PARTIAL REGULARITY FOR DEGENERATE VARIATIONAL PROBLEMS AND IMAGE RESTORATION MODELS IN BV

THOMAS SCHMIDT

ABSTRACT. We establish partial and local $\mathrm{C}^{1,\alpha}$-regularity results for vectorial almost-minimizers of convex variational integrals in BV. In particular, we investigate cases with a degenerate or singular behavior of p-Laplace type, and we cover (local) minimizers of the exemplary integrals

$$\int_\Omega (1+|\nabla w|^p)^{\frac{1}{p}}\,\mathrm{d}x$$

with $1<p<\infty$. We also treat some related models with lower-order terms, which are motivated by image restoration.

CONTENTS

1. INTRODUCTION

We are concerned with variational integrals of the type

$$\int_\Omega \big[f(\,\cdot\,,\nabla w)+g(\,\cdot\,,w)\big]\,\mathrm{d}x \qquad \text{for } w\colon \Omega\to\mathbb{R}^N\,,$$

where throughout this paper Ω denotes a non-empty bounded open set in $\mathbb{R}^n$, and the dimension $n\in\mathbb{N}$ and the codimension $N\in\mathbb{N}$ are arbitrary. Moreover, $f\colon\Omega\times\mathbb{R}^{Nn}\to\mathbb{R}$ and $g\colon\Omega\times\mathbb{R}^N\to\mathbb{R}$ are a suitable Borel integrands, and we will permanently assume that f is convex in the gradient variable and has linear growth in the sense of

$$|f(x,z)|\le C|z|+C \qquad \text{for all } (x,z)\in\Omega\times\mathbb{R}^{Nn}\,, \tag{1.1}$$

Date: May 23, 2014.
2010 *Mathematics Subject Classification.* Primary 49N60, 35J70, 35J75, 35J92; Secondary 26B30, 94A08, 35J93.
The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement GeMeThnES n° 246923.

where C is a fixed positive constant. Then, if w is in the space $\mathrm{BV}(\Omega)^N$ of $\mathbb{R}^N$-valued functions of bounded variation on Ω, we set

$$(1.2)\qquad F[w,\Omega] := \int_\Omega f(\,\cdot\,,\mathrm{D}w) \qquad\text{and}\qquad G[w,\Omega] := \int_\Omega g(\,\cdot\,,w)\,\mathrm{d}x\,.$$

Here, $G[w,\Omega]$ is only explained for $g(\,\cdot\,,w) \in \mathrm{L}^1(\Omega)$, and the definition of $F[w,\Omega]$ is understood in a generalized sense, namely as a convex functional of measures. For the moment we only remark that this interpretation of $F[w,\Omega]$ is quite natural and fruitful in the existence theory of minimizers, and we refer to Section 4.5 for the precise definition; see also Section 4 as a whole for further terminology.

We will work with the following notion of local minimizers, thus including minimizers of different boundary value problems and also unconstrained minimizers in our analysis.

Definition 1.1 (local minimizers)**.** *We say that $u \in \mathrm{BV}_{\mathrm{loc}}(\Omega)^N$ is a local minimizer of $F{+}G$ if we have $g(\,\cdot\,,u) \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$, and if*

$$(1.3)\qquad F[u,\mathrm{B}_\varrho(x_0)] + G[u,\mathrm{B}_\varrho(x_0)] \le F[u{+}\varphi,\mathrm{B}_\varrho(x_0)] + G[u{+}\varphi,\mathrm{B}_\varrho(x_0)]$$

holds for all balls $\mathrm{B}_\varrho(x_0) \subset\subset \Omega$ and all $\varphi \in \mathrm{BV}(\Omega)^N$ with $\operatorname{spt}\varphi \subset \mathrm{B}_\varrho(x_0)$ and $g(\,\cdot\,,u{+}\varphi) \in \mathrm{L}^1(\mathrm{B}_\varrho(x_0))$.

In this paper we study interior gradient regularity of local minimizers for functionals of the type $F{+}G$. We first state and discuss two exemplary results which illustrate our intentions and which are already new and interesting; compare the introductory discussions in [34, 7, 40].

1.1. **Degenerate model problems.** The first result guarantees almost-everywhere Hölder continuity of the gradient of a local minimizer — for some particularly simple choice of f and for $g \equiv 0$.

Theorem 1.2 (almost-everywhere gradient regularity)**.** *For $p \in (1,\infty)$ suppose that $u \in \mathrm{BV}_{\mathrm{loc}}(\Omega)^N$ is a local minimizer of*

$$(1.4)\qquad w \mapsto \int_\Omega (1+|\mathrm{D}w|^p)^{\frac{1}{p}}\,.$$

Then there exists an open subset Ω_0 of Ω with $\mathscr{L}^n(\Omega\setminus\Omega_0)=0$ such that u is of class $\mathrm{C}^{1,\alpha/\max\{p,2\}}$ locally on Ω_0 for every exponent $\alpha<\gamma_p$. Here, γ_p is a positive constant, depending only on n, N, and p, whose optimality will be discussed below.

The case $p=2$ of Theorem 1.2 (with $\gamma_2=2$) is a direct consequence of the local regularity result of Anzellotti & Giaquinta [8] (see Theorem 2.1 for a restatement in our setting). For $p \neq 2$ their result still yields the existence of an open and dense regular set Ω_0 on which u is $C^{1,\alpha/2}$ for all $\alpha \in (0,2)$. However, the crucial point in Theorem 1.2 is that the regular set Ω_0 has even full measure, and in the above generality this stronger assertion seems to be new even in the scalar[1] case $N=1$. In order to establish this improvement we will work near the zeros of $\mathrm{D}u$ (in the sense of Definition 4.13) — which is peculiar since the second derivative of $z \mapsto (1{+}|z|^p)^{\frac{1}{p}}$ either vanishes at $z=0$ ($p>2$) or becomes singular there ($p<2$). Our approach relies heavily on the observation that this degenerate or

[1]In the subcase $p=2$, $N=1$, everywhere regularity follows from the classical theory of the area integral, which is detailed in [35] (see also [51, 34] for extensions to broader classes of variational problems). The scalar Dirichlet problem with $p \neq 2$ has been discussed in [55], but only under restrictions on the domain and the boundary datum. However, imposing such restrictions everywhere regularity can be inferred via the boundary gradient estimates of [50, 57].

singular behavior resembles the one of the p-energy density $z \mapsto \frac{1}{p}|z|^p$, and indeed — adapting ideas of Duzaar & Mingione [26] — we will base our proof on the comparison of u with vector-valued p-harmonic functions (that are minimizers of the p-energy). Consequently, the limiting Hölder exponent $\gamma_p/\max\{p,2\}$ is the same one achievable for p-harmonic functions themselves (compare Theorem 4.18), but for $p \neq 2$ the optimal value of this exponent is not explicitly known.

We remark that for the particular integrands of Theorem 1.2 — contrary to the general settings of Sections 2 and 3 below — it remains possible that everywhere regularity $\Omega_0 = \Omega$ always holds. In fact, for $1 < p < 2$ we even prove this in a forthcoming joint work [10] with L. Beck, while for the limit case $p = 2$ we could at least show in [9] that the singular part of $\mathrm{D}u$ always vanishes (see also [12, 14, 41] for previous closely related results). However, for $p > 2$ it seems difficult to establish everywhere regularity, and in this case we consider Theorem 1.2 as most relevant.

1.2. **Image restoration models.** As a second exemplary case we discuss theoretical consequences of our results for some well-known image restoration models. These models are concerned with the recovery a good approximation u of an n-dimensional picture when only a corrupted recording S of this picture is available. Here, one should think of a corruption caused by noise and blur, and u and S are modeled as $\mathbb{R}^N$-valued functions on $\Omega \subset \mathbb{R}^n$, typically with $n \in \{2,3,4\}$. Moreover, the choice $N = 1$ is suitable for modeling the grayscales of monochrome pictures, while for colored pictures (or pictures with other special features) it might be relevant to take $N > 1$.

Given S is has been suggested by Rudin & Osher & Fatemi [46] to determine u from a constrained minimization problem. Reformulating this problem Chambolle & Lions [20] proposed to take u as the unique unconstrained minimizer of

$$w \mapsto \int_\Omega \left[|\mathrm{D}w| + \lambda|w-S|^\zeta \mathscr{L}^n\right]$$

with some (typically large) Lagrange multiplier $\lambda > 0$ and originally with $\zeta = 2$. Eventually, Chan & Esedoḡlu [21] studied the choice $\zeta = 1$ (leading to possible non-uniqueness of minimizers), and in the following we will consider arbitrary $\zeta > 0$. This approach is nowadays known as TV (L^ζ) regularization, since the leading regularization term $\int_\Omega |\mathrm{D}w|$ is the total variation of the gradient measure $\mathrm{D}w$. In fact, the basic idea is that the total variation has a *mild* regularizing effect: one may hope that u is smoother — and thus less noisy — than S, but that it also inherits jumps — and thus preserves or even sharpens edges; compare [18, 19]. Evidently, the fidelity term $\lambda \int_\Omega |w-S|^\zeta \,\mathrm{d}x$ forces u to stay close to S in addition.

Since the handling of the total variation near zeros of the gradient is quite difficult in both numerical and theoretical regards (nevertheless see [44, 43, 52, 2, 19] for some regularity results), it is convenient to use regularizations with similar properties. Among the various suggestions which have been made we here focus on the regularizing terms

$$\int_\Omega \mathrm{m}_p^\varepsilon(\mathrm{D}w) \qquad \text{and} \qquad \int_\Omega \widetilde{\mathrm{m}}_p^\varepsilon(\mathrm{D}w)\,,$$

where we have set

$$\mathrm{m}_p^\varepsilon(z) := (\varepsilon^p + |z|^p)^{\frac{1}{p}} - \varepsilon\,, \tag{1.5}$$

$$\widetilde{\mathrm{m}}_p^\varepsilon(z) := \begin{cases} \frac{1}{p\varepsilon^{p-1}}|z|^p & \text{for } |z| \le \varepsilon \\ |z| - \frac{p-1}{p}\varepsilon & \text{for } |z| \ge \varepsilon \end{cases} \tag{1.6}$$

for all $z \in \mathbb{R}^{Nn}$ with $p > 1$ and $\varepsilon > 0$; for $p = 2$, compare [21, Section7] and [20, formula (31)], respectively, for the usage of these integrals in image restoration and [39] for an occurrence of the latter one in a quite different context.

We would like to make the following two points about these regularizations. First, we stress that both $\mathrm{m}_p^\varepsilon(z)$ and $\widetilde{\mathrm{m}}_p^\varepsilon(z)$ converge to $|z|$ when we send *either* $p \searrow 1$ *or* $\varepsilon \searrow 0$. This double convergence gives some flexibility in the choices of p and ε when approximating the total variation. Second, in order to avoid the occurrence of piecewise constant minimizers, sometimes called the staircase effect, one may indeed want to strengthen the regularizing effect of the total variation slightly. Therefore, regularized models are also interesting in their own right. In this regard our results support the hypothesis that the usage of $\widetilde{\mathrm{m}}_p^\varepsilon$ is more promising than the one of m_p^ε: actually, in the latter case we have an almost-everywhere smoothing effect regardless of the size of the gradients (see Corollary 3.4), and even though this effect does not rule out jump discontinuous of u, it could result in blurred reconstructions. In contrast, in the former case we keep the desired milder regularization properties of the total variation near points of large gradient (thus retaining edges), while we also gain an additional selective smoothing property in regions of small gradient. Understanding Lebesgue points and Lebesgue values in the sense of Definition 4.13 below, this last effect can be made rigorous as follows.

Theorem 1.3 (selective smoothing)**.** *For $p \in (1, \infty)$, $\lambda \in [0, \infty)$, $\varepsilon, \zeta \in (0, \infty)$, and $S \in \mathrm{L}^\infty_{\mathrm{loc}}(\Omega)^N$ suppose that $u \in \mathrm{BV}_{\mathrm{loc}}(\Omega)^N$ is a local minimizer of*

$$w \mapsto \int_\Omega \left[\widetilde{\mathrm{m}}_p^\varepsilon(\mathrm{D}w) + \lambda |w - S|^\zeta \mathscr{L}^n\right],$$

where $\widetilde{\mathrm{m}}_p^\varepsilon$ is defined in (1.6). Then

$$\Omega_\varepsilon := \{x \in \Omega : x \text{ is a Lebesgue point of } \mathrm{D}u \text{ with Lebesgue value } z \text{ and } |z| < \varepsilon\}$$

is open and u is of class $\mathrm{C}^{1,\alpha/\max\{p,2\}}$ locally on Ω_ε for all those $\alpha < \gamma_p$ such that we have $\alpha \leq \min\{1, \zeta\}$.

We point out that for $p = 2$ and $\zeta \leq 1$ Theorem 1.3 follows immediately from [8, Section 6]. Moreover, motivated by image restoration Chen & Rao & Tonegawa & Wunderli [22, Theorem 1.2] established the scalar quadratic case $N = 1$, $p = \zeta = 2$ of the theorem. Some interest in the cases $p \neq 2$ has eventually been raised in [59], where however only certain non-degenerate approximations could be treated. To our knowledge Theorem 1.3 is new in all the remaining cases and provides the first regularity result for the degenerate and singular situations $p > 2$ and $1 < p < 2$, respectively.

We remark that in the case $p = 2$ of Theorem 1.3 (recall $\gamma_2 = 2$) the requirement $\alpha \leq \min\{1, \zeta\}$ and thus the Hölder exponent for the gradient of u on Ω_ε can indeed be improved: for $\zeta < 1$ the theorem gives the exponent $\frac{\zeta}{2}$, while a posteriori the optimal exponent $\frac{\zeta}{2-\zeta}$ of Phillips [45] can be obtained as in [31, 38, 48]. Similarly, for $\zeta \geq 1$ the theorem yields the exponent $\frac{1}{2}$, while a posteriori one reaches every exponent < 1 by standard estimates for the Poisson equation. Hence, it is likely that the condition $\alpha \leq \min\{1, \zeta\}$ can be improved for $p \neq 2$ as well, but we believe that a further discussion is rather pointless, as we also require $\alpha < \gamma_p$ and the optimal value of γ_p is not known for $p \neq 2$.

1.3. **Methodology of the proofs.** Our regularity proofs crucially rely on the localization method of Anzellotti & Giaquinta [8] and the comparison technique of

Duzaar & Mingione [26]. Let us comment on these techniques and their application to the model situation of Theorem 1.2 already at this preliminary stage.

To clarify notation and setup we first record that the integrand in (1.4) can be written as $1+\mathrm{m}_p^1(\mathrm{D}w)$ with the function m_p^1 from (1.5). Furthermore, we fix a local minimizer u of the integral (1.4) and a Lebesgue point $x_0 \in \Omega$ of $\mathrm{D}u$ with Lebesgue value $z_0 \in \mathbb{R}^{Nn}$, which implies in particular that the mean values $(\mathrm{D}u)_{\mathrm{B}_R(x_0)}$ converge to z_0 for $R \searrow 0$. In Section 4.7 we will a introduce an integral quantity $\Phi(u, \mathrm{B}_R(x_0))$, called the excess of u on $\mathrm{B}_R(x_0)$, which measures the deviation of u from being affine on $\mathrm{B}_R(x_0)$. The general strategy of proof is then to obtain estimates for the decay of $\Phi(u, \mathrm{B}_R(x_0))$ in R, which are equivalent with $\mathrm{C}^{1,\alpha}$-regularity near x_0. This common basic strategy has also been employed in [8], and regularity near x_0 is known, whenever $\nabla^2\mathrm{m}_p^1(z_0)$ is positive (see also Theorem 2.1 below). Thus, we will now confine ourselves to discussing the case $p \neq 2$, $z_0 = 0$, where $\nabla^2\mathrm{m}_p^1(z_0)$ vanishes or does not exist.

We are then in a degenerate[2] situation, and we will derive decay estimates by comparison with solutions of the p-Laplace system, which features a similar degenerate behavior. However, potentially the resulting estimates on balls $\mathrm{B}_R(x_0)$ may explode for $R \searrow 0$, and indeed this can happen when the ratio $|(\mathrm{D}u)_{\mathrm{B}_R(x_0)}|^p/\Phi(u, \mathrm{B}_R(x_0))$ blows up. As it turns out, when this ratio is sufficiently large, one should rather compare u with solutions of a linear system — even for $R + |(\mathrm{D}u)_{\mathrm{B}_R(x_0)}| \ll 1$. Therefore, we will actually distinguish between a truly degenerate case, based on comparison with p-Laplace systems, and a somehow non-degenerate case, based on comparison with linear systems. Notice that the latter situation is indeed a *non-uniformly non-degenerate* one, in which non-uniform factors like $|(\mathrm{D}u)_{\mathrm{B}_R(x_0)}|^{p-2}$ occur and need to be handled with care.

A comparison technique featuring the described case distinction has first been employed by Esposito & Mingione [29] in a blow-up argument, and eventually Duzaar & Mingione [26] implemented a similar strategy based on the p-harmonic approximation lemma of [27]; see also [47, 15, 16, 11] for further refinements and adaptions. While in all these references u is a priori in $\mathrm{W}^{1,p}$, in this paper u and its blow-ups are only in BV and are not easily approachable by the regularity theory for the p-Laplace system. This difficulty seems to rule out a direct usage of blow-up or harmonic approximation in our situation.

Consequently, we keep the general strategy of [26], but we rather implement the comparison arguments themselves via the direct localization method of Anzellotti & Giaquinta [8], which we customize to the degenerate situation. This method draws from earlier ideas of Schoen & Simon [49] in the setting of rectifiable currents, and it is based on a localization in the gradient variable z (which is highly non-trivial, as $\mathrm{D}u$ is not a continuous function). In addition, the usage of this method is also desirable for a couple of additional reasons: first, it works well with BV-functions; second, it yields localized statements in z; and third, it allows to ignore the non-uniformly elliptic behavior of $\nabla^2\mathrm{m}_p^1(z)$ for $|z| \to \infty$, that is the fact that the ratio between the largest and the smallest eigenvalue of $\nabla^2\mathrm{m}_p^1(z)$ blows up like $|z|^p$.

Finally, we stress that the localization procedure of [8] relies on up-to-the-boundary $\mathrm{C}^{1,\alpha}$-estimates for solutions h of the comparison systems. By these estimates, whenever the boundary values of h are $\mathrm{C}^{1,\alpha}$-close to an affine function, then ∇h is almost constant; thus, comparison with h can — to some extent —

[2] More precisely, the situation is degenerate only for $p > 2$, and it is singular for $p < 2$, but from now on we often use the word "degenerate" to summarize these cases.

be localized, in the sense that one can restrict the gradient variable z to a small ball. While the relevant up-to-the-boundary $\mathrm{C}^{1,\alpha}$-estimates are well-known for the linear comparison systems in [49, 8], they are not available for the p-Laplace system (where their validity is in fact a major open problem) and thus for our degenerate situation. We will however show that this difficulty can be overcome by a refinement of the localization method, which we obtain as a side benefit: our comparison procedure requires only up-to-the-boundary $\mathrm{W}^{1,q}$-estimates, and in some instances it avoids up-to-the-boundary estimates at all; compare Remarks 5.3 and 5.6.

2. Assumptions and statement of the main result

Postponing the treatment of the general integrals $F{+}G$ from (1.2), in the present section we restrict ourselves to autonomous integrals F; we thus consider

$$F[w,\Omega] := \int_\Omega f(\mathrm{D}w) \qquad \text{for } w \in \mathrm{BV}(\Omega)^N \tag{2.1}$$

with an integrand $f\colon \mathbb{R}^{Nn} \to \mathbb{R}$, for which we assume the following global hypothesis:

(H1) f is convex and Lipschitz continuous on $\mathbb{R}^{Nn}$ with Lipschitz constant $\le \Gamma$.

Here, global Lipschitz continuity is not restrictive in the sense that it follows from convexity and linear growth of f.

For comparison and future reference we now restate a version[3] of the local regularity result [8, Theorem 1.1], which we already mentioned in the introduction. Notice that in this statement and in the following we use (once more) the notions of Lebesgue points and Lebesgue values from Definition 4.13.

Theorem 2.1 (Anzellotti & Giaquinta [8])**.** *Assume that f satisfies (H1). Moreover, suppose that f is C^2 near some point $z_0 \in \mathbb{R}^{Nn}$ and that $\nabla^2 f(z_0)$ is positive. Then for every local minimizer $u \in \mathrm{BV}_{\mathrm{loc}}(\Omega)^N$ of F and every Lebesgue point $x_0 \in \Omega$ of $\mathrm{D}u$ with Lebesgue value z_0 there is some neighborhood of x_0 in which u is of class $\mathrm{C}^{1,\alpha/2}$ for every $\alpha \in (0,2)$.*

Our main regularity result provides a degenerate version of Theorem 2.1. In order to state it we impose the following localized hypotheses, which require that f behaves near some fixed point $z_0 \in \mathbb{R}^{Nn}$ similar to the p-energy density

$$\mathrm{e}_p(z) := \frac{1}{p}|z|^p \qquad \text{for } z \in \mathbb{R}^{Nn} \tag{2.2}$$

with $p \in (1,\infty)$:

(H2) ***p-growth near z_0:*** f is C^2 on $\mathrm{B}^{Nn}_\sigma(z_0) \setminus \{z_0\}$ and there holds

$$0 < |z| < \sigma \implies \sigma|z|^{p-2}|\xi|^2 \le \nabla^2 f(z_0{+}z)(\xi,\xi) \le \sigma^{-1}|z|^{p-2}|\xi|^2$$

for all $z, \xi \in \mathbb{R}^{Nn}$ and some constant $0 < \sigma \le \frac{1}{4}$;

[3] In fact, [8, Theorem 1.1] covers local minimizers of functionals with polynomial m-growth for any $m \in [1,\infty)$, and we believe that our results hold in this generality as well. However, in Theorem 2.1 and the whole present paper we restrict ourselves to the case $m = 1$ which we consider as the most interesting one. Notice also that [8, Theorem 1.1] imposes an additional coercivity condition of the type $f(z) \ge c|z|^m$. We can omit this condition for $m = 1$, since — after reduction to the case $f(z_0) > 0$, $\nabla f(z_0) = 0$ — it automatically follows from convexity of f.

(H3) **e_p-closeness near z_0:** f is differentiable at z_0 and resembles e_p there in the sense of

$$\lim_{z\to 0}\frac{f(z_0{+}z)-f(z_0)-\nabla f(z_0)z-\theta\mathrm{e}_p(z)}{|z|^p}=0$$

for some $\theta>0$.

We remark that in place of (H3) we could also require closeness to a general integrand of Uhlenbeck type (on which we would also impose suitable smoothness and p-growth conditions). However, to avoid further technicalities, we dispense with generalizations in this direction.

In the case $p\geq 2$ it follows from (H2) and (H3) that f is C^2 near z_0 with $\nabla^2 f(z_0)=\theta\nabla^2\mathrm{e}_p(0)\overset{p\geq 2}{=}0$. For $p<2$, in contrast, both $\nabla^2 f(z_0{+}z)$ and $\nabla^2\mathrm{e}_p(z)$ blow up as $z\to 0$. Furthermore, we record that the convergence in (H3) can be conveniently reformulated by saying that there exists some modulus $\eta_{\mathrm{d}}\colon(0,\infty)\to(0,\frac12]$ such that the following implication is true for all $\mu>0$ and $z\in\mathbb{R}^{Nn}$:

$$|z|\leq\eta_{\mathrm{d}}(\mu)\implies|f(z_0{+}z)-f(z_0)-\nabla f(z_0)z-\theta\mathrm{e}_p(z)|\leq\mu|z|^p\,. \tag{2.3}$$

In some sense the role of (2.3) on $\{z_0\}$ is taken over on $\mathrm{B}^{Nn}_{\sigma/2}(z_0)\setminus\{z_0\}$ by the following requirement with another modulus η_{n}:

(H4) **scaled uniform continuity property near z_0:** there is some constant $0<\sigma\leq\frac14$ (which we assume equals the one in (H2)) such that f is C^2 on $\mathrm{B}^{Nn}_\sigma(z_0)\setminus\{z_0\}$ and there exists some function $\eta_{\mathrm{n}}\colon(0,\infty)\to(0,\frac12]$ such that there holds

$$|\xi|<\frac12\sigma\,,\ 0<|z|\leq\eta_{\mathrm{n}}(\mu)|\xi|\implies|\nabla^2 f(z_0{+}\xi)-\nabla^2 f(z_0{+}\xi{+}z)|\leq\mu|\xi|^{p-2}$$

for all $\mu>0$ and $z,\xi\in\mathbb{R}^{Nn}$.

A sufficient criterion for the validity of (H4) is a scaled local Hölder condition $|\nabla^2 f(z_0{+}\xi)-\nabla^2 f(z_0{+}\xi{+}z)|\leq C|\xi|^{p-2-\beta}|z|^\beta$ for $0<2|z|\leq|\xi|\ll 1$ with $C,\beta>0$. Moreover, as noted in [47] there is a simple criterion for the validity of all three localized assumptions:

Remark 2.2. *The assumptions* (H2)*,* (H3)*, and* (H4) *all follow if f is C^2 on $\mathrm{B}^{Nn}_\sigma(z_0)\setminus\{z_0\}$ for some $\sigma>0$ with*

$$\lim_{z\to 0}\frac{\nabla^2 f(z_0{+}z)-\theta\nabla^2\mathrm{e}_p(z)}{|z|^{p-2}}=0 \tag{2.4}$$

for some $\theta>0$.

A proof of Remark 2.2 is sketched in Appendix B.

In order to cover several types of variational problems at once we will formulate our main result for almost-minimizers, thus following classical ideas in geometric measure theory; see for instance [3, 17, 53, 54, 5, 28]. In the present non-parametric setup, Anzellotti [6] and Duzaar & Gastel & Grotowski [23] employed almost-minimizers to treat different constraints — for instance obstacles and volume-constraints — in a unified way. Indeed, adapting [23, Definition 2.1] to the linear growth case, an adequate type of almost-minimizers $u\in\mathrm{BV}_{\mathrm{loc}}(\Omega)^N$ of F may be defined by requiring the inequality

$$F[u,\mathrm{B}_\varrho(x_0)]\leq F[u+\varphi,\mathrm{B}_\varrho(x_0)]+L\varrho^\alpha\int_{\mathrm{B}_\varrho(x_0)}(\mathscr{L}^n+|\mathrm{D}u|+|\mathrm{D}\varphi|) \tag{2.5}$$

for all balls $\mathrm{B}_\varrho(x_0) \subset\subset \Omega$ and all $\varphi \in \mathrm{BV}(\Omega)^N$ with $\operatorname{spt}\varphi \subset \mathrm{B}_\varrho(x_0)$, where $\alpha \in (0,\infty)$ and $L \in [0,\infty)$ are fixed. However, it turns out that the previous notion is still too restrictive for the purposes of Section 3 — where we want to include lower-order terms such as in Theorem 1.3. Thus, we will include almost-minimizers in the sense of (2.5) in our analysis, but in fact we will work with the following more general definitions, which will turn out to be particularly convenient in connection with Proposition 3.1.

Definition 2.3 (α-minimizers)**.** *For $\alpha \in (0,\infty)$ we say that $u \in \mathrm{BV}_{\mathrm{loc}}(\Omega)^N$ is an α-minimizer of F at $x_0 \in \Omega$ if there exists some function $\omega\colon [0,\infty) \to [0,\infty)$ such that the following property is valid: For all balls $\mathrm{B}_\varrho(x_0) \subset\subset \Omega$, all $\varphi \in \mathrm{BV}(\Omega)^N$ with $\operatorname{spt}\varphi \subset \mathrm{B}_\varrho(x_0)$, and all $M \in [0,\infty)$ with*

$$\fint_{\mathrm{B}_\varrho(x_0)} \big(|\mathrm{D}u| + |\mathrm{D}\varphi|\big) \le M \tag{2.6}$$

there holds

$$F[u,\mathrm{B}_\varrho(x_0)] \le F[u+\varphi,\mathrm{B}_\varrho(x_0)] + \omega(M)\varrho^\alpha \mathscr{L}^n(\mathrm{B}_\varrho)\,.$$

We call u an α-minimizer of F if it is an α-minimizer of F at every *$x_0 \in \Omega$ in such a way that the function ω can be chosen independent of x_0.*

Definition 2.4 (L^q-α-minimizers)**.** *For $q \in [0,\infty]$ we say that $u \in \mathrm{BV}_{\mathrm{loc}}(\Omega)^N$ is an L^q-α-minimizer of F (at $x_0 \in \Omega$) if the almost-minimizing property from Definition 2.3 holds for only those φ satisfying in addition to (2.6) also*

$$\|\varphi\|_{\mathrm{L}^q(\mathrm{B}_\varrho(x_0))^N} \le M\Big(\varrho^{1+\frac{n}{q}} + \|u - u_{\mathrm{B}_\varrho(x_0)}\|_{\mathrm{L}^q(\mathrm{B}_\varrho(x_0))^N}\Big)\,. \tag{2.7}$$

For $q \in [0,\frac{n}{n-1}]$ it turns out as a consequence of Sobolev's embedding that L^q-α-minimizers are the same as α-minimizers and thus in this range the notion is in fact independent of q. However, for $q > \frac{n}{n-1}$ the L^q-α-minimizers form an even more general class than the α-minimizers since for given M we have further restricted — at least on those balls where $\|u\|_{\mathrm{L}^q(\mathrm{B}_\varrho(x_0))^N}$ is finite — the class of test-functions φ, allowing only for those which are in some L^q-sense controlled by the minimizer itself.

Our main result for integrals of the type (2.1) reads:

Theorem 2.5 (local regularity for L^q-α-minimizers near degenerate points)**.** *Suppose that f satisfies the above assumptions* (H1)*,* (H2)*,* (H3)*, and* (H4) *for some $z_0 \in \mathbb{R}^{Nn}$ and $p \in (1,\infty)$. If $u \in \mathrm{BV}_{\mathrm{loc}}(\Omega,\mathbb{R}^N)$ is an L^q-α-minimizer of F with $q \in [0,\infty]$ and $0 < \alpha < \gamma_p$, then for every Lebesgue point $x_0 \in \Omega$ of* Du *with Lebesgue value z_0 there is some neighborhood of x_0 in which u is of class $\mathrm{C}^{1,\alpha/\max\{p,2\}}$;*

Here, the constant $\gamma_p(n,N,p) \in (0,2]$ is the one introduced in Theorem 4.18.

Remark 2.6 (the non-degenerate case)**.** *For $p = 2$ the theorem remains true with $\gamma_2 = 2$ if the role of $\theta\mathrm{e}_2$ in* (H3) *is taken by* any *positive symmetric bilinear form on $\mathbb{R}^{Nn}$. Then* (H2)*,* (H3)*, and* (H4) *altogether just correspond to saying that f is C^2 near z_0 and that $\nabla^2 f(z_0)$ is positive as in Theorem 2.1. An analogous remark applies to Proposition 2.7 and Corollary 3.3 below.*

We point out that in the case $p = 2 > \alpha$ Theorem 2.5 gives the Hölder exponent $\frac{\alpha}{2}$ for the gradient of u. It follows from [23, Example 3] that this exponent is optimal already for almost-minimizers in the sense of (2.5).

Assuming Theorem 2.5 we now provide a formal deduction of Theorem 1.2.

Proof of Theorem 1.2. We set $f(z) := (1+|z|^p)^{\frac{1}{p}}$. Then f satisfies (H1), and $\nabla^2 f(z_0)$ is positive for all $0 \neq z_0 \in \mathbb{R}^{Nn}$. Consequently, Theorem 2.1 implies $\mathrm{C}^{1,\alpha/\max\{p,2\}}$ regularity of u near the Lebesgue points of $\mathrm{D}u$ with Lebesgue value different from 0. Moreover, one easily checks (2.4) for the present f (with $\theta = 1$), and thus (H2), (H3), and (H4) are satisfied. Therefore, Theorem 2.5, specialized to local minimizers, yields $\mathrm{C}^{1,\alpha/\max\{p,2\}}$ regularity of u near the Lebesgue points of $\mathrm{D}u$ with Lebesgue value 0. In particular, it follows that the set of all Lebesgue points of $\mathrm{D}u$ is open, and choosing Ω_0 as this set we arrive at the claim. □

We remark that in a similar fashion Theorem 2.5 applies to integrands f with multiple degeneration points in $\mathbb{R}^{Nn}$. More precisely, we infer the following partial regularity statement: if $\nabla^2 f(z)$ is positive for all but at most countably many $z \in \mathbb{R}^{Nn}$, and if at those countably many $z \in \mathbb{R}^{Nn}$ it features a degenerate behavior of the above type, then every local minimizer u of F is C^1 near all Lebesgue points of $\mathrm{D}u$ (and thus $\mathscr{L}^n$-almost-everywhere).

For the purposes of the next section we record an additional slight refinement of Theorem 2.5. If $f\colon \Omega \times \mathbb{R}^{Nn} \to \mathbb{R}$ is such that $f(x_0, \cdot\,)$ satisfies (H1), we introduce the frozen functional F_{x_0} by

$$F_{x_0}[w,\Omega] := \int_\Omega f(x_0, \mathrm{D}w)$$

for $w \in \mathrm{BV}(\Omega)^N$ and $x_0 \in \Omega$. Then — inspired by [23] — we say that $u \in \mathrm{BV}_{\mathrm{loc}}(\Omega)^N$ is an α-minimizer of $(F_x)_{x\in\Omega}$ if u is an α-minimizer of F_{x_0} at every $x_0 \in \Omega$ such that the function ω in Definition 2.3 can be chosen independent of x_0. Analogously, we define L^q-α-minimizers of $(F_x)_{x\in\Omega}$.

Proposition 2.7. *For $f\colon \Omega \times \mathbb{R}^{Nn} \to \mathbb{R}$ suppose that $f(x, \cdot\,)$ satisfies the above assumptions (H1), (H2), (H3), and (H4) for all $x \in \Omega$, some $z_0 \in \mathbb{R}^{Nn}$, and some $p \in (1,\infty)$ in such a way that Γ, σ, η_{d}, and η_{n} can be chosen (locally) uniform in x. If $u \in \mathrm{BV}_{\mathrm{loc}}(\Omega,\mathbb{R}^N)$ is an L^q-α-minimizer of $(F_x)_{x\in\Omega}$ with $q \in [0,\infty]$ and $0 < \alpha < \gamma_p$, then for every Lebesgue point $x_0 \in \Omega$ of $\mathrm{D}u$ with Lebesgue value z_0 there is some neighborhood of x_0 in which u is of class $\mathrm{C}^{1,\alpha/\max\{p,2\}}$.*

The proof of Theorem 2.5, Remark 2.6, and Proposition 2.7 will be completed in Section 5.4.

3. Integrals with lower-order terms

In this section we return to the setting of (1.2), and we treat variational integrals $F{+}G$ with an explicit dependence on x and u. Indeed, we will show that regularity results for local minimizers of $F{+}G$ follow from Proposition 2.7, and as a particular case we will obtain Theorem 1.3.

Notice that in this section we use the notation $\mathrm{L}^{q,\alpha}(\Omega)$ from Definition 4.9 for Morrey spaces.

Proposition 3.1. *Suppose that $f\colon \Omega \times \mathbb{R}^{Nn} \to \mathbb{R}$ and $g\colon \Omega \times \mathbb{R}^N \to \mathbb{R}$ are Borel functions, that f is convex in its second argument and satisfies (1.1), and that the following Hölder conditions hold with a positive constant C, $0 < \alpha \le \beta \le 1$ and $\beta \le \zeta < \infty$:*

$$|f(x_2,z) - f(x_1,z)| \le C(1+|z|)|x_2-x_1|^\alpha\,, \tag{3.1}$$

$$|g(x,y_1) - g(x,y_2)| \le C(P(x)+|y_1|+|y_2|)^{\zeta-\beta}|y_2-y_1|^\beta \tag{3.2}$$

for all $x, x_1, x_2 \in \Omega$, $y_1, y_2 \in \mathbb{R}^N$, $z \in \mathbb{R}^{Nn}$. Moreover, assume that $P\colon \Omega \to [0,\infty)$ and $u \in \mathrm{BV}_{\mathrm{loc}}(\Omega)^N$ satisfy

$$|u|, P \in \mathrm{L}^{q,n-(\beta-\alpha)n_*}(\Omega)\,, \tag{3.3}$$

where we have set $n_* := \frac{n}{n-\beta n+\beta} \in [1,n]$ and $q := (\zeta-\beta)n_* \in [0,\infty[$. Then, if u is a local minimizer of $F{+}G$, it is also an L^q-α-minimizer of $(F_x)_{x\in\Omega}$.

Remark 3.2. *The proposition holds analogously for $\zeta = q = \infty$ if we understand $\mathrm{L}^{\infty,n-(\beta-\alpha)n_*} = \mathrm{L}^\infty$ and replace the Hölder condition for g with the requirement that $|g(x,y_1){-}g(x,y_2)| \le H(|y_1|{+}|y_2|)|y_2{-}y_1|^\beta$ holds for all $x \in \Omega$ and $y_1, y_2 \in \mathbb{R}^N$, and for some non-decreasing function $H\colon [0,\infty) \to [0,\infty)$. In this case the proof is a simplified variant of the one given below.*

Combining Propositions 2.7 and 3.1 we immediately get the following local regularity result for functionals of the type $F{+}G$.

Corollary 3.3. *Suppose that f satisfies the hypotheses of Proposition 2.7 for some $z_0 \in \mathbb{R}^{Nn}$ and $p \in (1,\infty)$, and that moreover f and g satisfy all the assumptions of Proposition 3.1 (including condition (3.3) for P) with $\alpha < \gamma_p$. If $u \in \mathrm{BV}_{\mathrm{loc}}(\Omega)^N$ is a local minimizer of $F{+}G$ satisfying the Morrey condition in (3.3), then for every Lebesgue point $x_0 \in \Omega$ of $\mathrm{D}u$ with Lebesgue value z_0 there is some neighborhood of x_0 in which u is of class $\mathrm{C}^{1,\alpha/\max\{p,2\}}$.*

Before proving Proposition 3.1 we make some comments on the meaning and justification of the extra Morrey assumption (3.3).

First and most importantly we stress that in many relevant cases boundedness of P is a reasonable assumption, which automatically implies (local) boundedness of u, see Appendix A. Thus, in many cases it suffices to impose an a priori assumption only on P, but not on u.

Moreover, we notice that the situation is particularly simple, if g is uniformly Hölder continuous in y, that is $\beta = \zeta \le 1$ and thus $q = 0$: in this case (3.3) is void.

In contrast, if $g(x,y)$ grows super-linearly in y, it cannot be uniformly Hölder continuous and we necessarily have $\beta < \zeta$. In this case we require the extra regularity (3.3) of P and u to compensate for the growth of g. The extra assumption is strongest (an L^∞-assumption), if we want to obtain an L^q-α-minimizer with $\alpha = \beta$. For $\alpha < \beta$ (3.3) is weaker, and we remark that it follows via (the limit case of) the embedding (4.10) from the condition

$$|u|, P \in \mathrm{L}^{\frac{\zeta-\beta}{\beta-\alpha}n}(\Omega) \tag{3.4}$$

on the scale of Lebesgue spaces. Moreover, sometimes (3.4) is trivially satisfied: for instance, if for $\beta = 1 < \zeta < \frac{n}{n-1}$ the natural integrability $|u|, P \in \mathrm{L}^\zeta(\Omega)$ is available (as in Corollaries 3.4 and 3.5 below with $P := |S|$), then we have $\frac{\zeta-\beta}{\beta}n < \zeta$, and thus (3.4) holds automatically for *some* positive α.

Proof of Proposition 3.1. We assume that u is a local minimizer of $F{+}G$. Then for $\mathrm{B}_\varrho(x_0) \subset\subset \Omega$ and $\varphi \in \mathrm{BV}(\Omega)^N$ with $\operatorname{spt}\varphi \subset \mathrm{B}_\varrho(x_0)$ we write

$$F_{x_0}[u, \mathrm{B}_\varrho(x_0)] - F_{x_0}[u+\varphi, \mathrm{B}_\varrho(x_0)] = I + II + III + IV \tag{3.5}$$

with

$$
\begin{aligned}
I &:= \int_{\mathrm{B}_\varrho(x_0)} f(x_0,\mathrm{D}u) - \int_{\mathrm{B}_\varrho(x_0)} f(\,\cdot\,,\mathrm{D}u)\,,\\
II &:= (F{+}G)[u,\mathrm{B}_\varrho(x_0)] - (F{+}G)[u+\varphi,\mathrm{B}_\varrho(x_0)]\,,\\
III &:= \int_{\mathrm{B}_\varrho(x_0)} f(\,\cdot\,,\mathrm{D}u+\mathrm{D}\varphi) - \int_{\mathrm{B}_\varrho(x_0)} f(x_0,\mathrm{D}u+\mathrm{D}\varphi)\,,\\
IV &:= \int_{\mathrm{B}_\varrho(x_0)} \big[g(\,\cdot\,,u+\varphi)-g(\,\cdot\,,u)\big]\,\mathrm{d}x\,.
\end{aligned}
$$

Here, we exploited that $g(\,\cdot\,,u)\in\mathrm{L}^1(\mathrm{B}_\varrho(x_0))$ by Definition 1.1, and we also used that $g(\,\cdot\,,u{+}\varphi)\in\mathrm{L}^1(\mathrm{B}_\varrho(x_0))$, which will follow from the below estimate showing the finiteness of IV. Relying on the latter condition once more we observe $II\le 0$ by the minimality property in (1.3). Moreover, if (2.6) holds for $M\in[0,\infty)$, then by the Hölder condition (3.1) we get

$$
I+III \le C\varrho^\alpha \int_{\mathrm{B}_\varrho(x_0)} \big[\mathscr{L}^n + |\mathrm{D}u| + |\mathrm{D}\varphi|\big] \le C(1{+}M)\varrho^\alpha\mathscr{L}^n(\mathrm{B}_\varrho)\,.
$$

For the treatment of the remaining term IV on the right-hand side of (3.5) we will work with constants C which may depend on n, α, β, ζ, the diameter of Ω, the constant C from (3.2), and the Morrey bounds for $|u|$ and P. We first deduce from the Hölder condition (3.2) and Hölder's inequality

$$
\begin{aligned}
IV &\le C\int_{\mathrm{B}_\varrho(x_0)} (P+|u|+|u+\varphi|)^{\zeta-\beta}|\varphi|^\beta\,\mathrm{d}x\\
&\le C\bigg(\int_{\mathrm{B}_\varrho(x_0)} (P+|u|+|\varphi|)^{(\zeta-\beta)n_*}\,\mathrm{d}x\bigg)^{\frac{1}{n_*}}\bigg(\int_{\mathrm{B}_\varrho(x_0)} |\varphi|^{\frac{n}{n-1}}\,\mathrm{d}x\bigg)^{\frac{n-1}{n}\beta}\,.
\end{aligned}
$$

Keeping $(\zeta-\beta)n_*=q$ in mind, involving also (2.7), and using a Sobolev-Poincaré inequality for φ we further get

$$
IV \le C(M^{\zeta-\beta}{+}1)\bigg(\varrho^{n+(\zeta-\beta)n_*} + \int_{\mathrm{B}_\varrho(x_0)} (P^q+|u|^q)\,\mathrm{d}x\bigg)^{\frac{1}{n_*}}\bigg(\int_{\mathrm{B}_\varrho(x_0)} |\mathrm{D}\varphi|\bigg)^{\beta}\,.
$$

Via (2.6) and the Morrey assumption (3.3) we arrive at

$$
\begin{aligned}
IV &\le C(M^{\zeta-\beta}{+}1)\Big(\varrho^{n+(\zeta-\beta)n_*} + \varrho^{n-(\beta-\alpha)n_*}\Big)^{\frac{1}{n_*}} M^\beta\varrho^{\beta n}\\
&= C(M^\zeta{+}M^\beta)\Big(\varrho^{(\zeta-\alpha)n_*}+1\Big)^{\frac{1}{n_*}}\varrho^{\frac{n}{n_*}+\beta n-\beta+\alpha}\,.
\end{aligned}
$$

Since ϱ is bounded by the diameter of Ω, in view of $\zeta\ge\alpha$ and $\frac{n}{n_*}+\beta n-\beta=n$ we find

$$
IV \le C(M^\zeta{+}M^\beta)\varrho^\alpha\mathscr{L}^n(\mathrm{B}_\varrho)\,.
$$

Collecting all the estimates we have shown that (2.6) and (2.7) imply

$$
F_{x_0}[u,\mathrm{B}_\varrho(x_0)] \le F_{x_0}[u+\varphi,\mathrm{B}_\varrho(x_0)] + C(1{+}M{+}M^\zeta{+}M^\beta)\varrho^\alpha\mathscr{L}^n(\mathrm{B}_\varrho)\,.
$$

Hence, u is an L^q-α-minimizer of F_{x_0} at x_0. □

To conclude this section we specialize Corollary 3.3 to the case where f is the integrand from either (1.5) or (1.6) and where g is given by $g(x,y)=\lambda|y-S(x)|^\zeta$ with a suitable function S, positive ζ, and an arbitrary real factor λ. We then get the following two statements.

Corollary 3.4. *For $p \in (1,\infty)$, $\lambda \in \mathbb{R}$, $\varepsilon,\zeta \in (0,\infty)$, and $S \in \mathrm{L}^\zeta_{\mathrm{loc}}(\Omega)^N$ suppose that $u \in \mathrm{BV}_{\mathrm{loc}}(\Omega)$ is a local minimizer of*

$$w \mapsto \int_\Omega \left[\mathrm{m}^\varepsilon_p(\mathrm{D}w) + \lambda|w-S|^\zeta \mathscr{L}^n\right].$$

Then the set Ω_0 of Lebesgue points of $\mathrm{D}u$ is open with $\mathscr{L}^n(\Omega \setminus \Omega_0) = 0$ and u is of class $\mathrm{C}^{1,\alpha/\max\{p,2\}}$ locally on Ω_0 for all those $\alpha < \gamma_p$ such that we have $\alpha \le \min\{1,\zeta\}$ and, in the case $\zeta > 1$, also

$$|u|, |S| \in \mathrm{L}^{(\zeta-1)n,\alpha n}_{\mathrm{loc}}(\Omega).$$

Corollary 3.5. *For $p \in (1,\infty)$, $\lambda \in \mathbb{R}$, $\varepsilon,\zeta \in (0,\infty)$, and $S \in \mathrm{L}^\zeta_{\mathrm{loc}}(\Omega)^N$ suppose that $u \in \mathrm{BV}_{\mathrm{loc}}(\Omega)^N$ is a local minimizer of*

$$w \mapsto \int_\Omega \left[\widetilde{\mathrm{m}}^\varepsilon_p(\mathrm{D}w) + \lambda|w-S|^\zeta \mathscr{L}^n\right].$$

Then

$$\Omega_\varepsilon := \{x \in \Omega : x \text{ is Lebesgue point of } \mathrm{D}u \text{ with Lebesgue value } z \text{ and } |z| < \varepsilon\}.$$

is open and u is of class $\mathrm{C}^{1,\alpha/\max\{p,2\}}$ locally on Ω_ε for all those $\alpha < \gamma_p$ such that we have $\alpha \le \min\{1,\zeta\}$ and, in the case $\zeta > 1$, also

$$|u|, |S| \in \mathrm{L}^{(\zeta-1)n,\alpha n}_{\mathrm{loc}}(\Omega)^N.$$

Proof of Corollaries 3.4 and 3.5. For the choices $f(x,z) = \mathrm{m}^\varepsilon_p(z)$ and $f(x,z) = \widetilde{\mathrm{m}}^\varepsilon_p(z)$ we first verify the assumptions of Proposition 2.7, which actually reduce to those of Theorem 2.5, since f is independent of x. First we notice that $\nabla^2\mathrm{m}^\varepsilon_p(z_0)$ is positive for $z_0 \ne 0$ and $\nabla^2\widetilde{\mathrm{m}}^\varepsilon_p(z_0)$ is positive for $0 < |z_0| < \varepsilon$. Thus, in these cases the required assumptions hold in the modified form of Remark 2.6. Moreover in the case $z_0 = 0$ the original assumptions follow — as in the proof of Theorem 1.2 in Section 2 — via Remark 2.2 (with $\theta = \frac{1}{\varepsilon^{p-1}}$).

Next we check that, locally on Ω, the assumptions of Proposition 3.1 are valid for the above f and for the choice $g(x,y) = \lambda|y-S(x)|^\zeta$. Evidently, f and g are Borel functions, and f satisfies (1.1). Moreover, (3.2) holds with $\beta := \min\{1,\zeta\}$ and $P := |S|$ (for $\zeta > 1$ this follows from the computation of $\nabla_y g(x,y)$, for $\zeta \le 1$ it is immediate). In addition, (3.3) is void for $\zeta \le 1$ (since $q = 0$) and reduces to the (localized) Morrey assumption of Corollaries 3.4 and 3.5 for $\zeta > 1$ (since $\beta = 1$, $n_* = n$). Finally, (3.1) is trivially valid for all $\alpha \le \min\{1,\zeta\}$.

Consequently, for $\alpha < \gamma_p$ Corollary 3.3 yields the claimed statements. □

Specializing to the case $\lambda \ge 0$, $S \in \mathrm{L}^\infty_{\mathrm{loc}}(\Omega)^N$ and involving Appendix A we finally obtain the second model result of the introduction.

Proof of Theorem 1.3. The claim follows from Corollary 3.5, once we show that the additional Morrey assumption of this corollary is satisfied. We can restrict ourselves to the case $\zeta > 1$, $\lambda > 0$ (as the assumption is only imposed for $\zeta > 1$ and for $\lambda = 0$ we can trivially reduce to $\zeta = 1$). Then, using $S \in \mathrm{L}^\infty_{\mathrm{loc}}(\Omega)^N$ and Lemma A.2 we see that $g(x,y) := \lambda|y-S(x)|^\zeta$ satisfies the assumptions of Theorem A.1. Moreover, we have $u \in \mathrm{L}^\zeta_{\mathrm{loc}}(\Omega)^N$ by the integrability assumption in Definition 1.1, and it is straightforward to check that also $f(x,z) := \widetilde{\mathrm{m}}^\varepsilon_p(z)$ satisfies the assumption of Theorem A.1. In conclusion, from this theorem we infer $u \in \mathrm{L}^\infty_{\mathrm{loc}}(\Omega)^N$, and we obtain the required Morrey assumption as a special case. □

4. Preliminaries

In this section we collect several preliminaries. Some of them are already closely related to the proof of the main result.

4.1. General notation. As most of our notation is standard, we just comment on a couple of features.

Throughout the paper, C and c denote positive constants, possibly varying from line to line, where we mostly use C for large and c for small constants. For $s \le t$ in $\mathbb{R} \cup \{\infty, -\infty\}$ we write (s,t) and $[s,t]$ for the open and the closed interval with endpoints s and t, respectively, and we also use $(s,t]$ and $[s,t)$ in the obvious meaning. We write $\mathrm{B}^n_\varrho(x) := \{y \in \mathbb{R}^n : |y-x| < \varrho\}$ for open balls in $\mathbb{R}^n$ (where the upper index n is often omitted), and we employ the abbreviation $S \subset\subset \Omega$ to indicate that a set S is relatively compact in Ω, in other words the closure of S is compact and is contained in Ω. Furthermore, $\mathbb{1}_S$ denotes the characteristic function of S and $\operatorname{spt}\eta$ the closure of the set of zeros of a function η. We usually identify the space of $(N{\times}n)$-matrices with $\mathbb{R}^{Nn}$, and we write $|\cdot|$ for several norms, namely the modulus of a real number, the Euclidean norm of a vector (consequently also for the Hilbert-Schmidt norm of a matrix), and the operator norm of a bilinear form. Finally, we adopt the convention that we think of first derivatives as vectors and of second derivatives as bilinear forms.

Further terminology is introduced below; see in particular Sections 4.3, 4.5, and 4.6 for notations related to measures and integrals.

4.2. Several inequalities for auxiliary functions. For this subsection we fix

$$1 < p < \infty\,.$$

We first record that the p-energy density e_p from (2.2) is convex with the following bound for its second derivatives.

Lemma 4.1. *For $z, \xi \in \mathbb{R}^{Nn}$ (with $z \neq 0$ if $p < 2$) one has*

$$\min\{1, p{-}1\}|z|^{p-2}|\xi|^2 \le \nabla^2\mathrm{e}_p(z)(\xi,\xi) \le \max\{p{-}1, 1\}|z|^{p-2}|\xi|^2$$

and consequently $|\nabla^2\mathrm{e}_p(z)| \le \max\{p{-}1, 1\}|z|^{p-2}$.

Proof. The claims follow easily from the calculation

$$\nabla^2\mathrm{e}_p(z)(\xi,\xi) = |z|^{p-2}|\xi|^2 + (p{-}2)|z|^{p-4}(z\cdot\xi)^2$$

and the Cauchy-Schwarz inequality. □

Next we restate another simple estimate which is useful in connection with $\nabla^2\mathrm{e}_p$; see [1, Lemma 2.1] for a proof in the case $p < 2$ which can easily be adapted to the general case.

Lemma 4.2. *For $z_0, z \in \mathbb{R}^{Nn}$ (with $|z_0|{+}|z| \neq 0$ if $p < 2$) there holds*

$$C^{-1}(|z_o| + |z|)^{p-2} \le \int_0^1 |z_0 + s(z{-}z_0)|^{p-2}\,\mathrm{d}s \le C(|z_o| + |z|)^{p-2}$$

with a positive constant C depending only on p.

Also the following slightly refined variant will be convenient.

Lemma 4.3. *For $z_0, z \in \mathbb{R}^{Nn}$ (with $|z_0|{+}|z| \neq 0$ if $p < 2$) there holds*

$$C^{-1}(|z_o| + |z|)^{p-2} \le \int_0^1\int_0^1 |z_0 + st(z{-}z_0)|^{p-2}\,\mathrm{d}s\,t\,\mathrm{d}t \le C(|z_o| + |z|)^{p-2}$$

with a positive constant C depending only on p.

Proof. We first provide a proof of the right-hand estimate in the case $p\le 2$: applying Lemma 4.2 twice we have

$$\begin{aligned}\int_0^1\int_0^1 |z_0+st(z{-}z_0)|^{p-2}\,\mathrm{d}s\,t\,\mathrm{d}t &\le C\int_0^1(|z_0|+|z_0+t(z{-}z_0)|)^{p-2}t\,\mathrm{d}t\\ &\le C\int_0^1|z_0+t(z{-}z_0)|^{p-2}\,\mathrm{d}t\le C(|z_0|+|z|)^{p-2}\,.\end{aligned}$$

Similarly, applying Lemma 4.2 just once in the last step we derive the left-hand estimate in the case $p\ge 2$:

$$\begin{aligned}(|z_0|+|z|)^{p-2} &\le 2^{p-1}\int_{\frac12}^1(t|z_0|+t|z|)^{p-2}t\,\mathrm{d}t\\ &\le 2^{p-1}\int_0^1(|z_0|+|tz|-|(1{-}t)z_0|)^{p-2}t\,\mathrm{d}t\\ &\le 2^{p-1}\int_0^1(|z_0|+|z_0+t(z{-}z_0)|)^{p-2}t\,\mathrm{d}t\\ &\le C\int_0^1\int_0^1|z_0+st(z{-}z_0)|^{p-2}\,\mathrm{d}s\,t\,\mathrm{d}t\,.\end{aligned}$$

The remaining cases are much simpler and can be treated directly without relying on Lemma 4.2; we omit further details. □

We also work with the auxiliary function A_p which is given by

$$\mathrm{A}_p(t):=(1+t)^{\frac1p}-1\qquad\text{for }t\ge 0\,. \tag{4.1}$$

We record $\mathrm{A}_p(t^p)=\mathrm{m}_p^1(t)$ holds for the function m_p^1 defined in (1.5), and we find it worth remarking that the role of A_p in the following lemmas and this entire paper could be taken over by $\widetilde{\mathrm{A}}_p$ with $\widetilde{\mathrm{A}}_p(t^p):=\widetilde{\mathrm{m}}_p^1(t)$ and $\widetilde{\mathrm{m}}_p^1$ defined in (1.6). Anyhow, retaining the choice from (4.1) we note that A_p is increasing and concave, while $z\mapsto\mathrm{A}_p(|z|^p)$ is convex. Moreover, we have

$$\mathrm{A}_p(|z|^p)\le\min\{|z|,|z|^p/p\}\qquad\text{for all }z\in\mathbb{R}^{Nn}\,. \tag{4.2}$$

Lemma 4.4. *For all $z,\xi\in\mathbb{R}^{Nn}$, and $C\ge 1$ there hold*

$$\begin{aligned}\mathrm{A}_p(C|z|^p)&\le C\mathrm{A}_p(|z|^p)\,,\\ \mathrm{A}_p(|z|^p+|\xi|^p)&\le 2[\mathrm{A}_p(|z|^p)+\mathrm{A}_p(|\xi|^p)]\,.\end{aligned}$$

Proof. The inequality $\mathrm{A}_p(C|z|^p)\le C\mathrm{A}_p(|z|^p)$ is a consequence of the concavity of A_p. Exploiting this inequality and the fact that A_p is increasing we estimate

$$\begin{aligned}\mathrm{A}_p(|z|^p+|\xi|^p)&\le\mathrm{A}_p(2\max\{|z|,|\xi|\}^p)\\ &\le 2\mathrm{A}_p(\max\{|z|,|\xi|\}^p)\le 2[\mathrm{A}_p(|z|^p)+\mathrm{A}_p(|\xi|^p)]\,.\end{aligned}$$ □

Lemma 4.5. *For all $z\in\mathbb{R}^{Nn}$ and $t>0$ we have the implications*

$$\begin{aligned}|z|\ge t &\implies \mathrm{A}_p(|z|^p)\ge c|z|\,,\\ |z|\le t &\implies \mathrm{A}_p(|z|^p)\ge c|z|^p\end{aligned}$$

with positive constants c depending only on p and t. For $t=2$ the first claim holds with $c=2^{-1}$ and the second one with $c=2^{-p}$.

Proof. For $|z| \geq t$ we have

$$(1+|z|^p)^{\frac{1}{p}} = (1-(1+t^p)^{-\frac{1}{p}})(1+|z|^p)^{\frac{1}{p}} + (1+t^p)^{-\frac{1}{p}}(1+|z|^p)^{\frac{1}{p}}$$
$$\geq (1-(1+t^p)^{-\frac{1}{p}})|z| + 1$$

and for $0 < |z| \leq t$ we deduce

$$\frac{\mathrm{A}_p(t^p)}{t^p}|z|^p = \frac{|z|^p}{t^p}\mathrm{A}_p\Big(\frac{t^p}{|z|^p}|z|^p\Big) \leq \mathrm{A}_p(|z|^p)$$

with the help of Lemma 4.4. Finally, the case $|z| = 0$ is trivial, and the claim for the case $t = 2$ is easily checked. □

Introducing yet other auxiliary functions we now turn to the well-studied quantities (see for instance [25, Section 3] for a discussion)

$$\mathrm{V}_p(z) := |z|^{\frac{p-2}{2}} z\,, \tag{4.3}$$

$$\mathrm{W}_p^\xi(z) := \begin{cases} (|\xi|^2+|z|^2)^{\frac{p-2}{2}}|z|^2 & \text{for } p \geq 2 \\ (|\xi|+|z|)^{p-2}|z|^2 & \text{for } 1 < p \leq 2 \end{cases}$$

for $\xi, z \in \mathbb{R}^{Nn}$. Here and in the following we understand $\mathrm{V}_p(0)$ and $\mathrm{W}_p^0(0)$ as 0 even for $1 < p \leq 2$, and without further mentioning we will adopt analogous reasonable conventions for similar singular expressions.

It can be checked by an explicit computation (a simplified version of those in the proof of Lemma 4.7 below) that W_p^ξ is a convex function on $\mathbb{R}^{Nn}$. Moreover, we record some useful estimates for W_p^ξ which are basically known. Nevertheless, we provide a brief proof.

Lemma 4.6. *For $z, \widetilde{z}, \xi \in \mathbb{R}^{Nn}$ there hold*

$$C_1^{-1}\big|\mathrm{V}_p(z) - \mathrm{V}_p(\xi)\big|^2 \leq \mathrm{W}_p^\xi(z-\xi) \leq C_1\big|\mathrm{V}_p(z) - \mathrm{V}_p(\xi)\big|^2\,, \tag{4.4}$$

$$\mathrm{W}_p^{\widetilde{z}}(z-\widetilde{z}) \leq C_1\big[\mathrm{W}_p^\xi(\widetilde{z}-\xi) + \mathrm{W}_p^\xi(z-\xi)\big]\,, \tag{4.5}$$

$$\mathrm{A}_p(\mathrm{W}_p^{\widetilde{z}}(z-\widetilde{z})) \leq C_1\big[\mathrm{A}_p(\mathrm{W}_p^\xi(\widetilde{z}-\xi)) + \mathrm{A}_p(\mathrm{W}_p^\xi(z-\xi))\big]\,, \tag{4.6}$$

$$\mathrm{W}_p^\xi(z+\widetilde{z}) \leq C_1\big[\mathrm{W}_p^\xi(z) + \mathrm{W}_p^\xi(\widetilde{z})\big] \tag{4.7}$$

with a positive constant C_1 depending only on p.

Proof. We first prove (4.4). To this end we set $\widetilde{p} := 2 + \frac{p-2}{2}$, and we observe

$$\mathrm{V}_p(z) - \mathrm{V}_p(\xi) = \nabla \mathrm{e}_{\widetilde{p}}(z) - \nabla \mathrm{e}_{\widetilde{p}}(\xi) = \int_0^1 \nabla^2 \mathrm{e}_{\widetilde{p}}(\xi + s(z-\xi))\,\mathrm{d}s\,(z-\xi,\cdot\,)\,. \tag{4.8}$$

From Lemmas 4.1 and 4.2 (with $\widetilde{p}$ in place of p) we thus infer

$$\big|\mathrm{V}_p(z) - \mathrm{V}_p(\xi)\big| \leq C(|\xi|+|z|)^{\frac{p-2}{2}}|z-\xi|\,.$$

Multiplying (4.8) with $(z-\xi)$ and using the lemmas again we also get

$$(|\xi|+|z|)^{\frac{p-2}{2}}|z-\xi|^2 \leq C\big|\mathrm{V}_p(z) - \mathrm{V}_p(\xi)\big|\,|z-\xi|\,.$$

In conclusion, we infer that $\big|\mathrm{V}_p(z) - \mathrm{V}_p(\xi)\big|$ is comparable to $(|\xi|+|z|)^{\frac{p-2}{2}}|z-\xi|$, and the last quantity is easily seen to be comparable to $\sqrt{\mathrm{W}_p^\xi(z-\xi)}$ as well. Thus, (4.4) is established. Now the claim (4.5) becomes obvious when we replace all three W_p-terms with the comparable quantities according to (4.4). Moreover, (4.6) is a consequence of (4.5) and the properties of A_p from Lemma 4.4. Finally, since $\mathrm{W}_p^\xi(z)$ is a non-increasing function of $|z|$, we get the estimate $\mathrm{W}_p^\xi(z+\widetilde{z}) \leq \mathrm{W}_p^\xi(2\max\{|z|,|\widetilde{z}|\}) \leq C\max\{\mathrm{W}_p^\xi(z),\mathrm{W}_p^\xi(\widetilde{z})\}$, and (4.7) follows. $\square$

Next we notice that $\mathrm{W}_p^\xi(z)$ behaves roughly like $|z|^p$ for $|\xi|\ll|z|\ll 1$ and like $|\xi|^{p-2}|z|^2$ for $|z|\ll|\xi|\ll 1$. This behavior will be crucial in our regularity proof in order to compare with both p-Laplacian and linear systems. However, in view of the BV-context we actually look for a quantity which, in addition, has linear growth and is convex in z, and it will turn out that a suitable choice is given by $\mathrm{A}_p(\mathrm{W}_p^\xi(z))$. The convexity of this composition in z is not immediate (A_p itself is indeed concave), but nevertheless it is guaranteed by the the following lemma, at least for sufficiently many ξ.

Lemma 4.7. *For all $\xi \in \mathbb{R}^{Nn}$ with $|\xi| \leq \frac{1}{2}$ the function*

$$\mathbb{R}^{Nn} \to \mathbb{R},\ z \mapsto \mathrm{A}_p(\mathrm{W}_p^\xi(z))$$

is convex.

Before proving the lemma, we briefly comment on the imposition of an upper bound for $|\xi|$. Indeed, for $p \leq 2$, the following computations show convexity for all $\xi \in \mathbb{R}^{Nn}$, and such a bound is not necessary. For $p > 2$, in contrast, it follows from the same computations that there exist sufficiently large ξ such that $\mathrm{A}_p(\mathrm{W}_p^\xi(z))$ is not convex in $z \in \mathbb{R}^{Nn}$, and thus some upper bound for $|\xi|$ seems unavoidable (though, needless to say, the stated bound $\frac{1}{2}$ is not optimal). We remark that this situation remains unchanged when we replace A_p with the function $\widetilde{\mathrm{A}}_p$ mentioned after (4.1).

Proof of Lemma 4.7. For $\xi = 0$ the claimed convexity has already been recorded at the beginning of this section, thus we assume $\xi \neq 0$. Since $\mathrm{A}_p(\mathrm{W}_p^\xi(z))$ is rotationally symmetric in z and non-decreasing in $|z|$, it suffices to check in the 1-dimensional case $Nn = 1$ that $\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{A}_p(\mathrm{W}_p^\xi(t)) \geq 0$ holds for $t \geq 0$. For $p \leq 2$ this requirement is immediate from the computation

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{A}_p(\mathrm{W}_p^\xi(t)) = \frac{1}{p}(1+\mathrm{W}_p^\xi(t))^{\frac{1-2p}{p}}(|\xi|+t)^{p-4}\bigg[&2\frac{2-p}{p}(|\xi|+t)^{p-2}t^2|\xi|^2\\ &+p(p-1)t^2+4(p-1)|\xi|t+2|\xi|^2\bigg].\end{aligned}$$

Similarly, for $p \geq 2$ we have

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{A}_p(\mathrm{W}_p^\xi(t)) = \frac{1}{p}(1+\mathrm{W}_p^\xi(t))^{\frac{1-2p}{p}}(|\xi|^2+t^2)^{\frac{p-6}{2}}\bigg[&\frac{p-2}{p}(|\xi|^2+t^2)^{\frac{p-2}{2}}t^2\big[p|\xi|^2t^2-2|\xi|^4\big]\\ &+p(p-1)t^4+(5p-6)|\xi|^2t^2+2|\xi|^4\bigg],\end{aligned}$$

and we can argue as follows: in case $t \geq |\xi|$ we evidently have $p|\xi|^2t^2 - 2|\xi|^4 \geq 0$ and the above expression is non-negative as required; in case $t \leq |\xi| \leq \frac{1}{2}$ we first observe $\frac{p-2}{p}(|\xi|^2+t^2)^{\frac{p-2}{2}}t^2 \leq 1$ and get the same conclusion. $\square$

Finally, we state a lemma which relates the integrands f of Section 2 to the quantity $\mathrm{A}_p(\mathrm{W}_p^\xi(z))$. This lemma will later be very convenient in order to compare the integral F with the excess Φ of Section 4.7 below.

Lemma 4.8. *Suppose that $f\colon \mathbb{R}^{Nn} \to \mathbb{R}$ satisfies (H1) and (H2) with $z_0 = 0$ and the fixed $p \in (1,\infty)$. Then for all $z, \xi \in \mathbb{R}^{Nn}$ with $|\xi| \leq \frac{1}{2}\sigma$ there holds*

$$C^{-1}\mathrm{A}_p(\mathrm{W}_p^{\xi}(z-\xi)) \leq f(z) - f(\xi) - \nabla f(\xi)(z-\xi) \leq C\mathrm{A}_p(\mathrm{W}_p^{\xi}(z-\xi))\,, \tag{4.9}$$

where C depends only p, Γ, and σ.

Proof. We first assume $|z| \leq \sigma$. In this case by (H2) and Lemma 4.3 we find

$$\begin{aligned} f(z) - f(\xi) - \nabla f(\xi)(z-\xi) &= \int_0^1 \int_0^1 \nabla^2 f(\xi + st(z-\xi))\,\mathrm{d}s\,t\,\mathrm{d}t\,(z-\xi, z-\xi) \\ &\leq \sigma^{-1} \int_0^1 \int_0^1 |\xi + st(z-\xi)|^{p-2}\,\mathrm{d}s\,t\,\mathrm{d}t\,|z-\xi|^2 \\ &\leq C(|\xi|+|z|)^{p-2}|z-\xi|^2 \\ &\leq C\,\mathrm{W}_p^{\xi}(z-\xi)\,. \end{aligned}$$

Since $\mathrm{W}_p^{\xi}(z-\xi)$ can be bounded by a constant depending only on σ and p, Lemma 4.5 gives $\mathrm{W}_p^{\xi}(z-\xi) \leq C\mathrm{A}_p(\mathrm{W}_p^{\xi}(z-\xi))$ and the right-hand inequality in (4.9) follows. Still in the case $|z| \leq \sigma$ the left-hand estimate in (4.9) is obtained by a completely analogous reasoning with (4.2) instead of Lemma 4.5.

Next we assume $|z| \geq \sigma$. In this case we have

$$|z-\xi| \geq |z| - |\xi| \geq \frac{1}{2}\sigma$$

and thus

$$f(z) - f(\xi) - \nabla f(\xi)(z-\xi) \leq 2\Gamma|z-\xi| \leq C\mathrm{A}_p(|z-\xi|^p) \leq C\mathrm{A}_p(\mathrm{W}_p^{\xi}(z-\xi))\,,$$

where for $p \geq 2$ we used only (H1), Lemma 4.5, and the fact that A_p is increasing. For $p < 2$ the last estimate follows as well, when we additionally notice $|z-\xi|^p \leq 2^{2-p}\mathrm{W}_p^{\xi}(z-\xi)$ and use Lemma 4.4 in the last step. Thus, the right-hand inequality in (4.9) is generally verified. Finally, to establish the remaining left-hand estimate in the case $|z| \geq \sigma$, we denote by $\widetilde{z}$ the unique point on the line segment from ξ to z with $|\widetilde{z}| = \frac{3}{4}\sigma$. Writing

$$\begin{aligned} f(z) - f(\xi) - \nabla f(\xi)(z-\xi) &= \big[f(z) - f(\widetilde{z}) - \nabla f(\widetilde{z})(z-\widetilde{z})\big] \\ &\quad + \big[f(\widetilde{z}) - f(\xi) - \nabla f(\xi)(\widetilde{z}-\xi)\big] + (\nabla f(\widetilde{z}) - \nabla f(\xi))(z-\widetilde{z}) \end{aligned}$$

the two terms in square brackets are non-negative by the convexity of f. Using this, the fact that $\widetilde{z}-\xi$ and $z-\xi$ point in the same direction, (H2), and Lemma 4.2 we infer

$$\begin{aligned} f(z) - f(\xi) - \nabla f(\xi)(z-\xi) &\geq \int_0^1 \nabla^2 f(\xi + s(\widetilde{z}-\xi))\,\mathrm{d}s\,(\widetilde{z}-\xi, z-\xi) \\ &\geq \int_0^1 |\xi + s(\widetilde{z}-\xi)|^{p-2}\,\mathrm{d}s\,|\widetilde{z}-\xi|\,|z-\xi| \\ &\geq c(|\xi|+|\widetilde{z}|)^{p-2}|\widetilde{z}-\xi|\,|z-\xi| \\ &\geq c|z-\xi|\,. \end{aligned}$$

By (4.2) and $|\xi| \leq |z-\xi|$ we moreover have

$$\mathrm{A}_p(\mathrm{W}_p^{\xi}(z-\xi)) \leq \mathrm{W}_p^{\xi}(z-\xi)^{\frac{1}{p}} \leq C|z-\xi|\,,$$

and combining the last two inequalities the proof is complete in all cases. ∎

4.3. **Signed and vector-valued measures.** In this paper a non-negative measure on Ω is a σ-additive function from the Borel-σ-algebra of Ω to $[0,\infty]$. Moreover, a signed or $\mathbb{R}^m$-valued measure ν on Ω is a σ-additive function which is initially defined only on the relatively compact Borel subsets of Ω and takes values in $(-\infty,\infty]$ or $\mathbb{R}^m$ (compare [4, Definition 1.40] for a similar method of approach). We write $|\nu|$ for the variation measure of ν which is given by

$$|\nu|(A) := \sup\left\{\sum_{i=1}^{\infty}|\nu(A_i)| \,:\, A_i \text{ disjoint relatively compact Borel subsets of } A\right\}$$

for *all* Borel subsets A of Ω. Evidently, $|\nu|$ is always a non-negative measure, and, in the particular case of a signed ν with only non-negative values, it provides a canonical extension to the full Borel-σ-algebra (this partially justifies our terminology). We call ν finite or locally finite if $|\nu|$ has the respective property in the usual sense, and we observe that also every finite ν has a unique extension — which we identify with ν in the following — to a σ-additive function on the full Borel-σ-algebra of Ω. In the locally finite case, in contrast, such an extension need not necessarily exist.

Integration with respect to (locally) finite measures ν is explained and notated as usual. Additionally, we agree on the less usual abbreviations $\int_A \nu := \int_A \mathrm{d}\nu = \nu(A)$ and $f\nu(A) = \int_A f\nu := \int_A f\,\mathrm{d}\nu$ for (relatively compact) Borel subsets A of Ω and Borel functions f on A. Finally, the symbol $\mathrm{d}x$ is used synonymous with $\mathrm{d}\mathscr{L}^n$ and indicates integration with respect to the Lebesgue measure $\mathscr{L}^n$.

For a non-negative measure μ and an $\mathbb{R}^m$-valued measure ν on Ω we say that ν is absolutely continuous with respect to μ if $|\nu|(A) = 0$ holds for all Borel subsets A of Ω with $\mu(A) = 0$. We record that if ν is (locally) finite and absolutely continuous with respect to μ, then by the Radon-Nikodým theorem there exists a Borel function $h\colon \Omega \to \mathbb{R}^m$ such that $\nu(A) = \int_A h\,\mathrm{d}\mu$ holds for all (relatively compact) Borel subsets A of Ω. The density h is μ-a. e. uniquely determined and is denoted by $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$. In addition, μ-a. e. we have $\left|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right| = \frac{\mathrm{d}|\nu|}{\mathrm{d}\mu}$. Finally, we say that ν is singular to μ if for some Borel subset A of Ω we have $|\nu|(A) = 0 = \mu(\Omega \setminus A)$.

By Lebesgue decomposition every (locally) finite $\mathbb{R}^m$-valued measure ν on Ω can be written as

$$\nu = \nu^{\mathrm{a}} + \nu^{\mathrm{s}}\,,$$

where ν^{a} and ν^{s} are uniquely determined (locally) finite $\mathbb{R}^m$-valued measures on Ω such that ν^{a} is absolutely continuous with respect to $\mathscr{L}^n$ and ν^{s} is singular to $\mathscr{L}^n$. Moreover, $|\nu| = |\nu^{\mathrm{a}}| + |\nu^{\mathrm{s}}|$ is the Lebesgue decomposition of $|\nu|$.

4.4. **Function spaces.**

4.4.1. *Lebesgue, Sobolev, and Hölder spaces.* Our notation for Lebesgue, Sobolev, and Hölder spaces (mainly L^q, $\mathrm{W}^{1,p}$, $\mathrm{C}^{1,\alpha}$) is quite standard and we just mention a few additional conventions.

We use $\mathrm{L}^q(\Omega)$ for all values $q \in [0,\infty]$: extending the common definition, $\mathrm{L}^q(\Omega)$ with $0 < q < 1$ is defined as the collection of all Lebesgue measurable functions $w\colon \Omega \to \mathbb{R}$ such that the integral $\int_\Omega |w|^q\,\mathrm{d}x$ is finite. Moreover, the qth root of this integral is denoted by $\|w\|_{\mathrm{L}^q(\Omega)}$ — regardless of the fact that for $0 < q < 1$ it is not a norm (but a quasinorm). Finally, $\mathrm{L}^0(\Omega)$ is the space of *all* Lebesgue measurable functions $\Omega \to \mathbb{R}$, and we adopt the convention that $\|w\|_{\mathrm{L}^0(\Omega)}$ stands for the essential infimum of $|w|$ on Ω.

We identify vector-valued functions with tuples of $\mathbb{R}$-valued functions. Consequently, notations like $\mathrm{L}^q(\Omega)^N$ correspond to spaces of $\mathbb{R}^N$-valued functions. Moreover, we use a subscript $_{\mathrm{loc}}$ to indicate that a function is of the required class on all balls $\mathrm{B}_\varrho(x_0) \subset\subset \Omega$.

4.4.2. *Functions of bounded variation.* The space $\mathrm{BV}(\Omega)$ of functions of bounded variation on Ω is defined as the collection of all functions $w \in \mathrm{L}^1(\Omega)$ whose distributional gradient can be represented by a finite $\mathbb{R}^n$-valued measure $\mathrm{D}w$. We also use the variants $\mathrm{BV}(\Omega)^N$ and $\mathrm{BV}_{\mathrm{loc}}(\Omega)^N$ as explained above, and in these cases we understand $\mathrm{D}w$ as a (locally) finite $\mathbb{R}^{Nn}$-valued measure.

In contrast to the notation $\mathrm{D}w$ for the gradient measure we use ∇w for the density $\frac{\mathrm{d}(\mathrm{D}w)^{\mathrm{a}}}{\mathrm{d}\mathscr{L}^n}$ of its absolutely continuous part $(\mathrm{D}w)^{\mathrm{a}}$. In particular, whenever the distributional derivative of w can be represented by an $(\mathrm{L}^1_{\mathrm{loc}}$-)function, we denote this function by ∇w.

4.4.3. *Morrey spaces.* Morrey spaces are only used in Section 3. Our terminology for them is as follows.

Definition 4.9 (Morrey spaces)**.** *Fix $q, \alpha \in [0,\infty)$. Writing $\operatorname{diam}\Omega$ for the diameter of Ω, the Morrey space $\mathrm{L}^{q,\alpha}(\Omega)$ is defined as the collection of all $w \in \mathrm{L}^q(\Omega)$ whose Morrey bound*

$$\sup_{\substack{x\in\Omega \\ 0<\varrho<\operatorname{diam}\Omega}} \varrho^{-\alpha} \int_{\Omega\cap \mathrm{B}_\varrho(x_0)} |w|^q \,\mathrm{d}x$$

is finite (where in the case $q = 0$ of this definition, we exceptionally adopt the convention $0^0 = 0$). The local Morrey spaces $\mathrm{L}^{q,\alpha}_{\mathrm{loc}}(\Omega)$ are then defined in the sense of the preceding convention.

We notice $\mathrm{L}^{0,\alpha}(\Omega) = \mathrm{L}^0(\Omega)$ for $\alpha \le n$, $\mathrm{L}^{q,0}(\Omega) = \mathrm{L}^q(\Omega)$, $\mathrm{L}^{q,n}(\Omega) = \mathrm{L}^\infty(\Omega)$ for $q > 0$, and $\mathrm{L}^{q,\alpha}(\Omega) = \{0\}$ for $\alpha > n$. Moreover, Hölder's inequality yields the embedding

$$\mathrm{L}^{\widetilde{q},n-\widetilde{\alpha}}(\Omega) \subset \mathrm{L}^{q,n-\alpha}(\Omega) \qquad \text{for } \widetilde{q} \ge \max\{1, \widetilde{\alpha}/\alpha\}q\,. \tag{4.10}$$

4.5. **Functionals of measures.** With the terminology of Section 4.3 we can give the following definitions in the spirit of Goffman & Serrin [36].

Definition 4.10 (recession function)**.** *If $\varphi\colon \Omega \times \mathbb{R}^m \to \mathbb{R}$ is convex in its second argument, we introduce the recession function $\varphi^\infty\colon \Omega\times\mathbb{R}^m \to (-\infty,\infty]$ of φ by*

$$\varphi^\infty(x,z) := \lim_{s\searrow 0} s\varphi(x,z/s) \qquad \textit{for } (x,z) \in \Omega\times\mathbb{R}^m\,.$$

Under the assumptions of Definition 4.10, φ^∞ is convex and 1-homogeneous in its second argument. Moreover, if φ is a Borel function, then φ^∞ is a Borel function as well.

Definition 4.11 (functionals of measures)**.** *Consider a locally finite $\mathbb{R}^m$-valued measure ν on Ω and a Borel function $\varphi\colon \Omega\times\mathbb{R}^m \to \mathbb{R}$ which is convex in its second argument. Then we introduce a new signed measure $\varphi(\,\cdot\,,\nu)$ on Ω by letting*

$$\varphi(\,\cdot\,,\nu)(A) = \int_A \varphi(\,\cdot\,,\nu) := \int_A \varphi\Big(\,\cdot\,,\frac{\mathrm{d}\nu^a}{\mathrm{d}\mathscr{L}^n}\Big)\,\mathrm{d}x + \int_A \varphi^\infty\Big(\,\cdot\,,\frac{\mathrm{d}\nu^{\mathrm{s}}}{\mathrm{d}|\nu^{\mathrm{s}}|}\Big)\,\mathrm{d}|\nu^{\mathrm{s}}|$$

for all relatively compact Borel subsets A of Ω.

We observe that, if φ has linear growth in the sense of (1.1), then $\varphi(\,\cdot\,,\nu)$ inherits (local) finiteness from ν. Furthermore, we remark that our notation for the variation measure is consistent with the choice $\varphi(x,z)=|z|$ in Definition 4.11.

In this paper we mostly use Definition 4.11 with $m=Nn$ for the finite $\mathbb{R}^{Nn}$-valued gradient measures $\mathrm{D}w$ of functions $w\in\mathrm{BV}(\Omega)^N$: in this setting the definition enables us to give a precise meaning to the functionals $F[w;\Omega]=\int_\Omega f(\,\cdot\,,\mathrm{D}w)$ in (1.2) — and as special case also to those from (2.1). For further motivation and discussion of such functionals of BV-functions as well as for corresponding existence results for minimizers we refer for instance to [34, 8, 13, 9]; notice however that we do not need to take into account the additional boundary terms in [34, 9], as in this paper we are exclusively concerned with interior regularity properties.

4.6. **Mean values, Lebesgue points, and Lebesgue values of measures.** For the mean values of a function $w\in\mathrm{L}^1(\Omega)^N$ and a finite $\mathbb{R}^m$-valued measure ν on Ω we introduce the notations

$$w_\Omega:=\fint_\Omega w\,\mathrm{d}x:=\frac{1}{\mathscr{L}^n(\Omega)}\int_\Omega w\,\mathrm{d}x \qquad\text{and}\qquad \nu_\Omega:=\fint_\Omega\nu:=\frac{\nu(\Omega)}{\mathscr{L}^n(\Omega)}\,.$$

Next we state a Jensen type inequality with mean values of measures. A similar inequality has been proved in [8, Proposition 2.4] via approximation, but here we suggest a different argument which works directly with the given measure.

Lemma 4.12 (Jensen inequality). *If ν is a finite $\mathbb{R}^m$-valued measure on Ω, then for any convex function $\varphi\colon\mathbb{R}^m\to\mathbb{R}$ and any $\xi\in\mathbb{R}^m$ there holds*

$$\varphi\Big(\fint_\Omega\nu\Big)\le\fint_\Omega\varphi(\nu)\,.$$

Proof. Following [34] we define a 1-homogeneous, lower semicontinuous, and convex function $\overline{\varphi}\colon[0,\infty)\times\mathbb{R}^{Nn}\to(-\infty,\infty]$ by $\overline{\varphi}(s,z):=s\varphi(z/s)$ for $s>0$ and by $\overline{\varphi}(s,z):=\varphi^\infty(z)$ for $s=0$. Then we set $\mu:=\mathscr{L}^n+|\nu^{\mathrm{s}}|$ and estimate

$$\begin{aligned}\varphi\Big(\fint_\Omega\nu\Big)=\frac{\mu(\Omega)}{\mathscr{L}^n(\Omega)}\overline{\varphi}\Big(\frac{\mathscr{L}^n(\Omega)}{\mu(\Omega)},\frac{\nu(\Omega)}{\mu(\Omega)}\Big)&=\frac{\mu(\Omega)}{\mathscr{L}^n(\Omega)}\overline{\varphi}\bigg(\frac{1}{\mu(\Omega)}\int_\Omega\Big(\frac{\mathrm{d}\mathscr{L}^n}{\mathrm{d}\mu},\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\Big)\,\mathrm{d}\mu\bigg)\\&\le\frac{1}{\mathscr{L}^n(\Omega)}\int_\Omega\overline{\varphi}\Big(\frac{\mathrm{d}\mathscr{L}^n}{\mathrm{d}\mu},\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\Big)\,\mathrm{d}\mu=:I\,,\end{aligned}$$

where we used the common version of Jensen's inequality in the last step. For a Borel subset B of Ω with $|\nu^{\mathrm{s}}|(B)=0=\mathscr{L}^n(\Omega\setminus B)$ we now integrate separately over B and $\Omega\setminus B$, and we exploit the above choice of μ. In this way I can be rewritten as

$$\begin{aligned}I&=\frac{1}{\mathscr{L}^n(\Omega)}\bigg[\int_B\overline{\varphi}\Big(1,\frac{\mathrm{d}\nu^{\mathrm{a}}}{\mathrm{d}\mathscr{L}^n}\Big)\,\mathrm{d}\mathscr{L}^n+\int_{\Omega\setminus B}\overline{\varphi}\Big(0,\frac{\mathrm{d}\nu^{\mathrm{s}}}{\mathrm{d}|\nu^{\mathrm{s}}|}\Big)\,\mathrm{d}|\nu^{\mathrm{s}}|\bigg]\\&=\frac{1}{\mathscr{L}^n(\Omega)}\bigg[\int_\Omega\varphi\Big(\frac{\mathrm{d}\nu^{\mathrm{a}}}{\mathrm{d}\mathscr{L}^n}\Big)\,\mathrm{d}\mathscr{L}^n+\int_\Omega\varphi^\infty\Big(\frac{\mathrm{d}\nu^{\mathrm{s}}}{\mathrm{d}|\nu^{\mathrm{s}}|}\Big)\,\mathrm{d}|\nu^{\mathrm{s}}|\bigg]=\fint_\Omega\varphi(\nu)\,,\end{aligned}$$

and we arrive at the claim. □

In a certain sense we now assign pointwise values to $\mathbb{R}^m$-valued measures. Once more we mostly use this concept with $m=Nn$ for the gradient measure $\mathrm{D}w$ of $w\in\mathrm{BV}_{\mathrm{loc}}(\Omega)^N$.

Definition 4.13 (Lebesgue points and Lebesgue values)**.** *Consider a locally finite $\mathbb{R}^m$-valued measure ν on Ω. We say that $x_0 \in \Omega$ is a Lebesgue point of ν if there exists some $z_0 \in \mathbb{R}^m$ such that*

$$\lim_{R \searrow 0} \fint_{\mathrm{B}_R(x_0)} |\nu - z_0 \mathscr{L}^n| = 0\,. \tag{4.11}$$

If x_0 is a Lebesgue point of ν, then z_0 is uniquely determined by (4.11) and is called the Lebesgue value of ν at x_0.

By the Lebesgue differentiation theorem for measures, $\mathscr{L}^n$-almost every $x_0 \in \Omega$ is a Lebesgue point of a locally finite $\mathbb{R}^m$-valued measure ν on Ω, and moreover the function which maps x_0 to z_0 is a representative of the density $\frac{\mathrm{d}\nu^{\mathrm{a}}}{\mathrm{d}\mathscr{L}^n}$.

4.7. **Energy and excess functionals.** Throughout this subsection we again fix $1 < p < \infty$. Relying on the auxiliary functions of Section 4.2 and the notation of Sections 4.3, 4.5, and 4.6 we now introduce for $u \in \mathrm{BV}(\Omega)^N$ the scaled energy

$$\mathrm{E}(u, \Omega) := \fint_\Omega \mathrm{A}_p(|\mathrm{D}u|^p)$$

and the excess

$$\Phi(u, \Omega) := \fint_\Omega \mathrm{A}_p(\mathrm{W}_p^{(\mathrm{D}u)_\Omega}(\mathrm{D}u - (\mathrm{D}u)_\Omega \mathscr{L}^n))\,.$$

Here, by Lemma 4.7 the excess $\Phi(u, \Omega)$ is well-defined (at least) for $|(\mathrm{D}u)_\Omega| \le \frac{1}{2}$.

Occasionally, we will also use the respective p-integrable quantities. Those are defined for $w \in \mathrm{W}^{1,p}(\Omega)^N$ as

$$\mathrm{E}^*(w, \Omega) := \fint_\Omega |\nabla w|^p \,\mathrm{d}x$$

and

$$\Phi^*(w, \Omega) := \fint_\Omega \mathrm{W}_p^{(\nabla w)_\Omega}(\nabla w - (\nabla w)_\Omega) \,\mathrm{d}x\,.$$

We now record a number of useful estimates for $\Phi(u, \Omega)$ and $\Phi^*(w, \Omega)$.

Lemma 4.14. *For all $u \in \mathrm{BV}(\Omega)^N$ with $|(\mathrm{D}u)_\Omega| \le \frac{1}{2}$ and all $\xi \in \mathbb{R}^{Nn}$ with $|\xi| \le \frac{1}{2}$ there holds*

$$\Phi(u, \Omega) \le C \fint_\Omega \mathrm{A}_p(\mathrm{W}_p^\xi(\mathrm{D}u - \xi \mathscr{L}^n))\,,$$

where C depends only on p.

Proof. Using (4.6) we find

$$\Phi(u, \Omega) \le C \left[\mathrm{A}_p(\mathrm{W}_p^\xi((\mathrm{D}u)_\Omega - \xi)) + \fint_\Omega \mathrm{A}_p(\mathrm{W}_p^\xi(\mathrm{D}u - \xi \mathscr{L}^n)) \right].$$

Moreover, the Jensen inequality of Lemma 4.12, applied with the convex function of Lemma 4.7, yields

$$\mathrm{A}_p(\mathrm{W}_p^\xi((\mathrm{D}u)_\Omega - \xi)) \le \fint_\Omega \mathrm{A}_p(\mathrm{W}_p^\xi(\mathrm{D}u - \xi \mathscr{L}^n))\,.$$

Combining the last two estimates we arrive at the claim. $\square$

Lemma 4.15. *For all $w \in \mathrm{W}^{1,p}(\Omega)^N$ and $\xi \in \mathbb{R}^{Nn}$ there hold*

$$\Phi^*(w,\Omega) \le C \fint_\Omega \mathrm{W}_p^\xi(\nabla w - \xi)\,\mathrm{d}x\,, \tag{4.12}$$

$$\Phi^*(w,\Omega) \le C \fint_\Omega \mathrm{W}_p^{(\nabla w)_\Omega}(\nabla w - \xi)\,\mathrm{d}x\,, \tag{4.13}$$

where C depends only on p.

Proof. Basically, the proof of Lemma 4.14 applies in a simplified version, where we just consider functions instead of measures, and where A_p does not occur. Accordingly, in place of (4.6) we apply either (4.5) (for the first claim) or (4.7) (for the second claim). Moreover, it suffices to use the standard version of Jensen's inequality for the convex function W_p^ξ. □

Lemma 4.16. *For all $w, h \in \mathrm{W}^{1,p}(\Omega)^N$ there holds*

$$\Phi^*(w,\Omega) \le C\left[\Phi^*(h,\Omega) + \fint_\Omega \mathrm{W}_p^{(\nabla w)_\Omega}(\nabla w - \nabla h)\,\mathrm{d}x\right]$$

with a constant C depending only on p.

Proof. We first assume that we have

$$\text{either } p \ge 2, |(\nabla w)_\Omega| \ge |(\nabla h)_\Omega| \text{ or } p \le 2, |(\nabla w)_\Omega| \le |(\nabla h)_\Omega|\,. \tag{4.14}$$

Then we use in turn (4.12), (4.7), and (4.14) to estimate

$$\begin{aligned}
\Phi^*(w,\Omega) &\le C \fint_\Omega \mathrm{W}_p^{(\nabla h)_\Omega}(\nabla w - (\nabla h)_\Omega)\,\mathrm{d}x \\
&\le C\left[\fint_\Omega \mathrm{W}_p^{(\nabla h)_\Omega}(\nabla h - (\nabla h)_\Omega)\,\mathrm{d}x + \fint_\Omega \mathrm{W}_p^{(\nabla h)_\Omega}(\nabla w - \nabla h)\,\mathrm{d}x\right] \\
&\le C\left[\Phi^*(h,\Omega) + \fint_\Omega \mathrm{W}_p^{(\nabla w)_\Omega}(\nabla w - \nabla h)\,\mathrm{d}x\right].
\end{aligned}$$

In the remaining cases we have

$$\text{either } p \ge 2, |(\nabla w)_\Omega| \le |(\nabla h)_\Omega| \text{ or } p \le 2, |(\nabla w)_\Omega| \ge |(\nabla h)_\Omega|\,. \tag{4.15}$$

In this situation we get via (4.13), (4.7), and (4.15)

$$\begin{aligned}
\Phi^*(w,\Omega) &\le C \fint_\Omega \mathrm{W}_p^{(\nabla w)_\Omega}(\nabla w - (\nabla h)_\Omega)\,\mathrm{d}x \\
&\le C\left[\fint_\Omega \mathrm{W}_p^{(\nabla w)_\Omega}(\nabla h - (\nabla h)_\Omega)\,\mathrm{d}x + \fint_\Omega \mathrm{W}_p^{(\nabla w)_\Omega}(\nabla w - \nabla h)\,\mathrm{d}x\right] \\
&\le C\left[\Phi^*(h,\Omega) + \fint_\Omega \mathrm{W}_p^{(\nabla w)_\Omega}(\nabla w - \nabla h)\,\mathrm{d}x\right].
\end{aligned}$$ □

4.8. **Regularity estimates for comparison systems.** We first record some basic estimates for weak solutions of linear systems; see for instance [32, Lemma 5.20] for the interior Campanato type estimate (4.17) and [32, Theorem 7.1] for the global $\mathrm{W}^{1,p}$-estimate (4.18). Finally, the global L^q-estimate (4.19) follows from (4.18) via a Poincaré inequality.

Proposition 4.17 (estimates for linear systems)**.** *For a positive symmetric bilinear form $\mathcal{B}$ on $\mathbb{R}^{Nn}$ and $w \in \mathrm{W}^{1,\infty}(\mathrm{B}_{R/2}(x_0))^N$ suppose that $h \in w + \mathrm{W}^{1,2}_0(\mathrm{B}_{R/2}(x_0))^N$ satisfies*

$$\int_{\mathrm{B}_{R/2}(x_0)} \mathcal{B}(\nabla h, \nabla \varphi)\,\mathrm{d}x = 0 \qquad \text{for all } \varphi \in \mathrm{W}^{1,2}_0(\mathrm{B}_{R/2}(x_0))^N\,. \tag{4.16}$$

Then there holds

$$\fint_{\mathrm{B}_{2\tau R}(x_0)} |\nabla h - (\nabla h)_{\mathrm{B}_{2\tau R}(x_0)}|^2\,\mathrm{d}x \le C\tau^2 \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla h - (\nabla h)_{\mathrm{B}_{R/2}(x_0)}|^2\,\mathrm{d}x \tag{4.17}$$

for all $\tau \in (0, \frac14]$. Moreover, for every $p \in [2,\infty)$ one has

$$\fint_{\mathrm{B}_{R/2}(x_0)} |\nabla h|^p\,\mathrm{d}x \le C \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w|^p\,\mathrm{d}x\,, \tag{4.18}$$

and for every $q \in [1,\infty]$ there holds

$$\|h - h_{\mathrm{B}_{R/2}(x_0)}\|_{\mathrm{L}^q(\mathrm{B}_{R/2}(x_0))^N} \le C R^{1+\frac{n}{q}} \sup_{\mathrm{B}_{R/2}(x_0)} |\nabla w|\,. \tag{4.19}$$

Here C depends only on n, N, an upper bound for the ellipticity ratio of $\mathcal{B}$ (that means the quotient of the largest and the smallest eigenvalue of $\mathcal{B}$), and in case of (4.18) also on p.

In the following theorem we collect similar estimates for weak solutions of the p-Laplace system, that is the Euler-Lagrange system associated with the p-energy density e_p from (2.2) (notice $\nabla \mathrm{e}_p(z) = |z|^{p-2}z$ for $z \in \mathbb{R}^{Nn}$). In particular, the first estimate of the theorem is essentially equivalent with interior $\mathrm{C}^{1,\alpha}$ regularity, first proved in Uhlenbeck's famous paper [58] for $p \ge 2$ and then extended to all $p > 1$ in [56].

Theorem 4.18 (estimates for the p-Laplace system)**.** *For $1 < p < \infty$ and $w \in \mathrm{W}^{1,\infty}(\mathrm{B}_{R/2}(x_0))^N$ suppose that $h \in w + \mathrm{W}^{1,p}_0(\mathrm{B}_{R/2}(x_0))^N$ satisfies*

$$\int_{\mathrm{B}_{R/2}(x_0)} \nabla \mathrm{e}_p(\nabla h)\nabla\varphi\,\mathrm{d}x = 0 \qquad \text{for all } \varphi \in \mathrm{W}^{1,p}_0(\mathrm{B}_{R/2}(x_0))^N\,. \tag{4.20}$$

Then one has

$$\Phi^*(h, \mathrm{B}_{2\tau R}(x_0)) \le C\tau^{\gamma_p}\,\Phi^*(h, \mathrm{B}_{R/2}(x_0)) \tag{4.21}$$

for all $\tau \in (0, \frac14]$ and

$$\fint_{\mathrm{B}_{R/2}(x_0)} |\nabla h|^{p+\kappa_{\mathrm{d}}}\,\mathrm{d}x \le C \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w|^{p+\kappa_{\mathrm{d}}}\,\mathrm{d}x\,. \tag{4.22}$$

Moreover, for all $q \in [1,\infty]$ there holds

$$\|h - h_{\mathrm{B}_{R/2}(x_0)}\|_{\mathrm{L}^q(\mathrm{B}_{R/2}(x_0))^N} \le C R^{1+\frac{n}{q}} \sup_{\mathrm{B}_{R/2}(x_0)} |\nabla w|\,. \tag{4.23}$$

Here, $\gamma_p \in (0,2]$ and $\kappa_{\mathrm{d}} \in (0,\infty)$ are constants, fixed for the whole paper, which depend only on n, N, and p. In addition, also C depends only on n, N, and p.

We now explain how the estimates (4.21), (4.22), and (4.23) can be extracted from the existing literature. To this end we first record that h is the unique minimizer of the p-energy in $w + \mathrm{W}^{1,p}_0(\mathrm{B}_{R/2}(x_0))^N$, and in particular we thus have

$$\int_{\mathrm{B}_{R/2}(x_0)} \mathrm{e}_p(\nabla h)\,\mathrm{d}x \le \int_{\mathrm{B}_{R/2}(x_0)} \mathrm{e}_p(\nabla w)\,\mathrm{d}x\,. \tag{4.24}$$

A decay estimate similar to (4.21) can be found in [33, Theorem 3.1] for the case $p \geq 2$ and in [37, Theorem 4.1] for all $1 < p < \infty$. Precisely, from these references we infer

$$
\begin{aligned}
(4.25)\quad & \fint_{\mathrm{B}_{2\tau R}(x_0)} \big|\mathrm{V}_p(\nabla h) - \big[\mathrm{V}_p(\nabla h)\big]_{\mathrm{B}_{2\tau R}(x_0)}\big|^2 \,\mathrm{d}x \\
& \qquad \leq C\tau^{\gamma_p} \fint_{\mathrm{B}_R(x_0)} \big|\mathrm{V}_p(\nabla h) - \big[\mathrm{V}_p(\nabla h)\big]_{\mathrm{B}_R(x_0)}\big|^2 \,\mathrm{d}x \,,
\end{aligned}
$$

with the function V_p defined in (4.3). In order to convert this estimate into (4.21) we argue as follows: on the right-hand side of (4.25) we estimate with the help of (4.4)

$$
\begin{aligned}
\fint_{\mathrm{B}_R(x_0)} \big|\mathrm{V}_p(\nabla h) - \big[\mathrm{V}_p(\nabla h)\big]_{\mathrm{B}_R(x_0)}\big|^2 \,\mathrm{d}x &\leq \fint_{\mathrm{B}_R(x_0)} \big|\mathrm{V}_p(\nabla h) - \mathrm{V}_p\big((\nabla h)_{\mathrm{B}_R(x_0)}\big)\big|^2 \,\mathrm{d}x \,, \\
&\leq C\Phi^*(h, \mathrm{B}_R(x_0)) \,,
\end{aligned}
$$

and for the left-hand side of (4.25) we use that by (4.12) and (4.4) we have

$$
\Phi^*(h, \mathrm{B}_{2\tau R}(x_0)) \leq C \fint_{\mathrm{B}_{2\tau R}(x_0)} \mathrm{W}_p^\xi(\nabla h - \xi) \,\mathrm{d}x \leq C \fint_{\mathrm{B}_{2\tau R}(x_0)} \big|\mathrm{V}_p(\nabla h) - \mathrm{V}_p(\xi)\big|^2 \,\mathrm{d}x \,,
$$

where $\xi \in \mathbb{R}^{Nn}$ is such that $\mathrm{V}_p(\xi) = \big[\mathrm{V}_p(\nabla h)\big]_{\mathrm{B}_{2\tau R}(x_0)}$. Combining the last three estimates we arrive at (4.21). Finally, we mention that in the case $p \geq 2$ the estimate (4.21) can be directly inferred from [30, Theorem 4.2].

The global higher gradient integrability (4.22) is proved in [24, Lemma 3.2] with an additional term on the right-hand side, but this extra term can easily be eliminated via (4.24).

Finally, we come to the global L^q-estimate (4.23): in the case[4] $q \leq p^*$ it follows easily from a Sobolev-Poincaré inequality and (4.24); otherwise we obtain (4.23) as a special case of [42, Theorem 2], first for $q = \infty$ and as a consequence for all q.

4.9. **Some estimates for mollifications.** For $u \in \mathrm{BV}(\mathrm{B}_R(x_0))$ and $0 < \lambda < R$ we will extensively work with the (mean-)mollifications

$$
u_\lambda(x) := u_{\mathrm{B}_\lambda(x)} \qquad \text{for } x \in \mathrm{B}_{R-\lambda}(x_0) \,.
$$

We notice that u_λ is in $\mathrm{W}^{1,\infty}(\mathrm{B}_{R-\lambda}(x_0))$ since we have

$$
|\nabla u_\lambda(x)| = |(\mathrm{D}u)_{\mathrm{B}_\lambda(x)}| \leq \frac{|\mathrm{D}u|(\mathrm{B}_R(x_0))}{\mathscr{L}^n(\mathrm{B}_\lambda)} \qquad \text{for all } x \in \mathrm{B}_{R-\lambda}(x_0) \,.
$$

Following [49, 8] we will choose the smoothing radius λ depending on the excess $\Phi(u, \mathrm{B}_R(x_0))$ on the ball under consideration. We provide some estimates for the corresponding mollifications, which are quite close to those in [8, Lemma 4.2].

Lemma 4.19. *For $u \in \mathrm{BV}(\mathrm{B}_R(x_0))^N$ set*

$$
\xi_0 := (\mathrm{D}u)_{\mathrm{B}_R(x_0)} \,,
$$

and assume $|\xi_0| \leq \frac{1}{2}$. If there holds

$$
\left\{ \begin{array}{l} \Phi(u, \mathrm{B}_R(x_0)) \leq 1 \\ \Phi(u, \mathrm{B}_R(x_0)) \leq |\xi_0|^p \neq 0 \end{array} \right\} ,
$$

[4] As usual we understand $p^* := \frac{np}{n-p}$ in the case $p < n$, while $q \leq p^*$ stands for $q < \infty$ in the case $p = n$ and for $q \leq \infty$ in the case $p > n$.

then choosing

$$\left\{\begin{aligned}\lambda &:= \frac{1}{2}\Phi(u,\mathrm{B}_R(x_0))^{\frac{1}{2n}}R\\ \lambda &:= \frac{1}{2}\big(|\xi_0|^{-p}\Phi(u,\mathrm{B}_R(x_0))\big)^{\frac{1}{2n}}R\end{aligned}\right\}$$

we have

$$\left\{\begin{aligned}&\sup_{\mathrm{B}_{R/2}(x_0)}|\nabla u_\lambda-\xi_0|\le C\Phi(u,\mathrm{B}_R(x_0))^{\frac{1}{2\max\{p,2\}}}\\ &\sup_{\mathrm{B}_{R/2}(x_0)}|\nabla u_\lambda-\xi_0|\le C\big(|\xi_0|^{-p}\Phi(u,\mathrm{B}_R(x_0))\big)^{\frac{1}{2\max\{p,2\}}}|\xi_0|\end{aligned}\right\}. \tag{4.26}$$

Here, the statement is valid choosing always the first alternative inside $\{\ldots\}$*, and it is also true choosing always second one. Moreover, in both cases we have the energy and excess estimates*

$$\mathrm{E}^*(u_\lambda,\mathrm{B}_{R/2}(x_0))\le C\mathrm{E}(u,\mathrm{B}_R(x_0))\,,$$
$$\fint_{\mathrm{B}_{R/2}(x_0)}\mathrm{W}_p^{\xi_0}(\nabla u_\lambda-\xi_0)\,\mathrm{d}x\le C\Phi(u,\mathrm{B}_R(x_0))\,,$$

and all the constants C in this lemma depend only on n and p.

Proof. Using the Jensen inequality of Lemma 4.12 for the convex function $z\mapsto \mathrm{A}_p(\mathrm{W}_p^{\xi_0}(z))$ of Lemma 4.7 we find for $x\in\mathrm{B}_{R/2}(x_0)$

$$\begin{aligned}\mathrm{A}_p(\mathrm{W}_p^{\xi_0}(\nabla u_\lambda(x)-\xi_0)) &= \mathrm{A}_p\bigg(\mathrm{W}_p^{\xi_0}\bigg(\fint_{\mathrm{B}_\lambda(x)}(\mathrm{D}u-\xi_0\mathscr{L}^n)\bigg)\bigg)\\ &\le \fint_{\mathrm{B}_\lambda(x)}\mathrm{A}_p(\mathrm{W}_p^{\xi_0}(\mathrm{D}u-\xi_0\mathscr{L}^n))\\ &\le \Big(\frac{R}{\lambda}\Big)^n\Phi(u,\mathrm{B}_R(x_0))\\ &= \left\{\begin{aligned}&2^n\Phi(u,\mathrm{B}_R(x_0))^{\frac12}\\ &2^n\Big(|\xi_0|^{-p}\Phi(u,\mathrm{B}_R(x_0))\Big)^{\frac12}|\xi_0|^p\end{aligned}\right\}.\end{aligned}$$

From the assumed smallness of $\Phi(u,\mathrm{B}_R(x_0))$ and the bound $|\xi_0|\le\frac12$ we infer that $\mathrm{W}_p^{\xi_0}(\nabla u_\lambda-\xi_0)$ is bounded on $\mathrm{B}_{R/2}(x_0)$ by a constant depending only on n and p. Thus, in view of Lemma 4.5 we also get

$$\sup_{\mathrm{B}_{R/2}(x_0)}\mathrm{W}_p^{\xi_0}(\nabla u_\lambda-\xi_0)\le\left\{\begin{aligned}&C\Phi(u,\mathrm{B}_R(x_0))^{\frac12}\\ &C\Big(|\xi_0|^{-p}\Phi(u,\mathrm{B}_R(x_0))\Big)^{\frac12}|\xi_0|^p\end{aligned}\right\}. \tag{4.27}$$

For $p<2$ we now employ (4.27) on the right-hand side of the elementary inequality $|\nabla u_\lambda-\xi_0|\le C\mathrm{W}_p^{\xi_0}(\nabla u_\lambda-\xi_0)^{\frac1p}+C\big(|\xi_0|^{2-p}\mathrm{W}_p^{\xi_0}(\nabla u_\lambda-\xi_0)\big)^{\frac12}$. Then simplifying via $|\xi_0|\le\frac12$ and the smallness of $\Phi(u,\mathrm{B}_R(x_0))$ the claim (4.26) follows. In the case $p\ge2$ the same conclusion is immediate by the trivial bound $|\nabla u_\lambda-\xi_0|\le\mathrm{W}_p^{\xi_0}(\nabla u_\lambda-\xi_0)^{\frac1p}$. Back to the general case we use the boundedness of $\mathrm{W}_p^{\xi_0}(\nabla u_\lambda-\xi_0)$ on $\mathrm{B}_{R/2}(x_0)$ together with Lemma 4.5, the Jensen inequality of Lemma 4.12, and the estimate

$$\int_{\mathrm{B}_{R/2}(x_0)}\bigg[\fint_{\mathrm{B}_\lambda(x)}\nu\bigg]\,\mathrm{d}x\le\int_{\mathrm{B}_R(x_0)}\nu$$

for non-negative measures ν (note $\lambda \le R/2$ by assumption). In this way we derive

$$\begin{aligned}\fint_{\mathrm{B}_{R/2}(x_0)} \mathrm{W}_p^{\xi_0}(\nabla u_\lambda - \xi_0)\,\mathrm{d}x &\le C\fint_{\mathrm{B}_{R/2}(x_0)} \mathrm{A}_p\bigg(\mathrm{W}_p^{\xi_0}\bigg(\fint_{\mathrm{B}_\lambda(x)}(\mathrm{D}u-\xi_0\mathscr{L}^n)\bigg)\bigg)\,\mathrm{d}x\\ &\le C\fint_{\mathrm{B}_{R/2}(x_0)}\bigg[\fint_{\mathrm{B}_\lambda(x)} \mathrm{A}_p(\mathrm{W}_p^{\xi_0}(\mathrm{D}u-\xi_0\mathscr{L}^n))\bigg]\,\mathrm{d}x\\ &\le C\fint_{\mathrm{B}_R(x_0)} \mathrm{A}_p(\mathrm{W}_p^{\xi_0}(\mathrm{D}u-\xi_0\mathscr{L}^n)) = C\Phi(u,\mathrm{B}_R(x_0))\end{aligned}$$

and we have established all claims apart from the estimate for E^*. This estimate is however obtained by an analogous calculation with 0 in place of ξ_0. $\square$

The following three elementary lemmas are based on the Jensen inequality of Lemma 4.12 and are essentially restatements of Lemma 5.1, Lemma 5.2, and Lemma 5.3 in [8, Section 5]. We restate them in a slightly adapted version with a single mollification step instead of the two-step procedure of [8]. Nevertheless, the proofs are almost unchanged (and even a bit simpler) in our situation, and we do not repeat them.

Lemma 4.20. *Consider a convex function $\varphi\colon \mathbb{R}^{Nn} \to \mathbb{R}$, $u \in \mathrm{BV}(\mathrm{B}_R(x_0))^N$, and positive numbers λ, t_*, and t^* such that there holds*

$$t_* < t^* \le R-\lambda\,.$$

Then there exists some t with $t_* < t < t^*$ such that one has

$$\int_{\mathrm{B}_t(x_0)} \varphi(\nabla u_\lambda)\,\mathrm{d}x - \int_{\mathrm{B}_t(x_0)} \varphi(\mathrm{D}u) \le \frac{2\lambda}{t^*-t_*}\int_{\mathrm{B}_R(x_0)} \varphi(\mathrm{D}u)\,.$$

Lemma 4.21. *Consider a convex function $\varphi\colon \mathbb{R}^{Nn} \to \mathbb{R}$, $u \in \mathrm{BV}(\mathrm{B}_R(x_0))^N$, and positive numbers λ, r_*, r^*, s_*, and s^* such that there holds*

$$r_* < r^* \le s_* < s^* \le R-\lambda\,.$$

Then there exist further radii $r_* < r < r^*$ and $s_* < s < s^*$ such that one has

$$\begin{aligned}\int_{\mathrm{B}_s(x_0)\setminus\mathrm{B}_r(x_0)} \varphi(\nabla u_\lambda)\,\mathrm{d}x &- \int_{\mathrm{B}_s(x_0)\setminus\mathrm{B}_r(x_0)} \varphi(\mathrm{D}u)\\ &\le \bigg(\frac{2\lambda}{r^*-r_*}+\frac{2\lambda}{s^*-s_*}\bigg)\int_{\mathrm{B}_R(x_0)}\varphi(\mathrm{D}u)\,.\end{aligned}$$

Lemma 4.22 (Poincaré type inequality)**.** *Let a non-decreasing, convex function $\varphi\colon [0,\infty) \to \mathbb{R}$, $u \in \mathrm{BV}(\mathrm{B}_R(x_0))^N$, and positive numbers λ, s, and t be given such that there hold*

$$\lambda \le s < t \le R-\lambda \qquad \text{and} \qquad t-s \ge \lambda\,.$$

Then one has

$$\int_{\mathrm{B}_t(x_0)\setminus\mathrm{B}_s(x_0)} \varphi\Big(\frac{|u-u_\lambda|}{t-s}\Big)\,\mathrm{d}x \le \int_{\mathrm{B}_{t+\lambda}(x_0)\setminus\mathrm{B}_{s-\lambda}(x_0)} \varphi(|\mathrm{D}u|)\,.$$

We mention the following difference between our statement of Lemma 4.22 and its counterpart [8, Lemma 5.3]: the extra assumption $t-s \ge \lambda$ does not occur in [8], but in contrast to [8] it allows us to state the conclusion of lemma without an additional constant factor on the right-hand side. This modification causes only marginal changes in the proof which is quite standard anyway: to establish the lemma for smooth functions u one recalls $u_\lambda(x) = u_{\mathrm{B}_\lambda(x)}$ and argues by integration along line segments and via Jensen's inequality. By approximation the claim then follows for arbitrary $u \in \mathrm{BV}(\mathrm{B}_R(x_0))^N$.

5. Proof of the main result

In this section we prove the main result stated in Theorem 2.5. Moreover, we establish the addenda of Remark 2.6 and Proposition 2.7.

Outline of the proof. Our reasoning is divided into the following four subsections:

5.1. **Estimates for competitors.** We derive estimates for certain competitors in the minimization problem for F. We distinguish between the degenerate case of Section 5.1.1 and the non-degenerate case of Section 5.1.2, which are based on the comparison with p-Laplace and linear systems, respectively.

5.2. **Estimates for almost-minimizers.** We turn to L^q-α-minimizers, and in Section 5.2.1 we establish some crucial estimates based on the minimality property. Distinguishing once more between the degenerate and the non-degenerate case, we combine these minimality estimates with the outcome of Section 5.1, and we state the respective corollaries in Sections 5.2.2 and 5.2.3.

5.3. **Iteration.** We implement an iteration procedure. Specifically, Section 5.3.1 is concerned with a purely non-degenerate case, while Section 5.3.2 finally combines all the previous considerations in the single excess-decay estimate of Proposition 5.12.

5.4. **Conclusion.** We deduce Theorem 2.5, Remark 2.6, and Proposition 2.7 from Proposition 5.12.

Setup and general conventions. In this section we permanently assume that (H1), (H2), (H3), and (H4) are satisfied for some $p \in (1,\infty)$ and $f\colon \mathbb{R}^{Nn} \to \mathbb{R}$ with

$$z_0 = 0\,, \qquad f(0) = 0\,, \qquad \nabla f(0) = 0\,, \qquad \text{and} \qquad \theta = 1\,. \tag{5.1}$$

The normalization (5.1) will be justified in Section 5.4 (but is anyway valid for the most relevant cases of the introduction). Moreover, in this section F always denotes the functional given by (2.1).

Notice that we do *not* generally assume in this section that u is L^q-α-minimizer. However, *when* we make this assumption, then we also suppose

$$q \geq 1\,,$$

which is possible by the discussion after Definition 2.4.

5.1. **Estimates for competitors.** In this subsection we do not use any minimality property, and in fact we supply two kinds of statements: the first one applies to arbitrary $w \in \mathrm{W}^{1,\infty}(\mathrm{B}_{R/2}(x_0))^N$, while the second one specializes to the case that w is the mollification (as in Section 4.9) of an arbitrary $u \in \mathrm{BV}(\mathrm{B}_R(x_0))^N$. The significance of these estimates depends crucially on a quantity $\Psi(w,\mathrm{B}_{R/2}(x_0))$, which measures the deviation of w from being minimizing and is defined as follows.

Definition 5.1 (deviation from minimality)**.** *For $q \in [1,\infty]$ we measure the (scaled) deviation of a function $w \in \mathrm{W}^{1,\infty}(\mathrm{B}_{R/2}(x_0))^N$ from being minimizing for F in terms of the quantity*

$$\Psi(w,\mathrm{B}_{R/2}(x_0)) := (R/2)^{-n}\Big(F[w,\mathrm{B}_{R/2}(x_0)] - \inf_{\mathcal{A}^q_w} F[\,\cdot\,,\mathrm{B}_{R/2}(x_0)]\Big)\,,$$

where the admissible class $\mathcal{A}^q_w$ *is defined as*

$$\Big\{h \in w+\mathrm{W}^{1,1}_0(\mathrm{B}_{R/2}(x_0))^N : \|h-h_{\mathrm{B}_{R/2}(x_0)}\|_{\mathrm{L}^q(\mathrm{B}_{R/2}(x_0))^N} \le C_{\mathcal{A}} R^{1+\frac{n}{q}} \sup_{\mathrm{B}_{R/2}(x_0)} |\nabla w|\Big\}.$$

Here, the constant $C_{\mathcal{A}}$ *is chosen as the maximum of the two constants* C *in* (4.19) *and* (4.23) *(when we use* σ^{-2} *as an upper bound for the ellipticity ratio in Proposition 4.17), and consequently* $C_{\mathcal{A}}$ *depends only on* n*,* N*,* p*, and* σ*.*

We point out that in order to deal just with α-minimizers or local minimizers it suffices to choose $\mathcal{A}^q_w$ as the Dirichlet class $w+\mathrm{W}^{1,1}_0(\mathrm{B}_\varrho(x_0))^N$. The previous more specific choice is only relevant for the treatment of L^q-α-minimizers, where it is used to verify (2.7) for a certain test-function; see the proof of Proposition 5.8 and in particular (5.30).

5.1.1. *Degenerate case.* We record that as a consequence of (H1), (H2), and (5.1) we have

$$|f(z)| \le C|z|^p \qquad \text{for all } z \in \mathbb{R}^{Nn} \tag{5.2}$$

with a constant C depending only on p, Γ, and σ.

Proposition 5.2. *Assume* $p \ge 2$*. If* $w \in \mathrm{W}^{1,\infty}(\mathrm{B}_{R/2}(x_0))^N$ *satisfies*

$$\sup_{\mathrm{B}_{R/2}(x_0)} |\nabla w| \le \mu^{1/\kappa_{\mathrm{d}}} \eta_{\mathrm{d}}(\mu)$$

for a given $\mu \in (0,1]$*, then there holds*

$$\Phi(w,\mathrm{B}_{2\tau R}(x_0)) \le C\tau^{\gamma_p}(1+\tau^{-(n+\gamma_p)\frac{p}{2}}\mu)\Big[\mathrm{E}^*(w,\mathrm{B}_{R/2}(x_0)) + \mu^{-1}\Psi(w,\mathrm{B}_{R/2}(x_0))\Big]$$

for all $0<\tau\le\frac14$*. Here, the constants* $\gamma_p=\gamma_p(n,N,p)$ *and* $\kappa_{\mathrm{d}}=\kappa_{\mathrm{d}}(n,N,p)$ *have been fixed in Theorem 4.18, and* C *depends only on* n*,* N*,* p*,* Γ*, and* σ*.*

Before proving the proposition, we remark that $\Phi(w,\mathrm{B}_{2\tau R}(x_0))$ is indeed well-defined, by the bound $|(\nabla w)_{\mathrm{B}_{2\tau R}(x_0)}| \le \eta_{\mathrm{d}}(\mu) \le \frac12$ and Lemma 4.7. Similar bounds keep Φ well-defined also in the sequel, but will not be highlighted anymore.

Proof. By standard results there exists a minimizer h of the p-energy in the Dirichlet class $w+\mathrm{W}^{1,p}_0(\mathrm{B}_{R/2}(x_0))^N$. This minimizer h is weakly p-harmonic (in the sense that it satisfies (4.20)) and all the estimates of Theorem 4.18 are available for h. In the following we now employ (4.2) and Lemma 4.16. In addition, we apply the easy estimate $\mathrm{W}^\xi_p(z) \le C\big[|\xi|^{p-2}|z|^2+|z|^p\big]$, which exploits the assumption $p\ge 2$, and Hölder's inequality. In this way we deduce

$$\begin{aligned}
&\Phi(w,\mathrm{B}_{2\tau R}(x_0)) \\
&\le \frac{1}{p}\Phi^*(w,\mathrm{B}_{2\tau R}(x_0)) \\
&\le C\Big[\Phi^*(h,\mathrm{B}_{2\tau R}(x_0)) + \fint_{\mathrm{B}_{2\tau R}(x_0)} \mathrm{W}_p^{(\nabla w)_{\mathrm{B}_{2\tau R}(x_0)}}(\nabla w-\nabla h)\,\mathrm{d}x\Big] \\
&\le C\Big[\Phi^*(h,\mathrm{B}_{2\tau R}(x_0)) + \mathrm{E}^*(w,\mathrm{B}_{2\tau R}(x_0))^{\frac{p-2}{p}}\Big(\fint_{\mathrm{B}_{2\tau R}(x_0)} |\nabla w-\nabla h|^p\,\mathrm{d}x\Big)^{\frac{2}{p}} \\
&\qquad + \fint_{\mathrm{B}_{2\tau R}(x_0)} |\nabla w-\nabla h|^p\,\mathrm{d}x\Big].
\end{aligned} \tag{5.3}$$

Using Young's inequality and recalling $\tau \leq \frac{1}{4}$ this estimate simplifies to

$$\Phi(w, \mathrm{B}_{2\tau R}(x_0)) \leq C\bigg[\Phi^*(h, \mathrm{B}_{2\tau R}(x_0)) + \tau^{\gamma_p}\mathrm{E}^*(w, \mathrm{B}_{R/2}(x_0)) \\ + \tau^{\gamma_p-(n+\gamma_p)\frac{p}{2}} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \nabla h|^p \,\mathrm{d}x\bigg]. \tag{5.4}$$

By (4.12) and (4.24) we furthermore get

$$\Phi^*(h, \mathrm{B}_{R/2}(x_0)) \leq C \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla h|^p \,\mathrm{d}x \leq C\mathrm{E}^*(w, \mathrm{B}_{R/2}(x_0)). \tag{5.5}$$

Combining (4.21), (5.4), and (5.5) we end up with

$$\Phi(w, \mathrm{B}_{2\tau R}(x_0)) \\ \leq C\tau^{\gamma_p}\bigg[\mathrm{E}^*(w, \mathrm{B}_{R/2}(x_0)) + \tau^{-(n+\gamma_p)\frac{p}{2}} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \nabla h|^p \,\mathrm{d}x\bigg]. \tag{5.6}$$

Thus, to finish the proof of the proposition it remains to deal with the term $\fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \nabla h|^p \,\mathrm{d}x$. Using Lemma 4.3, Lemma 4.1, and (4.20) we find

$$\begin{aligned}
&\fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \nabla h|^p \,\mathrm{d}x \\
&\leq C \fint_{\mathrm{B}_{R/2}(x_0)} \int_0^1 \int_0^1 |\nabla h + st(\nabla w - \nabla h)|^{p-2} \,\mathrm{d}s\, t \,\mathrm{d}t\, |\nabla w - \nabla h|^2 \,\mathrm{d}x \\
&\leq C \fint_{\mathrm{B}_{R/2}(x_0)} \int_0^1 \int_0^1 \nabla^2 \mathrm{e}_p(\nabla h + st(\nabla w - \nabla h)) \,\mathrm{d}s\, t \,\mathrm{d}t\, (\nabla w - \nabla h, \nabla w - \nabla h) \,\mathrm{d}x \\
&= C \fint_{\mathrm{B}_{R/2}(x_0)} \big[\mathrm{e}_p(\nabla w) - \mathrm{e}_p(\nabla h) - \nabla \mathrm{e}_p(\nabla h)(\nabla w - \nabla h)\big] \,\mathrm{d}x \\
&= C \fint_{\mathrm{B}_{R/2}(x_0)} \big[\mathrm{e}_p(\nabla w) - \mathrm{e}_p(\nabla h)\big] \,\mathrm{d}x \\
&= C\bigg(\fint_{\mathrm{B}_{R/2}(x_0)} \big[\mathrm{e}_p(\nabla w) - f(\nabla w)\big] \,\mathrm{d}x \\
&\qquad + \fint_{\mathrm{B}_{R/2}(x_0)} \big[f(\nabla w) - f(\nabla h)\big] \,\mathrm{d}x + \fint_{\mathrm{B}_{R/2}(x_0)} \big[f(\nabla h) - \mathrm{e}_p(\nabla h)\big] \,\mathrm{d}x\bigg) \\
&=: C(I + II + III)
\end{aligned}$$

Taking into account $\sup_{\mathrm{B}_{R/2}(x_0)} |\nabla w| \leq \mu^{1/\kappa_\mathrm{d}} \eta_\mathrm{d}(\mu) \leq \eta_\mathrm{d}(\mu)$, we get from (2.3) and (5.1)

$$I \leq \mu \mathrm{E}^*(w, \mathrm{B}_{R/2}(x_0)).$$

Moreover, by Definition 5.1 (recall in particular the choice of the constant $C_\mathcal{A}$) and (4.23) the function h is in the admissible class $\mathcal{A}^q_w$ on $\mathrm{B}_{R/2}(x_0)$ and we get

$$II \leq C\Psi(w, \mathrm{B}_{R/2}(x_0)).$$

Finally, we will estimate III taking advantage of the integrability improvement in (4.22). Splitting the domain of integration, and then using (5.2), (2.3), (5.1), (4.24), (4.22) we have

$$\begin{aligned}
III &\leq \frac{C}{R^n}\bigg[\int_{\{|\nabla h|<\eta_\mathrm{d}(\mu)\}} |f(\nabla h) - \mathrm{e}_p(\nabla h)|\,\mathrm{d}x + \int_{\{|\nabla h|\geq\eta_\mathrm{d}(\mu)\}} |\nabla h|^p\,\mathrm{d}x\bigg]\\
&\leq C\bigg[\mu\fint_{\mathrm{B}_{R/2}(x_0)} |\nabla h|^p\,\mathrm{d}x + \eta_\mathrm{d}(\mu)^{-\kappa_\mathrm{d}}\fint_{\mathrm{B}_{R/2}(x_0)} |\nabla h|^{p+\kappa_\mathrm{d}}\,\mathrm{d}x\bigg]\\
&\leq C\bigg[\mu\fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w|^p\,\mathrm{d}x + \eta_\mathrm{d}(\mu)^{-\kappa_\mathrm{d}}\fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w|^{p+\kappa_\mathrm{d}}\,\mathrm{d}x\bigg]\,.
\end{aligned}$$

Remembering $\sup_{\mathrm{B}_{R/2}(x_0)} |\nabla w| \leq \mu^{1/\kappa_\mathrm{d}}\eta_\mathrm{d}(\mu)$ we arrive at

$$III \leq C\mu\mathrm{E}^*(w,\mathrm{B}_{R/2}(x_0))\,,$$

and consequently we have

$$\fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \nabla h|^p\,\mathrm{d}x \leq C\mu\Big[\mathrm{E}^*(w,\mathrm{B}_{R/2}(x_0)) + \mu^{-1}\Psi(w,\mathrm{B}_{R/2}(x_0))\Big]\,.$$

Using this last estimate on the right-hand side of (5.6) we arrive at the claim. □

The case $p<2$ in Proposition 5.2: The statement of the proposition holds analogously for $p<2$ when we just replace the occurrence of $\tau^{-(n+\gamma_p)\frac{p}{2}}$ with $\tau^{-2(n+\gamma_p)}$.

In the proof we modify the last step of (5.3) by usage of the inequality

$$\mathrm{W}_p^{\xi_0}(z) \leq \mathrm{W}_p^{\xi}(z) + |\xi|^{\frac{p}{2}}\sqrt{\mathrm{W}_p^{\xi}(z)} \qquad \text{for } z,\xi,\xi_0 \in \mathbb{R}^{Nn} \tag{5.7}$$

(for $|\xi_0| \geq |\xi|$ this estimate holds trivially, while for $|\xi_0| \leq |\xi|$ one gets it elementarily via $\mathrm{W}_p^{\xi_0}(z) \leq (|\xi_0|+|z|)^{p-1}|z| \leq (|\xi|+|z|)^{p-1}|z| = \mathrm{W}_p^{\xi}(z)+(|\xi|+|z|)^{p-2}|\xi||z|$). Applying (5.7) pointwisely to $(\nabla w - \nabla h, \nabla w, (\nabla w)_{\mathrm{B}_{2\tau R}(x_0)})$ in place of (z,ξ,ξ_0) and reasoning as before, we find variants of (5.4) and (5.6), where — apart from the exponents of τ — only the last integrals have the slightly different shape $\fint_{\mathrm{B}_{2\tau R}(x_0)} \mathrm{W}_p^{\nabla w}(\nabla w - \nabla h)\,\mathrm{d}x$. Starting the following estimates from the comparable quantity $\fint_{\mathrm{B}_{2\tau R}(x_0)}(|\nabla w|+|\nabla h|)^{p-2}|\nabla w - \nabla h|^2\,\mathrm{d}x$, the remainder of the proof works just as before. □

Remark 5.3. *In some cases not all the estimates of Theorem 4.18 are really necessary for the proof of Proposition 5.2. Indeed, if we have $f(z) \leq \mathrm{e}_p(z)$ for all $z \in \mathbb{R}^{Nn}$, then the term III in the above proof is obviously non-positive, and thus (4.22) is not needed. For instance, this happens for $f = \mathrm{m}_p^1$ or $f = \widetilde{\mathrm{m}}_p^1$. Moreover, (4.23) is much simpler in the case $q \leq p^*$, and it is not needed at all in the case $q \leq \frac{n}{n-1}$ of α-minimizers; compare the end of Section 4.8 and the comment after Definition 5.1, respectively.*

Proposition 5.4. *For $u \in \mathrm{BV}(\mathrm{B}_R(x_0))^N$ set*

$$w := u_\lambda$$

with λ from the first alternative in Lemma 4.19. Then for every $\chi > 0$ and every $0 < \tau \leq \frac{1}{4}$ there exists some constant $\varepsilon_\mathrm{d}^* \in (0,1]$, depending only on n, N, p, η_d, χ, and τ, such that for all $Q \geq 0$ the conditions

$$\chi|(\mathrm{D}u)_{\mathrm{B}_R(x_0)}|^p \leq \Phi(u,\mathrm{B}_R(x_0)) + Q \leq \varepsilon_\mathrm{d}^*\,,$$

together imply

$$\Phi(w,\mathrm{B}_{2\tau R}(x_0)) \leq C\tau^{\gamma_p}\Big[\Phi(u,\mathrm{B}_R(x_0)) + Q + (\varepsilon_\mathrm{d}^*)^{-1}\Psi(w,\mathrm{B}_{R/2}(x_0))\Big]\,,$$

where C depends only on n, N, p, Γ, σ, and χ.

Here, the choice of λ and the bound $\Phi(u, \mathrm{B}_R(x_0)) \le \varepsilon^*_{\mathrm{d}} \le 1$ guarantee in particular that $\lambda \le R/2$ holds and hence that w is defined and of class $\mathrm{W}^{1,\infty}$ on $\mathrm{B}_{R/2}(x_0)$. Similarly, $\lambda \le R/2$ is also ensured in Propositions 5.7 and 5.8 below.

Proof. We first assume $p \ge 2$. Setting

$$\xi_0 := (\mathrm{D}u)_{\mathrm{B}_R(x_0)}$$

we have by assumption

$$\Phi(u, \mathrm{B}_R(x_0)) \le \varepsilon^*_{\mathrm{d}} \le 1 \qquad \text{and} \qquad |\xi_0| \le (\varepsilon^*_{\mathrm{d}}/\chi)^{\frac{1}{p}}, .$$

Thus, assuming

$$(\varepsilon^*_{\mathrm{d}}/\chi)^{\frac{1}{p}} \le \frac{1}{2} \tag{5.8}$$

the assumptions of Lemma 4.19 (first alternative) are satisfied, the resulting estimates are available, and in particular we have

$$\sup_{\mathrm{B}_{R/2}(x_0)} |\nabla w| \le |\xi_0| + C_2 \Phi(u, \mathrm{B}_R(x_0))^{\frac{1}{2p}} \le (\varepsilon^*_{\mathrm{d}}/\chi)^{\frac{1}{p}} + C_2 (\varepsilon^*_{\mathrm{d}})^{\frac{1}{2p}}$$

with a positive constant C_2 depending only on n and p. Assuming also

$$(\varepsilon^*_{\mathrm{d}}/\chi)^{\frac{1}{p}} + C_2 (\varepsilon^*_{\mathrm{d}})^{\frac{1}{2p}} \le \mu^{1/\kappa_{\mathrm{d}}} \eta_{\mathrm{d}}(\mu) \tag{5.9}$$

with $0 < \mu \le 1$ to be chosen later we may therefore apply Proposition 5.2. We infer

$$\Phi(w, \mathrm{B}_{2\tau R}(x_0)) \le C \tau^{\gamma_p} (1 + \tau^{-(n+\gamma_p)\frac{p}{2}} \mu) \Big[\mathrm{E}^*(w, \mathrm{B}_{R/2}(x_0)) + \mu^{-1} \Psi(w, \mathrm{B}_{R/2}(x_0)) \Big] .$$

By the energy estimate of Lemma 4.19 this reduces to

$$\Phi(w, \mathrm{B}_{2\tau R}(x_0)) \le C \tau^{\gamma_p} (1 + \tau^{-(n+\gamma_p)\frac{p}{2}} \mu) \Big[\mathrm{E}(u, \mathrm{B}_R(x_0)) + \mu^{-1} \Psi(w, \mathrm{B}_{R/2}(x_0)) \Big] ,$$

where via (4.6) and our assumption we control the energy term

$$\begin{aligned} \mathrm{E}(u, \mathrm{B}_R(x_0)) &= \fint_{\mathrm{B}_R(x_0)} \mathrm{A}_p(\mathrm{W}^0_p(\mathrm{D}u)) \\ &\le C \bigg[\mathrm{A}_p(\mathrm{W}^{\xi_0}_p(\xi_0)) + \fint_{\mathrm{B}_R(x_0)} \mathrm{A}_p(\mathrm{W}^{\xi_0}_p(\mathrm{D}u - \xi_0)) \bigg] \\ &\le C \big[|\xi_0|^p + \Phi(u, \mathrm{B}_R(x_0)) \big] \\ &\le C(\chi^{-1}{+}1) [\Phi(u, \mathrm{B}_R(x_0)) + Q] . \end{aligned}$$

Connecting the last two estimates we come out with

$$\begin{aligned} &\Phi(w, \mathrm{B}_{2\tau R}(x_0)) \\ &\quad \le C(\chi^{-1}{+}1) \tau^{\gamma_p} (1 + \tau^{-(n+\gamma_p)\frac{p}{2}} \mu) \Big[\Phi(u, \mathrm{B}_R(x_0)) + Q + \mu^{-1} \Psi(w, \mathrm{B}_{R/2}(x_0)) \Big] . \end{aligned}$$

Now we first fix

$$\mu = \tau^{(n+\gamma_p)\frac{p}{2}} ,$$

and then we choose $\varepsilon^*_{\mathrm{d}}$ such that the smallness assumptions (5.8) and (5.9) are valid and we moreover have

$$\varepsilon^*_{\mathrm{d}} \le \mu .$$

In view of these choices we conclude

$$\Phi(w, \mathrm{B}_{2\tau R}(x_0)) \le C(\chi^{-1}{+}1) \tau^{\gamma_p} \Big[\Phi(u, \mathrm{B}_R(x_0)) + Q + (\varepsilon^*_{\mathrm{d}})^{-1} \Psi(w, \mathrm{B}_{R/2}(x_0)) \Big] .$$

Finally, in the case $p < 2$ all arguments remain valid, if — according to the changes in (4.26) and Proposition 5.2 — we just replace the occurrences of the exponent $\frac{1}{2p}$ with $\frac{1}{4}$ and those of $(n+\gamma_p)\frac{p}{2}$ with $2(n+\gamma_p)$. □

5.1.2. *Non-degenerate case.*

Proposition 5.5. *Fix $p > 2$ and $\kappa_{\mathrm{n}} := p - 2 > 0$. If $w \in \mathrm{W}^{1,\infty}(\mathrm{B}_{R/2}(x_0))^N$ satisfies*

$$\sup_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \xi_0| \le \mu^{1/\kappa_{\mathrm{n}}} \eta_{\mathrm{n}}(\mu)|\xi_0| \tag{5.10}$$

for a given $\mu \in (0,1]$ and some $\xi_0 \in \mathbb{R}^{Nn}$ with $|\xi_0| < \frac{1}{2}\sigma$, then there holds

$$\begin{aligned}
&\Phi(w, \mathrm{B}_{2\tau R}(x_0)) \\
&\quad \le C\tau^2(1 + \tau^{-n-2}\mu)\bigg[|\xi_0|^{p-2} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \xi_0|^2 \,\mathrm{d}x + \mu^{-1}\Psi(w, \mathrm{B}_{R/2}(x_0))\bigg]
\end{aligned}$$

for all $0 < \tau \le \frac{1}{4}$, where C depends only on n, N, p, Γ, and σ.

Proof. We may assume $\xi_0 \neq 0$ since otherwise w is affine and the proposition holds trivially. In view of (H2) the symmetric bilinear form $\mathcal{B} := \nabla^2 f(\xi_0)$ is positive with ellipticity ratio controlled by σ^{-2}. By standard results we can thus find a weak solution $h \in w + \mathrm{W}^{1,2}_0(\mathrm{B}_{R/2}(x_0))^N$ of the linear system of second-order partial differential equations associated with $\mathcal{B}$, that is some function h satisfying (4.16). Consequently, all the estimates of Proposition 4.17 are available for h. Eventually we will also use these estimates with ∇h replaced by $(\nabla h - \xi_0)$; this is allowed, as h minus an affine function solves the same system. From (4.2) and the observation $\sup_{\mathrm{B}_{R/2}(x_0)} |\nabla w| \le |\xi_0| + \eta_{\mathrm{n}}(\mu)|\xi_0| \le 2|\xi_0|$ we now deduce

$$\begin{aligned}
&\Phi(w, \mathrm{B}_{2\tau R}(x_0)) \\
&\quad \le \frac{1}{p}\Phi^*(w, \mathrm{B}_{2\tau R}(x_0)) \\
&\quad \le C\bigg(\sup_{\mathrm{B}_{R/2}(x_0)} |\nabla w|\bigg)^{p-2} \fint_{\mathrm{B}_{2\tau R}(x_0)} |\nabla w - (\nabla w)_{\mathrm{B}_{2\tau R}(x_0)}|^2 \,\mathrm{d}x \\
&\quad \le C|\xi_0|^{p-2}\bigg[\fint_{\mathrm{B}_{2\tau R}(x_0)} |\nabla h - (\nabla h)_{\mathrm{B}_{2\tau R}(x_0)}|^2 \,\mathrm{d}x + \fint_{\mathrm{B}_{2\tau R}(x_0)} |\nabla w - \nabla h|^2 \,\mathrm{d}x\bigg].
\end{aligned} \tag{5.11}$$

Moreover, we have

$$\begin{aligned}
&\fint_{\mathrm{B}_{R/2}(x_0)} |\nabla h - (\nabla h)_{\mathrm{B}_{R/2}(x_0)}|^2 \,\mathrm{d}x \\
&\qquad\qquad \le \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla h - \xi_0|^2 \,\mathrm{d}x \\
&\qquad\qquad \le 2\bigg[\fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \xi_0|^2 \,\mathrm{d}x + \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla h - \nabla w|^2 \,\mathrm{d}x\bigg].
\end{aligned}$$

Combining the last two estimates with (4.17), and exploiting $\tau \leq \frac{1}{4}$ we end up with

$$
\begin{aligned}
\text{(5.12)}\quad \Phi(w, \mathrm{B}_{2\tau R}(x_0)) \leq C\tau^2 \bigg[&|\xi_0|^{p-2} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \xi_0|^2 \,\mathrm{d}x \\
&+ \tau^{-n-2} |\xi_0|^{p-2} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \nabla h|^2 \,\mathrm{d}x \bigg].
\end{aligned}
$$

In the remainder of the proof we will be concerned with an estimate for the last term on the right-hand side. Exploiting (H2), (4.16), and $h - w \in \mathrm{W}_0^{1,2}(\mathrm{B}_{R/2}(x_0))^N$ we get

$$
\begin{aligned}
&|\xi_0|^{p-2} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \nabla h|^2 \,\mathrm{d}x \\
&\leq \frac{1}{\sigma} \fint_{\mathrm{B}_{R/2}(x_0)} \nabla^2 f(\xi_0)(\nabla w - \nabla h, \nabla w - \nabla h) \,\mathrm{d}x \\
&= \frac{2}{\sigma} \bigg[\fint_{\mathrm{B}_{R/2}(x_0)} \tfrac{1}{2} \nabla^2 f(\xi_0)(\nabla w - \xi_0, \nabla w - \xi_0) \,\mathrm{d}x \\
&\qquad\qquad - \fint_{\mathrm{B}_{R/2}(x_0)} \tfrac{1}{2} \nabla^2 f(\xi_0)(\nabla h - \xi_0, \nabla h - \xi_0) \,\mathrm{d}x \bigg] \\
&= \frac{2}{\sigma} \bigg[\fint_{\mathrm{B}_{R/2}(x_0)} \big[\tfrac{1}{2} \nabla^2 f(\xi_0)(\nabla w - \xi_0, \nabla w - \xi_0) + \nabla f(\xi_0)(\nabla w - \xi_0) + f(\xi_0) - f(\nabla w) \big] \,\mathrm{d}x \\
&\qquad + \fint_{\mathrm{B}_{R/2}(x_0)} \big[f(\nabla w) - f(\nabla h) \big] \,\mathrm{d}x \\
&\qquad + \fint_{\mathrm{B}_{R/2}(x_0)} \big[f(\nabla h) - f(\xi_0) - \nabla f(\xi_0)(\nabla h - \xi_0) - \tfrac{1}{2} \nabla^2 f(\xi_0)(\nabla h - \xi_0, \nabla h - \xi_0) \big] \,\mathrm{d}x \bigg] \\
&=: \frac{2}{\sigma} \big[I + II + III \big].
\end{aligned}
$$

Integrating, taking into account $|\nabla w - \xi_0| \leq \eta_{\mathrm{n}}(\mu)|\xi_0|$ and $|\xi_0| < \frac{1}{2}\sigma$, and employing (H4) we control the first term via

$$
\begin{aligned}
I &= \fint_{\mathrm{B}_{R/2}(x_0)} \int_0^1 \int_0^1 \big[\nabla^2 f(\xi_0) - \nabla^2 f(\xi_0 + st(\nabla w - \xi_0)) \big] \,\mathrm{d}s\, t \,\mathrm{d}t\, (\nabla w - \xi_0, \nabla w - \xi_0) \,\mathrm{d}x \\
&\leq C\mu |\xi_0|^{p-2} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \xi_0|^2 \,\mathrm{d}x \,.
\end{aligned}
$$

Moreover, by Definition 5.1 (in particular the choice of $C_{\mathcal{A}}$) and (4.19) we have $h \in \mathcal{A}_w^q$ and

$$
II \leq C\Psi(w, \mathrm{B}_{R/2}(x_0))
$$

for the second term. Writing $[\dots]$ for the integrand in the integral III we next decompose

$$
\begin{aligned}
III &= \frac{1}{\mathscr{L}^n(\mathrm{B}_{R/2})} \int_{\{|\nabla h - \xi_0| < \eta_{\mathrm{n}}(\mu)|\xi_0|\}} [\dots] \,\mathrm{d}x \\
&\qquad + \frac{1}{\mathscr{L}^n(\mathrm{B}_{R/2})} \int_{\{|\nabla h - \xi_0| \geq \eta_{\mathrm{n}}(\mu)|\xi_0|\}} [\dots] \,\mathrm{d}x \\
&=: III_1 + III_2
\end{aligned}
$$

For III_1 we argue as we did before in order to estimate I and we exploit (4.18) (with exponent 2) coming out with

$$III_1 \leq C\mu|\xi_0|^{p-2} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla h - \xi_0|^2 \,\mathrm{d}x \leq C\mu|\xi_0|^{p-2} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \xi_0|^2 \,\mathrm{d}x\,.$$

In order to control III_2 we first use (4.9), (H2), and (4.2) to control the integrand by $\mathrm{W}_p^{\xi_0}(\nabla h - \xi_0)$. Then we take advantage from (4.18) (this time with exponent p) and finally conclude via the assumption $|\nabla w - \xi_0| \leq \mu^{1/\kappa_\mathrm{n}}\eta_\mathrm{n}(\mu)|\xi_0|$. We derive

$$\begin{aligned}
III_2 &\leq \frac{C}{\mathscr{L}(\mathrm{B}_{R/2})} \int_{\{|\nabla h - \xi_0| \geq \eta_\mathrm{n}(\mu)|\xi_0|\}} \mathrm{W}_p^{\xi_0}(\nabla h - \xi_0)\,\mathrm{d}x \\
&\leq C\eta_\mathrm{n}(\mu)^{2-p} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla h - \xi_0|^p \,\mathrm{d}x \\
&\leq C\eta_\mathrm{n}(\mu)^{2-p} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \xi_0|^p \,\mathrm{d}x \\
&\leq C\mu^{\frac{p-2}{\kappa_\mathrm{n}}} |\xi_0|^{p-2} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \xi_0|^2 \,\mathrm{d}x\,.
\end{aligned}$$

Recalling $\kappa_\mathrm{n} = p-2$ and collecting the above estimates we have

$$\begin{aligned}
&|\xi_0|^{p-2} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \nabla h|^2 \,\mathrm{d}x \\
&\qquad\qquad \leq C\mu \left[|\xi_0|^{p-2} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \xi_0|^2 \,\mathrm{d}x + \mu^{-1}\Psi(w, \mathrm{B}_{R/2}(x_0)) \right],
\end{aligned}$$

and plugging this into (5.12) we conclude the proof. □

The case $p \leq 2$ in Proposition 5.5: The statement of the proposition remains true for $p \leq 2$ if we replace κ_n by any positive constant, but in order to preserve the stated dependency of C let us assume in the following that $\kappa_\mathrm{n} > 0$ is fixed depending only on n, N, p, Γ, and σ.

In the proof of the proposition the estimate (5.11) does not follow analogously, but at least the resulting estimate remains true, as we have

$$\Phi^*(w, \mathrm{B}_{2\tau R}(x_0)) \leq \Big(\inf_{\mathrm{B}_{R/2}(x_0)} |\nabla w| \Big)^{p-2} \fint_{\mathrm{B}_{2\tau R}(x_0)} |\nabla w - (\nabla w)_{\mathrm{B}_{2\tau R}(x_0)}|^2 \,\mathrm{d}x$$

and $\inf_{\mathrm{B}_{R/2}(x_0)} |\nabla w| \geq |\xi_0| - \eta_\mathrm{n}(\mu)|\xi_0| \geq \frac{1}{2}|\xi_0|$. Otherwise the proof remains unchanged apart from the following modification of the estimate for III_2. Similar as above we control the integrand first by $\mathrm{W}_p^{\xi_0}(\nabla h - \xi_0) \leq |\xi_0|^{p-2}|\nabla h - \xi_0|^2$. Then, via (4.18) (with exponent $2{+}\kappa_\mathrm{n}$) and the assumption $|\nabla w - \xi_0| \leq \mu^{1/\kappa_\mathrm{n}}\eta_\mathrm{n}(\mu)|\xi_0|$, we infer

$$\begin{aligned}
III_2 &\leq \frac{C}{\mathscr{L}(\mathrm{B}_{R/2})} \int_{\{|\nabla h - \xi_0| \geq \eta_\mathrm{n}(\mu)|\xi_0|\}} |\xi_0|^{p-2}|\nabla h - \xi_0|^2 \,\mathrm{d}x \\
&\leq C\eta_\mathrm{n}(\mu)^{-\kappa_\mathrm{n}} |\xi_0|^{p-2-\kappa_\mathrm{n}} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla h - \xi_0|^{2+\kappa_\mathrm{n}} \,\mathrm{d}x \\
&\leq C\eta_\mathrm{n}(\mu)^{-\kappa_\mathrm{n}} |\xi_0|^{p-2-\kappa_\mathrm{n}} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \xi_0|^{2+\kappa_\mathrm{n}} \,\mathrm{d}x \\
&\leq C\mu|\xi_0|^{p-2} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \xi_0|^2 \,\mathrm{d}x\,.
\end{aligned}$$

□

Remark 5.6. *The previous reasoning for the case $p \le 2$ just requires a $\mathrm{W}^{1,2+\kappa_{\mathrm{n}}}$-estimate for h, possibly with arbitrarily small κ_{n}. The proof of such an estimate is in fact simpler (one may deduce it via Gehring's lemma) than the derivation of (4.18) for the full range of exponents, and thus it would be desirable to get by with small κ_{n} in all cases. However, for $p > 2$ I have not been able to avoid the usage of the slightly harder $\mathrm{W}^{1,p}$-estimate with the given exponent p from the assumptions on f. Furthermore, turning to the estimate (4.19) we notice that it is much simpler in the case $q \le 2^*$, and that it is not required at all for $q \le \frac{n}{n-1}$; compare Remark 5.3.*

Proposition 5.7. *For $u \in \mathrm{BV}(\mathrm{B}_R(x_0))^N$ with $0 < |(\mathrm{D}u)_{\mathrm{B}_R(x_0)}| < \frac{1}{2}\sigma$ we set*

$$w := u_\lambda$$

with λ from the second alternative in Lemma 4.19. Then for every $0 < \tau \le \frac{1}{4}$ there exists some constant $\varepsilon_{\mathrm{n}}^* \in (0,1]$, depending only on n, N, p, Γ, σ, η_{n}, and τ, such that the condition

$$\Phi(u, \mathrm{B}_R(x_0)) \le \varepsilon_{\mathrm{n}}^* |(\mathrm{D}u)_{\mathrm{B}_R(x_0)})|^p$$

implies

$$\Phi(w, \mathrm{B}_{2\tau R}(x_0)) \le C\tau^2 \Big[\Phi(u, \mathrm{B}_R(x_0)) + (\varepsilon_{\mathrm{n}}^*)^{-1} \Psi(w, \mathrm{B}_{R/2}(x_0))\Big] ,$$

where C depends only on n, N, p, Γ, and σ.

Proof. Setting

$$\xi_0 := (\mathrm{D}u)_{\mathrm{B}_R(x_0)}$$

we have by assumption

$$|\xi_0| \le \frac{1}{2}\sigma \le \frac{1}{2} \qquad \text{and} \qquad \Phi(u, \mathrm{B}_R(x_0)) \le \varepsilon_{\mathrm{n}}^* |\xi_0|^p \le |\xi_0|^p .$$

Thus, we may deduce from Lemma 4.19 (second alternative)

$$\sup_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \xi_0| \le C_3 \big(|\xi_0|^{-p} \Phi(u, \mathrm{B}_R(x_0))\big)^{\frac{1}{2\max\{p,2\}}} |\xi_0| \le C_3 (\varepsilon_{\mathrm{n}}^*)^{\frac{1}{2\max\{p,2\}}} |\xi_0|$$

with a positive constant C_3 depending only on n and p. Assuming

$$C_3 (\varepsilon_{\mathrm{n}}^*)^{\frac{1}{2\max\{p,2\}}} \le \mu^{1/\kappa_{\mathrm{n}}} \eta_{\mathrm{n}}(\mu) \tag{5.13}$$

with $0 < \mu \le 1$ to be chosen later we may therefore apply Proposition 5.5. In the resulting estimate we replace the term

$$|\xi_0|^{p-2} \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla w - \xi_0|^2 \,\mathrm{d}x \qquad \text{by} \qquad \fint_{\mathrm{B}_{R/2}(x_0)} \mathrm{W}_p^{\xi_0}(\nabla w - \xi_0) \,\mathrm{d}x$$

(for $p \ge 2$ this is trivially possible, while for $p < 2$ it relies on (5.10) which we have just verified), and thus we infer

$$\begin{aligned} &\Phi(w, \mathrm{B}_{2\tau R}(x_0)) \\ &\qquad \le C\tau^2 (1 + \tau^{-n-2}\mu) \bigg[\fint_{\mathrm{B}_{R/2}(x_0)} \mathrm{W}_p^{\xi_0}(\nabla w - \xi_0) \,\mathrm{d}x + \mu^{-1} \Psi(w, \mathrm{B}_{R/2}(x_0))\bigg] . \end{aligned}$$

By the excess estimate of Lemma 4.19 the preceding inequality reduces to

$$\Phi(w, \mathrm{B}_{2\tau R}(x_0)) \le C\tau^2 (1 + \tau^{-n-2}\mu) \Big[\Phi(u, \mathrm{B}_R(x_0)) + \mu^{-1} \Psi(w, \mathrm{B}_{R/2}(x_0))\Big]\Big] .$$

Now we first fix

$$\mu = \tau^{n+2} ,$$

and then we choose $\varepsilon_{\mathrm{n}}^*$ such that the smallness assumption (5.13) is valid and we moreover have

$$\varepsilon_{\mathrm{n}}^* \leq \mu\,.$$

In view of these choices we finally conclude

$$\Phi(w, \mathrm{B}_{2\tau R}(x_0)) \leq C\tau^2 \Big[\Phi(u, \mathrm{B}_R(x_0)) + (\varepsilon_{\mathrm{n}}^*)^{-1}\Psi(w, \mathrm{B}_{R/2}(x_0))\Big]\,. \qquad \square$$

5.2. **Estimates for almost-minimizers.**

5.2.1. *General minimality estimates.* In this subsection we follow the arguments of [8, Section 5]. However, though the basic ideas remain unchanged, the generalization to L^q-α-minimizers, the different shape of our excess, and the refined smoothing procedure lead to several non-trivial changes. Therefore, we give a rereading of the arguments from [8, Section 5] in our setting. Exactly as in Lemma 4.19 we shall partially provide two alternatives, where the first alternative always corresponds to the first choice of λ in Lemma 4.19, and the second alternative corresponds to the second choice.

Proposition 5.8. *Consider an L^q-α-minimizer $u \in \mathrm{BV}_{\mathrm{loc}}(\Omega)^N$ of F at $x_0 \in \Omega$, and set*

$$w := u_\lambda$$

with the choices of λ from Lemma 4.19, where we assume $(\mathrm{D}u)_{\mathrm{B}_R(x_0)} \neq 0$ in case of the second alternative. Then for every $0 < \tau \leq \frac{1}{4}$ and every $0 < \varepsilon \leq 1$ there exists some $0 < \delta \leq 1$, depending only on n, N, p, Γ, σ, ω, τ, and ε, such that for every ball $\mathrm{B}_R(x_0) \subset \Omega$ the smallness condition

$$\left\{\begin{array}{l} \Phi(u, \mathrm{B}_R(x_0)) \leq \delta \\ \Phi(u, \mathrm{B}_R(x_0)) \leq \delta |(\mathrm{D}u)_{\mathrm{B}_R(x_0)}|^p \end{array}\right\}$$

together with the bound

$$|(\mathrm{D}u)_{\mathrm{B}_R(x_0)}| \leq \frac{1}{2}\sigma$$

implies the estimates

$$\Phi(u, \mathrm{B}_{\tau R}(x_0)) \leq C\Big[\tau^{-n}R^\alpha + \Phi(w, \mathrm{B}_{2\tau R}(x_0))\Big] + \varepsilon\Phi(u, \mathrm{B}_R(x_0))\,,$$
$$\Psi(w, \mathrm{B}_{R/2}(x_0)) \leq CR^\alpha + \varepsilon\Phi(u, \mathrm{B}_R(x_0))\,,$$

where C depends only on n, N, p, Γ, σ, and ω.

Proof. We may assume $\Phi(u, \mathrm{B}_R(x_0)) > 0$, since otherwise u is affine on $\mathrm{B}_R(x_0)$ and equals w on $\mathrm{B}_{R-\lambda}(x_0)$ so that the claims are trivially valid. For radii

$$\tau R < r < 2\tau R \leq R/2 < s < t \leq R - \lambda$$

to be fixed later we set

$$\xi_0 := (\mathrm{D}u)_{\mathrm{B}_R(x_0)} \qquad \text{and} \qquad \xi := (\nabla w)_{\mathrm{B}_r(x_0)}\,,$$

and to avoid to many distinctions of cases we abbreviate

$$\widetilde{\Phi} := \left\{\begin{array}{l} \Phi(u, \mathrm{B}_R(x_0)) \\ |\xi_0|^{-p}\Phi(u, \mathrm{B}_R(x_0)) \end{array}\right\}\,,$$

depending on the alternative we are considering. We note $0<\widetilde{\Phi}\leq\delta$, and we record that by Lemma 4.19 (using also $|\xi_0|\leq\sigma\leq 1$ in case of the second alternative) there holds

$$|\xi-\xi_0|\leq \fint_{\mathrm{B}_r(x_0)}|\nabla w-\xi_0|\,\mathrm{d}x\leq C_4\widetilde{\Phi}^{\frac{1}{2\max\{p,2\}}}$$

with some positive constant C_4 depending only on n and p. Imposing the smallness condition

$$C_4\delta^{\frac{1}{2\max\{p,2\}}}<\frac{1}{2}\sigma \tag{5.14}$$

we consequently have

$$|\xi-\xi_0|<\frac{1}{2}\sigma\qquad\text{and}\qquad|\xi|<\sigma\leq\frac{1}{4}$$

and we may introduce the auxiliary non-negative convex function f_ξ by

$$f_\xi(z):=f(z)-f(\xi)-\nabla f(\xi)(z-\xi)\qquad\text{for }z\in\mathbb{R}^{Nn}\,.$$

Then Lemma 4.14 and (4.9) give

$$\Phi(u,\mathrm{B}_{\tau R}(x_0))\leq C\fint_{\mathrm{B}_{\tau R}(x_0)}\mathrm{A}_p(\mathrm{W}^\xi_p(\mathrm{D}u-\xi\mathscr{L}^n))\leq\frac{C}{(\tau R)^n}\int_{\mathrm{B}_r(x_0)}f_\xi(\mathrm{D}u)\,. \tag{5.15}$$

Exploiting $f_\xi\geq 0$, the fact that u is L^q-α-minimizing for F at x_0, and $\int_{\mathrm{B}_t(x_0)}\mathrm{D}\varphi=0$ we have

$$\begin{aligned}\int_{\mathrm{B}_r(x_0)}&f_\xi(\mathrm{D}u)\\&\leq\int_{\mathrm{B}_t(x_0)}f_\xi(\mathrm{D}u)-\int_{\mathrm{B}_s(x_0)\setminus\mathrm{B}_r(x_0)}f_\xi(\mathrm{D}u)\\&\leq\omega(M)t^\alpha\mathscr{L}^n(\mathrm{B}_t)+\int_{\mathrm{B}_t(x_0)}f_\xi(\mathrm{D}u{+}\mathrm{D}\varphi)-\int_{\mathrm{B}_s(x_0)\setminus\mathrm{B}_r(x_0)}f_\xi(\mathrm{D}u)\end{aligned} \tag{5.16}$$

for all $\varphi\in\mathrm{BV}(\Omega)^N$ with $\operatorname{spt}\varphi\subset\mathrm{B}_t(x_0)$ and all $M\in[0,\infty)$ with (2.6) and (2.7). Postponing the specification of M to the end of this proof we next choose

$$\varphi=\eta(w-u)\,,$$

where η is a smooth and compactly supported function on $\mathrm{B}_t(x_0)$ with

$$\mathbb{1}_{\mathrm{B}_s(x_0)}\leq\eta\leq\mathbb{1}_{\mathrm{B}_{(s+t)/2}(x_0)}\quad\text{and}\quad|\nabla\eta|\leq 3/(t{-}s)\qquad\text{on }\mathrm{B}_t(x_0)\,.$$

Then the right-hand side of (5.16) equals $T+I+II+III$, where we have set

$$\begin{aligned}T&:=\omega(M)t^\alpha\mathscr{L}^n(\mathrm{B}_t)\,,\\I&:=\int_{\mathrm{B}_r(x_0)}f_\xi(\nabla w)\,\mathrm{d}x\,,\\II&:=\int_{\mathrm{B}_s(x_0)\setminus\mathrm{B}_r(x_0)}f_\xi(\nabla w)\,\mathrm{d}x-\int_{\mathrm{B}_s(x_0)\setminus\mathrm{B}_r(x_0)}f_\xi(\mathrm{D}u)\,,\\III&:=\int_{\mathrm{B}_t(x_0)\setminus\mathrm{B}_s(x_0)}f_\xi((1{-}\eta)\mathrm{D}u+\eta\nabla w\mathscr{L}^n+(w{-}u)\otimes\nabla\eta\mathscr{L}^n)\,.\end{aligned}$$

Clearly, we have

$$T\leq CR^{n+\gamma_p}$$

with a constant C depending on n, M, and ω, and by (4.9) and the definition of ξ there holds

$$I\leq Cr^n\Phi(w,\mathrm{B}_r(x_0))\,.$$

we next derive estimates for II and III. Here, by (4.9) — once more — and Lemma 4.4 we may split

$$\begin{aligned} III &\le C\int_{\mathrm{B}_t(x_0)\setminus\mathrm{B}_s(x_0)} \mathrm{A}_p(\mathrm{W}_p^\xi((1-\eta)(\mathrm{D}u-\xi\mathscr{L}^n)+\eta(\nabla w-\xi)\mathscr{L}^n+(w-u)\otimes\nabla\eta\mathscr{L}^n))\\ &\le C\bigg[\int_{\mathrm{B}_t(x_0)\setminus\mathrm{B}_s(x_0)} \mathrm{A}_p(\mathrm{W}_p^\xi(\mathrm{D}u-\xi\mathscr{L}^n)) + \int_{\mathrm{B}_t(x_0)\setminus\mathrm{B}_s(x_0)} \mathrm{A}_p(\mathrm{W}_p^\xi(\nabla w-\xi))\,\mathrm{d}x\\ &\qquad\qquad + \int_{\mathrm{B}_t(x_0)\setminus\mathrm{B}_s(x_0)} \mathrm{A}_p\Big(\mathrm{W}_p^\xi\Big(\frac{u-w}{t-s}\Big)\Big)\,\mathrm{d}x\bigg]\\ &=: C\big[III_1+III_2+III_3\big]\,. \end{aligned}$$

Recalling $w=u_\lambda$ we apply the Jensen inequality of Lemma 4.12 with the convex function of Lemma 4.7. Using also a standard estimate for mollifications we infer

$$III_2 \le \int_{\mathrm{B}_{t+\lambda}(x_0)\setminus\mathrm{B}_{s-\lambda}(x_0)} \mathrm{A}_p(\mathrm{W}_p^\xi(\mathrm{D}u-\xi\mathscr{L}^n))\,.$$

When we assume

$$t-s\ge\lambda$$

(which will be satisfied by the choice of radii to come), we can also use the Poincaré inequality of Lemma 4.22, with $x\mapsto u(x){-}\xi x$ in place of u, to get

$$III_3 \le \int_{\mathrm{B}_{t+\lambda}(x_0)\setminus\mathrm{B}_{s-\lambda}(x_0)} \mathrm{A}_p(\mathrm{W}_p^\xi(\mathrm{D}u-\xi\mathscr{L}^n))\,.$$

Consequently, we may control

$$III \le C\int_{\mathrm{B}_{t+\lambda}(x_0)\setminus\mathrm{B}_{s-\lambda}(x_0)} \mathrm{A}_p(\mathrm{W}_p^\xi(\mathrm{D}u-\xi\mathscr{L}^n))\,.$$

We will now determine good radii r, s, and t. To this end we first choose

$$K := \lfloor \widetilde{\Phi}^{-\frac{1}{4n}}/8\rfloor - 1$$

(here, $\lfloor x\rfloor$ denotes the integer with $x{-}1 < \lfloor x\rfloor \le x$). Imposing the smallness condition

$$\delta \le 16^{-4n} \tag{5.17}$$

we ensure $K\ge 1$ and may thus deduce from Lemma 4.21 that for every $k\in\{1,2,3,\ldots,K\}$ there exist radii r_k and s_k with

$$\tau R < r_k < 2\tau R$$

and

$$R/2+(4k-1)\widetilde{\Phi}^{\frac{1}{4n}}R < s_k < R/2+4k\widetilde{\Phi}^{\frac{1}{4n}}R$$

(note that by choice of K the right-hand bound is at most $R-4\widetilde{\Phi}^{\frac{1}{4n}}R$ and is thus smaller than $R-\lambda = R-\frac{1}{2}\widetilde{\Phi}^{\frac{1}{2n}}R$) such that we may estimate

$$\begin{aligned} \int_{\mathrm{B}_{s_k}(x_0)\setminus\mathrm{B}_{r_k}(x_0)} f_\xi(\nabla u_\lambda)\,\mathrm{d}x - \int_{\mathrm{B}_{s_k}(x_0)\setminus\mathrm{B}_{r_k}(x_0)} f_\xi(\mathrm{D}u)&\\ \le \bigg(\frac{2\lambda}{\tau R}+\frac{2\lambda}{\widetilde{\Phi}^{\frac{1}{4n}}R}\bigg)\int_{\mathrm{B}_R(x_0)} f_\xi(\mathrm{D}u)\,.& \end{aligned} \tag{5.18}$$

Via Lemma 4.20 we choose further radii t_k with

$$s_k+\widetilde{\Phi}^{\frac{1}{4n}}R < t_k < s_k+2\widetilde{\Phi}^{\frac{1}{4n}}R$$

(again that the right-hand bound is always smaller than $R-\lambda$) such that the estimate

$$\int_{\mathrm{B}_{t_k}(x_0)} f_\xi(\nabla u_\lambda)\,\mathrm{d}x - \int_{\mathrm{B}_{t_k}(x_0)} f_\xi(\mathrm{D}u) \le \frac{2\lambda}{\widetilde{\Phi}^{\frac{1}{4n}}R}\int_{\mathrm{B}_R(x_0)} f_\xi(\mathrm{D}u) \tag{5.19}$$

is valid. Note that (5.19) will only be needed at another point of the proof, while for our momentary purposes it would also suffice to define t_k simply as $s_k+\widetilde{\Phi}^{\frac{1}{4n}}R$. Anyway, recalling $2\lambda=\widetilde{\Phi}^{\frac{1}{2n}}R\le\widetilde{\Phi}^{\frac{1}{4n}}R$ we observe that the sets

$$\mathrm{B}_{t_k+\lambda}(x_0)\setminus\mathrm{B}_{s_k-\lambda}(x_0)\qquad\text{with } k\in\{1,2,3,\dots,K\}$$

are pairwise disjoint in $\mathrm{B}_R(x_0)$. Thus, there is some $k_0\in\{1,2,3,\dots,K\}$ such that there holds

$$\int_{\mathrm{B}_{t_{k_0}+\lambda}(x_0)\setminus\mathrm{B}_{s_{k_0}-\lambda}(x_0)} \mathrm{A}_p(\mathrm{W}_p^\xi(\mathrm{D}u-\xi\mathscr{L}^n)) \le \frac{1}{K}\int_{\mathrm{B}_R(x_0)} \mathrm{A}_p(\mathrm{W}_p^\xi(\mathrm{D}u-\xi\mathscr{L}^n))\,.$$

Fixing $s:=s_{k_0}$ and $t:=t_{k_0}$ we have $t-s\ge\lambda$ as required before and for later use we record even $t-s\ge 2\lambda$. In view of $K\ge\frac{1}{3}(K+2)\ge\widetilde{\Phi}^{-\frac{1}{4n}}/24$ this leaves us with the estimate

$$III\le\frac{C}{K}\int_{\mathrm{B}_R(x_0)} \mathrm{A}_p(\mathrm{W}_p^\xi(\mathrm{D}u-\xi\mathscr{L}^n)) \le C\widetilde{\Phi}^{\frac{1}{4n}}\int_{\mathrm{B}_R(x_0)} \mathrm{A}_p(\mathrm{W}_p^\xi(\mathrm{D}u-\xi\mathscr{L}^n))\,.$$

Fixing also $r:=r_{k_0}$ and inserting $\lambda=\frac{1}{2}\widetilde{\Phi}^{\frac{1}{2n}}R$ into (5.18) we moreover have

$$\begin{aligned} II &\le (\tau^{-1}\widetilde{\Phi}^{\frac{1}{4n}}+1)\widetilde{\Phi}^{\frac{1}{4n}}\int_{\mathrm{B}_R(x_0)} f_\xi(\mathrm{D}u)\\ &\le C(\tau^{-1}\widetilde{\Phi}^{\frac{1}{4n}}+1)\widetilde{\Phi}^{\frac{1}{4n}}\int_{\mathrm{B}_R(x_0)} \mathrm{A}_p(\mathrm{W}_p^\xi(\mathrm{D}u-\xi\mathscr{L}^n))\,, \end{aligned}$$

where we used also (4.9). Combining the final estimates for T, I, II, and III with (5.15) and (5.16) we arrive at

$$\begin{aligned} \Phi(u,\mathrm{B}_{\tau R}(x_0)) \le C(\tau R)^{-n}\Big[&R^{n+\alpha}+r^n\Phi(w,\mathrm{B}_r(x_0))\\ &+(\tau^{-1}\widetilde{\Phi}^{\frac{1}{4n}}+1)\widetilde{\Phi}^{\frac{1}{4n}}\int_{\mathrm{B}_R(x_0)} \mathrm{A}_p(\mathrm{W}_p^\xi(\mathrm{D}u-\xi\mathscr{L}^n))\Big] \end{aligned} \tag{5.20}$$

For the second term on the right-hand side of (5.20) Lemma 4.14 yields

$$\begin{aligned} r^n\Phi(w,\mathrm{B}_r(x_0)) &\le C\int_{\mathrm{B}_r(x_0)} \mathrm{A}_p(\mathrm{W}_p^{(\nabla w)_{\mathrm{B}_{2\tau R}(x_0)}}(\nabla w-(\nabla w)_{\mathrm{B}_{2\tau R}(x_0)}))\,\mathrm{d}x\\ &\le C(2\tau R)^n\Phi(w,\mathrm{B}_{2\tau R}(x_0))\,, \end{aligned} \tag{5.21}$$

and our next aim is to control the last term on the right-hand side of (5.20) in terms of the excess $\Phi(u,\mathrm{B}_R(x_0))$ (which actually means to pass from ξ to ξ_0 in this term). To this end we estimate via (4.9)

$$\begin{aligned}
&\fint_{\mathrm{B}_R(x_0)} \mathrm{A}_p(\mathrm{W}_p^\xi(\mathrm{D}u-\xi\mathscr{L}^n))\\
&\qquad\le C\fint_{\mathrm{B}_R(x_0)} \Big(f(\mathrm{D}u)-f(\xi)\mathscr{L}^n-\nabla f(\xi)(\mathrm{D}u-\xi\mathscr{L}^n)\Big)\\
&\qquad= C\bigg[\fint_{\mathrm{B}_R(x_0)} \Big(f(\mathrm{D}u)-f(\xi_0)\mathscr{L}^n-\nabla f(\xi_0)(\mathrm{D}u-\xi_0\mathscr{L}^n)\Big)\\
&\qquad\qquad+\Big(f(\xi_0)+\nabla f(\xi_0)(\xi-\xi_0)-f(\xi)\Big)\\
&\qquad\qquad+\big(\nabla f(\xi_0)-\nabla f(\xi)\big)\fint_{\mathrm{B}_R(x_0)}(\mathrm{D}u-\xi\mathscr{L}^n)\bigg]\,.
\end{aligned}$$

For two terms on the right-hand side we easily derive

$$\fint_{\mathrm{B}_R(x_0)} \Big(f(\mathrm{D}u)-f(\xi_0)\mathscr{L}^n-\nabla f(\xi_0)(\mathrm{D}u-\xi_0\mathscr{L}^n)\Big)\le C\Phi(u,\mathrm{B}_R(x_0))$$

by (4.9) and the definition of ξ_0 and

$$f(\xi_0)+\nabla f(\xi_0)(\xi-\xi_0)-f(\xi)\le 0$$

by the convexity of f. To estimate the remainder term we use in turn the bounds $|\xi_0|\le\frac12\sigma\le\frac18$ and $|\xi|\le\sigma\le\frac14$, (H2), Lemma 4.2, the definition of ξ, the Jensen inequality from Lemma 4.12, and the definition of ξ_0. In this way we infer

$$\begin{aligned}
&\big(\nabla f(\xi_0)-\nabla f(\xi)\big)\fint_{\mathrm{B}_R(x_0)}(\mathrm{D}u-\xi\mathscr{L}^n)\\
&\qquad= \int_0^1 \nabla^2 f(\xi+s(\xi_0-\xi))\,\mathrm{d}s\,(\xi_0-\xi,\xi_0-\xi)\\
&\qquad\le C\int_0^1 |\xi+s(\xi_0-\xi)|^{p-2}\,\mathrm{d}s\,|\xi-\xi_0|^2\\
&\qquad\le C\mathrm{W}_p^{\xi_0}(\xi-\xi_0)\\
&\qquad\le C\mathrm{A}_p(\mathrm{W}_p^{\xi_0}(\xi-\xi_0))\\
&\qquad= C\mathrm{A}_p\left(\mathrm{W}_p^{\xi_0}\left(\fint_{\mathrm{B}_r(x_0)}(\mathrm{D}u-\xi_0\mathscr{L}^n)\right)\right)\\
&\qquad\le C\fint_{\mathrm{B}_r(x_0)}\mathrm{A}_p(\mathrm{W}_p^{\xi_0}(\mathrm{D}u-\xi_0\mathscr{L}^n))\\
&\qquad\le C\Big(\frac{R}{r}\Big)^n\Phi(u,\mathrm{B}_R(x_0))
\end{aligned}$$

and recalling $\tau R<r<R$ we have established

$$\fint_{\mathrm{B}_R(x_0)} \mathrm{A}_p(\mathrm{W}_p^\xi(\mathrm{D}u-\xi\mathscr{L}^n))\le C\tau^{-n}\Phi(u,\mathrm{B}_R(x_0))\,. \tag{5.22}$$

Plugging (5.21) and (5.22) into (5.20) we conclude

$$\begin{aligned}
&\Phi(u,\mathrm{B}_{\tau R}(x_0))\\
&\qquad\le C_5\Big[\tau^{-n}R^\alpha+\Phi(w,\mathrm{B}_{2\tau R}(x_0))+(\tau^{-1}\widetilde{\Phi}^{\frac{1}{4n}}+1)\tau^{-2n}\widetilde{\Phi}^{\frac{1}{4n}}\Phi(u,\mathrm{B}_R(x_0))\Big]
\end{aligned}$$

with a positive constant C_5 depending only on n, p, Γ, σ, M, and ω. Imposing still another smallness condition, namely

$$C_5(\tau^{-1}\delta^{\frac{1}{4n}}+1)\tau^{-2n}\delta^{\frac{1}{4n}}\le\varepsilon \tag{5.23}$$

this estimate indeed reduces to

$$\Phi(u,\mathrm{B}_{\tau R}(x_0))\le C\Big[\tau^{-n}R^\alpha+\Phi(w,\mathrm{B}_{2\tau R}(x_0))\Big]+\varepsilon\Phi(u,\mathrm{B}_R(x_0))$$

and we arrive at the first claim of the proposition.

To establish the second claim we consider an arbitrary number $0<\kappa\le 1$. Then by the definition of Ψ there is a function $\widetilde{w}\in\mathcal{A}^q_w$ such that we have

$$(R/2)^n\Psi(w,\mathrm{B}_{R/2}(x_0))\le(R/2)^n\kappa+F[w,\mathrm{B}_{R/2}(x_0)]-F[\widetilde{w},\mathrm{B}_{R/2}(x_0)]\,. \tag{5.24}$$

Using the notation introduced above and in particular the radii $R/2<s<t<R-\lambda$ we rewrite this inequality as

$$(R/2)^n\Psi(w,\mathrm{B}_{R/2}(x_0))\le(R/2)^n\kappa+\int_{\mathrm{B}_t(x_0)}f_\xi(\nabla w)\,\mathrm{d}x-\int_{\mathrm{B}_t(x_0)}f_\xi(\nabla\widetilde{w})\,\mathrm{d}x\,,$$

where we understand $\widetilde{w}=w$ outside $\mathrm{B}_{R/2}(x_0)$. Coming back to the cut-off η between $\mathrm{B}_s(x_0)$ and $\mathrm{B}_t(x_0)$, setting

$$\widetilde{\varphi}:=\eta(\widetilde{w}-u)$$

and exploiting the L^q-α-minimality of u once more, we arrive at

$$\begin{aligned}
&(R/2)^n\Psi(w,\mathrm{B}_{R/2}(x_0))\\
&\qquad\le(R/2)^n\kappa+\omega(M)t^\alpha\mathscr{L}^n(\mathrm{B}_t)+\int_{\mathrm{B}_t(x_0)}f_\xi(\nabla w)\,\mathrm{d}x-\int_{\mathrm{B}_t(x_0)}f_\xi(\mathrm{D}u)\\
&\qquad\qquad+\int_{\mathrm{B}_t(x_0)}f_\xi(\mathrm{D}u+\mathrm{D}\widetilde{\varphi})-\int_{\mathrm{B}_t(x_0)}f_\xi(\nabla\widetilde{w})\,\mathrm{d}x\\
&\qquad\le(R/2)^n\kappa+T+\widetilde{II}+III
\end{aligned}$$

for all M satisfying (2.6) and (2.7) with $\widetilde{\varphi}$ in place of φ. Here we have set

$$\widetilde{II}:=\int_{\mathrm{B}_t(x_0)}f_\xi(\nabla w)\,\mathrm{d}x-\int_{\mathrm{B}_t(x_0)}f_\xi(\mathrm{D}u)\,,$$

while $T=\omega(M)t^\alpha\mathscr{L}^n(\mathrm{B}_t)$ and

$$III=\int_{\mathrm{B}_t(x_0)\setminus\mathrm{B}_s(x_0)}f_\xi((1-\eta)\mathrm{D}u+\eta\nabla w\mathscr{L}^n+(w-u)\otimes\nabla\eta\mathscr{L}^n)$$

denote precisely the same terms as before. We now essentially follow the arguments used for the first claim. Bounding $\widetilde{II}$ via (5.19) (recall $t=t_{k_0}$) in place of (5.18) we deduce

$$(R/2)^n\Psi(w,\mathrm{B}_{R/2}(x_0))\le(R/2)^n\kappa+C\Big[R^{n+\alpha}+\widetilde{\Phi}^{\frac{1}{4n}}\int_{\mathrm{B}_R(x_0)}\mathrm{A}_p(\mathrm{W}^\xi_p(\mathrm{D}u-\xi\mathscr{L}^n))\Big]\,.$$

Combined with (5.22) this implies

$$\Psi(w,\mathrm{B}_{R/2}(x_0))\le\kappa+C_6\Big[R^\alpha+\tau^{-n}\widetilde{\Phi}^{\frac{1}{4n}}\Phi(u,\mathrm{B}_R(x_0))\Big]$$

with some positive constant C_6 depending only on n, p, Γ, σ, M, and ω. Sending $\kappa\to 0$ and assuming

$$C_6\tau^{-n}\delta^{\frac{1}{4n}}\le\varepsilon \tag{5.25}$$

the second claim of the proposition is verified.

To end the proof it remains to implement the choices of M and δ postponed before. To this aim we first collect a couple of related bounds. Indeed, recalling $R/2<t<R$ we first notice

$$\fint_{\mathrm{B}_t(x_0)} |\mathrm{D}u| \le 2^n |(\mathrm{D}u)_{\mathrm{B}_R(x_0)}| \le 2^{n-1}\sigma \,. \tag{5.26}$$

Now we essentially repeat the arguments which we used for the term III, but simply with the modulus as the convex function. Keeping in mind $\varphi=\eta(w{-}u)$ with $0\le\eta\le \mathbb{1}_{\mathrm{B}_{(s+t)/2}(x_0)}$ and $t-s\ge 2\lambda$ we thus obtain

$$\begin{aligned}
&\fint_{\mathrm{B}_t(x_0)} |\mathrm{D}\varphi| \\
&\le \frac{1}{\mathscr{L}^n(\mathrm{B}_t)}\bigg[\int_{\mathrm{B}_t(x_0)} |\mathrm{D}u| + \int_{\mathrm{B}_{t-\lambda}(x_0)} |\nabla w|\,\mathrm{d}x + 5\int_{\mathrm{B}_{t-\lambda}(x_0)\setminus \mathrm{B}_s(x_0)} \Big|\frac{u{-}w}{t-\lambda-s}\Big|\,\mathrm{d}x\bigg] \\
&\le 5\fint_{\mathrm{B}_t(x_0)} |\mathrm{D}u| \\
&\le 2^{n-1}5\sigma \,.
\end{aligned} \tag{5.27}$$

Furthermore, recalling $w=u_\lambda$ we have

$$\begin{aligned}
\|\varphi\|_{\mathrm{L}^q(\mathrm{B}_t(x_0))^N} &\le \|w-u_{\mathrm{B}_t(x_0)}\|_{\mathrm{L}^q(\mathrm{B}_{t-\lambda}(x_0))^N} + \|u-u_{\mathrm{B}_t(x_0)}\|_{\mathrm{L}^q(\mathrm{B}_t(x_0))^N} \\
&\le 2\|u-u_{\mathrm{B}_t(x_0)}\|_{\mathrm{L}^q(\mathrm{B}_t(x_0))^N} \,.
\end{aligned} \tag{5.28}$$

We also provide similar estimates for $\widetilde{\varphi}$ instead of φ. Using $|z|\le 1+\mathrm{A}_p(|z|^p)$, (4.9), (5.1), (5.24), $\kappa\le 1$ and the bound $\sup_{\mathrm{B}_{R/2}(x_0)}|\nabla w|\le C$ (which results from Lemma 4.19 and the assumptions of the proposition) we first get

$$\begin{aligned}
\fint_{\mathrm{B}_{R/2}(x_0)} |\nabla\widetilde{w}|\,\mathrm{d}x &\le 1+\fint_{\mathrm{B}_{R/2}(x_0)} \mathrm{A}_p(|\nabla\widetilde{w}|^p)\,\mathrm{d}x \\
&\le 1+C\fint_{\mathrm{B}_{R/2}(x_0)} f(\nabla\widetilde{w})\,\mathrm{d}x \\
&\le 1+C\bigg(\kappa+\fint_{\mathrm{B}_{R/2}(x_0)} f(\nabla w)\,\mathrm{d}x\bigg) \le C\,.
\end{aligned}$$

As $\widetilde{\varphi}=\widetilde{w}{-}u$ holds on $\mathrm{B}_{R/2}(x_0)$ and $\widetilde{\varphi}$ equals φ outside $\mathrm{B}_{R/2}(x_0)$, in combination with (5.26) and (5.27) we infer

$$\fint_{\mathrm{B}_t(x_0)} |\mathrm{D}\widetilde{\varphi}| \le \fint_{\mathrm{B}_{R/2}(x_0)} |\nabla\widetilde{w}|\,\mathrm{d}x + \fint_{\mathrm{B}_t(x_0)} |\mathrm{D}u| + \fint_{\mathrm{B}_t(x_0)} |\mathrm{D}\varphi| \le C_7\,, \tag{5.29}$$

where the positive constant C_7 depends only on n, p, Γ, and σ. Finally, we exploit $\widetilde{w}\in\mathcal{A}^q_w$, the choice of the admissible class $\mathcal{A}^q_w$ in Definition 5.1, and Sobolev-Poincaré inequalities (remember that $\widetilde{w}{-}w$ has zero boundary values on $\mathrm{B}_{R/2}(x_0)$).

Relying also the above observations we then get

$$
\begin{aligned}
(5.30)\quad & \|\widetilde{\varphi}\|_{\mathrm{L}^q(\mathrm{B}_t(x_0))^N} \\
&\le \|\widetilde{w}-w\|_{\mathrm{L}^q(\mathrm{B}_{R/2}(x_0))^N} + \|\varphi\|_{\mathrm{L}^q(\mathrm{B}_t(x_0))^N} \\
&\le \|\widetilde{w}-\widetilde{w}_{\mathrm{B}_{R/2}(x_0)}\|_{\mathrm{L}^q(\mathrm{B}_{R/2}(x_0))^N} + (R/2)^{\frac{n}{q}}|(\widetilde{w}-w)_{\mathrm{B}_{R/2}(x_0)}| \\
&\qquad + \|w-w_{\mathrm{B}_{R/2}(x_0)}\|_{\mathrm{L}^q(\mathrm{B}_{R/2}(x_0))^N} + \|\varphi\|_{\mathrm{L}^q(\mathrm{B}_t(x_0))^N} \\
&\le C(R/2)^{1+\frac{n}{q}}\left[-\!\!\!\!\!\!\int_{\mathrm{B}_{R/2}(x_0)} |\nabla\widetilde{w}|\,\mathrm{d}x + \sup_{\mathrm{B}_{R/2}(x_0)} |\nabla w| \right] + \|\varphi\|_{\mathrm{L}^q(\mathrm{B}_t(x_0))^N} \\
&\le C_8\Big[t^{1+\frac{n}{q}} + \|u-u_{\mathrm{B}_t(x_0)}\|_{\mathrm{L}^q(\mathrm{B}_t(x_0))^N} \Big]
\end{aligned}
$$

with a positive constant C_8 depending only on n, N, p, Γ, and σ. In summary, from (5.28), (5.29), and (5.30) we infer that all the bounds in (2.6) and (2.7), which we used in testing the L^q-α-minimizing property of u against $u+\varphi$ and $u+\widetilde{\varphi}$, are indeed valid on $\mathrm{B}_t(x_0)$ if we set $M := \max\{2, C_7, C_8\}$. At this stage we finally fix $0<\delta\le 1$ small enough that all the imposed conditions (5.14), (5.17), (5.23), and (5.25) are satisfied. □

5.2.2. *Degenerate case.* We combine Proposition 5.4 and Proposition 5.8.

Corollary 5.9. *Suppose that $u\in \mathrm{BV}_{\mathrm{loc}}(\Omega)^N$ is an L^q-α-minimizer of F at $x_0\in\Omega$. Then for every $0<\beta<\gamma_p$ and every $\chi>0$ there exist numbers $0<\tau_{\mathrm{d}}\le \frac14$ and $0<\varepsilon_{\mathrm{d}}\le 1$ such that for every ball $\mathrm{B}_R(x_0)\subset\Omega$ the conditions*

$$
\chi|(\mathrm{D}u)_{\mathrm{B}_R(x_0)}|^p \le \Phi(u,\mathrm{B}_R(x_0)) + R^\alpha \le \varepsilon_{\mathrm{d}}
$$

imply the estimate

$$
\Phi(u,\mathrm{B}_{\tau_{\mathrm{d}}R}(x_0)) \le \tau_{\mathrm{d}}^\beta \Phi(u,\mathrm{B}_R(x_0)) + \varepsilon_{\mathrm{d}}^{-1}R^\alpha .
$$

Here, the constant $\gamma_p(n,N,p)$ has been fixed in Theorem 4.18, and the dependencies are given by $\tau_{\mathrm{d}}(n,N,p,\Gamma,\sigma,\omega,\beta,\chi)$ and $\varepsilon_{\mathrm{d}}(n,N,p,\Gamma,\sigma,\eta_{\mathrm{d}},\omega,\beta,\chi)$.

Proof. We consider the number $\varepsilon_{\mathrm{d}}^*$ from Proposition 5.4 corresponding to the given χ and the choice $\tau=\tau_{\mathrm{d}}$ with $0<\tau_{\mathrm{d}}\le\frac14$ to be fixed later on. Moreover, we take the number δ from Proposition 5.8 corresponding once more to $\tau=\tau_{\mathrm{d}}$ and to the choice

$$
\varepsilon = \min\left\{\varepsilon_{\mathrm{d}}^*, \frac12\tau_{\mathrm{d}}^\beta\right\},
$$

and we set

$$
\varepsilon_{\mathrm{d}} = \min\{\delta, \varepsilon_{\mathrm{d}}^*, \chi(\sigma/2)^p, \tau_{\mathrm{d}}^{n+\gamma_p}\}\,.
$$

In view of this choice our assumption guarantees both

$$
\Phi(u,\mathrm{B}_R(x_0))\le\delta \qquad\text{and}\qquad |(\mathrm{D}u)_{\mathrm{B}_R(x_0)}|\le \frac12\sigma\,.
$$

Hence we may apply Proposition 5.8, and we come out with the estimates

$$
\begin{aligned}
(5.31)\qquad \Phi(u,\mathrm{B}_{\tau_{\mathrm{d}}R}(x_0)) &\le C\Big[\tau_{\mathrm{d}}^{-n}R^\alpha + \Phi(w,\mathrm{B}_{2\tau_{\mathrm{d}}R}(x_0))\Big] + \frac12\tau_{\mathrm{d}}^\beta\Phi(u,\mathrm{B}_R(x_0))\,,\\
\Psi(w,\mathrm{B}_{R/2}(x_0)) &\le CR^\alpha + \varepsilon_{\mathrm{d}}^*\Phi(u,\mathrm{B}_R(x_0))\,,
\end{aligned}
$$

where $w=u_\lambda$ is the mollification corresponding to the first alternative choice of λ in Lemma 4.19. By our assumption we are moreover in the position to apply

Proposition 5.4 with $Q := R^\alpha$ which we combine with the last inequality arriving at

$$\Phi(w, \mathrm{B}_{2\tau_{\mathrm{d}}R}(x_0)) \le C\tau_{\mathrm{d}}^{\gamma_p}\Big[\Phi(u, \mathrm{B}_R(x_0)) + (\varepsilon_{\mathrm{d}}^*)^{-1}R^\alpha\Big]\,. \tag{5.32}$$

Inserting (5.32) into (5.31) and exploiting the choice of ε_{d} we find the inequality

$$\Phi(u, \mathrm{B}_{\tau_{\mathrm{d}}R}(x_0)) \le \Big(C_9\tau_{\mathrm{d}}^{\gamma_p} + \frac{1}{2}\tau_{\mathrm{d}}^\beta\Big)\Big[\Phi(u, \mathrm{B}_R(x_0)) + \varepsilon_{\mathrm{d}}^{-1}R^\alpha\Big]$$

with a positive constant C_9 depending only on n, N, p, Γ, σ, ω, and χ. Finally, when we fix $0 < \tau_{\mathrm{d}} \le \frac{1}{4}$ small enough that there holds

$$C_9\tau_{\mathrm{d}}^{\gamma_p} \le \frac{1}{2}\tau_{\mathrm{d}}^\beta\,,$$

we arrive at the claim. □

5.2.3. *Non-degenerate case.* We combine Proposition 5.7 and Proposition 5.8.

Corollary 5.10. *Consider an* L^q*-*α*-minimizer* $u \in \mathrm{BV}_{\mathrm{loc}}(\Omega)^N$ *of* F *at* $x_0 \in \Omega$. *Then for every* $0 < \beta < 2$ *there exist numbers* $0 < \tau_{\mathrm{n}} \le \frac{1}{4}$ *and* $0 < \varepsilon_{\mathrm{n}} \le 1$ *such that for every ball* $\mathrm{B}_R(x_0) \subset \Omega$ *the condition*

$$\Phi(u, \mathrm{B}_R(x_0)) \le \varepsilon_{\mathrm{n}}|(\mathrm{D}u)_{\mathrm{B}_R(x_0)}|^p$$

together with $|(\mathrm{D}u)_{\mathrm{B}_R(x_0)}| < \frac{1}{2}\sigma$ *implies the estimate*

$$\Phi(u, \mathrm{B}_{\tau_{\mathrm{n}}R}(x_0)) \le \tau_{\mathrm{n}}^\beta\Phi(u, \mathrm{B}_R(x_0)) + \varepsilon_{\mathrm{n}}^{-1}R^\alpha.$$

Here, the dependencies are given by $\tau_{\mathrm{n}}(n, N, p, \Gamma, \sigma, \omega, \beta)$ *and* $\varepsilon_{\mathrm{n}}(n, N, p, \Gamma, \sigma, \eta_{\mathrm{n}}, \omega, \beta)$.

Proof. We assume $(\mathrm{D}u)_{\mathrm{B}_R(x_0)} \ne 0$ since otherwise also $\Phi(u, \mathrm{B}_R(x_0))$ vanishes, u is affine on $\mathrm{B}_R(x_0)$ and the claim is clearly valid. We consider the number $\varepsilon_{\mathrm{n}}^*$ from Proposition 5.7 corresponding to the choice $\tau = \tau_{\mathrm{n}}$ with $0 < \tau_{\mathrm{n}} \le \frac{1}{4}$ to be fixed later on. Moreover, we take the number δ from Proposition 5.8 corresponding once more to $\tau = \tau_{\mathrm{d}}$ and to the choice

$$\varepsilon = \min\Big\{\varepsilon_{\mathrm{n}}^*, \frac{1}{2}\tau_{\mathrm{n}}^\beta\Big\}\,,$$

and we set

$$\varepsilon_{\mathrm{n}} = \min\{\delta, \varepsilon_{\mathrm{n}}^*, \tau_{\mathrm{n}}^{n+2}\}\,.$$

In view of this choice our assumption guarantees

$$\Phi(u, \mathrm{B}_R(x_0)) \le \delta|(\mathrm{D}u)_{\mathrm{B}_R(x_0)}|^p$$

so that we may apply Proposition 5.8. We come out with the estimates

$$\begin{aligned} \Phi(u, \mathrm{B}_{\tau_{\mathrm{n}}R}(x_0)) &\le C\Big[\tau_{\mathrm{n}}^{-n}R^\alpha + \Phi(w, \mathrm{B}_{2\tau_{\mathrm{n}}R}(x_0)\Big] + \frac{1}{2}\tau_{\mathrm{n}}^\beta\Phi(u, \mathrm{B}_R(x_0))\,, \\ \Psi(w, \mathrm{B}_{R/2}(x_0)) &\le CR^\alpha + \varepsilon_{\mathrm{n}}^*\Phi(u, \mathrm{B}_R(x_0))\,, \end{aligned} \tag{5.33}$$

where $w = u_\lambda$ is the mollification corresponding to the second alternative choice of λ in Lemma 4.19. By our assumption we are moreover in the position to apply Proposition 5.7 which we combine with the last inequality arriving at

$$\Phi(w, \mathrm{B}_{2\tau_{\mathrm{n}}R}(x_0)) \le C\tau_{\mathrm{n}}^2\Big[\Phi(u, \mathrm{B}_R(x_0)) + (\varepsilon_{\mathrm{n}}^*)^{-1}R^\alpha\Big]\,. \tag{5.34}$$

We plug (5.34) into (5.33) and exploit the choice of ε_{n} to get

$$\Phi(u, \mathrm{B}_{\tau_{\mathrm{n}}R}(x_0)) \le \Big(C_{10}\tau_{\mathrm{n}}^2 + \frac{1}{2}\tau_{\mathrm{n}}^\beta\Big)\Big[\Phi(u, \mathrm{B}_R(x_0)) + \varepsilon_{\mathrm{n}}^{-1}R^\alpha\Big]$$

with a positive constant C_{10} depending only on n, N, p, Γ, σ, and ω. Finally, when we fix $0 < \tau_{\mathrm{n}} \leq \frac{1}{4}$ small enough that there holds

$$C_{10}\tau_{\mathrm{n}}^2 \leq \frac{1}{2}\tau_{\mathrm{n}}^{\beta}\,,$$

we arrive at the claim. □

5.3. **Iteration.** We now iterate Corollaries 5.9 and 5.10 in order to obtain a relation between the excess on two concentric balls with an arbitrary ratio of the radii.

5.3.1. *Non-degenerate case.* We start with the simpler consideration for the non-degenerate case.

Proposition 5.11. *Suppose that $u \in \mathrm{BV}(\Omega)^N$ is an L^q-α-minimizer of F at $x_0 \in \Omega$ with $\alpha \in (0,2)$. Then there exists a number $0 < \widetilde{\varepsilon}_{\mathrm{n}} \leq 1$, depending only on n, N, p, Γ, σ, η_{n}, ω, and α such that for every ball $\mathrm{B}_R(x_0) \subset \Omega$ the condition*

$$\Phi(u, \mathrm{B}_R(x_0)) + R^{\alpha} \leq \widetilde{\varepsilon}_{\mathrm{n}} |(\mathrm{D}u)_{\mathrm{B}_R(x_0)}|^p$$

together with $|(\mathrm{D}u)_{\mathrm{B}_R(x_0)}| < \frac{1}{3}\sigma$ implies the decay estimate

$$\Phi(u, \mathrm{B}_{\varrho}(x_0)) \leq C\left[\Big(\frac{\varrho}{R}\Big)^{\frac{\alpha+2}{2}}\Phi(u, \mathrm{B}_R(x_0)) + \varrho^{\alpha}\right] \qquad \textit{for all } 0 < \varrho \leq R\,.$$

Here, C depends only on n, N, p, Γ, σ, η_{n}, ω, and α.

Proof. We set

$$\beta := \frac{\alpha+2}{2} \in (\alpha, 2)$$

and choose

$$\sqrt{\widetilde{\varepsilon}_{\mathrm{n}}} := \min\{\varepsilon_{\mathrm{n}}(\tau_{\mathrm{n}}^{\alpha} - \tau_{\mathrm{n}}^{\beta}), 2^{-p}\varepsilon_{\mathrm{n}}, (100C_1)^{-1}2^{-p}\tau_{\mathrm{n}}^{np}(1-\tau_{\mathrm{n}}^{\alpha/2})^2\}\,,$$

with the constant C_1 in (4.4) and the numbers $0 < \tau_{\mathrm{n}} \leq \frac{1}{4}$ and $0 < \varepsilon_{\mathrm{n}} \leq 1$ from Corollary 5.10. We now prove by induction that there holds

$$\Phi(u, \mathrm{B}_{\tau_{\mathrm{n}}^k R}(x_0)) \leq \tau_{\mathrm{n}}^{k\beta}\Phi(u, \mathrm{B}_R(x_0)) + \frac{(\tau_{\mathrm{n}}^k R)^{\alpha}}{\varepsilon_{\mathrm{n}}\tau_{\mathrm{n}}^{\alpha}}\sum_{l=0}^{k-1}\tau_{\mathrm{n}}^{l(\beta-\alpha)} \tag{5.35}$$

for all $k \in \mathbb{N}_0$. Obviously, (5.35) is true for $k = 0$ and assuming it to be true for all $k \in \{0, 1, 2, \ldots, m\}$ we shall now establish it for $k = m+1 \in \mathbb{N}$. To this end we first deduce from this inductive assumption, the estimate

$$\frac{1}{\tau_{\mathrm{n}}^{\alpha}}\sum_{l=0}^{k-1}\tau_{\mathrm{n}}^{l(\beta-\alpha)} \leq \frac{1}{\tau_{\mathrm{n}}^{\alpha} - \tau_{\mathrm{n}}^{\beta}}$$

for the sum on the right-hand side of (5.35), and the smallness hypothesis of the corollary that there holds

$$\begin{aligned}\Phi(u, \mathrm{B}_{\tau_{\mathrm{n}}^k R}(x_0)) &\leq \tau_{\mathrm{n}}^{k\beta}\Phi(u, \mathrm{B}_R(x_0)) + \frac{\tau_{\mathrm{n}}^{k\alpha}R^{\alpha}}{\varepsilon_{\mathrm{n}}(\tau_{\mathrm{n}}^{\alpha} - \tau_{\mathrm{n}}^{\beta})}\\ &\leq \tau_{\mathrm{n}}^{k\alpha}\frac{\widetilde{\varepsilon}_{\mathrm{n}}}{(\tau_{\mathrm{n}}^{\alpha} - \tau_{\mathrm{n}}^{\beta})\varepsilon_{\mathrm{n}}}|(\mathrm{D}u)_{\mathrm{B}_R(x_0)}|^p \leq \tau_{\mathrm{n}}^{k\alpha}\sqrt{\widetilde{\varepsilon}_{\mathrm{n}}}|(\mathrm{D}u)_{\mathrm{B}_R(x_0)}|^p \leq 1\end{aligned}$$

for all $k \in \{0,1,2,\dots,m\}$. Using the abbreviation $\xi_k := (\mathrm{D}u)_{\mathrm{B}_{\tau_{\mathrm{n}}^k R}(x_0)}$, (4.4), and the inequality $|z|^p \le 2^p\big[\mathrm{A}_p(|z|^p) + \mathrm{A}_p(|z|^p)^p\big]$ from Lemma 4.5 we conclude

$$
\begin{aligned}
(5.36)\quad & \big|\mathrm{V}_p(\xi_m) - \mathrm{V}_p(\xi_0)\big| \\
&\le \sum_{k=0}^{m-1} \big|\mathrm{V}_p(\xi_{k+1}) - \mathrm{V}_p(\xi_k)\big| \\
&\le C_1^{\frac12} \sum_{k=0}^{m-1} \mathrm{W}_p^{\xi_k}(\xi_{k+1}-\xi_k)^{\frac12} \\
&\le (2^p C_1)^{\frac12} \sum_{k=0}^{m-1} \big[\mathrm{A}_p(\mathrm{W}_p^{\xi_k}(\xi_{k+1}-\xi_k))^{\frac12} + \mathrm{A}_p(\mathrm{W}_p^{\xi_k}(\xi_{k+1}-\xi_k))^{\frac{p}{2}}\big] \\
&\le (2^p C_1)^{\frac12} \sum_{k=0}^{m-1} \bigg[\bigg(\fint_{\mathrm{B}_{\tau_{\mathrm{n}}^{k+1}R}(x_0)} \mathrm{A}_p(\mathrm{W}_p^{\xi_k}(\mathrm{D}u - \xi_k \mathscr{L}^n))\bigg)^{\frac12} \\
&\qquad\qquad + \bigg(\fint_{\mathrm{B}_{\tau_{\mathrm{n}}^{k+1}R}(x_0)} \mathrm{A}_p(\mathrm{W}_p^{\xi_k}(\mathrm{D}u - \xi_k \mathscr{L}^n))\bigg)^{\frac{p}{2}}\bigg] \\
&\le (2^p C_1 \tau_{\mathrm{n}}^{-np})^{\frac12} \sum_{k=0}^{m-1} \Big[\Phi(u,\mathrm{B}_{\tau_{\mathrm{n}}^k R}(x_0))^{\frac12} + \Phi(u,\mathrm{B}_{\tau_{\mathrm{n}}^k R}(x_0))^{\frac{p}{2}}\Big] \\
&\le 2(2^p C_1 \tau_{\mathrm{n}}^{-np}\sqrt{\widetilde{\varepsilon}_{\mathrm{n}}})^{\frac12} \sum_{k=0}^{\infty} \tau_{\mathrm{n}}^{k\alpha/2} |(\mathrm{D}u)_{\mathrm{B}_R(x_0)}|^{p/2} \\
&= 2\frac{(2^p C_1 \tau_{\mathrm{n}}^{-np}\sqrt{\widetilde{\varepsilon}_{\mathrm{n}}})^{\frac12}}{1-\tau_{\mathrm{n}}^{\alpha/2}} |\mathrm{V}_p(\xi_0)| \le \frac15 |\mathrm{V}_p(\xi_0)|\,,
\end{aligned}
$$

where we exploited in the last line the definition of $\widetilde{\varepsilon}_{\mathrm{n}}$. The previous estimate yields

$$\frac45 |\mathrm{V}_p(\xi_0)| \le |\mathrm{V}_p(\xi_m)| \le \frac65 |\mathrm{V}_p(\xi_0)|\,,$$

and via the definition of V_p in (4.3) and the choice of ξ_k we infer

$$\frac12 |(\mathrm{D}u)_{\mathrm{B}_R(x_0)}| \le \Big(\frac45\Big)^{\frac2p} |(\mathrm{D}u)_{\mathrm{B}_R(x_0)}| \le |(\mathrm{D}u)_{\mathrm{B}_{\tau_{\mathrm{n}}^m R}(x_0)}| \le \Big(\frac65\Big)^{\frac2p} |(\mathrm{D}u)_{\mathrm{B}_R(x_0)}| < \frac12 \sigma\,.$$

Consequently, we get

$$
\begin{aligned}
\Phi(u,\mathrm{B}_{\tau_{\mathrm{n}}^m R}(x_0)) &\le \tau_{\mathrm{n}}^{m\alpha} \sqrt{\widetilde{\varepsilon}_{\mathrm{n}}} |(\mathrm{D}u)_{\mathrm{B}_R(x_0)}|^p \\
&\le 2^p \tau_{\mathrm{n}}^{m\alpha} \sqrt{\widetilde{\varepsilon}_{\mathrm{n}}} |(\mathrm{D}u)_{\mathrm{B}_{\tau_{\mathrm{n}}^m R}(x_0)}|^p \le \varepsilon_{\mathrm{n}} |(\mathrm{D}u)_{\mathrm{B}_{\tau_{\mathrm{n}}^m R}(x_0)}|^p\,,
\end{aligned}
$$

where we used the definition of $\widetilde{\varepsilon}_{\mathrm{n}}$ once more. In view of the preceding estimates we are in the position to apply Corollary 5.10 on the ball $\mathrm{B}_{\tau_{\mathrm{n}}^m R}(x_0)$. Using also the inductive hypothesis we come out with

$$
\begin{aligned}
\Phi(u,\mathrm{B}_{\tau_{\mathrm{n}}^{m+1} R}(x_0)) &\le \tau_{\mathrm{n}}^{\beta} \Phi(u,\mathrm{B}_{\tau_{\mathrm{n}}^m R}(x_0)) + \varepsilon_{\mathrm{n}}^{-1} (\tau_{\mathrm{n}}^m R)^\alpha \\
&\le \tau_{\mathrm{n}}^{\beta} \tau_{\mathrm{n}}^{m\beta} \Phi(u,\mathrm{B}_R(x_0)) + \tau_{\mathrm{n}}^{\beta} \frac{(\tau_{\mathrm{n}}^m R)^\alpha}{\varepsilon_{\mathrm{n}} \tau_{\mathrm{n}}^\alpha} \sum_{l=0}^{m-1} \tau_{\mathrm{n}}^{l(\beta-\alpha)} + \frac{(\tau_{\mathrm{n}}^m R)^\alpha}{\varepsilon_{\mathrm{n}}} \\
&= \tau_{\mathrm{n}}^{(m+1)\beta} \Phi(u,\mathrm{B}_R(x_0)) + \frac{(\tau_{\mathrm{n}}^{m+1} R)^\alpha}{\varepsilon_{\mathrm{n}} \tau_{\mathrm{n}}^\alpha} \sum_{l=0}^{m} \tau_{\mathrm{n}}^{l(\beta-\alpha)}\,,
\end{aligned}
$$

and the induction is completed.

Finally, for $0 < \varrho \le R$ we choose some $k \in \mathbb{N}_0$ with $\tau_{\rm n}^{k+1} R \le \varrho \le \tau_{\rm n}^k R$. Then using Lemma 4.14 and (5.35) we arrive at

$$\begin{aligned}\Phi(u, \mathrm{B}_\varrho(x_0)) &\le \Big(\frac{\tau_{\rm n}^k R}{\varrho}\Big)^n \Phi(u, \mathrm{B}_{\tau_{\rm n}^k R}(x_0)) \\ &\le \Big(\frac{R}{\varrho}\Big)^n \tau_{\rm n}^{k(n+\beta)} \Phi(u, \mathrm{B}_R(x_0)) + \Big(\frac{R}{\varrho}\Big)^n \frac{\tau_{\rm n}^{k(n+\alpha)} R^\alpha}{\varepsilon_{\rm n}(\tau_{\rm n}^\alpha - \tau_{\rm n}^\beta)} \\ &\le \frac{\tau_{\rm n}^{-n-\beta}}{\varepsilon_{\rm n}(\tau_{\rm n}^\alpha - \tau_{\rm n}^\beta)} \bigg[\Big(\frac{\varrho}{R}\Big)^\beta \Phi(u, \mathrm{B}_R(x_0)) + \varrho^\alpha\bigg]. \end{aligned} \tag{5.37}$$

∎

5.3.2. *General case.* Following ideas from [26] we finally merge our degenerate and non-degenerate estimates. Specifically, we combine Corollary 5.9 and Proposition 5.11.

Proposition 5.12. *Suppose that $u \in \mathrm{BV}(\Omega)^N$ is an L^q-α-minimizer of F at $x_0 \in \Omega$ with $\alpha \in (0, \gamma_p)$. Then there exists a number ε_0, depending only on n, N, p, Γ, σ, $\eta_{\rm d}$, $\eta_{\rm n}$, ω, and α, such that for every ball $\mathrm{B}_R(x_0) \subset \Omega$ the condition*

$$\Phi(u, \mathrm{B}_R(x_0)) + R^\alpha \le \varepsilon_0$$

together with $|(\mathrm{D}u)_{\mathrm{B}_R(x_0)}| < \frac{1}{8}\sigma$ implies the decay estimate

$$\Phi(u, \mathrm{B}_\varrho(x_0)) \le C\bigg[\Big(\frac{\varrho}{R}\Big)^{\frac{\alpha+\gamma_p}{2}} \Phi(u, \mathrm{B}_R(x_0)) + \varrho^\alpha\bigg] \qquad \text{for all } 0 < \varrho \le R\,.$$

Here, C depends only on n, N, p, Γ, σ, $\eta_{\rm d}$, $\eta_{\rm n}$, ω, and α.

Proof. We set

$$\beta := \frac{\alpha + \gamma_p}{2} \in (\alpha, \gamma_p)$$

and fix

$$\sqrt{\varepsilon_0} := \min\{\varepsilon_{\rm d}(\tau_{\rm d}^\alpha - \tau_{\rm d}^\beta), (4C_1)^{-1} 48^{-p} \tau_{\rm d}^{np} (1 - \tau_{\rm d}^{\alpha/2})^2 \sigma^p\}\,,$$

where C_1 is the constant in (4.4). Furthermore, $0 < \tau_{\rm d} \le \frac{1}{4}$ and $0 < \varepsilon_{\rm d} \le 1$ are the constants from Corollary 5.9 corresponding to the choice $\chi := \widetilde{\varepsilon}_{\rm n}$ with the constant $\widetilde{\varepsilon}_{\rm n}$ from Proposition 5.11. For the constant $\tau_{\rm d}$ from Corollary 5.9 we now denote by m the smallest number in $\mathbb{N}_0$ such that there holds

$$\Phi(u, \mathrm{B}_{\tau_{\rm d}^m R}(x_0)) + R^\alpha < \chi |(\mathrm{D}u)_{\mathrm{B}_{\tau_{\rm d}^m R}(x_0)}|^p,$$

where we set $m := \infty$ if no such number exists at all. Following a similar — but somewhat simpler — line of argument as in the proof of Proposition 5.11 we apply Corollary 5.9 inductively on the balls $\mathrm{B}_R(x_0)$, $\mathrm{B}_{\tau_{\rm d} R}(x_0)$, $\mathrm{B}_{\tau_{\rm d}^2 R}(x_0)$, $\ldots$, $\mathrm{B}_{\tau_{\rm d}^{m-1} R}(x_0)$ and infer

$$\Phi(u, \mathrm{B}_{\tau_{\rm d}^k R}(x_0)) \le \tau_{\rm d}^{k\beta} \Phi(u, \mathrm{B}_R(x_0)) + \frac{(\tau_{\rm d}^k R)^\alpha}{\varepsilon_{\rm d} \tau_{\rm d}^\alpha} \sum_{l=0}^{k-1} \tau_{\rm d}^{l(\beta-\alpha)} \qquad \text{for all } k \in \mathbb{N}_0 \text{ with } k \le m\,. \tag{5.38}$$

Here we exploited that the previous inequality implies

$$\Phi(u, \mathrm{B}_{\tau_{\rm d}^k R}(x_0)) + (\tau_{\rm d}^k R)^\alpha \le \frac{\varepsilon_0}{\varepsilon_{\rm d}(\tau_{\rm d}^\alpha - \tau_{\rm d}^\beta)} \le \sqrt{\varepsilon_0} \le \varepsilon_{\rm d}$$

by our choice of ε_0 and thus in every inductive step one hypothesis of Corollary 5.9 is available on a smaller ball (while the other hypothesis is ensured on all relevant balls by the definition of m). Moreover, if m is finite, then by essentially the same reasoning as for (5.36) above and again the choice of ε_0 we have

$$|\mathrm{V}_p((\mathrm{D}u)_{\mathrm{B}_{\tau_{\mathrm{d}}^m R(x_0)}}) - \mathrm{V}_p((\mathrm{D}u)_{\mathrm{B}_R(x_0)})| \le 2\frac{(2^p C_1 \tau_{\mathrm{d}}^{-np}\sqrt{\varepsilon_0})^{\frac{1}{2}}}{1-\tau_{\mathrm{d}}^{\alpha/2}} \le \big(\tfrac{1}{24}\sigma\big)^{\frac{p}{2}}\,,$$

and as a consequence we get

$$|(\mathrm{D}u)_{\mathrm{B}_{\tau_{\mathrm{d}}^m R(x_0)}}| < \Big[\big(\tfrac{1}{8}\sigma\big)^{\frac{p}{2}} + \big(\tfrac{1}{24}\sigma\big)^{\frac{p}{2}}\Big]^{\frac{2}{p}} < \frac{1}{3}\sigma\,.$$

By the definition of m and the choice of χ we may thus apply Proposition 5.11 on $\mathrm{B}_{\tau_{\mathrm{d}}^m R}(x_0)$ coming out with

$$\Phi(u,\mathrm{B}_{\tau_{\mathrm{d}}^k R}(x_0)) \le C\Big[\tau_{\mathrm{d}}^{(k-m)\beta}\Phi(u,\mathrm{B}_{\tau_{\mathrm{d}}^m R}(x_0)) + (\tau_{\mathrm{d}}^k R)^\alpha\Big] \qquad \text{for all } k\in\mathbb{N}_0 \text{ with } k \ge m\,, \tag{5.39}$$

where we also used $\frac{\alpha+2}{2} \ge \beta$. Putting together the decay estimates (5.38) and (5.39) we have

$$\Phi(u,\mathrm{B}_{\tau_{\mathrm{d}}^k R}(x_0)) \le C\left[\tau_{\mathrm{d}}^{k\beta}\Phi(u,\mathrm{B}_R(x_0)) + \frac{(\tau_{\mathrm{d}}^k R)^\alpha}{\varepsilon_{\mathrm{d}}(\tau_{\mathrm{d}}^\alpha - \tau_{\mathrm{d}}^\beta)}\right] \qquad \text{for all } k\in\mathbb{N}_0$$

and in any case. The proof of the proposition is now completed by a computation analogous to (5.37). □

5.4. **Conclusion.** We now assume that we are in the situation of either Theorem 2.5 or Proposition 2.7, and we finalize the proof of these results.

We first notice that Theorem 2.5 can indeed be reduced to the case where the normalization (5.1) is valid. Indeed, to this end it suffices to consider in place of u the L^q-α-minimizer $\widetilde{u}$ of $\widetilde{F}$, where we have set $\widetilde{u}(x) := u(x) - z_0 x$, $\widetilde{f}(z) := \theta^{-1}\big[f(z_0{+}z) - f(z_0) - \nabla f(z_0)z\big]$, and $\widetilde{F}[w] := \int_\Omega \widetilde{f}(\mathrm{D}w)$. In the situation of Proposition 2.7 we may analogously pass on to the case where (5.1) holds for each $f(x_0,\,\cdot\,)$ in place of f.

Having said this, the previous proposition is applicable, and it suffices to consider a Lebesgue point $x_0\in\Omega$ of $\mathrm{D}u$ with Lebesgue value $z_0 = 0$. By Definition 4.13 we can take a positive radius R with $\mathrm{B}_{2R}(x_0) \subset\subset \Omega$,

$$R^\alpha \le \frac{1}{2}\varepsilon_0\,, \qquad \text{and} \qquad |(\mathrm{D}u)_{\mathrm{B}_{2R}(x_0)}| < 2^{-n-3}\min\{\varepsilon_0,\sigma\}\,.$$

As a consequence we get

$$|(\mathrm{D}u)_{\mathrm{B}_R(y)}| \le 2^n |(\mathrm{D}u)_{\mathrm{B}_{2R}(x_0)}| < \frac{1}{8}\sigma\,,$$
$$\Phi(u,\mathrm{B}_R(y)) \le 4|(\mathrm{D}u)_{\mathrm{B}_R(y)}| \le 2^{n+2}|(\mathrm{D}u)_{\mathrm{B}_{2R}(x_0)}| \le \frac{1}{2}\varepsilon_0$$

for all $x\in\mathrm{B}_R(x_0)$. Hence, we are in the Position to apply Proposition 5.12 on the balls $\mathrm{B}_R(y)$, either using it for F or for the frozen functional F_y (in the later case notice that ε_0 and C do not depend on y). As we have $(\mathrm{A}_p\circ\mathrm{W}_p^\xi)^\infty(z) = |z|$, the proposition yields

$$\fint_{\mathrm{B}_\varrho(y)} |\mathrm{D}^{\mathrm{s}}u| \le \Phi(u,\mathrm{B}_\varrho(y)) \le C\Big(\frac{\varrho}{R}\Big)^\alpha\big[\Phi(u,\mathrm{B}_R(y)) + R^\alpha\big] \le C\Big(\frac{\varrho}{R}\Big)^\alpha\varepsilon_0$$

for all $0<\varrho\le R$ and $y\in\mathrm{B}_R(x_0)$. In particular, $\frac{|\mathrm{D}^{\mathrm{s}}u|(\mathrm{B}_\varrho(y))}{\mathscr{L}^n(\mathrm{B}_\varrho(y))}$ is uniformly bounded, and thus the singular part $\mathrm{D}^{\mathrm{s}}u$ in the Lebesgue decomposition of $\mathrm{D}u$ vanishes on

$\mathrm{B}_R(x_0)$. In the case $p \geq 2$ we now employ the estimate $|z| \leq 2\mathrm{A}_p(|z|^p)^{\frac{1}{p}} + 2\mathrm{A}_p(|z|^p)$ of Lemma 4.5 and Hölder's inequality, and we control $\Phi(u, \mathrm{B}_\varrho(y))$ as before to get

$$\begin{aligned}
&\fint_{\mathrm{B}_\varrho(y)} |\nabla u - (\nabla u)_{\mathrm{B}_\varrho(y)}| \,\mathrm{d}x \\
&\qquad \leq 2\fint_{\mathrm{B}_\varrho(y)} \mathrm{A}_p(|\nabla u - (\nabla u)_{\mathrm{B}_\varrho(y)}|^p)^{\frac{1}{p}} \,\mathrm{d}x + 2\fint_{\mathrm{B}_\varrho(y)} \mathrm{A}_p(|\nabla u - (\nabla u)_{\mathrm{B}_\varrho(y)}|^p) \,\mathrm{d}x \\
&\qquad \leq 2\Phi(u, \mathrm{B}_\varrho(y))^{\frac{1}{p}} + 2\Phi(u, \mathrm{B}_\varrho(y)) \\
&\qquad \leq C\Big(\frac{\varrho}{R}\Big)^{\frac{\alpha}{p}} \varepsilon_0^{\frac{1}{p}}
\end{aligned}$$

for all $0 < \varrho \leq R$ and still all $y \in \mathrm{B}_R(x_0)$. From this bound we deduce by the Campanato space characterization of Hölder continuity (see [32, Theorem 5.4]) that ∇u is Hölder continuous on $\mathrm{B}_R(x_0)$ with exponent α/p. In the case $1 < p < 2$ in contrast we use (4.4) and the inequality $|z|^{\frac{p}{2}} \leq C\big[\mathrm{A}_p(|z|^p)^{\frac{1}{2}} + \mathrm{A}_p(|z|^p)^{\frac{p}{2}}\big]$ to control similarly

$$\begin{aligned}
\fint_{\mathrm{B}_\varrho(y)} \big|\mathrm{V}_p(\nabla u) - \big[\mathrm{V}_p(\nabla u)\big]_{\mathrm{B}_\varrho(y)}\big| \,\mathrm{d}x &\leq 2\fint_{\mathrm{B}_\varrho(y)} \big|\mathrm{V}_p(\nabla u) - \mathrm{V}_p((\nabla u)_{\mathrm{B}_\varrho(y)})\big| \,\mathrm{d}x \\
&\leq C\fint_{\mathrm{B}_\varrho(y)} \Big[\mathrm{W}_p^{(\nabla u)_{\mathrm{B}_\varrho(y)}}(\nabla u - (\nabla u)_{\mathrm{B}_\varrho(y)})\Big]^{\frac{1}{2}} \,\mathrm{d}x \\
&\leq C\Big[\Phi(u, \mathrm{B}_\varrho(y))^{\frac{1}{2}} + \Phi(u, \mathrm{B}_\varrho(y))^{\frac{p}{2}}\Big] \\
&\leq C\Big(\frac{\varrho}{R}\Big)^{\frac{\alpha}{2}} \varepsilon_0^{\frac{1}{2}}\,.
\end{aligned}$$

The preceding estimate gives Hölder continuity of $\mathrm{V}_p(\nabla u)$ on $\mathrm{B}_R(x_0)$ with exponent $\alpha/2$, and in particular ∇u is bounded on $\mathrm{B}_R(x_0)$. Relying on (4.4) once more we therefore have

$$\begin{aligned}
\Big(3 \sup_{\mathrm{B}_r(x_0)} |\nabla u|\Big)^{p-2} |\nabla u(x) - \nabla u(y)|^2 &\leq \mathrm{W}_p^{\nabla u(x)}(\nabla u(y) - \nabla u(x)) \\
&\leq |\mathrm{V}_p(\nabla u(y)) - \mathrm{V}_p(\nabla u(x))|^2
\end{aligned}$$

for all $x, y \in \mathrm{B}_R(x_0)$, and ∇u is Hölder continuous with the same exponent $\alpha/2$ as $\mathrm{V}_p(\nabla u)$ on $\mathrm{B}_{R/2}(x_0)$. In summary, we arrive at the claimed $\mathrm{C}^{1,\alpha/\max\{p,2\}}$-regularity of u near x_0 in all cases. □

Finally, we justify Remark 2.6. To this end we first record that we exploited (H3) solely in the proof of Proposition 5.2, and even there the special form of e_p entered only through the estimate for the p-Laplace system. However, if for $p = 2$ any positive bilinear form takes the role of e_2 in (H3), then we can just use the estimate (4.17) from linear theory in place of (4.21), and we find that Proposition 5.2 and the whole reasoning of Section 5 remain valid. Needless to say, in this case many arguments can be extensively simplified, and actually the distinction between the degenerate and non-degenerate case is not necessary anymore. □

Appendix A. Local boundedness

In this appendix we consider the integrals from (1.2), and we provide interior L^∞-bounds for local minimizers of $F+G$, as they are needed for the purposes of Theorem 1.3. We will only sketch the proof, which resembles an argument from [9].

Theorem A.1. *Assume that the Borel functions* $f\colon \Omega \times \mathbb{R}^{Nn} \to \mathbb{R}$ *and* $g\colon \Omega \times \mathbb{R}^N \to \mathbb{R}$ *are* C^1 *in their second argument. Moreover, suppose that* f *is also convex in the second argument, and that the following set of assumptions holds for all* $x \in \Omega$, $y \in \mathbb{R}^N$, $z \in \mathbb{R}^{Nn}$, *some* $\zeta \in [1,\infty)$, *some positive* C, *and some function* $P \in \mathrm{L}^\infty_{\mathrm{loc}}(\Omega)$*:*

$$C^{-1}|z| - C \le f(x,z) \le C|z| + C\,,$$
$$\sum_{i=1}^n \sum_{k,l=1}^N y^k \frac{\partial f}{\partial z_i^k}(x,z) z_i^l y^l \ge -C|y|^2\,,$$
$$|\nabla_y g(x,y)| \le C|y|^{\zeta-1} + P(x)\,,$$
$$\nabla_y g(x,y) y \ge -C|y| - P(x)\,. \tag{A.1}$$

Then every local minimizer $u \in \mathrm{BV}_{\mathrm{loc}}(\Omega)^N \cap \mathrm{L}^\zeta_{\mathrm{loc}}(\Omega)^N$ *of* $F{+}G$ *is in* $\mathrm{L}^\infty_{\mathrm{loc}}(\Omega)^N$.

Sketch of proof. We may assume that $u \in \mathrm{BV}(\Omega)^N \cap \mathrm{L}^\zeta(\Omega)^N$ is a minimizer for the Dirichlet problem on a Lipschitz domain Ω. If g vanishes and $f(x,z)$ is independent of x, then the claim follows from [9, Theorem 1.11], which has been proved in [9, Section 4] by an adaption of Moser's iteration technique. Moreover, it has been pointed out in [9, Appendix C] that the x-dependence of f does in no way affect this approach. Now we briefly explain why the occurrence of g does not suspend the line of argument as well. Indeed, it suffices to cope with an extra term like

$$\int_\Omega \nabla_y g(\,\cdot\,,u_k)\varphi \,\mathrm{d}x \tag{A.2}$$

on the left-hand side of the approximative Euler equation in [9, Section 4]. Here, $\varphi \in \mathrm{W}^{1,1}_0(\Omega)^N \cap \mathrm{L}^\zeta(\Omega)^N$ is an arbitrary test-function, and the approximations $u_k \in \mathrm{W}^{1,1}(\Omega)^N \cap \mathrm{L}^\zeta(\Omega)^N$ of u are obtained via Ekeland's variational principle[5]. When we insert the specific φ from [9, Proof of Lemma 4.1], then (A.2) can be bounded from below via (A.1). The resulting terms, shifted to the right-hand side of the estimates, are basically the same ones already treated in [9], and thus the conclusion of [9, Lemma 4.1] remains valid (up to very minor adaptions, namely passing to $M_\eta := \max_\Omega(\eta{+}|\nabla\eta|)$ and replacing λ by $C{+}\|P\|_{\mathrm{L}^\infty(\operatorname{spt}\eta)}$). Once the lemma is available, the remaining arguments of [9, Section 4] apply unchanged. □

For further information and an extensive discussion of the above assumptions on f we refer to [9, Section 1.2]. Concerning the assumptions on g, some refinements are possible, but for simplicity we have limited ourselves to the above statement. Anyway, in view of the following lemma this statement suffices for our purposes.

Lemma A.2. *The assumptions of Theorem A.1 on* g *are satisfied in the case* $g(x,y) := \lambda|y{-}S(x)|^\zeta$ *with* $\lambda \in [0,\infty)$, $\zeta \in (1,\infty)$, *and* $S \in \mathrm{L}^\infty_{\mathrm{loc}}(\Omega)^N$.

Proof. Evidently, $g(x,y)$ is C^1 in y, and from $\nabla_y g(x,y) = \lambda\zeta|y{-}S(x)|^{\zeta-2}(y{-}S(x))$ we derive by Young's inequality

$$|\nabla_y g(x,y)| \le 2^{\zeta-1}\lambda\zeta\big[|y|^{\zeta-1} + |S(x)|^{\zeta-1}\big]$$

[5]The Ekeland principle is applied on the Dirichlet class $\mathcal{D}$ of u as in [9, Section 4], but for $\zeta > \frac{n}{n-1}$ a technical adaption is needed in order to guarantee $u_k \in \mathrm{L}^\zeta(\Omega)^N$. One way of doing this is to choose the w_k of [9, Section 4] additionally in $\mathrm{L}^\zeta(\Omega)^N$ (this is possible since $\mathrm{L}^\zeta(\Omega)^N$ is dense in $\mathcal{D}$) and to apply the principle to the functionals $F{+}G{+}\varepsilon_k\|\cdot\|^\zeta_{\mathrm{L}^\zeta(\Omega)^N}$ with $\varepsilon_k := k^{-2}(1{+}\|w_k\|^\zeta_{\mathrm{L}^\zeta(\Omega)^N})^{-1}$.

and

$$\nabla_y g(x,y)y = \lambda\zeta\big[|y-S(x)|^\zeta - |y-S(x)|^{\zeta-2}(y-S(x))S(x)\big] \geq -\lambda\zeta|S(x)|^\zeta\,. \qquad \square$$

Appendix B. Sketch of proof for Remark 2.2

We show that (2.4) implies the conditions (H2), (H3), and (H4).

Regarding (H2) we first note that the condition holds for $(\mathrm{e}_p,0)$ instead of (f,z_0) with a positive constant $\sigma_p \leq 1$ depending only on p; then by (2.4) we can pass to some positive $\sigma \leq \frac{1}{2}\sigma_p$ such that $|\nabla^2 f(z_0{+}z) - \nabla^2\mathrm{e}_p(z)| \leq \frac{1}{2}\sigma_p|z|^{p-2}$ holds for $|z| \leq \sigma$. In view of this estimate (H2) carries over to (f,z_0) in the required form.

Turning to (H3) we abbreviate $h(z) := f(z_0{+}z) - \theta\mathrm{e}_p(z)$ and estimate

$$\begin{aligned}\frac{|h(z)-h(0)-\nabla h(0)z|}{|z|^p} &\leq \int_0^1\int_0^1 \frac{|\nabla^2 h(stz)|}{|z|^{p-2}}\,\mathrm{d}s\,t\,\mathrm{d}t\\ &\leq \int_0^1\int_0^1 \frac{|\nabla^2 h(stz)|}{|stz|^{p-2}}\,\mathrm{d}s\,\mathrm{d}t\end{aligned}$$

for $|z| < \sigma$. By (2.4) the right-hand side of this estimate converges to 0 for $z \to 0$. Consequently, also the left-hand side vanishes in the limit which just corresponds to (H3).

Finally, for (H4) we argue as follows: One first verifies by homogeneity that for $2|z| \leq |\xi|$ there hold

$$\begin{aligned}|\nabla^2\mathrm{e}_p(\xi) - \nabla^2\mathrm{e}_p(\xi{+}z)| &\leq C|\xi|^{p-3}|z|\,,\\ \big||\xi{+}z|^{p-2} - |\xi|^{p-2}\big| &\leq C|\xi|^{p-3}|z|\end{aligned}$$

with some constant C depending only on p. Setting

$$h(z) := \frac{\nabla^2 f(z_0{+}z) - \theta\nabla^2\mathrm{e}_p(z)}{|z|^{p-2}} \qquad \text{for } |z| < \sigma$$

we thus get the estimate

$$\begin{aligned}|\nabla^2 f(z_0{+}\xi) - \nabla^2 f(z_0{+}\xi{+}z)| &\leq |\xi|^{p-2}|h(\xi)-h(\xi{+}z)|\\ &\qquad + \big||\xi|^{p-2} - |\xi{+}z|^{p-2}\big|\,|h(\xi{+}z)|\\ &\qquad + \theta|\nabla^2\mathrm{e}_p(\xi) - \nabla^2\mathrm{e}_p(\xi{+}z)|\\ &\leq |\xi|^{p-2}|h(\xi)-h(\xi{+}z)| + C\big[|h(\xi+z)|+\theta\big]|\xi|^{p-3}|z|\\ &=: A+B\end{aligned}$$

for $2|z| \leq |\xi| < \frac{1}{2}\sigma$. Now we notice that by the assumption (2.4) h is uniformly continuous and bounded near 0. Given $\mu > 0$ we can thus choose $\eta_{\mathrm{n}}(\mu) \in (0,\frac{1}{2}]$ such that $|z| \leq \eta_{\mathrm{n}}(\mu)$ implies $A \leq \frac{1}{2}\mu|\xi|^{p-2}$ and such that $|z| \leq \eta_{\mathrm{n}}(\mu)|\xi|$ implies $B \leq \frac{1}{2}\mu|\xi|^{p-2}$. At this point, (H4) easily follows. $\square$

References

1. E. Acerbi and N. Fusco, *Regularity for minimizers of non-quadratic functionals: the case $1<p<2$*, J. Math. Anal. Appl. **140** (1989), 115–135.
2. W.K. Allard, *Total variation regularization for image denoising. I: Geometric theory*, SIAM J. Math. Anal. **39** (2007), 1150–1190.
3. F.J Almgren, *Existence and regularity almost everywhere of solutions to elliptic variational problems with constraints*, Mem. Am. Math. Soc. **4** (1976), no. 165, 199 pages.
4. L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford University Press, Oxford, 2000.

5. L. Ambrosio and E. Paolini, *Partial regularity for quasi minimizers of perimeter*, Ric. Mat. **48** (1999), 167–186; correction in Ric. Mat. **50** (2001), 191–193.
6. G. Anzellotti, *On the $C^{1,\alpha}$-regularity of ω-minima of quadratic functionals*, Boll. Unione Mat. Ital., VI. Ser., C, Anal. Funz. Appl. **2** (1983), 195–212.
7. ———, *The Euler equation for functionals with linear growth*, Trans. Am. Math. Soc. **290** (1985), 483–501.
8. G. Anzellotti and M. Giaquinta, *Convex functionals and partial regularity*, Arch. Ration. Mech. Anal. **102** (1988), 243–272.
9. L. Beck and T. Schmidt, *On the Dirichlet problem for variational integrals in BV*, J. Reine Angew. Math. **674** (2013), 113–194.
10. ———, *Interior gradient regularity for* BV*-minimizers of singular variational problems*, Preprint (2014), 26 pages.
11. L. Beck and B. Stroffolini, *Regularity results for differential forms solving degenerate elliptic systems*, Calc. Var. Partial Differ. Equ., to appear (2013), 40 pages.
12. M. Bildhauer, *A priori gradient estimates for bounded generalized solutions of a class of variational problems with linear growth*, J. Convex Anal. **9** (2002), 117–137.
13. ———, *Convex variational problems. Linear, nearly linear and anisotropic growth conditions*, Springer Lecture Notes in Mathematics **1818**, Berlin, 2003.
14. ———, *Two dimensional variational problems with linear growth*, Manuscr. Math. **110** (2003), 325–342.
15. V. Bögelein, *Partial regularity for minimizers of discontinuous quasi-convex integrals with degeneracy*, J. Differ. Equations **252** (2012), 1052–1100.
16. V. Bögelein, F. Duzaar, and G. Mingione, *The regularity of general parabolic systems with degenerate diffusion*, Mem. Am. Math. Soc. **221** (2013), no. 1041, 143 pages.
17. E. Bombieri, *Regularity theory for almost minimal currents*, Arch. Ration. Mech. Anal. **78** (1982), 99–130.
18. V. Caselles, A. Chambolle, and M. Novaga, *The discontinuity set of solutions of the TV denoising problem and some extensions*, Multiscale Model. Simul. **6** (2007), 879–894.
19. ———, *Regularity for solutions of the total variation denoising problem*, Rev. Mat. Iberoam. **27** (2011), 233–252.
20. A. Chambolle and P.-L. Lions, *Image recovery via total variation minimization and related problems*, Numer. Math. **76** (1997), 167–188.
21. T.F. Chan and S. Esedoḡlu, *Aspects of total variation regularized L^1 function approximation*, SIAM J. Appl. Math. **65** (2005), 1817–1837.
22. Y. Chen, M. Rao, Y. Tonegawa, and T. Wunderli, *Partial regularity for a selective smoothing functional for image restoration in* BV *space*, SIAM J. Math. Anal. **37** (2005), 1098–1116.
23. F. Duzaar, A. Gastel, and J.F. Grotowski, *Partial regularity for almost minimizers of quasi-convex integrals*, SIAM J. Math. Anal. **32** (2000), 665–687.
24. F. Duzaar, J.F. Grotowski, and M. Kronz, *Partial and full boundary regularity for minimizers of functionals with nonquadratic growth*, J. Convex Anal. **11** (2004), 437–476.
25. ———, *Regularity of almost minimizers of quasi-convex variational integrals with sub-quadratic growth*, Ann. Mat. Pura Appl. **11** (2005), 421–448.
26. F. Duzaar and G. Mingione, *Regularity for degenerate elliptic problems via p-harmonic approximation*, Ann. Inst. Henri Poincaré, Anal. Non Linéaire **21** (2004), 735–766.
27. ———, *The p-harmonic approximation and the regularity of p-harmonic maps*, Calc. Var. **20** (2004), 235–256.
28. F. Duzaar and K. Steffen, *Optimal interior and boundary regularity for almost minimizers to elliptic variational integrals*, J. Reine Angew. Math. **546** (2002), 73–138.
29. L. Esposito and G. Mingione, *Partial regularity for minimizers of degenerate polyconvex energies*, J. Convex Anal. **8** (2001), 1–38.
30. N. Fusco and J.E. Hutchinson, *Partial regularity for minimisers of certain functionals having nonquadratic grwoth*, Ann. Mat. Pura Appl., IV. Ser. **155** (1989), 1–24.
31. M. Giaquinta and E. Giusti, *Sharp estimates for the derivatives of local minima of variational integrals*, Boll. Unione Mat. Ital., VI. Ser., A **3** (1984), 239–248.
32. M. Giaquinta and L. Martinazzi, *An introduction to the regularity theory for elliptic systems, harmonic maps and minimal graphs*, Edizioni della Normale, Pisa, 2005.
33. M. Giaquinta and G. Modica, *Remarks on the regularity of the minimizers of certain degenerate functionals*, Manuscr. Math. **57** (1986), 55–99.
34. M. Giaquinta, G. Modica, and J. Souček, *Functionals with linear growth in the calculus of variations*, Commentat. Math. Univ. Carol. **20** (1979), 143–156, 157–172.

35. E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*, Birkhäuser, Basel, 1984.
36. C. Goffman and J. Serrin, *Sublinear functions of measures and variational integrals*, Duke Math. J. **31** (1964), 159–178.
37. C. Hamburger, *Regularity of differential forms minimizing degenerate elliptic functionals*, J. Reine Angew. Math. **431** (1992), 7–64.
38. ———, *Optimal partial regularity of minimizers of quasiconvex variational integrals*, ESAIM, Control Optim. Calc. Var. **13** (2007), 639–656.
39. R. Hardt and D. Kinderlehrer, *Elastic plastic deformation*, Appl. Math. Optimization **10** (1983), 203–246.
40. ———, *Variational principles with linear growth*, Partial differential equations and the calculus of variations, Vol. II. Progr. Nonlinear Differential Equations Appl. **2** (1989), 633–659.
41. P. Marcellini and G. Papi, *Nonlinear elliptic systems with general growth*, J. Differ. Equations **221** (2006), 412–443.
42. M. Meier, *Boundedness and integrability properties of weak solutions of quasilinear elliptic systems*, J. Reine Angew. Math. **333** (1982), 191–220.
43. H.R. Parks, *Partial regularity of functions of least gradient in* $\mathbb{R}^8$, Ann. Sc. Norm. Super. Pisa, Cl. Sci., IV. Ser. **15** (1988), 447–451.
44. H.R. Parks and W.P. Ziemer, *Jacobi fields and regularity of functions of least gradients*, Ann. Sc. Norm. Super. Pisa, Cl. Sci., IV. Ser. **11** (1984), 505–527.
45. D. Phillips, *A minimization problem and the regularity of solutions in the presence of a free boundary*, Indiana Univ. Math. J. **32** (1983), 1–17.
46. L.I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D **60** (1992), 259–268.
47. T. Schmidt, *Regularity theorems for degenerate quasiconvex energies with* (p,q)*-growth*, Adv. Calc. Var. **1** (2008), 241–270.
48. ———, *A simple partial regularity proof for minimizers of variational integrals*, NoDEA, Nonlinear Differential Equations Appl. **16** (2009), 109–129.
49. R. Schoen and L. Simon, *A new proof of the regularity theorem for rectifiable currents which minimize parametric elliptic functionals*, Indiana Univ. Math. J. **31** (1982), 415–434.
50. J. Serrin, *The problem of Dirichlet for quasilinear elliptic differential equations with many independent variables*, Philos. Trans. R. Soc. Lond., Ser. A **264** (1969), 413–496.
51. L. Simon, *Interior gradient bounds for non-uniformly elliptic equations*, Indiana Univ. Math. J. **25** (1976), 821–855.
52. P. Sternberg, G. Williams, and W.P. Ziemer, *Existence, uniqueness, and regularity for functions of least gradient*, J. Reine Angew. Math. **430** (1992), 35–60.
53. I. Tamanini, *Boundaries of Caccioppoli sets with Hölder-continuous normal vector*, J. Reine Angew. Math. **334** (1982), 27–39.
54. ———, *Regularity results for almost minimal oriented hypersurfaces in* $\mathbb{R}^n$, Quaderni del Dipartimento di Matematica dell' Università di Lecce (1984), no. 1, 92 pages.
55. E. Tausch, *A class of variational problems with linear growth*, Math. Z. **164** (1978), 159–178.
56. P. Tolksdorf, *Everywhere-regularity for some quasilinear systems with a lack of ellipticity*, Ann. Mat. Pura Appl., IV. Ser. **134** (1983), 241–266.
57. N.S. Trudinger, *The boundary gradient estimate for quasilinear elliptic and parabolic differential equations*, Indiana Univ. Math. J. **21** (1972), 657–670.
58. K. Uhlenbeck, *Regularity for a class of nonlinear elliptic systems*, Acta Math. **138** (1977), 219–240.
59. T. Wunderli, *A partial regularity result for an anisotropic smoothing functional for image restoration in BV space*, J. Math. Anal. Appl. **339** (2008), 1169–1178.

Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland
E-mail address: thomas.schmidt@math.uzh.ch

SNS Pisa, Piazza dei Cavalieri 7, 56126 Pisa, Italy
E-mail address: thomas.schmidt@sns.it