples discussed in [6, 26, 56, 57, 58].

In order to get higher regularity of minima, assumption (1.6) are not sufficient already in the standard case  $a(x) \equiv 0$ . For this, after (1.6), we assume that  $F(\cdot)$  is a continuous integrand, of class  $C_{\text{loc}}^2(\mathbb{R}^n \setminus \{0\})$  in the  $z$ -variable, and we consider the following natural assumptions:

$$(1.8) \quad \left\{ \begin{array}{l} |\partial F(x, v, z)||z| + |\partial^2 F(x, v, z)||z|^2 \leq LH(x, z) \\ \nu(|z|^{p-2} + a(x)|z|^{q-2})|\xi|^2 \leq \langle \partial^2 F(x, v, z)\xi, \xi \rangle \\ |\partial F(x_1, v, z) - \partial F(x_2, v, z)||z| \leq L\omega(|x_1 - x_2|)[H(x_1, z) + H(x_2, z)] \\ \quad + L|a(x_1) - a(x_2)||z|^q \\ |F(x, v_1, z) - F(x, v_2, z)| \leq L\omega(|v_1 - v_2|)H(x, z). \end{array} \right.$$

These are assumed to hold whenever  $x, x_1, x_2 \in \Omega$ ,  $v, v_1, v_2 \in \mathbb{R}$ ,  $z \in \mathbb{R}^n \setminus \{0\}$ ,  $\xi \in \mathbb{R}^n$ , where  $0 < \nu \leq L$  are fixed constants, and

$$(1.9) \quad \omega(t) := \min \{t^\beta, 1\} \quad \text{for } t > 0$$

is the standard  $\beta$ -Hölder continuous modulus of continuity for  $\beta \in (0, 1]$ . In (1.8) the symbol  $\partial$  stands for the partial derivative with respect to the  $z$ -variable. It is not difficult to see that assumptions (1.8) are for instance satisfied by the model functional

$$(1.10) \quad W^{1,1}(\Omega) \ni w \mapsto \int_{\Omega} b(x, w)H(x, Dw) dx,$$

where  $0 < \nu_1 \leq b(x, v) \leq L_1$ , for some constants  $\nu_1, L_1$  and for some  $b(\cdot)$  Hölder continuous function. Further model cases are clearly given by integrals of the type

$$(1.11) \quad w \mapsto \int_{\Omega} [F_1(x, w, Dw) + a(x)F_2(x, w, Dw)] dx,$$

where  $F_1(\cdot)$  and  $F_2(\cdot)$  have  $p$ - and  $q$ -growth, respectively, and satisfy conditions suited to imply Hölder continuity of the gradient of minima when considered as single integrands. See for instance [33, Assumptions (1.1)] or just consider (1.8) with  $p = q$ . We remark that both the functional in (1.10) and the one in (1.11) cannot be covered by the present literature on functionals with  $(p, q)$ -growth as in (1.7). Note indeed that in both cases the functional considered is non-differentiable and therefore its treatment cannot pass through the analysis of the related Euler-Lagrange equation.

Sometimes we shall replace (1.9) by the weaker

$$(1.12) \quad \omega: [0, \infty) \rightarrow [0, \infty) \text{ is concave and such that } \omega(0) = 0 \text{ and } \omega(\cdot) \leq 1,$$

thereby considering functionals as in (1.10) with  $b(\cdot)$  being just continuous rather than Hölder continuous.

The first two main results of this paper, stated in the next Theorem 1.1 and 1.2, draw a complete parallel with the theory of functionals with standard polynomial growth [28, 41, 42] i.e., when  $a(\cdot) \equiv 0$ ; see also [35, 36] for more recent a priori estimates.

**Theorem 1.1** (Maximal regularity). *Let  $u \in W^{1,p}(\Omega)$  be a local minimiser of the functional  $\mathcal{F}$  defined in (1.5), under the assumptions (1.6) and (1.8)-(1.9). Moreover, assume that either (1.2) or (1.3) are satisfied. Then there exists  $\beta_0 \in (0, 1)$ , depending only on  $n, p, q, \nu, L, \alpha$  and  $\beta$ , such that  $Du \in C_{\text{loc}}^{0,\beta_0}(\Omega; \mathbb{R}^n)$ .*

The main contribution in this paper is essentially threefold. First, we extend the results of [16, 17] to the largest possible family of functionals exhibiting a double phase behaviour of the type in (1.1), that is those functionals that can be controlled by the one in (1.1) in the sense of (1.6); we further observe that we provide a unified and more transparent approach to the two different cases (1.2)-(1.3). Second, the methods we develop here can provide a guideline to face the more general problem of regularity of minima under general Orlicz-type conditions. This means dealing with general non-autonomous functionals of the type in (1.5) where

$$(1.13) \quad F(x, v, z) \approx \Phi(x, |z|)$$

and  $\Phi(\cdot)$  is a generalized Young function in the sense specified for instance in [19, 32]. An example is in fact given by the function  $H(x, z) = K(x, |z|) \equiv \Phi(x, |z|)$  defined in (1.4). The study of such problems has gained large attention over the last years, in particular with respect to the sharp interplay between the regularity of the function  $x \mapsto \Phi(x, \cdot)$  and the growth conditions of  $t \mapsto \Phi(\cdot, t)$ . An example of this is indeed already given in (1.2)-(1.3), while further results in this direction can be for instance found in [2, 3, 4, 8, 11, 9, 12, 13, 14, 20, 47, 48, 49, 52]; see also the work of Ragusa & Tachikawa for partial regularity [50, 51, 52]. In particular, we point out the recent paper [31], where a general setting for conditions (1.13) has been considered, with some unifying approaches and assumptions fitting several different contexts, as for instance those described in [6, 26, 56, 57, 58]. The issue is intriguing as assumptions that are relevant for regularity largely coincide with those that are necessary to get good functional theoretic properties of the spaces in questions; see for instance [6, 16, 19] for this interplay. In this paper we use an intrinsic approach, useful to deal with cases like (1.13). An instance of this is the Morrey type decay estimate (1.14) below, formulated in terms of the natural quantity  $H(\cdot, Du)$  and resembling the classical one valid for  $p$ -harmonic functions. It requires assumptions that are weaker than those considered in Theorem 1.1.

**Theorem 1.2** (Intrinsic Morrey decay). *Let  $u \in W^{1,p}(\Omega)$  be a local minimiser of the functional  $\mathcal{F}$  defined in (1.5), under the assumptions (1.6), (1.8) and (1.12). Moreover, assume that either (1.2) or (1.3) are satisfied. Then,  $u \in C_{\text{loc}}^{0,\theta}(\Omega)$  for every  $\theta < 1$ . Finally, for every  $\sigma \in (0, n)$ , there exists a positive constant  $c \equiv c(\mathbf{data}(\Omega_0), \sigma)$ , such that the decay estimate*

$$(1.14) \quad \int_{B_\rho} H(x, Du) dx \leq c \left( \frac{\rho}{R} \right)^{n-\sigma} \int_{B_R} H(x, Du) dx$$

*holds whenever  $B_\rho \subset B_R \Subset \Omega_0$  are concentric balls with  $R \leq 1$ .*

The meaning of  $\mathbf{data}(\Omega_0)$  is clarified in (2.2) below, see Section 2 for more notation. Theorem 1.2 has been obtained for the model case (1.1) in [16] under assumptions (1.2) (but only when  $q/p < 1 + \alpha/n$ ), but is new already for  $\mathcal{P}$  when conditions in (1.3) are considered. Another advantage of the methods presented

here is that they allow to avoid the use of fractional estimates previously employed in [16, 17, 26]. The use of such methods has a few drawbacks. As an example, it does not allow to extend the interior regularity results up to the boundary under the general assumptions one may wish to consider thinking to the classical case with  $p$ -growth. See for instance the recent interesting paper [12], where, due to the use of fractional estimates, additional regularity assumptions on the boundary have to be imposed in order to extend the interior results in [18].

**Remark 1.1.** The approach proposed here also allows to deal with the delicate borderline case in (1.2), i.e.,

$$(1.15) \quad \frac{q}{p} = 1 + \frac{\alpha}{n} .$$

This has been left open in several papers on non-autonomous functionals [9, 12, 16, 25, 48]. Borderline cases are hard to catch. For instance, in the case of general functionals of the type  $w \mapsto \int_{\Omega} F(Dw) dx$  with  $(p, q)$ -conditions of the type in (1.7) and suitable convexity assumptions, the bound available in the literature is  $q/p < 1 + 2/n$ . It is not clear how to deal with the related limiting case  $q/p = 1 + 2/n$ .

**Remark 1.2.** Another feature of the methods introduced here is that they open the way to treat parabolic equations of the type

$$u_t - \operatorname{div} (|Du|^{p-2} Du + a(x, t)|Du|^{q-2} Du) = 0 .$$

Such equations poses non-trivial additional difficulties as they generate new *double phase intrinsic geometries* and new methods must be developed; see [7].

The third goal of this paper is finally to disclose a new interpolative type phenomenon. In fact, we conjecture this should be a general principle when considering functionals with  $(p, q)$ -growth as in (1.7) and related non-uniformly elliptic problems. For this we shall consider simpler functionals of the type

$$(1.16) \quad W^{1,1}(\Omega) \ni w \mapsto \mathcal{F}(w, \Omega) := \int_{\Omega} F(x, Dw) dx ,$$

but the results are completely new already in the model case (1.1). The idea is that assuming more regularity of  $u$  allows to further relax the bound linking  $p$  and  $q$ . This is already visible in (1.3) vs (1.2). The right scale to further quantify this phenomenon is the one of Hölder continuity, as shown in the next

**Theorem 1.3** (New interpolative bound). *Let  $u \in W^{1,p}(\Omega)$  be a local minimiser of the functional  $\mathcal{F}$  defined in (1.16), under the assumptions (1.6), (1.8)-(1.9) and with  $a(\cdot) \in C^{0,\alpha}(\Omega)$ . Moreover assume that*

$$(1.17) \quad u \in C^{0,\gamma}(\Omega) \quad \text{and} \quad q < p + \frac{\alpha}{1-\gamma}, \quad \gamma \in (0, 1) .$$

*Then there exists  $\beta_1 \in (0, 1)$ , depending only on  $n, p, q, \nu, L, \alpha$  and  $\beta$ , such that  $Du \in C_{\text{loc}}^{0,\beta_1}(\Omega; \mathbb{R}^n)$ . Moreover, weakening assumption (1.9) by (1.12), leads to the conclusions of Theorem 1.2.*

The bound in (1.17) exhibits the correct asymptotic. It formally reduces to (1.3) when  $\gamma \rightarrow 0$ , while says that no bound on  $q - p$  is needed when  $\gamma \rightarrow 1$ . Indeed, in this case we approach the Lipschitz continuity of  $u$ , which in fact makes the functional uniformly elliptic at infinity. As expected, the bound in (1.17) improves the one in (1.2) only when  $p < n/(1 - \gamma)$ , that is, only when the exponent  $\gamma$  is better than the one naturally given by Sobolev-Morrey embedding theorem, since  $p > n/(1 - \gamma)$  implies that  $u \in C_{\text{loc}}^{0,\gamma}(\Omega)$ . Only in this case  $u \in C^{0,\gamma}(\Omega)$  becomes a real assumption and gives therefore an improvement.

Finally, the assumptions of Theorem 1.3 also allow to give new instances of absence of Lavrentiev phenomenon [37, 56, 57, 58]. According to the classical definition specialized to the present context, the Lavrentiev phenomenon for the functional  $\mathcal{F}$  in (1.5) under assumptions (1.6) occurs when

$$(1.18) \quad \inf_{w \in u_0 + W_0^{1,p}(B)} F(w, B) < \inf_{w \in u_0 + W_0^{1,p}(B) \cap W_{\text{loc}}^{1,q}(B)} F(w, B),$$

where  $B \Subset \Omega$  is a ball and  $u_0 \in W^{1,\infty}(B)$ . Considering the model functional  $\mathcal{P}$  in (1.1), as shown in [26], it can be proved that (1.18) actually occurs when (1.2)-(1.3) fail, and this eventually allows to give examples of irregular minima; see also [27]. In [26] this approach is reversed and the absence of Lavrentiev phenomenon - i.e., the non-occurrence of (1.18) - is used to prove regularity of minima. In particular, in [26] the absence of Lavrentiev phenomenon for general functionals as in (1.5) is proved under assumptions (1.2) and (1.6). When instead looking at (1.3), the absence of Lavrentiev phenomenon for the model case  $\mathcal{P}$  has been obtained in [17] as a consequence of the available higher regularity of minima (not available for functionals as in (1.5) under conditions (1.6)). Next theorem features a more general result, under less general assumptions already in the case (1.3) is considered. Indeed, it only assumes (1.6).

**Theorem 1.4** (Absence of Lavrentiev phenomenon). *Let  $u \in W^{1,p}(\Omega)$  be a local minimiser of the functional  $\mathcal{F}$  defined in (1.5), under assumptions (1.6) and with  $a(\cdot) \in C^{0,\alpha}(\Omega)$ . Assume that either (1.3) or*

$$(1.19) \quad u \in C^{0,\gamma}(\Omega) \quad \text{and} \quad q \leq p + \frac{\alpha}{1-\gamma}, \quad \gamma \in (0, 1)$$

*holds. Then, for every ball  $B \Subset \Omega$ , there exists a sequence  $\{u_k\}$  of  $W^{1,\infty}(B)$ -regular functions such that  $u_k \rightarrow u$  strongly in  $W^{1,p}(B)$  and such that*

$$(1.20) \quad \lim_k \mathcal{F}(u_k, B) = \mathcal{F}(u, B).$$

We finally conclude giving an outline of the proofs, starting with the one of Theorem 1.1. As already mentioned, a main new fact, allowing to catch the borderline case (1.15), is that we are avoiding the use of fractional estimates and actually of any differentiation of the Euler-Lagrange equation of the functional (1.5) (fractional or standard). These are very common and heavy tools in the setting of functionals with non-standard growth conditions (see for instance [11, 9, 12, 16, 17, 25, 26]), that in fact do not allow to reach optimal assumptions in several situations. We instead employ a suitable blow-up argument relying on a quantitative version of a certain nonlinear harmonic type approximation lemma (Lemma 5.1 below) that, in the original linear version, goes back to the classical work of De Giorgi [21]; see [24] for an overview. This allows a more efficient separation of phases, following the terminology introduced in [16, 17]. Specifically, on smaller and smaller scales/balls  $B_R$ , we shall distinguish between the  $p$ -phase, where an inequality of the type  $a(x) \lesssim [a]_{0,\alpha} R^\alpha$  holds, and the  $(p, q)$ -phase, that is when  $[a]_{0,\alpha} R^\alpha \ll a(x)$  instead occurs; see Section 6. In the  $p$ -phase we blow-up the original minimizer  $u$  from Theorem 1.1 in  $B_R$ , and see that, in a sense, it behaves as a minimizer of a functional with standard  $p$ -growth. From this we infer certain regularity estimates for  $u$  in  $B_R$ . This is done in Section 7. In the  $(p, q)$ -phase, instead, we use a direct freezing argument and see that  $u$  behaves in  $B_R$  as a minimizer of an anisotropic functional whose growth is controlled by  $w \mapsto \int_{B_R} [|Dw|^p + a_0 |Dw|^q] dx$ , for some constant  $a_0 > 0$ . From this we again infer some regularity estimates in  $B_R$ ; see Section 8. It is important to see that, in order to accelerate the blow-up, and especially to catch the borderline case (1.15), we have to use a few preliminary higher

integrability and Hölder continuity results (see Theorems 3.2-3.1 below). Some non-linear Calderón-Zygmund estimates in a non-standard setting are also employed (see Theorem 5.1 below). We then combine the treatment of the two phases via an exit-time argument described in Section 9. This leads to Theorem 1.2. We finally look at the gradient regularity. To achieve a good control of the constants we use a different separation of phases. We indeed consider a  $p$ -phase this time defined by  $a(x) \lesssim [a]_{0,\alpha} R^{\alpha-s}$  and a  $(p, q)$ -phase which is still of the type  $[a]_{0,\alpha} R^\alpha \ll a(x)$ . A suitable choice of the number  $s \in [0, \alpha)$  and of the constants involved will eventually determine the Hölder continuity exponent  $\beta_0$  of  $Du$  appearing in Theorem 1.1. The proof of Theorem 1.3 takes conceptually more effort and requires a double application of the blow-up lemma. The first time, this will be done in order to by-pass the fact that the assumed Hölder continuity in (1.3) is not quantitatively preserved under blow-up. This does not provide us with the higher integrability result (Theorem 3.1 below) allowing to apply Lemma 5.1 as for Theorem 1.1. For this we proceed in two steps. First we apply a more traditional, non-quantitative version of the harmonic approximation to get higher Hölder continuity of  $u$ . Once this is done, we recover the missing Hölder continuity in the blow-up procedure. We can then proceed as for Theorem 1.2, with the quantitative version of the harmonic approximation.

## 2. NOTATION AND PRELIMINARIES

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 emphasised using parentheses, i.e.,  $c_1 \equiv c_1(n, p, \nu, L)$  means that  $c_1$  depends on  $n, p, \nu, L$ . 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  $r > 0$ ; 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, when not otherwise stated, different balls in the same context will share the same center. We shall also denote  $B_1 = B_1(0)$  if not differently specified. Finally, with  $B$  being a given ball with radius  $r$  and  $\gamma$  being a positive number, we denote by  $\gamma B$  the concentric ball with radius  $\gamma r$ . With  $\mathcal{B} \subset \mathbb{R}^n$  being a measurable subset with 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. With  $f: \Omega \rightarrow \mathbb{R}$  and  $\mathcal{B} \subset \Omega$ , with  $\gamma \in (0, 1)$  being a given number, we shall denote

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

**Remark 2.1.** We note that, since all the results from Theorems 1.1-1.4 are local in nature, when considering assumptions (1.3) and (1.17), we can also assume that  $u \in L_{\text{loc}}^\infty(\Omega)$  and  $u \in C_{\text{loc}}^{0,\gamma}(\Omega)$ . This can be done up to passing to open subsets  $\Omega_0 \Subset \Omega$  and restating all the results with an additional dependence of the constants on  $\text{dist}(\Omega_0, \partial\Omega)$ . In the same way, with  $u$  being the minimizer of Theorems 1.1-1.3, we can assume that  $H(\cdot, Du) \in L_{\text{loc}}^1(\Omega)$ . We prefer to start with the global formulations for the ease of exposition.

In order to shorten the notation, we shall express the dependence of the constants on the various basic parameters using the symbol **data**. This is defined as a set of

objects that will vary according to the assumptions considered as follows

$$(2.1) \quad \mathbf{data} \equiv \begin{cases} n, p, q, \nu, L, \alpha, \omega(\cdot), [a]_{0,\alpha}, \|H(\cdot, Du)\|_{L^1(\Omega)} & \text{when (1.2) holds} \\ n, p, q, \nu, L, \alpha, \omega(\cdot), [a]_{0,\alpha}, \|u\|_{L^\infty(\Omega)} & \text{when (1.3) holds} \\ n, p, q, \nu, L, \alpha, \gamma, \omega(\cdot), [a]_{0,\alpha}, [u]_{0,\gamma} & \text{when (1.17) holds.} \end{cases}$$

We also need a local version of  $\mathbf{data}$ . For this, with  $\Omega_0 \Subset \Omega$  being a fixed open subset, we denote by  $\mathbf{data}(\Omega_0)$  the above set of parameters in addition to  $\text{dist}(\Omega_0, \partial\Omega)$ :

$$(2.2) \quad \mathbf{data}(\Omega_0) \equiv \mathbf{data}, \text{dist}(\Omega_0, \partial\Omega) .$$

**Remark 2.2.** We shall sometimes consider functionals of the type in (1.5) under the only assumptions (1.6). In this case  $\mathbf{data}$  and  $\mathbf{data}(\Omega_0)$  omit the specifications of  $\omega(\cdot)$ , since (1.8) are not considered.

In the following we shall often deal with the vector field

$$(2.3) \quad V_t(z) := |z|^{(t-2)/2} z, \quad t \in \{p, q\},$$

so that  $V_p(z) := |z|^{(p-2)/2} z$  and  $V_q(z) := |z|^{(q-2)/2} z$ . This vector field is of common use to formulate the monotonicity properties of operators of  $p$ -Laplacian type and related integral functionals (see for instance [33, 34]). In this respect we record the following property

$$(2.4) \quad |V_t(z_1) - V_t(z_2)|^2 \leq c \langle |z_1|^{t-2} z_1 - |z_2|^{t-2} z_2, z_1 - z_2 \rangle,$$

We shall also use the following equivalence:

$$(2.5) \quad |V_t(z_1) - V_t(z_2)| \approx (|z_1| + |z_2|)^{(t-2)/2} |z_1 - z_2|,$$

that holds with involved constants depending only on  $n, t$  (see for instance [33, 34]).

Finally, we summarize some basic terminology about so called generalized Orlicz-Sobolev spaces. Roughly speaking, these are Sobolev spaces defined by the fact that the distributional derivatives lie in a suitable Orlicz space (actually, a Orlicz-Musielak space according to the terminology in use [22]) rather than a Lebesgue spaces, as usual. Classical Sobolev spaces are then a particular case. Such spaces, and related variational problems, are for instance discussed in [19, 22, 32], to which we refer for more details. Here we shall simply consider spaces related to the Young type function defined in (1.4). We then define

$$(2.6) \quad W^{1,H}(\Omega) := \{u \in W^{1,1}(\Omega) : H(\cdot, Du) \in L^1(\Omega)\},$$

with the local variant being defined in the obvious way and  $W_0^{1,H}(\Omega) = W^{1,H}(\Omega) \cap W_0^{1,p}(\Omega)$ . The one in (2.6) is in a sense the natural energy space associated to variational problems under growth conditions (1.6); see for instance [31]. Finally, with  $a_0 \geq 0$  being a non-negative number, we shall use the frozen Young function

$$(2.7) \quad H_0(z) := |z|^p + a_0 |z|^q =: K_0(|z|),$$

with the notation fixed accordingly to the one used in (1.4). Then  $W^{1,H_0}(\Omega)$ , and its related variants, can be defined exactly as in (2.6) with the choice  $a(x) \equiv a_0$ . Note that  $[0, \infty) \ni s \mapsto K_0(s) \in [0, \infty)$  is a strictly convex, monotone smooth bijection of  $[0, \infty)$ . Denoting by  $\widetilde{K}_0$  its Young's (or convex) conjugate, i.e.,  $\widetilde{K}_0(s) := \sup_{\tau > 0} [s\tau - K_0(\tau)]$ , then the following standard property

$$(2.8) \quad \widetilde{K}_0\left(\frac{K_0(s)}{s}\right) \leq K_0(s)$$

holds for every  $s > 0$  (see [8] for more details). We shall later use the following Young type inequality, valid for every  $\varepsilon \in (0, 1)$  and  $s, t \geq 0$ :

$$(2.9) \quad st \leq \frac{K_0(s)}{\varepsilon^{q-1}} + \varepsilon \widetilde{K}_0(t).$$

We briefly report the proof. For  $0 < \kappa \leq 1$  and for all  $s \geq 0$ , since  $q \geq p$ , we have

$$\begin{aligned} \widetilde{K}_0(\kappa s) &= \sup_{\tau > 0} (\kappa s \tau - K_0(\tau)) \\ &\leq \sup_{\tau > 0} \left( \kappa^{\frac{q}{q-1}} s \kappa^{-\frac{1}{q-1}} \tau - \kappa^{\frac{q}{q-1}} K_0(\kappa^{-\frac{1}{q-1}} \tau) \right) \\ &= \sup_{\tilde{\tau} = \kappa^{-1/(q-1)} \tau > 0} \left( \kappa^{\frac{q}{q-1}} s \tilde{\tau} - \kappa^{\frac{q}{q-1}} K_0(\tilde{\tau}) \right) \\ &= \kappa^{\frac{q}{q-1}} \widetilde{K}_0(s). \end{aligned}$$

Thus, with  $\varepsilon \in (0, 1)$ , we have the standard Young's inequality for conjugate functions yields  $st \leq K_0(\varepsilon^{(1-q)/q} s) + \widetilde{K}_0(\varepsilon^{(q-1)/q} t)$  and (2.9) follows using the property in the above display and again that  $q \geq p$  while  $\varepsilon \in (0, 1)$ .

### 3. FIRST REGULARITY RESULTS

We review a few basic regularity results available for minimizers of functionals with double phase. The main references here are [5, 16, 17, 48] and the results are essentially proved there. We shall treat minimizers of functionals as in (1.5) under the only growth assumptions in (1.6). Here as in the rest of the paper, we shall denote by  $\Omega_0$  an open subset such that  $\Omega_0 \Subset \Omega$ . We start by a Caccioppoli type inequality from [16].

**Proposition 3.1.** *Let  $u \in W^{1,p}(\Omega)$  be a local minimiser of the functional  $\mathcal{F}$  defined in (1.5), under the assumptions (1.6), with  $a(\cdot) \in L^\infty_{\text{loc}}(\Omega)$  and  $u \in L^q_{\text{loc}}(\Omega)$ . Then there exists a constant depending only on  $n, p, q, \nu, L$  such that*

$$(3.1) \quad \int_{B_t} H(x, Du) dx \leq c \int_{B_s} K \left( x, \left| \frac{u - u_0}{s - t} \right| \right) dx,$$

holds whenever  $B_t \Subset B_s \Subset \Omega$  are concentric balls and  $u_0 \in \mathbb{R}$ , where  $K(\cdot)$  has been introduced in (1.4). In the same way, the following related inequality on level sets

$$(3.2) \quad \int_{B_t} H(x, D(u - k)_\pm) dx \leq c \int_{B_s} K \left( x, \left| \frac{(u - k)_\pm}{s - t} \right| \right) dx$$

holds for every  $k \in \mathbb{R}$ ,  $t < s$ , where  $c \equiv c(n, p, q, \nu, L)$  and

$$(3.3) \quad (u - k)_+ := \max\{u - k, 0\} \quad \text{and} \quad (u - k)_- := \max\{k - u, 0\}.$$

Notice that the condition  $u \in L^q_{\text{loc}}(\Omega)$  is always satisfied when one of the assumptions (1.2), (1.3) and (1.17) is in force.

Next, a higher integrability result.

**Theorem 3.1.** *Let  $u \in W^{1,p}(\Omega)$  be a local minimiser of the functional  $\mathcal{F}$  defined in (1.5), under the assumptions (1.6). Moreover, assume that one of the conditions in (1.2), (1.3) and (1.17) is satisfied. Then there exists a higher integrability exponent  $\delta \equiv \delta(\mathbf{data}) \in (0, 1)$  such that  $H(\cdot, Du) \in L^{1+\delta}_{\text{loc}}(\Omega)$ . Furthermore, the reverse type Hölder inequality*

$$(3.4) \quad \left( \int_{B_{R/2}} [H(x, Du)]^{1+\delta} dx \right)^{1/(1+\delta)} \leq c \int_{B_R} H(x, Du) dx$$

holds for a constant  $c \equiv c(\mathbf{data})$ , whenever  $B_R \Subset \Omega$  is a ball with  $R \leq 1$ .

*Proof.* This has already been obtained in [16, 17, 48] as far as (1.2)-(1.3) are used. Here we cover the missing case of (1.17), essentially recalling the arguments for [17, Theorem 1.2], to which we refer for more details. Consider a ball  $B_R \Subset \Omega$  with

$R \leq 1$  and consider the quantity  $a_i(B_R) := \min_{\overline{B_R}} a(x)$ . If  $a_i(B_R) > 4[a]_\alpha R^\alpha$  then, exactly as in [17, Theorem 1.2], we find that the reverse type inequality

$$(3.5) \quad \int_{B_{R/2}} H(x, Du) dx \leq c \left( \int_{B_R} [H(x, Du)]^d dx \right)^{1/d},$$

holds for  $d := \max\{n/(n+p), 1/p\} \in (0, 1)$ , where the constant  $c$  depends only on  $n, p, q, \nu, L$ . We now consider the remaining case  $a_i(B_R) \leq 4[a]_\alpha R^\alpha$ , where we estimate as follows:

$$\begin{aligned} a(x) \left| \frac{u - (u)_{B_R}}{R} \right|^{q-p} &\leq 8[a]_{0,\alpha} R^{p-q+\alpha} [\text{osc}_{B_R} u]^{q-p} \\ &\stackrel{(1.17)}{\leq} c[u]_{0,\gamma;B_R} R^{p-q+\alpha+\gamma(q-p)} \stackrel{(1.17)}{\leq} c(\mathbf{data}). \end{aligned}$$

The last inequality allows to conclude

$$\int_{B_R} a(x) \left| \frac{u - (u)_{B_R}}{R} \right|^q dx \leq c \int_{B_R} \left| \frac{u - (u)_{B_R}}{R} \right|^p dx$$

and combining this with (3.1) (with  $s = R$ ,  $t = R/2$ ,  $u_0 = (u)_{B_R}$ ), we get, by Sobolev-Poincaré inequality

$$\begin{aligned} \int_{B_{R/2}} H(x, Du) dx &\leq c \int_{B_R} \left| \frac{u - (u)_{B_R}}{R} \right|^p dx \\ &\leq c \left( \int_{B_R} |Du|^{pd} dx \right)^{1/d} \leq c \left( \int_{B_R} [H(x, Du)]^{pd} dx \right)^{1/d}, \end{aligned}$$

where  $c \equiv c(\mathbf{data})$  and  $d$  is the same appearing in (3.5). We conclude that  $H(\cdot, Du)$  satisfies a reverse type Hölder inequality, that is (3.5) holds for every ball  $B_R \Subset \Omega$  such that  $R \leq 1$ . At this point (3.4) follows using a variant of Gehring's lemma on reverse Hölder inequalities [29, Theorem 6.6].  $\square$

Finally, we collect some Hölder continuity assertions from [5, 16, 48]. Special emphasis is put on the precise dependence of the various constants.

**Theorem 3.2.** *Let  $u \in W^{1,p}(\Omega)$  be a local minimiser of the functional  $\mathcal{F}$  defined in (1.5), under the assumptions (1.2) and (1.6). Then  $u$  is locally Hölder continuous. Moreover, for every open subset  $\Omega_0 \Subset \Omega$ , there exists a Hölder continuity exponent*

$$(3.6) \quad \gamma \equiv \gamma(n, p, q, \nu, L, \alpha, [a]_{0,\alpha}, \|u\|_{L^\infty(\Omega_0)}) \in (0, 1)$$

such that

$$(3.7) \quad \|u\|_{L^\infty(\Omega_0)} + [u]_{0,\gamma;\Omega_0} \leq c(\mathbf{data}(\Omega_0)),$$

and the oscillation estimate

$$(3.8) \quad \text{osc}_{B_\varrho} u \leq c \left( \frac{\varrho}{r} \right)^\gamma \text{osc}_{B_r} u$$

holds for  $c \equiv c(n, p, q, \nu, L, \alpha, [a]_{0,\alpha}, \|u\|_{L^\infty(\Omega_0)})$  and all concentric balls  $B_\varrho \Subset B_r \Subset \Omega_0 \Subset \Omega$  with  $r \leq 1$ . Finally, the dependence of the exponent in (3.6) can be reformulated as

$$(3.9) \quad \gamma \equiv \gamma(n, p, q, \nu, L, \alpha, [a]_{0,\alpha}, \|H(\cdot, Du)\|_{L^1(\Omega)}, \text{dist}(\Omega_0, \partial\Omega)) \in (0, 1).$$

The same conclusions hold when (1.3) is assumed instead of (1.2); in this case the  $L^\infty$ -estimate in (3.7) becomes immaterial as  $u$  is assumed to be globally bounded. Moreover, the exponent  $\gamma$  depends only on  $\mathbf{data}$  and it is independent of the open subset  $\Omega_0 \Subset \Omega$  considered.

*Proof.* The results in [16, Section 10] give that if  $\|u\|_{L^\infty(\Omega)}$  is finite and  $q \leq p + \alpha$ , then  $u \in C_{\text{loc}}^{0,\gamma}$  and (3.7)-(3.8) hold. In this case  $\gamma$  depends only on the parameters  $n, p, q, \nu, L, \alpha, [a]_{0,\alpha}, \|u\|_{L^\infty(\Omega)}$  and no dependence on  $\text{dist}(\Omega_0, \partial\Omega)$  occurs. Moreover, estimate (3.8) follows from [48, Theorem 3.8] or directly from the Harnack inequality proved in [5]. In this respect, see how to derive estimates of the type in (3.8) from Harnack inequalities in [29, Section 7.9]. This fully covers the case when (1.3) comes into the play. We then consider (1.2), when  $p \leq n$ . By the results in [5, 16] and in particular by [48, Theorem 3.6], minimizers are locally bounded. Therefore, since  $q \leq p + \alpha$  is implied by the bound in  $q/p \leq 1 + \alpha/n$  when  $p \leq n$ , then we can reduce to the case when  $u$  is bounded and this gives (3.7)-(3.8) again; moreover, the dependence displayed in (3.6) follows in any case. To recast the alternative dependence in (3.9), we recall that the local bounds proved in [16, 48], together with standard covering arguments, imply an estimate of the type

$$\|u\|_{L^\infty(\Omega_0)} \leq c \left( \int_{\Omega} [H(x, Du) + 1] dx \right)^{1/p},$$

where  $c$  depends on  $n, p, q, \nu, L, \alpha, [a]_{0,\alpha}, \|H(\cdot, Du)\|_{L^1(\Omega)}$  and  $\text{dist}(\Omega_0, \partial\Omega)$ . Combining this fact with the dependence in (3.6) yields the required dependence in (3.9). Finally, when  $p > n$ , we can use the Harnack inequality proved in [5], and this yields again (3.8) for a constant  $c$  and an exponent  $\gamma$ , being independent of  $\text{dist}(\Omega_0, \partial\Omega)$ , but just depending on  $n, p, q, \nu, L, \alpha, [a]_{0,\alpha}, \|H(\cdot, Du)\|_{L^1(\Omega)}$ . Notice that in the Harnack statements from [5] we have that the constants also depend on the diameter of  $\Omega$ . This dependence does not occur when one restricts the Harnack inequalities on balls  $B_R$  with  $R \leq 1$ , which is the thing needed here in order to get (3.8).  $\square$

#### 4. INITIAL SETTING FOR THEOREMS 1.1-1.2

In this section we begin the proofs of Theorems 1.1-1.2 and fix the initial setting aimed at treating the cases (1.2)-(1.3) in a unified way. We first observe that  $p + \alpha < p(1 + \alpha/n)$  iff  $p > n$ , so that (1.2) provides a weaker assumption than (1.17) when  $p > n$ . On the other hand, in this last case minimizers are automatically bounded. When  $p = n$  we have that the relations between  $p$  and  $q$  in (1.2)-(1.3) do coincide and therefore (1.2) becomes again an assumption weaker than (1.3), since it does not require that minimizers are a priori bounded. In this respect, notice that assuming  $p = n \geq q - \alpha$  gives that minimizers are bounded by Theorem 3.2. On the contrary, in the case  $p < n$ , assumptions (1.2) imply (1.3) and we can always reduce to consider (1.3). Indeed when  $q/p \leq 1 + \alpha/n$ , minimizers are always locally bounded again by Theorem 3.2, and  $q/p \leq 1 + \alpha/n$  implies that  $q \leq p + \alpha$ . Summarizing, we shall always make a distinction in the forthcoming proofs: we shall consider the case (1.3) holds when  $p < n$  holds, and the case when (1.2) holds but only for  $p \geq n$  (and vice-versa). Notice that this is perfectly consistent with the notation in (2.1)-(2.2). In fact, when  $p < n$  and (1.3) covers also the case (1.2), it happens that by Theorem 3.2 we can locally bound  $\|u\|_{L^\infty}$  by  $\|H(\cdot, Du)\|_{L^1}$ , so that the final dependence is on this last quantity, exactly as prescribed in (2.1) when (1.2) is considered. See the proof of Theorem 3.2 for more.

#### 5. QUANTITATIVE APPROXIMATION

Harmonic type approximation lemmas give a way to perform blow-up procedures without passing to the limit. They are in use since the seminal work of De Giorgi [21], and we refer to [24] for an overview of the subject. For recent non-standard versions related to the setting of this paper we refer to [23], from which we will also borrow the direct approach used here (vs the indirect one presented in [21, 24]).

In this section we give a version where the control of the constants can be made explicit and becomes of polynomial type. This is the effect of assuming initial higher integrability (see (5.9) below). In the following we shall deal with a general vector field  $A_0: \mathbb{R}^n \rightarrow \mathbb{R}^n$ , which is assumed to be  $C_{\text{loc}}^1(\mathbb{R}^n \setminus \{0\})$ -regular and satisfying the following growth and ellipticity assumptions:

$$(5.1) \quad \begin{cases} |A_0(z)||z| + |\partial A_0(z)||z|^2 \leq LH_0(z) \\ \nu \frac{H_0(z)}{|z|^2} |\xi|^2 \leq \langle \partial A_0(z)\xi, \xi \rangle, \end{cases}$$

whenever  $z \in \mathbb{R}^n \setminus \{0\}$ ,  $\xi \in \mathbb{R}^n$ , where  $0 < \nu \leq 1 \leq L$  are fixed constants. The function  $H_0(\cdot)$  is the one defined in (2.7) for some  $a_0 \geq 0$ . No condition is considered here on the two exponents  $1 < p < q$ , which can be arbitrarily far from each other. We shall in the following use the strict monotonicity property:

$$(5.2) \quad |V_p(z_1) - V_p(z_2)|^2 + a_0 |V_q(z_1) - V_q(z_2)|^2 \leq c \langle A_0(z_1) - A_0(z_2), z_1 - z_2 \rangle$$

that holds for a constant  $c \equiv c(n, p, q, \nu) \geq 1$ , and for every  $z_1, z_2 \in \mathbb{R}^n$ ; the maps  $V_p(\cdot)$  and  $V_q(\cdot)$  are defined in (2.3). This is easily seen to be a consequence of (5.1)<sub>2</sub> and (2.5) (see [16, 33]). Related to the vector field  $A_0(\cdot)$ , we consider the following Dirichlet boundary value problem, which is naturally defined in the Sobolev space  $W^{1, H_0}(B)$ :

$$(5.3) \quad \begin{cases} -\operatorname{div} A_0(Dh) = 0 & \text{in } B \\ h \in v + W_0^{1, H_0}(B) \end{cases}$$

where  $B \subset \mathbb{R}^n$  is a given ball and  $v \in W^{1, H_0}(B)$  is a given boundary datum. Solutions are meant in the usual distributional sense. Such a problem has already been considered in [18, Section 5], to which we refer for existence. We then report the following Calderón-Zygmund type result:

**Theorem 5.1** ([18], Theorem 5.1). *Let  $h \in W^{1,1}(B)$  be a distributional solution to (5.3) such that  $H_0(Dv), H_0(Dh) \in L^1(B)$ , under the assumptions (5.1). Then*

$$H_0(Dv) \in L^t(B) \implies H_0(Dh) \in L^t(B) \quad \text{for every } t > 1.$$

Moreover, for every  $t > 1$ , there exists a constant  $c \equiv c(n, p, q, \nu, L, t)$ , which is in particular independent of  $a_0$ , such that the following inequality holds:

$$(5.4) \quad \int_B [H_0(Dh)]^t dx \leq c \int_B [H_0(Dv)]^t dx.$$

Before stating the main result of this section we still need another fact, namely, a by now classical truncation lemma due to Acerbi & Fusco [1]. The statement involves the Hardy-Littlewood maximal operator, defined as follows

$$(5.5) \quad M(f)(x) := \sup_{B_r(x) \subset \mathbb{R}^n} \int_{B_r} |f(y)| dy, \quad x \in \mathbb{R}^n,$$

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

**Theorem 5.2** ([1]). *Let  $B \subseteq \mathbb{R}^n$  be a ball and  $w \in W_0^{1,1}(B)$ . Then for every  $\lambda > 0$  there exists  $w_\lambda \in W_0^{1,\infty}(B)$  such that*

$$(5.6) \quad \|Dw_\lambda\|_{L^\infty(B)} \leq c\lambda$$

for some constant  $c$  depending only on  $n$ . Moreover, it holds that

$$(5.7) \quad \{w_\lambda \neq w\} \subseteq B \cap \{M(|\nabla w|) > \lambda\} \cup \text{negligible set}.$$

Note that, in view of the definition used in (5.5), in the above theorem we can assume  $w$  to be defined on the whole  $\mathbb{R}^n$  by setting  $w \equiv 0$  outside  $B$ .

We are now ready to state the main result of this section. We remark in advance that a crucial point is that all the constants must be independent of the number  $a_0$  appearing in (2.7).

**Lemma 5.1** (Harmonic type approximation). *Let  $B \equiv B_r \subseteq \mathbb{R}^n$  be a ball, let  $\sigma \in (0, 1)$  and let  $v \in W^{1,p}(2B)$  be a function satisfying the following estimates:*

$$(5.8) \quad \int_{2B} H_0(Dv) \, dx \leq \tilde{c}_1$$

and

$$(5.9) \quad \int_B [H_0(Dv)]^{1+\delta_0} \, dx \leq \tilde{c}_2 ,$$

where  $\tilde{c}_1, \tilde{c}_2 \geq 1$  and  $\delta_0 > 0$  are fixed constants. Moreover, assume that

$$(5.10) \quad \left| \int_B \langle A_0(Dv), D\varphi \rangle \, dx \right| \leq \sigma \|D\varphi\|_{L^\infty(B)} \quad \text{holds for all } \varphi \in C_0^\infty(B) .$$

Then there exists a function  $h \in v + W_0^{1,H_0}(B)$  such that the following conditions are satisfied:

$$(5.11) \quad \int_B \langle A_0(Dh), D\varphi \rangle \, dx = 0 \quad \text{holds for all } \varphi \in C_0^\infty(B),$$

$$(5.12) \quad \int_B [H_0(Dh)]^{1+\delta_0} \, dx \leq c(n, p, q, \nu, L, \delta_0) \tilde{c}_2 ,$$

$$(5.13) \quad \int_B (|V_p(Dv) - V_p(Dh)|^2 + a_0 |V_q(Dv) - V_q(Dh)|^2) \, dx \leq c \sigma^{s_1} ,$$

and

$$(5.14) \quad \int_B \left( \left| \frac{v-h}{r} \right|^p + a_0 \left| \frac{v-h}{r} \right|^q \right) \, dx \leq c \sigma^{s_0} .$$

In (5.13) and (5.14) the dependence of constants involved is as follows:  $s_1 := s_1(p, q, \delta_0) > 0$ ,  $s_0 := s_0(n, p, q, \delta_0) > 0$  and  $c \equiv c(n, p, q, \nu, L, \delta_0, \tilde{c}_1, \tilde{c}_2) \geq 1$ .

*Proof.* By a standard approximation argument we notice that, if (5.10) holds for every  $\varphi \in C_0^\infty(B)$ , then it holds also for every  $\varphi \in W_0^{1,\infty}(B)$ . We then divide the proof in three steps.

*Step 1: Truncation.* We define  $h \in W^{1,p}(B)$  as in (5.3). The standard energy estimate in this case (see [16, Theorem 3.1]) and (5.8) give

$$(5.15) \quad \int_B H_0(Dh) \, dx \leq c \int_B H_0(Dv) \, dx \leq c(n, p, q, \nu, L) \tilde{c}_1$$

for  $c \equiv c(n, p, q, \nu, L)$ . Similarly, but using first (5.4) and then (5.9), we get

$$(5.16) \quad \int_B [H_0(Dh)]^{1+\delta_0} \, dx \leq c \int_B [H_0(Dv)]^{1+\delta_0} \, dx \leq c(n, p, q, \nu, L, \delta_0) \tilde{c}_2 ,$$

and this proves (5.12). We now set  $w := v - h \in W_0^{1,H_0}(B)$  and we let  $\lambda \geq 1$  to be chosen later; we consider  $w_\lambda \in W_0^{1,\infty}(B)$  given by Theorem 5.2, which satisfies (5.6) and (5.7). Using such properties, Chebyshev's inequality and finally the maximal theorem, we deduce that

$$\frac{|\{w \neq w_\lambda\}|}{|B|} \leq \frac{|B \cap \{M(|Dw|) > \lambda\}|}{|B|}$$

$$\begin{aligned}
&\leq \frac{1}{[K_0(\lambda)]^{1+\delta_0}} \int_B [K_0(|M(Dw)|)]^{1+\delta_0} dx \\
&\leq \frac{c(n, p, q, \delta_0)}{[K_0(\lambda)]^{1+\delta_0}} \int_B [H_0(Dw)]^{1+\delta_0} dx \\
&\leq \frac{c}{[K_0(\lambda)]^{1+\delta_0}} \left[ \int_B [H_0(Dv)]^{1+\delta_0} dx + \int_B [H_0(Dh)]^{1+\delta_0} dx \right] \\
(5.17) \quad &\leq \frac{c \tilde{c}_2}{[K_0(\lambda)]^{1+\delta_0}}
\end{aligned}$$

with  $c \equiv c(n, p, q, \nu, L, \delta_0)$  and in the last line we have used (5.12) and (5.16). Now we test the weak formulation of (5.3)<sub>1</sub> with  $w_\lambda$  to get

$$\begin{aligned}
\mathcal{T}_1 &:= \int_B \langle A_0(Dv) - A_0(Dh), Dw_\lambda \rangle \chi_{\{w=w_\lambda\}} dx \\
&= \int_B \langle A_0(Dv), Dw_\lambda \rangle dx \\
(5.18) \quad &- \int_B \langle A_0(Dv) - A_0(Dh), Dw_\lambda \rangle \chi_{\{w \neq w_\lambda\}} dx =: \mathcal{T}_2 + \mathcal{T}_3.
\end{aligned}$$

Next, we estimate each term in the previous equality. By (5.2) we have

$$\mathcal{T}_1 \geq \frac{1}{c} \int_B (|V_p(Dv) - V_p(Dh)|^2 + a_0 |V_q(Dv) - V_q(Dh)|^2) \chi_{\{w=w_\lambda\}} dx$$

for  $c \equiv c(n, p, q, \nu)$ . Using (5.10) and then (5.6) we have

$$|\mathcal{T}_2| \leq \sigma \|Dw_\lambda\|_{L^\infty(B)} \leq c \sigma \lambda.$$

Finally, for  $\mathcal{T}_3$ , we fix  $\varepsilon \in (0, 1)$  to be chosen and estimate

$$\begin{aligned}
|\mathcal{T}_3| &\leq \int_B (|A_0(Dh)| + |A_0(Dv)|) |Dw_\lambda| \chi_{\{w \neq w_\lambda\}} dx \\
(5.1) \quad &\leq L \|Dw_\lambda\|_{L^\infty(B)} \int_B \left[ \frac{H_0(Dv)}{|Dv|} + \frac{H_0(Dh)}{|Dh|} \right] \chi_{\{w \neq w_\lambda\}} dx \\
(2.7) \quad &\stackrel{=}{=} L \|Dw_\lambda\|_{L^\infty(B)} \int_B \left[ \frac{K_0(|Dv|)}{|Dv|} + \frac{K_0(|Dh|)}{|Dh|} \right] \chi_{\{w \neq w_\lambda\}} dx \\
(2.9) \quad &\leq \varepsilon \int_B \left[ \widetilde{K}_0 \left( \frac{K_0(|Dv|)}{|Dv|} \right) + \widetilde{K}_0 \left( \frac{K_0(|Dh|)}{|Dh|} \right) \right] dx \\
&\quad + \frac{cK_0(\|Dw_\lambda\|_{L^\infty(B)})}{\varepsilon^{q-1}} \frac{|\{w \neq w_\lambda\}|}{|B|} \\
(2.8), (5.6) \quad &\leq \varepsilon \int_B [H_0(Dv) + H_0(Dh)] dx + \frac{cK_0(\lambda)}{\varepsilon^{q-1}} \frac{|\{w \neq w_\lambda\}|}{|B|} \\
(5.17) \quad &\leq \varepsilon \int_B [H_0(Dv) + H_0(Dh)] dx + \frac{c}{[K_0(\lambda)]^{\delta_0} \varepsilon^{q-1}}.
\end{aligned}$$

By further using (5.8), (5.15) and that  $K_0(\lambda) \geq \lambda^p$ , we conclude with

$$\mathcal{T}_3 \leq c\varepsilon + \frac{c}{\lambda^{p\delta_0} \varepsilon^{q-1}}.$$

Here it is  $c \equiv c(n, p, q, \nu, L, \tilde{c}_1, \tilde{c}_2, \delta_0)$ . Merging the estimates for  $\mathcal{T}_1, \mathcal{T}_2$  and  $\mathcal{T}_3$  with (5.18), we deduce that

$$\begin{aligned}
&\int_B (|V_p(Dv) - V_p(Dh)|^2 + a_0 |V_q(Dv) - V_q(Dh)|^2) \chi_{\{w=w_\lambda\}} dx \\
(5.19) \quad &\leq c \left[ \sigma \lambda + \varepsilon + \frac{1}{\lambda^{p\delta_0} \varepsilon^{q-1}} \right] =: cS(\sigma, \lambda, \varepsilon)
\end{aligned}$$

for a constant  $c$  depending on  $n, p, q, \nu, L, \tilde{c}_1, \tilde{c}_2, \delta_0$  and  $\varepsilon \in (0, 1)$  has still to be chosen. Let us use the short notation

$$(5.20) \quad \mathcal{V}^2 := |V_p(Dv) - V_p(Dh)|^2 + a_0 |V_q(Dv) - V_q(Dh)|^2$$

and fix  $\theta \in (0, 1)$ , again to be chosen. Hölder's inequality and (5.19) give

$$(5.21) \quad \left( \int_B \mathcal{V}^{2\theta} \chi_{\{w=w_\lambda\}} dx \right)^{1/\theta} \leq c S(\sigma, \lambda, \varepsilon).$$

Again using Hölder inequality, we estimate as

$$\begin{aligned} \left( \int_B \mathcal{V}^{2\theta} \chi_{\{w \neq w_\lambda\}} dx \right)^{1/\theta} &\leq c \left( \frac{|\{w \neq w_\lambda\}|}{|B|} \right)^{\frac{1-\theta}{\theta}} \int_B \mathcal{V}^2 dx \\ &\stackrel{(5.17)}{\leq} c [K_0(\lambda)]^{-\frac{(1-\theta)(1+\delta_0)}{\theta}} \int_B [H_0(Dv) + H_0(Dh)] dx \\ &\stackrel{(5.8), (5.15)}{\leq} c [K_0(\lambda)]^{-\frac{1-\theta}{\theta}} \leq c \lambda^{-\frac{p(1-\theta)}{\theta}}, \end{aligned}$$

for a constant  $c \equiv c(n, p, q, \nu, L, \tilde{c}_1, \tilde{c}_2, \delta_0)$ ; notice we have used that  $\lambda \geq 1$  and that  $K_0(\lambda) \geq \lambda^p$ . The inequalities in the last two displays give, after some manipulation

$$\left( \int_B \mathcal{V}^{2\theta} dx \right)^{1/\theta} \leq c S(\sigma, \lambda, \varepsilon) + c \lambda^{-\frac{p(1-\theta)}{\theta}},$$

and again  $c \equiv c(n, p, q, \nu, L, \tilde{c}_1, \tilde{c}_2, \delta_0)$ . Recalling the definition of  $S(\sigma, \lambda, \varepsilon)$  in (5.19), and choosing  $\lambda = \sigma^{-1/2}$  and  $\varepsilon = \sigma^{p\delta_0/[4(q-1)]}$ , the previous inequality implies that

$$(5.22) \quad \left( \int_B (|V_p(Dv) - V_p(Dh)|^2 + a_0 |V_q(Dv) - V_q(Dh)|^2)^\theta dx \right)^{1/\theta} \leq c \sigma^m$$

for  $c := c(n, p, q, \tilde{c}_1, \tilde{c}_2, \delta_0, \theta)$ , where

$$(5.23) \quad m \equiv m(\theta) := \min \left\{ \frac{1}{2}, \frac{p\delta_0}{4}, \frac{p\delta_0}{4(q-1)}, \frac{p(1-\theta)}{2\theta} \right\}.$$

Recall that  $\theta \in (0, 1)$  is yet to be chosen.

*Step 2: Proof of (5.13).* We show how to deduce (5.13) from (5.22) by properly choosing  $\theta \in (0, 1)$ . Hölder's inequality with conjugate exponents  $2(1+\delta_0)/(1+2\delta_0)$  and  $2(1+\delta_0)$  yields

$$(5.24) \quad \int_B |\mathcal{V}|^2 dx \leq \left( \int_B |\mathcal{V}|^{\frac{2(1+\delta_0)}{1+2\delta_0}} dx \right)^{\frac{1+2\delta_0}{2(1+\delta_0)}} \left( \int_B |\mathcal{V}|^{2(1+\delta_0)} dx \right)^{\frac{1}{2(1+\delta_0)}}.$$

We now chose  $\theta$  as

$$(5.25) \quad \theta \equiv \theta(\delta_0) := \frac{1+\delta_0}{1+2\delta_0} \in (0, 1).$$

This allows to estimate

$$\left( \int_B |\mathcal{V}|^{\frac{2(1+\delta_0)}{1+2\delta_0}} dx \right)^{\frac{1+2\delta_0}{2(1+\delta_0)}} \leq c \sigma^{m/2}.$$

On the other hand, recalling (5.20), we simply note that

$$\begin{aligned} \int_B |\mathcal{V}|^{2(1+\delta_0)} dx &= \int_B (|V_p(Dh) - V_p(Dv)|^2 + a_0 |V_q(Dh) - V_q(Dv)|^2)^{1+\delta_0} dx \\ &\leq c \int_B [H_0(Dv)]^{1+\delta_0} dx + c \int_B [H_0(Dh)]^{1+\delta_0} dx \\ &\leq c(n, p, q, \nu, L, \delta_0, \tilde{c}_2), \end{aligned}$$

where we have used (5.12) and (5.16) in the last line. Combining the content of the last display with (5.24) and recalling (5.23) yields (5.13) with

$$s_1 := \frac{m(\theta(\delta_0))}{2} \equiv \frac{1}{2} \min \left\{ \frac{1}{2}, \frac{p\delta_0}{4}, \frac{p\delta_0}{4(q-1)}, \frac{p\delta_0}{2(1+\delta_0)} \right\},$$

that in fact exhibits the announced dependence on  $p, q$  and  $\delta_0$ .

*Step 3: Proof of (5.14).* We again use (5.22) and show how this implies (5.14) for a suitable choice of  $\theta \in (0, 1)$ . If  $q \geq 2$ , by (2.5) we see that taking  $\theta \equiv \theta_q = \max\{n/(n+q), 1/q\}$

$$\begin{aligned} a_0 \int_B \left| \frac{v-h}{r} \right|^q dx &\leq c a_0 \left( \int_B |Dv - Dh|^{q\theta} dx \right)^{1/\theta} \\ (5.26) \qquad \qquad \qquad &\leq c \left( \int_B (a_0 |V_q(Dv) - V_q(Dh)|^2)^\theta dx \right)^{1/\theta} \leq c \sigma^{m(\theta_q)} \end{aligned}$$

for a constant  $c$  depending only on  $n, p$  and  $\theta$ . When  $q < 2$ , we instead argue as follows:

$$\begin{aligned} a_0 \int_B \left| \frac{v-h}{r} \right|^q dx &\leq a_0 \left( \int_B |Du - Dh|^{q\theta} dx \right)^{1/\theta} \\ &\stackrel{(2.5)}{\leq} c a_0 \left( \int_B (|V_q(Du) - V_q(Dh)|^q (|Du| + |D\tilde{h}|)^{q(2-q)/2})^\theta dx \right)^{1/\theta} \\ &\leq c \left( \int_B (a_0 |V_q(Du) - V_q(Dh)|^2)^\theta dx \right)^{q/(2\theta)} \\ (5.27) \qquad \cdot \left( \int_B (a_0 (|Du| + |Dh|)^q)^\theta dx \right)^{(2-q)/(2\theta)} &\leq c \sigma^{m(\theta_q)q/2}, \end{aligned}$$

where in the last line we have used also (5.8) and (5.12). Summarizing, we have

$$a_0 \int_B \left| \frac{v-h}{r} \right|^q dx \leq c \sigma^{m(\theta_q) \min\{1, q/2\}}.$$

In a completely similar way we can estimate

$$\int_B \left| \frac{v-h}{r} \right|^p dx \leq c \sigma^{m(\theta_p) \min\{1, p/2\}},$$

for  $\theta_p := \max\{n/(n+p), 1/p\}$ , and the proof of (5.14) follows for a suitable exponent  $s_0 = m(\theta_p) \min\{1, p/2\}$ , with the dependence described in the statement.  $\square$

Beside the above quantitative harmonic approximation, we report the following more standard version, which is close to the ones described in [23, 24].

**Lemma 5.2.** *Let  $B \equiv B_r \subseteq \mathbb{R}^n$  be a ball, and let  $v \in W^{1,p}(B)$  be a function satisfying*

$$\int_B H_0(Dv) dx \leq \tilde{c}_1$$

for some  $\tilde{c}_1 \geq 1$ . Fix  $\tilde{\varepsilon} \in (0, 1]$ . There exists  $\sigma \equiv \sigma(n, p, q, \nu, L, \tilde{\varepsilon})$ , but otherwise independent of  $a_0$ , such that if (5.10) is satisfied, then there exists  $h \in v + W_0^{1, H_0}(B)$  such that (5.11) holds together with the estimates

$$(5.28) \qquad \int_B H_0(Dh) dx \leq c(n, p, q, \nu, L) \tilde{c}_1$$

and

$$(5.29) \qquad \int_B \left( \left| \frac{v-h}{r} \right|^p + a_0 \left| \frac{v-h}{r} \right|^q \right) dx \leq \tilde{\varepsilon}.$$

*Proof.* This is an approximation lemma of the type already developed in [23, 24] and the proof can be obtained as in these papers, with some modifications from the proof of Lemma 5.1. This leads to establish that, actually, for every choice of  $\tilde{\varepsilon}, \theta \in (0, 1)$ , there exists  $\sigma \equiv \sigma(n, p, q, \nu, L, \tilde{\varepsilon}, \theta)$  such that if (5.10) holds, then there exists  $h \in v + W_0^{1, H_0}(B)$  as in the statement of Lemma 5.2, and such that

$$(5.30) \quad \int_B (|V_p(Dv) - V_p(Dh)|^2 + a_0|V_q(Dv) - V_q(Dh)|^2)^\theta dx \leq \sigma .$$

Once this is available, then (5.29) can be deduced from (5.30) as done using (5.22) in Step 3 from the proof of Lemma 5.1.  $\square$

## 6. TWO PHASES

From now, in the rest of the proof of Theorems 1.1 and 1.2, we consider a fixed open subset  $\Omega_0 \Subset \Omega$ ; accordingly, we consider  $\mathbf{data}(\Omega_0)$  as defined in (2.2). By the basic results of Theorem 3.1 and 3.2, we know  $H(\cdot, Du) \in L_{\text{loc}}^{1+\delta}(\Omega)$  and  $u \in C_{\text{loc}}^{0, \gamma}(\Omega)$  hold for higher integrability and Hölder continuity exponents  $\delta > 0$  and  $\gamma \in (0, 1)$ , depending on  $\mathbf{data}$  and  $\mathbf{data}(\Omega_0)$ , respectively (see in fact (3.6)-(3.9)). Notice that in the case we are assuming (1.3), then  $\gamma$  only depends on  $\mathbf{data}$  as defined in (2.1). Moreover, the following estimate holds:

$$(6.1) \quad [u]_{0, \gamma; \Omega_0} + \|H(\cdot, Du)\|_{L^{1+\delta}(\Omega_0)} \leq c(\mathbf{data}(\Omega_0))$$

as a consequence of (3.4), (3.7) and of standard covering arguments. To proceed, in the following we shall consider a number  $s$  such that

$$(6.2) \quad 0 \leq s < s_m := \begin{cases} \alpha - \frac{(q-p)n}{p} + \frac{n\delta(q-p)}{p(1+\delta)} & \text{if (1.2) is considered and } p > n \\ p + \alpha - q + \gamma(q-p) & \text{if (1.3) is considered .} \end{cases}$$

Needless to say, (1.2)-(1.3) imply that  $s_m > 0$ , and in the limiting equality cases too; moreover, it is  $s < s_m \leq \alpha$ .

Next, given a ball  $B \equiv B_R \Subset \Omega$ ,  $R \leq 1$ , and a number  $M \geq 1$ , and with

$$(6.3) \quad x_B \in \bar{B}, \quad a_i(B) := a(x_B) = \min_{\bar{B}} a(x) ,$$

we define the  $p$ -phase in  $B_R$  as the occurrence of the following inequality:

$$(6.4) \quad a_i(B_R) \leq M[a]_{0, \alpha} R^{\alpha-s} .$$

The  $(p, q)$ -phase is instead defined by the complementary condition, i.e.,

$$(6.5) \quad a_i(B_R) > M[a]_{0, \alpha} R^{\alpha-s} .$$

**Remark 6.1.** We record an immediate consequence of (6.4), that is

$$(6.6) \quad a_i(B_R) \leq M[a]_{0, \alpha} R^{\alpha-s} \implies \|a\|_{L^\infty(B_R)} \leq 3M[a]_{0, \alpha} R^{\alpha-s} .$$

Indeed, observe that

$$\|a\|_{L^\infty(B_R)} \leq a_i(B_R) + [a]_{0, \alpha} (2R)^\alpha \leq M[a]_{0, \alpha} R^{\alpha-s} + [a]_{0, \alpha} (2R)^\alpha \leq 3M[a]_{0, \alpha} R^{\alpha-s} .$$

**Remark 6.2.** We observe that if the condition (6.5) holds in  $B_R \equiv B_R(x_0)$ , then it also holds on all smaller concentric balls  $B_\varrho(x_0)$ , that is

$$(6.7) \quad a_i(B_R(x_0)) > M[a]_{0, \alpha} R^{\alpha-s} \implies a_i(B_\varrho(x_0)) > M[a]_{0, \alpha} \varrho^{\alpha-s} \quad \forall \varrho \leq R .$$

Indeed we have  $a_i(B_\varrho(x_0)) \geq a_i(B_R(x_0)) > M[a]_{0, \alpha} R^{\alpha-s} \geq M[a]_{0, \alpha} \varrho^{\alpha-s}$ .

**Remark 6.3.** We notice the property

$$(6.8) \quad a_i(B_R) > M[a]_{0, \alpha} R^{\alpha-s} \implies \|a\|_{L^\infty(B_R)} \leq [2/M + 1]a_i(B_R)$$

Indeed,  $a(x) \leq a(x) - a_i(B_R) + a_i(B_R) \leq 2[a]_{0, \alpha} R^\alpha + a_i(B_R) \leq [2/M + 1]a_i(B_R)$ .

We gather the following consequence of Proposition 3.1:

**Lemma 6.1.** *Let  $u \in W^{1,p}(\Omega)$  be a local minimiser of the functional  $\mathcal{F}$  defined in (1.5), under the assumptions (1.6). Moreover, assume that either (1.2) or (1.3) is satisfied. Consider a ball  $B_R \Subset \Omega_0 \Subset \Omega$ , with  $R \leq 1$ , such that (6.4) holds for some  $s$  as in (6.2) and some  $M \geq 1$ . Then*

$$(6.9) \quad \int_{B_t} H(x, Du) dx \leq c \left( \frac{R}{s-t} \right)^q \int_{B_s} \left| \frac{u - u_0}{R} \right|^p dx$$

holds for a constant  $c \equiv c(\mathbf{data}(\Omega_0), M)$ , whenever  $B_t \Subset B_s \Subset B_R$  are concentric balls and for every  $u_0 \in \mathbb{R}$  such that  $\inf u \leq u_0 \leq \sup u$ . Similarly, the inequality

$$(6.10) \quad \int_{B_t} H(x, D(u - k)_\pm) dx \leq c \left( \frac{R}{s-t} \right)^q \int_{B_s} \left| \frac{(u - k)_\pm}{s-t} \right|^p dx$$

holds for every  $k \in \mathbb{R}$  such that  $\inf u \leq k \leq \sup u$ , where  $c \equiv c(n, p, q, \nu, L, M)$ . Here  $(u - k)_+$  and  $(u - k)_-$  are defined as in (3.3).

*Proof.* We give the proof of (6.10), the one of (6.9) being completely similar. An easy manipulation of (3.2), which holds in every case/phase, gives

$$(6.11) \quad \begin{aligned} & \int_{B_t} H(x, D(u - k)_\pm) dx \\ & \leq c \left( \frac{R}{s-t} \right)^q \int_{B_s} \left( 1 + a(x) \left| \frac{(u - k)_\pm}{R} \right|^{q-p} \right) \left| \frac{(u - k)_\pm}{R} \right|^p dx. \end{aligned}$$

To proceed, we distinguish between the two cases, (1.2) and (1.3). For (1.3), recalling that  $\inf u \leq k \leq \sup u$  and using Remark 6.1, we can estimate

$$(6.12) \quad a(x) \left| \frac{(u - k)_\pm}{R} \right|^{q-p} \leq cM[a]_{0,\alpha}[u]_{0,\gamma;B_R}^{q-p} R^{p+\alpha-q+\gamma(q-p)-s} \stackrel{(6.1)}{\leq} c,$$

where  $c \equiv c(\mathbf{data}(\Omega_0), M)$  and we have used the relation in (6.2). Combining the inequalities in the last two displays yields (6.10). Then we consider (1.2). By the discussion in Section 4 we restrict to the case  $p \geq n$ . Indeed, by Sobolev-Morrey theorem, using again Remark 6.1, we have that

$$\begin{aligned} a(x) \left| \frac{(u - k)_\pm}{R} \right|^{q-p} & \leq 3M[a]_{0,\alpha} R^{p-q+\alpha-s} [\text{osc}_{B_r} u]^{q-p} \\ & \leq cR^{p-q+\alpha-s} \|Du\|_{L^{p(1+\delta)}(B_R)}^{q-p} R^{\left[1 - \frac{n}{p(1+\delta)}\right](q-p)} \\ & \leq c \|H(\cdot, Du)\|_{L^{1+\delta}(B_R)}^{q/p-1} R^{\alpha - \frac{(q-p)n}{p} + \frac{n\delta(q-p)}{p(1+\delta)} - s} \stackrel{(6.1)}{\leq} c, \end{aligned}$$

for  $c \equiv c(\mathbf{data}(\Omega_0), M)$ . Combining this last inequality with (6.11) yields (6.10) once again and the proof is complete.  $\square$

**Remark 6.4.** Estimates (6.9) and (6.10) continue to hold if, instead of assuming (1.2) or (1.3), we consider (1.17). In this case, instead of (6.2), we always take  $s_m = p + \alpha - q + \gamma(q-p)$ . The validity of (6.10) (and similarly of (6.9)) then follows by first writing (6.11) and then observing that (6.12) still holds under assumptions (1.17).

## 7. $p$ -PHASE

In the setting specified in the previous section, we shall here exploit the estimates implied by condition (6.4), where  $B_R \Subset \Omega_0 \Subset \Omega$  with  $R \leq 1$ , and  $s$  is a number

satisfying (6.2). We define, whenever  $u \in W^{1,1}(\Omega)$  is a local minimizer of (1.5) and  $w \in W^{1,1}(B_R)$

$$(7.1) \quad \mathcal{G}_B(w) := \int_{B_R} F(x, (u)_{B_R}, Dw) dx ,$$

and  $\tilde{v} \in u + W_0^{1,p}(B)$  as the solution to the variational Dirichlet problem:

$$(7.2) \quad \begin{cases} \tilde{v} \mapsto \min_w \mathcal{G}_B(w) \\ w \in u + W_0^{1,p}(B_R) . \end{cases}$$

We then have the following preliminary lemma, that essentially serves to get rid of the  $u$ -dependence in the original functional  $\mathcal{F}$  in (1.5).

**Lemma 7.1.** *Let  $u \in W^{1,p}(\Omega)$  be a local minimiser of the functional  $\mathcal{F}$  defined in (1.5), under assumptions (1.6), (1.8) and (1.12). Assume that either (1.2) or (1.3) is satisfied. Let  $B_R \Subset \Omega_0 \Subset \Omega$  be a ball with  $R \leq 1$ . Then there exists a positive constant  $c \equiv c(\mathbf{data}(\Omega_0))$  such that then the following estimate holds:*

$$(7.3) \quad \begin{aligned} & \int_{B_R} (|V_p(Du) - V_p(D\tilde{v})|^2 + a(x)|V_q(Du) - V_q(D\tilde{v})|^2) dx \\ & \leq c\omega(R^\gamma) \int_{B_R} H(x, Du) dx , \end{aligned}$$

where  $\gamma \equiv \gamma(\mathbf{data}(\Omega_0)) \in (0, 1)$  is the Hölder continuity exponent fixed in (6.1). Moreover, the following inequalities hold:

$$(7.4) \quad \|\tilde{v}\|_{L^\infty(B_R)} \leq \|u\|_{L^\infty(B_R)} , \quad \int_{B_R} H(x, D\tilde{v}) dx \leq \frac{L}{\nu} \int_{B_R} H(x, Du) dx$$

and

$$(7.5) \quad \text{osc}_{B_R} \tilde{v} \leq \text{osc}_{B_R} u .$$

Finally, for every  $\varepsilon \in (0, 1)$ , the estimate

$$(7.6) \quad \int_{B_R} \left| \frac{u - \tilde{v}}{R} \right|^p dx \leq c [c_\varepsilon \omega(R^\gamma) + \varepsilon] \int_{B_R} H(x, Du) dx$$

holds for  $c$  depending on  $\mathbf{data}(\Omega_0)$  and  $c_\varepsilon$  depending only on  $\varepsilon, p$  and  $q$ .

*Proof.* The Euler-Lagrange equation of the functional  $\mathcal{F}_B$ , that is

$$(7.7) \quad \int_{B_R} \langle \partial F(x, (u)_{B_R}, D\tilde{v}), D\varphi \rangle dx = 0 ,$$

is now satisfied for any choice  $\varphi \in W_0^{1,1}(B_R)$  such that  $H(x, D\varphi) \in L^1(B_R)$  (see also [18, Remark 6.1]). On the other hand, we notice that by minimality and growth conditions we also have that  $H(x, D\tilde{v}) \in L^1(B_R)$ :

$$(7.8) \quad \begin{aligned} \int_{B_R} H(x, D\tilde{v}) dx & \stackrel{(1.6)}{\leq} \frac{1}{\nu} \int_{B_R} F(x, (u)_{B_R}, D\tilde{v}) dx \\ & \leq \frac{1}{\nu} \int_{B_R} F(x, (u)_{B_R}, Du) dx \leq \frac{L}{\nu} \int_{B_R} H(x, Du) dx . \end{aligned}$$

This proves the second inequality in (7.4). Therefore we conclude with

$$(7.9) \quad \int_{B_R} \langle \partial F(x, (u)_{B_R}, D\tilde{v}), Du - D\tilde{v} \rangle dx = 0 .$$

Moreover, let us recall two basic consequences of (1.8)<sub>2</sub>. Using the content of [33, (4.13)], and arguing in a totally similar way to get that lemma, we this time have that

$$(7.10) \quad |V_p(z_2) - V_p(z_1)|^2 + a(x)|V_q(z_2) - V_q(z_1)|^2 \\ + c \langle \partial F(x, (u)_{B_R}, z_1), z_2 - z_1 \rangle \leq c [F(x, (u)_{B_R}, z_2) - F(x, (u)_{B_R}, z_1)]$$

holds for every choice of  $z_1, z_2 \in \mathbb{R}^n$  and  $x \in \Omega$ , for a constant  $c$  depending only on  $n, p, q, \nu$ . By using (7.7), (7.10) and the minimality of  $u$  and  $\tilde{v}$ , we then find

$$(7.11) \quad \int_{B_R} (|V_p(Du) - V_p(D\tilde{v})|^2 + a(x)|V_q(Du) - V_q(D\tilde{v})|^2) dx \\ \stackrel{(7.9)}{=} \int_{B_R} (|V_p(Du) - V_p(D\tilde{v})|^2 + a(x)|V_q(Du) - V_q(D\tilde{v})|^2) dx \\ + c \int_{B_R} \langle \partial F(x, (u)_{B_R}, D\tilde{v}), Du - D\tilde{v} \rangle dx \\ \stackrel{(7.10)}{\leq} c \int_{B_R} [F(x, (u)_{B_R}, Du) - F(x, (u)_{B_R}, D\tilde{v})] dx \\ = c \int_{B_R} [F(x, (u)_{B_R}, Du) - F(x, u, Du)] dx \\ + c \int_{B_R} [F(x, u, Du) - F(x, \tilde{v}, D\tilde{v})] dx \\ + c \int_{B_R} [F(x, \tilde{v}, D\tilde{v}) - F(x, (\tilde{v})_{B_R}, D\tilde{v})] dx \\ + c \int_{B_R} [F(x, (\tilde{v})_{B_R}, D\tilde{v}) - F(x, (u)_{B_R}, D\tilde{v})] dx =: c \sum_{i=1}^4 I_i .$$

We can now proceed estimating the terms  $I_1, \dots, I_4$ . We have

$$I_1 \stackrel{(1.8)}{\leq} c \int_{B_R} \omega(|u - (u)_{B_R}|) H(x, Du) dx \\ \stackrel{(6.5)}{\leq} c\omega(2[u]_{0,\gamma;\Omega_0} R^\gamma) \int_{B_R} H(x, Du) dx \\ \stackrel{(3.7)}{\leq} c(\mathbf{data}(\Omega_0)) \omega(R^\gamma) \int_{B_R} H(x, Du) dx ,$$

where we have also used that  $\omega(\cdot)$  is concave. On the other hand, the minimality of  $u$  yields  $I_2 \leq 0$ . As for  $I_3$ , we start proving (7.5). Indeed, observe that, with  $k := \sup_{B_R} u$ , the minimality of  $\tilde{v}$  yields  $\mathcal{G}_B(\tilde{v}) \leq \mathcal{G}_B(\min\{\tilde{v}, k\})$  and this, together with (1.6), in turn gives

$$\int_{\{\tilde{v} \geq k\} \cap B} H(x, D\tilde{v}) dx = 0 .$$

By coarea formula we then get  $\tilde{v} \leq k$  a.e. Arguing similarly, but with the choice  $k := \inf_{B_R} u$ , and using that  $\mathcal{G}_B(\tilde{v}) \leq \mathcal{G}_B(\max\{\tilde{v}, k\})$ , we get that  $\tilde{v} \geq k$  a.e. and (7.5) follows. This also proves the first inequality in (7.4). We therefore have

$$I_3 \stackrel{(1.8)}{\leq} c \int_{B_R} \omega(|\tilde{v} - (\tilde{v})_{B_R}|) H(x, D\tilde{v}) dx \\ \stackrel{(7.5)}{\leq} c\omega\left(\operatorname{osc}_{B_R} u\right) \int_{B_R} H(x, D\tilde{v}) dx$$

$$\begin{aligned}
&\stackrel{(7.8)}{\leq} c\omega(2[u]_{0,\gamma;\Omega_0}R^\gamma) \int_{B_R} H(x, Du) dx \\
&\stackrel{(3.7)}{\leq} c\omega(R^\gamma) \int_{B_R} H(x, Du) dx,
\end{aligned}$$

where  $c \equiv c(\mathbf{data}(\Omega_0))$ . In a similar way, as  $u \equiv \tilde{v}$  on  $\partial B$ , we first observe that

$$|(\tilde{v})_{B_R} - (u)_{B_R}|^p dx \leq c \int_{B_R} |\tilde{v} - u|^p dx \stackrel{(7.5)}{\leq} c \left( \operatorname{osc}_{B_R} u \right)^p$$

and then, as already done for  $I_3$ , we still find

$$I_4 \leq c\omega(R^\gamma) \int_{B_R} H(x, Du) dx.$$

Connecting the estimates found for the terms  $I_1, \dots, I_4$  to (7.11) completes the proof of (7.3). It remains to prove (7.6). For this we distinguish two cases; the first is when  $p \geq 2$ ; then by Poincaré inequality we have

$$\int_{B_R} \left| \frac{u - \tilde{v}}{R} \right|^p dx \leq c \int_{B_R} |Du - D\tilde{v}|^p dx \leq c \int_{B_R} |V_p(Du) - V_p(D\tilde{v})|^2 dx.$$

In this case (7.6) follows combining the inequalities in the last two displays with (7.3). The case  $p < 2$  is slightly more elaborate. Using also Young's inequality with conjugate exponents  $(2/p, 2/(2-p))$ , we have

$$\begin{aligned}
&\int_{B_R} \left| \frac{u - \tilde{v}}{R} \right|^p dx \leq \int_{B_R} |Du - D\tilde{v}|^p dx \\
&\stackrel{(2.5)}{\leq} c \int_{B_R} \left( |V_p(Du) - V_p(D\tilde{v})|^p (|Du| + |D\tilde{v}|)^{p(2-p)/2} \right) dx \\
&\leq c \left( \int_{B_R} |V_p(Du) - V_p(D\tilde{v})|^2 dx \right)^{p/2} \left( \int_{B_R} (|Du| + |D\tilde{v}|)^p dx \right)^{(2-p)/2} \\
&\stackrel{(7.4)}{\leq} c \left( \int_{B_R} |V_p(Du) - V_p(D\tilde{v})|^2 dx \right)^{p/2} \left( \int_{B_R} [H(x, Du)] dx \right)^{(2-p)/2} \\
&\leq c_\varepsilon \int_{B_R} |V_p(Du) - V_p(D\tilde{v})|^2 dx + \varepsilon \int_{B_R} [H(x, Du)] dx.
\end{aligned}$$

Once again (7.6) follows combining the last inequality with (7.3).  $\square$

To proceed, we gather some regularity information on  $\tilde{v}$ .

**Lemma 7.2** (Controlled transfer of regularity). *Under the assumptions of Lemma 7.1, let  $\tilde{v} \in W^{1,p}(B_R)$  be defined as in (7.2) and assume that (6.4) holds in  $B_R$  with some  $s$  as in (6.2) and some  $M \geq 1$ . Then the following inequality holds for a constant  $c \equiv c(\mathbf{data}(\Omega_0), M)$*

$$(7.12) \quad \sup_{B_{R/4}} |\tilde{v} - (\tilde{v})_{B_{R/2}}|^p \leq c \int_{B_{R/2}} |\tilde{v} - (\tilde{v})_{B_{R/2}}|^p dx,$$

while the following ones

$$(7.13) \quad [\tilde{v}]_{0,\gamma;B_{R/2}} \leq c[u]_{0,\gamma;B_R}, \quad \|H(\cdot, D\tilde{v})\|_{L^{1+\delta}(B_{R/2})} \leq c \|H(\cdot, Du)\|_{L^{1+\delta}(B_R)}$$

hold for constants  $c \equiv c(\mathbf{data}(\Omega_0))$ . In (7.13) the exponents  $\gamma \equiv \gamma(\mathbf{data}(\Omega_0))$  and  $\delta \equiv \delta(\mathbf{data})$  are those fixed in (6.1).

*Proof.* Looking at the proof in [5, Lemma 4.1], it is easy to see that the  $L^\infty$ -bound

$$(7.14) \quad \sup_{B_{R/4}} |\tilde{v}|^p \leq c \int_{B_{R/2}} |\tilde{v}|^p dx$$

holds whenever the Caccioppoli inequality

$$(7.15) \quad \int_{B_t} H(x, D(\tilde{v} - k)_\pm) dx \leq c_{cacc} \left( \frac{R}{s-t} \right)^q \int_{B_s} \left| \frac{(\tilde{v} - k)_\pm}{s-t} \right|^p dx$$

is satisfied for concentric balls  $B_t \Subset B_s \Subset B_{R/2}$  and  $\inf \tilde{v} \leq k \leq \sup \tilde{v}$ ; moreover, the constant  $c$  in (7.14) depends only on  $n, p, q$  and  $c_{cacc}$ . The validity of (7.15) follows observing that Lemma 6.1 applies to  $\tilde{v}$  since this minimizes a functional (the one in (7.1)) that satisfies the assumptions of Lemma 6.1. Indeed, here we are assuming (6.4). Furthermore, thanks to estimates (7.4), when applying Lemma 6.1, we can dispense from the dependence of the various constants on the quantities  $\|\tilde{v}\|_{L^\infty(B_R)}$  and  $\|H(\cdot, D\tilde{v})\|_{L^1(B_R)}$ . We therefore conclude that (7.15) holds for a constant  $c_{cacc}$  that ultimately depends only on  $\mathbf{data}(\Omega_0)$  and  $M$  but not on  $\tilde{v}$ ; from this it follows that also  $c$  appearing in (7.14) only depends on  $\mathbf{data}(\Omega_0)$  and  $M$ . To conclude with (7.12), notice that since  $\tilde{v} - (\tilde{v})_{B_{R/2}}$  is still a local minimizer of the same functional in (7.1), then (7.12) follows from (7.14) applied to  $\tilde{v} - (\tilde{v})_{B_{R/2}}$ . We then pass to the first estimate in (7.13), arguing in a similar way, but this time applying Theorem 3.2 to  $\tilde{v}$ ; in particular, we need the precise dependence of  $\gamma$  stated in (3.6). This and the first inequality in (7.4) give that  $\tilde{v} \in C_{loc}^{0,\gamma}(B_R)$  where  $\gamma$  is the same as in (6.1). Moreover, estimate (3.8) holds for  $\tilde{v}$  with  $c$  depending only on  $\mathbf{data}$ ; this is still a consequence of (7.4). Applying (3.8) to  $\tilde{v}$  and using a standard covering argument, we gain  $[\tilde{v}]_{0,\gamma;B_{R/2}} \leq c(\mathbf{data}(\Omega_0))R^{-\gamma} \text{osc}_{B_R} \tilde{v}$ . On the other hand, using also (7.5), it is  $\text{osc}_{B_R} \tilde{v} \leq \text{osc}_{B_R} u \leq 2R^\gamma [u]_{0,\gamma;B_R}$ , so that the first inequality in (7.13) follows. It remains to prove the second inequality in (7.13) and this time we apply Theorem 3.1 to  $\tilde{v}$ ; again by (7.4) the exponent  $\delta$  for  $v$  does not depend on the solution considered, but only on  $\mathbf{data}$ , and can be taken to be the same one for  $u$  fixed in (6.1). Therefore we have

$$\begin{aligned} \left( \int_{B_{R/2}} [H(x, D\tilde{v})]^{1+\delta} dx \right)^{1/(1+\delta)} &\stackrel{(3.4)}{\leq} c \int_{B_R} H(x, D\tilde{v}) dx \\ &\stackrel{(7.4)}{\leq} c \int_{B_R} H(x, Du) dx \leq c \left( \int_{B_R} [H(x, Du)]^{1+\delta} dx \right)^{1/(1+\delta)}, \end{aligned}$$

and the proof is complete.  $\square$

In the following we shall use the classical excess functional given by

$$(7.16) \quad E(w; B_r) := \left( \int_{B_r} |w - (w)_{B_r}|^p dx \right)^{1/p},$$

and defined whenever  $w \in L^p(B_r)$  is a possibly vector-valued map and  $B_r \subset \mathbb{R}^n$  is a ball. We shall several times use the following elementary property:

$$(7.17) \quad E(w; B_r) \leq 2 \left( \int_{B_r} |w - w_0|^p dx \right)^{1/p} \quad \text{for every } w_0 \in \mathbb{R}^k.$$

We now go to the main lemma of this section, featuring a first decay estimate:

**Lemma 7.3** (Blow-up). *Under the assumptions of Lemma 7.1, let  $\tilde{v} \in W^{1,p}(B_R)$  be defined as in (7.2) and assume also that (6.4) holds with some  $s$  as in (6.2) and some  $M \geq 1$ . Then, for every choice of  $\tilde{\varepsilon} \in (0, 1)$ , there exists a positive radius*

$$(7.18) \quad R^* \equiv R^*(\mathbf{data}(\Omega_0), M, \omega(\cdot), s_m - s, \tilde{\varepsilon}),$$

such that if  $R \leq R^*$ , then the inequality

$$(7.19) \quad \int_{B_{\tau R}} |\tilde{v} - (\tilde{v})_{B_{\tau R}}|^p dx \leq c \{ \tau^p + \tau^{-n} \tilde{\varepsilon} \} \int_{B_R} |\tilde{v} - (\tilde{v})_{B_R}|^p dx$$

holds for every  $\tau \in (0, 1/16)$ , where the constant  $c \equiv c(\mathbf{data}(\Omega_0), M)$  is otherwise independent of  $\tilde{\varepsilon}$  and  $\tau$ .

*Proof.* The proof goes in several steps and essentially aims at applying the harmonic approximation argument of Lemma 5.1 to a suitable blown-up function.

*Step 1: Blow-up.* We start blowing-up  $\tilde{v}$  and  $a(\cdot)$  in  $B_R \equiv B_R(x_0)$ , thereby defining

$$(7.20) \quad v(x) := \frac{\tilde{v}(x_0 + Rx) - (\tilde{v})_{B_{R/2}}}{E(\tilde{v}; B_{R/2})} \quad \text{and} \quad a_R(x) := a(x_0 + Rx), \quad x \in B_1.$$

Notice that we can always assume that  $E(\tilde{v}; B_{R/2}) > 0$ , otherwise (7.19) follows trivially. Abbreviating  $E(R) \equiv E(\tilde{v}; B_{R/2})$ , by (6.4), the Caccioppoli inequality in (6.9) applies to  $\tilde{v}$  in the ball  $B_{R/2}$  as well. The result is that

$$\int_{B_{R/4}} H(x, D\tilde{v}) dx \leq c \int_{B_{R/2}} \left| \frac{\tilde{v} - (\tilde{v})_{B_{R/2}}}{R} \right|^p dx$$

holds, where  $c \equiv c(\mathbf{data}(\Omega_0), M)$ . Notice that, as already done for Lemma 7.2, here we are using (7.4) to dispense from the dependence on  $\|\tilde{v}\|_{L^\infty(B_R)}$  and incorporate it in the one on  $\|u\|_{L^\infty(\Omega_0)}$  and, ultimately, in  $\mathbf{data}(\Omega_0)$ . In view of the definition in (7.16), the inequality in the above display can be re-written as

$$\int_{B_{R/4}} ([R/E(R)]^p |D\tilde{v}|^p + [R/E(R)]^p a(x) |D\tilde{v}|^q) dx \leq c$$

that in terms of  $v$  means

$$(7.21) \quad \int_{B_{1/4}} (|Dv|^p + [E(R)/R]^{q-p} a_R(x) |Dv|^q) dx \leq c,$$

again for  $c \equiv c(\mathbf{data}(\Omega_0), M)$ . It is now easy to see that  $v$  is a local minimizer of the functional defined by

$$(7.22) \quad w \mapsto \int_{B_1} \bar{F}(x, Dw) dx$$

with

$$(7.23) \quad \bar{F}(x, z) := \frac{F(x_0 + Rx, (u)_{B_R}, [E(R)/R]z)}{[E(R)/R]^p},$$

for every  $x \in B_1$  and  $z \in \mathbb{R}^n$ , and, as such, it satisfies its Euler-Lagrange equation. This means that

$$(7.24) \quad \int_{B_1} \langle \partial F(x_0 + Rx, (u)_{B_R}, [E(R)/R]Dv), D\varphi \rangle dx = 0$$

holds for every  $\varphi \in W_0^{1,\infty}(B_1)$ .

*Step 2: Universal energy bounds.* We check the properties of the rescaled integrand  $\bar{F}(\cdot)$  in (7.23). By defining the new control function

$$(7.25) \quad \bar{H}(x, z) := |z|^p + [E(R)/R]^{q-p} a_R(x) |z|^q.$$

the following conditions are now satisfied whenever  $x \in B_1$ , and  $z \in \mathbb{R}^n$ :

$$(7.26) \quad \nu \bar{H}(x, z) \leq \bar{F}(x, z) \leq L \bar{H}(x, z)$$

as a direct consequence of (1.6) and (7.23). We now want to check that  $\bar{F}(\cdot)$  satisfies the conditions allowing to apply Theorem 3.1 in  $B_{1/4}$ , this time to  $v$ . For this we have to check that  $v$  is bounded in  $B_{1/4}$  (in fact needed only when  $p < n$ ) and that, in every case, the new coefficient  $x \mapsto [E(R)/R]^{q-p} a_R(x)$  satisfies the needed Hölder continuity assumption in  $B_{1/4}$ , that is

$$(7.27) \quad [E(R)/R]^{q-p} [a_R]_{0,\alpha;B_{1/4}} \leq c(\mathbf{data}(\Omega_0)).$$

Recalling the definition of  $v$  in (7.20), and using Lemma 7.2, (7.12) yields

$$(7.28) \quad \|v\|_{L^\infty(B_{1/4})} \leq c(\mathbf{data}(\Omega_0), M).$$

For (7.27), we distinguish two cases.

*Case assumption (1.3) is in force.* By (6.1) and (7.13) we have

$$(7.29) \quad E(R) = \left( \int_{B_{R/2}} |\tilde{v} - (\tilde{v})_{B_{R/2}}|^p dx \right)^{1/p} \leq cR^\gamma$$

for a constant  $c \equiv c(\mathbf{data}(\Omega_0))$ . To prove (7.27), observe that, whenever  $x_1, x_2 \in B_1$ , it follows that

$$(7.30) \quad \begin{aligned} & |[E(R)/R]^{q-p} a_R(x_1) - [E(R)/R]^{q-p} a_R(x_2)| \\ & \leq c[a]_{0,\alpha} [E(R)/R]^{q-p} R^\alpha |x_1 - x_2|^\alpha \\ & \stackrel{(7.29)}{\leq} cR^{(p-q)(1-\gamma)+\alpha} |x_1 - x_2|^\alpha \stackrel{(1.3)}{\leq} c(\mathbf{data}(\Omega_0)) |x_1 - x_2|^\alpha. \end{aligned}$$

**Remark 7.1.** Notice that, in order to perform the last estimation, we essentially need the larger bound  $q \leq p + \alpha/(1 - \gamma)$  related to the one appearing in (1.17).

*Case assumption (1.2) is in force and  $p \geq n$ .* Recalling (6.1) and the second inequality in (7.13), by Sobolev-Morrey embedding theorem we then have

$$(7.31) \quad \begin{aligned} E(R) & \leq c(n, p, \delta) \|D\tilde{v}\|_{L^{p(1+\delta)}(B_{R/4})} R^{1-\frac{n}{p(1+\delta)}} \\ & \stackrel{(7.13)}{\leq} c(\mathbf{data}(\Omega_0)) \|H(\cdot, Du)\|_{L^{1+\delta}(B_{R/2})}^{1/p} R^{1-\frac{n}{p(1+\delta)}} \\ & \stackrel{(6.1)}{\leq} c(\mathbf{data}(\Omega_0)) R^{1-\frac{n}{p}+\frac{n\delta}{p(1+\delta)}}. \end{aligned}$$

Therefore this time we can estimate, again for  $x_1, x_2 \in B_1$

$$\begin{aligned} |[E(R)/R]^{q-p} a_R(x_1) - [E(R)/R]^{q-p} a_R(x_2)| & \leq c[a]_{0,\alpha} R^{n(1-q/p)+\alpha} |x_1 - x_2|^\alpha \\ & \leq c(\mathbf{data}(\Omega_0)) |x_1 - x_2|^\alpha, \end{aligned}$$

as now it is  $q/p \leq 1 + \alpha/n$ , and (7.27) follows in this case too.

With (7.27) and (7.28) at our disposal, and recalling that (7.26) holds, we are now able to apply Theorem 3.1 to  $v$ , thereby obtaining the existence of a higher integrability exponent  $\delta_1 > 0$  and a constant  $c$ , both depending on  $\mathbf{data}(\Omega_0)$  and  $M$ , but otherwise independent of  $R$ , such that

$$(7.32) \quad \left( \int_{B_{1/8}} [\bar{H}(x, Dv)]^{1+\delta_1} dx \right)^{1/(1+\delta_1)} \leq c \int_{B_{1/4}} \bar{H}(x, Dv) dx.$$

Moreover, recalling (7.21), we conclude with

$$(7.33) \quad \int_{B_{1/4}} \bar{H}(x, Dv) dx + \int_{B_{1/8}} [\bar{H}(x, Dv)]^{1+\delta_1} dx \leq c(\mathbf{data}(\Omega_0), M).$$

*Step 3: Frozen functional.* Let us now define

$$\bar{F}_0(z) := \frac{F(x_B, (u)_{B_R}, [E(R)/R]z)}{[E(R)/R]^p}$$

and the related control Young function

$$(7.34) \quad \bar{H}_0(z) := |z|^p + [E(R)/R]^{q-p} a_i(B_R) |z|^q$$

for every  $z \in \mathbb{R}^n$ . Later on, we shall use the functional

$$(7.35) \quad W^{1, \bar{H}_0}(B_{1/8}) \ni w \mapsto \int_{B_{1/8}} \bar{F}_0(Dw) dx.$$

We notice that the integrand  $\bar{F}_0(\cdot)$  satisfies the following growth and ellipticity conditions:

$$(7.36) \quad \begin{cases} \nu \bar{H}_0(z) \leq \bar{F}_0(z) \leq L \bar{H}_0(z) \\ |\partial \bar{F}_0(z)| |z| + |\partial^2 \bar{F}_0(z)| |z|^2 \leq L \bar{H}_0(z) \\ \nu \frac{\bar{H}_0(z)}{|z|^2} |\xi|^2 \leq \langle \partial^2 \bar{F}_0(z) \xi, \xi \rangle, \end{cases}$$

for every choice of  $z \in \mathbb{R}^n \setminus \{0\}$  and  $\xi \in \mathbb{R}^n$ .

*Step 4: Harmonic type approximation.* We notice that, with  $\varphi \in W_0^{1,\infty}(B_{1/8})$  being such that  $\|D\varphi\|_{L^\infty(B_{1/8})} \leq 1$ , by (7.24) we can write

$$(7.37) \quad \begin{aligned} & \left| \int_{B_{1/8}} \langle \partial \bar{F}_0(Dv), D\varphi \rangle dx \right| \\ &= \left| \int_{B_{1/8}} \left\langle \partial \bar{F}_0(Dv) - \frac{\partial F(x_0 + Rx, (u)_{B_R}, [E(R)/R] Dv)}{[E(R)/R]^{p-1}}, D\varphi \right\rangle dx \right| \\ &\leq \int_{B_{1/8}} \left| \partial \bar{F}_0(Dv) - \frac{\partial F(x_0 + Rx, (u)_{B_R}, [E(R)/R] Dv)}{[E(R)/R]^{p-1}} \right| dx := \mathcal{I}. \end{aligned}$$

The term  $\mathcal{I}$  can be estimated using (1.8)<sub>3</sub>. Recalling that  $a_i(B_R) \leq a_R(x)$  for every  $x \in B_1$ , this gives

$$(7.38) \quad \begin{aligned} \mathcal{I} &\leq c\omega(R) \int_{B_{1/8}} |Dv|^{p-1} dx + c\omega(R) \left[ \frac{E(R)}{R} \right]^{q-p} \int_{B_{1/8}} a_R(x) |Dv|^{q-1} dx \\ &\quad + c \left[ \frac{E(R)}{R} \right]^{q-p} \int_{B_{1/8}} |a_R(x) - a_i(B_R)| |Dv|^{q-1} dx \\ &=: \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3, \end{aligned}$$

with  $c \equiv c(p, q, L)$ . We then estimate separately the above three terms. In any case we have, by Hölder's inequality, we have that

$$\mathcal{I}_1 \leq c\omega(R) \left( \int_{B_{1/8}} |Dv|^p dx \right)^{1-1/p} \stackrel{(7.21)}{\leq} c\omega(R)$$

for  $c \equiv c(\text{data}(\Omega_0), M)$ . For the remaining terms we distinguish between two different cases, as already done Step 2.

*Case assumption (1.3) is in force.* We have

$$(7.39) \quad \begin{aligned} & \mathcal{I}_2 \\ &= c\omega(R) \left[ \frac{E(R)}{R} \right]^{\frac{q-p}{q}} \int_{B_{1/8}} [a_R(x)]^{\frac{1}{q}} [a_R(x)]^{\frac{q-1}{q}} \left[ \frac{E(R)}{R} \right]^{\frac{(q-p)(q-1)}{q}} |Dv|^{q-1} dx \\ &\stackrel{(7.29)}{\leq} c\omega(R) M^{1/q} R^{\frac{p+\alpha-q+\gamma(q-p)-s}{q}} \left( \int_{B_{1/8}} [E(R)/R]^{q-p} a_R(x) |Dv|^q dx \right)^{1-1/q} \\ &\stackrel{(7.21)}{\leq} c\omega(R), \end{aligned}$$

for  $c \equiv c(\text{data}(\Omega_0), M)$ , where we have used  $R \leq 1$  and (6.2). Similarly, we have

$$\mathcal{I}_3 \leq c[a]_{0,\alpha}^{1/q} \left[ \frac{E(R)}{R} \right]^{\frac{q-p}{q}} R^{\frac{\alpha}{q}} \int_{B_{1/8}} [a_R(x)]^{\frac{q-1}{q}} \left[ \frac{E(R)}{R} \right]^{\frac{(q-p)(q-1)}{q}} |Dv|^{q-1} dx$$

$$\begin{aligned}
& \stackrel{(7.29)}{\leq} cR^{\frac{p+\alpha-q+\gamma(q-p)-s}{q}} \left( \int_{B_{1/8}} [E(R)/R]^{q-p} a_R(x) |Dv|^q dx \right)^{1-1/q} \\
(7.40) \quad & \stackrel{(7.21)}{\leq} cR^{\frac{p+\alpha-q+\gamma(q-p)-s}{q}} .
\end{aligned}$$

Collecting the estimates found for the terms  $\mathcal{I}_1, \mathcal{I}_2$  and  $\mathcal{I}_3$  in the last three displays and merging them with (7.38), we conclude with

$$(7.41) \quad \mathcal{I} \leq c(\mathbf{data}(\Omega_0), M) \left[ \omega(R) + R^{\frac{p+\alpha-q+\gamma(q-p)-s}{q}} \right] .$$

*Case assumption (1.2) is in force and  $p \geq n$ .* Re-writing as in (7.39), but using this time (7.31), we have

$$\begin{aligned}
& \left[ \frac{E(R)}{R} \right]^{q-p} \int_{B_{1/8}} a_R(x) |Dv|^{q-1} dx \\
& \stackrel{(6.6)}{\leq} cM^{1/q} R^{\frac{\alpha-s}{q}} \left[ \frac{E(R)}{R} \right]^{\frac{q-p}{q}} \left( \int_{B_{1/8}} [E(R)/R]^{q-p} a_R(x) |Dv|^q dx \right)^{1-1/q} \\
& \stackrel{(7.21)}{\leq} cM^{1/q} R^{\frac{\alpha-s}{q}} \left[ \frac{E(R)}{R} \right]^{\frac{q-p}{q}} \\
(7.42) \quad & \stackrel{(7.31)}{\leq} c(\mathbf{data}(\Omega_0), M) R^{\frac{1}{q} \left[ \alpha - \frac{(q-p)n}{p} + \frac{n\delta(q-p)}{p(1+\delta)} - s \right]} .
\end{aligned}$$

Notice that the exponent of  $R$  in the last line of (7.42) is non-negative by (1.2) and (6.2). By using (7.42) we readily have that

$$\mathcal{I}_2 \leq cR^{\frac{1}{q} \left[ \alpha - \frac{(q-p)n}{p} + \frac{n\delta(q-p)}{p(1+\delta)} - s \right]} \omega(R) \leq c\omega(R) ,$$

again for  $c \equiv c(\mathbf{data}(\Omega_0), M)$ , while, recalling the definition of  $a_i(B_R)$  in (6.3), we finally estimate

$$\mathcal{I}_3 \leq c \left[ \frac{E(R)}{R} \right]^{q-p} \int_{B_{1/8}} a_R(x) |Dv|^{q-1} dx \leq cR^{\frac{1}{q} \left[ \alpha - \frac{(q-p)n}{p} + \frac{n\delta(q-p)}{p(1+\delta)} - s \right]} ,$$

again with  $c \equiv c(\mathbf{data}(\Omega_0), M)$ . Merging the estimates for the three terms  $\mathcal{I}_1, \mathcal{I}_2$  and  $\mathcal{I}_3$  with the one in (7.38), we this time get

$$(7.43) \quad \mathcal{I} \leq c(\mathbf{data}(\Omega_0), M) \left\{ \omega(R) + R^{\frac{1}{q} \left[ \alpha - \frac{(q-p)n}{p} + \frac{n\delta(q-p)}{p(1+\delta)} - s \right]} \right\} .$$

**Remark 7.2.** The analysis of the first case is still valid when assuming larger bound in (1.17). This is better than the one in (1.2) if  $p > n/(1-\gamma)$ .

Now, taking into account the estimates found for  $\mathcal{I}$  in (7.41) and (7.43), and recalling (7.37), by finally defining

$$(7.44) \quad \mathfrak{o}(R) := \begin{cases} \omega(R) + R^{\frac{1}{q} \left[ \alpha - \frac{(q-p)n}{p} + \frac{n\delta(q-p)}{p(1+\delta)} - s \right]} & \text{if we use (1.2)} \\ \omega(R) + R^{\frac{p+\alpha-q+\gamma(q-p)-s}{q}} & \text{if we use (1.3)} , \end{cases}$$

which is non-decreasing, we conclude with

$$(7.45) \quad \left| \int_{B_{1/8}} \langle \partial \bar{F}_0(Dv), D\varphi \rangle dx \right| \leq c_h \mathfrak{o}(R) \|D\varphi\|_{L^\infty}$$

that now holds for every  $\varphi \in W_0^{1,\infty}(B_{1/8})$ , where  $c_h \equiv c_h(\mathbf{data}(\Omega_0), M)$ . Observe also that, with the notation in (6.2), we have in any case - either we use (1.2) or (1.3) - that

$$(7.46) \quad \mathfrak{o}(R) = \omega(R) + R^{(s_m-s)/q} .$$

By (7.33), (7.45) and (7.36) we are now in position to apply Lemma 5.1 with the choice  $A_0(z) \equiv \partial \bar{F}_0(z)$ ,  $H_0(z) \equiv \bar{H}_0(z)$ , and  $a_0 \equiv [E(R)/R]^{q-p} a_i(B_R)$ , so that, in particular, assumptions (5.1) are satisfied. By Lemma 5.1 there exists a function  $h \in v + W_0^{1, \bar{H}_0}(B_{1/8})$  such that

$$(7.47) \quad \int_{B_{1/8}} \langle \partial \bar{F}_0(Dh), D\varphi \rangle dx = 0 \quad \text{for all } \varphi \in W_0^{1, \infty}(B_{1/8}),$$

$$(7.48) \quad \int_{B_{1/8}} \bar{H}_0(Dh) dx + \int_{B_{1/8}} [H_0(Dh)]^{1+\delta_0} dx \leq c,$$

$$(7.49) \quad \int_{B_{1/8}} (|V_p(Dv) - V_p(Dh)|^2 + [E(R)/R]^{q-p} a_i(B_R) |V_q(Dv) - V_q(Dh)|^2) dx \leq c[\mathfrak{o}(R)]^{s_1}$$

and finally

$$(7.50) \quad \int_{B_{1/8}} (|v - h|^p + [E(R)/R]^{q-p} a_i(B_R) |v - h|^q) dx \leq c_d [\mathfrak{o}(R)]^{s_0}.$$

In all the estimates above the constants  $c, c_d \geq 1$  and  $s_0, s_1 \in (0, 1)$  depend on  $\mathbf{data}(\Omega_0)$  and  $M$ , but are otherwise independent of  $R$ .

*Step 5: Choice of the radius  $R_*$  in (7.18).* Given the definition of  $\mathfrak{o}(R)$  in (7.46), and given  $\tilde{\varepsilon} \in (0, 1)$  as in the statement, we determine  $R_*$  small enough to get

$$c_d [\omega(R_*)]^{s_0} + c_d R_*^{s_0(s_m - s)/q} \leq \tilde{\varepsilon}.$$

This, given the dependence of  $s_0$  and  $c_d$  on  $\mathbf{data}(\Omega_0)$  and  $M$ , gives the final dependence displayed in (7.18). By (7.50) we conclude that

$$(7.51) \quad \int_{B_{1/8}} (|v - h|^p + [E(R)/R]^{q-p} a_i(B_R) |v - h|^q) dx \leq \tilde{\varepsilon}.$$

*Step 6: Proof of the decay estimate (7.19).* We first observe that by a standard density argument the relation in (7.47) continues to hold whenever  $\varphi \in W_0^{1,1}(B_{1/8})$  is such that  $\bar{H}_0(D\varphi) \in L^1(B_{1/8})$ . This means that  $h$  is a local minimizer of the functional

$$W^{1, \bar{H}_0}(B_{1/8}) \ni w \mapsto \int_{B_{1/8}} \bar{F}_0(Dw) dx.$$

Conditions (7.36) are satisfied by the integrand  $\bar{F}_0(\cdot)$ , so we are able to apply the results from [39] (see also [4]) to deduce the following a priori gradient bound:

$$(7.52) \quad \sup_{B_{1/16}} \bar{H}_0(Dh) \leq c \int_{B_{1/8}} \bar{H}_0(Dh) dx,$$

where this time  $c \equiv c(n, p, q, \nu, L)$ . By taking  $\tau \leq 1/16$ , we can estimate

$$\begin{aligned} \int_{B_\tau} |v - (v)_{B_\tau}|^p dx &\leq c \int_{B_\tau} |v - (h)_{B_\tau}|^p dx \\ &\leq c \int_{B_\tau} |h - (h)_{B_\tau}|^p dx + c\tau^{-n} \int_{B_{1/8}} |v - h|^p dx \\ (7.51) \quad &\leq c\tau^p \sup_{B_\tau} |Dh|^p + c\tau^{-n} \tilde{\varepsilon} \\ &\leq c\tau^p \sup_{B_\tau} \bar{H}_0(Dh) + c\tau^{-n} \tilde{\varepsilon} \\ (7.52) \quad &\leq c\tau^p \int_{B_{1/8}} \bar{H}_0(Dh) dx + c\tau^{-n} \tilde{\varepsilon} \end{aligned}$$

$$\stackrel{(7.48)}{\leq} c\tau^p + c\tau^{-n}\tilde{\varepsilon}.$$

By scaling back to  $\tilde{v}$  and recalling (7.20), we get (7.19) after some elementary manipulations.  $\square$

A crucial outgrowth of the previous proof is

**Lemma 7.4** (Comparison). *In the setting of Lemma 7.1, let  $\tilde{v} \in W^{1,p}(B_R)$  be defined as in (7.2) and assume also that (6.4) holds with some  $s$  as in (6.2) and some  $M \geq 1$ . Then, there exists a function  $\tilde{h} \in \tilde{v} + W_0^{1,p}(B_{R/8})$  which is a local minimizer of the functional*

$$(7.53) \quad \mathcal{F}_B(w) := \int_{B_{R/8}} F(x_B, (u)_{B_R}, Dw) dx$$

in the sense of Definition 1 and that satisfies

$$(7.54) \quad \begin{aligned} & \int_{B_{R/8}} \left( |V_p(Du) - V_p(D\tilde{h})|^2 + a_i(B_R) |V_q(Du) - V_q(D\tilde{h})|^2 \right) dx \\ & \leq c \left\{ [\omega(R^\gamma)]^{s_1} + R^{s_1(s_m-s)/q} \right\} \int_{B_R} H(x, Du) dx \end{aligned}$$

for a constant  $c$  and a positive exponent  $s_1$ , both depending only on  $\mathbf{data}(\Omega_0)$  and  $M$ . Finally, the inequality

$$(7.55) \quad \int_{B_{R/8}} \left( |D\tilde{h}|^p + a_i(B_R) |D\tilde{h}|^q \right) dx \leq c \int_{B_R} H(x, Du) dx$$

holds for  $c \equiv c(n, \nu, L)$ .

*Proof.* We go back to the proof of Lemma 7.3, Steps 3 and 4 and consider the function  $h \in v + W_0^{1,\tilde{H}_0}(B_{1/8})$  found by the application of Lemma 5.1 and solving (7.47). We blow-down  $h$  defining

$$(7.56) \quad \tilde{h}(x) := E(\tilde{v}; B_R) h \left( \frac{x - x_0}{R} \right), \quad x \in B_{R/8}(x_0).$$

We notice that  $\tilde{h} \in \tilde{v} + W_0^{1,\tilde{H}_0}(B_{R/8})$ , where now  $\tilde{H}_0(z) := |z|^p + a_i(B_R)|z|^q$ . The function  $h$  solves (7.47), which is the Euler-Lagrange equation of the functional in (7.35), and therefore, by strong convexity, a local minimizer too. More precisely, we have that

$$\int_{B_{1/8}} \bar{F}_0(Dh) dx \leq \int_{B_{1/8}} \bar{F}_0(Dh + D\varphi) dx$$

holds whenever  $\varphi \in W_0^{1,\tilde{H}_0}(B_{1/8})$  and  $\tilde{H}_0(\cdot)$  is defined as in (7.34). This implies that  $\tilde{h}$  in (7.56) minimizes the functional in (7.53) in the sense that  $\mathcal{F}_B(\tilde{h}) \leq \mathcal{F}_B(\tilde{h} + \varphi)$  holds for every  $\varphi \in W_0^{1,\tilde{H}_0}(B_{R/8})$ . By (7.36)  $\tilde{h}$  satisfies the energy estimate

$$(7.57) \quad \int_{B_{R/8}} \tilde{H}_0(D\tilde{h}) dx \leq \frac{L}{\nu} \int_{B_{R/8}} \tilde{H}_0(D\tilde{v}) dx \leq \frac{8^n L}{\nu} \int_{B_R} H(x, D\tilde{v}) dx$$

which can be obtained as in (7.8). Given the definitions of  $\tilde{h}$  and of  $v$  in (7.20), the comparison estimate in (7.49) gives, using also Poincaré inequality and (7.17)

$$\begin{aligned} & \int_{B_{R/8}} \left( |V_p(D\tilde{v}) - V_p(D\tilde{h})|^2 + a_i(B_R) |V_q(D\tilde{v}) - V_q(D\tilde{h})|^2 \right) dx \\ & \leq c [\mathfrak{o}(R)]^{s_1} \int_{B_R} \left| \frac{\tilde{v} - (\tilde{v})_{B_R}}{R} \right|^p dx \leq c [\mathfrak{o}(R)]^{s_1} \int_{B_R} |D\tilde{v}|^p dx \end{aligned}$$

$$(7.58) \quad \leq c [\mathfrak{o}(R)]^{s_1} \int_{B_R} H(x, D\tilde{v}) dx \stackrel{(7.4)}{\leq} c [\mathfrak{o}(R)]^{s_1} \int_{B_R} H(x, Du) dx .$$

Here it is  $c \equiv c(\mathbf{data}(\Omega_0), M)$  while  $\mathfrak{o}(R)$  has been defined in (7.44) and (7.46). Combining (7.58) with (7.3), and recalling again (7.46), we get (7.54) again after a few elementary manipulations. Finally, (7.55) follows using (7.57) together with the second inequality in (7.4).  $\square$

We finally conclude this section with another decay estimate.

**Lemma 7.5.** *Let  $u \in W^{1,p}(\Omega)$  be a local minimiser of the functional  $\mathcal{F}$  defined in (1.5), under the assumptions (1.8)-(1.9). Assume that either (1.2) or (1.3) is satisfied, and furthermore that*

$$(7.59) \quad a_i(B_{\tau R}) \leq M[a]_{0,\alpha}(\tau R)^{\alpha-s}$$

holds with some  $s$  as in (6.2) and some  $M \geq 1$ . Then, for every choice of  $\bar{\varepsilon} \in (0, 1)$ , there exists a positive radius  $R_*$ ,

$$(7.60) \quad R_* \equiv R_*(\mathbf{data}(\Omega_0), M, \omega(\cdot), s_m - s, \bar{\varepsilon}) ,$$

such that if  $R \leq R_*$ , then the inequality

$$(7.61) \quad \int_{B_{\tau R}} H(x, Du) dx \leq c_p \{ \tau^n + c_\varepsilon \tau^{-p} \bar{\varepsilon} + \tau^{-p} \varepsilon \} \int_{B_R} H(x, Du) dx ,$$

holds for every  $\tau \in (0, 1/32)$  and every  $\varepsilon \in (0, 1)$ , where the constants  $c_p \equiv c_p(\mathbf{data}(\Omega_0), M)$  and  $c_\varepsilon \equiv c_\varepsilon(\varepsilon, p, q)$  are otherwise independent of  $\bar{\varepsilon}$  and  $\tau$ .

*Proof.* We notice that (7.59) implies that also (6.4) holds, that is,

$$(7.62) \quad a_i(B_R) \leq M[a]_{0,\alpha} R^{\alpha-s}$$

holds true. Indeed, on the contrary, thanks to (6.7),  $a_i(B_R) > M[a]_{0,\alpha} R^{\alpha-s}$  would imply that  $a_i(B_{\tau R}) > M[a]_{0,\alpha}(\tau R)^{\alpha-s}$ , that is, a contradiction (7.59). The validity of (7.62) means we can apply Lemma 7.3 with  $\tilde{\varepsilon} \equiv \varepsilon_1$  to be chosen in a few lines, and we can use (7.19) provided  $R \leq R^* \equiv R^*(\mathbf{data}(\Omega_0), M, \omega(\cdot), s_m - s, \varepsilon_1)$  is determined via (7.18). We therefore get, with some elementary manipulations and using (7.17) repeatedly

$$(7.63) \quad \begin{aligned} \int_{B_{\tau R}} |u - (u)_{B_{\tau R}}|^p dx &\leq 2^p \int_{B_{\tau R}} |u - (\tilde{v})_{B_{\tau R}}|^p dx \\ &\leq c \int_{B_{\tau R}} |\tilde{v} - (\tilde{v})_{B_{\tau R}}|^p dx + c\tau^{-n} \int_{B_R} |u - \tilde{v}|^p dx \\ &\leq c \{ \tau^p + \tau^{-n} \varepsilon_1 \} \int_{B_R} |u - (u)_{B_R}|^p dx + c\tau^{-n} \int_{B_R} |u - \tilde{v}|^p dx . \end{aligned}$$

Using (7.6) to estimate the last integral in the above display, and performing some other standard manipulations, we further get

$$\begin{aligned} \int_{B_{\tau R}} \left| \frac{u - (u)_{B_{\tau R}}}{\tau R} \right|^p dx \\ \leq c \{ \tau^n + c_\varepsilon \tau^{-p} \omega(R^\gamma) + \tau^{-p} \varepsilon_1 + \tau^{-p} \varepsilon \} \int_{B_R} H(x, Du) dx . \end{aligned}$$

This holds for every  $\tau \in (0, 1/16)$  and  $c \equiv c(\mathbf{data}(\Omega_0), M)$ . Then we select  $\varepsilon_1 = \bar{\varepsilon}/2$  and finally, we take  $R_* \leq R^*$  such that  $\omega(R_*^\gamma) \leq \bar{\varepsilon}/2$ . This choice determines the dependence described in (7.60) and yields

$$(7.64) \quad \int_{B_{\tau R}} \left| \frac{u - (u)_{B_{\tau R}}}{\tau R} \right|^p dx \leq c \{ \tau^n + c_\varepsilon \tau^{-p} \bar{\varepsilon} + \tau^{-p} \varepsilon \} \int_{B_R} H(x, Du) dx .$$

On the other hand the validity (7.59) implies that applicability conditions of Lemma 6.1 are verified with  $B_R$  replaced by  $B_{\tau R}$ . We therefore deduce

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

Using this last inequality together with (7.64) (and renaming  $\tau/2$  in  $\tau$ ) yields (7.61) and the proof is complete.  $\square$

## 8. $(p, q)$ -PHASE

In this section we consider the condition which is complementary to (6.4), that is (6.5). This time we consider it with *the special choice*  $s = 0$ , i.e.,

$$(8.1) \quad a_i(B_R) > M[a]_{0,\alpha} R^\alpha \quad \text{for some } M \geq 1 .$$

Notice that we can actually always reduce to this case since (6.5) implies (8.1) for every  $s$  as in (6.2) (recall that we always take  $R \leq 1$ ). We consider the different frozen functional

$$(8.2) \quad \mathcal{F}_B(w) := \int_{B_R} F(x_B, (u)_{B_R}, Dw) dx ,$$

and define  $h \in u + W_0^{1,p}(B_R)$  as the solution to the following Dirichlet problem:

$$(8.3) \quad \begin{cases} \tilde{h} \mapsto \min_w \mathcal{F}_B(w) \\ w \in u + W_0^{1,p}(B_R) . \end{cases}$$

As for (7.8), we get

$$(8.4) \quad \int_{B_R} (|D\tilde{h}|^p + a_i(B_R)|D\tilde{h}|^q) dx \leq \frac{L}{\nu} \int_{B_R} (|Du|^p + a_i(B_R)|Du|^q) dx .$$

Notice that now it is  $\tilde{h} \in W^{1,q}(B_R)$ . We then have the following direct analog of Lemma 7.4.

**Lemma 8.1.** *Let  $u \in W^{1,p}(\Omega)$  be a local minimiser of the functional  $\mathcal{F}$  defined in (1.5), under the assumptions (1.6), (1.8) and (1.12). Assume that either (1.2) or (1.3) is satisfied. Let  $B_R \Subset \Omega_0 \Subset \Omega$  be a ball with  $R \leq 1$  and such that the condition (8.1) is in force. Then there exists a positive constant  $c \equiv c(\mathbf{data}(\Omega_0))$ , which is independent of  $M$ , such that the following estimate holds:*

$$(8.5) \quad \begin{aligned} & \int_{B_R} (|V_p(Du) - V_p(D\tilde{h})|^2 + a(x)|V_q(Du) - V_q(D\tilde{h})|^2) dx \\ & \leq c \left[ \omega(R^\gamma) + \frac{1}{M} \right] \int_{B_R} H(x, Du) dx . \end{aligned}$$

Here  $\gamma \equiv \gamma(\mathbf{data}(\Omega_0)) \in (0, 1)$  is the Hölder continuity exponent defined in Theorem 3.2. Moreover, it holds that

$$(8.6) \quad \int_{B_R} H(x, D\tilde{h}) dx \leq \frac{3L}{\nu} \int_{B_R} H(x, Du) dx .$$

*Proof.* First, let us observe that assumptions (1.8) also imply that

$$(8.7) \quad \begin{aligned} & |F(x_1, v, z) - F(x_2, v, z)| \\ & \leq L \omega(|x_1 - x_2|) [H(x_1, z) + H(x_2, z)] + L|a(x_1) - a(x_2)||z|^q \end{aligned}$$

holds for every choice of  $x_1, x_2 \in \Omega$ ,  $v \in \mathbb{R}$  and  $z \in \mathbb{R}^n$ . Indeed, we can write

$$|F(x_1, v, z) - F(x_2, v, z)| = \left| \int_0^1 \langle \partial F(x_1, v, tz) - \partial F(x_2, v, tz), z \rangle dt \right|$$

$$\leq \int_0^1 |\partial F(x_1, v, tz) - \partial F(x_2, v, tz)| dt |z| ,$$

so that (8.7) follows by (1.8). Notice that, in order to derive the equality in the above display, we have used that  $F(x_1, v, 0) = F(x_2, v, 0) = 0$ , which is in turn implied by (1.6). Next, observe the basic consequence of (6.5)

$$(8.8) \quad a(x) \leq a(x) - a_i(B_R) + a_i(B_R) \leq 2[a]_{0,\alpha} R^\alpha + a_i(B_R) \leq 3a_i(B_R) \leq 3a(x) ,$$

that holds for every  $x \in B_R$ . This and (8.4) easily imply (8.6). We now proceed as in Lemma 7.1 and arrive at (7.11), which is this time replaced by

$$\begin{aligned} & \int_{B_R} \left( |V_p(Du) - V_p(D\tilde{h})|^2 + a_i(B_R) |V_q(Du) - V_q(D\tilde{h})|^2 \right) dx \\ & \leq c \int_{B_R} \left[ F(x_B, (u)_{B_R}, Du) - F(x_B, (u)_{B_R}, D\tilde{h}) \right] dx \\ & = c \int_{B_R} [F(x_B, (u)_{B_R}, Du) - F(x, (u)_{B_R}, Du)] dx \\ & \quad + c \int_{B_R} [F(x, (u)_{B_R}, Du) - F(x, u, Du)] dx \\ & \quad + c \int_{B_R} [F(x, u, Du) - F(x, \tilde{h}, D\tilde{h})] dx \\ & \quad + c \int_{B_R} [F(x, \tilde{h}, D\tilde{h}) - F(x, (\tilde{h})_{B_R}, D\tilde{h})] dx \\ & \quad + c \int_{B_R} [F(x, (\tilde{h})_{B_R}, D\tilde{h}) - F(x, (u)_{B_R}, D\tilde{h})] dx \\ (8.9) \quad & + c \int_{B_R} [F(x, (u)_{B_R}, D\tilde{h}) - F(x_B, (u)_{B_R}, D\tilde{h})] dx =: c \sum_{i=1}^6 II_i . \end{aligned}$$

We can now proceed estimating the terms  $II_1, \dots, II_6$  essentially using (1.8)<sub>3,4</sub> and (6.5). We have

$$\begin{aligned} II_1 + II_6 & \stackrel{(8.7)}{\leq} c\omega(2R) \int_{B_R} [H(x_B, Du) + H(x_B, D\tilde{h})] dx \\ & \quad + c\omega(2R) \int_{B_R} [H(x, Du) + H(x, D\tilde{h})] dx \\ & \quad + c[a]_{0,\alpha} R^\alpha \int_{B_R} (|Du|^q + |D\tilde{h}|^q) dx \\ & \leq c\omega(2R) \int_{B_R} [H(x, Du) + H(x, D\tilde{h})] dx \\ & \quad + c[a]_{0,\alpha} R^\alpha \int_{B_R} (|Du|^q + |D\tilde{h}|^q) dx \\ & \stackrel{(8.1)}{\leq} c\omega(2R) \int_{B_R} [H(x, Du) + H(x, D\tilde{h})] dx \\ & \quad + \frac{c}{M} \int_{B_R} a(x) (|Du|^q + |D\tilde{h}|^q) dx \\ & \stackrel{(8.6)}{\leq} c \left[ \omega(R) + \frac{1}{M} \right] \int_{B_R} H(x, Du) dx . \end{aligned}$$

The estimates for the terms  $II_2, II_4$  and  $II_5$  can be now be obtained as in the case of the analogous terms  $I_1, I_3$  and  $I_4$ , respectively, considered in Lemma 7.1. This

yields, by also repeatedly using (8.6)

$$II_2 + II_4 + II_5 \leq c\omega(R^\gamma) \int_{B_R} H(x, Du) dx$$

for a constant  $c \equiv c(\mathbf{data}(\Omega_0))$ . Finally, by the minimality of  $u$  it follows that  $II_3 \leq 0$ . By combining the estimates found for the terms  $II_1, \dots, II_6$  with (8.9) and, taking (8.8) into account, we conclude with (8.5) and the proof is complete.  $\square$

We can now prove the following  $(p, q)$ -phase counterpart of Lemma 7.5:

**Lemma 8.2.** *Under the assumptions and notations of Lemma 8.1, for every number  $\tau \in (0, 1/2)$ , the inequality*

$$(8.10) \quad \int_{B_{\tau R}} H(x, Du) dx \leq c_{p,q} \left\{ \tau^n + \omega(R^\gamma) + \frac{1}{M} \right\} \int_{B_R} H(x, Du) dx ,$$

holds for a constant  $c_{p,q} \equiv c_{p,q}(\mathbf{data}(\Omega_0))$ , which is otherwise independent of  $M$  and  $\tau$ .

*Proof.* The integrand  $z \mapsto F(x_B, (u)_{B_R}, z)$  of the functional in (8.2) satisfies assumptions (1.6) and (1.8)<sub>1,2</sub> with  $a(x) \equiv a_i(B_R)$ . We are again in position to apply the results from [39], yielding the a priori estimate

$$\sup_{B_{R/2}} \left( |D\tilde{h}|^p + a_i(B_R)|D\tilde{h}|^q \right) \leq c \int_{B_R} \left( |D\tilde{h}|^p + a_i(B_R)|D\tilde{h}|^q \right) dx$$

for  $c$  depending only on  $n, p, q, \nu, L$ . Thanks to (8.8), the last inequality commutes into

$$(8.11) \quad \sup_{B_{R/2}} H(x, D\tilde{h}) \stackrel{(8.8)}{\leq} \int_{B_R} H(x, D\tilde{h}) dx .$$

Then we proceed comparing  $\tilde{h}$  and  $u$  as follows:

$$\begin{aligned} \int_{B_{\tau R}} H(x, Du) dx &\leq c \int_{B_{\tau R}} H(x, D\tilde{h}) dx \\ &\quad + c \int_{B_R} \left( |V_p(Du) - V_p(D\tilde{h})|^2 + a_i(B_R)|V_q(Du) - V_q(D\tilde{h})|^2 \right) dx \\ &\stackrel{(8.5), (8.11)}{\leq} c\tau^n \int_{B_R} H(x, D\tilde{h}) dx + c \left[ \omega(R^\gamma) + \frac{1}{M} \right] \int_{B_R} H(x, Du) dx , \end{aligned}$$

where  $c \equiv c(\mathbf{data}(\Omega_0))$ . Using (8.6) to estimate the first integral in the last line finally yields (8.10) and the proof is complete.  $\square$

## 9. MORREY DECAY, EXIT TIMES AND THEOREM 1.2

In this section we prove Theorem 1.2. The proof employs (1.12) but not (1.9), and follows combining Lemmas 7.5 and 8.2 from the last two sections. We shall use an exit time argument, that, in a different form, has been introduced in [16, 17] to treat the regularity of minimizers of the model functional  $\mathcal{P}$  in (1.1). The final outcome is the Morrey type decay estimate (1.14). This eventually implies the local Hölder continuity of  $u$  for every exponent  $\theta < 1$  claimed in Theorem 1.2 via the usual integral characterization of Campanato and Meyers (see for instance [29, Chapter 2]) and a standard covering argument. We start fixing a number  $\sigma \in (0, n)$ ; moreover, for every ball  $B_R \equiv B_R(x_0) \Subset \Omega$  we abbreviate as follows:

$$\mathcal{H}(\varrho) := \int_{B_\varrho(x_0)} H(x, Du) dx \quad \forall \varrho \leq R .$$

*Step 1: Iteration in the  $(p, q)$ -phase.* Let us consider a ball  $B_R \equiv B_R(x_0) \Subset \Omega_0 \Subset \Omega$ ,  $R \leq 1$  and assume that (6.5) holds for some  $M \geq 1$  to be chosen in a few lines. We start recalling the basic property in (6.7). This, in perspective, allows to iteratively apply Lemma 8.2 on all smaller scales once it can be applied on a fixed, initial scale. When using Lemma 8.2, inequality (8.10) can be rewritten as

$$(9.1) \quad \mathcal{H}(\tau R) \leq \tau^{n-\sigma} \left\{ c_{p,q} \tau^\sigma + c_{p,q} \tau^{\sigma-n} \omega(R^\gamma) + \frac{c_{p,q} \tau^{\sigma-n}}{M} \right\} \mathcal{H}(R),$$

where the constant  $c_{p,q}$  has been determined in Lemma 8.2 as a function of  $\mathbf{data}(\Omega_0)$  but not of  $M$ . We start taking  $\tau \equiv \tau_1$  small enough in order to get

$$(9.2) \quad \tau_1^\sigma \leq \frac{1}{3c_{p,q}} \implies \tau_1 \equiv \tau_1(\mathbf{data}(\Omega_0), \sigma).$$

With this choice we determine a (potentially small) radius  $R_1 \equiv R_1(\mathbf{data}(\Omega_0), \sigma) \leq 1$  and a (potentially large) number  $M \equiv M(\mathbf{data}(\Omega_0), \sigma) \geq 1$ , in order to satisfy the two inequalities

$$(9.3) \quad \omega(R_1^\gamma) \leq \frac{\tau_1^{n-\sigma}}{3c_{p,q}} \quad \text{and} \quad \frac{1}{M} \leq \frac{\tau_1^{n-\sigma}}{3c_{p,q}}.$$

With such choices (9.1) implies

$$(9.4) \quad \mathcal{H}(\tau_1 R) \leq \tau_1^{n-\sigma} \mathcal{H}(R),$$

for every  $R$  such that  $R \leq R_1$ . By using (6.7) we deduce that since (6.5) holds, then we also have  $a_i(B_{\tau^k R}) > M[a]_{0,\alpha}(\tau^k R)^{\alpha-s}$  holds for every integer  $k \geq 0$ . We can therefore iterate (9.4), thereby obtaining

$$(9.5) \quad \mathcal{H}(\tau_1^k R) \leq \tau_1^{(n-\sigma)k} \mathcal{H}(R) \quad \text{for every integer } k \geq 0.$$

*Step 2: Decay in the  $p$ -phase.* The choices in (9.2) and (9.3) fix, in particular, the value of the threshold constant  $M$  as a function of  $\mathbf{data}(\Omega_0)$  and  $\sigma$ . This is the value we are going to use here when applying Lemma 7.5; in particular, this choice determines  $c_p$  and  $s_0$  as functions of  $\mathbf{data}(\Omega_0)$  and  $\sigma$  only. Moreover, we now take a fixed choice of the exponent  $s$  in (7.59), say  $s := s_m/2$ . With  $M$  and  $s$  having been fixed, the radius  $R_*$  from (7.18) is now a function of  $\mathbf{data}(\Omega_0)$ ,  $\sigma$  and  $\bar{\varepsilon}$ . Again, let us consider, as in Step 1, a ball  $B_R \Subset \Omega_0 \Subset \Omega$ . This will be here considered with a radius  $R$  that will initially be such that  $R \leq R_1$  and  $R_1 \equiv R_1(\mathbf{data}(\Omega_0), \sigma) \leq 1$  has been determined in the previous step; further restrictions will be taken in a short while. In order to apply Lemma 7.5 with the mentioned choice of  $M$ , we look at estimate (7.61), that reads as

$$(9.6) \quad \mathcal{H}(\tau R) \leq \tau^{n-\sigma} \left\{ c_p \tau^\sigma + c_p c_\varepsilon \tau^{\sigma-n-p} \bar{\varepsilon} + c_p \tau^{\sigma-n-p} \varepsilon \right\} \mathcal{H}(R),$$

for every choice of  $\bar{\varepsilon}, \varepsilon \in (0, 1)$ , provided we are choosing  $R \leq R_*(\mathbf{data}(\Omega_0), \sigma, \bar{\varepsilon})$ . We choose  $\tau \equiv \tau_2$  such that

$$(9.7) \quad \tau_2^\sigma \leq \frac{1}{3c_p} \implies \tau_2 \equiv \tau_2(\mathbf{data}(\Omega_0), \sigma).$$

We then proceed with the choice of  $\varepsilon \equiv \varepsilon(\mathbf{data}(\Omega_0), \sigma)$  such that

$$(9.8) \quad \varepsilon \leq \frac{\tau_2^{n+p-\sigma}}{3c_p}.$$

This, in turn, determines the constant  $c_\varepsilon$  in (9.6) as a function of  $\mathbf{data}(\Omega_0)$  and  $\sigma$ . We then take  $\bar{\varepsilon} \equiv \bar{\varepsilon}(\mathbf{data}(\Omega_0), \sigma)$  small enough to guarantee that

$$(9.9) \quad \bar{\varepsilon} \leq \frac{\tau_2^{n+p-\sigma}}{3c_p c_\varepsilon}.$$

The choice in (9.9), via (7.18), in turn determines a radius

$$(9.10) \quad R_2 := R_*(\mathbf{data}(\Omega_0), \sigma) \leq R_1(\mathbf{data}(\Omega_0), \sigma) \leq 1$$

such that (9.6) is satisfied according to the prescription of Lemma 7.5.

Summarizing, with the choices of  $\tau_2$  and  $R_2$  displayed in (9.7)-(9.9) and used in (9.6), we can conclude that if  $R \leq R_2 \leq R_1$  and if the condition

$$(9.11) \quad a_i(B_{\tau_2 R}) \leq M[a]_{0,\alpha}(\tau_2 R)^\alpha,$$

is satisfied, then we have

$$(9.12) \quad \mathcal{H}(\tau_2 R) \leq \tau_2^{n-\sigma} \mathcal{H}(R).$$

*Step 3: Exit time and iteration.* We combine (9.5) and (9.12) via the announced exit time argument and take a general ball  $B_R \equiv B_R(x_0) \Subset \Omega_0 \Subset \Omega$  with  $R \leq R_2$  and  $R_2$  as in (9.10). By looking at (9.11), we consider conditions of the type

$$(9.13) \quad a_i(B_{\tau_2^{k+1} R}) \leq M[a]_{0,\alpha}(\tau_2^{k+1} R)^\alpha,$$

for every integer  $k \geq 0$ . We define the exit time index

$$(9.14) \quad m := \min\{k \in \mathbb{N} : \text{condition (9.13) fails}\}$$

(notice that it can happen that  $m = 0$ ). By the content of Step 2, and in particular applying (9.12) repeatedly, we have

$$(9.15) \quad \mathcal{H}(\tau_2^k R) \leq \tau_2^{(n-\sigma)k} \mathcal{H}(R) \quad \forall k \in \{0, \dots, m\}.$$

Condition (9.13) fails for  $k = m$  and this means

$$(9.16) \quad a_i(B_{\tau_2^{m+1} R}) > M[a]_{0,\alpha}(\tau_2^{m+1} R)^\alpha.$$

The occurrence of (9.16) allows to apply the argument of Step 1, and in particular allows to use (9.5) (with  $B_R$  now replaced by  $B_{\tau_2^{m+1} R}$ ); we therefore conclude that

$$(9.17) \quad \mathcal{H}(\tau_1^k \tau_2^{m+1} R) \leq \tau_1^{(n-\sigma)k} \mathcal{H}(\tau_2^{m+1} R) \quad \text{for every integer } k \geq 0.$$

*Step 4: Conclusion.* We are ready to establish (1.14). For this we first assume that  $0 < \varrho < R \leq R_2$  and  $R_2$  is in (9.10). We then distinguish various cases. The first occurs when  $\tau_2^{m+1} R \leq \varrho$  and  $m$  is in (9.14); in this situation we find  $\bar{k} \in \{0, \dots, m\}$  such that  $\tau_2^{\bar{k}+1} R \leq \varrho \leq \tau_2^{\bar{k}} R$  and therefore  $\tau_2^{\bar{k}+1} \leq \varrho/R$ . By using this last relation we can estimate

$$(9.18) \quad \begin{aligned} \mathcal{H}(\varrho) &\leq \mathcal{H}(\tau_2^{\bar{k}} R) \stackrel{(9.15)}{\leq} \tau_2^{(n-\sigma)\bar{k}} \mathcal{H}(R) \\ &\leq \tau_2^{\sigma-n} (\varrho/R)^{n-\sigma} \mathcal{H}(R) \stackrel{(9.7)}{\equiv} c(\mathbf{data}(\Omega_0), \sigma) (\varrho/R)^{n-\sigma} \mathcal{H}(R), \end{aligned}$$

where  $c \equiv c(\mathbf{data}(\Omega_0), \sigma)$ ; therefore (1.14) follows in this case provided  $R \leq R_2$ .

We then consider the occurrence of  $\varrho < \tau_2^{m+1} R$ , and make a further distinction. We first treat the case when  $\tau_1 \tau_2^{m+1} R \leq \varrho$ . In this situation we estimate

$$\begin{aligned} \mathcal{H}(\varrho) &\leq \mathcal{H}(\tau_2^{m+1} R) \stackrel{(9.18)}{\leq} c \tau_2^{(n-\sigma)(m+1)} \mathcal{H}(R) \\ &\leq c \tau_1^{\sigma-n} (\varrho/R)^{n-\sigma} \mathcal{H}(R) \\ &\stackrel{(9.2)}{\equiv} c(\mathbf{data}(\Omega_0), \sigma) (\varrho/R)^{n-\sigma} \mathcal{H}(R). \end{aligned}$$

Finally, in the case  $\varrho < \tau_1 \tau_2^{m+1} R$  we again find  $\tilde{k} \geq 1$  such that  $\tau_1^{\tilde{k}+1} \tau_2^{m+1} R \leq \varrho \leq \tau_1^{\tilde{k}} \tau_2^{m+1} R$  and therefore we have

$$\begin{aligned} \mathcal{H}(\varrho) &\leq \mathcal{H}(\tau_1^{\tilde{k}} \tau_2^{m+1} R) \stackrel{(9.17)}{\leq} \tau_1^{(n-\sigma)\tilde{k}} \mathcal{H}(\tau_2^{m+1} R) \\ &\stackrel{(9.18)}{\leq} c \left( \tau_1^{\tilde{k}} \tau_2^{m+1} \right)^{n-\sigma} \mathcal{H}(R) \end{aligned}$$

$$\leq c\tau_1^{\sigma-n}(\varrho/R)^{n-\sigma}\mathcal{H}(R) \equiv c(\varrho/R)^{n-\sigma}\mathcal{H}(R) ,$$

where again it is  $c \equiv c(\mathbf{data}(\Omega_0), \sigma)$  again by the dependence on the constants in (9.2). Summarizing the estimates in the last three displays, we have established

$$(9.19) \quad \mathcal{H}(\varrho) \leq c(\mathbf{data}(\Omega_0), \sigma)(\varrho/R)^{n-\sigma}\mathcal{H}(R) \quad 0 < \varrho \leq R \leq R_2 ,$$

so that (1.14) follows in this case too provided  $R \leq R_2$ ,  $R_2$  is as in (9.10). We finally settle the remaining case  $R_2 < R \leq 1$ . For  $R_2 \leq \varrho \leq R \leq 1$ , we can trivially estimate

$$\begin{aligned} \mathcal{H}(B_\varrho) &\leq (\varrho/R)^{\sigma-n}(\varrho/R)^{n-\sigma}\mathcal{H}(B_R) \\ &\leq R_2^{\sigma-n}(\varrho/R)^{n-\sigma}\mathcal{H}(B_R) \stackrel{(9.10)}{\equiv} c(\mathbf{data}(\Omega_0), \sigma)(\varrho/R)^{n-\sigma}\mathcal{H}(B_R) . \end{aligned}$$

On the other hand, if  $\varrho \leq R_2 \leq R \leq 1$ , we similarly have

$$\mathcal{H}(B_\varrho) \stackrel{(9.19)}{\leq} c(\varrho/R_2)^{n-\sigma}\mathcal{H}(B_{R_2}) \leq R_2^{\sigma-n}(\varrho/R)^{n-\sigma}\mathcal{H}(B_R) \equiv c(\varrho/R)^{n-\sigma}\mathcal{H}(B_R) .$$

We have therefore established the decay estimate in (1.14) for every possible choice of  $\varrho, R$  such that  $0 < \varrho \leq R \leq 1$  and the proof of Theorem 1.2 is complete.

## 10. GRADIENT CONTINUITY AND THEOREM 1.1

In this section we prove Theorem 1.1. The way we combine the  $p$ - and  $(p, q)$ -phases is different. Specifically, we shall use condition (6.5), again with  $s = 0$ , but with  $M$  depending on  $R$ , in a way that makes it blow-up when  $R \rightarrow 0$ . This would create a bad dependence on the constants when using Lemma 7.3 with the same choice of  $s$ . To eliminate this bad dependence we then consider the  $p$ -phase (6.4) with a suitable positive exponent  $s$ , depending on the way  $M$  has been chosen. Iterating this argument along with a comparison scheme finally leads to desired Hölder continuity. To proceed, we first appeal to Theorem 1.2. This, via a standard covering argument, gives that for every open subset  $\Omega_0 \Subset \Omega$  and  $\kappa > 0$ , there exists a constant  $c \equiv c(\mathbf{data}(\Omega_0), \kappa)$  such that

$$(10.1) \quad \int_{B_r} H(x, Du) \, dx \leq cr^{-\kappa}$$

holds for every ball  $B_r \Subset \Omega_0 \subset \Omega$ . We then consider a ball  $B_R \Subset \Omega_0$ ,  $R \leq 1$ , and assume that condition (6.5) holds with the choice  $s = 0$  and  $M \equiv M(R)$  as follows:

$$(10.2) \quad a_i(B_R) > M(R)[a]_{0,\alpha}R^\alpha \quad \text{with} \quad M \equiv M(R) := R^{-s_m/2} ,$$

where  $s_m$  has been defined in (6.2). Using (8.5) and recalling (1.9), we conclude that the inequality

$$(10.3) \quad \begin{aligned} &\int_{B_R} \left( |V_p(Du) - V_p(D\tilde{h})|^2 + a_i(B_R)|V_q(Du) - V_q(D\tilde{h})|^2 \right) dx \\ &\leq cR^{\min\{\gamma\beta, s_m/2\}} \int_{B_R} H(x, Du) \, dx , \end{aligned}$$

holds for a constant  $c \equiv c(\mathbf{data}(\Omega_0))$ . Here  $\tilde{h}$  is defined as in (7.2). Then, we consider the case (10.2) does not hold, and this leads to

$$(10.4) \quad a_i(B_R) \leq M(R)[a]_{0,\alpha}R^\alpha = [a]_{0,\alpha}R^{\alpha-s_m/2} .$$

This means that condition (6.4) holds with the choice  $M = 1$  and  $s = s_m/2$ . We are therefore in position to apply Lemma 7.4, the advantage being now that, since now  $M = 1$ , then the constant  $c$  and the exponent  $s_1$  appearing in (7.54) are determined

as functions  $\mathbf{data}(\Omega_0)$  while no dependence with respect to  $R$  occurs. Inequality (7.54) now gives

$$(10.5) \quad \begin{aligned} & \int_{B_{R/8}} \left( |V_p(Du) - V_p(D\tilde{h})|^2 + a_i(B_R) |V_q(Du) - V_q(D\tilde{h})|^2 \right) dx \\ & \leq cR^{\min\{\gamma\beta s_1, s_1 s_m / (2q)\}} \int_{B_R} H(x, Du) dx . \end{aligned}$$

Summarizing the contents of (10.3) and (10.5) and combining them with (10.1) for the choice

$$(10.6) \quad \kappa \equiv \kappa_1 \equiv \kappa_1(\mathbf{data}(\Omega_0)) := \frac{1}{2} \min \left\{ \gamma\beta, \gamma\beta s_1, \frac{s_m}{2}, \frac{s_1 s_m}{2q} \right\} ,$$

we conclude with

$$(10.7) \quad \int_{B_{R/8}} \left( |V_p(Du) - V_p(D\tilde{h})|^2 + a_i(B_R) |V_q(Du) - V_q(D\tilde{h})|^2 \right) dx \leq cR^{\kappa_1} ,$$

and  $c \equiv c(\mathbf{data}(\Omega_0))$ . Notice that here  $\tilde{h}$  is defined either via (8.3) or via Lemma 7.4 (and minimizes the functional in (7.53)), accordingly to which inequality (10.2) or (10.4) comes into the play, respectively. Therefore it satisfies

$$(10.8) \quad \int_{B_{R/8}} \left( |D\tilde{h}|^p + a_i(B_R) |D\tilde{h}|^q \right) dx \leq c \int_{B_R} H(x, Du) dx ,$$

for  $c \equiv c(n, \nu, L)$ . This follows from estimates (8.4) and (7.55). We can conclude with a comparison scheme, which follows with some variants the one in [16]. Notice that, arguing exactly as in (5.26)-(5.27) (take  $\theta \equiv 1$  and  $a_0 \equiv a_i(B_R)$  there), from the relation (10.7) it follows that

$$(10.9) \quad \int_{B_{R/8}} \left( |Du - D\tilde{h}|^p + a_i(B_R) |Du - D\tilde{h}|^q \right) dx \leq cR^{\kappa_2} ,$$

for a new positive exponent  $\kappa_2 \equiv \kappa_2(\kappa_1, n, p, q) \in (0, 1)$ , computable as in (5.26)-(5.27) once  $\kappa_1$  is fixed as in (10.6); all in all we have  $\kappa_2 \equiv \kappa_2(\mathbf{data}(\Omega_0))$ . We observe that in both cases (7.2) and (8.3),  $\tilde{h}$  is a minimizer of a functional whose integrand depends only on the gradient and satisfies assumptions (1.8)<sub>1,2</sub> with  $a(x) \equiv a_i(B_R)$ . The theory in [39] applies and provides the following gradient Hölder decay estimate, valid whenever  $0 < \varrho \leq R/8$ :

$$(10.10) \quad \begin{aligned} & \int_{B_\varrho} \left( |D\tilde{h} - (D\tilde{h})_{B_\varrho}|^p + a_i(B_R) |D\tilde{h} - (D\tilde{h})_{B_\varrho}|^q \right) dx \\ & \leq c \left( \frac{\varrho}{R} \right)^{p\tilde{\beta}} \int_{B_R} \left( |D\tilde{h}|^p + a_i(B_R) |D\tilde{h}|^q \right) dx \stackrel{(10.8)}{\leq} c \left( \frac{\varrho}{R} \right)^{p\tilde{\beta}} \int_{B_R} H(x, Du) dx , \end{aligned}$$

and where  $c$  and  $\tilde{\beta}$  both depend only on  $n, p, q, \nu, L$ , but are otherwise independent of the specific value of  $a_i(B_R)$ ; with no loss of generality we assume it is  $p\tilde{\beta} < 1$ . Then, by also using (7.17), for  $0 < \varrho \leq R/8$  we have

$$(10.11) \quad \begin{aligned} & \int_{B_\varrho} |Du - (Du)_{B_\varrho}|^p dx \leq c \int_{B_\varrho} |D\tilde{h} - (D\tilde{h})_{B_\varrho}|^p dx + c \int_{B_\varrho} |Du - D\tilde{h}|^p dx \\ & \stackrel{(10.10)}{\leq} c \left( \frac{\varrho}{R} \right)^{p\tilde{\beta}} \int_{B_R} H(x, Du) dx + c \left( \frac{R}{\varrho} \right)^n \int_{B_{R/8}} |Du - D\tilde{h}|^p dx \\ & \stackrel{(10.1), (10.9)}{\leq} c \left( \frac{\varrho}{R} \right)^{p\tilde{\beta}} R^{-\kappa} + c \left( \frac{R}{\varrho} \right)^n R^{\kappa_2} \end{aligned}$$

for  $c \equiv c(\mathbf{data}(\Omega_0), \kappa)$ ; notice that  $\kappa \in (0, 1)$  can still be chosen, while  $\kappa_2$  has been fixed through (10.9) and depends on  $\mathbf{data}(\Omega_0)$ . Choosing  $\kappa \equiv \kappa_2 p \tilde{\beta} / (8n)$ , and

taking  $\varrho = (R/8)^{1+\kappa_2/(4n)}$  in (10.11), after a few elementary manipulations we get that

$$(10.12) \quad \int_{B_{\varrho}} |Du - (Du)_{B_{\varrho}}|^p dx \leq c\varrho^{\frac{\kappa_2 p \tilde{\beta}}{16n}}$$

holds for every  $\varrho \in (0, 1/8)$ , provided  $B_{8\varrho} \Subset \Omega_0$ . By the already invoked integral characterisation of Hölder continuity due to Campanato and Meyers, and a standard covering argument, (10.12) implies that  $Du \in C_{loc}^{0, \beta_0}(\Omega)$  for  $\beta_0 = \kappa_2 \tilde{\beta}/(16n)$ . This proves the local Hölder continuity of  $Du$  but not yet the full statement of Theorem 1.1, that claims that the Hölder continuity exponent  $\beta_0$  of  $Du$  depends only on  $n, p, q, \nu, L, \alpha$  and  $\beta$ , while we have that the exponent found  $\kappa_2 \tilde{\beta}/(16n)$  depends on  $\mathbf{data}(\Omega_0)$ . This exponent can be upgraded to reach the required dependence. Indeed, once we know that the gradient is locally bounded, the non-uniform ellipticity of the functional  $\mathcal{F}$  becomes immaterial and we can apply some standard perturbation methods in order to obtain full statement of Theorem 1.1. We briefly summarize the steps. We go back to Section 8 and perform the same procedure as for Lemma 8.1; we do this in any case, independently of the occurrence of condition (6.5). Let us observe that the functional in (7.53) satisfies the Bounded Slope Condition (see for instance [10]) and therefore there exists a constant  $c \equiv c(n, p, q, \nu, L, \|Du\|_{L^\infty(B_R)})$  such that  $\|D\tilde{h}\|_{L^\infty(B_R)} \leq c$ . Using this fact, the estimation of the terms  $II_1, \dots, II_6$  in Lemma 8.1 simplifies and leads to replace (8.5) by the stronger

$$(10.13) \quad \int_{B_R} |V_p(Du) - V_p(D\tilde{h})|^2 dx \leq cR^{\min\{\beta, \alpha\}},$$

where the constant now depends on  $n, p, q, \nu, L, \|Du\|_{L^\infty(\Omega_0)}, \|a\|_{L^\infty(\Omega_0)}$ . Notice also that we can take  $\gamma = 1$  in Lemma 8.1 as now  $u$  is locally Lipschitz. Eventually, (10.13) gives

$$(10.14) \quad \int_{B_R} |Du - D\tilde{h}|^2 dx \leq cR^\kappa,$$

where  $\kappa = \min\{\beta, \alpha\}$  if  $p \geq 2$  and  $2\kappa = p \min\{\beta, \alpha\}$  otherwise. Similarly, (10.10) now simplifies in

$$(10.15) \quad \int_{B_{\varrho}} |D\tilde{h} - (D\tilde{h})_{B_{\varrho}}|^p dx \leq c \left(\frac{\varrho}{R}\right)^{p\tilde{\beta}}$$

where  $\tilde{\beta}$  depends only on  $n, p, q, \nu, L$ , and the constant  $c$  depends also on  $\|Du\|_{L^\infty(\Omega_0)}$  and  $\|a\|_{L^\infty(\Omega_0)}$ . The inequalities in (10.14)-(10.15) can be now combined in the same way (10.11) and (10.12) have been combined above to deduce the gradient Hölder continuity. This time all the exponents involved depend only on  $n, p, q, \nu, L, \alpha, \beta$ . This means that the local Hölder continuity of  $Du$  follows with an exponent  $\beta_0$  with the dependence on the various constants described in the statement and the proof is finally complete.

## 11. INTERPOLATIVE EFFECTS AND THEOREM 1.3

The proof of Theorem 1.3 upgrades the one for Theorem 1.1. In the following, the definition in (6.2) will be replaced by  $s_m := p + \alpha - q + \gamma(q - p)$ , where  $\gamma \in (0, 1]$  is the Hölder continuity exponent coming from assumptions (1.17) and not any longer the one coming from Theorem 3.2, whose assumptions are in general not implied by (1.17). In this respect, we shall restrict to the case  $p \leq n/(1 - \gamma)$ . This is the range for which (1.17) is a weaker assumption than (1.2). See also the comments in the Introduction after Theorem 1.3.

**11.1. Step 1: Blow-up, first time.** We start revisiting the proofs in Section 7 in order to get Lemma 7.5 in this case, with obviously a different function  $R_*$  defined in (7.18). We notice that since the functional of Theorem 1.3 does not depend on the  $u$ -variable, then there is no reason to go through Lemma 7.1 and in the following we shall revise the proofs of Section 7 with  $\tilde{v} \equiv u$ . For this reason from Lemma 7.2 we only retain (7.12). We then go through Lemma 7.3, where  $v$  defined in (7.20), is actually the blow-up of the original minimizer  $u$  in  $B_R \equiv B_R(x_0)$ :

$$(11.1) \quad v(x) := \frac{u(x_0 + Rx) - (u)_{B_{R/2}}}{E(R)}, \quad E(R) := \left( \int_{B_{R/2}} |u - (u)_{B_{R/2}}|^p dx \right)^{1/p}.$$

The main point is that we shall use the non-quantitative harmonic approximation of Lemma 5.2 instead of the quantitative one in Lemma 5.1. This essentially depends on the fact that at this stage we are not yet able to verify the conditions to apply Theorem 3.1 to  $v$  and get the higher integrability in (7.32), which is essential for Lemma 5.1. For the same reason, we can by-pass Step 2 of the proof of Lemma 7.3. We therefore go to Step 3 for Lemma 7.3 and note that the computations for in (7.39)-(7.40) keep on being true in this setting; see also Remark 7.2. These lead to establish (7.45) also in this case. Now, fix  $\tilde{\varepsilon} \in (0, 1)$  and determine the corresponding  $\sigma$  from Lemma 5.2. We can find

$$(11.2) \quad R^* \equiv R^*(\mathbf{data}(\Omega_0), \omega(\cdot), M, \tilde{\varepsilon}) > 0$$

such that  $c_h[o(R)] \leq \sigma$  whenever  $R \leq R^*$ , so that (7.45) gives that

$$\left| \int_{B_{1/8}} \langle \partial \bar{F}_0(Dv), D\varphi \rangle dx \right| \leq \sigma \|D\varphi\|_{L^\infty}$$

holds for every function  $\varphi \in W_0^{1,\infty}(B_{1/8})$ . Recall that in the present setting  $\bar{F}_0(\cdot)$  is defined as  $z \mapsto [R/E(R)]^p F(x_B, [E(R)/R]z)$ . We are therefore in condition to apply Lemma 5.2. This yields the existence of a function  $h \in u + W_0^{1,\bar{H}_0}(B_{1/8})$ , with  $\bar{H}_0(\cdot)$  as in (7.34), such that (7.47) holds for all  $\varphi \in W_0^{1,\infty}(B_{1/8})$  and such that

$$(11.3) \quad \int_{B_{1/8}} \bar{H}_0(Dh) dx \leq c(\mathbf{data}(\Omega_0), M)$$

is true and (7.51) holds. These are exactly the ingredients to proceed as in Step 6, thereby getting the validity of Lemma 7.3. We can then proceed as in Lemma 7.5 - actually in an easier way as  $\tilde{v} \equiv u$ . Indeed, (7.63) does not any longer take place and (7.19) (recall  $\tilde{v} \equiv u$ ) directly implies

$$\int_{B_{\tau R}} \left| \frac{u - (u)_{B_{\tau R}}}{\tau R} \right|^p dx \leq c \{ \tau^n + \tau^{-p\bar{\varepsilon}} \} \int_{B_R} H(x, Du) dx,$$

for every  $\bar{\varepsilon} \in (0, 1)$ ,  $\tau \in (0, 1/32)$ , where  $c \equiv c(\mathbf{data}(\Omega_0), M)$  and provided  $R \leq R^*$ , with  $R^*$  being defined in (11.2). Arguing exactly as after (7.64) we get the assertion of Lemma 7.5, with this time  $R_* \equiv R^*$ . This concludes the analysis of the  $p$ -phase. It is now not very difficult to see that all the arguments of Sections 8 and 9 can be reproduced verbatim, thereby leading to the analog of Theorem 1.2 under assumptions (1.17), as announced in the statement of Theorem 1.3. In particular, estimate (1.14) holds and this will be exploited in the next step.

**11.2. Step 2:  $C^{0,\gamma}$ -regularity in rigid form.** We now consider a ball  $B_R \Subset \Omega_0 \Subset \Omega$  with  $R \leq 1$  and fix  $\theta \in (0, 1)$ . In estimate (1.14) we take  $\sigma = p - p\theta$ . Using

Poincaré inequality in a standard way we get that

$$(11.4) \quad \int_{B_\rho} |u - (u)_{B_\rho}|^p dx \leq c \left[ r^{p-p\theta} \int_{B_r} H(x, Du) dx \right] \rho^{p\theta}$$

holds whenever  $B_\rho \subset B_r$  are concentric balls that are contained in  $B_R$  (but this time not necessarily concentric to  $B_R$ ), and where  $c \equiv c(\mathbf{data}(\Omega_0), \theta)$ . We are now able to apply the standard integral characterization of Hölder continuity due to Campanato and Meyers, that gives, together with a standard covering argument yields the following local  $C^{0,\theta}$ -estimate for  $u$ :

$$(11.5) \quad R^\theta [u]_{0,\theta;B_{R/4}} \leq c \left[ R^p \int_{B_{3R/8}} H(x, Du) dx \right]^{1/p}.$$

The constant  $c$  depends on  $\mathbf{data}(\Omega_0)$ . We now further assume that condition (6.4) holds with  $s = 0$  and for some  $M \geq 1$ . This immediately gives the validity of inequality (6.9) from Lemma 6.1; see also Remark 6.4. Taking  $\theta = \gamma$  in (11.5), and using (6.9) - with  $t = 3R/8$  and  $s = R/2$  - we conclude with

$$(11.6) \quad R^\gamma [u]_{0,\gamma;B_{R/4}} \leq c \left( \int_{B_{R/2}} |u - (u)_{B_{R/2}}|^p dx \right)^{1/p} = cE(R),$$

where  $c \equiv c(\mathbf{data}(\Omega_0), M)$ . This is the estimate we were looking for. Notice that this estimate, from a qualitative viewpoint, tells no new about the regularity of  $u$ , which is indeed assumed to be  $C^{0,\gamma}$ -regular. The main point here is the specific form of the a priori estimate (11.6), that will eventually allow to implement a blow-up procedure totally similar to the one of Lemma 7.3.

**11.3. Step 3: Blow-up, second time.** The only missing part for the proof of Theorem 1.1 is the one of Section 10, which is in turn based on Lemma 7.4, with the crucial quantitative estimate (7.54) from Lemma 7.3. This, in turn, comes from estimate (7.49), which is a direct consequence of the application of Lemma 5.1. To apply Lemma 5.1 in the setting of Lemma 7.3, we need to verify the higher integrability information (7.32). This was the missing information in Step 1, forcing us to apply the weaker Lemma 5.2. Estimate (7.32) comes from the application of Theorem 3.1 to  $v$ , that minimizes the rescaled functional in (7.22); this satisfies the new growth assumptions in (7.25)-(7.26). Under the new assumptions (1.17), the application of Theorem 3.1 requires essentially two ingredients: the  $\alpha$ -Hölder continuity of the rescaled coefficient  $[E(R)/R]^{q-p} a_R(\cdot)$  and the fact that  $[v]_{0,\gamma;B}$  is bounded; both information must be uniform with respect to  $R$ . The former can be checked exactly as in (7.30), see Remark 7.1. For the latter estimate (11.6) becomes crucial. Indeed, we have

$$[v]_{0,\gamma;B_{1/4}} \stackrel{(11.1)}{=} \frac{R^\gamma [u]_{0,\gamma;B_{R/4}}}{E(R)} \stackrel{(11.6)}{\leq} c(\mathbf{data}(\Omega_0), M).$$

Summarizing, we have that  $v$  is a local minimizer of the functional in (7.22) to which we can apply Theorem 3.1, with all the constants involved being independent of  $R$ . We get (7.32) and (7.33), with  $\delta_1$  depending only on  $\mathbf{data}(\Omega_0)$  and  $M$ . This - see also Remark 7.2 - eventually allows to get (7.47)-(7.49) with constants  $c$  being completely independent of  $R$ . This is sufficient to reprove Lemma 7.4 with the same dependence of the constants of the original statement, and in particular allows to get the comparison estimate (7.54). These are the essential ingredients to run the proof in Section 10 and eventually leads to the assertion of Theorem 1.3. The rest of the proof now goes exactly as in Lemma 7.3 - see also Remark 7.2 - and in the case of Theorem 1.1.

## 12. LAVRENTIEV PHENOMENON AND THEOREM 1.4

In this section we prove Theorem 1.4. The proof is unified for the two cases (1.3) and (1.19). In the first one we shall formally take the parameter  $\gamma$  considered in (1.19) as  $\gamma = 0$ . We start fixing a ball  $B \Subset \Omega$  as in the statement of Theorem 1.4. We notice that we can reduce ourselves to prove that there exists a sequence  $\{u_k\}$  of  $W^{1,\infty}(B)$ -regular functions such that  $u_k \rightarrow u$  strongly in  $W^{1,p}(B)$  and such that

$$(12.1) \quad \lim_k \int_B H(x, Du_k) dx = \int_B H(x, Du) dx .$$

In fact, once (12.1) is at hand, then (1.20) follows, up to a not relabelled subsequence, by a standard variant of Lebesgue's dominated convergence theorem since (1.6) is in force. For  $\varepsilon \in (0, 1/2)$  sufficiently small to have  $\varepsilon \leq \text{dist}(B, \partial\Omega)/20$ , we consider the mollified functions  $u_\varepsilon := u * \rho_\varepsilon$ . Here  $\{\rho_\varepsilon\}$  is a family of standard mollifiers generated by a smooth non-negative and radial function  $\rho \in C_0^\infty(B_1)$  such that  $\|\rho\|_{L^1(\mathbb{R}^n)} = 1$ , via  $\rho_\varepsilon(x) = \varepsilon^{-n} \rho(x/\varepsilon)$ . Next, following the notation introduced in Section 6, for every  $x \in B$  we consider

$$(12.2) \quad a_i(B_{2\varepsilon}(x)) := \min_{y \in B_{2\varepsilon}(x)} a(y) \quad \text{and} \quad H_\varepsilon(x, z) := |z|^p + a_i(B_{2\varepsilon}(x))|z|^q .$$

We now again distinguish two different phases as done in Section 6.  
*p-phase.* Here we assume that

$$(12.3) \quad a_i(B_{2\varepsilon}(x)) \leq 2[a]_{0,\alpha} \varepsilon^\alpha .$$

As seen in Remark 6.1, this also implies that

$$(12.4) \quad \|a\|_{L^\infty(B_{2\varepsilon}(x))} \leq 6[a]_{0,\alpha} \varepsilon^\alpha .$$

Moreover, in the case  $q \leq p + \alpha$  and we are considering (1.3), the Caccioppoli inequality (6.9) from Lemma 6.1 applies by (12.3) and gives, in particular

$$(12.5) \quad \int_{B_\varepsilon(x)} |Du|^p dy \leq c \int_{B_{2\varepsilon}(x)} \left| \frac{u - (u)_{B_{2\varepsilon}(x)}}{\varepsilon} \right|^p dy .$$

This inequality continues to hold when assuming (1.19). Indeed, in any case, by (3.1) we have

$$\int_{B_\varepsilon(x)} |Du|^p dy \leq c \int_{B_{2\varepsilon}(x)} \left[ \left| \frac{u - (u)_{B_{2\varepsilon}(x)}}{\varepsilon} \right|^p + a(y) \left| \frac{u - (u)_{B_{2\varepsilon}(x)}}{\varepsilon} \right|^q \right] dy .$$

In turn we estimate, for  $y \in B_\varepsilon(x)$

$$\begin{aligned} a(y) \left| \frac{u - (u)_{B_{2\varepsilon}(x)}}{\varepsilon} \right|^q &\stackrel{(12.4)}{\leq} c \left[ \text{osc}_{B_{2\varepsilon}(x)} u \right]^{q-p} \varepsilon^{p-q+\alpha} \left| \frac{u - (u)_{B_{2\varepsilon}(x)}}{\varepsilon} \right|^p \\ &\stackrel{(1.19)}{\leq} c \varepsilon^{(p-q)(1-\gamma)+\alpha} \left| \frac{u - (u)_{B_{2\varepsilon}(x)}}{\varepsilon} \right|^p \\ &\stackrel{(1.19)}{\leq} c \left| \frac{u - (u)_{B_{2\varepsilon}(x)}}{\varepsilon} \right|^p , \end{aligned}$$

with  $c$  being independent of  $\varepsilon$ . Combining the last two inequalities gives (12.5) also in the case. In turn, by using the very definition of convolution, we can estimate

$$(12.6) \quad |Du_\varepsilon(x)| \leq c \int_{B_\varepsilon(x)} |Du| dy \leq c \left( \int_{B_\varepsilon(x)} |Du|^p dy \right)^{1/p} \stackrel{(12.5)}{\leq} \frac{c}{\varepsilon^{1-\gamma}} .$$

In the last line we have used the fact that  $u \in C^{0,\gamma}(\Omega)$  by assumption (1.19) (when this is considered) and we have incorporated the case  $u \in L^\infty$  and (1.3) in the

occurrence  $\gamma = 0$ . Recalling the definition in (12.2), we continue to estimate

$$\begin{aligned} H(x, Du_\varepsilon(x)) &\leq [a(x) - a_i(B_{2\varepsilon}(x))] |Du_\varepsilon(x)|^q + H_\varepsilon(x, Du_\varepsilon(x)) \\ &\leq c[a]_{0,\alpha} \varepsilon^\alpha |Du_\varepsilon(x)|^{q-p} |Du_\varepsilon(x)|^p + H_\varepsilon(x, Du_\varepsilon(x)) \\ &\stackrel{(12.6)}{\leq} c\varepsilon^{\alpha+(q-p)(\gamma-1)} |Du_\varepsilon(x)|^p + H_\varepsilon(x, Du_\varepsilon(x)) \\ &\stackrel{(1.19)}{\leq} c|Du_\varepsilon(x)|^p + H_\varepsilon(x, Du_\varepsilon(x)) . \end{aligned}$$

We therefore conclude that

$$(12.7) \quad H(x, Du_\varepsilon(x)) \leq cH_\varepsilon(x, Du_\varepsilon(x)) ,$$

where  $c$  is independent of  $\varepsilon$ . Finally, Jensen's inequality and (12.2) yield

$$\begin{aligned} H_\varepsilon(x, Du_\varepsilon(x)) &\leq \int_{B_\varepsilon(x)} H_\varepsilon(x, Du(y)) \rho_\varepsilon(x-y) dy \\ &\leq \int_{B_\varepsilon(x)} H(y, Du(y)) \rho_\varepsilon(x-y) dy = [H(\cdot, Du(\cdot)) * \rho_\varepsilon](x) . \end{aligned}$$

The inequalities in the last two displays give that

$$(12.8) \quad H(x, Du_\varepsilon(x)) \leq c[H(\cdot, Du(\cdot)) * \rho_\varepsilon](x)$$

holds for every  $x \in B$  and for a constant  $c \equiv c(\mathbf{data})$  which is independent of  $\varepsilon$ .

*(p, q)-phase.* This is when  $a_i(B_{2\varepsilon}(x)) > 2[a]_{0,\alpha} \varepsilon^\alpha$ . By (6.8) we have that  $\|a(x)\|_{L^\infty(B_{2\varepsilon}(x))} \leq 3a_i(B_{2\varepsilon}(x))$ . This means that  $H(y, z) \leq 3H_\varepsilon(x, z)$  for every  $y \in B_{2\varepsilon}(x)$  and  $z \in \mathbb{R}^n$ , and in particular that  $H(x, z) \leq 3H_\varepsilon(x, z)$ . We conclude again with (12.7) and, proceeding as in the subsequent displays, we again arrive at (12.8), which is established in every case/phase.

With (12.8) at our disposal we are now ready to conclude the proof of (1.20) and therefore of the whole Theorem 1.4. We let  $u_k := u_{\varepsilon_k}$ , where  $\{\varepsilon_k\} \subset (0, 1/2)$  is a sequence such that  $\varepsilon_k \rightarrow 0$  and  $\varepsilon_k \leq \text{dist}(B, \partial\Omega)/20$  for every  $k \in \mathbb{N}$ . Obviously  $u_k \in W^{1,\infty}(B)$  and  $u_k \rightarrow u$  in  $W^{1,p}(B)$  since  $u \in W^{1,p}(B)$ . Now, up to passing to non-relabelled subsequences, we may assume that  $H(x, Du_k(x)) \rightarrow H(x, Du(x))$  a.e. in  $B$ . Moreover, since  $H(\cdot, Du(\cdot)) \in L^1(B)$ , we have that  $[H(\cdot, Du(\cdot)) * \rho_{\varepsilon_k}](x) \rightarrow H(x, Du(x))$  in  $L^1(B)$ . Therefore, by (12.8) we can use a well-known variant of Lebesgue's dominated convergence theorem that implies that  $H(x, Du_k(x)) \rightarrow H(x, Du(x))$  in  $L^1(B)$  that, in turn, gives (1.20). The proof is complete.

## REFERENCES

- [1] Acerbi E. & Fusco N.: An approximation lemma for  $W^{1,p}$  functions. *Material instabilities in continuum mechanics (Edinburgh, 1985–1986)*, 1–5, *Oxford Sci. Publ., Oxford Univ. Press, New York, 1988*.
- [2] Acerbi E. & Mingione G.: Regularity results for a class of functionals with non-standard growth. *Arch. Ration. Mech. Anal.* 156 (2001), 121–140.
- [3] Acerbi E. & Mingione G.: Regularity results for electrorheological fluids: the stationary case. *C. R. Math. Acad. Sci. Paris* 334 (2002), 817–822.
- [4] Baroni P.: Riesz potential estimates for a general class of quasilinear equations. *Calc. Var. & PDE* 53 (2015), 803–846.
- [5] Baroni P. & Colombo M. & Mingione G.: Harnack inequalities for double phase functionals. *Nonlinear Anal.* 121 (2015), 206–222 (Special Issue for E. Mitidieri).
- [6] Baroni P. & Colombo M. & Mingione G.: Non-autonomous functionals, borderline cases and related function classes. *St. Petersburg Math. J.* 27 (2016), 347–379 (Special Issue for N. Uraltseva).
- [7] Baroni P. & Colombo M. & Mingione G.: Flow of double phase functionals. *Forthcoming*.
- [8] Baroni P. & Lindfors C.: The Cauchy-Dirichlet problem for a general class of parabolic equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 34 (2017), 593–624.

- [9] Bildhauer M. & Fuchs M.:  $C^{1,\alpha}$ -solutions to non-autonomous anisotropic variational problems. *Calc. Var. & PDE* 24 (2005), 309–340.
- [10] Bousquet P. & Brasco L.: Global Lipschitz continuity for minima of degenerate problems. *Math. Ann.* 366 (2016), 1403–1450.
- [11] Breit D.: New regularity theorems for non-autonomous variational integrals with  $(p, q)$ -growth. *Calc. Var. & PDE* 44 (2012), 101–129.
- [12] 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.
- [13] Byun S.S. & Oh J.: Global gradient estimates for the borderline case of double phase problems with BMO coefficients in nonsmooth domains. *J. Diff. Equ.* 263 (2017), 1643–1693.
- [14] Byun S.S. & Ok J. & Ryu S.: Global gradient estimates for elliptic equations of  $p(x)$ -Laplacian type with BMO nonlinearity. *J. Reine Angew. Math. (Crelles J.)* 715 (2016), 1–38.
- [15] Choe H. J.: Interior behaviour of minimizers for certain functionals with nonstandard growth. *Nonlinear Anal.* 19 (1992), 933–945.
- [16] Colombo M. & Mingione G.: Regularity for double phase variational problems. *Arch. Ration. Mech. Anal.* 215 (2015), 443–496.
- [17] Colombo M. & Mingione G.: Bounded minimisers of double phase variational integrals. *Arch. Ration. Mech. Anal.* 218 (2015), 219–273.
- [18] Colombo M. & Mingione G.: Calderón-Zygmund estimates and non-uniformly elliptic operators. *J. Funct. Anal.* 270 (2016), 1416–1478.
- [19] Cruz-Uribe D. & Hästö P.: Extrapolation and interpolation in generalized Orlicz spaces. *Trans. Amer. Math. Soc.*, to appear.
- [20] Cupini G. & Marcellini P. & Mascolo E.: Regularity of minimizers under limit growth conditions. *Nonlinear Anal.* 153 (2017), 294–310.
- [21] De Giorgi E.: Frontiere orientate di misura minima. *Seminario di Matematica della Scuola Normale Superiore di Pisa, 1960-61*
- [22] Diening L.: Maximal function on Musielak-Orlicz spaces and generalized Lebesgue spaces. *Bull. Sci. Math.* 129 (2005), 657–700.
- [23] Diening L. & Stroffolini B. & Verde A.: The  $\varphi$ -harmonic approximation and the regularity of  $\varphi$ -harmonic maps. *J. Diff. Equ.* 253 (2012), 1943–1958.
- [24] Duzaar F. & Mingione G.: Harmonic type approximation lemmas. *J. Math. Anal. Appl.* 352 (2009), 301–335.
- [25] Eleuteri M. & Marcellini P. & Mascolo E.: Lipschitz estimates for systems with ellipticity conditions at infinity. *Ann. Mat. Pura Appl. (IV)* 195 (2016), 1575–1603.
- [26] Esposito L. & Leonetti F. & Mingione G.: Sharp regularity for functionals with  $(p, q)$ -growth. *J. Diff. Equ.* 204 (2004), 5–55.
- [27] Fonseca I. & Malý J. & Mingione G.: Scalar minimizers with fractal singular sets. *Arch. Rat. Mech. Anal.* 172 (2004), 295–307.
- [28] Giacinta M. & Giusti E.: Differentiability of minima of nondifferentiable functionals. *Invent. Math.* 72 (1983), 285–298.
- [29] Giusti E.: *Direct methods in the calculus of variations*. World Scientific Publishing Co., Inc., River Edge, NJ, 2003.
- [30] Harjulehto P. & Hästö P.: Boundary regularity under generalized growth conditions. *Preprint 2017*.
- [31] Harjulehto P. & Hästö P. & Toivanen O.: Hölder regularity of quasiminimizers under generalized growth conditions. *Calc. Var. & PDE* 56 (2017), no. 2, Paper No. 22, 26 pp.
- [32] Hästö P.: The maximal operator on generalized Orlicz spaces. *J. Funct. Anal.* 269 (2015), 4038–4048.
- [33] Kristensen J. & Mingione G.: The singular set of minima of integral functionals. *Arch. Ration. Mech. Anal.* 180 (2006), 331–398.
- [34] Kristensen J. & Mingione G.: Boundary regularity in variational problems. *Arch. Ration. Mech. Anal.* 198 (2010), 369–455.
- [35] Kuusi T. & Mingione G.: Linear potentials in nonlinear potential theory. *Arch. Ration. Mech. Anal.* 207 (2013), 215–246.
- [36] Kuusi T. & Mingione G.: Vectorial nonlinear potential theory. *J. Europ. Math. Soc.*, to appear.
- [37] Lavrentiev M.: Sur quelques problèmes du calcul des variations. *Ann. Mat. Pura Appl.* 4 (1926), 7–28
- [38] Lieberman G. M.: Boundary regularity for solutions of degenerate elliptic equations. *Nonlinear Anal.* 12 (1988), 1203–1219.
- [39] Lieberman G.M.: The natural generalization of the natural conditions of Ladyzhenskaya and Ural'tseva for elliptic equations. *Comm. PDE* 16 (1991), 311–361.

- [40] Lieberman G.M.: Gradient estimates for a new class of degenerate elliptic and parabolic equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (IV)* 21 (1994), 497–522.
- [41] Manfredi J.J.: Regularity for minima of functionals with  $p$ -growth. *J. Diff. Equ.* 76 (1988), 203–212.
- [42] 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.
- [43] 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.
- [44] Marcellini P.: Regularity and existence of solutions of elliptic equations with  $p, q$ -growth conditions. *J. Diff. Equ.* 90 (1991), 1–30.
- [45] Marcellini P.: Everywhere regularity for a class of elliptic systems without growth conditions *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (IV)* 23 (1996), 1–25.
- [46] Mingione G.: Regularity of minima: an invitation to the Dark Side of the Calculus of Variations. *Appl. Math.* 51 (2006), 355–425.
- [47] Ok J.: Gradient estimates for elliptic equations with  $L^{p(\cdot)} \log L$  growth. *Calc. Var. & PDE* 55 (2016), no. 2, Paper No. 26, 30 pp.
- [48] Ok J.: Regularity of  $\omega$ -minimizers for a class of functionals with non-standard growth. *Calc. Var. & PDE* 56 (2017), no. 2, Paper No. 48, 31 pp.
- [49] Pereira K. & Squassina M.: Existence results for double phase problems via Morse theory. *Comm. Cont. Math.*, to appear.
- [50] Ragusa M.A. & Tachikawa A.: Boundary regularity of minimizers of  $p(x)$ -energy functionals *Ann. IHP, Nonlinear Anal.* 33 (2017) 451–476.
- [51] Ragusa M.A. & Tachikawa A.: Partial regularity of  $p(x)$ -harmonic maps. *Trans. Amer. Math. Soc.* 365 (2013), 3329–3353.
- [52] Tachikawa A. & Usuba S.: Regularity results up to the boundary for minimizers of  $p(x)$ -energy with  $p(x) > 1$ . *manu. math.* 152 (2017), 127–151.
- [53] Uhlenbeck K.: Regularity for a class of non-linear elliptic systems. *Acta Math.* 138 (1977), 219–240.
- [54] Ural'tseva N.N.: Degenerate quasilinear elliptic systems. *Zap. Na. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)* 7 (1968), 184–222.
- [55] 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.
- [56] Zhikov V.V.: Averaging of functionals of the calculus of variations and elasticity theory. *Izv. Akad. Nauk SSSR Ser. Mat.* 50 (1986), 675–710.
- [57] Zhikov V.V.: On Lavrentiev's Phenomenon. *Russian J. Math. Phys.* 3 (1995), 249–269.
- [58] Zhikov V.V.: On some variational problems. *Russian J. Math. Phys.* 5 (1997), 105–116.

PAOLO BARONI, DEPARTMENT OF MATHEMATICAL, PHYSICAL AND COMPUTER SCIENCES, UNIVERSITY OF PARMA, PARCO AREA DELLE SCIENZE 53/A, CAMPUS, 43124 PARMA, ITALY  
*E-mail address:* [paolo.baroni@unipr.it](mailto:paolo.baroni@unipr.it)

MARIA COLOMBO, INSTITUTE FOR THEORETICAL STUDIES, ETH ZÜRICH,, CLAUDIUSSTRASSE 47, CH-8092 ZÜRICH, SWITZERLAND  
*E-mail address:* [maria.colombo@eth-its.ethz.ch](mailto:maria.colombo@eth-its.ethz.ch)

GIUSEPPE MINGIONE, DIPARTIMENTO DI MATEMATICA E INFORMATICA, UNIVERSITÀ DI PARMA, PARCO AREA DELLE SCIENZE 53/A, CAMPUS, 43124 PARMA, ITALY  
*E-mail address:* [giuseppe.mingione@unipr.it](mailto:giuseppe.mingione@unipr.it)