Calculus of Variations

CrossMark

On the Lipschitz character of orthotropic p-harmonic functions

P. Bousquet[1] · L. Brasco[2,3] · C. Leone[4] · A. Verde[4]

Abstract We prove that local weak solutions of the orthotropic p-harmonic equation are locally Lipschitz, for every $p \geq 2$ and in every dimension. More generally, the result holds true for more degenerate equations with orthotropic structure, with right-hand sides in suitable Sobolev spaces.

Mathematics Subject Classification 35J70 · 35B65 · 49K20

Contents

Communicated by L. Ambrosio.

✉ C. Leone
  chiara.leone@unina.it

  P. Bousquet
  pierre.bousquet@math.univ-toulouse.fr

  L. Brasco
  lorenzo.brasco@unife.it

  A. Verde
  anna.verde@unina.it

1 Institut de Mathématiques de Toulouse, CNRS UMR 5219, Université de Toulouse, 31062 Toulouse Cedex 9, France

2 Dipartimento di Matematica e Informatica, Università degli Studi di Ferrara, Via Machiavelli 35, 44121 Ferrara, Italy

3 CNRS, Centrale Marseille, I2M, UMR 7373, 39 Rue Frédéric Joliot Curie, Aix-Marseille Université, 13453 Marseille, France

4 Dipartimento di Matematica "R. Caccioppoli", Università degli Studi di Napoli "Federico II", Via Cinthia, Complesso Universitario di Monte S. Angelo, 80126 Naples, Italy

1 Introduction

1.1 The problem

In this paper, we pursue the study of the regularity of local minimizers of degenerate *functionals with orthotropic structure*, that we already considered in [1–4]. More precisely, for $p \geq 2$, we consider local minimizers of the functional

$$\mathfrak{F}_0(u, \Omega') = \sum_{i=1}^{N} \frac{1}{p} \int_{\Omega'} |u_{x_i}|^p \, dx, \qquad \Omega' \Subset \Omega, \ u \in W^{1,p}_{\mathrm{loc}}(\Omega'), \qquad (1.1)$$

and more generally of the functional

$$\mathfrak{F}_\delta(u, \Omega') = \sum_{i=1}^{N} \frac{1}{p} \int_{\Omega'} (|u_{x_i}| - \delta_i)_+^p \, dx + \int_{\Omega'} f \, u \, dx, \qquad \Omega' \Subset \Omega, \ u \in W^{1,p}_{\mathrm{loc}}(\Omega').$$

Here, $\Omega \subset \mathbb{R}^N$ is an open set, $N \geq 2$, and $\delta_1, \ldots, \delta_N$ are nonnegative numbers.

A local minimizer u of the functional $\mathfrak{F}_0$ defined in (1.1) is a local weak solution of the *orthotropic p-Laplace* equation

$$\sum_{i=1}^{N} \left(|u_{x_i}|^{p-2} u_{x_i} \right)_{x_i} = 0. \qquad (1.2)$$

For $p = 2$, this is just the Laplace equation, which is uniformly elliptic. For $p > 2$, this looks quite similar to the usual p-Laplace equation

$$\sum_{i=1}^{N} \left(|\nabla u|^{p-2} u_{x_i} \right)_{x_i} = 0,$$

whose local weak solutions are local minimizers of the functional

$$\mathfrak{I}(u, \Omega') = \frac{1}{p} \int_{\Omega'} |\nabla u|^p \, dx, \qquad \Omega' \Subset \Omega, \ u \in W^{1,p}_{\mathrm{loc}}(\Omega'). \qquad (1.3)$$

However, as explained in [1,2], equation (1.2) is much more degenerate. Consequently, as for the regularity of ∇u (i.e. boundedness and continuity), the two equations are dramatically different.

In order to understand this discrepancy between the p-Laplacian and its orthotropic version, let us observe that the map $\xi \mapsto |\xi|^p$ occuring in the definition (1.3) of $\mathfrak{I}$ degenerates

only at the origin, in the sense that its Hessian is positive definite on $\mathbb{R}^N \setminus \{0\}$. On the contrary, the definition of the orthotropic functional $\mathfrak{F}_0$ in (1.1) is related to the map $\xi \mapsto \sum_{i=1}^{N} |\xi_i|^p$, which *degenerates on an unbounded* set, namely the N hyperplanes orthogonal to the coordinate axes of $\mathbb{R}^N$.

The situation is even worse when

$$\max\{\delta_i \,:\, i = 1, \dots, N\} > 0, \tag{1.4}$$

for the lack of ellipticity of the *degenerate p-orthotropic functional* arises on the larger set

$$\bigcup_{i=1}^{N} \{\xi \in \mathbb{R}^N : |\xi_i| \leq \delta_i\}.$$

As a matter of fact, the regularity theory for these very degenerate functionals is far less understood than the corresponding theory for the standard case (1.3) and its variants.

Under suitable integrability conditions on the function f, we can use the classical theory for functionals with p-growth and ensure that the local minimizers of $\mathfrak{F}_\delta$ are locally bounded and Hölder continuous, see for example [11, Theorems 7.5 & 7.6]. This theory also assures that the gradients of local minimizers lie in $L_{\mathrm{loc}}^r(\Omega)$ for some $r > p$, see [11, Theorem 6.7].

We also point out that for $f \in L_{\mathrm{loc}}^\infty(\Omega)$, local minimizers of $\mathfrak{F}_\delta$ are contained in $W_{\mathrm{loc}}^{1,q}(\Omega)$, for every $q < +\infty$ (see [3, Main Theorem]).

1.2 Main result

In this paper, we establish the optimal regularity expected for the minimizers of $\mathfrak{F}_\delta$, namely the *Lipschitz regularity*.[1] More precisely, we establish the following result.

Theorem 1.1 *Let* $p \geq 2$, $f \in W_{\mathrm{loc}}^{1,h}(\Omega)$ *for some* $h > N/2$ *and let* $U \in W_{\mathrm{loc}}^{1,p}(\Omega)$ *be a local minimizer of the functional* $\mathfrak{F}_\delta$. *Then* U *is locally Lipschitz in* Ω.

Moreover, in the case $\delta_1 = \cdots = \delta_N = 0$, *we have the following local scaling invariant estimate: for every ball* $B_{2R_0} \Subset \Omega$, *it holds*

$$\|\nabla U\|_{L^\infty(B_{R_0/2})} \leq C \left(\fint_{B_{R_0}} |\nabla U|^p \, dx \right)^{\frac{1}{p}} + C \left[R_0^2 \left(\fint_{B_{R_0}} |\nabla f|^h \, dx \right)^{\frac{1}{h}} \right]^{\frac{1}{p-1}}, \tag{1.5}$$

for some $C = C(N, p, h) > 1$.

Remark 1.2 (Comparison with previous results) This result unifies and substantially extends the results on the orthotropic functional $\mathfrak{F}_\delta$ contained in [2], where it has been established that the local minimizers of $\mathfrak{F}_\delta$ are locally Lipschitz, provided that:

- $p \geq 2$, $N = 2$ and $f \in W_{\mathrm{loc}}^{1,p'}(\Omega)$, see [2, Theorem A];
- $p \geq 4$, $N \geq 2$ and $f \in W_{\mathrm{loc}}^{1,\infty}(\Omega)$, see [2, Theorem B].

The second result was based on the so-called *Bernstein's technique*, see for example [12, Proposition 2.19]. This technique had already been exploited in the pioneering paper [17] by Uralt'seva and Urdaletova, for a class of functionals which contains the orthotropic functional $\mathfrak{F}_0$ defined in (1.1), but not its more degenerate version $\mathfrak{F}_\delta$. Namely, the result of [17] does not cover the case when condition (1.4) is in force.

---

[1] Observe that when $f \equiv 0$, any Lipschitz function u with $|\nabla u| \leq \min\{\delta_i \,:\, i = 1, \dots, N\}$ is a local minimizer of $\mathfrak{F}_\delta$. Thus in general Lipschitz continuity is the best regularity one can hope for.

Still for the case $\delta_1 = \cdots = \delta_N = 0$, an entirely different approach relying on viscosity methods has been developed in [6]. To our knowledge, both methods are limited to (at least) *bounded* lower order terms f.

On the contrary, [2, Theorem A] can be considered as the true ancestor to Theorem 1.1 above. Indeed, they both follow the *Moser's iteration technique*, originally introduced in [16] to establish regularity for uniformly elliptic problems. However, going beyond the two-dimensional setting requires new ideas, that we will explain in Sect. 1.3 below.

In contrast to the partial results of [2, Theorems A & B], the proof of Theorem 1.1 does not depend on the dimension and does not need any additional restriction on p, apart from $p \geq 2$. It allows unbounded lower order terms, even if the condition $f \in W^{1,h}_{loc}(\Omega)$ for some $h > N/2$ is certainly not sharp. On this point, it is useful to observe that by Sobolev's embedding we have[2]

$$W^{1,h} \hookrightarrow L^{h^*},$$

with h^* larger than N and as close to N as desired, provided h is close to $N/2$. This means that, in terms of summability, our assumption on f amounts to $f \in L^q_{loc}(\Omega)$ for some $q > N$. This is exactly the sharp expected condition on f for the local minimizers to be locally Lipschitz, at least if one nurtures the (optimistic) hope that the regularity for the orthotropic p-Laplacian agrees with that for the standard p-Laplacian.[3]

Our strategy to prove Theorem 1.1 relies on energy methods and integral estimates, and more precisely on *ad hoc* Caccioppoli-type inequalities. This only requires growth assumptions on the Lagrangian and its derivatives and can be adapted to a large class of functionals. For instance, we briefly explain in "Appendix" how to adapt our poof to the case of *nonlinear lower order terms*, i.e. when $f u$ is replaced by a term of the form $G(x, u)$.

Remark 1.3 We collect in this remark some interesting open issues:

(1) one word about the assumption $p \geq 2$: as explained in [1,2], when $\delta_1 = \cdots = \delta_N = 0$, the subquadratic case $1 < p < 2$ is simpler in a sense. In this case, the desired Lipschitz regularity can be inferred from [8, Theorem 2.2] (see also [9, Theorem 2.7]). However, the more degenerate case (1.4) is open;

(2) in [1, Main Theorem], local minimizers were proven to be C^1, in the two-dimensional case, for $1 < p < \infty$ and when $\delta_1 = \cdots = \delta_N = 0$. We also refer to the very recent paper [14], where a modulus of continuity for the gradient of local mimizers is exhibited. We do not know whether such a result still holds in higher dimensions;

(3) in [4, Theorem 1.4], local Lipschitz regularity is established in the two-dimensional case for an orthotropic functional, with anisotropic growth conditions; that is, for the

---

[2] We recall that

$$h^* = \begin{cases} N h/(N - h), & \text{if } h < N, \\ \text{any } q < +\infty, & \text{if } h = N, \\ +\infty, & \text{if } h > N. \end{cases}$$

[3] In the case of the standard p-Laplacian, the sharp assumption to have Lipschitz regularity is that f belongs to the Lorentz space $L^{N,1}_{loc}$. This sharp condition has been first detected by Duzaar and Mingione in [7, Theorem 1.2], see also [13, Corollary 1.6] for a more general and refined result. This sharp result is obtained by using potential estimates techniques. We recall that $L^q_{loc} \subset L^{N,1}_{loc}$ for every $q > N$ and under this slightly stronger assumption on f, Lipschitz regularity for the p-Laplacian can be proved by more standard techniques based on Moser's iteration, see for example [5].

functional

$$\sum_{i=1}^{2} \frac{1}{p_i} \int (|u_{x_i}| - \delta_i)_+^{p_i}\, dx + \int f\, u\, dx, \qquad \text{with } 2 \le p_1 \le p_2.$$

For such a functional, Lipschitz regularity is open in higher dimensions, even for the case $\delta_1 = \cdots = \delta_N = 0$, i.e. for the functional

$$\sum_{i=1}^{2} \frac{1}{p_i} \int |u_{x_i}|^{p_i}\, dx + \int f\, u\, dx, \qquad \text{with } 2 \le p_1 \le p_2 \le \cdots \le p_N.$$

We point out that in this case, Lipschitz regularity in every dimension has been obtained in [17, Theorem 1] for *bounded* local minimizers, under the additional restrictions

$$p_1 \ge 4 \qquad \text{and} \qquad p_N < 2\, p_1.$$

Though these restrictions are not optimal, we recall that regularity can not be expected when p_N and p_1 are too far apart, due to the well-known counterexamples by Giaquinta [10] and Marcellini [15].

1.3 Technical novelties of the proof

Our main result is obtained by considering a regularized problem having a unique smooth solution converging to our local minimizer, and proving a local Lipschitz estimate independent of the regularization parameter.

At first sight, the strategy to prove such an estimate may seem quite standard:

(a) differentiate equation (1.2);
(b) obtain Caccioppoli-type inequalities for convex powers of the components u_{x_k} of the gradient;
(c) derive an iterative scheme of reverse Hölder's inequalities;
(d) iterate and obtain the desired local L^∞ estimate on ∇u.

However, steps **(b)** and **(c)** are quite involved, due to the degeneracy of our equation. This makes their concrete realization fairly intricate. Thus in order to smoothly introduce the reader to the proof, we prefer to spend some words.

We point out that our proof is not just a mere adaption of techniques used for the p-Laplace equation. Moreover, it does not even rely on the ideas developed in [2] for the two-dimensional case. In a nutshell, we need new ideas to deal with our functional in full generality.

In order to obtain "good" Caccioppoli-type inequalities for the gradient, we exploit an idea introduced in nuce in [1]. This consists in differentiating (1.2) in the direction x_j and then testing the resulting equation with a test function of the form[4]

$$u_{x_j} |u_{x_j}|^{2s-2} |u_{x_k}|^{2m},$$

with $1 \le s \le m$. This leads to an estimate of the type (see Proposition 4.1)

$$\sum_{i=1}^{N} \int |u_{x_i}|^{p-2}\, u_{x_i x_j}^2\, |u_{x_j}|^{2\,s-2}\, |u_{x_k}|^{2m}\, dx$$

---

[4] This test function is not really admissible, since it is not compactly supported. Actually, to make it admissible, we have to multiply it by a cut-off function. However, this gives unessential modifications and we prefer to avoid it in order to neatly present the idea of the proof.

$$\leq C \sum_{i=1}^{N} \int |u_{x_i}|^{p-2} \left(|u_{x_j}|^{2s+2m} + |u_{x_k}|^{2s+2m} \right) dx$$

$$+ \sum_{i=1}^{N} \int |u_{x_i}|^{p-2} u_{x_i x_j}^2 |u_{x_j}|^{4s-2} |u_{x_k}|^{2m-2s} dx. \tag{1.6}$$

Then the idea is the following: let us suppose that we are interested in improving the summability of the component u_{x_k}. Ideally, we would like to take $s = 1$ in (1.6), since in this case the left-hand side boils down to

$$\sum_{i=1}^{N} \int |u_{x_i}|^{p-2} u_{x_i x_j}^2 |u_{x_k}|^{2m} dx \geq \int |u_{x_k}|^{p-2} u_{x_k x_j}^2 |u_{x_k}|^{2m} dx$$

$$\simeq \int \left| \left(|u_{x_k}|^{\frac{p}{2}+m} \right)_{x_j} \right|^2 dx.$$

If we now sum over $j = 1, \ldots, N$, this would give a control on the $W^{1,2}$ norms of convex powers of u_{x_k}. *But there is a drawback here*: indeed, this $W^{1,2}$ norm is estimated still in terms of the Hessian of u, which is contained in the right-hand side of (1.6). Observe that (1.6) has the following form

$$\mathcal{I}(s-1, m) \leq C \sum_{i=1}^{N} \int |u_{x_i}|^{p-2} \left(|u_{x_j}|^{2s+2m} + |u_{x_k}|^{2s+2m} \right) dx$$

$$+ \mathcal{I}(2s-1, m-s), \tag{1.7}$$

where

$$\mathcal{I}(s, m) = \sum_{i=1}^{N} \int |u_{x_i}|^{p-2} u_{x_i x_j}^2 |u_{x_j}|^{2s} |u_{x_k}|^{2m} dx.$$

This suggests to perform a finite iteration of (1.7) for $s = s_i$ and $m = m_i$ such that

$$\begin{cases} 2s_i - 1 = s_{i+1} - 1 \\ s_0 \quad\quad = 1 \end{cases} \quad \text{and} \quad m_i - s_i = m_{i+1}, \quad \text{for } i = 0, \ldots, \ell.$$

The number ℓ is chosen so that we stop the iteration when we reach $m_\ell = 0$. The above conditions imply that for every $i = 0, \ldots, \ell$, we have

$$m_i + s_i = m_0 + s_0 = 2^\ell.$$

In this way, after a finite number of steps (comparable to ℓ), the coupling between u_{x_k} and the Hessian of u contained in the term $\mathcal{I}$ will disappear from the right-hand side. In other words, we will end up with an estimate of the type

$$\int \left| \nabla |u_{x_k}|^{2^\ell + \frac{p-2}{2}} \right|^2 dx \leq C \sum_{i,j=1}^{N} \int |u_{x_i}|^{p-2} \left(|u_{x_j}|^{2^{\ell+1}} + |u_{x_k}|^{2^{\ell+1}} \right) dx$$

$$+ \sum_{i=1}^{N} \int |u_{x_i}|^{p-2} u_{x_i x_j}^2 |u_{x_j}|^{2(2^\ell - 1)} dx. \tag{1.8}$$

Observe that we still have the Hessian of u in the right-hand side (this is the second term), but this time it is harmless. It is sufficient to use the standard Caccioppoli inequality (3.3) for

the gradient, which reads

$$\sum_{i=1}^{N} \int |u_{x_i}|^{p-2} u_{x_i x_j}^2 |u_{x_j}|^{2(2^\ell - 1)} \, dx \lesssim \sum_{i=1}^{N} \int |u_{x_i}|^{p-2} |u_{x_j}|^{2^{\ell+1}} \, dx,$$

and the last term is already contained in the right-hand side of (1.8). All in all, by applying the Sobolev inequality in the left-hand side of (1.8), we get the following type of self-improving information

$$\nabla u \in L^{2\gamma}(B_R) \quad \Longrightarrow \quad \nabla u \in L^{2^*\gamma}(B_r), \qquad \text{where we set } \gamma = \frac{p-2}{2} + 2^\ell.$$

In this way, we obtain an iterative scheme of reverse Holder's inequalities. This is **Step 1** in the proof of Proposition 5.1 below. Thus, apparently, we safely land in step (**c**) of the strategy described above.

We now want to pass to step (**d**) and iterate infinitely many times the previous information. The goal would be to define the diverging sequence of exponents γ_ℓ by

$$\gamma_\ell = \frac{p-2}{2} + 2^\ell, \qquad \ell \geq 1,$$

and conclude by iterating

$$\nabla u \in L^{2\gamma_\ell}(B_R) \quad \Longrightarrow \quad \nabla u \in L^{2^*\gamma_\ell}(B_r). \qquad (1.9)$$

Once again, *there is a drawback*. Indeed, observe that by definition

$$\frac{2^*}{2} \gamma_\ell \neq \gamma_{\ell+1}.$$

One may think that this is not a big issue: indeed, it would be sufficient to have

$$\gamma_{\ell+1} \leq \frac{2^*}{2} \gamma_\ell, \qquad (1.10)$$

then an application of Hölder's inequality in (1.9) would lead us to

$$\nabla u \in L^{2\gamma_\ell}(B_R) \quad \Longrightarrow \quad \nabla u \in L^{2\gamma_{\ell+1}}(B_r),$$

and we could enchain all the estimates. However, since the ratio $2^*/2$ tends to 1 as the dimension N goes to ∞, it is easy to see that (1.10) cannot be true in general. More precisely, such a condition holds only up to dimension $N = 4$.

The idea is then to go back to (1.9) and use interpolation in Lebesgue spaces in order to construct a Moser's scheme "*without holes*". In a nutshell, we control the term

$$\int_{B_R} |\nabla u|^{2\gamma_\ell} \, dx,$$

with

$$\int_{B_R} |\nabla u|^{2\gamma_{\ell-1}} \, dx \qquad \text{and} \qquad \int_{B_R} |\nabla u|^{2^*\gamma_\ell} \, dx,$$

and use an iteration over shrinking radii in order to absorb the last term, see **Step 2** of the proof of Proposition 5.1. Once this is done, we end up with the updated self-improving information

$$\nabla u \in L^{2\gamma_{\ell-1}}(B_R) \quad \Longrightarrow \quad \nabla u \in L^{2^*\gamma_\ell}(B_r).$$

What we gain is that now $2^* \gamma_\ell > 2 \gamma_\ell > 2 \gamma_{\ell-1}$, thus by using Hölder's inequality we obtain

$$\nabla u \in L^{2\gamma_{\ell-1}}(B_R) \quad \Longrightarrow \quad \nabla u \in L^{2\gamma_\ell}(B_r).$$

The information comes with a precise iterative estimate and a good control on the relevant constants. We can thus launch the Moser's iteration procedure and obtain the desired L^∞ estimate, see **Step 3** of the proof of Proposition 5.1.

There is still a small detail that needs some care: the first exponent of the iteration is

$$2\gamma_0 = p + 2,$$

which means that on ∇u we obtain a $L^\infty - L^{p+2}$ local estimate. Finally, in order to obtain the desired $L^\infty - L^p$ estimate, one can simply use an interpolation argument (this is **Step 4** of the proof of Proposition 5.1).

1.4 Plan of the paper

In Sect. 2, we define the approximation scheme and settle all the needed machinery. We have dedicated Sect. 3 to the new Caccioppoli inequalities which mix together the derivatives of the gradient with respect to 2 orthogonal directions. In Sect. 4, we exploit these Caccioppoli inequalities to establish integrability estimates on power functions of the gradient. In the subsequent section, we rely on these estimates to construct a Moser's iteration scheme which finally leads to the uniform a priori estimate of Proposition 5.1.

For ease of readability, both in Sects. 4 and 5, we first consider the case $f = 0$ and $\delta = 0$, in order to emphasize the main ideas and novelties of our approach. We explain subsequently in Sects. 4.2 and 5.2 respectively the technicalities to cover the general case $f \in W^{1,h}_{\text{loc}}(\Omega)$ and $\max\{\delta_i : i = 1, \ldots, N\} > 0$.

Finally, in "Appendix", we generalize Theorem 1.1 to nonlinear lower order terms.

2 Preliminaries

We will use the same approximation scheme as in [2, Section 2]. We introduce the notation

$$g_i(t) = \frac{1}{p} (|t| - \delta_i)_+^p, \qquad t \in \mathbb{R}, \ i = 1, \ldots, N,$$

where $0 \leq \delta_1, \ldots, \delta_N$ are given real numbers and we also set

$$\delta = 1 + \max\{\delta_i : i = 1, \ldots, N\}. \tag{2.1}$$

We are interested in local minimizers of the following variational integral

$$\mathfrak{F}_\delta(u; \Omega') = \sum_{i=1}^N \int_{\Omega'} g_i(u_{x_i}) \, dx + \int_{\Omega'} f \, u \, dx, \qquad u \in W^{1,p}_{\text{loc}}(\Omega),$$

where $\Omega' \Subset \Omega$ and $f \in W^{1,h}_{\text{loc}}(\Omega)$ for some $h > N/2$. The latter implies that

$$f \in L^{h^*}_{\text{loc}}(\Omega) \subset L^N_{\text{loc}}(\Omega) \subset L^{p'}_{\text{loc}}(\Omega).$$

The last inclusion is a consequence of the fact that $p \geq 2$ and $N \geq 2$. The condition $f \in L^{p'}_{\text{loc}}$ is exactly the one required in [2, Section 2] to justify the approximation scheme that we now describe.

We set

$$g_{i,\varepsilon}(t) = g_i(t) + \frac{\varepsilon}{2} t^2 = \frac{1}{p} \left(|t| - \delta_i \right)_+^p + \frac{\varepsilon}{2} t^2, \qquad t \in \mathbb{R}. \tag{2.2}$$

Remark 2.1 For $p = 2$ and $\delta_i > 0$, we have $g_i \in C^{1,1}(\mathbb{R}) \cap C^\infty(\mathbb{R} \setminus \{\delta_i, -\delta_i\})$, but g_i is not C^2. In this case, like in [3, Section 2] one would need to replace g_i by a regularized version, in particular for the C^2 regularity result of Lemma 2.2 below. In order not to overburden the presentation, we prefer to avoid to explicitly write down this regularization and keep on using the same symbol g_i.

From now on, *we fix U a local minimizer of* $\mathfrak{F}_\delta$. We also fix a ball

$$B \Subset \Omega \quad \text{such that} \quad 2B \Subset \Omega \text{ as well.}$$

Here λB denotes the ball having the same center as B, scaled by a factor $\lambda > 0$.

For every $0 < \varepsilon \ll 1$ and every $x \in \overline{B}$, we set $U_\varepsilon(x) = U * \varrho_\varepsilon(x)$, where ϱ_ε is a smooth convolution kernel, supported in a ball of radius ε centered at the origin.

Finally, we define

$$\mathfrak{F}_{\delta,\varepsilon}(v; B) = \sum_{i=1}^N \int_B g_{i,\varepsilon}(v_{x_i}) \, dx + \int_B f_\varepsilon \, v \, dx,$$

where $f_\varepsilon = f * \varrho_\varepsilon$. The following preliminary result is standard, see [2, Lemma 2.5 and Lemma 2.8].

Lemma 2.2 (Basic energy estimate) *There exists $\varepsilon_0 > 0$ such that for every $0 < \varepsilon \le \varepsilon_0 < 1$, the problem*

$$\min \left\{ \mathfrak{F}_\varepsilon(v; B) \, : \, v - U_\varepsilon \in W_0^{1,p}(B) \right\}, \tag{2.3}$$

admits a unique solution u_ε. Moreover, there exists a constant $C = C(N, p) > 0$ such that the following uniform estimate holds

$$\int_B |\nabla u_\varepsilon|^p \, dx \le C \left[\int_{2B} |\nabla U|^p \, dx + |B|^{\frac{p'}{N}} \int_{2B} |f|^{p'} \, dx + (\varepsilon_0 + (\delta - 1)^p) |B| \right].$$

Finally, $u_\varepsilon \in C^2(B)$.

We also rely on the following stability result, which is slightly more precise than [2, Lemma 2.9].

Lemma 2.3 (Convergence to a minimizer) *With the same notation as before, there exists a sequence $\{\varepsilon_k\}_{k \in \mathbb{N}} \subset (0, \varepsilon_0)$ converging to 0, such that*

$$\lim_{k \to \infty} \|u_{\varepsilon_k} - \widetilde{u}\|_{L^p(B)} = 0,$$

where $\widetilde{u}$ is a solution of

$$\min \left\{ \mathfrak{F}_\delta(v; B) \, : \, v - U \in W_0^{1,p}(B) \right\}.$$

We also have

$$\left| \widetilde{u}_{x_i} - U_{x_i} \right| \le 2\,\delta_i, \qquad \text{for a. e. } x \in B, \quad i = 1, \dots, N. \tag{2.4}$$

In the case $\delta = 1$, i.e. when $\delta_1 = \cdots = \delta_N = 0$, then $\widetilde{u} = U$ and we have the stronger convergence

$$\lim_{k \to \infty} \|u_{\varepsilon_k} - U\|_{W^{1,p}(B)} = 0. \tag{2.5}$$

Proof The first part is proven in [2, Lemma 2.9], while (2.4) is proven in [2, Lemma 2.3]. For the case $\delta = 1$, we observe that $\widetilde{u} = U$ follows from the strict convexity of the functional, together with the local minimality of U. In order to prove (2.5), we observe that

$$\left| \sum_{i=1}^{N} \frac{1}{p} \int_B \left| (u_{\varepsilon_k})_{x_i} \right|^p dx - \sum_{i=1}^{N} \frac{1}{p} \int_B \left| U_{x_i} \right|^p dx \right| \leq \left| \mathfrak{F}_{\delta,\varepsilon_k}(u_{\varepsilon_k}; B) - \mathfrak{F}_\delta(U; B) \right|$$

$$+ \frac{\varepsilon_k}{2} \int_B |\nabla u_{\varepsilon_k}|^2 dx$$

$$+ \left| \int_B f_{\varepsilon_k} u_{\varepsilon_k} dx - \int_B f U dx \right|.$$

We now use that $\{u_{\varepsilon_k}\}_{k \in \mathbb{N}}$ strongly converges in $L^p(B)$, is bounded in $W^{1,p}(B)$ and that $\{f_{\varepsilon_k}\}_{k \in \mathbb{N}}$ strongly converges in $L^{p'}(B)$ to f. By further using that (see the proof of [2, Lemma 2.9])

$$\lim_{k \to \infty} \left| \mathfrak{F}_{\delta,\varepsilon_k}(u_{\varepsilon_k}; B) - \mathfrak{F}_\delta(U; B) \right| = 0,$$

we finally get

$$\lim_{k \to \infty} \sum_{i=1}^{N} \int_B \left| (u_{\varepsilon_k})_{x_i} \right|^p dx = \sum_{i=1}^{N} \int_B \left| U_{x_i} \right|^p dx, \quad i = 1, \ldots, N. \tag{2.6}$$

Observe that by Clarkson's inequality for $p \geq 2$, we obtain

$$\sum_{i=1}^{N} \left\| \frac{(u_{\varepsilon_k})_{x_i} + U_{x_i}}{2} \right\|_{L^p(B)}^p + \sum_{i=1}^{N} \left\| \frac{(u_{\varepsilon_k})_{x_i} - U_{x_i}}{2} \right\|_{L^p(B)}^p$$

$$\leq \frac{1}{2} \left(\sum_{i=1}^{N} \left\| (u_{\varepsilon_k})_{x_i} \right\|_{L^p(B)}^p + \sum_{i=1}^{N} \left\| U_{x_i} \right\|_{L^p(B)}^p \right).$$

By using this and (2.6), we eventually get (2.5). $\qquad\square$

Remark 2.4 Observe that the functional $\mathfrak{F}_\delta$ is not strictly convex when $\delta > 1$. Thus property (2.4) is useful in order to transfer a Lipschitz estimate for the minimizer $\widetilde{u}$ selected in the limit, to the chosen one U.

Finally, we will repeatedly use the following classical result, see [11, Lemma 6.1] for a proof.

Lemma 2.5 *Let $0 < r < R$ and let $Z(t) : [r, R] \to [0, \infty)$ be a bounded function. Assume that for $r \leq t < s \leq R$ we have*

$$Z(t) \leq \frac{\mathcal{A}}{(s-t)^{\alpha_0}} + \frac{\mathcal{B}}{(s-t)^{\beta_0}} + \mathcal{C} + \vartheta \, Z(s),$$

with $\mathcal{A}, \mathcal{B}, \mathcal{C} \geq 0$, $\alpha_0 \geq \beta_0 > 0$ and $0 \leq \vartheta < 1$. Then we have

$$Z(r) \leq \left(\frac{1}{(1-\lambda)^{\alpha_0}} \frac{\lambda^{\alpha_0}}{\lambda^{\alpha_0} - \vartheta} \right) \left[\frac{\mathcal{A}}{(R-r)^{\alpha_0}} + \frac{\mathcal{B}}{(R-r)^{\beta_0}} + \mathcal{C} \right],$$

where λ is any number such that

$$\vartheta^{\frac{1}{\alpha_0}} < \lambda < 1.$$

3 Caccioppoli-type inequalities

The solution u_ε of the regularized problem (2.3) satisfies the Euler-Lagrange equation

$$\sum_{i=1}^{N} \int g'_{i,\varepsilon}((u_\varepsilon)_{x_i}) \, \varphi_{x_i} \, dx + \int f_\varepsilon \, \varphi \, dx = 0, \qquad \varphi \in W_0^{1,p}(B). \tag{3.1}$$

From now on, in order to simplify the notation, we will systematically forget the subscript ε on u_ε and f_ε and *simply write u and f respectively*.

We now insert a test function of the form $\varphi = \psi_{x_j} \in W_0^{1,p}(B)$ in (3.1), compactly supported in B. Then an integration by parts yields

$$\sum_{i=1}^{N} \int g''_{i,\varepsilon}(u_{x_i}) \, u_{x_i x_j} \, \psi_{x_i} \, dx + \int f_{x_j} \, \psi \, dx = 0, \tag{3.2}$$

for $j = 1, \ldots, N$. This is the equation solved by u_{x_j}.

We refer to [2, Lemma 3.2] for a proof of the following Caccioppoli inequality:

Lemma 3.1 *Let $\Phi : \mathbb{R} \to \mathbb{R}^+$ be a C^1 convex function. Then there exists a universal constant $C > 0$ such that for every function $\eta \in C_0^\infty(B)$ and every $j = 1, \ldots, N$, we have*

$$\sum_{i=1}^{N} \int g''_{i,\varepsilon}(u_{x_i}) \left| \left(\Phi(u_{x_j}) \right)_{x_i} \right|^2 \eta^2 \, dx$$

$$\leq C \sum_{i=1}^{N} \int g''_{i,\varepsilon}(u_{x_i}) \, |\Phi(u_{x_j})|^2 \, \eta_{x_i}^2 \, dx + C \int |f_{x_j}| \, |\Phi'(u_{x_j})| \, |\Phi(u_{x_j})| \, \eta^2 \, dx. \tag{3.3}$$

We need a more elaborate Caccioppoli-type inequality for the gradient, which is reminiscent of [1, Proposition 3.1].

Proposition 3.2 (Weird Caccioppoli inequality) *Let $\Phi, \Psi : [0, +\infty) \to [0, +\infty)$ be two non-decreasing continuous functions. We further assume that Ψ is convex and C^1. Then there exists a universal constant $C > 0$ such that for every $\eta \in C_0^\infty(B)$, $0 \leq \theta \leq 2$ and $k, j = 1, \ldots, N$,*

$$\sum_{i=1}^{N} \int g''_{i,\varepsilon}(u_{x_i}) \, u_{x_i x_j}^2 \, \Phi(u_{x_j}^2) \, \Psi(u_{x_k}^2) \, \eta^2 \, dx$$

$$\leq C \sum_{i=1}^{N} \int g''_{i,\varepsilon}(u_{x_i}) \, u_{x_j}^2 \, \Phi(u_{x_j}^2) \, \Psi(u_{x_k}^2) \, |\nabla \eta|^2 \, dx$$

$$+ C \left(\sum_{i=1}^{N} \int g''_{i,\varepsilon}(u_{x_i}) \, u_{x_i x_j}^2 \, u_{x_j}^2 \, \Phi(u_{x_j}^2)^2 \, \Psi'(u_{x_k}^2)^\theta \, \eta^2 \, dx \right)^{\frac{1}{2}} \tag{3.4}$$

$$\times \left[\left(\sum_{i=1}^{N} \int g''_{i,\varepsilon}(u_{x_i}) \, |u_{x_k}|^{2\theta} \, \Psi(u_{x_k}^2)^{2-\theta} \, |\nabla \eta|^2 \, dx \right)^{\frac{1}{2}} + \mathcal{E}_1(f)^{\frac{1}{2}} \right] + C \, \mathcal{E}_2(f)$$

where

$$\mathcal{E}_1(f) := \int |f_{x_k}| \, |u_{x_k}|^{\theta+1} \left| \Psi(u_{x_k}^2) \, \Psi'(u_{x_k}^2) \right|^{1-\frac{\theta}{2}} \eta^2 \, dx,$$

$$\mathcal{E}_2(f) := \int |f_{x_j}| \, |u_{x_j}| \, \Phi(u_{x_j}^2) \, \Psi(u_{x_k}^2) \, \eta^2 \, dx.$$

Proof By a standard approximation argument, one can assume that Φ is C^1 as well. We take in (3.2)

$$\varphi = u_{x_j} \, \Phi(u_{x_j}^2) \, \Psi(u_{x_k}^2) \, \eta^2.$$

This gives

$$\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j}^2 \left(\Phi(u_{x_j}^2) + 2u_{x_j}^2 \, \Phi'(u_{x_j}^2) \right) \Psi(u_{x_k}^2) \, \eta^2 \, dx$$

$$= -2 \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j} \, u_{x_j} \, \Phi(u_{x_j}^2) \, \Psi(u_{x_k}^2) \, \eta \, \eta_{x_i} \, dx$$

$$- 2 \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j} \, u_{x_j} \, u_{x_i x_k} \, u_{x_k} \, \Psi'(u_{x_k}^2) \, \Phi(u_{x_j}^2) \, \eta^2 \, dx$$

$$- \int f_{x_j} \, u_{x_j} \, \Phi(u_{x_j}^2) \, \Psi(u_{x_k}^2) \, \eta^2 \, dx =: \mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3. \tag{3.5}$$

We now proceed to estimating the three terms $\mathcal{A}_\ell$. We have

$$\mathcal{A}_1 \le \frac{1}{2} \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j}^2 \, \Phi(u_{x_j}^2) \, \Psi(u_{x_k}^2) \, \eta^2 \, dx$$

$$+ 2 \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_j}^2 \, \Phi(u_{x_j}^2) \, \Psi(u_{x_k}^2) \, \eta_{x_i}^2 \, dx$$

and the integral containing the Hessian of u can be absorbed in the left-hand side of (3.5). Using also that $2 u_{x_j}^2 \, \Phi'(u_{x_j}^2) \ge 0$, this yields

$$\frac{1}{2} \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j}^2 \, \Phi(u_{x_j}^2) \, \Psi(u_{x_k}^2) \, \eta^2 \, dx$$

$$\le 2 \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_j}^2 \, \Phi(u_{x_j}^2) \, \Psi(u_{x_k}^2) \, \eta_{x_i}^2 \, dx + \mathcal{A}_2 + \mathcal{A}_3. \tag{3.6}$$

We now estimate $\mathcal{A}_2$, which is the most delicate term: writing $\Psi'(u_{x_k}^2) = \Psi'(u_{x_k}^2)^{\frac{\theta}{2}} \, \Psi'(u_{x_k}^2)^{1-\frac{\theta}{2}}$ and using the Cauchy-Schwarz inequality, we get

$$\mathcal{A}_2 \le 2 \left(\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j}^2 \, u_{x_j}^2 \, \Phi(u_{x_j}^2)^2 \, \Psi'(u_{x_k}^2)^{\theta} \, \eta^2 \, dx \right)^{\frac{1}{2}}$$

$$\times \left(\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_k}^2 \, u_{x_k}^2 \, \Psi'(u_{x_k}^2)^{2-\theta} \, \eta^2 \, dx \right)^{\frac{1}{2}}. \tag{3.7}$$

We observe that

$$\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_k}^2 \, u_{x_k}^2 \, \Psi'(u_{x_k}^2)^{2-\theta} \, \eta^2 \, dx = \frac{1}{4} \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, \left| (G(u_{x_k}))_{x_i} \right|^2 \eta^2 \, dx,$$

where G is the convex nonnegative C^1 function defined by

$$G(t) = \int_0^{t^2} \Psi'(\tau)^{1-\frac{\theta}{2}} \, d\tau.$$

Thus by the Caccioppoli inequality (3.3) with x_k in place of x_j and

$$\Phi(t) = G(t), \qquad t \in \mathbb{R},$$

we get

$$\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_k}^2 \, u_{x_k}^2 \, \Psi'(u_{x_k}^2)^{2-\theta} \, \eta^2 \leq C \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, G(u_{x_k})^2 \, \eta_{x_i}^2 \, dx$$

$$+ \, C \int |f_{x_k}| \left| G(u_{x_k}) \, G'(u_{x_k}) \right| \eta^2 \, dx.$$

By Jensen's inequality

$$0 \leq G(u_{x_k}) \leq |u_{x_k}|^{\theta} \left(\int_0^{u_{x_k}^2} \Psi'(\tau) \, d\tau \right)^{1-\frac{\theta}{2}} \leq |u_{x_k}|^{\theta} \, \Psi(u_{x_k}^2)^{1-\frac{\theta}{2}}.$$

Together with the fact that $G'(u_{x_k}) = 2 \, u_{x_k} \Psi'(u_{x_k}^2)^{1-\frac{\theta}{2}}$, this implies

$$\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_k}^2 \, u_{x_k}^2 \, \Psi'(u_{x_k}^2)^{2-\theta} \, \eta^2 \leq C \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, |u_{x_k}|^{2\theta} \, \Psi(u_{x_k}^2)^{2-\theta} \, \eta_{x_i}^2 \, dx$$

$$+ \, C \int |f_{x_k}| \, |u_{x_k}|^{\theta+1} \left| \Psi(u_{x_k}^2) \, \Psi'(u_{x_k}^2) \right|^{1-\frac{\theta}{2}} \eta^2 \, dx,$$

which in turn yields by (3.6) and (3.7),

$$\frac{1}{2} \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j}^2 \, \Phi(u_{x_j}^2) \, \Psi(u_{x_k}^2) \, \eta^2 \, dx$$

$$\leq 2 \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_j}^2 \, \Phi(u_{x_j}^2) \, \Psi(u_{x_k}^2) \, \eta_{x_i}^2 \, dx$$

$$+ \, C \left(\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j}^2 \, u_{x_j}^2 \, \Phi(u_{x_j}^2)^2 \, \Psi'(u_{x_k}^2)^{\theta} \, \eta^2 \, dx \right)^{\frac{1}{2}}$$

$$\times \left[\left(\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, |u_{x_k}|^{2\theta} \, \Psi(u_{x_k}^2)^{2-\theta} \, \eta_{x_i}^2 \, dx \right)^{\frac{1}{2}} \right.$$

$$\left. + \left(\int |f_{x_k}| \, |u_{x_k}|^{\theta+1} \left| \Psi(u_{x_k}^2) \, \Psi'(u_{x_k}^2) \right|^{1-\frac{\theta}{2}} \eta^2 \, dx \right)^{\frac{1}{2}} \right] + \mathcal{A}_3.$$

Here, we have also used the inequality $(A + B)^{1/2} \leq A^{1/2} + B^{1/2}$.

Finally,

$$\mathcal{A}_3 \leq C \int |f_{x_j}| \, |u_{x_j}| \, \Phi(u_{x_j}^2) \, \Psi(u_{x_k}^2) \, \eta^2 \, dx.$$

This completes the proof.   □

4 Local energy estimates for the regularized problem

In order to emphasize the main ideas of the proof, we have divided this section in two parts. In the first one, we explain how (3.4) leads to higher integrability estimates for the gradient when $f = 0$ and $\delta = 1$. This allows to ignore a certain amount of technicalities. In the second part, we then detail the modifications of the proof to obtain the corresponding estimates in the general case.

4.1 The homogeneous case

In this subsection, we assume that $f = 0$ and $\delta = 1$. Then the two terms $\mathcal{E}_1(f)$ and $\mathcal{E}_2(f)$ in (3.4) vanish. Also observe that in this case from (2.2) we have

$$g''_{i,\varepsilon}(t) = (p - 1) \, |t|^{p-2} + \varepsilon.$$

Let us single out a particular case of Proposition 3.2 by taking

$$\Phi(t) = t^{s-1} \quad \text{and} \quad \Psi(t) = t^m, \qquad \text{for } t \geq 0, \tag{4.1}$$

with $1 \leq s \leq m$.

Proposition 4.1 (Staircase to the full Caccioppoli) *Let $p \geq 2$ and let $\eta \in C_0^\infty(B)$, then for every $k, j = 1, \ldots, N$ and $1 \leq s \leq m$*

$$\sum_{i=1}^N \int g''_{i,\varepsilon}(u_{x_i}) \, u_{x_i x_j}^2 \, |u_{x_j}|^{2s-2} \, |u_{x_k}|^{2m} \, \eta^2 \, dx$$

$$\leq C \sum_{i=1}^N \int g''_{i,\varepsilon}(u_{x_i}) \, |u_{x_j}|^{2s+2m} \, |\nabla \eta|^2 \, dx$$

$$+ C \, (m + 1) \sum_{i=1}^N \int g''_{i,\varepsilon}(u_{x_i}) \, |u_{x_k}|^{2s+2m} \, |\nabla \eta|^2 \, dx \tag{4.2}$$

$$+ \sum_{i=1}^N \int g''_{i,\varepsilon}(u_{x_i}) \, u_{x_i x_j}^2 \, |u_{x_j}|^{4s-2} \, |u_{x_k}|^{2m-2s} \, \eta^2 \, dx.$$

Proof We use (3.4) with the choices (4.1) above and

$$\theta = \begin{cases} \dfrac{m - s}{m - 1} \in [0, 1] & \text{if } m > 1, \\[2ex] 1 & \text{if } m = 1. \end{cases}$$

This gives

$$\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, u_{x_i x_j}^2 \, |u_{x_j}|^{2s-2} \, |u_{x_k}|^{2m} \, \eta^2 \, dx$$

$$\leq C \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, |u_{x_j}|^{2s} \, |u_{x_k}|^{2m} \, |\nabla\eta|^2 \, dx$$

$$+ C \left(m^\theta \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, u_{x_i x_j}^2 \, |u_{x_j}|^{4s-2} \, |u_{x_k}|^{2m-2s} \, \eta^2 \, dx \right)^{\frac{1}{2}}$$

$$\times \left(\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, |u_{x_k}|^{2m+2s} \, |\nabla\eta|^2 \, dx \right)^{\frac{1}{2}}.$$

We use Young's inequality in the form $C\sqrt{a\,b} \leq C^2\, b/4 + a$ for the product in the right-hand side to get

$$\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, u_{x_i x_j}^2 \, |u_{x_j}|^{2s-2} \, |u_{x_k}|^{2m} \, \eta^2 \, dx$$

$$\leq C \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, |u_{x_j}|^{2s} \, |u_{x_k}|^{2m} \, |\nabla\eta|^2 \, dx$$

$$+ C\, m^\theta \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, |u_{x_k}|^{2m+2s} \, |\nabla\eta|^2 \, dx$$

$$+ \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, u_{x_i x_j}^2 \, |u_{x_j}|^{4s-2} \, |u_{x_k}|^{2m-2s} \, \eta^2 \, dx.$$

In the first term of the right-hand side, we use Young's inequality with the exponents

$$\frac{2m+2s}{2s} \quad , \quad \frac{2m+2s}{2m}.$$

We also observe for the second term that $m^\theta \leq m$. This gives the desired estimate. $\qquad\square$

Proposition 4.2 (Caccioppoli for power functions of the gradient) *We fix an exponent*

$$q = 2^{\ell_0} - 1, \qquad \text{for a given } \ell_0 \in \mathbb{N} \setminus \{0\}.$$

Let $\eta \in C_0^\infty(B)$, then for every $k = 1, \ldots, N$ we have

$$\int \left| \nabla \left(|u_{x_k}|^{q+\frac{p-2}{2}} u_{x_k} \right) \right|^2 \eta^2 \, dx \leq C\, q^5 \sum_{i,j=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, |u_{x_j}|^{2q+2} \, |\nabla\eta|^2 \, dx$$

$$+ C\, q^5 \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, |u_{x_k}|^{2q+2} \, |\nabla\eta|^2 \, dx, \tag{4.3}$$

for some $C = C(N, p) > 0$.

Proof We define the two finite families of indices $\{s_\ell\}$ and $\{m_\ell\}$ such that

$$s_\ell = 2^\ell, \qquad m_\ell = q + 1 - 2^\ell, \qquad \ell \in \{0, \ldots, \ell_0\}.$$

Observe that

$$
\begin{aligned}
&1 \leq s_\ell \leq m_\ell, \qquad \ell \in \{0, \ldots, \ell_0 - 1\}, \\
&s_\ell + m_\ell = q + 1, \qquad \ell \in \{0, \ldots, \ell_0\}, \\
&4\, s_\ell - 2 = 2\, s_{\ell+1} - 2, \qquad 2\, m_\ell - 2\, s_\ell = 2\, m_{\ell+1},
\end{aligned}
$$

and

$$s_0 = 1, \qquad m_0 = q, \qquad s_{\ell_0} = 2^{\ell_0}, \qquad m_{\ell_0} = 0.$$

In terms of these families, inequality (4.2) implies for every $\ell \in \{0, \ldots, \ell_0 - 1\}$

$$
\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, u_{x_i x_j}^2\, |u_{x_j}|^{2\,s_\ell - 2}\, |u_{x_k}|^{2\,m_\ell}\, \eta^2\, dx
$$

$$
\leq C \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, |u_{x_j}|^{2\,q+2}\, |\nabla \eta|^2\, dx
$$

$$
+ C\,(m_\ell + 1) \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, |u_{x_k}|^{2\,q+2}\, |\nabla \eta|^2\, dx
$$

$$
+ \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, u_{x_i x_j}^2\, |u_{x_j}|^{2\,s_{\ell+1} - 2}\, |u_{x_k}|^{2\,m_{\ell+1}}\, \eta^2\, dx,
$$

for some $C > 0$ universal. By starting from $\ell = 0$ and iterating the previous estimate up to $\ell = \ell_0 - 1$, we then get

$$
\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, u_{x_i x_j}^2\, |u_{x_k}|^{2\,q}\, \eta^2\, dx \leq C\, q^2 \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, |u_{x_j}|^{2\,q+2}\, |\nabla \eta|^2\, dx
$$

$$
+ C\, q^2 \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, |u_{x_k}|^{2\,q+2}\, |\nabla \eta|^2\, dx
$$

$$
+ \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i})\, u_{x_i x_j}^2\, |u_{x_j}|^{2\,q}\, \eta^2\, dx,
$$

for a universal constant $C > 0$. For the last term, we apply the Caccioppoli inequality (3.3) with

$$\Phi(t) = \frac{|t|^{q+1}}{q+1}, \qquad t \in \mathbb{R},$$

thus we get

$$\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j}^2 \, |u_{x_k}|^{2q} \, \eta^2 \, dx \leq C q^2 \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, |u_{x_j}|^{2q+2} \, |\nabla \eta|^2 \, dx$$

$$+ C q^2 \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, |u_{x_k}|^{2q+2} \, |\nabla \eta|^2 \, dx$$

$$+ \frac{C}{(q+1)^2} \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, |u_{x_j}|^{2q+2} \, |\nabla \eta|^2 \, dx;$$

that is,

$$\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j}^2 \, |u_{x_k}|^{2q} \, \eta^2 \, dx \leq C q^2 \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, |u_{x_j}|^{2q+2} \, |\nabla \eta|^2 \, dx$$

$$+ C q^2 \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, |u_{x_k}|^{2q+2} \, |\nabla \eta|^2 \, dx, \tag{4.4}$$

possibly for a different universal constant $C > 0$.

We now observe that $g_{i,\varepsilon}''(u_{x_i}) = \left((p-1) \, |u_{x_i}|^{p-2} + \varepsilon \right)$ and thus

$$\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j}^2 \, |u_{x_k}|^{2q} \, \eta^2 \, dx \geq \int |u_{x_k}|^{p-2} \, u_{x_k x_j}^2 \, |u_{x_k}|^{2q} \, \eta^2 \, dx$$

$$= \left(\frac{2}{2q+p} \right)^2 \int \left| \left(|u_{x_k}|^{q + \frac{p-2}{2}} u_{x_k} \right)_{x_j} \right|^2 \eta^2 \, dx.$$

When we sum over $j = 1, \ldots, N$, we get

$$\sum_{i,j=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j}^2 \, |u_{x_k}|^{2q} \, \eta^2 \, dx \geq \left(\frac{2}{2q+p} \right)^2 \int \left| \nabla \left(|u_{x_k}|^{q + \frac{p-2}{2}} u_{x_k} \right) \right|^2 \eta^2 \, dx.$$

This proves the desired inequality. $\qquad\square$

4.2 The non-homogeneous case

In the general case where $f \neq 0$ and/or $\delta > 1$, we can prove the following analogue of (4.2), in a similar way:

$$\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j}^2 \, |u_{x_j}|^{2s-2} \, |u_{x_k}|^{2m} \, \eta^2 \, dx$$

$$\leq \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j}^2 \, |u_{x_j}|^{4s-2} \, |u_{x_k}|^{2m-2s} \, \eta^2 \, dx$$

$$+ C (m+1) \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \left(|u_{x_j}|^{2s+2m} + |u_{x_k}|^{2s+2m} \right) |\nabla \eta|^2 \, dx \tag{4.5}$$

$$+ C m^2 \int |\nabla f| \left(|u_{x_k}|^{2s+2m-1} + |u_{x_j}|^{2s+2m-1} \right) \eta^2 \, dx.$$

We then deduce the following analogue of Proposition 4.2:

Proposition 4.3 *We fix an exponent*

$$q = 2^{\ell_0} - 1, \qquad \text{for a given } \ell_0 \in \mathbb{N} \setminus \{0\}.$$

Let $\eta \in C_0^\infty(\Omega)$, then for every $k = 1, \ldots, N$ we have

$$\int \left| \nabla \left((|u_{x_k}| - \delta_k)_+^{\frac{p}{2}} |u_{x_k}|^q \right) \right|^2 \eta^2 \, dx$$

$$\leq C \, q^5 \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \left(|u_{x_k}|^{2q+2} + \sum_{j=1}^{N} |u_{x_j}|^{2q+2} \right) |\nabla \eta|^2 \, dx \tag{4.6}$$

$$+ C \, q^5 \int |\nabla f| \left(|u_{x_k}|^{2q+1} + \sum_{j=1}^{N} |u_{x_j}|^{2q+1} \right) \eta^2 \, dx,$$

for some $C = C(N, p) > 0$.

Proof Using the same notation and the same strategy as in the proof of (4.3), except that we start from (4.5) instead of (4.2), we get the following analogue of (4.4):

$$\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j}^2 \, |u_{x_k}|^{2q} \, \eta^2 \, dx$$

$$\leq C \, q^2 \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) (|u_{x_j}|^{2q+2} + |u_{x_k}|^{2q+2}) \, |\nabla \eta|^2 \, dx$$

$$+ C \, q^3 \int |\nabla f| (|u_{x_k}|^{2q+1} + |u_{x_j}|^{2q+1}) \, \eta^2 \, dx.$$

We now observe that

$$\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j}^2 \, |u_{x_k}|^{2q} \, \eta^2 \, dx \geq (p-1) \int (|u_{x_k}| - \delta_k)_+^{p-2} \, u_{x_k x_j}^2 \, |u_{x_k}|^{2q} \, \eta^2 \, dx.$$

Noting that

$$(|u_{x_k}| - \delta_k)_+^p \leq (|u_{x_k}| - \delta_k)_+^{p-2} |u_{x_k}|^2,$$

we have

$$\left| \left((|u_{x_k}| - \delta_k)_+^{\frac{p}{2}} |u_{x_k}|^q \right)_{x_j} \right|^2 \leq 2 \left| \left((|u_{x_k}| - \delta_k)_+^{\frac{p}{2}} \right)_{x_j} \right|^2 |u_{x_k}|^{2q}$$

$$+ 2 (|u_{x_k}| - \delta_k)_+^p \left| \left(|u_{x_k}|^q \right)_{x_j} \right|^2$$

$$\leq C \, q^2 \, (|u_{x_k}| - \delta_k)_+^{p-2} \, |u_{x_k}|^{2q} \, u_{x_k x_j}^2.$$

We deduce therefrom

$$\sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j}^2 \, |u_{x_k}|^{2q} \, \eta^2 \, dx \geq \frac{C}{q^2} \int \left| \left((|u_{x_k}| - \delta_k)_+^{\frac{p}{2}} |u_{x_k}|^q \right)_{x_j} \right|^2 \eta^2 \, dx,$$

thus when we sum over $j = 1, \dots, N$, we get

$$\sum_{i,j=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, u_{x_i x_j}^2 \, |u_{x_k}|^{2q} \, \eta^2 \, dx \geq \frac{C}{q^2} \int \left| \nabla \left((|u_{x_k}| - \delta_k)_+^{\frac{p}{2}} \, |u_{x_k}|^q \right) \right|^2 \eta^2 \, dx.$$

This proves the desired inequality (4.6). $\qquad\square$

5 Proof of Theorem 1.1

Proof The core of the proof of Theorem 1.1 is the uniform Lipschitz estimate of Proposition 5.1 below. Its proof, which is postponed for ease of readability, uses the integrability estimates of Sect. 4. Once we have this uniform estimate, we can reproduce the proof of [2, Theorem A] and prove that $\nabla U \in L^\infty(\Omega')$, for every $\Omega' \Subset \Omega$.

We now detail how to obtain the scaling invariant local estimate (1.5) in the case $\delta_1 = \cdots = \delta_N = 0$. We take $0 < r_0 < R_0 \leq 1$ and a ball $B_{2R_0} \Subset \Omega$. We then consider the sequence of miminizers $\{u_{\varepsilon_k}\}_{k \in \mathbb{N}}$ of (2.3) obtained in Lemma 2.3, with B a ball slightly larger than B_{R_0} so that $2B \Subset \Omega$. By using the uniform Lipschitz estimate (5.3) below, taking the limit as k goes to ∞ and using the strong convergence of Lemma 2.3, we obtain

$$\|\nabla U\|_{L^\infty(B_{r_0})} \leq \frac{C}{(R_0 - r_0)^{\sigma_2}} \left(1 + \|\nabla f\|_{L^h(B_{R_0})}^{\sigma_2} \right) \left(\|\nabla U\|_{L^p(B_{R_0})}^{\sigma_1} + 1 \right).$$

Without loss of generality, we can assume that $\|\nabla U\|_{L^p(B_{R_0})} > 0$. Hence, by Young's inequality,

$$\|\nabla U\|_{L^\infty(B_{r_0})} \leq \frac{C}{(R_0 - r_0)^{\sigma_2}} \left(1 + \|\nabla f\|_{L^h(B_{R_0})}^{2\sigma_2} + \|\nabla U\|_{L^p(B_{R_0})}^{2\sigma_1} \right), \tag{5.1}$$

possibly for a different $C = C(N, p, h) > 0$. We now observe that for every $\lambda > 0$, λU is still a solution of the orthotropic $p-$Laplace equation, with the right hand side f replaced by $\lambda^{p-1} f$. We can use (5.1) for λU and get

$$\lambda \|\nabla U\|_{L^\infty(B_{r_0})} \leq \frac{C}{(R_0 - r_0)^{\sigma_2}} \left(1 + \lambda^{2\sigma_2(p-1)} \|\nabla f\|_{L^h(B_{R_0})}^{2\sigma_2} + \lambda^{2\sigma_1} \|\nabla U\|_{L^p(B_{R_0})}^{2\sigma_1} \right).$$

Dividing by λ, we obtain

$$\|\nabla U\|_{L^\infty(B_{r_0})} \leq \frac{C}{(R_0 - r_0)^{\sigma_2}} \left(\frac{1}{\lambda} + \lambda^{2\sigma_2(p-1)-1} \|\nabla f\|_{L^h(B_{R_0})}^{2\sigma_2} + \lambda^{2\sigma_1-1} \|\nabla U\|_{L^p(B_{R_0})}^{2\sigma_1} \right).$$

We take

$$\lambda := \frac{1}{\|\nabla U\|_{L^p(B_{R_0})} + \|\nabla f\|_{L^h(B_{R_0})}^{\frac{1}{p-1}}},$$

and observe that if $\|\nabla f\|_{L^h(B_{R_0})} > 0$, then

$$\lambda^{2\sigma_2(p-1)-1} \|\nabla f\|_{L^h(B_{R_0})}^{2\sigma_2} \leq \frac{1}{\left(\|\nabla f\|_{L^h(B_{R_0})}^{\frac{1}{p-1}} \right)^{2\sigma_2(p-1)-1}} \|\nabla f\|_{L^h(B_{R_0})}^{2\sigma_2} = \|\nabla f\|_{L^h(B_{R_0})}^{\frac{1}{p-1}}$$

while the inequality is obvious when $\|\nabla f\|_{L^h(B_{R_0})} = 0$. Similarly,

$$\lambda^{2\sigma_1-1} \|\nabla U\|_{L^p(B_{R_0})}^{2\sigma_1} \leq \frac{1}{\|\nabla U\|_{L^p(B_{R_0})}^{2\sigma_1-1}} \|\nabla U\|_{L^p(B_{R_0})}^{2\sigma_1} = \|\nabla U\|_{L^p(B_{R_0})}.$$

It thus follows that

$$\|\nabla U\|_{L^\infty(B_{r_0})} \leq \frac{C}{(R_0 - r_0)^{\sigma_2}} \left(\|\nabla f\|_{L^h(B_{R_0})}^{\frac{1}{p-1}} + \|\nabla U\|_{L^p(B_{R_0})} \right). \tag{5.2}$$

We now make this estimate dimensionally correct. Given $R_0 > 0$, we consider a ball $B_{2R_0} \Subset \Omega$. Then the rescaled function

$$U_{R_0}(x) = U(R_0 \, x), \qquad \text{for } x \in R_0^{-1} \, \Omega,$$

is a solution of the orthotropic p-Laplace equation, with right-hand side $f_{R_0}(x) := R_0^p \, f(R_0 \, x)$. We can use for it the estimate (5.2) with radii 1 and 1/2. By scaling back, we thus obtain

$$R_0 \|\nabla U\|_{L^\infty(B_{R_0/2})} \leq C \left(R_0^{-\frac{N}{p}+1} \|\nabla U\|_{L^p(B_{R_0})} + R_0^{\frac{h(p+1)-N}{h(p-1)}} \|\nabla f\|_{L^h(B_{R_0})}^{\frac{1}{p-1}} \right),$$

for some constant $C = C(N, p, h) > 1$. Dividing by R_0, we get

$$\|\nabla U\|_{L^\infty(B_{R_0/2})} \leq C \left(\fint_{B_{R_0}} |\nabla U|^p \, dx \right)^{\frac{1}{p}} + C \, R_0^{\frac{2}{p-1} - \frac{N}{h(p-1)}} \left(\fint_{B_{R_0}} |\nabla f|^h \, dx \right)^{\frac{1}{h(p-1)}}.$$

This concludes the proof. $\qquad\qquad\square$

Proposition 5.1 (Uniform Lipschitz estimate) *Let $p \geq 2$, $h > N/2$ and $0 < \varepsilon \leq \varepsilon_0$. For every $B_{r_0} \subset B_{R_0} \Subset B$ with $0 < r_0 < R_0 \leq 1$, we have*

$$\|\nabla u_\varepsilon\|_{L^\infty(B_{r_0})} \leq C \left(\frac{1 + \|\nabla f_\varepsilon\|_{L^h(B_{R_0})}^{\sigma_2}}{(R_0 - r_0)^{\sigma_2}} \right) \left(\|\nabla u_\varepsilon\|_{L^p(B_{R_0})}^{\sigma_1} + 1 \right), \tag{5.3}$$

where $C = C(N, p, h, \delta) > 1$ and $\sigma_i = \sigma_i(N, p, h) > 0$, for $i = 1, 2$.

5.1 Proof of Proposition 5.1: the homogeneous case

In this subsection, we assume that $f = 0$ and $\delta = 1$.

For simplicity, we assume throughout the proof that $N \geq 3$, so in this case the Sobolev exponent 2^* is finite. The case $N = 2$ can be treated with minor modifications and is left to the reader. For ease of readability, we divide the proof into four steps.

Step 1: a first iterative scheme We add on both sides of inequality (4.3) the term

$$\int |\nabla \eta|^2 \, |u_{x_k}|^{2q+p} \, dx.$$

We thus obtain

$$\int \left| \nabla \left(\left(|u_{x_k}|^{q+\frac{p-2}{2}} u_{x_k} \right) \eta \right) \right|^2 dx \le C q^5 \sum_{i,j=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) |u_{x_j}|^{2q+2} |\nabla \eta|^2 dx$$

$$+ C q^5 \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) |u_{x_k}|^{2q+2} |\nabla \eta|^2 dx$$

$$+ C \int |\nabla \eta|^2 |u_{x_k}|^{2q+p} dx.$$

An application of the Sobolev inequality leads to

$$\left(\int |u_{x_k}|^{\frac{2^*}{2}(2q+p)} \eta^{2^*} dx \right)^{\frac{2}{2^*}} \le C q^5 \sum_{i,j=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) |u_{x_j}|^{2q+2} |\nabla \eta|^2 dx$$

$$+ C q^5 \sum_{i=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) |u_{x_k}|^{2q+2} |\nabla \eta|^2 dx$$

$$+ C \int |\nabla \eta|^2 |u_{x_k}|^{2q+p} dx.$$

We now sum over $k = 1, \ldots, N$ and use that by the Minkowski inequality,

$$\sum_{k=1}^{N} \left(\int |u_{x_k}|^{\frac{2^*}{2}(2q+p)} \eta^{2^*} dx \right)^{\frac{2}{2^*}} = \sum_{k=1}^{N} \left\| |u_{x_k}|^{2q+p} \eta^2 \right\|_{L^{\frac{2^*}{2}}} \ge \left\| \sum_{k=1}^{N} |u_{x_k}|^{2q+p} \eta^2 \right\|_{L^{\frac{2^*}{2}}}.$$

This implies

$$\left(\int \left| \sum_{k=1}^{N} |u_{x_k}|^{2q+p} \right|^{\frac{2^*}{2}} \eta^{2^*} dx \right)^{\frac{2}{2^*}} \le C q^5 \sum_{i,k=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) |u_{x_k}|^{2q+2} |\nabla \eta|^2 dx$$

$$+ C \int |\nabla \eta|^2 \sum_{k=1}^{N} |u_{x_k}|^{2q+p} dx. \tag{5.4}$$

We now introduce the function

$$\mathcal{U}(x) := \max_{k=1,\ldots,N} |u_{x_k}(x)|.$$

We use that

$$\mathcal{U}^{2q+p} \le \sum_{k=1}^{N} |u_{x_k}|^{2q+p} \le N \mathcal{U}^{2q+p},$$

and also that $g_{i,\varepsilon}''(u_{x_i}) |u_{x_k}|^{2q+2} \le C \mathcal{U}^{2q+p} + \varepsilon \mathcal{U}^{2q+2}$ for every $1 \le i, k \le N$. This yields

$$\left(\int \mathcal{U}^{\frac{2^*}{2}(2q+p)} \eta^{2^*} \right)^{\frac{2}{2^*}} \le C q^5 \int \mathcal{U}^{2q+p} |\nabla \eta|^2 dx + C q^5 \varepsilon \int \mathcal{U}^{2q+2} |\nabla \eta|^2 dx$$

for a possibly different $C = C(N, p) > 1$. By using that $\mathcal{U}^{2q+2} \leq 1 + \mathcal{U}^{2q+p}$, we obtain (for $\varepsilon < 1$)

$$\left(\int \mathcal{U}^{\frac{2^*}{2}(2q+p)} \eta^{2^*} \, dx \right)^{\frac{2}{2^*}} \leq C \, q^5 \int |\nabla \eta|^2 \left(\mathcal{U}^{2q+p} + 1 \right) dx. \tag{5.5}$$

We fix two concentric balls $B_r \subset B_R \Subset B$ and $0 < r < R \leq 1$. Let us assume for simplicity that all the balls are centered at the origin. Then for every pair of radius $r \leq t < s \leq R$ we take in (5.5) a standard cut-off function

$$\eta \in C_0^\infty(B_s), \quad \eta \equiv 1 \text{ on } B_t, \quad 0 \leq \eta \leq 1, \quad \|\nabla \eta\|_{L^\infty} \leq \frac{C}{s-t}. \tag{5.6}$$

This yields

$$\left(\int_{B_t} \mathcal{U}^{\frac{2^*}{2}(2q+p)} \, dx \right)^{\frac{2}{2^*}} \leq C \frac{q^5}{(s-t)^2} \int_{B_s} \left(\mathcal{U}^{2q+p} + 1 \right) dx. \tag{5.7}$$

We define the sequence of exponents

$$\gamma_j = p + 2^{j+2} - 2, \quad j \in \mathbb{N},$$

and take in (5.7) $q = 2^{j+1} - 1$. This gives

$$\left(\int_{B_t} \mathcal{U}^{\frac{2^*}{2}\gamma_j} \, dx \right)^{\frac{2}{2^*}} \leq C \frac{2^{5j}}{(s-t)^2} \int_{B_s} \left(\mathcal{U}^{\gamma_j} + 1 \right) dx, \tag{5.8}$$

for a possibly different constant $C = C(N, p) > 1$.

Step 2: filling the gaps We now observe that

$$\gamma_{j-1} < \gamma_j < \frac{2^*}{2} \gamma_j, \quad \text{for every } j \in \mathbb{N} \setminus \{0\}.$$

By interpolation in Lebesgue spaces, we obtain

$$\int_{B_t} \mathcal{U}^{\gamma_j} \, dx \leq \left(\int_{B_t} \mathcal{U}^{\gamma_{j-1}} \, dx \right)^{\frac{\tau_j \gamma_j}{\gamma_{j-1}}} \left(\int_{B_t} \mathcal{U}^{\frac{2^*}{2}\gamma_j} \, dx \right)^{\frac{(1-\tau_j)2}{2^*}}$$

where $0 < \tau_j < 1$ is given by

$$\tau_j = \frac{\frac{2^*}{2} - 1}{\frac{2^*}{2} \frac{\gamma_j}{\gamma_{j-1}} - 1}.$$

We now rely on (5.8) to get

$$\int_{B_t} \mathcal{U}^{\gamma_j} \, dx \leq \left(\int_{B_t} \mathcal{U}^{\gamma_{j-1}} \, dx \right)^{\frac{\tau_j \gamma_j}{\gamma_{j-1}}} \left(C \frac{2^{5j}}{(s-t)^2} \int_{B_s} \left(\mathcal{U}^{\gamma_j} + 1 \right) dx \right)^{1-\tau_j}$$

$$= \left[\left(C \frac{2^{5j}}{(s-t)^2} \right)^{\frac{1-\tau_j}{\tau_j}} \left(\int_{B_t} \mathcal{U}^{\gamma_{j-1}} \, dx \right)^{\frac{\gamma_j}{\gamma_{j-1}}} \right]^{\tau_j} \left(\int_{B_s} \left(\mathcal{U}^{\gamma_j} + 1 \right) dx \right)^{1-\tau_j}.$$

The sequence $(\tau_j)_{j \geq 1}$ is decreasing, which implies

$$\tau_j > \lim_{n \to \infty} \tau_n = \frac{1}{2} \frac{2^* - 2}{2^* - 1} =: \underline{\tau} \quad \text{for every } j \in \mathbb{N} \setminus \{0\}.$$

Hence,

$$\frac{1 - \tau_j}{\tau_j} \leq \frac{1 - \underline{\tau}}{\underline{\tau}} =: \beta.$$

Using that $s \leq R \leq 1$ and $C > 1$, this implies that

$$\left(C \, \frac{2^{5j}}{(s-t)^2} \right)^{\frac{1-\tau_j}{\tau_j}} \leq \left(C \, \frac{2^{5j}}{(s-t)^2} \right)^{\beta}.$$

By Young's inequality,

$$\int_{B_t} \mathcal{U}^{\gamma_j} \, dx \leq (1 - \tau_j) \int_{B_s} \left(\mathcal{U}^{\gamma_j} + 1 \right) dx + \tau_j \left(C \, \frac{2^{5j}}{(s-t)^2} \right)^{\beta} \left(\int_{B_t} \mathcal{U}^{\gamma_{j-1}} \, dx \right)^{\frac{\gamma_j}{\gamma_{j-1}}}$$

$$\leq (1 - \underline{\tau}) \int_{B_s} \mathcal{U}^{\gamma_j} \, dx + C \, \frac{2^{5j\beta}}{(s-t)^{2\beta}} \left(\int_{B_R} \mathcal{U}^{\gamma_{j-1}} \, dx \right)^{\frac{\gamma_j}{\gamma_{j-1}}} + |B_R|.$$

By applying Lemma 2.5 with

$$Z(t) = \int_{B_t} \mathcal{U}^{\gamma_j} \, dx, \qquad \alpha_0 = 2 \, \beta, \qquad \text{and} \qquad \vartheta = 1 - \underline{\tau},$$

we finally obtain

$$\int_{B_r} \mathcal{U}^{\gamma_j} \, dx \leq C \left(2^{5j\beta} \, (R - r)^{-2\beta} \left(\int_{B_R} \mathcal{U}^{\gamma_{j-1}} \, dx \right)^{\frac{\gamma_j}{\gamma_{j-1}}} + 1 \right), \tag{5.9}$$

for some $C = C(N, p) > 1$.

Step 3: Moser's iteration We now want to iterate the previous estimate on a sequence of shrinking balls. We fix two radii $0 < r < R \leq 1$, then we consider the sequence

$$R_j = r + \frac{R - r}{2^{j-1}}, \qquad j \in \mathbb{N} \setminus \{0\},$$

and we apply (5.9) with $R_{j+1} < R_j$ instead of $r < R$. Thus we get

$$\int_{B_{R_{j+1}}} \mathcal{U}^{\gamma_j} \, dx \leq C \left(2^{7j\beta} \, (R - r)^{-2\beta} \left(\int_{B_{R_j}} \mathcal{U}^{\gamma_{j-1}} \, dx \right)^{\frac{\gamma_j}{\gamma_{j-1}}} + 1 \right) \tag{5.10}$$

where the constant $C > 1$ only depends on N and p.

We introduce the notation

$$Y_j = \int_{B_{R_j}} \mathcal{U}^{\gamma_{j-1}} \, dx,$$

thus (5.10) rewrites as

$$Y_{j+1} \leq C \left(2^{7j\beta} \, (R-r)^{-2\beta} Y_j^{\frac{\gamma_j}{\gamma_{j-1}}} + 1 \right) \leq 2 \, C \, 2^{7j\beta} \, (R-r)^{-2\beta} \, (Y_j + 1)^{\frac{\gamma_j}{\gamma_{j-1}}}.$$

Here, we have used again that $R \leq 1$, so that the term multiplying Y_j is larger than 1. By iterating the previous estimate starting from $j = 1$ and using some standard manipulations, we obtain

$$Y_{n+1} \leq \left(C \, 2^{7\beta} \, (R - r)^{-2\beta}\right)^{\sum_{j=0}^{n-1}(n-j)\frac{\gamma_n}{\gamma_{n-j}}} \left[Y_1 + 1\right]^{\frac{\gamma_n}{\gamma_0}},$$

possibly for a different constant $C = C(N, p) > 1$. We now take the power $1/\gamma_n$ on both sides:

$$Y_{n+1}^{\frac{1}{\gamma_n}} \leq \left(C \, 2^{7\beta} \, (R - r)^{-2\beta}\right)^{\sum_{j=0}^{n-1}\frac{n-j}{\gamma_{n-j}}} \left[Y_1 + 1\right]^{\frac{1}{\gamma_0}}$$

$$= \left(C \, 2^{7\beta} \, (R - r)^{-2\beta}\right)^{\sum_{j=1}^{n}\frac{j}{\gamma_j}} \left[Y_1 + 1\right]^{\frac{1}{\gamma_0}}.$$

We observe that $\gamma_j \sim 2^{j+2}$ as j goes to ∞. This implies the convergence of the series above and we thus get

$$\|\mathcal{U}\|_{L^\infty(B_r)} = \lim_{n\to\infty} \left(\int_{B_{R_{n+1}}} \mathcal{U}^{\gamma_{n+1}} \, dx\right)^{\frac{1}{\gamma_{n+1}}} \leq C \, (R - r)^{-\beta'} \left(\int_{B_R} \mathcal{U}^{p+2} \, dx + 1\right)^{\frac{1}{p+2}},$$

for some $C = C(N, p) > 1$ and $\beta' = \beta'(N, p) > 0$. We also used that $\gamma_0 = p + 2$. By recalling the definition of $\mathcal{U}$, we finally obtain

$$\|\nabla u\|_{L^\infty(B_r)} \leq C \, (R - r)^{-\beta'} \left(\int_{B_R} |\nabla u|^{p+2} \, dx + 1\right)^{\frac{1}{p+2}}. \tag{5.11}$$

Step 4 $L^\infty - L^p$ estimate We fix two concentric balls $B_{r_0} \subset B_{R_0} \Subset B$ with $R_0 \leq 1$. Then for every $r_0 \leq t < s \leq R_0$ from (5.11) we have

$$\|\nabla u\|_{L^\infty(B_t)} \leq \frac{C}{(s - t)^{\beta'}} \left(\int_{B_s} |\nabla u|^{p+2} \, dx\right)^{\frac{1}{p+2}} + \frac{C}{(s - t)^{\beta'}},$$

where we also used the subadditivity of $\tau \mapsto \tau^{1/(p+2)}$. We now observe that

$$\frac{C}{(s - t)^{\beta'}} \left(\int_{B_s} |\nabla u|^{p+2} \, dx\right)^{\frac{1}{p+2}} \leq \frac{C}{(s - t)^{\beta'}} \left(\int_{B_s} |\nabla u|^p \, dx\right)^{\frac{1}{p+2}} \|\nabla u\|_{L^\infty(B_s)}^{\frac{2}{p+2}}$$

$$\leq \frac{2}{p + 2} \|\nabla u\|_{L^\infty(B_s)}$$

$$+ \frac{p}{p + 2} \left(\frac{C}{(s - t)^{\beta'}}\right)^{\frac{p+2}{p}} \left(\int_{B_s} |\nabla u|^p \, dx\right)^{\frac{1}{p}}.$$

We can apply again Lemma 2.5, this time with the choices

$$Z(t) = \|\nabla u\|_{L^\infty(B_t)}, \quad \mathcal{A} = \frac{p}{p + 2} \, C^{\frac{p+2}{p}} \left(\int_{B_{R_0}} |\nabla u|^p \, dx\right)^{\frac{1}{p}}, \quad \alpha_0 = \frac{p + 2}{p \, \beta'}, \quad \beta_0 = \beta'.$$

This yields

$$\|\nabla u\|_{L^\infty(B_{r_0})} \leq C \left[\frac{1}{(R_0 - r_0)^{\beta' \frac{p+2}{p}}} \left(\int_{B_{R_0}} |\nabla u|^p \, dx\right)^{\frac{1}{p}} + \frac{1}{(R_0 - r_0)^{\beta'}}\right],$$

for every $R_0 \leq 1$. This readily implies the desired estimate (5.3) in the homogeneous case. $\square$

5.2 Proof of Proposition 5.1: the non-homogeneous case

We follow step by step the proof of the homogeneous case and we only indicate the main changes, which essentially occur in **Step 1** and **Step 2**.
Step 1: a first iterative scheme This time, we add on both sides of inequality (4.6) the term

$$\int |\nabla \eta|^2 \, (|u_{x_k}| - \delta_k)_+^p \, |u_{x_k}|^{2q} \, dx.$$

Then the left-hand side is greater, up to a constant, than

$$\int \left| \nabla \left((|u_{x_k}| - \delta_k)_+^{\frac{p}{2}} \, |u_{x_k}|^q \, \eta \right) \right|^2 \, dx.$$

The latter in turn, by the Sobolev inequality is greater, up to a constant, than

$$\left(\int (|u_{x_k}| - \delta_k)_+^{\frac{2^* p}{2}} \, |u_{x_k}|^{2^* q} \, \eta^{2^*} \, dx \right)^{\frac{2}{2^*}}.$$

By summing over $k = 1, \ldots, N$ and using the Minkowski inequality, we obtain the analogue of (5.4), namely

$$\left(\int \left| \sum_{k=1}^{N} (|u_{x_k}| - \delta_k)_+^p \, |u_{x_k}|^{2q} \right|^{\frac{2^*}{2}} \eta^{2^*} \, dx \right)^{\frac{2}{2^*}} \leq C q^5 \sum_{i,k=1}^{N} \int g_{i,\varepsilon}''(u_{x_i}) \, |u_{x_k}|^{2q+2} \, |\nabla \eta|^2 \, dx$$

$$+ C q^5 \sum_{k=1}^{N} \int |\nabla f| \, |u_{x_k}|^{2q+1} \, \eta^2 \, dx$$

$$+ C \int |\nabla \eta|^2 \sum_{k=1}^{N} (|u_{x_k}| - \delta_k)_+^p \, |u_{x_k}|^{2q} \, dx.$$

We now introduce the function

$$\mathcal{U}(x) := \frac{1}{2\,\delta} \max_{k=1,\ldots,N} |u_{x_k}(x)|,$$

where the parameter δ is defined in (2.1). We use that

$$\sum_{k=1}^{N} (|u_{x_k}| - \delta_k)_+^p \, |u_{x_k}|^{2q} \geq (2\,\delta\,\mathcal{U} - \delta)_+^p \, |2\,\delta\,\mathcal{U}|^{2q} \geq (2\,\delta)^{2q+p} \left(\mathcal{U} - \frac{1}{2} \right)_+^p \mathcal{U}^{2q},$$

and also that for every $1 \leq i \leq N$,

$$g_{i,\varepsilon}''(u_{x_i}) = (p-1)\,(|u_{x_i}| - \delta_i)_+^{p-2} + \varepsilon \leq C\,\delta^{p-2}\,\mathcal{U}^{p-2} + \varepsilon.$$

This yields

$$\left(\int \left(u - \frac{1}{2} \right)_+^{\frac{2^*}{2} p} u^{2^* q} \, \eta^{2^*} \, dx \right)^{\frac{2}{2^*}} \leq C \, q^5 \int u^{2q+p} \, |\nabla \eta|^2 \, dx$$

$$+ C \, q^5 \varepsilon \int u^{2q+2} \, |\nabla \eta|^2 \, dx$$

$$+ C \, q^5 \int |\nabla f| u^{2q+1} \, \eta^2 \, dx$$

for a possibly different $C = C(N, p, \delta) > 1$.

With the concentric balls $B_r \subset B_t \subset B_s \subset B_R$ and the function η as defined in (5.6), an application of Hölder's inequality leads to

$$\left(\int_{B_t} \left(u - \frac{1}{2} \right)_+^{\frac{2^*}{2} p} u^{2^* q} \, dx \right)^{\frac{2}{2^*}} \leq C \, \frac{q^5}{(s-t)^2} \int_{B_s} u^{2q+p} \, dx$$

$$+ C \, \frac{q^5}{(s-t)^2} \, \varepsilon \int_{B_s} u^{2q+2} \, dx \qquad (5.12)$$

$$+ C \, q^5 \, \|\nabla f\|_{L^h(B_R)} \left(\int_{B_s} u^{(2q+1) \, h'} \, dx \right)^{\frac{1}{h'}}.$$

From now on, we assume that

$$q \geq \max \left\{ \frac{p - 2 \, h'}{2 \, (h' - 1)}, \, \frac{2^* \, p}{2 \, h'} - 1 \right\}. \qquad (5.13)$$

This in particular implies that

$$2 \, q + 2 \leq 2 \, q + p \leq (2 \, q + 2) \, h',$$

then by using Hölder's inequality and taking into account that $s \leq 1$, we get

$$\left(\int_{B_t} \left(u - \frac{1}{2} \right)_+^{\frac{2^*}{2} p} u^{2^* q} \, dx \right)^{\frac{2}{2^*}} \leq C \, \frac{q^5}{(s-t)^2} \left(\int_{B_s} u^{(2q+2) \, h'} \, dx \right)^{\frac{2q+p}{(2q+2) \, h'}}$$

$$+ C \, \frac{q^5}{(s-t)^2} \, \varepsilon \left(\int_{B_s} u^{(2q+2) \, h'} \, dx \right)^{\frac{1}{h'}}$$

$$+ C \, q^5 \, \|\nabla f\|_{L^h(B_R)} \left(\int_{B_s} u^{(2q+2) \, h'} \, dx \right)^{\frac{2q+1}{(2q+2) \, h'}}.$$

Thanks to the relation on the exponents, this gives (recall that $\varepsilon < 1$ and $s \leq 1$)

$$\left(\int_{B_t} \left(u - \frac{1}{2} \right)_+^{\frac{2^*}{2} p} u^{2^* q} \, dx \right)^{\frac{2}{2^*}} \leq \frac{C \, q^5}{(s-t)^2} \left(1 + \|\nabla f\|_{L^h(B_R)} \right)$$

$$\times \left(\int_{B_s} u^{(2q+2) \, h'} \, dx + 1 \right)^{\frac{2q+p}{(2q+2) \, h'}}. \qquad (5.14)$$

We now estimate

$$\int_{B_s} \mathcal{U}^{(2q+2)\,h'}\,dx = \int_{B_s \cap \{\mathcal{U} \geq 1\}} \mathcal{U}^{(2q+2)\,h'}\,dx + \int_{B_s \cap \{\mathcal{U} \leq 1\}} \mathcal{U}^{(2q+2)\,h'}\,dx$$

$$\leq \int_{B_s \cap \{\mathcal{U} \geq 1\}} \mathcal{U}^{(2q+2)\,h'}\,dx + C.$$

Observe that on the set $\{\mathcal{U} \geq 1\}$, we have $\mathcal{U} \leq 2\,(\mathcal{U} - 1/2)_+$. Hence,

$$\int_{B_s} \mathcal{U}^{(2q+2)\,h'}\,dx \leq C \int_{B_s} \left(\mathcal{U} - \frac{1}{2}\right)_+^{\frac{2^*}{2}\,p} \mathcal{U}^{(2q+2)\,h' - \frac{2^*}{2}\,p}\,dx + C, \tag{5.15}$$

where the exponent $(2q+2)\,h' - (2^* p)/2$ is positive, thanks to the choice (5.13) of q. We deduce from (5.14) that

$$\left(\int_{B_t} \left(\mathcal{U} - \frac{1}{2}\right)_+^{\frac{2^*}{2}\,p} \mathcal{U}^{2^* q}\,dx\right)^{\frac{2}{2^*}} \leq \frac{C\,q^5}{(s-t)^2}\left(1 + \|\nabla f\|_{L^h(B_R)}\right)$$

$$\times \left(\int_{B_s} \left(\mathcal{U} - \frac{1}{2}\right)_+^{\frac{2^*}{2}\,p} \mathcal{U}^{(2q+2)\,h' - \frac{2^*}{2}\,p}\,dx + 1\right)^{\frac{2q+p}{(2q+2)\,h'}}, \tag{5.16}$$

for a constant $C = C(N, p, h, \delta) > 1$. We now take $q = 2^{j+1} - 1$ for $j \geq j_0 - 1$, where $j_0 \in \mathbb{N}$ is chosen so as to ensure condition (5.13). Then we define the sequence of positive exponents

$$\gamma_j = (2q+2)\,h' - \frac{2^*}{2}\,p = 2^{j+2}\,h' - \frac{2^*}{2}\,p, \qquad j \geq j_0,$$

and

$$\widehat{\gamma}_j = 2^*\,q = 2^*\,(2^{j+1} - 1), \qquad j \geq j_0.$$

In order to simplify the notation, we also introduce the absolutely continuous measure

$$d\mu := \left(\mathcal{U} - \frac{1}{2}\right)_+^{\frac{2^*}{2}\,p}\,dx.$$

From (5.16), we get

$$\left(\int_{B_t} \mathcal{U}^{\widehat{\gamma}_j}\,d\mu\right)^{\frac{2}{2^*}} \leq \frac{C\,2^{5j}}{(s-t)^2}\left(1 + \|\nabla f\|_{L^h(B_R)}\right)\left(\int_{B_s} \mathcal{U}^{\gamma_j}\,d\mu + 1\right)^{\frac{2}{2^*}\frac{\widehat{\gamma}_j + \frac{2^*}{2}\,p}{\gamma_j + \frac{2^*}{2}\,p}}.$$

We now observe that $h > N/2$ implies $h' < 2^*/2$. By recalling that $p \geq 2$, we thus have $2\,h' < (2^* p)/2$, which in turn implies

$$\frac{\widehat{\gamma}_j}{\gamma_j} \geq \frac{2^*}{2\,h'} > 1, \qquad j \geq j_0. \tag{5.17}$$

It follows that

$$\frac{\widehat{\gamma}_j + \frac{2^*}{2}\,p}{\gamma_j + \frac{2^*}{2}\,p} \leq \frac{\widehat{\gamma}_j}{\gamma_j}.$$

Hence, we obtain

$$\left(\int_{B_t} \mathcal{U}^{\widehat{\gamma}_j} \, d\mu \right)^{\frac{2}{2^*}} \leq \frac{C \, 2^{5j}}{(s-t)^2} \left(1 + \|\nabla f\|_{L^h(B_R)} \right) \left(\int_{B_s} \mathcal{U}^{\gamma_j} \, d\mu + 1 \right)^{\frac{2}{2^*} \frac{\widehat{\gamma}_j}{\gamma_j}}. \tag{5.18}$$

Step 2: filling the gaps Since

$$\gamma_{j-1} < \gamma_j < \widehat{\gamma}_j, \qquad \text{for every } j \geq j_0 + 1,$$

we obtain by interpolation in Lebesgue spaces,

$$\int_{B_t} \mathcal{U}^{\gamma_j} \, d\mu \leq \left(\int_{B_t} \mathcal{U}^{\gamma_{j-1}} \, d\mu \right)^{\frac{\tau_j \gamma_j}{\gamma_{j-1}}} \left(\int_{B_t} \mathcal{U}^{\widehat{\gamma}_j} \, d\mu \right)^{\frac{(1-\tau_j)\gamma_j}{\widehat{\gamma}_j}},$$

where $0 < \tau_j < 1$ is given by

$$\tau_j = \frac{\dfrac{\widehat{\gamma}_j}{\gamma_j} - 1}{\dfrac{\widehat{\gamma}_j}{\gamma_j} \dfrac{\gamma_j}{\gamma_{j-1}} - 1}. \tag{5.19}$$

We now rely on (5.18) to get

$$\int_{B_t} \mathcal{U}^{\gamma_j} \, d\mu \leq \left(\int_{B_t} \mathcal{U}^{\gamma_{j-1}} \, d\mu \right)^{\frac{\tau_j \gamma_j}{\gamma_{j-1}}}$$

$$\times \left[\left(C \frac{2^{5j}}{(s-t)^2} (1 + \|\nabla f\|_{L^h(B_R)}) \right)^{\frac{2^* \gamma_j}{2 \widehat{\gamma}_j}} \left(\int_{B_s} \mathcal{U}^{\gamma_j} \, d\mu + 1 \right) \right]^{1-\tau_j}$$

$$= \left[\left(C \frac{2^{5j}}{(s-t)^2} (1 + \|\nabla f\|_{L^h(B_R)}) \right)^{\frac{2^* \gamma_j (1-\tau_j)}{2 \widehat{\gamma}_j \tau_j}} \left(\int_{B_t} \mathcal{U}^{\gamma_{j-1}} \, d\mu \right)^{\frac{\gamma_j}{\gamma_{j-1}}} \right]^{\tau_j}$$

$$\times \left(\int_{B_s} \mathcal{U}^{\gamma_j} \, d\mu + 1 \right)^{1-\tau_j}. \tag{5.20}$$

We claim that

$$\tau_j \geq \underline{\tau} := \frac{2^* - 2h'}{4 \cdot 2^* - 2h'} \qquad \text{for every } j \geq j_0 + 1. \tag{5.21}$$

We already know by (5.17) that $(\widehat{\gamma}_j / \gamma_j) \geq 2^*/(2h')$. Moreover, relying on the fact that $(2^* p)/2 \leq 2^{j_0} h'$ (this follows from the definition of j_0), we also have

$$2 \leq \frac{\gamma_j}{\gamma_{j-1}} \leq 4, \qquad j \geq j_0 + 1.$$

By recalling the definition (5.19) of τ_j, we get

$$\tau_j = \zeta \left(\frac{\widehat{\gamma}_j}{\gamma_j}, \frac{\gamma_j}{\gamma_{j-1}} \right), \qquad \text{where } \zeta(x, y) = \frac{x-1}{x\,y-1}.$$

Observe that on $[2^*/(2h'), +\infty) \times [2, 4]$, the function $x \mapsto \zeta(x, y)$ is increasing, while $y \mapsto \zeta(x, y)$ is decreasing. Thus we get

$$\tau_j \geq \zeta\left(\frac{2^*}{2h'}, 4\right),$$

which is exactly claim (5.21). We deduce from (5.21) and (5.17) that

$$\frac{2^* \gamma_j (1 - \tau_j)}{2 \widehat{\gamma}_j \tau_j} \leq \frac{1 - \underline{\tau}}{\underline{\tau}} h' =: \beta.$$

In particular, we have

$$\left(C \frac{2^{5j}}{(s-t)^2}(1 + \|\nabla f\|_{L^h(B_R)})\right)^{\frac{2^* \gamma_j (1-\tau_j)}{2\widehat{\gamma}_j \tau_j}} \leq \left(C \frac{2^{5j}}{(s-t)^2}(1 + \|\nabla f\|_{L^h(B_R)})\right)^{\beta},$$

since the quantity inside the parenthesis is larger than 1 (here, we use again that $s \leq 1$). In view of (5.20), this implies

$$\int_{B_t} \mathcal{U}^{\gamma_j} \, d\mu \leq \left[\left(C \frac{2^{5j}}{(s-t)^2}(1 + \|\nabla f\|_{L^h(B_R)})\right)^{\beta} \left(\int_{B_t} \mathcal{U}^{\gamma_{j-1}} \, d\mu\right)^{\frac{\gamma_j}{\gamma_{j-1}}}\right]^{\tau_j}$$
$$\times \left(\int_{B_s} \mathcal{U}^{\gamma_j} \, d\mu + 1\right)^{1-\tau_j}.$$

By Young's inequality,

$$\int_{B_t} \mathcal{U}^{\gamma_j} \, d\mu \leq (1 - \tau_j) \left(\int_{B_s} \mathcal{U}^{\gamma_j} \, d\mu + 1\right)$$
$$+ \tau_j \left(C \frac{2^{5j}}{(s-t)^2}(1 + \|\nabla f\|_{L^h(B_R)})\right)^{\beta} \left(\int_{B_t} \mathcal{U}^{\gamma_{j-1}} \, d\mu\right)^{\frac{\gamma_j}{\gamma_{j-1}}}$$
$$\leq (1 - \underline{\tau}) \int_{B_s} \mathcal{U}^{\gamma_j} \, d\mu$$
$$+ C \frac{2^{5j\beta}}{(s-t)^{2\beta}} (1 + \|\nabla f\|_{L^h(B_R)})^{\beta} \left(\int_{B_R} \mathcal{U}^{\gamma_{j-1}} \, d\mu\right)^{\frac{\gamma_j}{\gamma_{j-1}}} + 1,$$

where $C = C(N, p, h, \delta) > 1$ as usual. By applying again Lemma 2.5, this times with the choices

$$Z(t) = \int_{B_t} \mathcal{U}^{\gamma_j} \, d\mu, \qquad \alpha_0 = 2\beta, \qquad \text{and} \qquad \vartheta = 1 - \underline{\tau},$$

we finally obtain

$$\int_{B_r} \mathcal{U}^{\gamma_j} \, d\mu \leq C \frac{2^{5j\beta}}{(R-r)^{2\beta}} (1 + \|\nabla f\|_{L^h(B_R)})^{\beta} \left(\int_{B_R} \mathcal{U}^{\gamma_{j-1}} \, d\mu\right)^{\frac{\gamma_j}{\gamma_{j-1}}} + C. \qquad (5.22)$$

Step 3: Moser's iteration Estimate (5.22) is the analogue of (5.9), except that the Lebesgue measure dx is now replaced by the measure $d\mu$, and the index j is assumed to be larger than

some $j_0 + 1$, instead of $j \geq 0$ as in (5.9). Following the same iteration argument and starting from $j = j_0 + 1$, we are led to

$$\|\mathcal{U}\|_{L^\infty(B_r, d\mu)} \leq C \left(\frac{1 + \|\nabla f\|_{L^h(B_R)}}{R - r} \right)^{\beta'} \left(\int_{B_R} \mathcal{U}^{\gamma_{j_0}} \, d\mu + 1 \right)^{\frac{1}{\gamma_{j_0}}}, \tag{5.23}$$

for some $C = C(N, p, h, \delta) > 1$, $\beta' = \beta'(N, p, h) > 0$.

Step 4 $L^\infty - L^p$ **estimate** We now want to replace the norm $L^{\gamma_{j_0}}(B_R, d\mu)$ of $\mathcal{U}$ in the right-hand side of (5.23) by its norm $L^p(B_R, dx)$. Let $q_1 := 2^{j_1 + 1} - 1$ where

$$j_1 := \min \left\{ j \geq j_0 : j + 1 \geq \log_2 \left(1 + \frac{\gamma_{j_0}}{2^*} \right) \right\}.$$

Then $\gamma_{j_0} \leq 2^* q_1$ and thus, by using that

$$\mathcal{U}^{\gamma_{j_0}} \leq 2^{2^* q_1 - \gamma_{j_0}} \mathcal{U}^{2^* q_1}, \qquad \text{whenever } \mathcal{U} \geq \frac{1}{2},$$

we have

$$\|\mathcal{U}\|_{L^{\gamma_{j_0}}(B_R, d\mu)} \leq C \|\mathcal{U}\|_{L^{2^* q_1}(B_R, d\mu)}^{\frac{2^* q_1}{\gamma_{j_0}}}. \tag{5.24}$$

We rely on (5.14) with $q = q_1$ to get for every $0 < r < t < s < R$

$$\|\mathcal{U}\|_{L^{2^* q_1}(B_t, d\mu)}^{2 q_1} \leq \frac{C}{(s - t)^2} \left(1 + \|\nabla f\|_{L^h(B_R)} \right) \left(\|\mathcal{U}\|_{L^{2(q_1 + 1) h'}(B_s)}^{2 q_1 + p} + 1 \right), \tag{5.25}$$

for some new constant $C = C(N, p, h, \delta) > 1$.

Since $j_1 \geq j_0$, we have $p < (2 q_1 + 2) h' < (2 q_1 + p) \frac{2^*}{2}$, and thus, by interpolation in Lebesgue spaces

$$\|\mathcal{U}\|_{L^{2(q_1 + 1) h'}(B_s)} \leq \|\mathcal{U}\|_{L^{2^* q_1 + \frac{2^*}{2} p}(B_s)}^{\theta} \|\mathcal{U}\|_{L^p(B_s)}^{1 - \theta}, \tag{5.26}$$

where $\theta \in (0, 1)$ is determined as usual by scale invariance. As in the proof of (5.15), we have

$$\|\mathcal{U}\|_{L^{2^* q_1 + \frac{2^*}{2} p}(B_s)} \leq C \|\mathcal{U}\|_{L^{2^* q_1}(B_s, d\mu)}^{\frac{2 q_1}{2 q_1 + p}} + C.$$

Inserting this last estimate into (5.26), we obtain

$$\|\mathcal{U}\|_{L^{2(q_1 + 1) h'}(B_s)}^{2 q_1 + p} \leq C \|\mathcal{U}\|_{L^{2^* q_1}(B_s, d\mu)}^{2 q_1 \theta} \|\mathcal{U}\|_{L^p(B_s)}^{(1 - \theta)(2 q_1 + p)} + C \|\mathcal{U}\|_{L^p(B_s)}^{(1 - \theta)(2 q_1 + p)},$$

up to changing the constant $C = C(N, p, h, \delta) > 1$. In view of (5.25), this gives

$$\|\mathcal{U}\|_{L^{2^* q_1}(B_t, d\mu)}^{2 q_1} \leq \frac{C}{(s - t)^2} \left(1 + \|\nabla f\|_{L^h(B_R)} \right)$$
$$\times \left(\|\mathcal{U}\|_{L^{2^* q_1}(B_s, d\mu)}^{2 q_1 \theta} \|\mathcal{U}\|_{L^p(B_s)}^{(1 - \theta)(2 q_1 + p)} + \|\mathcal{U}\|_{L^p(B_s)}^{(1 - \theta)(2 q_1 + p)} + 1 \right).$$

By Young's inequality, we get

$$\|\mathcal{U}\|_{L^{2^* q_1}(B_t, d\mu)}^{2 q_1} \leq \theta \|\mathcal{U}\|_{L^{2^* q_1}(B_s, d\mu)}^{2 q_1}$$
$$+ (1 - \theta) \left(\frac{C}{(s - t)^2} (1 + \|\nabla f\|_{L^h(B_R)}) \right)^{\frac{1}{1 - \theta}} \|\mathcal{U}\|_{L^p(B_R)}^{(2 q_1 + p)}$$
$$+ \frac{C}{(s - t)^2} \left(1 + \|\nabla f\|_{L^h(B_R)} \right) \left(\|\mathcal{U}\|_{L^p(B_R)}^{(1 - \theta)(2 q_1 + p)} + 1 \right).$$

By Lemma 2.5, this implies

$$\|\mathcal{U}\|_{L^{2^* q_1}(B_r, d\mu)}^{2 q_1} \leq C \left(\frac{1}{(R-r)^2} (1 + \|\nabla f\|_{L^h(B_{R_0})}) \right)^{\frac{1}{1-\theta}} \left(\|\mathcal{U}\|_{L^p(B_R)}^{(2 q_1 + p)} + 1 \right),$$

after some standard manipulations. Coming back to (5.23) and taking into account (5.24), we obtain

$$\|\mathcal{U}\|_{L^\infty(B_{r_0}, d\mu)} \leq C \left(\frac{1 + \|\nabla f\|_{L^h(B_{R_0})}}{R_0 - r_0} \right)^{\sigma_2} \left(\|\mathcal{U}\|_{L^p(B_{R_0})}^{\sigma_1} + 1 \right),$$

where $C = C(N, p, h, \delta) > 1$ and $\sigma_i = \sigma_i(N, p, h) > 0$, for $i = 1, 2$. By definition of $\mathcal{U}$, we have

$$|\nabla u| \leq 2 \delta \sqrt{N} \, \mathcal{U} \leq \sqrt{N} \, |\nabla u|.$$

Since $\|\mathcal{U}\|_{L^\infty(B_{r_0}, d\mu)} + 1 \geq \|\mathcal{U}\|_{L^\infty(B_{r_0})}$, it follows that

$$\|\nabla u\|_{L^\infty(B_{r_0})} \leq C \left(\frac{1 + \|\nabla f\|_{L^h(B_{R_0})}}{R_0 - r_0} \right)^{\sigma_2} \left(\|\nabla u\|_{L^p(B_{R_0})}^{\sigma_1} + 1 \right),$$

possibly for a different constant $C = C(N, p, h, \delta) > 1$. This completes the proof. □

Appendix: Lipschitz regularity with a nonlinear lower order term

In this section, we consider the functional

$$\mathfrak{G}_\delta(u, \Omega') = \sum_{i=1}^N \int_{\Omega'} \left[g_i(u_{x_i}) + G(x, u) \right] dx, \qquad \Omega' \Subset \Omega, \ u \in W^{1,p}_{\text{loc}}(\Omega').$$

The lower order term $f u$ of the functional $\mathfrak{F}_\delta$ is thus replaced by a more general term $G(x, u)$. We assume that G is a Carathéodory function and that for almost every $x \in \Omega$, the map

$$\xi \mapsto G(x, \xi) \qquad \text{is } C^1 \text{ and convex.}$$

We denote $f(x, \xi) := G_\xi(x, \xi)$ and we assume that $f \in W^{1,h}_{\text{loc}}(\Omega \times \mathbb{R})$, for some $h > N/2$. Finally, we assume that $G(x, \xi)$ satisfies the inequality

$$|G(x, \xi)| \leq b(x) |u|^\gamma + a(x) \tag{A.1}$$

where $1 < p \leq \gamma < p^*$ and a, b are two non-negative functions belonging respectively to $L^s_{\text{loc}}(\Omega)$ and $L^\sigma_{\text{loc}}(\Omega)$ with $s > N/p$ and $\sigma > p^*/(p^* - \gamma)$.

Under assumption (A.1), all the local minimizers of $\mathfrak{G}_\delta$ are locally bounded, see [11, Theorem 7.5] and moreover, for every such minimizer u, for every $B_{r_0} \Subset B_{R_0} \Subset \Omega$,

$$\|u\|_{L^\infty(B_{r_0})} \leq M,$$

where M depends on $\|u\|_{W^{1,p}(B_{R_0})}$, r_0, R_0, $\|b\|_{L^\sigma(R_0)}$, and $\|a\|_{L^s(B_{R_0})}$.

Then we have:

Theorem A.1 *Let $p \geq 2$ and let $U \in W^{1,p}_{loc}(\Omega)$ be a local minimizer of the functional $\mathfrak{G}_\delta$. Then U is locally Lipschitz in Ω.*

Proof We only explain the main differences with respect to the proof of Theorem 1.1. Since G is convex with respect to the second variable, the functional $\mathfrak{G}$ is still convex. This implies that Lemma 2.3 remains true with the same proof. We then introduce the approximation of G:

$$G_\varepsilon(x, \xi) = \int_{\mathbb{R}^N \times \mathbb{R}} G(x - y, \xi - \zeta)\, \rho_\varepsilon(y)\, \widetilde{\rho}_\varepsilon(\zeta)\, dy\, d\zeta,$$

where ρ_ε is the same regularization kernel as before, while $\widetilde{\rho}_\varepsilon$ is a regularization kernel on $\mathbb{R}$.

Given a local minimizer $U \in W^{1,p}_{loc}(\Omega)$ and a ball $B \subset 2B \Subset \Omega$, there exists a unique C^2 solution u_ε to the regularized problem

$$\min\left\{\mathfrak{G}_\varepsilon(v; B) \;:\; v - U_\varepsilon \in W^{1,p}_0(B)\right\},$$

where

$$\mathfrak{G}_\varepsilon(v; B) = \sum_{i=1}^N \int_B g_{i,\varepsilon}(v_{x_i})\, dx + \int_B G_\varepsilon(x, v)\, dx$$

and $U_\varepsilon = U * \rho_\varepsilon$. Moreover, by [11, Remark 7.6] we have $u_\varepsilon \in L^\infty(B)$, with a bound on the L^∞ norm uniform in $\varepsilon > 0$. In order to simplify the notation, we simply write as usual u and f instead of u_ε and f_ε. The Euler equation is now

$$\sum_{i=1}^N \int g'_{i,\varepsilon}(u_{x_i})\, \varphi_{x_i}\, dx + \int f(x, u)\, \varphi\, dx = 0, \qquad \varphi \in W^{1,p}_0(B).$$

When we differentiate the Euler equation with respect to some direction x_j, we obtain

$$\sum_{i=1}^N \int g''_{i,\varepsilon}(u_{x_i})\, u_{x_i x_j}\, \psi_{x_i}\, dx + \int \left(f_{x_j}(x, u) + f_\xi(x, u)\, u_{x_j}\right) \psi\, dx = 0, \qquad \psi \in W^{1,p}_0(B).$$

We can then repeat the proof of Proposition 5.1 with this additional term $f_\xi(x, u)u_{x_j}$ which leads to the following analogue of (5.12):

$$\left(\int_{B_t} \left(\mathcal{U} - \frac{1}{2}\right)_+^{\frac{2^*}{2}p} \mathcal{U}^{2^* q}\, dx\right)^{\frac{2}{2^*}} \leq C\, \frac{q^5}{(s-t)^2} \int_{B_s} \mathcal{U}^{2q+p}\, dx$$

$$+ C\, \frac{q^5}{(s-t)^2}\, \varepsilon \int_{B_s} \mathcal{U}^{2q+2}\, dx$$

$$+ C\, q^5\, \|\nabla_x f\|_{L^h} \left(\int_{B_s} \mathcal{U}^{(2q+1)h'}\, dx\right)^{\frac{1}{h'}}$$

$$+ C\, q^5\, \|f_\xi\|_{L^h} \left(\int_{B_s} \mathcal{U}^{(2q+2)h'}\, dx\right)^{\frac{1}{h'}}.$$

Using again Hölder's inequality for the first three terms, we obtain inequality (5.14) where $\|\nabla f\|$ now represents the full gradient of f with respect to both x and ξ. The rest of the proof is the same and leads to a uniform Lipschitz estimate, as desired. $\qquad\square$

References

1. Bousquet, P., Brasco, L.: C^1 regularity of orthotropic p-harmonic functions in the plane. Anal. PDE **11**, 813–854 (2018)
2. Bousquet, P., Brasco, L., Julin, V.: Lipschitz regularity for local minimizers of some widely degenerate problems. Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **26**, 1–40 (2016)
3. Brasco, L., Carlier, G.: On certain anisotropic elliptic equations arising in congested optimal transport: local gradient bounds. Adv. Calc. Var. **7**, 379–407 (2014)
4. Brasco, L., Leone, C., Pisante, G., Verde, A.: Sobolev and Lipschitz regularity for local minimizers of widely degenerate anisotropic functionals. Nonlinear Anal. **153**, 169–199 (2017)
5. Brasco, L.: Global L^∞ gradient estimates for solutions to a certain degenerate elliptic equation. Nonlinear Anal. **74**, 516–531 (2011)
6. Demengel, F.: Lipschitz interior regularity for the viscosity and weak solutions of the pseudo p-Laplacian equation. Adv. Differ. Equ. **21**, 373–400 (2016)
7. Duzaar, F., Mingione, G.: Local Lipschitz regularity for degenerate elliptic systems. Ann. Inst. H. Poincaré Anal. Non Linéaire **27**, 1361–1396 (2010)
8. Fonseca, I., Fusco, N.: Regularity results for anisotropic image segmentation models. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **24**, 463–499 (1997)
9. Fonseca, I., Fusco, N., Marcellini, P.: An existence result for a nonconvex variational problem via regularity. ESAIM Control Optim. Calc. Var. **7**, 69–95 (2002)
10. Giaquinta, M.: Growth conditions and regularity, a counterexample. Manuscr. Math. **59**, 245–248 (1987)
11. Giusti, E.: Direct Methods in the Calculus of Variations. World Scientific Publishing Co. Inc., River Edge (2003)
12. Han, Q., Lin, F.: Elliptic Partial Differential Equations, 2nd edn. In: Courant Lecture Notes in Mathematics, Vol. 1. Courant Institute of Mathematical Sciences, New York. AMS, Providence (2011)
13. Kuusi, T., Mingione, G.: Linear potentials in nonlinear potential theory. Arch. Ration. Mech. Anal. **207**, 215–246 (2013)
14. Lindqvist, P., Ricciotti, D.: Regularity for an anisotropic equation in the plane. Nonlinear Anal. (2018). https://doi.org/10.1016/j.na.2018.02.002
15. Marcellini, P.: Un exemple de solution discontinue d'un problème variationnel dans le cas scalaire, preprint n. **11**, Ist. Mat. Univ. Firenze (1987). http://web.math.unifi.it/users/marcell/lavori
16. Moser, J.: A new proof of De Giorgi's theorem concerning the regularity problem for elliptic differential equations. Commun. Pure Appl. Math. **13**, 457–468 (1960)
17. Ural'tseva, N., Urdaletova, N.: The boundedness of the gradients of generalized solutions of degenerate quasilinear nonuniformly elliptic equations. Vest. Leningr. Univ. Math. **16**, 263–270 (1984)