Regularity for double phase variational problems

MARIA COLOMBO & GIUSEPPE MINGIONE

Abstract

We prove sharp regularity theorems for minimisers of a class of variational integrals whose integrand switches between two different types of degenerate elliptic phases, according to the zero set of a modulating coefficient $a(\cdot)$. The model case is given by the functional

$$w \mapsto \int (|Dw|^p + a(x)|Dw|^q)\, dx \ ,$$

where $q > p$ and $a(\cdot) \geq 0$.

Contents

1. Introduction and results . . . 1
2. Notation and preliminaries . . . 9
3. Fractional operators and Theorem 1.6 . . . 12
4. Sharpness, and Lavrentiev phenomenon . . . 17
5. A reverse Hölder inequality in the p-phase . . . 18
6. Estimates in the two phases . . . 31
7. Proof of Theorem 1.3 . . . 35
8. Proof of Theorem 1.2 . . . 39
9. Proof of Theorem 1.1: (1.8) . . . 42
10. Proof of Theorem 1.1: (1.9) . . . 44
11. Estimates for frozen functionals . . . 50

1. Introduction and results

In this paper we prove sharp regularity results for a class of integral functionals, that, originally connected to Homogenization theory [49], [53] and to the Lavrentiev phenomenon [50], [51], [53], present very new and interesting features from the viewpoint of regularity theory. They fall in the realm of those non-uniformly elliptic problems characterised by having non-standard growth conditions [27], [29], [30], [34], [35], [36], [37] as described in Section 1.1 below. In this respect they provide basic examples of energies that cannot be dealt with via the currently available regularity methods and whose treatment has remained an open problem for a while. The primary model we have in mind is given by the functional

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

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

$$1 < p \leq q\,, \qquad 0 \leq a(\cdot) \in C^{0,\alpha}(\Omega)\,, \quad \alpha \in (0,1]\,. \tag{1.2}$$

The significant case occurs when $p < q$. The main feature of the functional $\mathcal{P}$ is the change of ellipticity/growth type occurring on the zero set $\{a(x) = 0\}$. Indeed, while in those points x where $a(x)$ is positive the energy density of $\mathcal{P}$ exhibits a growth/ellipticity in the gradient which is of order q, on $\{a(x) = 0\}$ the energy density has p-growth in the gradient. In his seminal works [48], [49], [51], [52], Zhikov was the first to introduce and study functionals whose integrands change their ellipticity rate according to the point, and, in particular, the one in (1.1). Such functionals provide a useful paradigm for describing the behaviour of strongly anisotropic materials whose hardening properties - linked to the exponent ruling the growth of the gradient variable - drastically change with the point. The coefficient $a(\cdot)$ serves to regulate the mixture between two different materials, with p and q hardening, respectively. In this class of functionals $\mathcal{P}$ appears to be the one exhibiting the most dramatic phase-transition and therefore the most difficult to treat.

The functional $\mathcal{P}$ appears to be a very interesting one also from the point of view of regularity theory. Indeed, while in the standard situation $p = q$ the coefficient $a(\cdot)$ acts in the energy density as a local perturbation of the main elliptic terms, this is not obviously the case when $q > p$, since it is $a(\cdot)$ to dictate the ellipticity rate of the energy density. Coefficients are no longer a perturbation and a new phenomenon emerges: the rate of Hölder continuity of $a(\cdot)$ interacts with the ratio q/p in a crucial yet precise way. Indeed, as shown in [15], [17], when $q/p > 1 + \alpha/n$, minimisers, which are initially only in $W^{1,p}$, are in general not even locally $W^{1,q}$-regular; moreover, the so called Lavrentiev phenomenon appears (see Theorem 4.1 below). Finally, it is possible to construct a minimiser of the functional in (1.1) such that the set of its (essential) discontinuity points has Hausdorff dimension arbitrarily close to $n - p$ when (1.3) fails. This means that minimisers can be nearly as bad as any other $W^{1,p}$-function and any type of regularity is lost. On the other hand, even basic regularity issues like continuity or gradient Hölder continuity of minimisers have remained unsolved due to substantial technical difficulties. In this paper we shall therefore consider the natural condition for the regularity of minima

$$\frac{q}{p} < 1 + \frac{\alpha}{n}\,, \tag{1.3}$$

and prove that this is also a sufficient condition. Indeed, assuming (1.3) we provide a complete regularity theory for general integral functionals of the type

$$W^{1,1}(\Omega) \ni w \mapsto \mathcal{F}(w,\Omega) := \int_\Omega F(x,w,Dw)\,dx \tag{1.4}$$

modelled on (1.1), drawing a full parallel with the classical theory available when $p = q$. In fact, when considering the particular case $p = q$, or $a(x) \equiv 0$, the results here recover the known ones valid in the standard case [31], [32], [33].

Specifically, we shall consider functionals as in (1.4) where the energy density $F\colon \Omega \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is initially only assumed to be a Carathéodory function satisfying the following growth conditions:

$$\nu H(x,z) \le F(x,v,z) \le LH(x,z) \tag{1.5}$$

whenever $z \in \mathbb{R}^n$, $v \in \mathbb{R}$ and $x \in \Omega$, where $0 < \nu \le L$. Here and in the rest of the paper we denote

$$H(x,z) := |z|^p + a(x)|z|^q\,. \tag{1.6}$$

The function $H(x,z)$, with some ambiguity of notation, will be considered both in the case $z \in \mathbb{R}^n$ and $z \in \mathbb{R}$ (and $z \in \mathbb{R}^{N\times n}$ in the case we are considering vector valued minimisers, as for instance in Theorems 1.4-1.5 below). In our situation, that is assuming (1.5), (local) minimisers of $\mathcal{F}$ can be then defined as follows:

Definition 1. A function $u \in W^{1,1}(\Omega)$ is a local minimiser of the functional $\mathcal{F}$ defined in (1.4) if and only if $H(\cdot, Du) \in L^1(\Omega)$ and the minimality condition $\mathcal{F}(u, \operatorname{supp}(u-v)) \le \mathcal{F}(v, \operatorname{supp}(u-v))$ is satisfied whenever $v \in W^{1,1}_{\mathrm{loc}}(\Omega)$ is such that $\operatorname{supp}(u-v) \subset \Omega$.

Similar definitions apply to the case when minimisers are vector valued $u\colon \Omega \to \mathbb{R}^N$, $N \ge 1$. The above definition and the structure of the function $H(\cdot)$ imply that $u \in W^{1,p}(\Omega)$. We shall derive several a priori estimates on local minimisers and for brevity we shall just appeal to them as minimisers. The constants in the estimates will depend only on the starting quantities assigned in the problem, that is on $n, p, q, \nu, L, \alpha, [a]_{0,\alpha}, \|a\|_{L^\infty}$, and, as in all other non-uniformly elliptic problems, on the energy controlled by the coercivity of the functional, that is $\|Du\|_{L^p}$. No dependence will appear on $\|Du\|_{L^q}$. We shall therefore denote in the following

$$\texttt{data} \equiv (n,p,q,\nu,L,\alpha,[a]_{0,\alpha},\|a\|_{L^\infty},\|Du\|_{L^p})\ .$$

Further notation is in Section 2. Our first results are obtained considering only assumptions (1.5). In particular, measurable dependence of $F(\cdot)$ on the variable x is allowed, while no convexity in the gradient variable is assumed. Moreover, the first of the regularity results in the next theorem applies directly to the case we are dealing with vector valued minimisers.

Theorem 1.1 (Basic regularity). *Let $u \in W^{1,p}(\Omega)$ be a local minimiser of the functional $\mathcal{F}$ defined in* (1.4)*, under the assumptions* (1.2), (1.3) *and* (1.5)*. Then:*

– *(Gehring's theory) There exists a positive integrability exponent δ_g, depending only on* `data`*, such that*

$$H(x, Du) \in L^{1+\delta_g}_{\rm loc}(\Omega)\ . \tag{1.7}$$

Moreover, there exists a constant c, again depending only on `data`*, such that the following reverse inequality holds for every ball $B_R \subset \Omega$:*

$$\left(\fint_{B_{R/2}} [H(x, Du)]^{1+\delta_g}\, dx\right)^{1/(1+\delta_g)} \le c \fint_{B_R} H(x, Du)\, dx\,. \tag{1.8}$$

In particular, if $p > n/(1 + \delta_g)$, then u is locally Hölder continuous. Finally, the result extends to the case the minimiser u is vector valued.

– *(De Giorgi's theory) u is locally bounded. Moreover, when $p \le n/(1 + \delta_g)$, for every open subset $\Omega' \Subset \Omega$ there exists $\beta \in (0, 1)$, depending only on $n, p, q, \nu, L, [a]_{0,\alpha}$ and $\|u\|_{L^\infty(\Omega')}$, such that*

$$u \in C^{0,\beta}_{\rm loc}(\Omega')\,. \tag{1.9}$$

Theorem 1.1 is sharp both with respect to the assumptions and with respect to the results obtained. Indeed, when $q/p > 1 + \alpha/n$ minimisers can be discontinuous (see Section 4). In particular, on the contrary of what happens in the classical De Giorgi's theory, a measurable coefficient $a(\cdot)$ does not ensure the continuity of minimisers. We observe that (1.8) reduces to the usual reverse Hölder type inequalities when $a(\cdot) \equiv 0$ or when $p = q$. We also remark that in the standard case $p = q$, the Hölder continuity exponent β in (1.9) does not depend on the solution u. It is not the case here, due to the fact that the functional $\mathcal{F}$ is not uniformly elliptic and exhibits measurable dependence on coefficients (x, v). Model examples covered by Theorem 1.1 are given by functionals of the type

$$w \mapsto \int_\Omega (f_1(x, w, Dw) + a(x) f_2(x, w, Dw))\ dx$$

where $f_1, f_2 \colon \Omega \times \mathbb{R} \times \mathbb{R}^n \to [0, \infty]$ are Carathéodory functions satisfying double sided p and q-growth conditions, respectively. All in all, Theorem 1.1 provides a complete parallel to the classical theory of Giaquinta & Giusti developed in [19] and based on [12], where $p = q$.

We now turn to the maximal regularity; assumptions must be stronger, since under measurable dependence on coefficients and/or with no convexity in the gradient variable z of the energy density $F(\cdot)$, we cannot expect to have more than Theorem 1.1, already when $p = q$. For this reason, and also in order to highlight the main new ideas, we shall this time consider functionals of the type

$$\mathcal{G}(w, \Omega) := \int_\Omega [f(Dw) + a(x) g(Dw)]\, dx\,. \tag{1.10}$$

Here $f, g\colon \mathbb{R}^n \to [0,\infty]$ are $C^1(\mathbb{R}^n)\cap C^2(\mathbb{R}^n\setminus\{0\})$-regular and satisfy the following growth and ellipticity assumptions:

$$\begin{cases} \nu|z|^p \le f(z) \le L|z|^p \\ |\partial^2 f(z)| \le L|z|^{p-2} \\ \nu|z|^{p-2}|\xi|^2 \le \langle \partial^2 f(z)\xi,\xi\rangle \end{cases} \quad \text{and} \quad \begin{cases} \nu|z|^q \le g(z) \le L|z|^q \\ |\partial^2 g(z)| \le L|z|^{q-2} \\ \nu|z|^{q-2}|\xi|^2 \le \langle \partial^2 g(z)\xi,\xi\rangle \end{cases} \tag{1.11}$$

for every $z \in \mathbb{R}^n\setminus\{0\}, \xi \in \mathbb{R}^n$, actually every $z \in \mathbb{R}^n$ when $p,q \ge 2$, while $0<\nu\le L$. For the functional $\mathcal{G}$ our main result is the following:

Theorem 1.2 (Maximal regularity). *Let $u \in W^{1,p}(\Omega)$ be a local minimiser of the functional $\mathcal{G}$ defined in* (1.10)*, under the assumptions* (1.2)*,* (1.3) *and* (1.11)*. There exists $\tilde\beta \in (0,1)$, depending only on n,p,q,ν,L and α, such that $Du \in C^{0,\tilde\beta}_{\mathrm{loc}}(\Omega;\mathbb{R}^n)$.*

Theorem 1.2 is again the best possible both with respect to the bound (1.3) considered and to the type of regularity obtained. Indeed, by a fundamental result of Uraltseva [26], [46], when $a(\cdot)\equiv 0$ the gradient Hölder continuity is the best possible regularity obtainable for minimisers. The analogy goes further; we indeed obtain a few priori regularity estimates which are the exact counterpart of those available for the classical p-Laplacean case, and that, as in the case of (1.8), can be intrinsically formulated using the function $H(\cdot)$. An instance is the following decay estimate, which is actually a key point in the proof of Theorem 1.2, and that allows to get further regularity results, as shown in forthcoming work of the authors:

Theorem 1.3 (Morrey type estimate). *Let $u \in W^{1,p}(\Omega)$ be a local minimiser of the functional $\mathcal{G}$ defined in* (1.10)*, under the assumptions* (1.2)*,* (1.3) *and* (1.11)*. For every $\delta \in (0,n)$, there exists a positive constant c, depending only on* `data` *and δ, such that the decay estimate*

$$\int_{B_\varrho} H(x,Du)\,dx \le c\left(\frac{\varrho}{R}\right)^{n-\delta}\int_{B_R} H(x,Du)\,dx \tag{1.12}$$

holds whenever $0<\varrho\le R\le 1$ and $B_R\subset\Omega$.

Theorem 1.3 already implies that $u \in C^{0,\beta}_{\mathrm{loc}}(\Omega)$ for every $\beta\in(0,1)$. Many of the methods developed for Theorems 1.2-1.3 are general enough to cover vector valued minimisers $u\colon\Omega\to\mathbb{R}^N$ for $N\ge 1$:

Theorem 1.4 (Vectorial maximal regularity). *Let $u \in W^{1,p}(\Omega;\mathbb{R}^N)$ be a local minimiser of the functional $\mathcal{P}$ defined in* (1.1)*, under the assumptions* (1.2) *and* (1.3)*. There exists $\tilde\beta\in(0,1)$, depending only on n,N,p,q,ν,L and α, such that $Du \in C^{0,\tilde\beta}_{\mathrm{loc}}(\Omega;\mathbb{R}^{N\times n})$.*

In the vectorial case dependence on the modulus is necessary since otherwise solutions are known to be discontinuous already in the standard case $p=q$ (see [41]). Theorem 1.4 is the sharp counterpart of the classical result of Uhlenbeck [45] valid for the case $a(\cdot)\equiv 0$. The proof of Theorem 1.4 is essentially the same of the one for Theorem 1.2; see Remark 6 below. When turning to general functionals, we have the following result:

Theorem 1.5 (Higher differentiability for general vectorial functionals). *Let $u \in W^{1,p}(\Omega;\mathbb{R}^N)$ be a local minimiser of the functional $\mathcal{G}$ defined in* (1.10)*, under the assumptions* (1.2)*,* (1.3) *and* (1.11)*. Then it holds that*

$$|Du|^{(p-2)/2}Du \in \mathcal{N}^{\alpha,2}_{\mathrm{loc}}(\Omega,\mathbb{R}^{N\times n})$$

and

$$Du \in L^{np/(n-2\beta)}_{\mathrm{loc}}(\Omega;\mathbb{R}^{N\times n}) \qquad \textit{for every } \beta \in (0,\alpha)\,.$$

The previous theorem - involving the Nikolski space $\mathcal{N}^{\alpha,2}$ - sharply extends previously known differentiability and integrability theorems valid in the p-Laplacean case [8], [24], [40], [41]. It also improves the differentiability results contained in [15] up to the optimal fractional differentiability exponent, which is α. The proof of Theorem 1.5 in the scalar case $N = 1$ is implicitly contained in the one of Theorem 5.1 in Section 5 below. The proof in the genuine vectorial case $N > 1$ follows verbatim and this time the dependence on the modulus of the gradient is not required on the energy density of the functional. Concerning the proof of Theorem 1.5, this is based on a careful difference quotient argument, where the Hölder continuity of $a(\cdot)$ is read as fractional differentiability. We remark that this contains a few technical novelties in the case $p < 2$. Indeed, already in the standard case $p = q$ the classical difference quotient technique was not known to work in the fractional case (i.e. when coefficients are non-differentiable) since it involved an integration by parts argument that required differentiability of coefficients [2]. We overcome this point via a delicate mollification argument that seems to be new already in the classical setting.

The methods developed here are the starting point for more further developments; in particular, in the forthcoming paper [5], we shall study the regularity of parabolic problems having as underlying energy the functional $\mathcal{P}$; the gradient flow of $\mathcal{P}$ will be included. Moreover, when proving the gradient Hölder continuity, more general functionals of the type

$$w \mapsto \int_\Omega (\gamma(x,w)|Dw|^p + a(x)|Dw|^q)\,dx$$

will be considered under specific assumptions on minimisers. Finally, cases in which equality in (1.3) can be reached are treated in [10].

1.1. Connections with non-uniformly elliptic problems

The functionals $\mathcal{P}$, $\mathcal{F}$ and $\mathcal{G}$ above belong to the class of integral functionals having so called (p,q)-growth conditions. These are functionals of the type in (1.4), where the energy density satisfies $\nu|z|^p \leq F(x,v,z) \leq L(|z|^q+1)$. The fundamentals of the corresponding regularity theory have been laid in Marcellini's pioneering papers [35], [36], [37], [38]; further relevant contributions are for instance in [8], [9], [27], [28], [29], [30], [43], [44], [47]. See also [41] for a survey. A main point in the theory is the lack of regularity results for functionals whose integrand depends on x, possibly in a non-smooth way. In this respect, functionals as in (1.1) or (1.4) are the prototype of the worst kind of interplay between the coefficient x and the (p,q)-growth. Up to now the only cases that have been treated are those for which the presence of x causes a modest change in the ellipticity [4], [11], or when special geometries are considered [3]. More can be said when the energy density $F(\cdot)$ is smoother [6], [7] (differentiability with respect to x, non-degeneracy with respect to the gradient). Similar difficulties occur in relaxation and semicontinuity problems for integral functionals (see for instance [1], [23], [34] and in particular [16, Remark 5.4]). Another connection occurs with the non-uniformly elliptic equations, those for which the *ellipticity ratio* between the largest and the smallest eigenvalue might become unbounded. By looking at the Euler-Lagrange equation of the functional in (1.1), that is

$$-\text{div}\,(|Du|^{p-2}Du + (p/q)a(x)|Du|^{q-2}Du) = 0\ ,$$

we notice that the ellipticity ratio is proportional to $1 + a(x)|Du|^{q-p}$. Therefore, it blows-up when $|Du|$ blows-up. So, possible loss of uniform ellipticity links to a subtle interplay between $a(\cdot)$ and Du. It is precisely from this fact that condition (1.3) stems from. Smallness of $a(\cdot)$ around its zero set (that is taking α large as in (1.3)) serves now to compensate the potential blow-up of $|Du|$. In other words, the transition between the two phases (materials) must be fast enough.

1.2. Gradient continuity, universal threshold and two phases

We shall provide here a brief sketch of the proofs. In this section we first describe Theorems 1.2-1.3 and refer to the model case (1.1); in these cases ideas become indeed more transparent. Since singularities of minima can only occur on $\{a(x) = 0\}$, as typical in phase transitions, the natural thing would be in this case to distinguish between points x_0 such that $a(x_0) = 0$ and those for which $a(x_0) > 0$ holds. In the first case - the p-phase - it would be natural to look at the functional

$$w \mapsto \int_{\Omega} |Dw|^p\,dx\,, \tag{1.13}$$

and to try to use the corresponding regularity theory of minima. In the case we have $a(x_0) > 0$ - the (p,q)-phase - one is the led to use the regularity of minima of the "frozen" functional

$$\mathcal{P}_0(w,\Omega) := \int_{\Omega} (|Dw|^p + a_0|Dw|^q)\,dx\,, \tag{1.14}$$

for $a_0 = a(x_0)$. Now, while this approach is natural, how to provide a quantitative version of it is far from being clear. In fact, as the counterexamples show, the two phases can match together only if (1.3) holds, a fact that must be used in a sharp way in the estimates. This also tells that a *naive perturbation argument would fail*. Moreover, since there is in general a very poor control on the zero set of Hölder continuous functions, we cannot take a pointwise path. The idea is to reformulate this alternative on suitable scales, i.e. balls B_R. Indeed we show there exists a *universal threshold*

$$T_s \equiv T_s\,(n,p,q,\nu,L,[a]_{0,\alpha}) > 0$$

independent of the solution u such that, if

$$\sup_{x \in B_R} \frac{a(x)}{R^\alpha} \leq T_s \tag{1.15}$$

holds, then the functional behaves as the one in (1.13) *at the single scale* B_R. In this case the term with q-growth in the gradient can be controlled by the one with p-growth. This is encoded in the fact that a reverse Hölder's inequality with quantitative exponent surprisingly holds on B_R, that is

$$\left(\fint_{B_{R/2}} |Du|^{2q-p} \, dx \right)^{1/(2q-p)} \lesssim \left(\fint_{B_R} |Du|^p \, dx \right)^{1/p} . \tag{1.16}$$

For functionals with (p,q)-growth reverse inequalities of this type do not in general hold, for basic homogeneity reasons. Indeed (1.16) holds provided (1.15) is in force. The drawback is that (1.16) only holds at the scale B_R, and it is not possible to replace B_R by another smaller ball $B_\varrho \subset B_R$, as it happens in the standard case $p = q$. Indeed, (1.15) will be considered in a inductive procedure where (1.15) is checked at every step. On the other hand if (1.15) fails, then the q-component of the energy density is large enough, and we are in the (p,q)-phase. Combining the two alternatives requires that, for a sequence of nested balls

$$.... B_{\tau^k R_0} \subset B_{\tau^{k-1} R_0} \subset .... \subset B_{\tau R_0} \subset B_{R_0} \, , \tag{1.17}$$

we for each ball verify condition (1.15) and perform the related perturbation estimates around the p-Laplacean functional (1.13). If this process never ends, we are approaching the zero set of $a(\cdot)$ and we are done. If this process stops at the exit time ball $B_{\tau^m R_0}$, we enter the (p,q)-phase; we can use the regularity theory available for (1.14) for $a_0 = a(x_0)$ and some $x_0 \in B_{\tau^m R_0}$ (see Section 11 below). At this stage we use the fact that the (p,q)-phase is stable: the functional remains in the (p,q)-phase for all subsequent scales $B_{\tau^{m+h} R_0}$ and we again conclude the proof. With this argument we prove that minimisers are $C^{0,\beta}$-regular *for every* exponent $\beta < 1$ and Theorem 1.3. By using this result and another higher integrability estimate, that this time holds independently of (1.15) - see (5.3) below - we can prove the Hölder continuity of Du and Theorem 1.2. This again involves building-up another version of the same alternative based on (1.15). The implementation of this heuristic scheme involves a certain number of very delicate technicalities and it is not straightforward. For instance, in order to prove (1.16), we will need a few estimates leading to Caccioppoli inequalities in fractional Sobolev spaces taking precisely into account the size of the coefficient $a(\cdot)$ when (1.15) holds. We will also use in a direct way the absence of Lavrentiev phenomenon implied by (1.3); see Theorem 4.1 below.

1.3. Low regularity, Sobolev inequalities, and intrinsic approach

For Theorem 1.1 we introduce an intrinsic approach according to which the function

$H(x, Du)$ plays for $\mathcal{F}$ the same role that $|Du|^p$ plays when considering problems with standard p-growth. This shorter path has two advantages: it encodes several features that can be useful in similar situations, and it emphasises the interconnections between several aspects of the problem. In fact, we first prove a sort of intrinsic Sobolev-Poincaré inequality for the quantity in (1.1). This in turn relies on a maximal inequality - see (3.4) below - which in general is true if and only if (1.3) holds; see Remark 4 below.

Theorem 1.6 (Sobolev-Poincaré inequality). *Let* $1 < p \leq q$ *and* $\alpha \in (0,1]$ *verifying* (1.3). *Then there exist a constant* c *depending only on* $n, p, q, [a]_{0,\alpha}$ *and* $\|Dw\|_{L^p(B_R)}$, *and exponents* $d_1 > 1 > d_2$, *depending only on* n, p, q, α, *such that*

$$\left(\fint_{B_R}\left[H\left(x, \frac{w-(w)_{B_R}}{R}\right)\right]^{d_1} dx\right)^{1/d_1} \leq c\left(\fint_{B_R}[H(x, Dw)]^{d_2}\, dx\right)^{1/d_2} \tag{1.18}$$

holds whenever $w \in W^{1,p}(B_R)$, *and whenever* $B_R \subset \Omega$ *is such that* $R \leq 1$.

For the proof of (1.18) both the specific structure of $H(\cdot)$ and condition (1.3) reveal to be crucial. Inequalities similar to the one in (1.18) can be proved for fractional operators too, see Remark 3 below. A crucial and peculiar feature of such inequalities is that they seem to lie half-way between classical Sobolev type inequalities and weighed inequalities involving Muckenhoupt weights. With Theorem 1.6 available, a modification of the known techniques allows to rapidly get (1.8) and that minimisers are locally bounded. We then proceed with the oscillation reduction of u by implementing an exit time argument around $\{a(x) = 0\}$ as the one described in Section 1.2, but based this time on very different estimates. We consider a series of shrinking balls as in (1.17) and we find here another interesting twist: condition (1.15) also ensures that minimisers satisfy Caccioppoli type inequalities of the type that hold in the standard case $p = q$, modulo correction terms; see Lemma 10.1 below. These terms can be then controlled along the iterations and therefore we reduce the oscillation of u along the chain (1.17) until we (possibly never) reach the exit time ball $B_{\tau^m R_0}$. At the exit time, if any, we cannot perform perturbation arguments around the functional $\mathcal{P}_0$ defined in (1.14), since we are now dealing with $\mathcal{F}$, which has measurable coefficients. We instead observe that on $B_{\tau^m R_0}$ the original minimiser u becomes a so called Q-minimiser of $\mathcal{P}_0$ and use the related theory for this notion (in the version due to Lieberman [27]) to infer the needed estimates and conclude with the proof of Hölder continuity. In fact, we take the opportunity to remark that Theorem 1.1 continues to hold for quasi-minima of the functional $\mathcal{F}$; see Section 11 for the definition.

2. Notation and preliminaries

In what follows we denote by c a general positive constant, possibly varying from line to line; special occurrences will be denoted by $c_1, c_2, c_*, \bar{c}$ or the like. All such constants will always be *larger or equal than one*; moreover 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,ν,L (or *at least* on n,p,ν,L if further dependences are specified). 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: $B_r \equiv B_r(x_0)$. Unless otherwise stated, different balls in the same context will have the same center. We shall also denote $B_1 = B_1(0)$ if not differently specified. With $\mathcal{B} \subset \mathbb{R}^n$ being a measurable subset with positive measure $|\mathcal{B}| > 0$, and with $g\colon \mathcal{B} \to \mathbb{R}^k$, $k \geq 1$, being a measurable map, we shall denote by

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

its integral average. With $a(\cdot)$ being fixed in (1.2), in the following we shall as usual denote

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

For a vector valued function $G : \mathbb{R}^n \to \mathbb{R}^k$ and a vector $h \in \mathbb{R}^n$, we define the finite difference operator

$$\tau_h G(x) \equiv (\tau_h G)(x) = G(x+h) - G(x)\,. \tag{2.1}$$

Given concentric balls $B_{\varrho} \subset B_R$, we shall use the following basic property of finite differences:

$$\int_{B_{\rho}} |\tau_h G|^{\gamma}\,dx \leq |h|^{\gamma} \int_{B_R} |DG|^{\gamma}\,dx\,, \tag{2.2}$$

that holds whenever $\gamma \geq 1$, $G \in W^{1,\gamma}(B_R)$ and $|h| \leq R - \varrho$.

Let us now record a few simple consequences of assumptions (1.11). The last line of (1.11) implies that f and g are convex; this, together with the first line also implies a bound on the first derivatives

$$|\partial f(z)| \leq c_f |z|^{p-1} \quad \text{and} \quad |\partial g(z)| \leq c_g |z|^{q-1}\,, \tag{2.3}$$

for constants $c_f \equiv c_f(n,p,L)$ and $c_g \equiv c_g(n,q,L)$. For this see [21, Lemma 5.2]. We shall extensively use the auxiliary vector fields $V_p, V_q\colon \mathbb{R}^n \to \mathbb{R}^n$ defined by

$$V_p(z) := |z|^{(p-2)/2} z \qquad \text{and} \qquad V_q(z) := |z|^{(q-2)/2} z \tag{2.4}$$

whenever $z \in \mathbb{R}^n$. These maps are very convenient to formulate the monotonicity properties of the vector fields $\partial f(\cdot)$ and $\partial g(\cdot)$, and more in general, of the vector field $z \mapsto |z|^{\gamma-2}z$ for $\gamma > 1$. Indeed, whenever $z_1, z_2 \in \mathbb{R}^n$ it holds that

$$\begin{cases} |V_p(z_1) - V_p(z_2)|^2 \leq c\langle \partial f(z_1) - \partial f(z_2), z_1 - z_2\rangle \\ |V_q(z_1) - V_q(z_2)|^2 \leq c\langle \partial g(z_1) - \partial g(z_2), z_1 - z_2\rangle \\ |V_{\gamma}(z_1) - V_{\gamma}(z_2)|^2 \leq c\langle |z_1|^{\gamma-2}z_1 - |z_2|^{\gamma-2}z_2, z_1 - z_2\rangle\,, \end{cases} \tag{2.5}$$

with constants depending on n, ν and p, q, γ, respectively. We shall often use the inequalities implicit in

$$|V_\gamma(z_1) - V_\gamma(z_2)| \approx (|z_1| + |z_2|)^{(\gamma-2)/2}|z_1 - z_2| \tag{2.6}$$

where the implied constants still depend on n, γ (see [21]). We finally recall that the upper bounds in $(1.11)_2$ together with standard algebraic lemmas (see for instance [21, Lemma 8.3]) imply the following local Lipschitz continuity properties, valid whenever $z_1, z_2 \in \mathbb{R}^n$ (not simultaneously null, otherwise by (2.3) there is nothing to prove):

$$\begin{cases} |\partial f(z_1) - \partial f(z_2)| \leq c(n,p,L)\,(|z_1| + |z_2|)^{p-2}\,|z_2 - z_1| \\ |\partial g(z_1) - \partial g(z_2)| \leq c(n,q,L)\,(|z_1| + |z_2|)^{q-2}\,|z_2 - z_1| \\ ||z_1|^{\gamma-2}z_1 - |z_2|^{\gamma-2}z_2| \leq c(n,\gamma)\,(|z_1| + |z_2|)^{\gamma-2}\,|z_2 - z_1|\,. \end{cases} \tag{2.7}$$

We end this section reporting two by now classical iteration lemmas; see [21, Lemma 7.3] and [21, Lemma 6.1], respectively.

Lemma 2.1. *Let $\phi\colon [0, \tilde R] \to [0, \infty)$ be a non-decreasing function, such that the following inequality holds for some $\varepsilon \geq 0$ and whenever $0 < \varrho \leq R \leq \tilde R$*

$$\phi(\varrho) \leq \tilde c \left[\left(\frac{\varrho}{R}\right)^n + \varepsilon\right] \phi(R)\ .$$

Then for every $\delta \in (0, n)$ there exists $\bar\varepsilon \equiv \bar\varepsilon(n, \delta, \tilde c) > 0$ such that if $\varepsilon \leq \bar\varepsilon$, then

$$\phi(\varrho) \leq \bar c \left(\frac{\varrho}{R}\right)^{n-\delta} \phi(R)$$

holds whenever $0 < \varrho \leq R \leq \tilde R$ and for a constant $\bar c \equiv \bar c(n, \delta, \tilde c)$.

Lemma 2.2. *Let $h : [\rho_0, \rho_1] \to \mathbb{R}$ be a nonnegative and bounded function, and let $\theta \in (0, 1)$ and $A, B \geq 0, \gamma_1, \gamma_2 \geq 0$ be numbers. Assume that*

$$h(t) \leq \theta h(s) + \frac{A}{(s-t)^{\gamma_1}} + \frac{B}{(s-t)^{\gamma_2}}$$

holds for $\rho_0 \leq t < s \leq \rho_1$. Then the following inequality holds with $c \equiv c(\theta, \gamma_1, \gamma_2)$:

$$h(\rho_0) \leq \frac{cA}{(\rho_1 - \rho_0)^{\gamma_1}} + \frac{cB}{(\rho_1 - \rho_0)^{\gamma_2}}\ .$$

Remark 1. In the following we shall use several times a scaling procedure. In order to prove assertions in a ball $B_R \equiv B_R(x_0)$ for a minimiser u of $\mathcal{G}$ (or of $\mathcal{F}$) it will be convenient to reduce the proof to the case $B_R(x_0) = B_1(0)$ by introducing the rescaled functions:

$$\tilde u(x) := \frac{u(x_0 + Rx) - (u)_{B_R}}{R} \qquad \text{and} \qquad \tilde a(x) := a(x_0 + Rx)\,,$$

for $x \in B_1$. Indeed, in the case $u\colon B_R(x_0) \mapsto \mathbb{R}$ is for instance a minimiser of the functional $\mathcal{G}(\cdot, B_R(x_0))$ defined in (1.10), it is then easy to see that $\tilde{u}$ minimises the functional

$$W^{1,1}(B_1) \ni w \mapsto \int_{B_1} [f(Dw) + \tilde{a}(x)g(Dw)]\, dx$$

and that

$$\begin{cases} [\tilde{a}]_{0,\alpha;B_1} = R^{\alpha}[a]_{0,\alpha;B_R}, \quad \|\tilde{a}\|_{L^\infty(B_1)} = \|a\|_{L^\infty(B_R)} \\ \|D\tilde{u}\|_{L^p(B_1)} = \dfrac{\|Du\|_{L^p(B_R)}}{R^{n/p}}, \quad \|\tilde{u}\|_{L^p(B_1)} = \dfrac{\|u-(u)_{B_R}\|_{L^p(B_R)}}{R^{n/p+1}}\,. \end{cases} \tag{2.8}$$

3. Fractional operators and Theorem 1.6

For the proof of Theorem 1.6 we prefer to give a fully intrinsic approach that shows the global interplay between Lavrentiev phenomenon, boundedness of maximal and fractional operators, and regularity; see also next section. Indeed, we start with a proposition building on some hidden facts developed when studying the Lavrentiev phenomenon [15], [42], [51] related to functionals as in (1.1). We shall in the following report the full details.

Proposition 3.1. *Given a function $f \in L^p(B_R(x))$, where $B_R(x) \subset \mathbb{R}^n$ is a ball such that $R \leq 2$, if* (1.3) *holds, then there exists a constant c, depending only on n, p, q, such that the following inequality holds:*

$$H\left(x, (f)_{B_R(x)}\right) \leq c\left(1 + [a]_{0,\alpha}\|f\|_{L^p(B_R(x))}^{q-p}\right)\left(H(\cdot, f(\cdot))\right)_{B_R(x)}\,. \tag{3.1}$$

Proof. We set

$$a_R(x) := \inf_{y\in B_R(x)} a(y) \qquad \text{and} \qquad H_R(x,z) := |z|^p + a_R(x)|z|^q$$

so that Jensen's inequality gives

$$H_R\left(x, (f)_{B_R(x)}\right) \leq \fint_{B_R(x)} H_R\left(x, f(y)\right)\, dy \leq \fint_{B_R(x)} H(y, f(y))\, dy\,. \tag{3.2}$$

Now, we observe that

$$\begin{aligned} H\left(x, (f)_{B_R(x)}\right) &\leq |a(x)-a_R(x)||(f)_{B_R(x)}|^q + H_R\left(x,(f)_{B_R(x)}\right) \\ &\leq [a]_{0,\alpha}R^{\alpha}|(f)_{B_R(x)}|^{q-p}|(f)_{B_R(x)}|^p + H_R\left(x,(f)_{B_R(x)}\right)\,. \end{aligned}$$

Since by Hölder's inequality we have

$$|(f)_{B_R(x)}| \leq \fint_{B_R(x)} |f|\, dy \leq cR^{-n/p}\|f\|_{L^p(B_R(x))}$$

we continue to estimate

$$
\begin{aligned}
&H\left(x,(f)_{B_R(x)}\right)\\
&\qquad \le c[a]_{0,\alpha}\|f\|_{L^p(B_R(x))}^{q-p}R^{\frac{p\alpha-n(q-p)}{p}}|(f)_{B_R(x)}|^p+H_R\left(x,(f)_{B_R(x)}\right)\\
&\qquad \le c\left(1+[a]_{0,\alpha}\|f\|_{L^p(B_R(x))}^{q-p}\right)H_R\left(x,(f)_{B_R(x)}\right)
\end{aligned}
$$

where we have used (1.3) and that $R \le 2$. Last estimate and (3.2) yield (3.1).

We are now ready for Theorem 1.6.

Proof (of Theorem 1.6). *Step 1: Maximal estimate*. We define

$$
M(f)(x) := M_\Omega(f)(x) := \sup_{B_\varrho(x)\subset\Omega\,,\varrho\le 2} \fint_{B_\varrho(x)} |f(y)|\,dy\,, \tag{3.3}
$$

i.e. the restricted maximal operator, for any open and bounded domain $\Omega \subset \mathbb{R}^n$ and maps $f \in L^1(\Omega;\mathbb{R}^k)$. We notice that $M(f) = M(|f|)$. Then, for every $t \ge 1$ the maximal inequality

$$
\int_\Omega [H(x,M(f))]^t\,dx \le c\left(1+[a]_{0,\alpha}^t\|f\|_{L^p(\Omega)}^{t(q-p)}\right)\int_\Omega [H(x,f)]^t\,dx \tag{3.4}
$$

holds whenever $f \in L^p(\Omega)$, for a constant c depending only on the quantities n,p,q,α and t. We remark that (3.4) in general fails when (1.3) is not satisfied; for this see Remark 4 below. To prove (3.4), let us define the new function

$$
\bar H(x,z) := |z|^{p/\gamma}+\bar a(x)|z|^{q/\gamma}\,, \qquad \text{where } \bar a(x) := [a(x)]^{1/\gamma} \tag{3.5}
$$

and where the exponent $\gamma \equiv \gamma(n,p,q,\alpha) \in (1,p)$ is chosen in a such a way that

$$
\frac{q/\gamma}{p/\gamma} = \frac{q}{p} < 1+\frac{\alpha}{\gamma n} \tag{3.6}
$$

still holds. This is possible since (1.3) is in force. We note that the new function $\bar H(\cdot)$ is of the type of the one in (1.6), but carries now the new coefficient $[a(\cdot)]^{1/\gamma}$, which is Hölder continuous with exponent α/γ; it holds that

$$
[a^{1/\gamma}]_{0,\alpha/\gamma} \le [a]_{0,\alpha}^{1/\gamma}\,. \tag{3.7}
$$

Moreover, when passing from $H(\cdot)$ to $\bar H(\cdot)$, we pass from p and q to p/γ and q/γ, respectively. We notice that

$$
[H(x,z)]^{1/\gamma} \le \bar H(x,z) \le 2^{1-1/\gamma}[H(x,z)]^{1/\gamma}\,. \tag{3.8}
$$

By (3.6) we can now apply (3.1) to $\bar H(\cdot)$ and $|f|$ in order to get that

$$
\bar H(x,Mf(x)) \le c\left(1+[a]_{0,\alpha}^{1/\gamma}\|f\|_{L^p(\Omega)}^{(q-p)/\gamma}\right)M(\bar H(\cdot,f(\cdot)))(x) \tag{3.9}
$$

holds for every $x \in \Omega$, with $c \equiv c(n,p,q)$; we have used that the function $t \mapsto H(\cdot,t)$ is increasing. Integrating (3.9) over Ω, and applying Hardy-Littlewood maximal theorem in $L^{\gamma t}$ (recall that $\gamma > 1$ so that $t\gamma > 1$ holds too) yields

$$\begin{aligned}\int_\Omega [\bar{H}(x, M(f)(x))]^{t\gamma}\,dx &\leq c\left(1+[a]_{0,\alpha}^t \|f\|_{L^p(\Omega)}^{t(q-p)}\right)\int_\Omega [M(\bar{H}(\cdot, f(\cdot))(x)]^{t\gamma}\,dx \\ &\leq c\left(1+[a]_{0,\alpha}^t \|f\|_{L^p(\Omega)}^{t(q-p)}\right)\int_\Omega [\bar{H}(x, f(x))]^{t\gamma}\,dx\,,\end{aligned}$$

so that (3.4) follows by (3.8).

Step 2: A first Sobolev-Poincaré type inequality. Here we prove

$$\begin{aligned}&\left(\fint_{B_R}\left[H\left(x, \frac{w-(w)_{B_R}}{R}\right)\right]^{\frac{p+q(n-1)}{q(n-1)}}\,dx\right)^{\frac{q(n-1)}{p+q(n-1)}} \\ &\qquad\leq c\left(1+[a]_{0,\alpha}\|Dw\|_{L^p(B_R)}^{q-p}\right)\fint_{B_R} H(x, Dw)\,dx\,, \quad (3.10)\end{aligned}$$

with $c \equiv c(n,p,q)$ and $B_R \subset \Omega$ with $R \leq 1$; we assume that the right-hand side is finite, otherwise there is nothing to prove. We shall use a few arguments due to Hedberg [22]. The following classical formula holds for almost every $x \in B_R$:

$$|\tilde{w}(x)| := \left|\frac{w(x)-(w)_{B_R}}{R}\right| \leq \frac{c(n)}{R}\int_{B_R}\frac{|Dw(y)|}{|x-y|^{n-1}}\,dy\,. \quad (3.11)$$

See for instance [20, Lemma 7.16]. We now define $\tilde{D}\colon \mathbb{R}^n \to \mathbb{R}^n$ as

$$\tilde{D}(y) := \begin{cases} Dw(y) & \text{if } \quad y \in B_R \\ 0 & \text{if } y \in \mathbb{R}^n \setminus B_R\,. \end{cases} \quad (3.12)$$

Taking $\varepsilon \in (0,1]$ and considering the annuli $A_i(x) := B_{2^{-i}\varepsilon R}(x) \setminus B_{2^{-(i+1)}\varepsilon R}(x)$ for integers $i \geq 0$, in (3.11) we split as follows:

$$\begin{aligned}|\tilde{w}(x)| &\leq \frac{c}{R}\int_{B_{\varepsilon R}(x)}\frac{|\tilde{D}(y)|}{|x-y|^{n-1}}\,dy + \frac{c}{R}\int_{B_R\setminus B_{\varepsilon R(x)}}\frac{|\tilde{D}(y)|}{|x-y|^{n-1}}\,dy \\ &\leq \frac{c}{R}\left[\sum_{i=0}^{\infty}\left(\frac{2^i}{\varepsilon R}\right)^{n-1}\int_{A_i(x)}|\tilde{D}(y)|\,dy + \frac{1}{(\varepsilon R)^{n-1}}\int_{B_{2R}(x)}|\tilde{D}(y)|\,dy\right] \\ &\leq c\left[\sum_{i=0}^{\infty}\frac{\varepsilon}{2^i}\fint_{B_{2^{-i}\varepsilon R}(x)}|\tilde{D}(y)|\,dy + \frac{1}{\varepsilon^{n-1}}\fint_{B_{2R}(x)}|\tilde{D}(y)|\,dy\right] \\ &\leq c\varepsilon M(\tilde{D})(x) + \frac{c}{\varepsilon^{n-1}}\fint_{B_{2R}(x)}|\tilde{D}(y)|\,dy\,,\end{aligned}$$

where the maximal operator $M(\cdot) \equiv M_{B_{3R}}(\cdot)$ has been defined in (3.3), and we are taking $\Omega \equiv B_{3R}$. Since $t \mapsto H(\cdot,t)$ is an increasing function and since

$$H(x, \tilde{a}s + \tilde{b}t) \leq c[\tilde{a}^p H(x,s) + \tilde{b}^q H(x,t)] \quad (3.13)$$

holds for every $0 < \tilde{a} \le 1$, $\tilde{b} \ge 1$, $s, t > 0$, then we have

$$H\left(x, \tilde{w}(x)\right) \le c\varepsilon^p H(x, M(\tilde{D})(x)) + c\varepsilon^{q(1-n)} H(x, (|\tilde{D}|)_{B_{2R}(x)})$$

for a constant $c \equiv c(n,p,q)$. We equalise the two terms on the right hand side of the above inequality by choosing

$$\varepsilon := \left[\frac{H(x, (|\tilde{D}|)_{B_{2R}(x)})}{H(x, M(\tilde{D})(x))}\right]^{1/[p+q(n-1)]} .$$

This choice of ε is admissible; indeed, by the definition of maximal operator in (3.3) and by the fact that $t \mapsto H(\cdot, t)$ is increasing, it follows that

$$H(x, (|\tilde{D}|)_{B_{2R}(x)}) \le H(x, M(\tilde{D})(x))$$

(recall that $2R \le 2$), so that $\varepsilon \le 1$. This yields

$$H(x, \tilde{w}(x)) \le c[H(x, M(\tilde{D})(x))]^{\frac{q(n-1)}{p+q(n-1)}} [H(x, (|\tilde{D}|)_{B_{2R}(x)})]^{\frac{p}{p+q(n-1)}} .$$

By (3.1) and (3.12), and recalling that $\tilde{D} \equiv 0$ outside B_R by (3.12), we have

$$\begin{aligned}
&[H(x, \tilde{w}(x))]^{\frac{p+q(n-1)}{q(n-1)}} \\
&\le cH(x, M(\tilde{D})(x)) \left[\left(1 + [a]_{0,\alpha} \|\tilde{D}\|^{q-p}_{L^p(B_{2R}(x))}\right) \fint_{B_{2R}(x)} H(y, \tilde{D}(y))\, dy\right]^{\frac{p}{q(n-1)}} \\
&\le cH(x, M(\tilde{D})(x)) \left[\left(1 + [a]_{0,\alpha} \|Dw\|^{q-p}_{L^p(B_R)}\right) \fint_{B_R} H(y, Dw(y))\, dy\right]^{\frac{p}{q(n-1)}} .
\end{aligned}$$

Integrating on B_R and using (3.4) with $t = 1$ and $\Omega = B_{3R}$, and again (3.12), yields

$$\begin{aligned}
\fint_{B_R} [H(x, \tilde{w})]^{\frac{p+q(n-1)}{q(n-1)}}\, dx &\le c\left(1 + [a]_{0,\alpha} \|\tilde{D}\|^{q-p}_{L^p(B_{3R})}\right) \fint_{B_{3R}} H(x, \tilde{D})\, dx \\
&\quad \cdot \left[\left(1 + [a]_{0,\alpha} \|Dw\|^{q-p}_{L^p(B_R)}\right) \fint_{B_R} H(x, Dw)\, dx\right]^{\frac{p}{q(n-1)}} \\
&\le c\left(1 + [a]_{0,\alpha} \|Dw\|^{q-p}_{L^p(B_R)}\right)^{\frac{p+q(n-1)}{q(n-1)}} \left[\fint_{B_R} [H(x, Dw)]\, dx\right]^{\frac{p+q(n-1)}{q(n-1)}}
\end{aligned}$$

from which (3.10) follows.

Step 3: Improved Sobolev-Poincaré inequality and conclusion. We consider $\bar{H}(\cdot)$ defined in (3.5), for $\gamma \equiv \gamma(n,p,q,\alpha) \in (1,p)$ satisfying (3.6) and such that

$$d_1 := \frac{p + q(n-1)}{\gamma q(n-1)} > 1 . \tag{3.14}$$

We apply (3.10), but replacing $H(\cdot)$ by $\bar{H}(\cdot)$; this is possible as (3.6) is satisfied (keep also (3.7) in mind). Recalling (3.8), we have

$$\begin{aligned}
&\left(\fint_{B_R}\left[H\left(x,\frac{w-(w)_{B_R}}{R}\right)\right]^{d_1}\,dx\right)^{1/d_1}\\
&\qquad\leq c\left(\fint_{B_R}\left[\bar{H}\left(x,\frac{w-(w)_{B_R}}{R}\right)\right]^{\frac{p+q(n-1)}{q(n-1)}}\,dx\right)^{\frac{\gamma q(n-1)}{p+q(n-1)}}\\
&\qquad\leq c\left(1+[a]_{0,\alpha}\|Dw\|_{L^p(B_R)}^{q-p}\right)\left(\fint_{B_R}\bar{H}(x,Dw)\,dx\right)^{\gamma}\,.
\end{aligned}$$

Using again (3.8) to estimate the last integral in the above display, (1.18) follow with $d_1>1$ as in (3.14) and $d_2:=1/\gamma<1$.

Remark 2. Under the assumptions and notation of Theorem 1.6 the inequality

$$\begin{aligned}
&\left(\fint_{B_R}\left[H\left(x,\frac{w}{R}\right)\right]^{d_1}\,dx\right)^{1/d_1}\\
&\qquad\leq c\left(1+[a]_{0,\alpha}\|Dw\|_{L^p(B_R)}^{q-p}\right)\left(\fint_{B_R}[H(x,Dw)]^{d_2}\,dx\right)^{1/d_2}
\end{aligned}\tag{3.15}$$

holds whenever $w\in W^{1,1}_0(B_R)$, for $d_1>1>d_2$ being the same numbers appearing in (1.18). The arguments are essentially the same of Theorem 1.6 and rest on the fact that (3.11) also holds in the case $w\in W^{1,1}_0(B_R)$ (see this time [20, Lemma 7.14]) with $\tilde{w}(x):=w(x)/R$.

Remark 3. The method of proof of Theorem 1.6 allows to prove various inequalities for fractional operators that can be useful when studying energies modelled on the one in (1.1). For instance, consider the standard Riesz potential operator defined by

$$I_\beta(f)(x):=\int_{\mathbb{R}^n}\frac{f(y)}{|x-y|^{n-\beta}}\,dy\,,\qquad 0<\beta\leq n$$

and let $d:=[p\beta+q(n-\beta)]/[q(n-\beta)]>1$. It is possible to prove that whenever $B_R\subset\mathbb{R}^n$ with $R\leq 1$, and $f\in L^1(B_R)$ has compact support in B_R, the following inequality:

$$\left(\fint_{B_R}\left[H\left(x,\frac{I_\beta(f)(x)}{R^\beta}\right)\right]^{d}\,dx\right)^{1/d}\leq c\fint_{B_R}H(x,f)\,dx$$

holds, where c depends on $n,p,q,[a]_{0,\alpha},\|f\|_{L^p(B_R)}$ and β.

4. Sharpness, and Lavrentiev phenomenon

In the sequel we shall use the following:

Theorem 4.1 (Lavrentiev phenomenon [15], [17], [51]). *Let $\mathcal{P}, \mathcal{F}$ be the functionals defined in* (1.1) *and* (1.4)*, respectively, under the assumptions* (1.2) *and* (1.5)*. The following holds:*

- *If $q/p \leq 1+\alpha/n$, then for every function $w \in W^{1,1}_{\mathrm{loc}}(\Omega)$ and balls $B \Subset \tilde{B} \Subset \Omega$ such that $\mathcal{F}(w, \tilde{B}) < \infty$, there exists a sequence $\{w_m\}$ of $W^{1,\infty}$-functions such that $w_m \to w$ strongly in $W^{1,p}(B)$ and $\mathcal{F}(w_m, B) \to \mathcal{F}(w, B)$.*
- *For every $\varepsilon > 0$ and ball $B \subset \Omega$, there exist p, q satisfying $\varepsilon > q/p - 1 - \alpha/n > 0$, a coefficient $a(\cdot) \in C^{0,\alpha}$ and a boundary datum $u_0 \in W^{1,q}(B) \cap L^\infty(B)$, such that*

$$\inf_{w \in u_0 + W^{1,p}_0(B)} \mathcal{P}(w, B) < \inf_{w \in u_0 + W^{1,p}_0(B) \cap W^{1,q}_{\mathrm{loc}}(B)} \mathcal{P}(w, B)\,,$$

  where $\mathcal{P}$ has been defined in (1.1)*. That is, the Lavrentiev phenomenon occurs between $W^{1,p}(B)$ and $W^{1,q}_{\mathrm{loc}}(B)$. In particular, local minimisers are in general not in $W^{1,q}_{\mathrm{loc}}(B)$. Moreover, they can be discontinuous.*

Now, while we shall use the first part of Theorem 4.1 to prove Theorems 1.2-1.3 via Theorem 5.1 below, the second part can be used to prove their sharpness.

Remark 4. Inequality (3.4) fails in general when (1.3) is not satisfied. This can be seen by showing that the Lavrentiev phenomenon is absent provided (3.4) holds. This also gives a proof of the first part of Theorem 4.1. For this, take $w \in W^{1,p}(\Omega)$ such that $\mathcal{F}(w, \tilde{B})$ is finite; this in particular means that $H(x, Dw) \in L^1(\tilde{B})$ by (1.5). Then, for $\varepsilon \in (0,1)$ such that $(1+\varepsilon)B \subset \tilde{B}$, we consider the mollified sequence $\{w_\varepsilon := w * \rho_\varepsilon\}$, where $\{\rho_\varepsilon\}$ is a family of standard mollifiers; we also consider the maximal operator $M(\cdot) \equiv M_{\tilde{B}}(\cdot)$ as defined in (3.3). It follows that $|Dw_\varepsilon(x)| \leq cM(Dw)(x)$, for every $x \in B$ and for c independent of ε, so that

$$H(x, Dw_\varepsilon(x)) \leq cH(x, M(Dw)(x)) \tag{4.1}$$

also holds for such ε and x. By (3.4) we know that $H(x, M(Dw)) \in L^1(\tilde{B})$ too and therefore we conclude, by (4.1) and Lebesgue's dominated convergence theorem, that

$$H(x, Dw_\varepsilon(x)) \to H(x, Dw(x)) \qquad \text{in } L^1(B) \tag{4.2}$$

(with abuse of notation, with the subscript ε we here denote a not relabelled subsequence). In turn, again by (1.5), (4.2) and a variant of Lebesgue's dominated convergence theorem, we have that $F(x, w_\varepsilon, Dw_\varepsilon) \to F(x, w, Dw)$ in $L^1(B)$, that is $\mathcal{F}(w_\varepsilon, B) \to \mathcal{F}(w, B)$.

5. A reverse Hölder inequality in the p-phase

Here we prove a theorem asserting that reverse Hölder inequalities persist when the functional is in the p-phase. The statement will be also useful for future developments, in order to obtain further regularity results for the class of operators considered.

Theorem 5.1. *Let $u \in W^{1,p}(\Omega)$ be a local minimiser of the functional $\mathcal{G}$ defined in* (1.10) *under the assumptions* (1.2), (1.3) *and* (1.11). *Let $B_R \subset \Omega$ be a ball with $R \le 1$ and such that the inequality*

$$\sup_{x\in B_R} a(x) \le M[a]_{0,\alpha}R^{\alpha} \tag{5.1}$$

is satisfied for some number $M \ge 1$. Then for every $\tilde q < np/(n-2\alpha)$ there exists a positive constant c, depending only on $\mathtt{data}, M$ and $\tilde q$, such that the following reverse Hölder inequality holds:

$$\left(\fint_{B_{R/2}} |Du|^{\tilde q}\,dx\right)^{1/\tilde q} \le c\left(\fint_{B_R} |Du|^{p}\,dx\right)^{1/p}. \tag{5.2}$$

Moreover, there exist an exponent $\tilde b \equiv \tilde b(n,p,q,\alpha) \ge 1$ and a constant $c \equiv c(\mathtt{data}, \tilde q, \operatorname{diam}(\Omega))$, such that the following inequality holds for every ball $B_R \subset \Omega$ (no matter (5.1) *is satisfied or not):*

$$\left(\fint_{B_{R/2}} |Du|^{\tilde q}\,dx\right)^{1/\tilde q} \le c\left(\fint_{B_R} (|Du|^{p}+1)\,dx\right)^{\tilde b/p}. \tag{5.3}$$

In particular, both (5.2) *and* (5.3) *hold with $\tilde q \equiv 2q-p$ and $u \in W^{1,2q-p}_{\mathrm{loc}}(\Omega)$.*

Proof. The proof is rather delicate since we have to keep track of the role of the coefficient $a(\cdot)$ in order to use condition (5.1) in an efficient way and get the reverse inequality (5.2). We will derive higher integrability estimates of the gradient via fractional Sobolev embedding theorem, after proving that $V_p(Du) = |Du|^{(p-2)/2}Du$ belongs to a suitable Nikolski space; this will also involve a careful convolution argument to treat certain non-differentiable terms. To get the final estimate we first use an approximation procedure based on the absence of Lavrentiev phenomenon and then an interpolation argument; both work if and only if (1.3) holds. To shorten formulas, we shall use the symbol $\lesssim$ to write an inequality where the implied constant c only depends only on n, p, q, L, ν. The proof falls in seven steps. For the proof of (5.3) see Step 1 while the proof of (5.2) is in Step 7; that is the only place where (5.1) will be used. We just remark that for (5.2) it is not really necessary to have that $R \le 1$; when such a bound is not assumed the inequality still holds but this time the constant c depends also on the diameter of Ω. Moreover, from the proof below we note that when $p = q$ we have $\tilde b = 1$ and (5.1) is not needed to obtain a reverse type inequality.

Step 1: Scaling and proof of (5.3). Given a ball $B_R \equiv B_R(x_0)$ as in the statement of Theorem 5.1, we begin by proving that for every $\beta \in (n(q/p-1), \alpha)$ there exists a positive constant $c \equiv c(n, p, q, \nu, L, \beta)$ such that the inequality

$$\left(\fint_{B_{R/2}(x_0)} |Du|^{\frac{np}{n-2\beta}} \, dx \right)^{\frac{n-2\beta}{n}} \leq c \fint_{B_R(x_0)} |Du|^p \, dx + \left(\|a\|^2_{L^\infty(B_R)} + R^{2\alpha}[a]^2_{0,\alpha;B_R} \right)^{b_1} \left(\fint_{B_R(x_0)} |Du|^p \, dx \right)^{b_2} \quad (5.4)$$

holds for (large) exponents

$$b_1 = \frac{\beta p}{\beta p - n(q-p)} \geq 1, \qquad b_2 = \frac{\beta(2q-p) - n(q-p)}{\beta p - n(q-p)} \geq 1 \,. \quad (5.5)$$

Notice that the range $(n(q/p-1), \alpha)$ of admissible values for β is non-empty since (1.3) is in force. We observe that (5.3) directly follows from (5.4), by finding $\beta \in (n(q/p-1), \alpha)$ - close enough to α - in order to guarantee that $\tilde{q} \leq np/(n-2\beta)$. On there other hand observe that

$$\beta > n\left(\frac{q}{p} - 1\right) \Longleftrightarrow q < p\left(1 + \frac{\beta}{n}\right) \Longrightarrow q < \frac{p(n-\beta)}{n-2\beta} \quad (5.6)$$

and

$$q < \frac{p(n-\beta)}{n-2\beta} \Longleftrightarrow 2q - p < \frac{np}{n-2\beta} < \frac{np}{n-2\alpha}\,, \quad (5.7)$$

so that we can take $\tilde{q} \equiv 2q - p$ in (5.2)-(5.3). Now, in order to prove (5.4), we first use the scaling argument of Remark 1, that is, we rescale the minimiser in B_1 and prove the corresponding inequality

$$\|D\tilde{u}\|^p_{L^{\frac{np}{n-2\beta}}(B_{1/2})} \leq c \left[\|D\tilde{u}\|^p_{L^p(B_1)} + (\|\tilde{a}\|^2_{L^\infty(B_1)} + [\tilde{a}]^2_{0,\alpha;B_1})^{b_1} \|D\tilde{u}\|^{pb_2}_{L^p(B_1)} \right], \quad (5.8)$$

from which (5.4) follows scaling back to the original minimiser u. We observe that in proving (5.4) we can without loss of generality assume that $B_R(x_0) \Subset \Omega$ since all the constants involved in (5.4) are independent of $B_R(x_0)$. At this point we can consider another ball $B_{R'}(x_0) \Subset \Omega$ such that $R < R'$ so that scaling as in Remark 1 actually gives that the functions $\tilde{u}$ and $\tilde{a}$ can be considered in $B_{R'/R}(0)$ and

$$|D\tilde{u}|^p + \tilde{a}(x)|D\tilde{u}|^q \in L^1(B_{R'/R}(0))\,. \quad (5.9)$$

Summarising, Steps 2-6 are now devoted to the proof of (5.8) and therefore of (5.4).

Step 2: Approximation via absence of Lavrentiev phenomenon. By (5.9) we can apply the first part of Theorem 4.1 with $B \equiv B_1(0)$ and $\tilde{B} \equiv B_{R'/R}(0)$. Therefore we find a sequence $\{u_m\} \subset W^{1,2q-p}(B_1)$ such that

$$u_m \to \tilde{u} \text{ strongly in } W^{1,p}(B_1) \quad \text{and} \quad \mathcal{F}(u_m, B_1) \to \mathcal{F}(\tilde{u}, B_1)\,. \quad (5.10)$$

To proceed, we define, for any $m \in \mathbb{N}$, $x \in B_1$ and $z \in \mathbb{R}^n$, the numbers σ_m and the integrands $F, F_m\colon B_1 \times \mathbb{R}^n \to \mathbb{R}$ as follows:

$$\begin{cases} \sigma_m := \left(1 + m + \|Du_m\|_{L^{2q-p}(B_1)}^{2(2q-p)}\right)^{-1} \\ F(x,z) := f(z) + \tilde{a}(x)g(z)\,, \quad F_m(x,z) := F(x,z) + \sigma_m |z|^{2q-p}\ . \end{cases} \tag{5.11}$$

Direct Methods of the Calculus of Variations allow us to define the function $v_m \in u_m + W_0^{1,2q-p}(B_1)$ as the unique solution to the following Dirichlet problem:

$$\begin{cases} v_m \mapsto \min\limits_{w} \displaystyle\int_{B_1} F_m(x, Dw)\, dx \\ w \in u_m + W_0^{1,2q-p}(B_1)\ . \end{cases}$$

In the next step we shall derive uniform estimates for the sequence $\{V_p(Dv_m)\}$ in Nikolski spaces. We notice that functional appearing in the above display has standard polynomial growth of order $2q-p$, therefore by standard regularity theory (see for instance [18], [25]) we know that

$$v_m \in W^{1,\infty}_{\mathrm{loc}}(B_1)\ . \tag{5.12}$$

To simplify the notation we shall denote $v_m \equiv v, \tilde{a} \equiv a$, eventually restoring the full notation in Step 4.

Step 3: Fractional Caccioppoli inequality and integrability. Here we prove that for every $\beta \in (0, \alpha)$ there exists a positive constant $c \equiv c(n,p,q,\nu,L,\beta)$ such that

$$\begin{aligned} \|Dv\|_{L^{np/(n-2\beta)}(B_{1/2})}^p &\leq c\|Dv\|_{L^p(B_1)}^p \\ &\quad + c(\|a\|_{L^\infty(B_1)}^2 + [a]_{0,\alpha;B_1}^2 + \sigma_m)\|Dv\|_{L^{2q-p}(B_1)}^{2q-p} \end{aligned} \tag{5.13}$$

holds. For this we shall use the Euler-Lagrange equation:

$$\int_{B_1} \langle \partial F_m(x, Dv), D\varphi \rangle\, dx = 0 \qquad \forall\, \varphi \in W_0^{1,2q-p}(B_1)\ . \tag{5.14}$$

Let $\eta \in C_0^\infty(B_{3/4})$ be a cut-off function such that $0 \leq \eta \leq 1$, $\eta \equiv 1$ in $B_{2/3}$, and $|D\eta|^2 + |D^2\eta| \leq 10^4$. In the following we take $h \in \mathbb{R}^n$ such that $0 < |h| \leq 10^{-4}$ holds and denote

$$S(x,h) := |Dv(x+h)| + |Dv(x)|\,. \tag{5.15}$$

This function will be in fact considered only for $x \in B_{3/4}$. We take the function $\varphi = \tau_{-h}(\eta^2 \tau_h v)$ in (5.14); we recall that the operator τ_h has been defined in (2.1). Integration by parts for finite differences yields

$$\int_{B_1} \langle \tau_h(\partial F_m(\cdot, Dv)), D(\eta^2 \tau_h v) \rangle\, dx = 0\ . \tag{5.16}$$

Using the decomposition

$$\begin{aligned} \tau_h(\partial F_m(\cdot, Dv))(x) &= [\partial F_m(x+h, Dv(x+h)) - \partial F_m(x+h, Dv(x))] \\ &\quad + [\partial F_m(x+h, Dv(x)) - \partial F_m(x, Dv(x))] \\ &=: A_1 + A_2\ , \end{aligned} \tag{5.17}$$

we rewrite (5.16) as

$$\begin{aligned} I_0 := \int_{B_1} \eta^2 \langle A_1, \tau_h Dv\rangle \, dx &= -\int_{B_1} \eta^2 \langle A_2, \tau_h Dv\rangle \, dx \\ &\quad -2\int_{B_1} \eta\langle A_2, D\eta\rangle \tau_h v \, dx - 2\int_{B_1} \eta\langle A_1, D\eta\rangle \tau_h v \, dx \\ &=: I_1 + I_2 + I_3 \, . \end{aligned} \tag{5.18}$$

We estimate each term in (5.18). For I_0 we apply (2.5) with $z_1 = Dv(x+h)$, $z_2 = Dv(x)$, $\gamma = p, q, 2q-p$ and discarding the q-term we obtain

$$\int_{B_1} \eta^2 |\tau_h[V_p(Dv)]|^2 \, dx + \sigma_m \int_{B_1} \eta^2 |\tau_h[V_{2q-p}(Dv)]|^2 \, dx \lesssim I_0 \, . \tag{5.19}$$

To estimate I_1 in (5.18) we use the Hölder continuity of $a(\cdot)$ and (2.3). This yields

$$\begin{aligned} |I_1| &\lesssim [a]_{0,\alpha} |h|^\alpha \int_{B_1} \eta^2 |Dv|^{q-1} |\tau_h Dv| \, dx \\ &\leq [a]_{0,\alpha} |h|^\alpha \int_{B_1} \eta^2 [S(x,h)]^{q-1} |\tau_h Dv| \, dx \, , \end{aligned}$$

with $S(x,h)$ that has been defined in (5.15); note that here we take into account that $\eta \equiv 0$ outside $B_{3/4}$ and that $|h| \leq 10^{-4}$ to use $S(x,h)$. In turn, by using Young's inequality, for $\varepsilon \in (0,1)$ we estimate

$$\begin{aligned} |I_1| &\lesssim \varepsilon \int_{B_1} \eta^2 [S(x,h)]^{p-2} |\tau_h Dv|^2 \, dx + \frac{|h|^{2\alpha}[a]_{0,\alpha;B_1}^2}{\varepsilon} \int_{B_{3/4}} [S(x,h)]^{2q-p} \, dx \\ &\lesssim \varepsilon \int_{B_1} \eta^2 |\tau_h[V_p(Dv)]|^2 \, dx + \frac{|h|^{2\alpha}[a]_{0,\alpha;B_1}^2}{\varepsilon} \int_{B_1} |Dv|^{2q-p} \, dx \, . \end{aligned}$$

We have also used (2.6) and that $|h| \leq 10^{-4}$. Recalling (5.19), we conclude with

$$|I_1| \lesssim \varepsilon I_0 + \frac{|h|^{2\alpha}[a]_{0,\alpha;B_1}^2}{\varepsilon} \int_{B_1} |Dv|^{2q-p} \, dx \, . \tag{5.20}$$

I_2 can be estimated via (2.3) and then by Young's and Hölder's inequalities

$$\begin{aligned} |I_2| &\lesssim [a]_{0,\alpha;B_1} |h|^\alpha \int_{B_{3/4}} |Dv|^{q-1} |\tau_h v| |D\eta| \, dx \\ &\lesssim |h|^\alpha \left(\int_{B_{3/4}} |Dv|^{p-1} |\tau_h v| \, dx + [a]_{0,\alpha;B_1}^2 \int_{B_{3/4}} |Dv|^{2q-p-1} |\tau_h v| \, dx \right) \\ &\lesssim |h|^\alpha \left(\int_{B_1} |Dv|^p \, dx \right)^{\frac{p-1}{p}} \left(\int_{B_{3/4}} |\tau_h v|^p \, dx \right)^{\frac{1}{p}} \\ &\quad + [a]_{0,\alpha;B_1}^2 |h|^\alpha \left(\int_{B_1} |Dv|^{2q-p} \, dx \right)^{\frac{2q-p-1}{2q-p}} \left(\int_{B_{3/4}} |\tau_h v|^{2q-p} \, dx \right)^{\frac{1}{2q-p}} \\ &\lesssim |h|^{2\alpha} \left(\int_{B_1} |Dv|^p \, dx + [a]_{0,\alpha;B_1}^2 \int_{B_1} |Dv|^{2q-p} \, dx \right) \, . \end{aligned} \tag{5.21}$$

In the last line we have used (2.2) and that $|h|^{1+\alpha} \leq |h|^{2\alpha}$ since $|h| \leq 10^{-4}$. We are left to estimate I_3 in (5.18); this requires much more effort and we are forced to treat differently the case $p \geq 2$ from the case $p < 2$. If $p \geq 2$ we have

$$|I_3| \lesssim \|D\eta\|_{L^\infty(B_1)}[I_{3,p} + \|a\|_{L^\infty(B_1)} I_{3,q} + \sigma_m I_{3,2q-p}] \,, \tag{5.22}$$

where, for $\gamma \in \{p, q, 2q-p\}$, we denote

$$I_{3,\gamma} := \int_{B_1} \eta [S(x,h)]^{\gamma-2} |\tau_h Dv||\tau_h v| \, dx \,. \tag{5.23}$$

Notice that in making the estimate in (5.22) we have used (2.7) with $\gamma = 2q-p$. We confine ourselves to show the estimates for $I_{3,q}$, the treatment of the other terms being completely similar. By Young's inequality and (2.6) we have, for $\varepsilon \in (0,1)$

$$\begin{aligned} \|a\|_{L^\infty(B_1)} I_{3,q} &\lesssim \varepsilon \int_{B_{3/4}} \eta^2 [S(x,h)]^{p-2} |\tau_h Dv|^2 \, dx \\ &\qquad + \frac{\|a\|^2_{L^\infty(B_1)}}{\varepsilon} \int_{B_{3/4}} [S(x,h)]^{2q-p-2} |\tau_h v|^2 \, dx \\ &\lesssim \varepsilon \int_{B_1} \eta^2 |\tau_h [V_p(Dv)]|^2 \, dx \\ &\qquad + \frac{\|a\|^2_{L^\infty(B_1)}}{\varepsilon} \int_{B_{3/4}} [S(x,h)]^{2q-p-2} |\tau_h v|^2 \, dx \,. \end{aligned}$$

In turn, by Hölder's inequality and (2.2), and recalling that $|h| \leq 10^{-4}$, we have

$$\begin{aligned} \int_{B_{3/4}} [S(x,h)]^{2q-p-2} |\tau_h v|^2 \, dx &\leq \|S(x,h)\|^{2q-p-2}_{L^{2q-p}(B_{3/4})} \|\tau_h v\|^2_{L^{2q-p}(B_{3/4})} \\ &\lesssim |h|^2 \|Dv\|^{2q-p}_{L^{2q-p}(B_1)} \,. \end{aligned}$$

By merging the content of the last two displays and using (5.19) we conclude with

$$\|a\|_{L^\infty(B_1)} I_{3,q} \lesssim \varepsilon I_0 + \frac{|h|^2 \|a\|^2_{L^\infty(B_1)}}{\varepsilon} \int_{B_1} |Dv|^{2q-p} \, dx \,. \tag{5.24}$$

The terms $I_{3,p}$ and $I_{3,2q-p}$ can be estimated in exactly the same way:

$$I_{3,p} + \sigma_m I_{3,2q-p} \lesssim \varepsilon I_0 + \frac{|h|^2}{\varepsilon} \int_{B_1} |Dv|^p \, dx + \frac{|h|^2 \sigma_m}{\varepsilon} \int_{B_1} |Dv|^{2q-p} \, dx \,. \tag{5.25}$$

Connecting the last display with (5.24) and (5.22), and eventually to (5.18), (5.20) and (5.21) yields, for every $\varepsilon \in (0,1)$

$$\begin{aligned} I_0 &\lesssim \varepsilon I_0 + \frac{|h|^{2\alpha}}{\varepsilon} \int_{B_1} |Dv|^p \, dx \\ &\qquad + \frac{|h|^{2\alpha}}{\varepsilon} (\|a\|^2_{L^\infty(B_1)} + [a]^2_{0,\alpha;B_1} + \sigma_m) \int_{B_1} |Dv|^{2q-p} \, dx \,, \end{aligned} \tag{5.26}$$

where the implied constant depends as usual only on n, p, q, ν, L. Therefore, choosing $\varepsilon \equiv \varepsilon(n, p, q, \nu, L)$ small enough and recalling (5.19), we conclude with the following *fractional Caccioppoli type inequality*:

$$\int_{B_1} \eta^2 |\tau_h[V_p(Dv)]|^2 \, dx \lesssim |h|^{2\alpha} \int_{B_1} |Dv|^p \, dx + |h|^{2\alpha} (\|a\|^2_{L^\infty(B_1)} + [a]^2_{0,\alpha;B_1} + \sigma_m) \int_{B_1} |Dv|^{2q-p} \, dx \,. \quad (5.27)$$

Remark 5. In (5.19)-(5.21) an ambiguity appears when for instance considering terms of the type $[S(x,h)]^{p-2}$ and $p < 2$, since it could be $S(x,h) = 0$. On the other hand notice that these terms always multiply $|\tau_h Dv|^2$ so that in the case $S(x,h) = 0$, there is no problem in considering $[S(x,h)]^{p-2}|\tau_h Dv|^2 = 0$. Alternatively, when $S(x,h) = 0$ all the integrands considered in (5.17) vanish and therefore we can always restrict the domain of integration of I_0, I_1, I_2 to $B_1 \cap \{S(x,h) > 0\}$.

We now proceed in proving (5.27) for the case $p < 2$. To this purpose we need to estimate I_3 in a different way and begin by splitting as follows:

$$I_3 = -2J_{3,p} - 2J_{3,q} - 2J_{3,2q-p} \,, \quad (5.28)$$

where

$$\begin{cases} J_{3,p} := \displaystyle\int_{B_1} \eta \langle \tau_h[\partial f(Dv)], D\eta \rangle \tau_h v \, dx \\ J_{3,q} := \displaystyle\int_{B_1} \eta(x) \langle \tau_h[\partial g(Dv)](x), D\eta(x) \rangle a(x+h)(\tau_h v)(x) \, dx \\ J_{3,2q-p} := \sigma_m \displaystyle\int_{B_1} \eta \langle \tau_h[|Dv|^{2q-p-2} Dv], D\eta \rangle \tau_h v \, dx \,. \end{cases}$$

As for the case $p \geq 2$, we shall give all the details for $J_{3,q}$, with short remarks on the modifications needed for $J_{3,p}$ and $J_{3,2q-p}$. First, let us notice a general fact; if $G \in L^{\gamma/(\gamma-1)}(B_1; \mathbb{R}^n)$ and $\mathcal{H} \in W^{1,\gamma}_0(B_{3/4}; \mathbb{R}^n)$ for some $\gamma > 1$, we have that for every $h \in \mathbb{R}^n$ with $0 < |h| \leq 10^{-4}$ it holds

$$\int_{B_1} \langle \tau_h G, \mathcal{H} \rangle \, dx = -|h| \int_{B_1} \int_0^1 \langle G(x+th), \partial_{h/|h|} \mathcal{H}(x) \rangle \, dt \, dx \,. \quad (5.29)$$

Indeed, approximating both G and $\mathcal{H}$ with smooth functions, we reduce to prove that (5.29) holds when $G \in C^\infty(B_1; \mathbb{R}^n)$ and $\mathcal{H} \in C^\infty_c(B_{3/4}; \mathbb{R}^n)$; in this case it is

$$G(x+h) - G(x) = \int_0^1 \frac{d}{dt} \big[G(x+th) \big] \, dt = |h| \int_0^1 \partial_{h/|h|} G(x+th) \, dt \,.$$

Plugging this formula on the left-hand side of (5.29) and integrating by parts we obtain the right-hand side of (5.29). We would like to apply (5.29) with the choice $G \equiv \partial g(Dv) \in L^{q/(q-1)}(B_1; \mathbb{R}^n)$ and $\mathcal{H} \equiv a \tau_h v \eta D\eta$, but this is not possible, since the function $a(\cdot)$ is not differentiable. We overcome this point by replacing $a(\cdot)$ with new functions $a_{|h|}(\cdot)$, which are obtained by $a(\cdot)$ via convolution with a mollifier related to the same scale of discrete differentiation $|h|$; we then estimate the corresponding correction term. Specifically, we consider a standard non-negative symmetric mollifier $\rho \in C_c^\infty(B_1)$, such that $\|\rho\|_{L^1(\mathbb{R}^n)} = 1$, and let $\rho_{|h|}(x) = |h|^{-n}\rho(x/|h|)$ for $|h| \leq 10^{-4}$. Finally, we define $a_{|h|} := a * \rho_{|h|} \in C^\infty(B_{15/16})$ (for reasons that are going to be clear in a few lines it is here sufficient to consider $a_{|h|}(\cdot)$ defined only on $B_{15/16}$, essentially because it will always appear multiplied by η, which is zero outside $B_{3/4}$). The main point here is that the blow-up rate of $\|Da_{|h|}\|_{L^\infty}$ can be quantified when $|h| \to 0$ and indeed we have that

$$\begin{cases} |a_{|h|}(x) - a(x)| \leq [a]_{0,\alpha,B_1}|h|^\alpha \\ \\ |Da_{|h|}(x)| \leq c[a]_{0,\alpha,B_1}|h|^{\alpha-1} \end{cases} \tag{5.30}$$

hold for every $x \in B_{15/16}$. Indeed, the first inequality in (5.30) is a direct consequence of the definition of convolution, while, as for the second, we have

$$\begin{aligned} |Da_{|h|}(x)| &= \left| -\frac{1}{|h|^{n+1}} \int_{B_{|h|}(x)} a(y)(D\rho)\left(\frac{y-x}{|h|}\right) dy \right| \\ &= \left| \frac{1}{|h|} \int_{B_1} a(x+|h|y)(D\rho)(y)\, dy \right| \\ &= \left| \frac{1}{|h|} \int_{B_1} [a(x+|h|y) - a(x)](D\rho)(y)\, dy \right| \\ &\leq \frac{|B_1|\|D\rho\|_{L^\infty(B_1)}}{|h|}[a]_{0,\alpha;B_1}|h|^\alpha \\ &\leq c(n)[a]_{0,\alpha;B_1}|h|^{\alpha-1}\,. \end{aligned}$$

To proceed with the estimates, we further decompose as

$$\begin{aligned} |J_{3,q}| \;\leq\; & \int_{B_1} \eta|\tau_h[\partial g(Dv)](x)||D\eta(x)||a(x+h) - a_{|h|}(x)||\tau_h v(x)|\, dx \\ & + \left| \int_{B_1} \eta\langle \tau_h[\partial g(Dv)](x), D\eta(x)\rangle a_{|h|}(x)(\tau_h v)(x)\, dx \right| \\ =:\; & J_{3,q,1} + J_{3,q,2}\,. \end{aligned} \tag{5.31}$$

Observing that by $(5.30)_1$ we have

$$|a(x+h) - a_{|h|}(x)| \leq |a(x+h) - a(x)| + |a(x) - a_{|h|}(x)| \leq 2[a]_{0,\alpha;B_1}|h|^\alpha,$$

and applying in sequence (2.3), Young's inequality and finally (2.2), yields

$$\begin{aligned} J_{3,q,1} &\lesssim [a]_{0,\alpha;B_1}|h|^\alpha \int_{B_{3/4}} |Dv|^{q-1}|\tau_h v||D\eta|\, dx \\ &\leq [a]_{0,\alpha;B_1}|h|^\alpha \int_{B_{3/4}} |S(x,h)|^{q-1}|\tau_h v||D\eta|\, dx \end{aligned}$$

$$
\begin{aligned}
&\lesssim |h|^{\alpha}\left(\int_{B_{3/4}} |S(x,h)|^{p-1}|\tau_h v|\,dx\right.\\
&\qquad\qquad \left.+[a]_{0,\alpha;B_1}^2 \int_{B_{3/4}} |S(x,h)|^{2q-p-1}|\tau_h v|\,dx\right) \qquad (5.32)\\
&\lesssim |h|^{\alpha}\left(\int_{B_{3/4}} |S(x,h)|^{p}\,dx\right)^{\frac{p-1}{p}}\left(\int_{B_{3/4}} |\tau_h v|^{p}\,dx\right)^{\frac{1}{p}}\\
&\qquad +[a]_{0,\alpha;B_1}^2 |h|^{\alpha}\left(\int_{B_{3/4}} |S(x,h)|^{2q-p}\,dx\right)^{\frac{2q-p-1}{2q-p}}\\
&\qquad\qquad \cdot\left(\int_{B_{3/4}} |\tau_h v|^{2q-p}\,dx\right)^{\frac{1}{2q-p}}\\
&\lesssim |h|^{2\alpha}\left(\int_{B_1} |Dv|^{p}\,dx + [a]_{0,\alpha;B_1}^2\int_{B_1} |Dv|^{2q-p}\,dx\right).
\end{aligned}
$$

Applying (5.29) with $G \equiv \partial g(Dv) \in L^{q/(q-1)}(B_1;\mathbb{R}^n)$ and $\mathcal{H} \equiv \eta a_{|h|}\tau_h v D\eta \in W^{1,q}_0(B_{3/4},\mathbb{R}^n)$ we estimate the second term in (5.31) as follows:

$$
\begin{aligned}
J_{3,q,2} &= |h|\left|\int_0^1\int_{B_1} \langle \partial g(Dv(x+th)), \partial_{h/|h|}[\eta a_{|h|}\tau_h v D\eta]\rangle\,dx\,dt\right|\\
&\lesssim |h|\int_0^1\int_{B_1} |Dv(x+th)|^{q-1}\big[(|D\eta|^2+|D^2\eta|)|\tau_h v|a_{|h|}\\
&\qquad\qquad +\eta|D\eta||\tau_h Dv|a_{|h|} + \eta|D\eta||\tau_h v||Da_{|h|}|\big]\,dx\,dt\\
&\lesssim \big(|h|\|a\|_{L^\infty(B_1)} + |h|^{\alpha}[a]_{0,\alpha;B_1}\big)\int_0^1\int_{B_{3/4}} |Dv(x+th)|^{q-1}|\tau_h v|\,dx\,dt\\
&\qquad\qquad +|h|\|a\|_{L^\infty(B_1)}\int_0^1\int_{B_1} \eta|Dv(x+th)|^{q-1}|\tau_h Dv|\,dx\,dt\\
&=: K_1 + K_2\,.
\end{aligned}
$$

Noticed that we have used the second inequality in (2.3) and (5.30). In turn, by using Young's and Hölder's inequalities, and yet (2.2), we estimate K_1 as follows:

$$
\begin{aligned}
K_1 &\lesssim |h|^{\alpha}\left[\int_0^1\int_{B_{3/4}} |Dv(x+th)|^{p-1}|\tau_h v|\,dx\,dt\right.\\
&\qquad \left.+(\|a\|^2_{L^\infty(B_1)} + [a]^2_{0,\alpha;B_1})\int_0^1\int_{B_{3/4}} |Dv(x+th)|^{2q-p-1}|\tau_h v|\,dx\,dt\right]\\
&\lesssim |h|^{2\alpha}\left[\int_{B_1} |Dv|^p\,dx + (\|a\|^2_{L^\infty(B_1)} + [a]^2_{0,\alpha;B_1})\int_{B_1} |Dv|^{2q-p}\,dx\right].
\end{aligned}
$$

Notice that the two integrals in the square bracket above have been estimated as those appearing in the estimate for $J_{3,q,1}$, and more precisely as those in (5.32).

As for K_2, with $\varepsilon, \kappa \in (0,1)$, we use Young's inequality to get

$$
\begin{aligned}
K_2 &\lesssim \varepsilon \int_{B_1} \eta^2 [\kappa + S(x,h)]^{p-2} |\tau_h Dv|^2 \, dx \\
&\quad + \frac{|h|^2 \|a\|^2_{L^\infty(B_1)}}{\varepsilon} \int_0^1 \int_{B_{3/4}} [\kappa + S(x,h)]^{2-p} |Dv(x+th)|^{2(q-1)} \, dx \, dt \, .
\end{aligned}
$$

Again using Young's inequality, this time with conjugate exponents $(2q-p)/(2-p)$ and $(2q-p)/(2q-2)$, letting $\kappa \to 0$ and finally using (2.6) with $\gamma = p$, yields

$$
\begin{aligned}
K_2 &\lesssim \varepsilon \int_{B_1} \eta^2 |\tau_h [V_p(Dv)]|^2 \, dx + \frac{|h|^2 \|a\|^2_{L^\infty(B_1)}}{\varepsilon} \int_{B_{3/4}} [S(x,h)]^{2q-p} \, dx \\
&\qquad + \frac{|h|^2 \|a\|^2_{L^\infty(B_1)}}{\varepsilon} \int_0^1 \int_{B_{3/4}} |Dv(x+th)|^{2q-p} \, dx \, dt \\
&\lesssim \varepsilon \int_{B_1} \eta^2 |\tau_h [V_p(Dv)]|^2 \, dx + \frac{|h|^2 \|a\|^2_{L^\infty(B_1)}}{\varepsilon} \int_{B_1} |Dv|^{2q-p} \, dx \, .
\end{aligned}
$$

Collecting the estimates found for $K_1, K_2, J_{3,q,1}, J_{3,q,2}$ to (5.31) and recalling (5.19), we finally conclude with

$$
\begin{aligned}
|J_{3,q}| &\lesssim \varepsilon I_0 + \frac{|h|^{2\alpha}}{\varepsilon} \int_{B_1} |Dv|^p \, dx \\
&\quad + \frac{|h|^{2\alpha} (\|a\|^2_{L^\infty(B_1)} + [a]^2_{0,\alpha;B_1})}{\varepsilon} \int_{B_1} |Dv|^{2q-p} \, dx \, .
\end{aligned} \tag{5.33}
$$

Estimating $J_{3,p}$ is now very much similar, and actually simpler since the function $a(\cdot)$ is not involved. Indeed, applying identity (5.29) with the choice $G \equiv \partial f(Dv) \in L^{p/(p-1)}(B_1;\mathbb{R}^n)$ and $\mathcal{H} \equiv \eta \tau_h v D\eta \in W^{1,p}_0(B_{3/4}, \mathbb{R}^n)$, we deduce that

$$
|J_{3,p}| \leq |h| \left| \int_0^1 \int_{B_1} \langle \partial f(Dv(x+th)), \partial_{h/|h|} [\eta \tau_h v D\eta] \rangle \, dx \, dt \right| \, .
$$

This term can be now estimated exactly as $J_{3,q,2}$, but taking $a_{|h|} \equiv 1$ and $q = p$. The final estimate is

$$
|J_{3,p}| \lesssim \varepsilon I_0 + \frac{|h|^2}{\varepsilon} \int_{B_1} |Dv|^p \, dx \, . \tag{5.34}
$$

We finally come to $J_{3,2q-p}$, and again distinguish between two cases. The first is when $2q-p<2$; in this case we proceed as for $J_{3,p}$, thereby getting

$$
|J_{3,2q-p}| \lesssim \varepsilon I_0 + \frac{|h|^2 \sigma_m}{\varepsilon} \int_{B_1} |Dv|^{2q-p} \, dx \, . \tag{5.35}
$$

In the remaining case $2q-p \geq 2$, by using (2.7) we can estimate

$$
|J_{3,2q-p}| \lesssim \|D\eta\|_{L^\infty(B_1)} \sigma_m I_{3,2q-p}
$$

(see (5.23) for the definition of $I_{3,2q-p}$). Using the corresponding estimate contained in (5.25), we again arrive at (5.35). Using (5.33)-(5.35) in (5.28) finally yields

$$|I_3| \lesssim \varepsilon I_0 + \frac{|h|^{2\alpha}}{\varepsilon}\int_{B_1} |Dv|^p \, dx + \frac{|h|^{2\alpha}(\|a\|^2_{L^\infty(B_1)} + [a]^2_{0,\alpha;B_1} + \sigma_m)}{\varepsilon}\int_{B_1} |Dv|^{2q-p} \, dx \,,$$

whenever $\varepsilon \in (0,1)$. Using this last inequality together with (5.18), (5.20) and (5.21) (which are valid whenever $p > 1$), we again arrive at (5.26) and eventually at (5.27), which therefore holds in the full range of considered exponents $1 < p < q$. We now notice that (5.27) implies that $V_p(Dv)$ belongs to the Nikolski space $\mathcal{N}^{\alpha,2}(B_{2/3})$; see [24] and references therein for the relevant definitions and properties. Then, the embedding in fractional Sobolev spaces $\mathcal{N}^{\alpha,2} \hookrightarrow W^{\beta,2} \cap L^{2n/(n-2\beta)}$ holds for every $\beta < \alpha$. We report a local quantitative form of it: if a map $G\colon B_1 \to \mathbb{R}^n$ satisfies

$$\sup_{0<|h|\leq 10^{-4}} |h|^{-2\alpha}\int_{B_{2/3}} |\tau_h G|^2 \, dx \leq K^2 \tag{5.36}$$

for some $K \geq 0$, then there exists a positive constant $c \equiv c(n, \alpha-\beta)$, independent of K and G, such that the following inequality holds:

$$\|G\|_{L^{2n/(n-2\beta)}(B_{1/2})} + \|G\|_{W^{\beta,2}(B_{1/2})} \leq c\left(K + \|G\|_{L^2(B_1)}\right).$$

This statement can be retrieved for instance combining [24, Lemma 2.3] with the classical fractional Sobolev embedding theorem (notice here that we have $\beta < \alpha \leq 1$). By (5.27), and recalling that $\eta \equiv 1$ on $B_{2/3}$, we can take $G \equiv V_p(Dv)$ in (5.36) and therefore conclude that for every $\beta \in (0,\alpha)$ there exists a positive constant $c \equiv c(n,p,q,\nu,L,\beta)$ such that

$$\|V_p(Dv)\|^2_{L^{\frac{2n}{n-2\beta}}(B_{1/2})} \leq c\|Dv\|^p_{L^p(B_1)} + c\|V_p(Dv)\|^2_{L^2(B_1)} + c(\|a\|^2_{L^\infty(B_1)} + [a]^2_{0,\alpha;B_1} + \sigma_m)\|Dv\|^{2q-p}_{L^{2q-p}(B_1)}$$

is true. Now (5.13) follows observing that $|V_p(Du)|^2 = |Du|^p$.

Step 4: Scaling and covering. We now recast the full notation $v \leftrightarrow v_m$ and $a \leftrightarrow \tilde{a}$; see Steps 1 and 2, and define

$$T_m := \|\tilde{a}\|^2_{L^\infty(B_1)} + [\tilde{a}]^2_{0,\alpha;B_1} + \sigma_m \,. \tag{5.37}$$

Here we prove that for every $\beta \in (0,\alpha)$ there exists $c_* \equiv c_*(n,p,q,\nu,L,\beta)$ such that for every $s,t \in [1/2,1]$, $t < s$, the inequality

$$\|Dv_m\|^p_{L^{np/(n-2\beta)}(B_t)} \leq \frac{c\|Dv_m\|^p_{L^p(B_s)}}{(s-t)^{2\beta}} + \frac{c_* T_m\|Dv_m\|^{2q-p}_{L^{2q-p}(B_s)}}{(s-t)^{2\beta}} \,, \tag{5.38}$$

holds for balls $B_t \subset B_s$ concentric to $B_1 \equiv B_1(0)$. To this end, take s,t as specified and set $r := (s-t)/8$. We then consider a covering of B_t with a family of balls $\{B_{r/2}(y)\}$, $y \in B_t$, made of at most $c(n)r^{-n}$ balls, where $c(n)$ is a constant depending only on n, and such that both the family $\{B_{r/2}(y)\}$ and the enlarged one $\{B_r(y)\}$ have the finite intersection property. This means that every ball of a considered family touches at most $\tilde{c}(n)$ different balls belonging to the same family. In every ball $B_r \equiv B_r(y)$, we apply (5.13) after scaling as in Remark 1, i.e. we apply (5.13) to $v_m(y+rx)/r$, thereby getting (recalling that $r^{2\alpha} \leq 1$ and (2.8))

$$\|Dv_m\|^p_{L^{\frac{np}{n-2\beta}}(B_{r/2})} \leq \frac{c\|Dv_m\|^p_{L^p(B_r)}}{r^{2\beta}} + \frac{cT_m\|Dv_m\|^{2q-p}_{L^{2q-p}(B_r)}}{r^{2\beta}}$$

for $c \equiv c(n,p,q,\nu,L,\beta)$. Now (5.38) follows summing up the inequalities in the last display with respect to the covering $\{B_{r/2}(y)\}$, using the finite intersection property, and the fact that, by construction, we have that the inclusion $B_r(y) \subset B_s$ holds for every ball. When summing, we have also used the elementary inequality

$$\left(\sum_B a_B\right)^{(n-2\beta)/n} \leq \sum_B a_B^{(n-2\beta)/n} \quad \text{with} \quad a_B \equiv \|Dv_m\|^{np/(n-2\beta)}_{L^{np/(n-2\beta)}(B_{r/2}(y))}.$$

Step 5: Interpolation. We prove that for every $\beta \in (n(q/p-1),\alpha)$ there exists a positive constant $c \equiv c(n,p,q,\nu,L,\beta)$ such that

$$\|Dv_m\|^p_{L^{\frac{np}{n-2\beta}}(B_{1/2})} \leq c\|Dv_m\|^p_{L^p(B_1)} + cT_m^{b_1}\|Dv_m\|^{pb_2}_{L^p(B_1)} \tag{5.39}$$

where the exponents b_1, b_2 have been introduced in (5.5) and T_m is as in (5.37). Notice that we can always assume that $q > p$ here, since otherwise the previous inequality coincides with (5.13) (note that $b_1 = b_2 = 1$ for $p = q$). Anyway, all the forthcoming estimates involve constants that remain stable when $q \to p$. Now, observe that $p < 2q-p$ holds since $q > p$ while $2q-p < np/(n-2\beta)$ is proved in (5.6)-(5.7). Therefore the interpolation inequality

$$\|Dv_m\|_{L^{2q-p}(B_s)} \leq \|Dv_m\|^{1-\theta}_{L^p(B_s)}\|Dv_m\|^{\theta}_{L^{np/(n-2\beta)}(B_s)} \tag{5.40}$$

holds for $\theta \in (0,1)$, found by solving

$$\frac{1}{2q-p} = \frac{1-\theta}{p} + \frac{(n-2\beta)\theta}{np}\,, \qquad \text{namely} \qquad \theta = \frac{n(q-p)}{\beta(2q-p)}\,. \tag{5.41}$$

We notice that $\beta > n(q/p-1)$ implies that $(2q-p)\theta < p$. Hence we apply Young's inequality with exponents $p/[\theta(2q-p)]$ and $p/[p-(2q-p)\theta]$ in (5.40) to have

$$\begin{aligned}\frac{T_m\|Dv_m\|^{2q-p}_{L^{2q-p}(B_s)}}{(s-t)^{2\beta}} &\leq \frac{T_m}{(s-t)^{2\beta}}\|Dv_m\|^{(1-\theta)(2q-p)}_{L^p(B_s)}\|Dv_m\|^{\theta(2q-p)}_{L^{np/(n-2\beta)}(B_s)}\\ &\leq \varepsilon\|Dv_m\|^p_{L^{np/(n-2\beta)}(B_s)}\\ &\quad +c(\varepsilon)\left[\frac{T_m}{(s-t)^{2\beta}}\|Dv_m\|^{(1-\theta)(2q-p)}_{L^p(B_s)}\right]^{\frac{p}{p-(2q-p)\theta}}\end{aligned} \tag{5.42}$$

Recalling the definitions of b_1, b_2 in (5.5) and that of θ in (5.41), we notice that

$$b_1 = \frac{p}{p-(2q-p)\theta} \qquad \text{and} \qquad b_2 = \frac{(1-\theta)(2q-p)}{p-(2q-p)\theta}\,,$$

so that using (5.42) in (5.38), and taking $\varepsilon \equiv \varepsilon(n,p,q,\nu,L,\beta) := 1/(2c_*)$, gives

$$\|Dv_m\|^p_{L^{\frac{np}{n-2\beta}}(B_t)} \leq \frac{1}{2}\|Dv_m\|^p_{L^{\frac{np}{n-2\beta}}(B_s)} + \frac{c\|Dv_m\|^p_{L^p(B_1)}}{(s-t)^{2\beta}} + \frac{cT_m^{b_1}\|Dv_m\|^{pb_2}_{L^p(B_1)}}{(s-t)^{2\beta b_1}}$$

again for $c \equiv c(n,p,q,\nu,L,\beta)$. We are now in position to apply Lemma 2.2 with $\rho_0 = 1/2 < 1-\varepsilon =: \rho_1$, $\theta = 1/2$, $\gamma_1 = 2\beta b_1$, $\gamma_2 = 2\beta$, $h(s) = \|Dv_m\|^p_{L^{np/(n-2\beta)}(B_s)}$; upon letting $\varepsilon \to 0$, the final outcome is exactly (5.39). Notice that Lemma 2.2 is applicable since by (5.12) we have that the function $h(\cdot)$ is bounded on the interval $(0, 1-\varepsilon)$, for every $\varepsilon \in (0,1)$.

Step 6: Letting $m \to \infty$. Using $(1.11)_1$, the minimality of v_m and (5.11), we find

$$\begin{aligned}\nu\int_{B_1}|Dv_m|^p\,dx &\leq \int_{B_1} F_m(x,Dv_m)\,dx\\ &\leq \int_{B_1} F_m(x,Du_m)\,dx\\ &= \mathcal{F}(u_m,B_1) + \sigma_m\|Du_m\|^{2q-p}_{L^{2q-p}(B_1)}\,, \end{aligned} \tag{5.43}$$

so that by (5.10) it follows that the sequence $\{v_m\}$ is bounded in $W^{1,p}(B_1)$. Up to not relabelled subsequences, we may therefore assume there exists $w \in \tilde{u} + W^{1,p}_0(B_1)$ such that $Dv_m \rightharpoonup Dw$, weakly in $L^p(B_1)$. Using in (5.43) the convergence of energies from (5.10) and that $\sigma_m\|Du_m\|^{2q-p}_{L^{2q-p}(B_1)} \to 0$, together with standard lower semicontinuity theorems for convex functionals, we find

$$\int_{B_1} F(x,Dw)\,dx \leq \liminf_m \int_{B_1} F(x,Dv_m)\,dx \leq \int_{B_1} F(x,D\tilde{u})\,dx\,.$$

In turn, the original minimality of u, and thereby the one of $\tilde{u}$ in B_1, implies that there is equality in the previous chain of inequalities. Finally, the strict convexity of the single integrand $f(\cdot)$ with respect to the variable z implies that $\tilde{u} \equiv w$ on B_1. In particular we have

$$\int_{B_1} F(x,D\tilde{u})\,dx = \liminf_m \int_{B_1} F(x,Dv_m)\,dx$$

and therefore

$$\int_{B_1} F(x,D\tilde{u})\,dx \geq \liminf_m \int_{B_1} f(Dv_m)\,dx + \liminf_m \int_{B_1} a(x)g(Dv_m)\,dx\,. \tag{5.44}$$

Again by semicontinuity (and using that $\tilde{u} = w$) we have

$$\int_{B_1} f(D\tilde{u})\,dx \le \liminf_m \int_{B_1} f(Dv_m)\,dx\,,$$
$$\int_{B_1} a(x)g(D\tilde{u})\,dx \le \liminf_m \int_{B_1} a(x)g(Dv_m)\,dx\,.$$

By (5.44) and the last two inequalities we obtain in particular that

$$\liminf_m \int_{B_1} f(Dv_m)\,dx = \int_{B_1} f(D\tilde{u})\,dx\,.$$

In turn, this and $(1.11)_1$ imply

$$\liminf_m \int_{B_1} |Dv_m|^p\,dx \le \frac{1}{\nu}\int_{B_1} f(D\tilde{u})\,dx \le \frac{L}{\nu}\int_{B_1} |D\tilde{u}|^p\,dx\,.$$

Using this last information and the semicontinuity of the $L^{np/(2-2\beta)}$-norm, inequality (5.8) follows passing to the liminf in (5.39).

Step 7: Proof of (5.2). As in the case of (5.3), see Step 1, it is sufficient to prove (5.2) in the case $\tilde{q} = np/(n-2\beta)$ for $\beta \in (n(q/p-1), \alpha)$. Now, with $B_R \subset \Omega$ such that $R \le 1$, we rewrite (5.4) as

$$\left(\fint_{B_{R/2}} |Du|^{\frac{np}{n-2\beta}}\,dx\right)^{\frac{n-2\beta}{np}} \le cB\left(\fint_{B_R} |Du|^p\,dx\right)^{1/p}, \tag{5.45}$$

where $c \equiv c(n,p,q,\nu,L,\beta)$ and

$$\begin{aligned} B &:= \left[1 + \left(\|a\|^2_{L^\infty(B_R)} + R^{2\alpha}[a]^2_{0,\alpha}\right)^{b_1}\left(\fint_{B_R} |Du|^p\,dx\right)^{b_2-1}\right]^{1/p} \\ &\le c\left[1 + (M^2+1)^{b_1}[a]^{2b_1}_{0,\alpha}\frac{R^{2\alpha b_1}}{R^{n(b_2-1)}}\left(\int_{B_R} |Du|^p\,dx\right)^{b_2-1}\right]^{1/p}. \end{aligned}$$

We have used (5.1) to achieve the last estimate. Recalling the expressions for the exponents b_1, b_2 appearing in (5.5) we compute

$$2\alpha b_1 - n(b_2-1) = \left(\frac{2\beta pn}{\beta p - n(q-p)}\right)\left(1 + \frac{\alpha}{n} - \frac{q}{p}\right),$$

which is a positive exponent by (1.3). Using that $R \le 1$ yields

$$B \le c + c(M^2+1)^{b_1/p}[a]^{2b_1/p}_{0,\alpha}\|Du\|^{b_2-1}_{L^p(B_R)}$$

again with $c \equiv c(n,p,q,\nu,L,\beta)$. This last inequality and (5.45) finally yield (5.2) with the required dependence on the constants. The proof is complete.

6. Estimates in the two phases

In this section we consider a ball B_R and give some rather delicate comparison and decay lemmas, of different nature, involving a minimiser u of the functional $\mathcal{G}$ defined in (1.10). In Lemmas 6.1 and 6.3, we assume that the functional $\mathcal{G}$ is in the (p,q)-phase, and we make perturbation around a functional modelled by $\mathcal{P}_0$ in (1.14). In Lemmas 6.2 and 6.4 we instead assume that the functional $\mathcal{F}$ is in the p-phase, through an assumption as (5.1), and we use in a quantitative way the reverse inequality (5.2) to make perturbation around functionals as in (1.13).

Lemma 6.1 (**$(p,q)-$phase comparison).** *Let $u \in W^{1,p}(\Omega)$ be a local minimiser of $\mathcal{G}$ under the assumptions* (1.2), (1.3), *and* (1.11) *and let $B_{\tilde{R}}(x_0) \subset \Omega$ be a ball such that $u \in W^{1,q}(B_{\tilde{R}}(x_0))$. With $a_0 \geq 0$, let $v \in W^{1,p}(B_R)$ be the solution to the following Dirichlet problem:*

$$\begin{cases} v \mapsto \min\limits_{w} \displaystyle\int_{B_R} \left(f(Dw) + a_0 g(Dw)\right)\, dx \\ \qquad\qquad w \in u + W^{1,p}_0(B_R)\,, \end{cases} \tag{6.1}$$

where $B_R \equiv B_R(x_0) \subset B_{\tilde{R}}(x_0)$ with $R \leq \tilde{R}$. If

$$a_0 > M[a]_{0,\alpha}\tilde{R}^{\alpha} \qquad \textit{and} \qquad \sup_{x\in B_R} |a(x) - a_0| \leq 2[a]_{0,\alpha}\tilde{R}^{\alpha} \tag{6.2}$$

hold for a positive constant M, then $v \in W^{1,q}(B_R)$ and the inequality

$$\int_{B_R} \left(|V_p(Du) - V_p(Dv)|^2 + a_0|V_q(Du) - V_q(Dv)|^2\right) dx \leq \frac{c_1}{M}\int_{B_R} \left(|Du|^p + a_0|Du|^q\right) dx \tag{6.3}$$

holds for a constant c_1 depending only on n, p, q, ν, L.

Proof. Let us preliminary observe that the minimality of v and (1.11) easily give

$$\int_{B_R} \left(|Dv|^p + a_0|Dv|^q\right) dx \leq \frac{L}{\nu}\int_{B_R} \left(|Du|^p + a_0|Du|^q\right) dx\,. \tag{6.4}$$

Since $a_0 > 0$ by (6.2) and $u \in W^{1,q}(B_R)$, it follows in particular that $v \in W^{1,q}(B_R)$. Since both u and v are minimisers, using the corresponding Euler-Lagrange equations we compute

$$\begin{aligned} \mathcal{D}_1 &:= \int_{B_R} \langle \partial f(Du) - \partial f(Dv) + a_0(\partial g(Du) - \partial g(Dv)), Du - Dv\rangle\, dx \\ &= \int_{B_R} \langle \partial f(Du) + a_0\partial g(Du), Du - Dv\rangle\, dx \\ &= \int_{B_R} [a_0 - a(x)]\langle \partial g(Du), Du - Dv\rangle\, dx := \mathcal{D}_2\,. \end{aligned}$$

We remark that taking $u-v$ as test function in the Euler-Lagrange equations of the considered functionals is legal by a standard density argument, since $u, v \in W^{1,q}(B_R)$ and $u-v \in W^{1,q}_0(B_R)$. Then (2.5) yields

$$\int_{B_R} (|V_p(Du)-V_p(Dv)|^2 + a_0|V_q(Du)-V_q(Dv)|^2)\,dx \leq c\mathcal{D}_1 \leq c|\mathcal{D}_2|\,, \tag{6.5}$$

where $c \equiv c(n,p,q,\nu,L)$. As for $\mathcal{D}_2$, by (2.3) and (6.2), we have

$$\begin{aligned}
|\mathcal{D}_2| &\leq c\int_{B_R} |a(x)-a_0||Du|^{q-1}|Du-Dv|\,dx \\
&\leq c[a]_{0,\alpha}\tilde{R}^\alpha \int_{B_R} |Du|^{q-1}|Du-Dv|\,dx \\
&\leq \frac{ca_0}{M}\int_{B_R} (|Du|^q + |Dv|^q)\,dx \\
&\leq \frac{c}{M}\int_{B_R} (|Du|^p + a_0|Du|^q)\,dx\,,
\end{aligned}$$

where in the last line we have also used (6.4). Now (6.3) follows by merging the content of the last two displays.

Lemma 6.2 (**p-phase comparison**). *Let $u \in W^{1,p}(\Omega)$ be a local minimiser of $\mathcal{G}$ under the assumptions* (1.2), (1.3) *and* (1.11) *and let $B_R \subset \Omega$ be a ball such that $R \leq 1$. Let $v \in W^{1,p}(B_R)$ be the solution to the following Dirichlet problem:*

$$\begin{cases} v \mapsto \min\limits_{w} \displaystyle\int_{B_{R/2}} f(Dw)\,dx \\ w \in u + W^{1,p}_0(B_{R/2}) \end{cases} \tag{6.6}$$

and assume that for some constant $M \geq 1$ the bound

$$\sup_{x\in B_R} a(x) \leq M[a]_{0,\alpha}R^\alpha \tag{6.7}$$

is satisfied. Then

$$\int_{B_{R/2}} |V_p(Du)-V_p(Dv)|^2\,dx \leq c(M)R^\sigma \int_{B_R} |Du|^p\,dx \tag{6.8}$$

holds with

$$\sigma := 2n\left(1+\frac{\alpha}{n}-\frac{q}{p}\right) > 0\,, \tag{6.9}$$

and for a constant $c(M)$ depending only on `data` *and M. In particular, $c(M)$ is a non-decreasing function of M and* (1.3) *implies that σ is positive. Moreover, it follows that $v \in W^{1,2q-p}(B_{R/2})$ and that the inequality*

$$\int_{B_{R/2}} |V_p(Du)-V_p(Dv)|^2 \leq cM^2[a]^2_{0,\alpha}R^{2\alpha}\int_{B_{R/2}} |Du|^{2q-p}\,dx \tag{6.10}$$

holds for a constant $c \equiv c(n,p,q,\nu,L)$, but otherwise independent of M.

Proof. This time we premise a result on the boundary higher integrability of minima of functionals as the one in (6.6). From [24, Theorem 7.7] we gain that there exists a positive constant $c \equiv c(n,p,\nu,L)$ such that

$$\int_{B_{R/2}} |Dv|^{2q-p}\,dx \le c\int_{B_{R/2}} |Du|^{2q-p}\,dx \tag{6.11}$$

holds; notice that [24, Theorem 7.7] applies here since $2q-p < np/(n-2\alpha)$ holds by (5.6)-(5.7). Then by Theorem 5.1 we have that $u \in W^{1,2q-p}(B_{R/2})$ so that the right-hand side in (6.11) is finite; in particular, this proves that $v \in W^{1,2q-p}(B_{R/2})$. The minimality of both u and v, and (2.5), (2.3) and (6.7), yield

$$\begin{aligned}
\int_{B_{R/2}} |V_p(Du)-V_p(Dv)|^2\,dx &\le c\int_{B_{R/2}} \langle \partial f(Dv)-\partial f(Du), Dv-Du\rangle\,dx\\
&= c\int_{B_{R/2}} \langle \partial f(Du), Du-Dv\rangle\,dx\\
&= c\int_{B_{R/2}} a(x)\langle \partial g(Du), Dv-Du\rangle\,dx\\
&\le c_* M[a]_{0,\alpha}R^{\alpha}\int_{B_{R/2}} |Du|^{q-1}|Du-Dv|\,dx\,,
\end{aligned}$$

for a constant c_* depending only on n,p,q,ν,L. Recalling (2.6) and applying Young's inequality we estimate (for $\varepsilon \in (0,1)$)

$$\begin{aligned}
&M[a]_{0,\alpha}R^{\alpha}|Du|^{q-1}|Du-Dv|\\
&\quad\le M[a]_{0,\alpha}R^{\alpha}(|Du|+|Dv|)^{q-1}|Du-Dv|\\
&\quad\le cM[a]_{0,\alpha}R^{\alpha}(|Du|+|Dv|)^{(2q-p)/2}|V_p(Du)-V_p(Dv)|\\
&\quad\le \varepsilon|V_p(Du)-V_p(Dv)|^2 + \frac{cM^2[a]_{0,\alpha}^2R^{2\alpha}}{\varepsilon}(|Du|+|Dv|)^{2q-p}\,.
\end{aligned}$$

Combining the content of the last two displays, choosing $\varepsilon \equiv \varepsilon(n,p,q,\nu,L) = 1/(2c_*)$ and reabsorbing terms on the left-hand side yields

$$\int_{B_{R/2}} |V_p(Du)-V_p(Dv)|^2 \le cM^2[a]_{0,\alpha}^2R^{2\alpha}\int_{B_{R/2}} (|Du|+|Dv|)^{2q-p}\,dx$$

for $c \equiv c(n,p,q,\nu,L)$. Estimating the right hand side of the previous inequality by mean of (6.11) yields (6.10). Finally, since (5.1) is in force here, we can apply (5.2) with the choice $\tilde{q} = 2q-p$ and this yields

$$\begin{aligned}
&R^{2\alpha}\int_{B_{R/2}} |Du|^{2q-p}\,dx\\
&\quad\le c(M)R^{n+2\alpha}\left(\fint_{B_R} |Du|^p\,dx\right)^{\frac{2q-p}{p}}\\
&\quad\le c(M)R^{2\alpha-\frac{2n(q-p)}{p}}\left(\int_{B_R} |Du|^p\,dx\right)^{\frac{2q-2p}{p}}\int_{B_R} |Du|^p\,dx
\end{aligned}$$

$$\leq c(M)R^{\sigma}\int_{B_R}|Du|^p\,dx\,.$$

Combining this last inequality with (6.10) yields (6.8) and the proof is complete.

Accordingly to the two lemmas above, we have two decay estimates. The first holds for all scales $\varrho \leq R \leq \tilde{R}$ for which (6.2) holds; the second instead holds for the scales $\varrho \leq R$ for which (6.13) below hold.

Lemma 6.3 (**(p,q)-phase decay at all scales).** *Let $u \in W^{1,p}(\Omega)$ be a local minimiser of $\mathcal{G}$ under the assumptions* (1.2)*,* (1.3) *and* (1.11)*, and let $B_{\tilde{R}}(x_0) \subset \Omega$ be such that $u \in W^{1,q}(B_{\tilde{R}}(x_0))$. Assume that for every ball $B_R(x_0) \subset B_{\tilde{R}}(x_0)$ with $R \leq \tilde{R}$, the conditions in* (6.2) *are satisfied for some $M \geq 1$ and some $a_0 \geq 0$. Then the inequality*

$$\int_{B_\varrho}(|Du|^p + a_0|Du|^q)\,dx \leq c_2\left[\left(\frac{\varrho}{R}\right)^n + \frac{c_1}{M}\right]\int_{B_R}(|Du|^p + a_0|Du|^q)\,dx \tag{6.12}$$

holds whenever $0 < \varrho \leq R \leq \tilde{R}$, where $c_2 \equiv c_2(n,p,q,\nu,L)$. The constant c_1 has been introduced in Lemma 6.1 and depends only on n,p,q,ν,L.

Proof. It is sufficient to prove (6.12) for $\varrho \leq R/2$. We go back to the setting of Lemma 6.1 and consider the function v defined in (6.1). By using (11.2) below and the fact that $|V_\gamma(Du)|^2 = |Du|^\gamma$ for $\gamma = p,q$, we obtain

$$\begin{aligned}
\int_{B_\varrho}(|Du|^p + a_0|Du|^q)\,dx &\leq c\int_{B_\varrho}(|V_p(Dv)|^2 + a_0|V_q(Dv)|^2)\,dx\\
&\quad + c\int_{B_R}(|V_p(Du)-V_p(Dv)|^2 + a_0|V_q(Du)-V_q(Dv)|^2)\,dx\\
&\leq c\sup_{B_\varrho}(|Dv|^p + a_0|Dv|^q)\,\varrho^n\\
&\quad + c\int_{B_R}(|V_p(Du)-V_p(Dv)|^2 + a_0|V_q(Du)-V_q(Dv)|^2)\,dx\\
&\leq c\left(\frac{\varrho}{R}\right)^n\int_{B_R}(|Du|^p + a_0|Du|^q)\,dx\\
&\quad + c\int_{B_R}(|V_p(Du)-V_p(Dv)|^2 + a_0|V_q(Du)-V_q(Dv)|^2)\,dx\,.
\end{aligned}$$

The assertion follows using Lemma 6.1 to estimate the last integral.

Lemma 6.4 (**p-phase decay at one scale).** *Let $u \in W^{1,p}(\Omega)$ be a local minimiser of $\mathcal{G}$ under the assumptions* (1.2)*,* (1.3) *and* (1.11) *and let $B_\varrho \subset B_R \subset \Omega$ be concentric balls with $0 < \varrho \leq R \leq 1$ such that*

$$\sup_{x\in B_\varrho} a(x) \leq M[a]_{0,\alpha}\varrho^\alpha \quad \textit{and} \quad \sup_{x\in B_R} a(x) \leq M[a]_{0,\alpha}R^\alpha \tag{6.13}$$

hold for some $M \geq 1$. Then it follows that

$$\int_{B_\varrho} H(x, Du)\,dx \leq c_3(M)\left[\left(\frac{\varrho}{R}\right)^n + R^\sigma\right]\int_{B_R} H(x, Du)\,dx \tag{6.14}$$

holds for a constant $c_3(M)$ depending on `data` *and M, and with $\sigma > 0$ as in* (6.9).

Proof. We observe that

$$\int_{B_\varrho} |Du|^p\,dx \leq c\left[\left(\frac{\varrho}{R}\right)^n + c(M)R^\sigma\right]\int_{B_R} |Du|^p\,dx \tag{6.15}$$

holds for $c \equiv c(n, p, q, \nu, L)$ whenever $0 < \varrho \leq R$, and for this we only need the first inequality in (6.13). This follows via the same arguments employed for Lemma 6.3; it this time rests on (6.8). We now prove (6.14) and we can restrict to the case $\varrho \leq R/2$. Now, by (6.13) and the Hölder continuity of $a(\cdot)$ we observe that

$$\sup_{x\in B_{2\varrho}} a(x) \leq \sup_{x\in B_\varrho} a(x) + 3[a]_{0,\alpha}\varrho^\alpha \leq 4M[a]_{0,\alpha}(2\varrho)^\alpha\,.$$

We can therefore apply Theorem 5.1 (with $4M$ instead of M and with $B_R \equiv B_{2\varrho}$); at this point, since $p\alpha - n(q-p) > 0$ by (1.3) and $\varrho < 1$, we have

$$\begin{aligned}
\int_{B_\varrho} a(x)|Du|^q\,dx &\leq 4M[a]_{0,\alpha}\varrho^\alpha|B_\varrho| \fint_{B_\varrho} |Du|^q\,dx \\
&\leq c\varrho^{n+\alpha}\left(\fint_{B_{2\varrho}} |Du|^p\,dx\right)^{q/p} \\
&\leq c\varrho^{\frac{p\alpha-n(q-p)}{p}}\left(\int_{B_{2\varrho}} |Du|^p\,dx\right)^{(q-p)/p}\int_{B_{2\varrho}} |Du|^p\,dx \\
&\leq c\int_{B_{2\varrho}} |Du|^p\,dx \\
&\leq c\left[\left(\frac{\varrho}{R}\right)^n + c(M)R^\sigma\right]\int_{B_R} H(x, Du)\,dx\,.
\end{aligned} \tag{6.16}$$

To make the last estimate above we have once again used (6.15) an that $2\varrho \leq R$; the constant c depends on `data` and M. Combining (6.16) with (6.15) yields (6.14).

7. Proof of Theorem 1.3

Step 1: Universal choice of the constants. In the following, we shall apply Lemmas 6.1-6.4 for values of a_0 being such that $a_0 \leq \|a\|_{L^\infty}$. We shall also apply Lemma 2.1 with the choice

$$\phi(\varrho) \equiv \int_{B_\varrho} (|Du|^p + a_0|Du|^q)\,dx\,, \tag{7.1}$$

with $\tilde{c} = c_2 \equiv c_2(n,p,q,\nu,L)$, where c_2 is the constant appearing in Lemma 6.3, and where the number $\delta \in (0,n)$ is fixed as in the statement of Theorem 1.3. This allows to determine the corresponding number $\bar{\varepsilon} \equiv \bar{\varepsilon}(n,p,q,\nu,L,\delta) > 0$ given by Lemma 2.1. We now consider $c_1 \equiv c_1(n,p,q,\nu,L)$ coming from Lemma 6.1 and determine $M \geq 4$ large enough in order to have

$$\frac{c_1}{\bar{\varepsilon}} \leq M\,. \tag{7.2}$$

Recalling the dependence on n,p,q,ν,L,δ of $\bar{\varepsilon}$ and c_1, this fixes M as a quantity depending only on n,p,q,ν,L,δ. With this value of M we consider Lemma 6.4 and determine the corresponding constant $c_3(M)$. This quantity depends only on $\mathtt{data}$ and δ. Accordingly, we take $\tau \equiv \tau(\mathtt{data},\delta) \in (0,1/4)$ such that

$$2c_3(M)\tau^{\delta} \leq 1 \Longleftrightarrow \tau \leq \left(\frac{1}{2c_3(M)}\right)^{1/\delta}\,. \tag{7.3}$$

Finally, we determine the positive radius $R_1 \leq 1$ in such a way that

$$R_1 \leq \tau^{n/\sigma}\,. \tag{7.4}$$

Taking the previous dependences into account, R_1 depends only on $\mathtt{data},\delta$. In the next step we shall prove that (1.12) holds whenever $R \leq R_1$; once this is proved, estimate (1.12) with any $R \leq 1$ follows by enlarging the constant of a factor $R_1^{\delta-n}$, and recalling again that R_1 depends on $\mathtt{data}$ and δ.

Step 2: Exit time argument. With R_1,τ and M fixed in the previous step as quantities depending only on $\mathtt{data},\delta$, we take $R_0 \leq R_1$ and a ball $B_{R_0} \subset \Omega$, and finally build the sequence of nested balls as in (1.17) (all the balls considered are concentric to B_{R_0}). We consider the condition

$$\sup_{x\in B_{\tau^k R_0}} a(x) \leq M[a]_{0,\alpha}\tau^{k\alpha}R_0^{\alpha}\,, \tag{7.5}$$

and, accordingly, we define the following exit time index:

$$m := \min\,\{k \in \mathbb{N}\cup\{\infty\}\,:\,(7.5)\,\text{fails}\}\,. \tag{7.6}$$

Notice that m is allowed to be infinite. The idea is now as follows: if $m < \infty$, after the exit time the functional enters the (p,q)-phase, and then we can use (6.12) from Lemma 6.3. Before the exit time argument the functional is in the p-phase, and we shall inductively prove a decay estimate based on (6.14) from Lemma 6.4.

Step 3: Iteration in the (p,q)-phase. Here we assume that $m < \infty$, otherwise go directly to Step 4 (on the other hand, if $m = 0$, then Step 4 can be avoided). By the very definition in (7.6) we have now that

$$\sup_{x\in B_{\tau^m R_0}} a(x) > M[a]_{0,\alpha}\tau^{m\alpha}R_0^{\alpha}\,.$$

We are then led to apply Lemma 6.3 (keep also in mind Lemma 6.1) with the choice

$$a_0 := \sup_{x\in B_{\tau^m R_0}} a(x) \qquad \text{and} \qquad \tilde{R} := \tau^m R_0$$

and these make conditions (6.2) satisfied whenever $R \leq \tau^m R_0$, when considering a smaller ball B_R concentric to $B_{\tau^m R_0}$. Moreover, for $x \in B_{\tau^m R_0}$ we have

$$\begin{aligned} a(x) &\geq a_0 - |a(x) - a_0| \\ &\geq a_0/2 + (M/2)[a]_{0,\alpha}(\tau^m R_0)^\alpha - 2[a]_{0,\alpha}(\tau^m R_0)^\alpha \geq a_0/2 \end{aligned}$$

(recall we picked $M \geq 4$) so that in any case we have

$$a_0 \geq a(x) \geq a_0/2 \qquad \text{for every } x \in B_{\tau^m R_0} \,. \tag{7.7}$$

In particular, we have that $u \in W^{1,q}(B_{\tau^m R_0})$ (recall that in the case $m = 0$ the ball $B_{\tau^m R_0}$ can touch $\partial\Omega$). Lemma 6.3 is therefore applicable and we come up with (6.12), that with the notation in (7.1) reads as

$$\phi(\varrho) \leq c_2 \left[\left(\frac{\varrho}{R} \right)^n + \frac{c_1}{M} \right] \phi(R) \,,$$

and this holds whenever $0 < \varrho \leq R \leq \tau^m R_0$. Due to the choice in (7.2), we can now apply Lemma 2.1 thereby concluding with the following decay estimate:

$$\int_{B_\varrho} (|Du|^p + a_0|Du|^q)\, dx \leq c \left(\frac{\varrho}{\tau^m R_0} \right)^{n-\delta} \int_{B_{\tau^m R_0}} (|Du|^p + a_0|Du|^q)\, dx \,,$$

that again holds whenever $\varrho \leq \tau^m R_0$, and for a constant $c \equiv c(n, p, q, \nu, L, \delta)$. By (7.7) we can conclude with

$$\int_{B_\varrho} H(x, Du)\, dx \leq c \left(\frac{\varrho}{\tau^m R_0} \right)^{n-\delta} \int_{B_{\tau^m R_0}} H(x, Du)\, dx \tag{7.8}$$

for $0 < \varrho \leq \tau^m R_0$ and $c \equiv c(n, p, q, \nu, L, \delta)$. This completes the proof of (1.12) when $m = 0$, upon renaming R_0 by R. When $m > 0$ it remains to estimate the last integral appearing in the above display; this will be done in the next step.

Step 4: Iteration in the p-phase. By the previous step we can assume that $m \geq 1$. Notice that m here can be infinite and therefore we distinguish two cases, the first being the one when $m < \infty$, the second when $m = \infty$. We proceed treating the former, the modifications for the latter will be done after. We start proving by induction that for every $k \in \{0, \ldots, m-1\}$ it holds that

$$\int_{B_{\tau^k R_0}} H(x, Du)\, dx \leq \tau^{k(n-\delta)} \int_{B_{R_0}} H(x, Du)\, dx \,, \tag{7.9}$$

and we can assume that $m \geq 2$ otherwise there is nothing to prove. For $k = 0$ - induction basis - (7.9) is trivial. Assume that now (7.9) holds for a certain k such that $0 \leq k < m - 1$. Since $k + 1 < m$ then

$$\sup_{x \in B_{\tau^{k+1} R_0}} a(x) \leq M[a]_{0,\alpha} \left[\tau^{k+1} R_0 \right]^\alpha$$

and

$$\sup_{x \in B_{\tau^k R_0}} a(x) \leq M[a]_{0,\alpha} [\tau^k R_0]^\alpha \,.$$

We can therefore apply Lemma 6.4 with the choice $R = \tau^k R_0$, $\varrho = \tau^{k+1} R_0$, and we obtain (7.8); using (7.3)-(7.4) and that $R_0 \leq R_1$ we have

$$\begin{aligned}\int_{B_{\tau^{k+1}R_0}} H(x, Du)\, dx &\leq c_3(M)\,[\tau^n + R_1^\sigma] \int_{B_{\tau^k R_0}} H(x, Du)\, dx \\ &\leq 2c_3(M)\tau^\delta \tau^{n-\delta} \int_{B_{\tau^k R_0}} H(x, Du)\, dx \\ &\leq \tau^{n-\delta} \int_{B_{\tau^k R_0}} H(x, Du)\, dx\,.\end{aligned}$$

Using the induction assumption (7.9) we conclude with

$$\int_{B_{\tau^{k+1}R_0}} H(x, Du)\, dx \leq \tau^{(k+1)(n-\delta)} \int_{B_{R_0}} H(x, Du)\, dx$$

so that, after induction, the family of inequalities in (7.9) holds whenever $k \in \{0, \ldots, m-1\}$. In turn the previous inequality implies that

$$\int_{B_\varrho} H(x, Du)\, dx \leq c \left(\frac{\varrho}{R_0}\right)^{n-\delta} \int_{B_{R_0}} H(x, Du)\, dx \tag{7.10}$$

holds whenever $\tau^m R_0 \leq \varrho \leq R_0$, and again for $c \equiv c(\texttt{data}, \delta)$. Indeed, let us take $\varrho \in [\tau^m R_0, R_0)$; then there exists an index k such that $1 \leq k \leq m$ and $\tau^k R_0 \leq \varrho < \tau^{k-1} R_0$. We use (7.9) to estimate

$$\begin{aligned}\int_{B_\varrho} H(x, Du)\, dx &\leq \int_{B_{\tau^{k-1}R_0}} H(x, Du)\, dx \\ &\leq \frac{\tau^{k(n-\delta)}}{\tau^{n-\delta}} \int_{B_{R_0}} H(x, Du)\, dx \\ &\leq \frac{1}{\tau^{n-\delta}} \left(\frac{\varrho}{R_0}\right)^{n-\delta} \int_{B_{R_0}} H(x, Du)\, dx\end{aligned} \tag{7.11}$$

so that (7.10) follows by taking $c := \tau^{\delta - n}$, which is a constant that still depends on $\texttt{data}$ and δ since $\tau \equiv \tau(\texttt{data}, \delta)$ - see (7.3). Therefore, upon renaming R_0 by R, we have established (1.12) in the range $\tau^m R_0 \leq \varrho \leq R_0$. In order to prove (1.12) for the range $0 < \varrho < \tau^m R_0$ it is sufficient to estimate the right hand side of (7.8) by (7.10) written in the case $\varrho = \tau^m R_0$. This completes the proof in the case the exit time m is finite. The case $m = \infty$ is actually simpler: the functional always remains in the p-phase and by induction, exactly as when $m < \infty$, we prove that (7.9) holds for every $k \in \mathbb{N}$. Then, estimating as in (7.11), we prove that (7.10) holds for $\varrho \in (0, R_0)$, so that (1.12) follows (with R_0 renamed by R).

8. Proof of Theorem 1.2

Step 1: Morrey type estimate. With $\Omega' \Subset \Omega$ being a open subset, using a standard covering argument, by Theorem 1.3 for every $\delta \in (0,n)$ we can find a constant c, and a radius $R_1 \leq \min\{1, \operatorname{dist}(\Omega', \partial\Omega)/100\}$, both depending only on $\texttt{data}, \delta, \operatorname{dist}(\Omega', \partial\Omega)$ and $\|H(\cdot, Du)\|_{L^1(\Omega)}$, such that the estimate

$$\int_{B_s} |Du|^p \, dx \leq \int_{B_s} H(x, Du) \, dx \leq cs^{n-\delta} \tag{8.1}$$

holds whenever $B_s \subset \Omega$ is a ball with center in Ω' and $s \leq R_1$.

Step 2: A family of frozen problems. Now, let us take $x_0 \in \Omega'$ and consider concentric balls $B_R \equiv B_R(x_0)$ such that $2R \leq R_1$; it follows in particular that $B_{2R} \Subset \Omega$. We set $a_0 := \sup_{x \in B_R} a(x)$ and analyse an alternative on B_R. In the first case we look at Lemma 6.1 and assume that a_0 satisfies (6.2) with

$$\tilde{R} \equiv R \qquad \text{and} \qquad M \equiv 8/R^{\alpha/2} \, . \tag{8.2}$$

Notice that $M \geq 8$ since $R \leq 1$. Observe that by the definition of a_0, the second inequality in (6.2) is always verified and therefore if (6.2) is satisfied this means

$$\sup_{x \in B_R} a(x) > \left(8/R^{\alpha/2}\right) [a]_{0,\alpha} R^\alpha \, . \tag{8.3}$$

We also notice that since $B_{2R} \Subset \Omega$ by Theorem 5.1 we have in particular that $u \in W^{1,q}(B_R)$. In this case we define $v \in W^{1,q}(B_R)$ as in (6.1) and by Lemma 6.1 with (8.2) we come up with (6.3), that can now be restated as

$$\int_{B_R} (|V_p(Du) - V_p(Dv)|^2 + a_0 |V_q(Du) - V_q(Dv)|^2) \, dx \leq cR^{\frac{\alpha}{2}} \int_{B_R} (|Du|^q + 1) \, dx \, ,$$

for a constant c depending only on n, p, q, ν, L. The other case of the alternative is when (6.2) does not hold and this in turn means that (8.3) fails. Therefore we can define $v \in W^{1,q}(B_{R/2})$ as in (6.6) and apply (6.10) from Lemma 6.2 with (8.2). This gives

$$\int_{B_{R/2}} |V_p(Du) - V_p(Dv)|^2 \leq cR^\alpha \int_{B_{R/2}} |Du|^{2q-p} \, dx$$

for $c \equiv c(n, p, q, \nu, L, [a]_{0,\alpha})$. All in all, by defining for every $R \leq R_1/2$ the coefficient

$$a_0(R) := \begin{cases} \|a\|_{L^\infty(B_R)} & \text{if (8.3) holds} \\ 0 & \text{if (8.3) does not hold} \, , \end{cases}$$

we have built a local minimiser $v \equiv v_R \in W^{1,q}(B_{R/2}(x_0))$ of the functional

$$W^{1,p}(B_{R/2}) \ni w \mapsto \int_{B_{R/2}} (f(Dw) + a_0(R) g(Dw)) \, dx \tag{8.4}$$

such that, for $c \equiv c(n,p,q,\nu,L,[a]_{0,\alpha})$, the following inequality is true:

$$\begin{aligned}\fint_{B_{R/2}} (|V_p(Du)-V_p(Dv_R)|^2 + a_0(R)|V_q(Du)-V_q(Dv_R)|^2)\,dx \\ \leq cR^{\alpha/2}\fint_{B_R}(|Du|^{2q-p}+1)\,dx\,. \quad (8.5)\end{aligned}$$

Moreover, by the way v_R has been constructed, using its minimality and the growth conditions on $f(\cdot)$ and $g(\cdot)$ in (1.11), it follows

$$\int_{B_{R/2}} (|Dv_R|^p + a_0(R)|Dv_R|^q)\,dx \leq \frac{L}{\nu}\int_{B_R} (|Du|^p + a_0(R)|Du|^q)\ dx\,. \quad (8.6)$$

Step 3: Using (8.1) *and higher integrability.* We estimate the right hand side in (8.5) by mean of (5.3) and then (8.1) as follows

$$\fint_{B_R}(|Du|^{2q-p}+1)\,dx \leq c\left(\fint_{B_{2R}}(|Du|^p+1)\,dx\right)^b \leq cR^{-\delta b} \quad (8.7)$$

with $c \equiv c\left(\texttt{data},\delta,\operatorname{dist}(\Omega',\partial\Omega),\|H(\cdot,Du)\|_{L^1(\Omega)}\right)$ and $b \equiv b(n,p,q,\alpha) \geq 1$; we shall initially take $\delta < \alpha/(10b)$, making further restrictions later. Then (8.5) yields

$$\fint_{B_{R/2}} (|V_p(Du)-V_p(Dv_R)|^2 + a_0(R)|V_q(Du)-V_q(Dv_R)|^2)\,dx \leq cR^{\alpha/2-b\delta}\,. \quad (8.8)$$

Now observe that in the case $q \geq 2$, (2.6) immediately yields

$$a_0(R)\fint_{B_{R/2}} |Du-Dv_R|^q\,dx \leq cR^{\alpha/2-b\delta}\,.$$

Instead, when $q<2$, by using again (2.6) and Hölder's inequality with exponents $2/q$ and $2/(2-q)$, and also recalling that $R \leq 1$, we estimate as

$$\begin{aligned}
a_0(R)&\fint_{B_{R/2}} |Du-Dv_R|^q\,dx \\
&\leq ca_0(R)\fint_{B_{R/2}} |V_q(Du)-V_q(Dv_R)|^q(|Du|+|Dv_R|)^{q(2-q)/2}\,dx \\
&\leq c\left(\fint_{B_{R/2}} a_0(R)|V_q(Du)-V_q(Dv_R)|^2\,dx\right)^{q/2} \\
&\qquad \cdot \left(\fint_{B_{R/2}} a_0(R)(|Du|+|Dv_R|)^q\,dx\right)^{(2-q)/2} \\
&\leq cR^{\alpha/4-b\delta/2}\cdot\left(\fint_{B_R}(|Du|^q+1)\,dx\right)^{1/2} \\
&\leq cR^{\frac{\alpha}{4}-\frac{b\delta(3q-p)}{4q-2p}}\,.
\end{aligned}$$

Here we have used (8.6), (8.7) and (8.8). Arguing in a completely similar for the homologous integrals involving $|Du - Dv_R|^p$, in any case we conclude that, for $c \equiv c\left(\texttt{data}, \delta, \operatorname{dist}(\Omega', \partial\Omega), \|H(\cdot, Du)\|_{L^1(\Omega)}\right)$, the following inequality holds:

$$\fint_{B_{R/2}} (|Du - Dv_R|^p + a_0(R)|Du - Dv_R|^q)\, dx \leq cR^{\alpha/4 - b\delta}\,. \tag{8.9}$$

Step 4: Conclusion. For $0 < \varrho \leq R/2 \leq R_1/4$ we apply (11.3) from Theorem 11.1 (with B_R there replaced by $B_{R/2}$ here and with $a_0 \equiv a_0(R)$) to each v_R introduced in Step 2, which is a minimiser of the functional in (8.4); we have

$$\begin{aligned}
&\fint_{B_\varrho} |Du - (Du)_{B_\varrho}|^p\, dx \\
&\quad \leq 2^p \fint_{B_\varrho} |Du - (Dv_R)_{B_\varrho}|^p\, dx \\
&\quad \leq c \fint_{B_\varrho} |Dv_R - (Dv_R)_{B_\varrho}|^p\, dx + c \fint_{B_\varrho} |Du - Dv_R|^p\, dx \\
&\quad \leq c \left(\frac{\varrho}{R}\right)^{p\tilde{\alpha}} \fint_{B_{R/2}} (|Dv_R|^p + a_0(R)|Dv_R|^q)\, dx \\
&\qquad\quad + c\left(\frac{R}{\varrho}\right)^n \fint_{B_{R/2}} |Du - Dv_R|^p\, dx \\
&\quad \leq c \left(\frac{\varrho}{R}\right)^{p\tilde{\alpha}} \fint_{B_R} (|Du|^p + a_0(R)|Du|^q)\, dx \\
&\qquad\quad + c\left(\frac{R}{\varrho}\right)^n \fint_{B_{R/2}} (|Du - Dv_R|^p + a_0(R)|Du - Dv_R|^q)\, dx \\
&\quad \leq c \left(\frac{\varrho}{R}\right)^{p\tilde{\alpha}} \fint_{B_R} |Du|^p\, dx + c\left(\frac{\varrho}{R}\right)^{p\tilde{\alpha}} \left(\fint_{B_R} |Du|^{2q-p}\, dx\right)^{q/(2q-p)} \\
&\qquad\quad + c\left(\frac{R}{\varrho}\right)^n \fint_{B_{R/2}} (|Du - Dv_R|^p + a_0(R)|Du - Dv_R|^q)\, dx\,,
\end{aligned}$$

for a constant c depending on n, p, q, ν, L and $\|a\|_{L^\infty}$. We now use (8.1), (8.7) and (8.9) to estimate the last three integrals, respectively. Recalling that $R \leq 1$ and that $q/(2q-p) \leq 1$, a few elementary manipulations yield, for $0 < \varrho \leq R/2$

$$\fint_{B_\varrho} |Du - (Du)_{B_\varrho}|^p\, dx \leq c\left(\frac{\varrho}{R}\right)^{p\tilde{\alpha}} \left[R^{-\delta} + R^{-\delta b}\right] + c\left(\frac{R}{\varrho}\right)^n R^{\alpha/4 - b\delta}\,,$$

again for $c \equiv c\left(\texttt{data}, \delta, \operatorname{dist}(\Omega', \partial\Omega), \|H(\cdot, Du)\|_{L^1(\Omega)}\right)$. Again with the help of (8.1), the previous inequality continue to hold whenever $0 < \varrho \leq R \leq R_1/2 \leq 1/2$, with B_{2R} as above. We start taking ϱ and R linked by $\varrho = R^{1+\varepsilon}$ for some positive $\varepsilon \in (0,1)$. This yields, again via elementary manipulations

$$\fint_{B_\varrho} |Du - (Du)_{B_\varrho}|^p\, dx \leq c\varrho^{\frac{\varepsilon p\tilde{\alpha} - \delta b}{1+\varepsilon}} + c\varrho^{\frac{\alpha/4 - n\varepsilon - b\delta}{1+\varepsilon}}\,.$$

Choosing, for instance, $\varepsilon := \alpha/(16n)$ and $\delta := \tilde{\alpha}\alpha/(32bn)$, yields

$$\fint_{B_\varrho} |Du - (Du)_{B_\varrho}|^p \, dx \leq c\varrho^{\tilde{\alpha}\alpha/(64n)} .$$

Summarising, the previous estimate holds for any ball B_ϱ with centre in Ω' and whenever $\varrho \leq R_1^{1+\varepsilon}$ with R_1 as above. Since $\Omega' \Subset \Omega$ is arbitrary, by a well-known characterisation of Hölder continuity due to Campanato and Meyers, the previous inequality implies that $Du \in C^{0,\beta}_{\text{loc}}(\Omega)$ for $\beta = \tilde{\alpha}\alpha/(64np)$ and the proof is complete.

Remark 6 (Proof of Theorem 1.4). The proof of Theorem 1.4 can be obtained exactly as the proof of Theorem 1.2. Indeed, this last one is essentially based on a perturbation argument, that extends to the vectorial case verbatim. The only difference relies in the use of the reference estimates: instead of using Theorem 11.1 we have to use Theorem 11.2. It is precisely here that the structure condition comes into the play: direct dependence on the modulus of the gradient of the integrand is required. It is indeed possible to prove that Theorem 1.4 extends to minimisers of functionals of the type

$$W^{1,p}(\Omega;\mathbb{R}^N) \ni w \mapsto \int_\Omega [\tilde{f}(|Dw|) + a(x)\tilde{g}(|Dw|)] \, dx ,$$

where the non-negative functions $f(t)$ and $g(t)$ are close to t^p and t^q respectively, in a suitable $C^{2,\beta}$-sense.

9. Proof of Theorem 1.1: (1.8)

We prove (1.8), from which

$$u \in C^{0,(p-n+p\delta_g)/(p+p\delta_g)}_{\text{loc}}(\Omega)$$

follows when $p > n/(1+\delta_g)$ by Morrey's embedding theorem, as (1.7) implies $u \in W^{1,p(1+\delta_g)}_{\text{loc}}(\Omega)$. For (1.8) we follow the methods of [19] and [21, Chapter 7]. We start taking a ball $B_R \subset \Omega$ with $R \leq 1$; all the subsequent balls will be concentric to this one. With $\varrho \leq t < s \leq R$, we determine a cut-off function $\eta \in C^\infty_0(B_s)$ such that $|D\eta| \leq 4/(s-t)$, $0 \leq \eta \leq 1$ and $\eta \equiv 1$ on B_t. We set $w := u - \eta(u-(u)_{B_R})$, to be used as a competitor. Notice that $H(\cdot, Du) \in L^1(\Omega)$ (which follows by the minimality of u) and (3.13), imply also

$$a(x)|Dw|^q + |u|^q \in L^1(\Omega) \tag{9.1}$$

by (1.3) and Sobolev embedding theorem (recall that $u \in W^{1,p}(\Omega)$ and that $q \leq p + p\alpha/n \leq p^*$). The minimality of u then gives $\mathcal{F}(u, B_s) \leq \mathcal{F}(w, B_s)$. In turn

(1.5) yields $\nu\mathcal{P}(u, B_s) \le L\mathcal{P}(w, B_s)$, with $\mathcal{P}$ as in (1.1). Since $Dw \equiv 0$ on B_t we have

$$\begin{aligned}\int_{B_s} H(x, Du)\,dx &\le c_* \int_{B_s\setminus B_t} H(x, Du)\,dx \\ &\quad + \frac{c}{(s-t)^p}\int_{B_s} |u-(u)_{B_R}|^p\,dx + \frac{c}{(s-t)^q}\int_{B_s} a(x)|u-(u)_{B_R}|^q\,dx\,,\end{aligned}$$

with c_* and c both depending on n, p, q, ν, L. The right hand side of the previous inequality is finite by (9.1). By "filling the hole", i.e. adding to both sides the quantity

$$c_* \int_{B_t} H(x, Du)\,dx$$

we come up with

$$\begin{aligned}\int_{B_t} H(x, Du)\,dx &\le \theta \int_{B_s} H(x, Du)\,dx \\ &\quad + \frac{c}{(s-t)^p}\int_{B_R} |u-(u)_{B_R}|^p\,dx + \frac{c}{(s-t)^q}\int_{B_R} a(x)|u-(u)_{B_R}|^q\,dx\,,\end{aligned}$$

where $\theta = c_*/(c_*+1) < 1$, and for a new constant c. We can now apply Lemma 2.2 with $h(t) \equiv \int_{B_t} H(x, Du)\,dx$, $\gamma_1 \equiv p$ and $\gamma_2 \equiv q$, in order to conclude with the following intrinsic Caccioppoli estimate:

$$\int_{B_\varrho} H(x, Du)\,dx \le c\int_{B_R} H\left(x, \frac{u-(u)_{B_R}}{R-\varrho}\right)\,dx\,. \tag{9.2}$$

Taking $\varrho = R/2$ and passing to averages yields

$$\fint_{B_{R/2}} H(x, Du)\,dx \le c\fint_{B_R} H\left(x, \frac{u-(u)_{B_R}}{R}\right)\,dx\ ,$$

while using (1.18) we conclude that whenever $B_R \subset \Omega$ is such that with $R \le 1$

$$\fint_{B_{R/2}} H(x, Du)\,dx \le c\left(\fint_{B_R} [H(x, Du)]^{d_2}\ dx\right)^{1/d_2}$$

holds for a constant $c \equiv c(n, p, q, \nu, L, \alpha, [a]_{0,\alpha}, \|Du\|_{L^p})$. Recalling that $d_2 < 1$, we have that (1.8) follows by a suitable localised version of Gehring's lemma (see for instance [21, Chapter 6]). We just remark that the above proof only uses the growth assumptions (1.5) and the Sobolev-Poincaré inequality of Theorem 1.6. Therefore the proof works verbatim in the case we are considering vector-valued variational problems and minimisers $u\colon \Omega \to \mathbb{R}^N$, $N > 1$.

10. Proof of Theorem 1.1: (1.9)

Step 1: Intrinsic Caccioppoli inequality in terms of $H(\cdot)$. Indeed, u satisfies

$$\int_{B_t} H(x, D(u-k)_\pm)\, dx \leq c \int_{B_s} H\left(x, \frac{(u-k)_\pm}{s-t}\right) dx \tag{10.1}$$

for every $k \in \mathbb{R}$, $t < s$, where $c \equiv c(n, p, q, \nu, L)$, whenever $B_t \subset B_s$ are concentric balls with $t < s$. We are using the standard notation

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

The proof of (10.1) follows along the lines of (9.2), by this time taking as a competitor $w := u - \eta(u-k)_+$ in the case we are interested in the version with $(u-k)_+$. As for the version with $(u-k)_-$, it is sufficient to note that $-u$ is still a local minimiser of a functional as $\mathcal{F}$ with integrand $F(x, v, z)$ replaced by $F(x, -v, -z)$, that still satisfies (1.5), and to apply the version with $_+$ to this case. See also [19] for more details. Accordingly, we shall also use the standard notation

$$A(k, s) := B_s \cap \{u > k\} \qquad \text{and} \qquad B(k, s) := B_s \cap \{u < k\}\ .$$

Step 2: u *is locally bounded*. We prove that the local estimate

$$\begin{aligned} &\|u\|_{L^\infty(B_{R/2})} \\ &\quad \leq c\left(n, p, q, \nu, L, \alpha, [a]_{0,\alpha}, R, \|Du\|_{L^p(B_R)}, \|H(\cdot, u)\|_{L^1(B_R)}\right) \end{aligned} \tag{10.2}$$

holds whenever $B_R \subset \Omega$ is a ball. As usual we may assume that $B_R \equiv B_1$, scaling as in Remark 1 and considering the new minimiser $\tilde{u}$. Next, take numbers $0 \leq h < k$, $1/2 \leq \varrho < s \leq 1$ and concentric balls $B_\varrho \subset B_t \subset B_s$ with $t := (s+\varrho)/2$. Then take a related cut-off function η such that $\eta \in C^\infty_0(B_t)$, $\eta \equiv 1$ on B_ϱ and $|D\eta| \leq c/(s-\varrho)$. Applying (3.15) and Hölder's inequality yields

$$\begin{aligned} &\int_{B_t} H\left(x, \eta(\tilde{u}-k)_+\right)\, dx \\ &\qquad \leq \frac{c}{(s-\varrho)^{q-p}} \left(\int_{B_t} \left[H\left(x, D(\eta(\tilde{u}-k)_+)\right)\right]^{d_2}\, dx\right)^{1/d_2} \\ &\qquad \leq \frac{c}{(s-\varrho)^{q-p}} \int_{B_t} H\left(x, D(\eta(\tilde{u}-k)_+)\right)\, dx\, |A(k,t)|^{d_3}\ , \end{aligned}$$

where $d_2 \equiv d_2(n, p, q, \alpha) \in (0, 1)$, $d_3 := (1-d_2)/d_2 > 0$ and with c depending on $n, p, q, \alpha, [a]_{0,\alpha}, R^{-n/p}\|Du\|_{L^p}, R^{-n/p-1}\|u\|_{L^p}$ (keep (2.8) in mind). On the other hand, using the properties of $H(\cdot)$ and η we have

$$\begin{aligned} &\int_{B_t} H\left(x, D(\eta(\tilde{u}-k)_+)\right)\, dx \\ &\quad \leq c \int_{B_t} H\left(x, \frac{(\tilde{u}-k)_+}{s-\varrho}\right) dx + c \int_{B_t} H(x, D(\tilde{u}-k)_+)\, dx\, . \end{aligned} \tag{10.3}$$

Combining the last two displays with (10.1) and recalling that $h < k$ yields

$$\int_{A(k,\varrho)} H\left(x, \tilde{u} - k\right)\, dx \leq \frac{c}{(s-\varrho)^{q-p}} \int_{A(h,s)} H\left(x, \frac{\tilde{u}-h}{s-\varrho}\right)\, dx\, |A(k,s)|^{d_3}\,, \qquad (10.4)$$

where $c \equiv c(n,p,q,\nu,L,\alpha,[a]_{0,\alpha}, R^{-n/p-1}\|u\|_{W^{1,p}})$. Consider now a sequence of concentric (to B_1) nested balls $\{B_{\varrho_i}\}$, where $\varrho_i := (1+2^{-i})/2$ for $i \geq 0$ and for $T > 0$ to be chosen in a few lines, define the levels $k_i := 2T(1-2^{-(i+1)})$. We apply (10.4) with $\varrho \equiv \varrho_{i+1}$, $s \equiv \varrho_i$, $k \equiv k_{i+1}$ and $h \equiv k_i$. This gives, for $i \geq 0$

$$\int_{A(k_{i+1},\varrho_{i+1})} H\left(x, \tilde{u} - k_{i+1}\right)\, dx \leq c4^{iq} \int_{A(k_i,\varrho_i)} H\left(x, \tilde{u} - k_i\right)\, dx\, |A(k_{i+1},\varrho_i)|^{d_3}\,.$$

Moreover, we have

$$\Psi_i := T^{-p} \int_{A(k_i,\varrho_i)} H\left(x, \tilde{u} - k_i\right)\, dx \geq T^{-p}(k_{i+1}-k_i)^p |A(k_{i+1},\varrho_i)|\,.$$

Combining the last two displays and making obvious manipulations, we obtain

$$\Psi_{i+1} \leq c_4 4^{(1+d_3)qi} \Psi_i^{1+d_3}\,,$$

for every $i \geq 0$, for a constant c_4 depending only on $n,p,q,\nu,L,\alpha,[a]_{0,\alpha}$ and $R^{-n/p-1}\|u\|_{W^{1,p}}$. We can now use a standard iteration lemma (see [21, Lemma 7.1]) that tells that $\Psi_i \to 0$ provided there exists a positive constant $c \geq 1$, depending only on c_4, d_3, q, such that $\Psi_0 \leq 1/c$; in this way c finally depends only on $n,p,q,\nu,L,\alpha,[a]_{0,\alpha}$ and $R^{-n/p-1}\|u\|_{W^{1,p}}$. This condition can be satisfied by choosing

$$T := c\left[\|H\left(\cdot, \tilde{u}_+\right)\|_{L^1(B_1)}\right]^{1/p}$$

for a suitable constant c depending only on $n,p,q,\nu,L,\alpha,[a]_{0,\alpha}$ and the norm $R^{-n/p-1}\|u\|_{W^{1,p}}$. It follows that $\tilde{u} \leq 2T$ a.e. on $B_{1/2}$:

$$\|\tilde{u}_+\|^p_{L^\infty(B_{1/2})} \leq c\|H\left(\cdot, \tilde{u}_+\right)\|_{L^1(B_1)}.$$

Repeating the same argument for $-\tilde{u}$ - which is a local minimiser of a functional as $\mathcal{F}$ with integrand $F(x,v,z)$ replaced by $F(x,-v,-z)$, that still satisfies (1.5) - estimate (10.2) follows after rescaling.

Step 3: Oscillation reduction in the p-phase. From now on we assume $p \leq n/(1+\delta_g)$ and prove the Hölder continuity of u. All the balls B_R considered in the rest of the proof of Theorem 1.1 will be such that $R \leq 1$; we shall recall this fact several times later. Step 2 implies that $u \in L^\infty_{\rm loc}(\Omega)$. The key fact here is that in the p-regime then minimisers satisfy a Caccioppoli type inequality similar to the one that holds for standard functionals with p-growth. The difference appears in terms of a controllable quantity.

Lemma 10.1 (Almost standard Caccioppoli's inequality). *Let $B_R \Subset \Omega$ be a ball such that $R \leq 1$ and*

$$\sup_{B_R} a(x) \leq 4[a]_{0,\alpha} R^{\alpha} \tag{10.5}$$

is satisfied. Then the following Caccioppoli's inequality holds on concentric balls $B_s \subset B_t \subset B_R$, whenever $0 < t < s \leq R$ and $k \in \mathbb{R}$ with $|k| \leq 2\|u\|_{L^\infty(B_R)}$:

$$\int_{B_t} |D(u-k)_\pm|^p \, dx \leq c \left[1 + \left(\frac{R}{s-t}\right)^q\right] \int_{B_s} \left|\frac{(u-k)_\pm}{R}\right|^p \, dx \tag{10.6}$$

where $c \equiv c(n, p, q, \nu, L, [a]_{0,\alpha}, \|u\|_{L^\infty(B_R)})$.

Proof. We shall manipulate (10.1); this trivially yields

$$\begin{aligned}
\int_{B_t} |D(u-k)_\pm|^p \, dx &\leq \int_{B_t} H(x, D(u-k)_\pm) \, dx \\
&\leq c \left(\frac{R}{s-t}\right)^p \int_{B_s} \left|\frac{(u-k)_\pm}{R}\right|^p \, dx \\
&\quad + c \left(\frac{R}{s-t}\right)^q \int_{B_s} a(x) \left|\frac{(u-k)_\pm}{R}\right|^q \, dx \, .
\end{aligned} \tag{10.7}$$

To estimate the last term, we observe that the inequality

$$a(x) \left|\frac{(u(x)-k)_\pm}{R}\right|^q \leq c \left|\frac{(u(x)-k)_\pm}{R}\right|^p \tag{10.8}$$

holds whenever $x \in B_R$, for some constant c depending on $[a]_{0,\alpha}, \|u\|_{L^\infty(B_R)}$. Indeed, since in this setting we are assuming that $p < n$, then by (1.3) we have $q - \alpha < p + p\alpha/n - \alpha < p$ and therefore by (10.5) and $R \leq 1$ we can estimate

$$a(x) \left|\frac{(u-k)_\pm}{R}\right|^q \leq \frac{c[a]_{0,\alpha} \|u\|_{L^\infty(B_R)}^{q-p} |(u-k)_\pm|^p}{R^{q-\alpha}} \leq c \left|\frac{(u-k)_\pm}{R}\right|^p \, ,$$

so that (10.8) is proved. Now (10.6) follows using the last inequality in (10.7).

We can now obtain two density lemmas which are typical of De Giorgi's theory in the standard case. In the following, with $B_R \Subset \Omega$, we shall denote

$$M(R) := \sup_{B_R} u \, , \quad m(R) := \inf_{B_R} u \quad \text{and} \quad \operatorname{osc}(u, R) := M(R) - m(R) \, .$$

All the balls appearing in the same context will be concentric.

Lemma 10.2. *Let $B_R \Subset \Omega$ be a ball such that $R \leq 1$ (10.5) holds. Moreover, assume that the density condition*

$$\frac{|A(k_0, R)|}{|B_R|} \leq \frac{1}{2} \quad \left(\textit{resp.} \ \frac{|B(k_0, R)|}{|B_R|} \leq \frac{1}{2}\right)$$

holds for $k_0 := (M(R) + m(R))/2$*. Then there exists a positive constant* c*, depending only on* $n, p, q, \nu, L, [a]_{0,\alpha}, \|u\|_{L^\infty(B_R)}$*, such that*

$$\frac{|A(k_m, R/2)|}{|B_{R/2}|} \leq \frac{c}{m^{\frac{n(p-1)}{p(n-1)}}} \quad \left(\textit{resp.}\ \frac{|B(k_m, R/2)|}{|B_{R/2}|} \leq \frac{c}{m^{\frac{n(p-1)}{p(n-1)}}}\right)$$

whenever m *is a positive integer and* $k_m := M(R) - 2^{-m}\operatorname{osc}(u, R)$ *(respect.* $k_m := m(R) + 2^{-m}\operatorname{osc}(u, R)$*).*

Proof. The proof is now identical to the one for the standard case $p = q$ (see for instance [21, Lemma 7.2]). Indeed, taking $t = R/2$ and $s = R$ in (10.6), one obtains the kind of Caccioppoli's inequality which is needed in [21] to prove the lemma.

We finally come to the needed oscillation reduction:

Lemma 10.3. *Let* $B_R \Subset \Omega$ *be a ball such that* $R \leq 1$*; there exists a number* $\theta \in (0, 1)$ *depending only on* $n, p, q, \nu, L, [a]_{0,\alpha}$ *and* $\|u\|_{L^\infty(B_R)}$ *such that if* (10.5) *is verified, then it holds that*

$$\operatorname{osc}(u, R/4) \leq \theta \operatorname{osc}(u, R)\ .$$

Proof. The lemma is a standard consequence (see for instance [13, page 358]) of Lemma 10.2 and of the following fact: For every $\kappa \in (0, 1)$, there exists $\sigma \in (0, 1)$ depending only on $n, p, q, \nu, L, [a]_{0,\alpha}, \|u\|_{L^\infty(B_R)}$ and κ, such that is for some $\varepsilon > 0$ the density condition

$$\frac{|A(M(R) - \varepsilon \operatorname{osc}(u, R), R)|}{|B_R|} \leq \sigma$$

$$\left(\text{resp.}\ \frac{|B(m(R) + \varepsilon \operatorname{osc}(u, R), R)|}{|B_R|} \leq \sigma\right) \tag{10.9}$$

then

$$u(x) \leq M(R) - \kappa\varepsilon \operatorname{osc}(u, R)\,, \quad (\text{resp. } u(x) \geq m(R) + \kappa\varepsilon \operatorname{osc}(u, R)) \tag{10.10}$$

holds for a.e. $x \in B_{R/2}$. For this, we prove the super-level sets part of (10.9) using $(u - k)_+$, the treatment for the other ones appearing in parentheses being completely analogous and using $(u - k)_-$. Once Lemma 10.1 is available, the proof becomes essentially the same as in the case $q = p$; we briefly report the details for completeness. Consider a sequence of nested balls $\{B_{\varrho_i}\}$ concentric to B_R for $i \geq 0$, where $\varrho_i := R(1 + 2^{-i})/2$ and define also $\bar{\varrho}_i := (\varrho_i + \varrho_{i+1})/2$. Accordingly, we use related cut-off functions $\eta_i \in C_0^\infty(B_{\bar{\varrho}_i})$ such that $\eta_i \equiv 1$ on $B_{\varrho_{i+1}}$ and $|D\eta_i| \leq c2^i/R$. We define the levels

$$k_i = M(R) - \kappa\varepsilon \operatorname{osc}(u, R) - (1 - \kappa)\varepsilon \operatorname{osc}(u, R)/2^i$$

so that

$$\left[\frac{(1-\kappa)\varepsilon \operatorname{osc}(u,R)}{2^{i+1}}\right]^p |A(k_{i+1},\varrho_{i+1})| = (k_{i+1}-k_i)^p|A(k_{i+1},\varrho_{i+1})| \leq c\int_{A(k_i,\bar{\varrho}_i)} [\eta_i(u-k_i)_+]^p\,dx\,. \tag{10.11}$$

By Sobolev embedding theorem - recall here it is $p < n/(1+\delta_g) < n$ - and using (10.6) (similarly to the treatment of (10.3)), we then get, for $p^* = np/(n-p)$,

$$\begin{aligned}\int_{A(k_i,\bar{\varrho}_i)} [\eta_i(u-k_i)_+]^p\,dx &\leq \left(\int_{B_{\bar{\varrho}_i}} [\eta_i(u-k_i)_+]^{p^*}\,dx\right)^{p/p^*} |A(k_i,\bar{\varrho}_i)|^{p/n}\\ &\leq c\int_{B_{\bar{\varrho}_i}} |D[\eta_i(u-k_i)]_+|^p\,dx\,|A(k_i,\bar{\varrho}_i)|^{p/n}\\ &\leq c2^{qi}\int_{A(k_i,\varrho_i)} \left|\frac{u-k_i}{R}\right|^p\,dx\,|A(k_i,\varrho_i)|^{p/n}\\ &\leq c2^{qi}\,[\varepsilon \operatorname{osc}(u,R)]^p\,\frac{|A(k_i,\varrho_i)|^{1+p/n}}{|B_{\varrho_i}|^{p/n}}\,.\end{aligned}$$

Using this last estimate together with (10.11), dividing both sides of the resulting inequality by $|B_{\varrho_i}|$ and setting $\Xi_i := |A(k_i,\varrho_i)|/|B_{\varrho_i}|$, we obtain the recursive estimate

$$\Xi_{i+1} \leq \frac{c_f 4^{qi}}{(1-\kappa)^p}\Xi_i^{1+p/n}\,,$$

for a constant c_f depending only on $n,p,q,\nu,L,\alpha,[a]_{0,\alpha}$ and $\|u\|_{L^\infty(B_R)}$, and for every integer $i \geq 0$. We can now conclude as in standard De Giorgi's theory: in order to prove (10.10) we need to check that $\Xi_i \to 0$. In turn this follows by choosing $\Xi_0 < \sigma$, for some small σ depending on c_f and κ, and therefore ultimately again on $n,p,q,\nu,L,[a]_{0,\alpha},\|u\|_{L^\infty(B_R)}$ and κ. Since $\Xi_0 \leq \sigma$ is precisely (10.9), then (10.10) follows by choosing the constant σ small enough.

Step 4: Separation of phases and conclusion. Let us fix $\Omega' \Subset \Omega$ as in the statement of Theorem 1.1. We consider a ball $B_{R_0} \subset \Omega'$ with $R_0 \leq 1$ and the conditions

$$\sup_{x\in B_{R_0/4^k}} a(x) \leq 4[a]_{0,\alpha}\left(R_0/4^k\right)^\alpha\,. \tag{10.12}$$

Accordingly, the exit time index is defined by

$$m := \min\left\{k\in\mathbb{N}\cup\{\infty\}\,:\,\text{(10.12) fails}\right\}\,. \tag{10.13}$$

We can apply Lemma 10.3 m (possibly $m = 0$) times to obtain, by induction, that

$$\operatorname{osc}(u,R_0/4^k) \leq \theta^k \operatorname{osc}(u,R_0) \qquad \text{for every } k\in\{0,\ldots,m\}\,, \tag{10.14}$$

where $\theta\in(0,1)$ depends only on $n,p,q,\nu,L,[a]_{0,\alpha},\|u\|_{L^\infty(\Omega')}$. In the case $m < \infty$, condition (10.12) fails at a point $x_0 \in B_{R_0/4^m}$, i.e.

$$a(x_0) > 4[a]_{0,\alpha}\left(R_0/4^m\right)^\alpha\,. \tag{10.15}$$

We now want to prove that u is a Q-minimiser - in $B_{R_0/4^m}$ - of the functional $\mathcal{P}_0$ defined in (1.14) for $Q = 4L/\nu$ and for the choice $a_0 = a(x_0)$. This means that

$$|Du|^p + a(x_0)|Du|^q \in L^1(B_{R_0/4^m}) \tag{10.16}$$

and that $\mathcal{P}_0(u,K) \leq (4L/\nu)\mathcal{P}_0(w,K)$ holds whenever $w \in W^{1,1}(B_{R_0/4^m})$ with $|Dw|^p + a(x_0)|Dw|^q \in L^1(B_{R_0/4^m})$, and $K \subset B_{R_0/4^m}$ is a compact set such that $\operatorname{supp}(u-w) \subset K$. For this, note that by (10.15), for any $x \in B_{R_0/4^m}$ we have

$$2a(x) \geq a(x_0) + 4[a]_{0,\alpha}\,(R_0/4^m)^\alpha - 2|a(x)-a(x_0)| \geq a(x_0)$$

and

$$2a(x_0) \geq a(x) + 4[a]_{0,\alpha}\,(R_0/4^m)^\alpha - |a(x)-a(x_0)| \geq a(x)\,,$$

so that (10.16) follows immediately. The minimality of u and (1.5) then yield

$$\begin{aligned}
\mathcal{P}_0(u,K) &\leq \int_K (|Du|^p + 2a(x)|Du|^q)\,dx \\
&\leq \frac{2}{\nu}\mathcal{F}(u,K) \\
&\leq \frac{2}{\nu}\mathcal{F}(w,K) \\
&\leq \frac{2L}{\nu}\int_K (|Dw|^p + a(x)|Dw|^q)\,dx \\
&\leq \frac{2L}{\nu}\int_K (|Dw|^p + 2a(x_0)|Dw|^q)\,dx \\
&\leq \frac{4L}{\nu}\mathcal{P}_0(w,K)\,.
\end{aligned}$$

We can therefore apply Theorem 11.3 below and find $\tau_0 \in (0,1/4)$, depending only on $n,p,q,\nu,L,[a]_{0,\alpha},\|u\|_{L^\infty(\Omega')}$, such that

$$\operatorname{osc}(u,\tau_0^h R_0/4^m) \leq \theta^h \operatorname{osc}(u,R_0/4^m)$$

holds for every $h \in \mathbb{N}$. Using this last inequality together with (10.14) yields

$$\operatorname{osc}(u,\tau_0^h R_0) \leq \theta^h \operatorname{osc}(u,R_0) \qquad \text{for every } h \in \mathbb{N}\,.$$

This last inequality also holds in the case $m = \infty$, directly with $\tau_0 = 1/4$, by (10.14). The Hölder continuity (1.9) of u now follows with $\beta = \log\theta/\log\tau_0$ via a standard covering argument, since all the previous considerations are independent of the starting ball B_{R_0} as long as this is contained in Ω' and since both τ_0 and θ depend only on $n,p,q,\nu,L,[a]_{0,\alpha},\|u\|_{L^\infty(\Omega')}$.

11. Estimates for frozen functionals

Here we collect the regularity results for minima of frozen functionals used in the previous sections; they all come from the seminal work of Lieberman [27] as long as the scalar case is concerned. The results in the vectorial case come instead from [14]. We start by higher regularity, thereby considering functionals of the type

$$\mathcal{G}_0(w,\Omega) := \int_\Omega (f(Dw) + a_0 g(Dw))\,dx \tag{11.1}$$

where $f, g\colon \mathbb{R}^n \to \mathbb{R}$ satisfy assumptions (1.11) and $a_0 \geq 0$ is a constant.

Theorem 11.1 ([27]). *Let $v \in W^{1,p}(\Omega)$ be a local minimiser of the functional $\mathcal{G}_0$ defined in* (11.1)*. There exists $\tilde{\alpha} \in (0,1)$, depending only on n, p, q, ν, L, but otherwise independent of a_0 and of the minimiser v, such that $Dv \in C^{0,\tilde{\alpha}}_{\rm loc}(\Omega)$. Moreover, whenever $B_R \subset \Omega$ the following inequalities hold:*

$$\sup_{B_{R/2}} (|Dv|^p + a_0|Dv|^q) \leq c \fint_{B_R} (|Dv|^p + a_0|Dv|^q)\,dx\,, \tag{11.2}$$

and, for every $0 < \varrho \leq R$

$$\fint_{B_\varrho} |Dv - (Dv)_{B_\varrho}|^p\,dx \leq c\left(\frac{\varrho}{R}\right)^{p\tilde{\alpha}} \fint_{B_R} (|Dv|^p + a_0|Dv|^q)\,dx\,, \tag{11.3}$$

where again c depends only on n, p, q, ν, L.

Proof. We use the results of [27] when - with the notation adopted there - $g(t) = pt^{p-1} + a_0 q t^{q-1}$, and, accordingly $G(t) = t^p + a_0 t^q$. The equation div $\bar{A}(Dv) = 0$ considered in [27] is determined by $\bar{A}(z) = \partial f(z) + a_0 \partial g(z)$. The main point here is that the crucial uniform ellipticity assumption considered in [27, (4)], that is

$$0 < \delta \leq \frac{g'(t)t}{g(t)} \leq g_0\ ,$$

is satisfied for constants δ, g_0 which are independent of $a_0 \geq 0$. Estimate (11.2) is then [27, (5.3a)] with this particular choice of the functions and of the equation. Estimate (11.3) becomes instead implicit in [27, (5.3b)] and in turn implies that $Dv \in C^{0,\tilde{\alpha}}_{\rm loc}(\Omega)$ via a standard characterisation of Hölder continuity due to Campanato and Meyers. We notice that the estimates proposed in [27] are given under the form of a priori estimates for bounded solutions of more regular equations; anyway a standard approximation augment (still explained in [27] and that reduces to a very simple one in our special case) allows to state the results in the general case we are considering here.

The vectorial version of the previous result is instead contained in the following:

Theorem 11.2 ([27]). *Let $v \in W^{1,p}(\Omega;\mathbb{R}^N)$ be a vector-valued local minimiser of the functional*

$$W^{1,p}(\Omega;\mathbb{R}^N) \ni w \mapsto \int_\Omega (|Dw|^p + a_0|Dw|^q)\ dx$$

where $\Omega \subset \mathbb{R}^n$ is a bounded open domain and $N \geq 1$ and $a_0 \geq 0$. There exists $\tilde{\alpha} \in (0,1)$, depending only on n, N, p, q, ν, L, but otherwise independent of a_0 and of the minimiser v, such that $Dv \in C^{0,\tilde{\alpha}}_{\mathrm{loc}}(\Omega)$. Moreover, whenever $B_R \subset \Omega$, the inequalities in (11.2)-(11.3) *continue to hold in this case.*

Proof. The results follow from [14] upon taking, in the notation of that paper, $\varphi(t) = t^p + a_0 t^q$. In particular, estimate (11.2) is a consequence of [14, Lemma 2.8], while estimate (11.3) is instead an easy corollary of the one contained in [14, Theorem 6.4].

The next result deals with the Hölder continuity of so called quasi-minima (see [21]) of the functional $\mathcal{P}_0$ defined in (1.14). In this setting, a function $v \in W^{1,1}(\Omega)$ is a Q-minimiser of $\mathcal{P}_0$ for $Q \geq 1$, iff $|Dv|^p + a_0|Dv|^q \in L^1(\Omega)$ and the quasi-minimality condition

$$\mathcal{P}_0(v,K) \leq Q\mathcal{P}_0(w,K)$$

holds for every $w \in W^{1,1}(\Omega)$ and compact set $K \subset \Omega$, such that $\operatorname{supp}(v-w) \subset K$ and $|Dw|^p + a_0|Dw|^q \in L^1(\Omega)$. We then have the next theorem, which follows from [27, Section 6], for the choice $G(t) = t^p + a_0 t^q$ (see also the main result of [39]):

Theorem 11.3 ([27]). *Let $v \in W^{1,p}(\Omega)$ be a Q-minimiser of the functional $\mathcal{P}_0$ defined in* (1.14)*. Then there exists a positive constant c and an exponent $\beta_0 \in (0,1)$, both depending on n, p, q, Q, but otherwise independent of a_0 and of the Q-minimiser v, such that*

$$\operatorname{osc}(v, B_\varrho) \leq c\left(\frac{\varrho}{R}\right)^{\beta_0} \operatorname{osc}(v, B_R)$$

holds whenever $B_\varrho \subset B_R \subset \Omega$ are concentric balls.

Acknowledgments. The work for this paper started while both the authors were visiting the Department of Mathematics of the University of Texas at Austin. They would like to thank Luis Caffarelli and Alessio Figalli for the invitation and the kind and warm hospitality. This work is partially supported by INDAM-GNAMPA.

References

1. Acerbi E. & Bouchitté G. & Fonseca I.: Relaxation of convex functionals: the gap problem. *Ann. Inst. H. Poincaré Anal. Non Linèaire* 20 (2003), 359–390.
2. Acerbi E. & Fusco N.: Regularity for minimizers of nonquadratic functionals: the case $1 < p < 2$. *J. Math. Anal. Appl.* 140 (1989), 115–135

3. Acerbi E. & Fusco N.: A transmission problem in the calculus of variations. *Calc. Var. & PDE* 2 (1994), 1–16.
4. Acerbi E. & Mingione G.: Regularity results for a class of functionals with non-standard growth. *Arch. Rat. Mech. Anal.* 156 (2001), 121–140.
5. Baroni P. & Colombo M. & Mingione G.: Flow of double phase functionals. *To appear*.
6. Bildhauer M. & Fuchs M.: $C^{1,\alpha}$-solutions to non-autonomous anisotropic variational problems. *Calc. Var. & PDE* 24 (2005), 309–340.
7. Breit D.: New regularity theorems for non-autonomous variational integrals with (p, q)-growth. *Calc. Var. & PDE* 44 (2012), 101–129.
8. Carozza M. & Kristensen J. & Passarelli di Napoli A.: Higher differentiability of minimizers of convex variational integrals. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 28 (2011), 395–411.
9. Carozza M. & Kristensen J. & Passarelli di Napoli A.: Regularity of minimizers of autonomous convex variational integrals. *Ann. Scu. Sup. Pisa. Cl. Sci. (V)*, doi: 10.2422/2036-2145.201208_005.
10. Colombo M. & Mingione G.: Bounded minimisers of double phase variational integrals. *Preprint 2014*.
11. Coscia A. & Mingione G.: Hölder continuity of the gradient of $p(x)$-harmonic mappings. *C. R. Acad. Sci. Paris Sér. I Math.* 328 (1999), 363–368.
12. De Giorgi E.: Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari. *Mem. Accad. Sci. Torino Cl. Sci. Fis. Mat. Nat. (III)* 125 3 (1957), 25–43.
13. DiBenedetto E.: *Partial differential equations*. Second edition. Cornerstones. Birkhäuser, Inc., Boston, MA, 2010.
14. Diening L. & Verde A. & Stroffolini B.: Everywhere regularity of functionals with φ-growth. *manuscripta math.* 129 (2009), 449–481.
15. Esposito L. & Leonetti F. & Mingione G.: Sharp regularity for functionals with (p, q) growth. *J. Diff. Equ.* 204 (2004), 5–55.
16. Fonseca I. & Malý J.: Relaxation of multiple integrals below the growth exponent. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 14 (1997), 309–338.
17. Fonseca I. & Malý J. & Mingione G.: Scalar minimizers with fractal singular sets. *Arch. Rat. Mech. Anal.* 172 (2004), 295–307.
18. Foss M.: Global regularity for almost minimizers of nonconvex variational problems. *Ann. Mat. Pura e Appl. (IV)* 187 (2008), 263-231.
19. Giaquinta M. & Giusti E.: On the regularity of the minima of variational integrals. *Acta Math.* 148 (1982), 31–46.
20. Gilbarg D. & Trudinger N.S.: *Elliptic partial differential equations of second order.* Grundlehren der Mathematischen Wissenschaften, Vol. 224. Springer-Verlag, Berlin-New York, 1977.
21. Giusti E.: *Direct methods in the calculus of variations.* World Scientific Publishing Co., Inc., River Edge, NJ, 2003.
22. Hedberg L.I.: On certain convolution inequalities. *Proc. Amer. Math. Soc.* 36 (1972), 505–510.
23. Kristensen J.: Lower semicontinuity in Sobolev spaces below the growth exponent of the integrand. *Proc. Roy. Soc. Edinburgh Sect. A* 127 (1997), 797–817.
24. Kristensen J. & Mingione G.: The singular set of minima of integral functionals. *Arch. Ration. Mech. Anal.* 180 (2006), 331–398.
25. Kuusi T. & Mingione G.: New perturbation methods for nonlinear parabolic problems. *J. Math. Pures Appl. (IX)* 98 (2012), 390–427.
26. Ladyzhenskaya O.A. & Ural'tseva N.N.: *Linear and quasilinear elliptic equations. Second edition* (in Russian) "Nauka", Moscow 1973.
27. Lieberman G.M.: The natural generalization of the natural conditions of Ladyzhenskaya and Ural'tseva for elliptic equations. *Comm. PDE* 16 (1991), 311–361.
28. Lieberman G.M.: Gradient estimates for a class of elliptic systems. *Ann. Mat. Pura Appl. (IV)* 164 (1993), 103–120.
29. 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.

30. Lieberman G.M.: Gradient estimates for anisotropic elliptic equations. *Adv. Diff. Equ.* 10 (2005), 767–182.

31. Lindqvist P.: Notes on the p-Laplace equation. *Univ. Jyväskylä, Report* 102, (2006).

32. Manfredi J.J.: Regularity for minima of functionals with p-growth. *J. Diff. Equ.* 76 (1988), 203–212.

33. 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.

34. Marcellini P.: On the definition and the lower semicontinuity of certain quasiconvex integrals. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 3 (1986), 391–409.

35. 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.

36. Marcellini P.: Regularity and existence of solutions of elliptic equations with p, q-growth conditions. *J. Diff. Equ.* 90 (1991), 1–30.

37. Marcellini P.: Regularity for elliptic equations with general growth conditions *J. Diff. Equ.* 105 (1993), 296–333.

38. 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.

39. Mascolo E. & Papi G.: Harnack inequality for minimizers of integral functionals with general growth conditions. *Nonlinear Differential Equations Appl.* 3 (1996), 231–244.

40. Mingione G.: Bounds for the singular set of solutions to non linear elliptic systems. *Calc. Var. & PDE* 18 (2003), 373–400.

41. Mingione G.: Regularity of minima: an invitation to the Dark side of the Calculus of Variations. *Appl. Math.* 51 (2006), 355-425.

42. Mingione G. & Mucci D.: Integral functionals and the gap problem: sharp bounds for relaxation and energy concentration. *SIAM J. Math. Anal.* 36 (2005), 1540–1579.

43. Schmidt T.: Regularity of minimizers of $W^{1,p}$-quasiconvex variational integrals with (p, q)-growth. *Calc. Var. & PDE* 32 (2008), 1–24.

44. Schmidt T.: Regularity of relaxed minimizers of quasiconvex variational integrals with (p, q)-growth. *Arch. Rat. Mech. Anal.* 193 (2009), 311–337.

45. Uhlenbeck K.: Regularity for a class of non-linear elliptic systems. *Acta Math.* 138 (1977), 219–240.

46. Ural'tseva N.N.: Degenerate quasilinear elliptic systems. *Zap. Na. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)* 7 (1968), 184–222.

47. Ural'tseva N.N. & Urdaletova A.B.: The boundedness of the gradients of generalized solutions of degenerate quasilinear nonuniformly elliptic equations. *Vestnik Leningrad Univ. Math. 19* (1983) (russian) english. tran.: 16 (1984), 263-270.

48. Zhikov V.V.: Problems of convergence, duality, and averaging for a class of functionals of the calculus of variations. *Dokl. Akad. Nauk SSSR* 267 (1982), 524–528.

49. Zhikov V.V.: Averaging of functionals of the calculus of variations and elasticity theory. *Izv. Akad. Nauk SSSR Ser. Mat.* 50 (1986), 675–710.

50. Zhikov V.V.: Lavrentiev phenomenon and homogenization for some variational problems. *C. R. Acad. Sci. Paris Sér. I Math.* 316 (1993), 435–439.

51. Zhikov V.V.: On Lavrentiev's Phenomenon. *Russian J. Math. Phys.* 3 (1995), 249–269.

52. Zhikov V.V.: On some variational problems. *Russian J. Math. Phys.* 5 (1997), 105–116.

53. Zhikov V.V. & Kozlov S. M. & Oleinik O. A.: *Homogenization of differential operators and integral functionals*. Springer-Verlag, Berlin, 1994.

Scuola Normale Superiore di Pisa
p.za dei Cavalieri, 7
I-56126 Pisa, Italy
e-mail: maria.colombo@sns.it

and

Dipartimento di Matematica e Informatica
Università di Parma
Parco Area delle Scienze 53/a, I-43124, Parma
email:giuseppe.mingione@math.unipr.it