

# Morrey estimates for the gradient in non-linear variational transmission problems

Luca Esposito

*Dipartimento di Matematica, University of Salerno  
Via Giovanni Paolo II, 84084 Fisciano (SA), Italy  
e-mail: luesposi@unisa.it*

Lorenzo Lamberti

*Institute Élie Cartan, Université de Lorraine, CNRS  
F-54000 Nancy, France  
e-mail: lorenzo.lamberti@univ-lorraine.fr*

November 2025

## Abstract

We study a class of variational transmission problems driven by nonlinear energies with discontinuous coefficients across a prescribed interface. The model setting consists of integral functionals of the form

$$\mathcal{F}(u; E) = \int_{\Omega} \sigma_E(x) F(\nabla u) dx,$$

where the coefficient  $\sigma_E$  takes two constant values on complementary regions separated by a  $C^1$  hypersurface, and the integrand  $F$  satisfies standard  $p$ -growth and monotonicity conditions with  $p > 2$ .

In this nonlinear variational framework, we establish local Morrey-space regularity for the gradient of local minimizers, proving that  $\nabla u \in L_{loc}^{2,\lambda}(\Omega)$  for every  $0 \leq \lambda < n$ , provided  $2 < p < \frac{2n}{n-2}$ . The proof is based on quantitative decay estimates for the energy near the interface, first obtained in a flat configuration and then extended to the general case by a suitable approximation argument.

## 1 Introduction

Transmission problems arise naturally in the study of diffusive processes in heterogeneous media, where the governing laws change abruptly across fixed interfaces. Typical applications include elasticity theory, composite materials, and conductivity phenomena. This kind of problems appear also in the study of thermal insulation of bodies under sources and prescribed boundary conditions (see for example [1, 4]) From a mathematical perspective, such problems are characterised by the presence of piecewise-defined operators or energy densities, together with suitable transmission conditions prescribed along the separating interface. The first systematic investigation of transmission problems in elasticity theory dates back to the seminal work of M. Picone in the 1950s (see [17]).

In this paper, we investigate a class of *variational transmission problems* associated with nonlinear integral functionals exhibiting discontinuous coefficients across a prescribed interface. Our focus is on the regularity properties of local minimizers, with particular attention to the behaviour of the gradient near the interface.

Let  $\Omega \subset \mathbb{R}^n$  be a bounded domain and let  $E \subset \Omega$  be a measurable subset. We assume that the interface  $\partial E$  is a hypersurface of class  $C^1$ . We fix two constants  $0 < \alpha < \beta < +\infty$  and define

$$\sigma_E := \beta \mathbb{1}_E + \alpha \mathbb{1}_{E^c},$$

where  $\mathbb{1}_E$  denotes the characteristic function of the set  $E$  and  $E^c$  is the complementary of  $E$ . We consider the integral functional

$$\mathcal{F}(u; E) := \int_{\Omega} \sigma_E(x) F(\nabla u) dx, \tag{1}$$

defined for functions  $u \in W^{1,p}(\Omega)$ .

The integrand  $F: \mathbb{R}^n \rightarrow \mathbb{R}$  is assumed to belong to  $C^2(\mathbb{R}^n)$  and to satisfy standard  $p$ -growth and monotonicity conditions: for all  $\xi, \eta \in \mathbb{R}^n$ ,

$$\langle \nabla F(\xi) - \nabla F(\eta), \xi - \eta \rangle \geq \nu (\mu^2 + |\xi|^2 + |\eta|^2)^{\frac{p-2}{2}} |\xi - \eta|^2, \tag{H1}$$

$$|\nabla F(\xi) - \nabla F(\eta)| \leq L(\mu^2 + |\xi|^2 + |\eta|^2)^{\frac{p-2}{2}} |\xi - \eta|, \quad (\text{H2})$$

for some constants  $\nu, L > 0$ ,  $\mu \in (0, 1]$ , and  $p > 2$ .

Local minimizers of the functional  $\mathcal{F}$  are understood in the usual variational sense.

**Definition 1.1.** *Let  $u \in W^{1,p}(\Omega)$ . We say that  $u$  is a local minimizer of  $\mathcal{F}(\cdot; \Omega)$  in  $\Omega$  if, for every ball  $B_r(x_0) \subset \Omega$  and every  $\phi \in W_0^{1,p}(B_r(x_0))$ , one has*

$$\mathcal{F}(u; B_r(x_0)) \leq \mathcal{F}(u + \phi; B_r(x_0)).$$

Before stating our main result, let us place it in the context of the existing literature. In the linear case  $p = 2$ , transmission problems with piecewise constant coefficients have been extensively studied, and gradient regularity across sufficiently smooth interfaces is by now classical (see, e.g., [12, 13, 15]). Such results follow from the elliptic regularity theory for divergence-form operators with discontinuous coefficients (see for instance [11, 12, 14, 15] and [2, Theorem 7.53]). In this setting, local  $W^{1,2}$ - and Hölder-type regularity for the gradient of minimizers is well understood under mild geometric assumptions on the interface.

More recently, fine regularity properties for linear transmission problems across  $C^{1,\alpha}$  interfaces have been obtained by Caffarelli, Soria-Carro, and Stinga [10].

By contrast, the nonlinear case  $p \neq 2$  remains far less developed. To the best of our knowledge, the only available regularity results for nonlinear transmission problems concern equations subject to *additive* transmission conditions on the interface. In particular, in the recent work [9], BMO regularity of the gradient is obtained for degenerate quasilinear equations under transmission conditions of the form

$$g(|\nabla u_1|) \frac{\partial u_1}{\partial \nu} - g(|\nabla u_2|) \frac{\partial u_2}{\partial \nu} = f \quad \text{on } \partial E.$$

In contrast, the transmission condition naturally associated with the variational functional (1) is of *multiplicative* type, namely

$$g_1(|\nabla u_1|) \frac{\partial u_1}{\partial \nu} = g_2(|\nabla u_2|) \frac{\partial u_2}{\partial \nu} \quad \text{on } \partial E,$$

which arises intrinsically from the Euler–Lagrange equations and expresses the continuity of the nonlinear flux across the interface. We stress that this multiplicative transmission condition is intrinsic to the variational structure of the functional (1), as it arises directly from the Euler–Lagrange equations and cannot be prescribed independently. The presence of a discontinuous coefficient combined with nonlinear growth prevents a direct application of standard techniques such as freezing arguments or difference quotient estimates across the interface. To the best of our knowledge, no Morrey-space regularity results for gradients of local minimizers are available in this genuinely nonlinear variational transmission setting.

Denoting by  $L^{p,\lambda}$  the classical Morrey spaces, we can now state the main result of this paper, which establishes local Morrey regularity for the gradient of local minimizers under minimal geometric assumptions on the interface.

**Theorem 1.2** (Gradient regularity in Morrey spaces). *Let  $u \in W^{1,p}(\Omega)$  be a local minimizer of the functional  $\mathcal{F}(\cdot; \Omega)$  defined in (1), and assume that  $\partial E$  is a hypersurface of class  $C^1$ . Suppose that assumptions (H1)–(H2) are satisfied where  $2 < p < \frac{2n}{n-2}$ . Then, for every  $0 \leq \lambda < n$ ,*

$$\nabla u \in L_{\text{loc}}^{p,\lambda}(\Omega).$$

Moreover, for every  $\Omega' \Subset \Omega$ , there exists a constant  $C = C(n, p, \nu, L, \beta, E, \text{diam}(\Omega), \text{dist}(\Omega', \partial\Omega)) > 0$  such that

$$\|\nabla u\|_{L^{p,\lambda}(\Omega')} \leq C.$$

**Remark 1.3.** *Within the present variational transmission framework, Morrey regularity represents the natural level of regularity that can be expected for the gradient of local minimizers, in view of the discontinuous coefficients and the intrinsic transmission condition.*

The proof of Theorem 1.2 is obtained as a consequence of the following decay estimate for the energy.

**Proposition 1.4.** *Let  $u$  be a local minimizer of the functional  $\mathcal{F}(\cdot; \Omega)$  defined in (1) under the assumption of Theorem 1.2. Then there exists a constant  $0 < \tau_0 < 1$  such that the following holds: for every  $\tau \in (0, \tau_0)$  there exists  $\varepsilon_0 = \varepsilon_0(\tau) > 0$  with the property that, if  $B_r(x_0) \Subset \Omega$  and one of the following conditions is satisfied,*

- (i)  $|E \cap B_r(x_0)| < \varepsilon_0 |B_r|$ ,
- (ii)  $|B_r(x_0) \setminus E| < \varepsilon_0 |B_r|$ ,

(iii) there exists a half-space  $H$  such that

$$\frac{|(E\Delta H) \cap B_r(x_0)|}{|B_r|} < \varepsilon_0,$$

then, for every  $0 < \delta < n$ , the estimate

$$\int_{B_{\tau r}(x_0)} |\nabla u|^p dx \leq C_0 \tau^{n-\delta} \left( \int_{B_r(x_0)} |\nabla u|^p dx + r^n \right)$$

holds, where the constant  $C_0$  depends only on  $n, \nu, L, \alpha, \beta, \delta$  and  $\|\nabla u\|_{L^2(\Omega)}$ .

**Remark 1.5.** For the proof of Theorem 1.2, it would be sufficient to state the proposition only in case (i). We nevertheless include the other two cases for the sake of completeness.

We give an outline of the proof of Proposition 1.4. It relies on decay estimates for the gradient near the interface, first established in a flat configuration, i.e. when the interface coincides with a hyperplane. Section 3 is devoted to this “flat case”.

The starting point is a local boundedness estimate for the tangential gradient  $\nabla' u$ , where  $\nabla' = (\partial_{x_1}, \dots, \partial_{x_{n-1}})$  denotes the vector of derivatives parallel to the interface. This is proved in Proposition 3.2 via a difference-quotient argument with test functions involving only tangential increments. A key ingredient is a Moser iteration applied to  $Z := |\Delta_{i,h} u|^{m/p}$ . During the estimates, the full gradient  $\nabla u$  enters through the structure conditions (H1)–(H2) and is treated as a weight; this is the origin of the restriction  $2 < p < \frac{2n}{n-2}$ .

The bound obtained in Proposition 3.2 is preliminary, since the decay exponent in the estimates is not optimal and the constant may depend on  $\mu$ . To obtain the scale-invariant Lipschitz estimate for  $\nabla' u$  stated in Proposition 3.4, we perform a rescaling argument, which requires uniformity with respect to  $\mu$ . This uniformity is proved in Lemma 3.3, exploiting the boundedness of  $\nabla' u$  and a Poincaré-type inequality for  $Z := |\partial_i u|^{m/p}$ ,  $i = 1, \dots, n-1$ , which is an admissible Sobolev function by Theorem 3.1.

Once the tangential Lipschitz bound is available, we prove the Morrey estimate in the flat case in Theorem 3.5 by a comparison argument (Section 3.2). Finally, the general case follows from a standard flattening/approximation procedure using the  $C^1$  regularity of the interface, along the lines of the quadratic theory (see, e.g., [5, 6]).

## 2 Auxiliary results and notation

Throughout the paper we set

$$V^2 := \mu^2 + |\nabla u|^2,$$

where  $\nabla u$  denotes the weak gradient of  $u$ .

**Definition 2.1.** Let  $\Omega \subset \mathbb{R}^n$  be an open set,  $1 \leq p < \infty$ , and  $0 \leq \lambda \leq n$ . The Morrey space  $L^{p,\lambda}(\Omega)$  is the set of all functions  $f \in L^p_{\text{loc}}(\Omega)$  such that

$$\|f\|_{L^{p,\lambda}(\Omega)} := \sup_{x \in \Omega, r > 0} \left( r^{-\lambda} \int_{\Omega \cap B(x,r)} |f|^p dy \right)^{1/p} < \infty,$$

where  $B(x, r) = \{y \in \mathbb{R}^n : |y - x| < r\}$ .

We will make use of the following iteration lemma that can be found in [3, Lemma 2.VIII].

**Lemma 2.2.** Let  $\phi: (0, r] \rightarrow \mathbb{R}$  be a nonnegative function such that for every  $\sigma \in (0, r]$ ,  $t \in (0, 1)$  and  $\epsilon > 0$

$$\phi(t\sigma) \leq (At^\lambda + B\epsilon)\phi(\sigma) + \epsilon^{-\beta} K\sigma^\mu,$$

where  $0 < \mu < \lambda$ ,  $A > 0$ ,  $B \geq 0$ ,  $K \geq 0$  and  $\beta \geq 0$ . Then for every  $\delta < \lambda - \mu$  it holds that

$$\phi(t\sigma) \leq (1 + A)t^{\lambda-\delta}\phi(\sigma) + KM(t\sigma)^\mu,$$

where  $M = M(A, B, \delta, \lambda, \mu, \beta)$  is a positive constant.

**Lemma 2.3** (Standard  $p$ -growth bounds). Assume (H1)–(H2). Then there exists a constant  $C = C(n, p, \nu, L) > 0$  such that for all  $\xi \in \mathbb{R}^n$

$$C^{-1}(\mu^2 + |\xi|^2)^{\frac{p}{2}} - C\mu^p \leq F(\xi) \leq C(\mu^2 + |\xi|^2)^{\frac{p}{2}} + C\mu^p.$$

The following results on higher integrability and local Hölder continuity of minimizers are classical. Since their proofs are standard, we omit them and refer the reader to [7, 8].

**Theorem 2.4** (Higher integrability). *Let  $u \in W^{1,p}(\Omega)$  be a local minimizer of the functional (1). There exist  $s > 1$  and  $C = C(n, p, \nu, L, \beta)$  such that for any ball  $B_{2r}(x_0) \subset\subset \Omega$*

$$\int_{B_r(x_0)} |\nabla u|^{sp} dx \leq C \left( \int_{B_{2r}(x_0)} (|\nabla u|^p + \mu^p) dx \right)^s.$$

**Theorem 2.5** (Hölder continuity). *Let  $u \in W^{1,p}(\Omega)$  be a local minimizer of the functional (1). Then*

- (i) *For any open set  $\Omega' \Subset \Omega$  the quantity  $\|u\|_{L^\infty(\Omega')}$  is bounded by a constant  $C = C(n, p, \nu, L, \beta) \|u\|_{L^2(\Omega)}$ .*
- (ii)  *$u$  is locally Hölder continuous in  $\Omega$ .*

### 3 Flat case

In this section, we establish a local boundedness estimate for the gradient of a minimizer of the functional (1), hereafter denoted by  $\mathcal{F}$ , in a ball  $B_r(x_0)$  centered at a point  $x_0 \in \partial E$ , assuming that the interface  $\partial E$  is flat in  $B_r(x_0)$ .

Up to a rotation, we may assume that the inner normal  $\nu_E$  is aligned with the  $e_n$ -direction. In particular,

$$E = \{x \in \Omega : \langle x - x_0, e_n \rangle > 0\}.$$

For  $B_r(x_0) \subset \Omega$ , we define

$$B_r(x_0)^\pm = B_r(x_0) \cap \{x_n \gtrless (x_0)_n\}, \quad \Gamma_r = B_r(x_0) \cap \{x_n = (x_0)_n\}.$$

Then it follows that

$$B_r(x_0) \cap E = B_r(x_0)^+.$$

For the sake of readability we denote

$$A = \nabla F \quad \text{and} \quad \sigma = \sigma_E.$$

Moreover, we use the notation

$$\nabla' := (\partial_{x_1}, \dots, \partial_{x_{n-1}})$$

for the tangential gradient, i.e., the gradient with respect to the first  $n - 1$  variables.

#### 3.1 Lipschitz estimate for $\nabla' u$

The technique involved in the proof makes use of difference quotients. Let us set the notations.

##### Difference quotients.

For  $s \in \{1, \dots, n\}$  and  $h \in \mathbb{R} \setminus \{0\}$  we define the difference quotient

$$\Delta_{s,h} f(x) := f(x + h e_s) - f(x), \quad \delta_{s,h} f(x) := \frac{\Delta_{s,h} f(x)}{h}.$$

The translate of  $f$  in the direction of  $e_s$  will be denoted by

$$f_{h,e_s}(x) := f(x + h e_s).$$

In order to guarantee that  $x$  and  $x + h e_s$  belong to  $B_r(x_0)$  we work in

$$B_{r-|h|}(x_0) := \{x \in \mathbb{R}^n : |x - x_0| < r - |h|\} \subset B_r(x_0).$$

We remark that a local minimizer  $u \in W^{1,p}(B_r(x_0))$  of  $\mathcal{F}$  satisfies the following weak Euler-Lagrange equation:

$$\int_{B_r(x_0)} \sigma(x) \langle A(\nabla u), \nabla \phi \rangle dx = 0, \quad \forall \phi \in W_0^{1,p}(B_r(x_0)). \quad (2)$$

In the flat case, the existence of second derivatives of the minimizer  $u$ , separately in  $B_r^+$  and  $B_r^-$ , up to the boundary  $\Gamma_r$ , can be established by standard arguments; we refer the reader to [7, Section 8.4]. In our setting, one can additionally prove a further property — useful in what follows — namely that the tangential gradient  $\nabla' u$  admit a gradient also across the interface  $\Gamma_r$ .

**Theorem 3.1.** *Let  $u$  be a minimizer of  $\mathcal{F}$  in  $B_r(x_0)$ . Then  $\partial_i u \in W_{\text{loc}}^{1,2}(B_r)$  for every  $i = 1 \dots n - 1$ . Moreover the following properties hold:*

i)  $u \in W_{\text{loc}}^{2,2}(B_r^+) \text{ e } u \in W_{\text{loc}}^{2,2}(B_r^-)$ ;

ii) for every  $x_0 \in \Gamma_r$  and  $R > 0$  such that  $B_{4R}(x_0) \subset B_r$ , it holds that

$$\int_{B_{R/2}} V^{p-2} |\nabla \nabla' u|^2 dx \leq \frac{C}{R^2} \int_{B_{2R}} V^p dx,$$

where  $C = C(\frac{L}{\nu}, n, p, \alpha, \beta, \mu)$  is a positive constant.

*Proof.* Without loss of generality, we may assume that  $x_0 = 0$ . Let  $\rho \in (0, \frac{r}{4})$  and choose a cutoff function  $\eta \in C_c^\infty(B_\rho)$  such that

$$\eta \equiv 1 \text{ on } B_{\rho/2}, \quad 0 \leq \eta \leq 1, \quad |\nabla \eta| \leq \frac{C}{\rho},$$

for a positive constant  $C$ . Fix  $i \in \{1, \dots, n - 1\}$  and let  $h \in \mathbb{R} \setminus \{0\}$  satisfy  $|h| < \rho/2$ . We define the test function

$$\phi := \Delta_{i,-h}(\eta^2 \Delta_{i,h} u) \in W_0^{1,p}(B_\rho).$$

By the Euler–Lagrange equation satisfied by  $u$ , using the change of variables  $y = x - he_i$  and observing that  $\sigma$  depends only on  $x_n$  (hence  $\Delta_{i,h}\sigma = 0$  for  $i \in \{1, \dots, n - 1\}$ ), we obtain

$$\begin{aligned} 0 &= \int_{B_\rho} \sigma \langle A(\nabla u), \Delta_{i,-h} \nabla(\eta^2 \Delta_{i,h} u) \rangle dx \\ &= \int_{B_\rho} \sigma \langle \Delta_{i,h} A(\nabla u), \nabla(\eta^2 \Delta_{i,h} u) \rangle dx \\ &= \int_{B_\rho} \sigma \langle \Delta_{i,h} A(\nabla u), \eta^2 \nabla \Delta_{i,h} u + 2\eta \Delta_{i,h} u \nabla \eta \rangle dx \\ &= \sum_{j=1}^n \int_{B_\rho} \sigma \Delta_{i,h} A^j(\nabla u) (\eta^2 \partial_j \Delta_{i,h} u + 2\eta \Delta_{i,h} u \partial_j \eta) dx. \end{aligned} \tag{3}$$

We represent the difference quotient of  $A$  using the fundamental theorem of calculus:

$$\begin{aligned} \Delta_{i,h} A^j(\nabla u) &= \frac{1}{h} \int_0^1 \frac{d}{dt} A^j(\nabla u + th \Delta_{i,h} \nabla u) dt \\ &= \sum_{k=1}^n \left( \int_0^1 \partial_k A^j(\nabla u + th \Delta_{i,h} \nabla u) dt \right) \partial_k \Delta_{i,h} u. \end{aligned} \tag{4}$$

We introduce the auxiliary quantity

$$W_{i,h} := (\mu^2 + |\nabla u(x)|^2 + |\nabla u(x + he_i)|^2)^{1/2}.$$

By the growth assumptions on  $\nabla A$  and by [7, Lemma 8.3], we infer that

$$\left| \int_0^1 \partial_k A^j(\nabla u + th \Delta_{i,h} \nabla u) dt \right| \leq C(p, L) W_{i,h}^{p-2},$$

and, by the monotonicity assumption,

$$\sum_{j,k=1}^n \left( \int_0^1 \partial_k A^j(\nabla u + th \Delta_{i,h} \nabla u) dt \right) \xi_j \xi_k \geq C(p, \nu) W_{i,h}^{p-2} |\xi|^2 \tag{5}$$

for every  $\xi \in \mathbb{R}^n$ .

Combining (3), (4) and (5) and applying Hölder’s and Young’s inequalities, we obtain

$$\begin{aligned} \int_{B_\rho} W_{i,h}^{p-2} \eta^2 |\nabla \Delta_{i,h} u|^2 dx &\leq C \int_{B_\rho} W_{i,h}^{p-2} |\nabla \Delta_{i,h} u| |\Delta_{i,h} u| |\nabla \eta| \eta dx \\ &\leq C \left( \varepsilon \int_{B_\rho} W_{i,h}^{p-2} \eta^2 |\nabla \Delta_{i,h} u|^2 dx + \frac{1}{\varepsilon \rho^2} \int_{B_\rho} W_{i,h}^{p-2} |\Delta_{i,h} u|^2 dx \right), \end{aligned}$$

where  $C = C\left(n, p, \frac{\beta}{\alpha}, \frac{L}{\nu}\right) > 0$ . Choosing  $\varepsilon > 0$  sufficiently small and absorbing the first term into the left-hand side, we deduce

$$\begin{aligned} \int_{B_{\rho/2}} W_{i,h}^{p-2} \eta^2 |\nabla \Delta_{i,h} u|^2 dx &\leq \frac{C}{\rho^2} \int_{B_\rho} W_{i,h}^{p-2} |\Delta_{i,h} u|^2 dx \\ &\leq \frac{C h^2}{\rho^2} \int_{B_{2\rho}} V^p dx, \end{aligned} \quad (6)$$

where we have used the estimate

$$\int_{B_\rho} W_{i,h}^{p-2} |\Delta_{i,h} u|^2 dx \leq C h^2 \left( \int_{B_\rho} W_{i,h}^p dx + \int_{B_\rho} |\delta_{i,h} u|^p dx \right) \leq C h^2 \int_{B_{2\rho}} V^p dx.$$

Dividing (6) by  $h^2$  we obtain

$$\int_{B_{\rho/2}} W_{i,h}^{p-2} \eta^2 |\nabla \delta_{i,h} u|^2 dx \leq \frac{C}{\rho^2} \int_{B_{2\rho}} V^p dx. \quad (7)$$

Since  $\mu^{p-2} \leq W_{i,h}^{p-2}$ , estimate (7) yields a uniform bound for  $\{\delta_{i,h} \nabla u\}_{|h| < \rho/2}$  in  $L^2(B_{\rho/2})$ . By the standard characterization of Sobolev spaces via difference quotients (see [7, Lemma 8.2] or [8, Theorem 7.11]), applied to  $u$  and  $\partial_j u$ , it follows that  $\partial_i u \in W^{1,2}(B_{\rho/2})$  and that, up to a subsequence,

$$\delta_{i,h} u \rightarrow \partial_i u \quad \text{strongly in } L^2(B_{\rho/2}), \quad \nabla \delta_{i,h} u \rightarrow \nabla \partial_i u \quad \text{strongly in } L^2(B_{\rho/2}) \quad \text{and a.e.}$$

We now pass to the limit in (7). Since  $W_{i,h} \rightarrow V$  a.e. in  $B_{\rho/2}$ , by Fatou's lemma we infer

$$\int_{B_{\rho/2}} V^{p-2} \eta^2 |\nabla \partial_i u|^2 dx \leq \liminf_{h \rightarrow 0} \int_{B_{\rho/2}} W_{i,h}^{p-2} \eta^2 |\nabla \delta_{i,h} u|^2 dx.$$

Combining the previous inequality with (7), we get

$$\int_{B_{\rho/2}} V^{p-2} \eta^2 |\nabla \partial_i u|^2 dx \leq \frac{C}{\rho^2} \int_{B_{2\rho}} V^p dx.$$

Summing over  $i \in \{1, \dots, n-1\}$  and recalling that  $\eta \equiv 1$  on  $B_{\rho/2}$  yields

$$\int_{B_{\rho/2}} V^{p-2} |\nabla \nabla' u|^2 dx \leq \frac{C}{\rho^2} \int_{B_{2\rho}} V^p dx,$$

which proves (ii) (up to the relabelling  $\rho = 2R$ ).

Finally, since  $\sigma$  is constant in each of the sets  $B_r^\pm$ , the same argument can be repeated inside  $B_r^\pm$  also for  $i = n$ , yielding  $\partial_n u \in W_{\text{loc}}^{1,2}(B_r^\pm)$ . Therefore  $u \in W_{\text{loc}}^{2,2}(B_r^+)$  and  $u \in W_{\text{loc}}^{2,2}(B_r^-)$ , proving (i).  $\square$

In what follows we assume that  $n > 2$  and  $2 < p < 2^* := \frac{2n}{n-2}$ . Since the computation leading to the desired estimate is rather lengthy, we divide the proof into several steps. We first establish the local boundedness of the tangential gradient  $\nabla' u$ , without keeping track of the precise dependence of the resulting estimate on the constants and the exponents. This preliminary result allows us, in the subsequent propositions, to work with powers of the gradient  $\nabla' u$  in place of difference quotients, thereby simplifying the exposition and focusing on the improvement of the estimate and on the precise dependence on the constants.

**Proposition 3.2.** *Let  $u$  be a minimizer of  $\mathcal{F}$  in  $B_r(x_0)$ ,  $x_0 \in \Gamma_r$ . Then  $\nabla_i u \in L^\infty(B_{\frac{r}{2}}(x_0))$ , for every  $\rho \in (0, r)$  and  $i \in \{1, \dots, n-1\}$ .*

*Proof.* Without loss of generality, we can assume that  $x_0 = 0$ . We divide the proof in three steps.

**Step 1:** *A Caccioppoli type inequality combined with Sobolev–Poincaré inequality.*

We start writing an equation involving  $u$  and  $\Delta_{i,h} u$ . We show that the following equation holds:

$$\int_{B_{\rho-|h|}} \sigma(x) \langle A(\nabla u_{h,e_i}) - A(\nabla u), \nabla \phi \rangle dx = 0, \quad \forall \phi \in W_0^{1,p}(B_{\rho-|h|}). \quad (8)$$

For any  $\phi \in W_0^{1,p}(B_{\rho-|h|})$ , the translated function  $\phi_{-h,e_i} \in W_0^{1,p}(B_\rho)$  can be used in the Euler–Lagrange equation (2). Being  $\Delta_{i,h}\sigma(x) = 0$ , a tangential translation then yields

$$\int_{B_{\rho-|h|}} \sigma(x) \langle A(\nabla u_{h,e_i}), \nabla \phi \rangle dx = 0.$$

Subtracting the same identity with  $A(\nabla u)$ , we obtain

$$\int_{B_{\rho-|h|}} \sigma(x) \langle A(\nabla u_{h,e_i}) - A(\nabla u), \nabla \phi \rangle dx = 0,$$

that is, (8).

Now we choose a suitable test function in (8). For any  $m \geq p$  we denote  $\tilde{m} := \frac{2m}{p} \geq 2$  and define  $\psi = \eta^2 |\Delta_{i,h} u|^{\tilde{m}-2} \Delta_{i,h} u \in W_0^{1,p}(B_{\rho-|h|})$ . It holds that

$$\nabla \psi = 2\eta |\Delta_{i,h} u|^{\tilde{m}-2} \Delta_{i,h} u \nabla \eta + (\tilde{m} - 1) \eta^2 |\Delta_{i,h} u|^{\tilde{m}-2} \nabla (\Delta_{i,h} u).$$

Plugging  $\psi$  in (8), we get

$$\begin{aligned} & (\tilde{m} - 1) \int_{B_{\rho-|h|}} \sigma(x) \eta^2 |\Delta_{i,h} u|^{\tilde{m}-2} \langle A(\nabla u_{h,e_i}) - A(\nabla u), \nabla (\Delta_{i,h} u) \rangle dx \\ &= -2 \int_{B_{\rho-|h|}} \sigma(x) \eta |\Delta_{i,h} u|^{\tilde{m}-2} \Delta_{i,h} u \langle A(\nabla u_{h,e_i}) - A(\nabla u), \nabla \eta \rangle dx. \end{aligned}$$

Setting

$$W_{i,h} = (\mu^2 + |\nabla u|^2 + |\nabla u_{h,e_i}|^2)^{\frac{1}{2}},$$

and observing that  $\tilde{m} - 1 \geq 1$ , we apply (H1), (H2), Holder's and Young's inequalities and infer that

$$\begin{aligned} & \alpha \nu \int_{B_{\rho-|h|}} W_{i,h}^{p-2} \eta^2 |\Delta_{i,h} u|^{\tilde{m}-2} |\nabla (\Delta_{i,h} u)|^2 dx \\ & \leq 2\beta L \int_{B_{\rho-|h|}} \eta |\Delta_{i,h} u|^{\tilde{m}-1} W_{i,h}^{p-2} |\nabla (\Delta_{i,h} u)| |\nabla \eta| dx \\ & \leq 2\beta L \left[ \frac{\varepsilon}{2} \int_{B_{\rho-|h|}} W_{i,h}^{p-2} \eta^2 |\Delta_{i,h} u|^{\tilde{m}-2} |\nabla (\Delta_{i,h} u)|^2 dx + \frac{1}{2\varepsilon} \int_{B_{\rho-|h|}} W_{i,h}^{p-2} |\Delta_{i,h} u|^{\tilde{m}} |\nabla \eta|^2 dx \right], \end{aligned}$$

for any  $\varepsilon > 0$ . Choosing

$$\varepsilon := \frac{\alpha \nu}{2\beta L},$$

so that  $\alpha \nu - \beta L \varepsilon = \alpha \nu / 2$ , we can absorb the first term on the right-hand side into the left-hand side and obtain

$$\int_{B_{\rho-|h|}} W_{i,h}^{p-2} \eta^2 |\Delta_{i,h} u|^{\tilde{m}-2} |\nabla (\Delta_{i,h} u)|^2 dx \leq C \left( \frac{\beta}{\alpha}, \frac{L}{\nu} \right) \int_{B_{\rho-|h|}} W_{i,h}^{p-2} |\Delta_{i,h} u|^{\tilde{m}} |\nabla \eta|^2 dx, \quad (9)$$

where

$$C = 4 \left( \frac{\beta}{\alpha} \right)^2 \left( \frac{L}{\nu} \right)^2.$$

Defining

$$Z := |\Delta_{i,h} u|^{\frac{m}{p}},$$

so that

$$|\nabla Z|^2 = \left( \frac{m}{p} \right)^2 |\Delta_{i,h} u|^{\tilde{m}-2} |\nabla (\Delta_{i,h} u)|^2,$$

using (9) we get a Caccioppoli type inequality for  $Z$ ,

$$\begin{aligned} \int_{B_{\rho-|h|}} W_{i,h}^{p-2} |\nabla (\eta Z)|^2 dx & \leq 2 \int_{B_{\rho-|h|}} W_{i,h}^{p-2} |\nabla \eta|^2 Z^2 dx + 2 \int_{B_{\rho-|h|}} W_{i,h}^{p-2} |\nabla Z|^2 \eta^2 dx \\ & \leq C \left( p, \frac{\beta}{\alpha}, \frac{L}{\nu} \right) (1 + m^2) \int_{B_{\rho-|h|}} W_{i,h}^{p-2} |\nabla \eta|^2 Z^2 dx. \end{aligned}$$

We then combine this inequality with the Sobolev–Poincaré inequality. Let us fix  $\frac{\rho}{2} \leq \rho'' < \rho' \leq \rho$ . We choose  $\eta \in C_c^1(B_{\rho'})$  such that  $\eta = 1$  on  $B_{\rho''}$  and  $|\nabla \eta| \leq \frac{C}{\rho' - \rho''}$ , for some positive constant  $C$ . Applying the Sobolev–Poincaré inequality, remarking that  $\mu^{p-2} \leq W_{i,h}^{p-2}$  and making use of Hölder’s inequality, we get

$$\begin{aligned} \left( \int_{B_{\rho''}} |Z|^{2^*} dx \right)^{\frac{1}{2^*}} &\leq \left( \int_{B_{\rho'}} |\eta Z|^{2^*} dx \right)^{\frac{1}{2^*}} \leq C(n) \rho' \left( \int_{B_{\rho'}} |\nabla(\eta Z)|^2 dx \right)^{\frac{1}{2}} \\ &\leq C(n) \mu^{\frac{2-p}{2}} \rho' \left( \int_{B_{\rho'}} W_{i,h}^{p-2} |\nabla(\eta Z)|^2 dx \right)^{\frac{1}{2}} \\ &\leq C \left( n, p, \frac{\alpha}{\beta}, \frac{\nu}{L} \right) (1+m^2)^{\frac{1}{2}} \mu^{\frac{2-p}{2}} \rho' \left( \int_{B_{\rho'}} W_{i,h}^{p-2} |\nabla \eta|^2 Z^2 dx \right)^{\frac{1}{2}} \\ &\leq C \left( n, p, \frac{\alpha}{\beta}, \frac{\nu}{L} \right) (1+m^2)^{\frac{1}{2}} \mu^{\frac{2-p}{2}} \frac{\rho'}{\rho' - \rho''} \left( \int_{B_{\rho'}} W_{i,h}^{p-2} Z^2 dx \right)^{\frac{1}{2}} \\ &\leq C \left( n, p, \frac{\alpha}{\beta}, \frac{\nu}{L} \right) (1+m^2)^{\frac{1}{2}} \mu^{\frac{2-p}{2}} \frac{\rho'}{\rho' - \rho''} \left( \int_{B_{\rho'}} W_{i,h}^p dx \right)^{\frac{p-2}{2p}} \left( \int_{B_{\rho'}} Z^p dx \right)^{\frac{1}{p}}. \end{aligned}$$

Since

$$\left( \int_{B_{\rho'}} W_{i,h}^p dx \right)^{\frac{p-2}{2p}} \leq C(n) \left( \frac{\rho}{\rho'} \right)^n \Phi_p(\rho)^{\frac{p-2}{2p}},$$

where

$$\Phi_p(\rho) := \int_{B_\rho} (1 + |\nabla u|^2)^{\frac{p}{2}} dx,$$

we infer that

$$\left( \int_{B_{\rho''}} |Z|^{2^*} dx \right)^{\frac{1}{2^*}} \leq C \left( n, p, \frac{\beta}{\alpha}, \frac{L}{\nu} \right) (1+m^2)^{\frac{1}{2}} \mu^{\frac{2-p}{2}} \frac{\rho'}{\rho' - \rho''} \left( \frac{\rho}{\rho'} \right)^n \Phi_p(\rho)^{\frac{p-2}{2p}} \left( \int_{B_{\rho'}} Z^p dx \right)^{\frac{1}{p}}. \quad (10)$$

**Step 2:** *Obtaining the starting inequality for Moser’s iteration.* Let  $\chi = \frac{2^*}{p} > 1$  and

$$\rho_k = \rho \left( \frac{1}{2} + \frac{1}{2^{k+1}} \right), \quad \forall k \in \mathbb{N}_0.$$

We remark that  $\{\rho_k\}_{k \in \mathbb{N}_0}$  is a decreasing sequence such that  $\rho_k \rightarrow \frac{\rho}{2}$ . We set  $m_0 = p$  and

$$m_k = p \chi^k, \quad \forall k \in \mathbb{N}_0,$$

so that  $m_{k+1} = \chi m_k$ ,  $\{m_k\}_{k \in \mathbb{N}_0}$  is an increasing sequence and  $m_k \rightarrow +\infty$ . Furthermore, we define

$$Y_k = \left( \int_{B_{\rho_k}} |\Delta_{i,h} u|^{m_k} dx \right)^{\frac{1}{m_k}}, \quad \forall k \in \mathbb{N}_0.$$

Applying (10) with  $m = m_k$ ,  $Z = |\Delta_{i,h} u|^{\frac{m_k}{p}}$  (so that  $Z^{2^*} = |\Delta_{i,h} u|^{m_{k+1}}$ ),  $\rho' = \rho_k$  and  $\rho'' = \rho_{k+1}$ , we deduce that

$$Y_{k+1}^{\frac{m_{k+1}}{2^*}} \leq \Lambda_k \Phi_p(\rho)^{\frac{p-2}{2p}} Y_k^{\frac{m_k}{p}}, \quad \text{where } \Lambda_k = C \left( n, p, \frac{\beta}{\alpha}, \frac{L}{\nu} \right) (1+m_k^2)^{\frac{1}{2}} \mu^{\frac{2-p}{2}} \frac{\rho'}{\rho' - \rho''} \left( \frac{\rho}{\rho'} \right)^n,$$

for every  $k \in \mathbb{N}_0$ . Raising to the power  $\frac{2^*}{m_{k+1}} = \frac{p}{m_k} = \frac{1}{\chi^k}$ , we get

$$Y_{k+1} \leq \Lambda_k^{\frac{1}{\chi^k}} \Phi_p(\rho)^{\frac{p-2}{2p} \frac{1}{\chi^k}} Y_k, \quad \forall k \in \mathbb{N}_0. \quad (11)$$

**Step 3:** *Moser’s iteration.* Iterating (11)  $N$  times starting from  $k = 0$ , we get

$$Y_N \leq \left( \prod_{k=0}^{N-1} \Lambda_k^{\frac{1}{\chi^k}} \right) \Phi_p(\rho)^{\frac{p-2}{2p} \sum_{k=0}^{N-1} \frac{1}{\chi^k}} Y_0.$$

Letting  $N \rightarrow +\infty$ , we estimate the infinite product

$$\prod_{k=0}^{\infty} \Lambda_k^{1/\chi^k} = \exp\left(\sum_{k=0}^{\infty} \frac{1}{\chi^k} \log \Lambda_k\right).$$

Recall that

$$\Lambda_k = C_0 \mu^{\frac{2-p}{2}} (1 + m_k^2)^{1/2} \frac{\rho_k}{\rho_k - \rho_{k+1}} \left(\frac{\rho}{\rho_k}\right)^n, \quad m_k = p\chi^k,$$

where  $C_0 = C_0\left(n, p, \frac{\beta}{\alpha}, \frac{L}{\nu}\right)$ . Since  $\rho_k = \rho\left(\frac{1}{2} + \frac{1}{2^{k+1}}\right)$ , we have

$$\rho_k - \rho_{k+1} = \frac{\rho}{2^{k+2}} \implies \frac{\rho_k}{\rho_k - \rho_{k+1}} \leq c 2^k, \quad \left(\frac{\rho}{\rho_k}\right)^n \leq c,$$

for a constant  $c = c(n) > 0$  independent of  $k$ . Hence, for a new constant  $C_1 = C_1\left(n, p, \frac{\beta}{\alpha}, \frac{L}{\nu}\right)$ ,

$$\Lambda_k \leq C_1 \mu^{\frac{2-p}{2}} (1 + m_k^2)^{1/2} 2^k.$$

Taking logarithms yields

$$\log \Lambda_k \leq \log C_1 + \frac{2-p}{2} \log \mu + \frac{1}{2} \log(1 + m_k^2) + k \log 2.$$

Therefore

$$\sum_{k=0}^{\infty} \frac{1}{\chi^k} \log \Lambda_k \leq \left(\log C_1 + \frac{2-p}{2} \log \mu\right) \sum_{k=0}^{\infty} \frac{1}{\chi^k} + \frac{1}{2} \sum_{k=0}^{\infty} \frac{\log(1 + m_k^2)}{\chi^k} + (\log 2) \sum_{k=0}^{\infty} \frac{k}{\chi^k}.$$

Since  $\chi > 1$ , the geometric series  $\sum_{k \geq 0} \chi^{-k}$  converges and  $\sum_{k \geq 0} k \chi^{-k}$  converges as well. Moreover, using  $m_k = p\chi^k$  we have

$$\log(1 + m_k^2) \leq \log(2p^2) + 2k \log \chi,$$

hence  $\sum_{k \geq 0} \chi^{-k} \log(1 + m_k^2) < \infty$ . It follows that

$$\sum_{k=0}^{\infty} \frac{1}{\chi^k} \log \Lambda_k < \infty, \quad \text{and thus} \quad \prod_{k=0}^{\infty} \Lambda_k^{1/\chi^k} \leq C,$$

where  $C = C\left(n, p, \frac{\beta}{\alpha}, \frac{L}{\nu}, \mu\right) > 0$ .

and computing

$$\Phi_p(\rho)^{\frac{p-2}{2p} \sum_{k=0}^{+\infty} \frac{1}{\chi^k}} = \Phi_p(\rho)^{\frac{p-2}{2p} \frac{\chi}{\chi-1}},$$

we obtain

$$\|\Delta_{i,h} u\|_{L^\infty(B_{\frac{\rho}{2}})} \leq C \Phi_p(\rho)^{\frac{p-2}{2p} \frac{\chi}{\chi-1}} \|\Delta_{i,h} u\|_{L^p(B_\rho)}.$$

Dividing the previous inequality by  $h$  and letting  $h \rightarrow 0$ , we conclude that

$$\begin{aligned} \|\nabla_i u\|_{L^\infty(B_{\frac{\rho}{2}})} &\leq C \Phi_p(\rho)^{\frac{p-2}{2p} \frac{\chi}{\chi-1}} \|\nabla_i u\|_{L^p(B_\rho)} \leq C \Phi_p(\rho)^{\frac{p-2}{2p} \frac{\chi}{\chi-1}} \left( \int_{B_\rho} (\mu^2 + |Du|^2)^{\frac{p}{2}} dx \right)^{\frac{1}{p}} \\ &\leq C \Phi_p(\rho)^{\frac{p-2}{2p} \frac{\chi}{\chi-1} + \frac{1}{p}}, \end{aligned}$$

which is the thesis.  $\square$

Having obtained local boundedness for  $\nabla' u$ , we can now proceed by using an appropriate test function to derive a uniform boundedness estimate for  $\nabla' u$  independent of  $\mu$ .

**Lemma 3.3.** *Let  $u$  be a minimizer of  $\mathcal{F}$  in  $B_r(x_0)$ ,  $x_0 \in \Gamma_r$ . Then, setting  $\chi = \frac{2^*}{p}$ , for every  $i \in \{1, \dots, n-1\}$  it holds that*

$$\sup_{B_{\frac{\rho}{2}}(x_0)} |\nabla_i u| \leq C \left( \int_{B_\rho(x_0)} (\mu^2 + |\nabla u|^2)^{\frac{p}{2}} dx \right)^{\frac{p-2}{2p} \frac{\chi}{\chi-1} + \frac{1}{p}},$$

for every  $\rho \in (0, \frac{r}{2})$  and for some positive constant  $C = C\left(n, p, \frac{\beta}{\alpha}, \frac{L}{\nu}\right)$  independent of  $\mu$ .

*Proof.* Let us fix  $i \in \{1, \dots, n-1\}$  and  $\rho < r$ . Without loss of generality we may assume that  $x_0 = 0$ . Let  $m \geq p$  and set  $\tilde{m} := \frac{2m}{p} \geq 2$ . Reasoning as in Proposition 3.2, we obtain the inequality (9), which we recall here for the reader's convenience:

$$\int_{B_{\rho-|h|}} W_{i,h}^{p-2} \eta^2 |\Delta_{i,h} u|^{\tilde{m}-2} |\nabla(\Delta_{i,h} u)|^2 dx \leq C \left( \frac{\beta}{\alpha}, \frac{L}{\nu} \right) \int_{B_{\rho-|h|}} W_{i,h}^{p-2} |\Delta_{i,h} u|^{\tilde{m}} |\nabla \eta|^2 dx.$$

By the boundedness of  $\partial_i u$  proved in Proposition 3.2, we can pass to the limit as  $h \rightarrow 0$  in the previous inequality. More precisely, by Proposition 3.2 and Theorem 3.1 (ii), for every  $i \in \{1, \dots, n-1\}$  we have, as  $h \rightarrow 0$ ,

$$\delta_{i,h} u \rightarrow \partial_i u \quad \text{a.e. in } B_\rho, \quad \delta_{i,h} u \rightarrow \partial_i u \quad \text{in } L_{\text{loc}}^2(B_\rho),$$

and moreover

$$\nabla(\delta_{i,h} u) \rightarrow \nabla \partial_i u \quad \text{in } L_{\text{loc}}^2(B_\rho).$$

Since  $\partial_i u \in L_{\text{loc}}^\infty(B_\rho)$  and  $V^{p-2} \leq 1 + V^p \in L_{\text{loc}}^1(B_\rho)$ , the right-hand side integrand is dominated by an  $L_{\text{loc}}^1$  function, hence dominated convergence applies. Therefore, by weak lower semicontinuity on the left-hand side and dominated convergence on the right-hand side, we may pass to the limit in (9) as  $h \rightarrow 0$ , obtaining

$$\int_{B_\rho} V^{p-2} \eta^2 |\nabla_i u|^{\tilde{m}-2} |\nabla \nabla_i u|^2 dx \leq C \int_{B_\rho} V^{p-2} |\nabla_i u|^{\tilde{m}} |\nabla \eta|^2 dx. \quad (12)$$

We now introduce the notation

$$Z := |\nabla_i u|^{\frac{m}{p}},$$

which is well defined in  $B_{\rho/2}$  by Proposition 3.2. Moreover, by Theorem 3.1 we have

$$|\nabla Z|^2 = \left( \frac{m}{p} \right)^2 |\nabla_i u|^{\tilde{m}-2} |\nabla \nabla_i u|^2 \quad \text{a.e. in } B_{\rho/2}.$$

With this notation, inequality (12) can be rewritten as

$$\int_{B_\rho} V^{p-2} \eta^2 |\nabla Z|^2 dx \leq C m^2 \int_{B_\rho} V^{p-2} |\nabla \eta|^2 Z^2 dx, \quad (13)$$

where  $C = C(p, \frac{\beta}{\alpha}, \frac{L}{\nu}) > 0$ .

Now we want to obtain the starting inequality for Moser's iteration. Define

$$Y := 1 + Z, \quad \Phi_p(s) := \int_{B_s} (\mu^2 + |Du|^2)^{\frac{p}{2}} dx.$$

Fix  $\frac{\rho}{2} \leq \rho'' < \rho' \leq \rho$  and choose  $\eta \in C_c^1(B_{\rho'})$  such that  $\eta \equiv 1$  on  $B_{\rho''}$  and  $|\nabla \eta| \leq c(\rho' - \rho'')^{-1}$ . Set  $Z_1 := (Z - 1)_+$ . Then  $Z_1 \leq Z$  and  $\nabla Z_1 = \nabla Z$  a.e. on  $\{Z > 1\}$ . Moreover,

$$\{Z > 1\} = \{|\nabla_i u| > 1\} \subset \{|Du| > 1\}.$$

Since  $p > 2$ , this implies  $V \geq 1$  and hence  $V^{p-2} \geq 1$  on  $\text{spt } Z_1$ . Using this observation and (13), we obtain

$$\begin{aligned} \int_{B_{\rho'}} |\nabla(\eta Z_1)|^2 dx &\leq 2 \int_{B_{\rho'}} |\nabla \eta|^2 Z_1^2 dx + 2 \int_{B_{\rho'}} \eta^2 |\nabla Z_1|^2 dx \\ &\leq 2 \int_{B_{\rho'}} V^{p-2} |\nabla \eta|^2 Z^2 dx + 2 \int_{B_{\rho'}} V^{p-2} \eta^2 |\nabla Z|^2 dx \\ &\leq C(1 + m^2) \int_{B_{\rho'}} V^{p-2} |\nabla \eta|^2 Z^2 dx. \end{aligned}$$

By Sobolev's inequality and Hölder's inequality we deduce

$$\begin{aligned} \left( \int_{B_{\rho'}} |\eta Z_1|^{2^*} dx \right)^{\frac{1}{2^*}} &\leq C \rho' \left( \int_{B_{\rho'}} |\nabla(\eta Z_1)|^2 dx \right)^{\frac{1}{2}} \\ &\leq C \frac{\rho'}{\rho' - \rho''} \sqrt{1 + m^2} \Phi_p(\rho')^{\frac{p-2}{2p}} \left( \int_{B_{\rho'}} Z^p dx \right)^{\frac{1}{p}}. \end{aligned}$$

Finally, since  $Y \leq 2 + Z_1$  and  $Y \geq 1$ , we have  $Y^{2^*} \leq C(1 + Z_1^{2^*})$  and  $Y^p \leq C(1 + Z^p)$ . Using also the standard scaling inequality

$$\left( \int_{B_{\rho''}} f^{2^*} dx \right)^{\frac{1}{2^*}} \leq C \left( \frac{\rho'}{\rho''} \right)^{\frac{n-2}{2}} \left( \int_{B_{\rho'}} (\eta f)^{2^*} dx \right)^{\frac{1}{2^*}}, \quad f \geq 0,$$

we conclude that

$$\left( \int_{B_{\rho''}} Y^{2^*} dx \right)^{\frac{1}{2^*}} \leq C \left( \frac{\rho'}{\rho''} \right)^{\frac{n-2}{2}} \frac{\rho'}{\rho' - \rho''} \sqrt{1 + m^2} \Phi_p(\rho')^{\frac{p-2}{2p}} \left( \int_{B_{\rho'}} Y^p dx \right)^{\frac{1}{p}}. \quad (14)$$

With the previous inequality we can start the Moser iteration. Let  $\chi := \frac{2^*}{p} > 1$  and define

$$\rho_k := \rho \left( \frac{1}{2} + \frac{1}{2^{k+1}} \right), \quad m_k := p\chi^k, \quad Y_k := 1 + |\nabla_i u|^{\frac{m_k}{p}}, \quad N_k := \left( \int_{B_{\rho_k}} Y_k^p dx \right)^{\frac{1}{m_k}}.$$

Since

$$Y_{k+1}^p = \left( 1 + |\nabla_i u|^{\frac{m_{k+1}}{p}} \right)^p \leq \left( 1 + |\nabla_i u|^{\frac{m_k}{p}} \right)^{p\chi} = Y_k^{2^*},$$

applying (14) with  $m = m_k$ ,  $\rho' = \rho_k$  and  $\rho'' = \rho_{k+1}$  and raising to the power  $\frac{2^*}{m_{k+1}} = \frac{p}{m_k} = \frac{1}{\chi^k}$ , we have

$$N_{k+1} \leq \Lambda_k^{\frac{1}{\chi^k}} \Phi_p(\rho)^{\frac{p-2}{2p} \frac{1}{\chi^k}} N_k,$$

where  $\Lambda_k \leq C 2^k \sqrt{1 + m_k^2}$ . Iterating and using the convergence of  $\prod_k \Lambda_k^{1/\chi^k}$ , we obtain

$$\|\nabla_i u\|_{L^\infty(B_{\rho/2})} \leq C \Phi_p(\rho)^{\frac{p-2}{2p} \frac{\chi}{\chi-1}} \left( \int_{B_\rho} Y_0^p dx \right)^{\frac{1}{p}},$$

which concludes the proof of the lemma.  $\square$

Finally, exploiting a rescaling argument, we refine the Lipschitz estimate of the previous lemma, obtaining the correct power-law decay of the energy.

**Proposition 3.4.** *Let  $u$  be a minimizer of  $\mathcal{F}$  in  $B_r(x_0)$ ,  $x_0 \in \Gamma_r$ . Then for every  $i \in \{1, \dots, n-1\}$  it holds that*

$$\sup_{B_{\frac{r}{2}}(x_0)} |\nabla_i u| \leq C \left( \int_{B_\rho(x_0)} (\mu^2 + |\nabla u|^2)^{\frac{p}{2}} dx \right)^{\frac{1}{p}},$$

for every  $0 < \rho < \frac{r}{2}$  and for some positive constant  $C = C(n, p, \frac{\beta}{\alpha}, \frac{L}{\nu})$  independent of  $\mu$ .

*Proof.* Let  $0 < \rho < r$  and set

$$E := \left( \int_{B_\rho(x_0)} (\mu^2 + |\nabla u|^2)^{\frac{p}{2}} dx \right)^{\frac{1}{p}}, \quad v(y) := \frac{u(x_0 + \rho y)}{\rho E}, \quad y \in B_1.$$

Then  $v \in W^{1,p}(B_1)$  and

$$\nabla v(y) = E^{-1} \nabla u(x_0 + \rho y) \quad \text{for a.e. } y \in B_1.$$

**Step 1: Basic bounds and the rescaled coefficient.** By definition of  $E$  we have

$$E^p = \int_{B_\rho(x_0)} (\mu^2 + |\nabla u|^2)^{\frac{p}{2}} dx \geq \mu^p,$$

hence  $E \geq \mu$  and therefore

$$\tilde{\mu} := \frac{\mu}{E} \in (0, 1].$$

Moreover, since

$$(\mu^2 + t^2)^{\frac{p}{2}} \geq \frac{1}{2}(\mu^p + t^p) \quad \forall t \geq 0,$$

we infer

$$E^p \geq \frac{1}{2} \int_{B_\rho(x_0)} (\mu^p + |\nabla u|^p) dx \implies E^{-p} \int_{B_\rho(x_0)} |\nabla u|^p dx \leq 2.$$

Define the rescaled coefficient

$$\sigma_\rho(y) := \sigma_E(x_0 + \rho y), \quad y \in B_1.$$

In the flat case,  $\sigma_\rho$  still takes only the two values  $\alpha, \beta$  and the interface is  $\{y_n = 0\}$  (up to the fixed rotation), so we keep the same notation  $\sigma$  for  $\sigma_\rho$  in what follows.

**Step 2: The rescaled integrand  $G$  and the Euler-Lagrange equation.** We introduce the rescaled integrand

$$G(\xi) := E^{-p} F(E\xi), \quad \xi \in \mathbb{R}^n,$$

and its gradient

$$\tilde{A}(\xi) := \nabla G(\xi) = E^{1-p} A(E\xi), \quad A := \nabla F.$$

A standard change of variables shows that, up to the constant factor  $\rho^n E^p$ , the functional  $\mathcal{F}$  on  $B_\rho(x_0)$  is equivalent to the functional

$$\tilde{\mathcal{F}}(w; B_1) := \int_{B_1} \sigma(y) G(\nabla w) dy.$$

In particular, the minimality of  $u$  in  $B_\rho(x_0)$  implies that  $v$  is a minimizer of  $\tilde{\mathcal{F}}(\cdot; B_1)$  in  $B_1$ , hence  $v$  satisfies the weak Euler-Lagrange equation

$$\int_{B_1} \sigma(y) \langle \tilde{A}(\nabla v), \nabla \varphi \rangle dy = 0, \quad \forall \varphi \in W_0^{1,p}(B_1).$$

**Step 3: Structure conditions for  $\tilde{A}$ .** For every  $\xi, \eta \in \mathbb{R}^n$  we have

$$\begin{aligned} \langle \tilde{A}(\xi) - \tilde{A}(\eta), \xi - \eta \rangle &= E^{-p} \langle A(E\xi) - A(E\eta), E\xi - E\eta \rangle \\ &\geq \nu E^{-p} (\mu^2 + |E\xi|^2 + |E\eta|^2)^{\frac{p-2}{2}} E^2 |\xi - \eta|^2 \\ &= \nu (\tilde{\mu}^2 + |\xi|^2 + |\eta|^2)^{\frac{p-2}{2}} |\xi - \eta|^2, \end{aligned}$$

and similarly

$$\begin{aligned} |\tilde{A}(\xi) - \tilde{A}(\eta)| &= E^{1-p} |A(E\xi) - A(E\eta)| \\ &\leq L E^{1-p} (\mu^2 + |E\xi|^2 + |E\eta|^2)^{\frac{p-2}{2}} E |\xi - \eta| \\ &= L (\tilde{\mu}^2 + |\xi|^2 + |\eta|^2)^{\frac{p-2}{2}} |\xi - \eta|. \end{aligned}$$

Therefore  $\tilde{A}$  satisfies the same structural conditions with the same constants  $\nu, L$  and with parameter  $\tilde{\mu} \in (0, 1]$ .

**Step 4: Uniform energy bound for  $v$ .** Using the identity  $\nabla v = E^{-1} \nabla u(x_0 + \rho y)$ , we obtain

$$\begin{aligned} \int_{B_1} (\tilde{\mu}^2 + |\nabla v|^2)^{\frac{p}{2}} dy &= \int_{B_1} \left( \frac{\mu^2}{E^2} + \frac{|\nabla u(x_0 + \rho y)|^2}{E^2} \right)^{\frac{p}{2}} dy \\ &= E^{-p} \int_{B_\rho(x_0)} (\mu^2 + |\nabla u|^2)^{\frac{p}{2}} dx \\ &= E^{-p} E^p = 1. \end{aligned}$$

**Step 5: Apply Lemma 3.3 and scale back.** Applying Lemma 3.3 (with radius  $\rho = 1$ ) to  $v$  we find a constant  $C = C(n, p, \frac{\beta}{\alpha}, \frac{L}{\nu})$  independent of  $\tilde{\mu}$  (and hence of  $\mu$ ) such that, for every  $i \in \{1, \dots, n-1\}$ ,

$$\sup_{B_{1/2}} |\partial_i v| \leq C \left( \int_{B_1} (\tilde{\mu}^2 + |\nabla v|^2)^{\frac{p}{2}} dy \right)^{\frac{p-2}{2p} \frac{\chi}{\chi-1} + \frac{1}{p}} = C.$$

Finally, since  $\partial_i u(x_0 + \rho y) = E \partial_i v(y)$ , changing variables back gives

$$\sup_{B_{\rho/2}(x_0)} |\partial_i u| = E \sup_{B_{1/2}} |\partial_i v| \leq C E = C \left( \int_{B_\rho(x_0)} (\mu^2 + |\nabla u|^2)^{\frac{p}{2}} dx \right)^{\frac{1}{p}},$$

which is the desired estimate.  $\square$

### 3.2 Morrey estimate in the flat case

**Theorem 3.5.** *Let  $u$  be a minimizer of  $\mathcal{F}$  in  $B_r(x_0)$ . Then for every  $\delta \in (0, 1)$  it holds that*

$$\int_{B_s^+} |\nabla u|^p dx \leq C \left( \frac{s}{\rho} \right)^{n-\delta} \int_{B_\rho^+} (1 + |\nabla u|^2)^{\frac{p}{2}} dx,$$

for every  $0 < s < \frac{\rho}{2} < \frac{r}{4}$  and for some positive constant  $C = C(n, p, \frac{\rho}{\alpha}, \frac{L}{\nu}, \delta)$  independent of  $\mu$ .

*Proof.* Without loss of generality, we may assume that  $x_0 = 0$ . Since  $u \in W^{1,p}(B_r)$  and is locally Hölder continuous up to  $\Gamma_r$ , the trace  $u(\cdot, 0)$  is well defined pointwise and coincides with the Sobolev trace on  $\Gamma_r$ . Moreover, by Proposition 3.4,  $u(\cdot, 0) \in W^{1,\infty}(B_\rho')$ , for every  $\rho < r/2$  and its constant extension in the  $x_n$ -direction belongs to  $W^{1,\infty}(B_\rho)$ , with distributional gradient  $(\nabla' u(\cdot, 0), 0)$ . We define

$$U(x) := u(x) - u(x', 0), \quad \forall x \in B_\rho,$$

so that  $U = 0$  on  $\Gamma_\rho$ . The function  $U$  defined above vanishes identically on the flat boundary  $\Gamma_\rho$ , which allows us to exploit classical boundary regularity theory up to the flat boundary. At the same time,  $U$  satisfies a perturbed Euler–Lagrange equation, where the perturbation is encoded in the tangential gradient term  $b$  defined below, and will be estimated accordingly.

Recalling that the tangential gradient  $\nabla' u$  admits a trace on  $\Gamma_\rho$  (e.g. by Theorem 3.1(ii) and Proposition 3.4), we set

$$b(x') := (\nabla' u(x', 0), 0) \in \mathbb{R}^n, \quad x' \in \Gamma_\rho,$$

and extend  $b$  to  $B_\rho^+$  by  $b(x', x_n) := b(x')$ , so that  $\|b\|_{L^\infty(B_\rho^+)} = \|\nabla' u\|_{L^\infty(\Gamma_\rho)}$ . Let  $0 < \sigma < \rho < \frac{r}{2}$  with  $B_\rho^+ \subset E$  (hence the coefficient equals  $\beta$  in  $B_\rho^+$ , and in particular in  $B_\sigma^+$ ).  $U$  solves

$$\int_{B_\sigma^+} \langle A(\nabla U + b), \nabla \varphi \rangle dx = 0, \quad \forall \varphi \in W_0^{1,p}(B_\sigma^+). \quad (15)$$

Now let  $U_0 \in W^{1,p}(B_\sigma^+)$  be the unique weak solution to the unperturbed Euler–Lagrange equation associated with  $F$  in  $B_\sigma^+$ , with boundary datum  $U$ , namely

$$\int_{B_\sigma^+} \langle A(\nabla U_0), \nabla \varphi \rangle dx = 0 \quad \forall \varphi \in W_0^{1,p}(B_\sigma^+). \quad (16)$$

Equivalently,  $U_0$  is the unique minimizer of  $\mathcal{F}_0$  in  $U + W_0^{1,p}(B_\sigma^+)$ .

Set  $W := U - U_0 \in W_0^{1,p}(B_\sigma^+)$ . Subtracting (15) from the above equation and testing with  $\varphi = W$ , we obtain

$$\int_{B_\sigma^+} \langle A(\nabla U_0) - A(\nabla U + b), \nabla W \rangle dx = 0. \quad (17)$$

**Step 1: a comparison estimate.** Set  $\xi := \nabla U_0$  and  $\eta := \nabla U + b$ . Since

$$\nabla W = \nabla U - \nabla U_0 = (\eta - b) - \xi = -(\xi - \eta + b),$$

from (17) we infer

$$\int_{B_\sigma^+} \langle A(\xi) - A(\eta), \xi - \eta \rangle dx = \int_{B_\sigma^+} \langle A(\xi) - A(\eta), b \rangle dx.$$

By monotonicity (H1) and growth (H2), for a.e.  $x \in B_\sigma^+$ ,

$$\langle A(\xi) - A(\eta), \xi - \eta \rangle \geq \nu(\mu^2 + |\xi|^2 + |\eta|^2)^{\frac{p-2}{2}} |\xi - \eta|^2,$$

and

$$|\langle A(\xi) - A(\eta), b \rangle| \leq L(\mu^2 + |\xi|^2 + |\eta|^2)^{\frac{p-2}{2}} |\xi - \eta| |b|.$$

Applying Young's inequality to the right-hand side yields, for every  $\varepsilon \in (0, 1)$ ,

$$|\langle A(\xi) - A(\eta), b \rangle| \leq \varepsilon \nu(\mu^2 + |\xi|^2 + |\eta|^2)^{\frac{p-2}{2}} |\xi - \eta|^2 + C(\varepsilon)(\mu^2 + |\xi|^2 + |\eta|^2)^{\frac{p-2}{2}} |b|^2.$$

Integrating and absorbing the  $\varepsilon$ -term, we obtain

$$\int_{B_\sigma^+} (\mu^2 + |\nabla U_0|^2 + |\nabla U + b|^2)^{\frac{p-2}{2}} |\nabla U_0 - (\nabla U + b)|^2 dx \leq C \int_{B_\sigma^+} (\mu^2 + |\nabla U_0|^2 + |\nabla U + b|^2)^{\frac{p-2}{2}} |b|^2 dx. \quad (18)$$

We now estimate  $\int_{B_\sigma^+} |\nabla W|^p$ . Note that

$$-\nabla W = (\nabla U_0 - (\nabla U + b)) + b.$$

Hence, we have

$$\int_{B_\sigma^+} |\nabla W|^p dx \leq C \int_{B_\sigma^+} |\nabla U_0 - (\nabla U + b)|^p dx + C \sigma^n \|b\|_{L^\infty(B_\sigma^+)}^p. \quad (19)$$

Accordingly, by (18) and (19) we deduce that

$$\begin{aligned} \int_{B_\sigma^+} |\nabla W|^p &\leq C \left( \int_{B_\sigma^+} (\mu^2 + |\nabla U_0|^2 + |\nabla U + b|^2)^{\frac{p-2}{2}} |b|^2 dx + \sigma^n \|b\|_{L^\infty(B_\sigma^+)}^p \right) \\ &\leq C \left( \varepsilon \int_{B_\sigma^+} (\mu^2 + |\nabla U_0|^2 + |\nabla U|^2)^{\frac{p}{2}} dx + C(\varepsilon) \sigma^n \|b\|_{L^\infty(B_\sigma^+)}^p \right), \end{aligned}$$

where  $C(\varepsilon) = C\varepsilon^{-1}$ .

**Step 2: control of  $\nabla U_0$  by  $\nabla U$ .** By minimality of  $U_0$  (using  $U$  as a competitor) we have

$$\int_{B_\sigma^+} F(\nabla U_0) dx \leq \int_{B_\sigma^+} F(\nabla U) dx.$$

By virtue of Lemma 2.3 we infer

$$\int_{B_\sigma^+} (\mu^2 + |\nabla U_0|^2)^{\frac{p}{2}} dx \leq C \int_{B_\sigma^+} (\mu^2 + |\nabla U|^2)^{\frac{p}{2}} dx + C \sigma^n \mu^p. \quad (20)$$

**Step 3: boundary gradient bound and energy decay for  $U_0$ .** Since  $U_0$  is a weak solution to the homogeneous quasilinear elliptic equation in divergence form (16) in  $B_\sigma^+$  and satisfies the homogeneous Dirichlet condition  $U_0 = 0$  on the flat portion of the boundary  $\Gamma_\sigma$ , boundary  $C^{1,\alpha}$  regularity theory applies. In particular, by Lieberman [16, Theorem 1, p. 1203], there exist constants  $\alpha_0 \in (0, 1/2)$  and  $C = C(n, p, \nu, L)$  such that

$$\|\nabla U_0\|_{L^\infty(B_{\alpha_0\sigma}^+)} \leq C \left( \int_{B_\sigma^+} (\mu^2 + |\nabla U_0|^2)^{\frac{p}{2}} dx \right)^{\frac{1}{p}}. \quad (21)$$

We point out that the constants  $C, \alpha_0$  in (21) depend only on  $n, p, \nu, L$  and are uniform in  $\mu \in (0, 1]$ . Relying on this result, we now deduce a decay estimate for  $\nabla U_0$ , which is a direct consequence of (21).

Fix  $\tau \in (0, 1)$ . If  $\tau \leq \alpha_0$ , then by (21) and the inclusion  $B_{\tau\sigma}^+ \subset B_{\alpha_0\sigma}^+$  we have

$$\begin{aligned} \int_{B_{\tau\sigma}^+} |\nabla U_0|^p dx &\leq |B_{\tau\sigma}^+| \|\nabla U_0\|_{L^\infty(B_{\alpha_0\sigma}^+)}^p \\ &\leq C \tau^n \int_{B_\sigma^+} (\mu^2 + |\nabla U_0|^2)^{\frac{p}{2}} dx, \end{aligned} \quad (22)$$

where we used that  $|B_{\tau\sigma}^+| = \tau^n |B_\sigma^+|$  and possibly enlarged the constant  $C$ .

If instead  $\tau > \alpha_0$ , then trivially

$$\int_{B_{\tau\sigma}^+} |\nabla U_0|^p dx \leq \int_{B_\sigma^+} |\nabla U_0|^p dx,$$

and since  $\tau^n \geq \alpha_0^n$  we obtain

$$\int_{B_{\tau\sigma}^+} |\nabla U_0|^p dx \leq \alpha_0^{-n} \tau^n \int_{B_\sigma^+} |\nabla U_0|^p dx \leq C \tau^n \int_{B_\sigma^+} (\mu^2 + |\nabla U_0|^2)^{\frac{p}{2}} dx, \quad (23)$$

where the last inequality follows from  $|\nabla U_0|^p \leq (\mu^2 + |\nabla U_0|^2)^{p/2}$ .

Combining (22) and (23), we conclude that for every  $\tau \in (0, 1)$  there exists a constant  $C = C(n, p, \nu, L)$  such that

$$\int_{B_{\tau\sigma}^+} |\nabla U_0|^p dx \leq C \tau^n \int_{B_\sigma^+} (\mu^2 + |\nabla U_0|^2)^{\frac{p}{2}} dx. \quad (24)$$

**Step 4: decay for  $\nabla U$  and Morrey iteration.** Using  $\nabla U = \nabla U_0 + \nabla W$ , for any  $\tau \in (0, 1)$  we have

$$\int_{B_{\tau\sigma}^+} |\nabla U|^p dx \leq C \int_{B_{\tau\sigma}^+} |\nabla U_0|^p dx + C \int_{B_{\tau\sigma}^+} |\nabla W|^p dx.$$

By (24) and (20) we obtain, for every  $\varepsilon \in (0, 1)$ ,

$$\begin{aligned} \int_{B_{\tau\sigma}^+} |\nabla U|^p dx &\leq C \tau^n \int_{B_{\sigma}^+} (\mu^2 + |\nabla U_0|^2)^{\frac{p}{2}} dx + C\varepsilon \int_{B_{\sigma}^+} (\mu^2 + |\nabla U_0|^2 + |\nabla U|^2)^{\frac{p}{2}} dx + C(\varepsilon)\sigma^n \|b\|_{L^\infty(B_{\sigma}^+)}^p \\ &\leq C(\tau^n + \varepsilon) \int_{B_{\sigma}^+} (\mu^2 + |\nabla U|^2)^{\frac{p}{2}} dx + C(\tau^n + \varepsilon)\sigma^n \mu^p + C(\varepsilon)\sigma^n \|b\|_{L^\infty(B_{\sigma}^+)}^p \\ &\leq C(\tau^n + \varepsilon) \int_{B_{\sigma}^+} |\nabla U|^p dx + C(\tau^n + \varepsilon)\sigma^n + C(\varepsilon)\sigma^n \|b\|_{L^\infty(B_{\sigma}^+)}^p. \end{aligned} \quad (25)$$

Fix  $\delta \in (0, 1)$ . Since  $\sigma \leq \rho$ , we may write  $\sigma^n = \sigma^{n-\delta} \sigma^\delta \leq \sigma^{n-\delta} \rho^\delta$ . Therefore (25) implies

$$\int_{B_{\tau\sigma}^+} |\nabla U|^p dx \leq C(\tau^n + \varepsilon) \int_{B_{\sigma}^+} |\nabla U|^p dx + C\varepsilon^{-1} \sigma^{n-\delta} \rho^\delta \left(1 + \|b\|_{L^\infty(B_{\sigma}^+)}^p\right),$$

for a suitable positive constant  $C$  depending only on  $n, p, \nu, L$ . Choosing  $\varepsilon > 0$  sufficiently small (depending only on  $C, \delta$ ) and applying Lemma 2.2 to  $\phi(\sigma) := \int_{B_{\sigma}^+} |\nabla U|^p dx$  (with  $\lambda = n$  and  $\gamma = n - \delta$ ) we infer that

$$\int_{B_{\tau\sigma}^+} |\nabla U|^p dx \leq C \tau^{n-\delta/2} \int_{B_{\sigma}^+} |\nabla U|^p dx + C(\tau\sigma)^{n-\delta} \rho^\delta \left(1 + \|b\|_{L^\infty(B_{\sigma}^+)}^p\right), \quad (26)$$

for all  $\tau \in (0, 1)$  and  $0 < \sigma < \rho$ .

**Step 5: conclusion for  $\nabla u$ .** Choose  $\sigma = \rho/2$  and  $\tau = 2s/\rho$  (so that  $B_s^+ = B_{\tau\sigma}^+$  and  $\tau \in (0, 1)$  whenever  $0 < s < \rho/2$ ). Using (26) and the inequality  $|\nabla u|^p \leq C(|\nabla U|^p + |b|^p)$ , we obtain

$$\int_{B_s^+} |\nabla u|^p dx \leq C \left(\frac{s}{\rho}\right)^{n-\delta} \int_{B_{\rho/2}^+} |\nabla U|^p dx + C s^{n-\delta} \rho^\delta \left(1 + \|b\|_{L^\infty(B_{\rho/2}^+)}^p\right).$$

Moreover, since  $\nabla U = \nabla u - b$ , we have

$$\int_{B_{\rho/2}^+} |\nabla U|^p dx \leq C \int_{B_{\rho/2}^+} |\nabla u|^p dx + C\rho^n \|b\|_{L^\infty(B_{\rho/2}^+)}^p \leq C \int_{B_{\rho/2}^+} (1 + |\nabla u|^2)^{\frac{p}{2}} dx + C\rho^n \|b\|_{L^\infty(B_{\rho/2}^+)}^p.$$

By the tangential Lipschitz estimate (Proposition 3.4),

$$\|b\|_{L^\infty(B_{\rho/2}^+)} = \|\nabla' u\|_{L^\infty(\Gamma_{\rho/2})} \leq C \left( \int_{B_\rho} (\mu^2 + |\nabla u|^2)^{\frac{p}{2}} dx \right)^{\frac{1}{p}}.$$

Inserting these bounds and absorbing the lower-order contributions into the right-hand side, we conclude that for every  $0 < s < \rho/2 < \frac{r}{4}$ ,

$$\int_{B_s^+} |\nabla u|^p dx \leq C \left(\frac{s}{\rho}\right)^{n-\delta} \int_{B_{\rho/2}^+} (1 + |\nabla u|^2)^{\frac{p}{2}} dx,$$

with  $C = C(n, p, \frac{\beta}{\alpha}, \frac{L}{\nu}, \delta)$  independent of  $\mu$ . This concludes the proof.  $\square$

**Remark 3.6.** Let the assumptions of Theorem 3.5 be satisfied. Then the estimate

$$\int_{B_s^+} |\nabla u|^p dx \leq C \left(\frac{s}{\rho}\right)^{n-\delta} \int_{B_{\rho/2}^+} (1 + |\nabla u|^2)^{\frac{p}{2}} dx$$

holds for every  $0 < s < \rho < r$ , possibly with a different constant  $C > 0$ .

More precisely, the restriction  $0 < s < \rho/2 < r/4$  in Theorem 3.5 can be removed at the price of modifying the constant  $C$  by a universal factor depending only on  $n$  and  $\delta$ . In particular, the resulting constant still depends only on  $n, p, \frac{\beta}{\alpha}, \frac{L}{\nu}, \delta$  and is independent of  $\mu$ .

## 4 Proof of Proposition 1.3

*Proof.* Fix  $\tau \in (0, \tau_0)$  (to be chosen small only depending on  $n, p, \nu, L, \beta/\alpha$ ) and assume without loss of generality that  $x_0 = 0$ . We prove that there exists  $\varepsilon_0 = \varepsilon_0(\tau) > 0$  such that if  $B_r(x_0) \Subset \Omega$  and one of (i)–(iii) holds, then for every  $0 < \delta < n$ , then

$$\int_{B_{\tau r}} |\nabla u|^p dx \leq C_0 \tau^{n-\delta} \left( \int_{B_r} |\nabla u|^p dx + r^n \right),$$

where  $C_0 = C_0(n, p, \nu, L, \beta, \delta, \|\nabla u\|_{L^2(\Omega)}) > 0$ .

**Step 1: the flat comparison in case (iii).** Assume that (iii) holds true. Up to a rigid motion we may assume that  $H = \{x_n > 0\}$  and set

$$E_H := H \cap B_r, \quad \sigma_H := \beta \mathbb{1}_{E_H} + \alpha \mathbb{1}_{B_r \setminus E_H}.$$

Let  $u_H \in u + W_0^{1,p}(B_{\frac{r}{2}})$  be the unique minimizer of the functional

$$\mathcal{F}_H(w; B_{\frac{r}{2}}) := \int_{B_{\frac{r}{2}}} \sigma_H(x) F(\nabla w) dx.$$

Subtracting the Euler-Lagrange equation satisfied by  $u$  from the one satisfied by  $u_H$  and testing them with  $u - u_H$ , we get

$$\int_{B_{\frac{r}{2}}} \sigma_E (DF(\nabla u) - DF(\nabla u_H)) \cdot (\nabla u - \nabla u_H) dx = \int_{B_{\frac{r}{2}}} (\sigma_H - \sigma_E) DF(\nabla u_H) \cdot (\nabla u - \nabla u_H) dx$$

Using assumptions (H1) and (H2) and Hölder's inequality, for every  $\varepsilon > 0$  it holds that

$$\begin{aligned} \int_{B_{\frac{r}{2}}} |\nabla u_H - \nabla u|^p dx &\leq C \int_{(E \Delta E_H) \cap B_{\frac{r}{2}}} (\mu^2 + |\nabla u_H|^2)^{\frac{p-2}{2}} |\nabla u - \nabla u_H| dx \\ &\leq C \left( \frac{1}{\varepsilon} \int_{(E \Delta E_H) \cap B_{\frac{r}{2}}} (\mu^2 + |\nabla u_H|^2)^{\frac{p}{2}} dx + \varepsilon \int_{B_{\frac{r}{2}}} |\nabla u - \nabla u_H|^p dx \right). \end{aligned}$$

Choosing  $\varepsilon$  sufficiently small, we get

$$\int_{B_{\frac{r}{2}}} |\nabla u_H - \nabla u|^p dx \leq C \int_{(E \Delta E_H) \cap B_{\frac{r}{2}}} (\mu^2 + |\nabla u_H|^2)^{\frac{p}{2}} dx.$$

Thus, the minimality of  $u_H$  with respect to  $u$ , Hölder's inequality and Theorem 2.4 yield

$$\begin{aligned} \int_{B_{\frac{r}{2}}} |\nabla u_H - \nabla u|^p dx &\leq \int_{(E \Delta E_H) \cap B_{\frac{r}{2}}} (1 + |\nabla u|^2)^{\frac{p}{2}} dx \\ &\leq \left( \frac{|(E \Delta E_H) \cap B_r|}{|B_r|} \right)^{1-\frac{1}{s}} |B_r| \left( \int_{B_r} (1 + |\nabla u|^2)^{\frac{sp}{2}} dx \right)^{\frac{1}{s}} \\ &\leq C \varepsilon_0^{1-\frac{1}{s}} \int_{B_r} (1 + |\nabla u|^2)^{\frac{p}{2}} dx, \end{aligned} \tag{27}$$

where we have used assumption *iii*) and  $C = C(n, p, \beta/\alpha, \nu/L)$  is a positive constant.

Since  $E_H$  has flat interface in  $B_1$ , Theorem 3.5, Remark 3.6 (applied in  $B_r^+$  and, similarly, in  $B_r^-$ ) and the minimality of  $u_H$  with respect to  $u$  yield that for every  $0 < \delta < n$ ,

$$\int_{B_{\tau r}} |\nabla u_H|^p dx \leq C \tau^{n-\delta} \int_{B_r} (1 + |\nabla u_H|^2)^{\frac{p}{2}} dx \leq C \tau^{n-\delta} \int_{B_r} (1 + |\nabla u|^2)^{\frac{p}{2}} dx, \tag{28}$$

with  $C = C(n, p, \beta/\alpha, L/\nu, \delta)$ . Combining (28) and (27) we get

$$\begin{aligned} \int_{B_{\tau r}} |\nabla u|^p dx &\leq C \int_{B_{\tau r}} |\nabla u_H|^p dx + C \int_{B_{\frac{r}{2}}} |\nabla u - \nabla u_H|^p dx \\ &\leq C \left( \tau^{n-\delta} + \varepsilon_0^{1-\frac{1}{s}} \right) \int_{B_r} (1 + |\nabla u|^2)^{\frac{p}{2}} dx \end{aligned} \tag{29}$$

Choose  $\varepsilon_0 = \varepsilon_0(\tau) > 0$  so small that  $C \varepsilon_0^{1-\frac{1}{s}} \leq \tau^{n-\delta}$ . Then (29) simplifies to

$$\int_{B_{\tau r}} |\nabla u|^p dx \leq C \tau^{n-\delta} \left( \int_{B_r} |\nabla u|^p dx + r^n \right).$$

**Step 3: the one-phase cases (i) and (ii).** Assume (i):  $|E \cap B_r| < \varepsilon_0 |B_r|$  (the case (ii) is analogous). Let us denote by  $v \in u + W_0^{1,p}(B_{\frac{r}{2}})$  be the unique minimizer of the functional

$$\mathcal{G}(w) := \int_{B_{\frac{r}{2}}} F(\nabla w) dx.$$

We can proceed as in step (iii) subtracting the Euler-Lagrange equation satisfied by  $u$  from the one satisfied by  $v$  and testing them with  $u - v$ , getting

$$\int_{B_{\frac{r}{2}}} \sigma_E (DF(\nabla u) - DF(\nabla v)) \cdot (\nabla u - \nabla v) dx = \int_{B_{\frac{r}{2}}} (\alpha - \sigma_E) DF(\nabla v) \cdot (\nabla u - \nabla v) dx$$

Using assumptions (H1) and (H2) and Hölder's inequality, for every  $\varepsilon > 0$  it holds that

$$\begin{aligned} \int_{B_{\frac{r}{2}}} |\nabla v - \nabla u|^p dx &\leq C \int_{E \cap B_{\frac{r}{2}}} (\mu^2 + |\nabla v|^2)^{\frac{p-2}{2}} |\nabla u - \nabla v| dx \\ &\leq C \left( \frac{1}{\varepsilon} \int_{E \cap B_{\frac{r}{2}}} (\mu^2 + |\nabla v|^2)^{\frac{p}{2}} dx + \varepsilon \int_{B_{\frac{r}{2}}} |\nabla u - \nabla v|^p dx \right). \end{aligned}$$

Choosing  $\varepsilon$  sufficiently small, we get

$$\int_{B_{\frac{r}{2}}} |\nabla v - \nabla u|^p dx \leq C \int_{E \cap B_{\frac{r}{2}}} (\mu^2 + |\nabla v|^2)^{\frac{p}{2}} dx.$$

Using the minimality of  $v$  compared with  $u$  and Theorem 2.4 we deduce,

$$\begin{aligned} \int_{B_{\frac{r}{2}}} |\nabla v - \nabla u|^p dx &\leq \int_{E \cap B_{\frac{r}{2}}} (1 + |\nabla u|^2)^{\frac{p}{2}} dx \\ &\leq \left( \frac{|E \cap B_r|}{|B_r|} \right)^{1-\frac{1}{s}} |B_r| \left( \int_{B_r} (1 + |\nabla u|^2)^{\frac{sp}{2}} dx \right)^{\frac{1}{s}} \\ &\leq C \varepsilon_0^{1-\frac{1}{s}} \int_{B_r} (1 + |\nabla u|^2)^{\frac{p}{2}} dx, \end{aligned}$$

where we have used assumption  $i$ ) and  $C = C(n, p, \beta/\alpha, \nu/L)$  is a positive constant. Thereafter we can argue as in the previous step using the classical decay estimate for minimizers of regular integrals in the Calculus of Variations

$$\int_{B_{r\rho}} |\nabla v|^p dx \leq C \tau^n \int_{B_\rho} |\nabla u|^p dx,$$

for every  $\rho \leq \frac{r}{2}$ . □

## 5 Proof of Theorem 1.2

*Proof.* Let  $\Omega' \Subset \Omega$ . We show that for every  $\varepsilon > 0$  there exists  $0 < r_*(\varepsilon) < d := \text{dist}(\Omega', \partial\Omega)$  such that for every  $x_0 \in \partial E \cap \Omega'$  there exists a halfspace  $H = H(x_0) \subset \mathbb{R}^n$  satisfying

$$\frac{|(E \Delta H) \cap B_r(x_0)|}{|B_r|} < \varepsilon, \quad \text{for all } 0 < r < r_*(\varepsilon). \quad (30)$$

Let us choose  $H(x_0)$  as the halfspace determined by the tangent hyperplane  $T_{x_0} \partial E$  and the exterior normal to  $E$  at  $x_0$ . Since  $\partial E \cap \Omega$  is a  $C^1$ -hypersurface, for each  $x_0 \in \partial E \cap \Omega'$  (after a rigid motion)  $\partial E$  can be represented in a neighborhood of  $x_0$  as the graph of a  $C^1$  function. The halfspace  $H(x_0)$  approximates  $E$  in measure at small scales, and by compactness of  $\partial E \cap \overline{\Omega'}$  and continuity of the unit normal, the approximation can be chosen uniformly with respect to  $x_0 \in \partial E \cap \Omega'$ . This yields the claim (30).

Fix  $\Omega' \Subset \Omega$  and let  $\lambda \in [0, n)$ . Set

$$\delta := n - \lambda \in (0, n].$$

Let  $\tau_0 \in (0, 1)$  be given by Proposition 1.3 and fix  $\tau \in (0, \tau_0)$ . Let  $\varepsilon_0 = \varepsilon_0(\tau) > 0$  be the corresponding threshold in Proposition 1.3. We take  $\varepsilon = \varepsilon_0/2^n$  and obtain  $r_0 := r_*(\varepsilon_0/2^n) \in (0, d)$  such that for every  $x \in \partial E \cap \Omega'$  and every  $0 < r < r_0$  there exists a halfspace  $H$  with

$$\frac{|(E \Delta H) \cap B_r(x)|}{|B_r|} < \frac{\varepsilon_0}{2^n}. \quad (31)$$

Let  $x \in \Omega'$  and  $0 < r < r_0/2$ . We distinguish two cases.

(a) *One-phase case.* If  $B_r(x) \cap \partial E = \emptyset$ , then either  $B_r(x) \subset E$  or  $B_r(x) \subset E^c$ . Hence  $|E \cap B_r(x)| = 0$  or  $|B_r(x) \setminus E| = 0$ , and condition (i) or (ii) of Proposition 1.3 holds on  $B_r(x)$ . Actually, we could even get better regularity, the function  $u$  being  $p$ -harmonic in  $B_r(x)$  (see also Remark 1.5). (b) *Two-phase case.* If

$B_r(x) \cap \partial E \neq \emptyset$ , choose  $x^* \in \partial E \cap B_r(x)$ . Then  $B_r(x) \subset B_{2r}(x^*)$  and  $B_{2r}(x^*) \Subset \Omega$  since  $2r < r_0 < d$ . Applying (31) at  $x^*$  with radius  $2r$ , we find a halfspace  $H$  such that

$$|(E\Delta H) \cap B_{2r}(x^*)| < \frac{\varepsilon_0}{2^n} |B_{2r}| = \varepsilon_0 |B_r|.$$

Therefore,

$$|(E\Delta H) \cap B_r(x)| \leq |(E\Delta H) \cap B_{2r}(x^*)| < \varepsilon_0 |B_r|,$$

i.e. condition (iii) of Proposition 1.3 holds on  $B_r(x)$ .

In either case, Proposition 1.3 applies on  $B_r(x)$  and yields the existence of a positive constant  $C_0$  such that

$$\int_{B_{\tau r}(x)} |\nabla u|^p dy \leq C_0 \tau^{n-\delta} \left( \int_{B_r(x)} |\nabla u|^p dy + r^n \right),$$

for all  $x \in \Omega'$ ,  $\tau < \tau_0$  and all  $0 < r < r_0/2$ . This inequality implies that

$$\sup_{\substack{x \in \Omega' \\ \rho \in (0, \tau_0 r)}} \rho^{\delta-n} \int_{B_\rho(x)} |\nabla u|^p dx \leq C \left( r^{\delta-n} \int_{B_r(x)} |\nabla u|^p dx + r^n \right).$$

For every  $x \in \Omega'$  and  $\rho \in [\tau_0 r, d)$  it holds that

$$\rho^{\delta-n} \int_{B_\rho(x)} |\nabla u|^p dx \leq (\tau_0 r)^{\delta-n} \int_{B_d(x)} |\nabla u|^p dx.$$

Thus, taking into account the previous two estimates, it holds that

$$\sup_{\substack{x \in \Omega' \\ \rho \in (0, d)}} \rho^{\delta-n} \int_{B_\rho(x)} |\nabla u|^p dx < +\infty,$$

which leads to the thesis. □

*Acknowledgements* The authors are members of the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM).

**Data Availability** Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

**Conflicts of interest** The authors declare that they have no conflicts of interest.

## References

- [1] P. Acampora and E. Cristoforoni, On the asymptotic behavior of a diffraction problem with a thin layer, *Adv. Calc. Var.* **18**(4), 2025, 1011–1037.
- [2] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, 2000.
- [3] S. Campanato, *Elliptic Systems with Non-Linearity  $q$  Greater or Equal to Two. Regularity of the Solution of the Dirichlet Problem*, *Annali Mat. Pura Appl.* **147**(4) (1987), 117–150.
- [4] F. Della Pietra, C. Nitsch, R. Scala and C. Trombetti, An optimization problem in thermal insulation with Robin boundary conditions, *Commun. Partial Differ. Equations* **46**(12) (2021), 2288–2304.
- [5] L. Esposito, Density lower bound estimate for local minimizer of free interface problem with volume constraint, *Ric. Mat.* **68**, no. 2 (2019), 359–373.
- [6] N. Fusco and V. Julin, On the regularity of critical and minimal sets of a free interface problem, *Interfaces Free Boundaries* **17** (1) (2015), 117–142.
- [7] E. Giusti, *Direct Methods in the Calculus of Variations*, World Scientific, 2003.
- [8] M. Giaquinta and L. Martinazzi, *An Introduction to the Regularity Theory for Elliptic Systems, Harmonic Maps and Minimal Graphs*, 2nd ed., Springer, 2012.

- [9] V. Bianca, E. A. Pimentel, and J. M. Urbano, *BMO-regularity for a degenerate transmission problem*, Anal. Math. Phys. **14** (2024), Art. 9.
- [10] L. Caffarelli, M. Soria-Carro, and P. R. Stinga, *Regularity for  $C^{1,\alpha}$  interface transmission problems*, Arch. Ration. Mech. Anal. **240**(1) (2021), 265–294.
- [11] L. C. Evans, *Partial Differential Equations*, 2nd ed., Graduate Studies in Mathematics, Vol. 19, American Mathematical Society, 2010.
- [12] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer, 2001.
- [13] Y. Y. Li and M. Vogelius, *Gradient estimates for solutions to divergence form elliptic equations with discontinuous coefficients*, Arch. Rational Mech. Anal. **153** (2000), 91–151.
- [14] J.-L. Lions and E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications*, Springer, 1972.
- [15] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman, 1985.
- [16] G. M. Lieberman, *Boundary regularity for solutions of degenerate elliptic equations*, Nonlinear Anal. **12** (1988), 1203–1219.
- [17] M. Picone, *Sur un problème nouveau pour l'équation linéaire aux dérivées partielles de la théorie mathématique classique de l'élasticité*, Colloque sur les équations aux dérivées partielles, Bruxelles, May 1954.