Higher regularity in congested traffic dynamics

Verena Bögelein[1] · **Frank Duzaar[2]** · **Raffaella Giova[3]** ·
Antonia Passarelli di Napoli[4]

Abstract

In this paper, we consider minimizers of integral functionals of the type

$$\mathcal{F}(u) := \int_{\Omega} \big[\tfrac{1}{p}\big(|Du| - 1\big)_+^p + f \cdot u \big] \mathrm{d}x$$

for $p > 1$ in the vectorial case of mappings $u \colon \mathbb{R}^n \supset \Omega \to \mathbb{R}^N$ with $N \geq 1$. Assuming that f belongs to $L^{n+\sigma}$ for some $\sigma > 0$, we prove that $\mathcal{H}(Du)$ is continuous in Ω for any continuous function $\mathcal{H} \colon \mathbb{R}^{Nn} \to \mathbb{R}^{Nn}$ vanishing on $\{\xi \in \mathbb{R}^{Nn} : |\xi| \leq 1\}$. This extends previous results of Santambrogio and Vespri (Nonlinear Anal 73:3832–3841, 2010) when $n = 2$, and Colombo and Figalli (J Math Pures Appl (9) 101(1):94–117, 2014) for $n \geq 2$, to the vectorial case $N \geq 1$.

Mathematics Subject Classification 49N60 · 35D10 · 35J70

Communicated by Y. Giga.

✉ Frank Duzaar
 frank.duzaar@fau.de

 Verena Bögelein
 verena.boegelein@sbg.ac.at

 Raffaella Giova
 raffaella.giova@uniparthenope.it

 Antonia Passarelli di Napoli
 antpassa@unina.it

[1] Fachbereich Mathematik, Universität Salzburg, Hellbrunner Str. 34, 5020 Salzburg, Austria

[2] Department Mathematik, Universität Erlangen-Nürnberg, Cauerstrasse 11, 91058 Erlangen, Germany

[3] Università degli Studi di Napoli "Parthenope", Palazzo Pacanowsky-Via Generale Parisi, 13, 80132 Naples, Italy

[4] Università degli studi di Napoli "Federico II", Dipartimento di Matematica e Appl. "R. Caccioppoli", Via Cintia 80126, Naples, Italy

Contents

1 Introduction and main result

In this paper, we study C^1-regularity of minimizers of integral functionals of the Calculus of Variations with widely degenerate convex integrands of the form

$$\mathcal{F}(u) := \int_\Omega \big[F(Du) + f \cdot u\big]\mathrm{d}x, \tag{1.1}$$

where $\Omega \subset \mathbb{R}^n$, $n \geq 2$, is a bounded domain and $u\colon \Omega \to \mathbb{R}^N$, $N \geq 1$, a possibly vector valued function. We concentrate ourself on the study of the prototype integrand

$$F(\xi) := \tfrac{1}{p}(|\xi| - 1)_+^p, \tag{1.2}$$

for some $p > 1$. The datum f is required to belong to $L^{n+\sigma}$ for some $\sigma > 0$. The functional $\mathcal{F}$ with the specific integrand F from (1.2) is the prototype for a class of more general functionals where F is a convex function vanishing inside some convex set, and satisfying specific growth and ellipticity assumptions. For sake of clarity, the results in this paper are stated and proved for the functionals $\mathcal{F}(u)$ as in (1.1)–(1.2). However, we expect our techniques to apply to a general class of integrands with a widely degenerate structure as well. The functional $\mathcal{F}(u)$ and its associated Euler–Lagrange system

$$\mathrm{div}\left(\big(|Du| - 1\big)_+^{p-1} \frac{Du}{|Du|}\right) = f \tag{1.3}$$

naturally arise in problems of optimal transport with congestion effects. In fact, minimizing (1.1) with $N = 1$ and the integrand from (1.2) is equivalent to the dual minimization problem

$$\min \left\{ \int_\Omega \mathcal{H}(\sigma) \, dx : \sigma \in L^q(\Omega, \mathbb{R}^n), \ \operatorname{div} \sigma = f, \ \sigma \cdot \nu_{\partial\Omega} = 0 \right\}, \qquad (1.4)$$

where the integrand

$$\mathcal{H}(\sigma) = H(|\sigma|), \qquad \text{with } H(t) = t + \tfrac{1}{q} t^q \quad \text{and} \quad \tfrac{1}{p} + \tfrac{1}{q} = 1$$

is the convex conjugate of F, or equivalently $F = \mathcal{H}^*$, and σ represents the traffic flow. The function $g(t) = H'(t)$ models the congestion effect. Note that $\sigma \mapsto g(\sigma)$ is increasing and $g(0) = 1 > 0$, so that moving in an empty street has nonzero cost. As shown in [4] the unique minimizer $\sigma(x)$ of (1.4) is given by $D_\xi F(Du(x))$. We refer to [2–4, 6, 7, 36] and the references therein for detailed motivations and for the physical meaning of the regularity of minimizers. It would be interesting to investigate if there are applications of the vectorial problem. However, our main motivation to consider the very degenerate system (1.3) was from a mathematical point of view.

In connection with congested traffic dynamic problems the regularity of minimizers, as well as the regularity of weak solutions of the associated autonomous Euler–Lagrange system has been an active field of research in recent years. For instance, in [2, 4, 9, 24] Lipschitz regularity of minimizers has been established under suitable assumptions on the datum f.

At this point it is worthwhile to observe that, in general, no more than Lipschitz regularity can be expected for solutions of equations or systems as in (1.3). Indeed when $f = 0$, every 1-Lipschitz continuous function solves (1.3). On the other hand, in the scalar case $N = 1$ assuming $f \in L^{n+\sigma}$ for some $\sigma > n$, it was shown by Santambrogio and Vespri [36] for $n = 2$ and Figalli and Colombo [10, 11] for $n \geq 2$, that the composition of an arbitrary continuous function vanishing on the set $\{|\nabla u| \leq 1\}$ with ∇u is continuous.

Our aim in this paper is to investigate C^1-regularity of minimizers in the vectorial case $N \geq 1$. In general, regularity in the vectorial case is much more delicate and minimizers may be irregular although the integrand is smooth, cf. [15, 38]. In this respect, regularity can be expected only for integrands with special structure. The first result in this direction has been obtained by Uhlenbeck [40] for the p-Laplace system when $p \geq 2$. She proved that weak solutions are of class $C^{1,\alpha}$. The scalar case had previously been established by Ural'ceva [41], while the case $p \in (1, 2)$ was obtained by Tolksdorf [39]. As already mentioned, we cannot hope for a $C^{1,\alpha}$-regularity result for the elliptic system (1.3), since Lipschitz continuity is optimal. However, we are able to establish in the vectorial setting that the composition $\mathcal{K}(Du)$ is continuous for any continuous function $\mathcal{K} \colon \mathbb{R}^{Nn} \to \mathbb{R}$ vanishing on $\{\xi \in \mathbb{R}^{Nn} : |\xi| \leq 1\}$. This phenomenon is somewhat reminiscent of comparable results for the Stefan problem, in which the continuity of the energy cannot been shown, but the temperature shows a logarithmic type continuity, cf. [16, 31].

1.1 Statement of the main result

Before formulating the main regularity result, we need to introduce a few notations. The natural energy space to deal with (local) minimizers of the integral functional $\mathcal{F}$ is the Sobolev space $W^{1,p}(\Omega, \mathbb{R}^N)$. Then, (local) minimizers in $W^{1,p}_{\text{loc}}(\Omega, \mathbb{R}^N)$ of the functional $\mathcal{F}$ are weak solutions of the Euler–Lagrange system

$$\operatorname{div} \mathbf{A}(Du) = f \tag{1.5}$$

and vice versa, where

$$\mathbf{A}(\xi) := h(|\xi|)\xi, \quad \text{with } h(t) := \frac{(t-1)^{p-1}_+}{t} \quad \text{for } t \in \mathbb{R}_+,$$

for some $p > 1$. A function $u \in W^{1,p}_{\text{loc}}(\Omega, \mathbb{R}^N)$ is a *weak solution* of the Euler–Lagrange system (1.5) if and only if

$$\int_\Omega \mathbf{A}(Du) \cdot D\varphi \, dx = -\int_\Omega f \cdot \varphi \, dx$$

holds true for any testing function $\varphi \in C^\infty_0(\Omega, \mathbb{R}^N)$. Our main result proves the continuity of the composition $\mathcal{K}(Du)$ explained above.

Theorem 1.1 *Let $p > 1$, $f \in L^{n+\sigma}(\Omega, \mathbb{R}^N)$ for some $\sigma > 0$ and $u \in W^{1,p}_{\text{loc}}(\Omega, \mathbb{R}^N)$ be a weak solution of (1.5) in Ω. Then,*

$$\mathcal{K}(Du) \in C^0(\Omega)$$

for any continuous function $\mathcal{K}: \mathbb{R}^{Nn} \to \mathbb{R}$ vanishing on $\{\xi \in \mathbb{R}^{Nn} : |\xi| \leq 1\}$.

By carefully tracing the dependence of constants on the parameter δ in the proof of Theorem 3.6, one could determine an explicit modulus of continuity of $\mathcal{G}(Du)$, where $\mathcal{G}$ is defined in (2.2). However, it is not clear if $\mathcal{G}(Du)$ is Hölder continuous in general. For a different very degenerate elliptic equation a counterexample to Hölder continuity is provided in [11].

Theorem 1.1 can be regarded as the vectorial analog of the regularity results of Santambrogio and Vespri [36, Theorem 11] and of Figalli and Colombo [11, Theorem 1.1] as far as the model type integral functional is considered. The vectorial case cannot be treated with the methods from [11, 36], since these are tailored to the scalar case. Nevertheless, some steps in our proof are similar, for example, the approximation procedure by some sequence of uniformly elliptic problems. The main difference in our proof is that we are establishing a variant of DiBenedetto's and Manfredi's proofs of $C^{1,\alpha}$-regularity of minimizers to p-energy type functionals [17, 32]. It is also inspired by the arguments from DiBenedetto and Friedman's pioneering results on $C^{1,\alpha}$-regularity for parabolic p-Laplacian systems [18, 20]. Roughly speaking, our strategy is the adaptation of De Giorgi's approach to the level of gradients in

combination with Campanato type comparison arguments. The past has shown that De Giorgi's approach is extremely flexible. Therefore, we expect that our approach can be transferred to larger classes of widely degenerate functionals in the vectorial case. However, and in order to keep the individual steps as simple as possible, we limit ourselves to treating the model case.

1.2 Strategy of the proof

Concerning the overall strategy of proof a few words are in order. First, we observe that weak solutions of (1.5) are Lipschitz continuous. This has been proved in [2, 4, 9]. Moreover, functionals as in (1.1) fit into the broader context of *asymptotically convex* functionals, i.e. functionals having a *p*-Laplacian type structure only at infinity. This class of functionals has been widely studied, since the local Lipschitz regularity result by Chipot and Evans [8]. In particular we mention generalizations allowing super- and sub-quadratic growth [28, 30, 35], lower order terms [34]. Extensions to various other settings can be found in the non-complete list [12–14, 23–27, 37].

The proof of Theorem 1.1 is divided into several steps and starts by an approximation procedure. Indeed, by replacing $h(t)$ by $h_\varepsilon(t) := h(t) + \varepsilon$ for $\varepsilon > 0$ and considering instead of (1.5) the Dirichlet-problem on a ball compactly contained in Ω associated to the regularized coefficients h_ε and with Dirchlet boundary datum u we obtain a sequence of approximating more regular mappings u_ε. In particular, u_ε has second weak derivatives in L^2_{loc}. In Sect. 3.1 we summarize the most important properties, i.e. uniform energy bounds, uniform quantitative interior L^∞-gradient bounds, uniform quantitative higher differentiability $W^{2,2}$-estimates, and finally strong L^p-convergence of $\mathcal{G}_\delta(Du_\varepsilon) \to \mathcal{G}_\delta(Du)$ in the limit $\delta \downarrow 0$. The nonlinear mapping $\mathcal{G}_\delta \colon \mathbb{R}^{Nn} \to \mathbb{R}^{Nn}$ with $\delta \in (0, 1]$ is defined by

$$\mathcal{G}_\delta(\xi) := \frac{(|\xi| - 1 - \delta)_+}{|\xi|} \, \xi, \quad \text{for } \xi \in \mathbb{R}^{Nn}.$$

Observe that $\mathcal{G}_\delta$ vanishes on the larger set $\{|\xi| \leq 1 + \delta\}$. The reason for considering $\mathcal{G}_\delta$ is that on the complement of $\{|\xi| > 1 + \delta\}$ the system (1.5) behaves non-degenerate in the sense that the vector field **A** admits a uniform ellipticity bound from below, of course, with constants depending on δ. This point of view has already been exploited in [11, 36]. As a first main result we prove that $\mathcal{G}_\delta(Du_\varepsilon)$ is Hölder continuous uniformly with respect to ε. However, the constants in the quantitative estimate, i.e. the Hölder exponent and the Hölder norm, may blow up when $\delta \downarrow 0$. We distinguish between two different regimes: the *degenerate* and *non-degenerate regime*. The degenerate regime is characterized by the fact that the measure of those points in a ball in which $|\mathcal{G}_\delta(Du_\varepsilon)|$ is far from its supremum is large, while the non-degenerate regime is characterized by the opposite. In the non-degenerate regime we compare u_ε with a solution of a linearized system. This allows us to derive a quantitative L^2-excess-improvement for $\mathcal{G}_{2\delta}(Du_\varepsilon)$ on some smaller ball (see Proposition 3.4). This step utilizes a suitable comparison estimate and the higher integrability of u_ε. On the smaller ball we are again in the non-degenerate regime, so that the argument can be iterated yielding a Campanato-type estimate for the L^2-excess of $\mathcal{G}_{2\delta}(Du_\varepsilon)$. In the degenerate regime we

establish that $U_\varepsilon := (|Du_\varepsilon| - 1 - \delta)_+^2$ is a subsolution to a linear uniformly elliptic equation with measurable coefficients; of course the ellipticity constants depend on δ and blow up as $\delta \downarrow 0$. At this stage a De Giorgi type argument allows a reduction of the modulus of $\mathcal{G}_\delta(Du_\varepsilon)$ on some smaller ball (see Proposition 3.5). However, on this smaller scale it is not clear whether or not we are in the degenerate or non-degenerate regime. Therefore one needs to distinguish between these two regimes again. In the non-degenerate regime we can conclude as above, while in the degenerate regime the reduction of the modulus of $\mathcal{G}_\delta(Du_\varepsilon)$ applies again. This argument can be iterated as long as we stay in the degenerate regime. However, if at a certain scale the switching from degenerate to non-degenerate occurs, the above Campanato type decay applies. If no switching occurs, we have at any scale of the iteration process a reduction of the modulus of $\mathcal{G}_\delta(Du_\varepsilon)$. This, however, shows that the supremum of $|\mathcal{G}_\delta(Du_\varepsilon)|$—and hence also the one of $|\mathcal{G}_{2\delta}(Du_\varepsilon)|$—on shrinking concentric balls converges to 0. Altogether this leads to a quantitative Hölder estimate for $\mathcal{G}_{2\delta}(Du_\varepsilon)$ which remains stable under the already established convergence $u_\varepsilon \to u$ as $\varepsilon \downarrow 0$. The final step consits in passing to the limit $\delta \downarrow 0$ and conclude that $\mathcal{G}(Du) := \frac{(|Du|-1)_+}{|Du|} Du$ is continuous. This can be achieved by an application of Ascoli–Arzela's theorem. It is here where we loose control on the quantitative Hölder exponent. At this point the continuity of $\mathcal{K}(Du)$ for any continuous function $\mathcal{K}$ vanishing on $\{|\xi| \leq 1\}$ is an immediate consequence.

2 Notation and preliminary results

2.1 Notation

For the open ball of radius $\varrho > 0$ and center $x_o \in \mathbb{R}^n$ we write $B_\varrho(x_o) \subset \mathbb{R}^n$. The mean value of a function $v \in L^1(B_\varrho(x_o), \mathbb{R}^k)$ is defined by

$$(v)_{x_o,\varrho} := \fint_{B_\varrho(x_o)} v(x)\, dx.$$

If the center is clear from the context we omit the reference to the center and write B_ϱ respectively $(v)_\varrho$ for short. For the standard scalar product on Euclidean spaces $\mathbb{R}^k$ as well as the space $\mathbb{R}^{kn}$ of $k \times n$ matrices, we use the notation $\xi \cdot \eta$. Finally, we use the notion ∇u for the gradient of a scalar function u, while we use Du for a vector field u.

Throughout this paper we abbreviate

$$g(t) := \frac{(t-1)_+^p}{t} \qquad \text{for } t \in \mathbb{R}_+ \tag{2.1}$$

and

$$\mathcal{G}(\xi) := \frac{(|\xi| - 1)_+}{|\xi|} \xi \qquad \text{for } \xi \in \mathbb{R}^k,\ k \in \mathbb{N}. \tag{2.2}$$

Observe that

$$g(|\xi|)\xi = |\mathcal{G}(\xi)|^{p-1}\mathcal{G}(\xi) \quad \text{for any } \xi \in \mathbb{R}^k.$$

Moreover, for $\delta \in (0, 1]$ we define

$$\mathcal{G}_\delta(\xi) := \frac{(|\xi| - 1 - \delta)_+}{|\xi|}\xi \quad \text{for } \xi \in \mathbb{R}^k, \ k \in \mathbb{N} \tag{2.3}$$

and note that $\mathcal{G}_0 \equiv \mathcal{G}$.

Generic constants are denoted by c. They may vary from line to line. Relevant dependencies on parameters and special constants will be suitably emphasized using parentheses or subscripts.

2.2 Algebraic inequalities

In this section, we summarize the relevant algebraic inequalities that will be needed later on. The first lemma follows from an elementary computation.

Lemma 2.1 *For $\eta, \zeta \in \mathbb{R}^k_{\neq 0}$, $k \in \mathbb{N}$ we have*

$$\left| \frac{\eta}{|\eta|} - \frac{\zeta}{|\zeta|} \right| \leq \frac{2}{|\eta|}|\eta - \zeta|.$$

The next lemma can be deduced as in [29, Lemma 8.3].

Lemma 2.2 *For any $\alpha > 0$, there exists a constant $c = c(\alpha)$ such that, for all $\eta, \zeta \in \mathbb{R}^k_{\neq 0}$, $k \in \mathbb{N}$, we have*

$$\frac{1}{c}\left| |\eta|^{\alpha-1}\eta - |\zeta|^{\alpha-1}\zeta \right| \leq \left(|\eta| + |\zeta| \right)^{\alpha-1}|\eta - \zeta| \leq c\left| |\eta|^{\alpha-1}\eta - |\zeta|^{\alpha-1}\zeta \right|.$$

Lemma 2.3 *Let $\delta \geq 0$ and $\eta, \zeta \in \mathbb{R}^k$, $k \in \mathbb{N}$. Then, for $\mathcal{G}_\delta$ as defined in (2.3) we have*

$$|\mathcal{G}_\delta(\eta) - \mathcal{G}_\delta(\zeta)| \leq 3|\eta - \zeta|.$$

Moreover, if $\delta > 0$ and $|\eta| \geq 1 + \delta$ there holds

$$|\eta - \zeta| \leq \left(1 + \tfrac{2}{\delta} \right) |\mathcal{G}(\eta) - \mathcal{G}(\zeta)|.$$

Proof We distinguish between different cases. If $|\eta|, |\zeta| \leq 1 + \delta$ the inequality holds trivially since $\mathcal{G}_\delta(\eta) = 0 = \mathcal{G}_\delta(\zeta)$. If $|\eta|, |\zeta| > 1 + \delta$, we apply Lemma 2.1 and obtain

$$
\begin{aligned}
|\mathcal{G}_\delta(\eta) - \mathcal{G}_\delta(\zeta)| &= \left| \frac{|\eta| - 1 - \delta}{|\eta|} \eta - \frac{|\zeta| - 1 - \delta}{|\zeta|} \zeta \right| \\
&\leq |\eta - \zeta| + (1 + \delta) \left| \frac{\eta}{|\eta|} - \frac{\zeta}{|\zeta|} \right| \\
&\leq 3 |\eta - \zeta|.
\end{aligned}
$$

If $|\eta| > 1 + \delta$ and $|\zeta| \leq 1 + \delta$, we have

$$
|\mathcal{G}_\delta(\eta) - \mathcal{G}_\delta(\zeta)| = |\mathcal{G}_\delta(\eta)| = |\eta| - 1 - \delta \leq |\eta| - |\zeta| \leq |\eta - \zeta|.
$$

The case when $|\eta| \leq 1 + \delta$ and $|\zeta| > 1 + \delta$ is similar, we just have to interchange the role of η and ζ. Joining the three cases gives the first assertion of the Lemma.

Now, we come to the proof of the second assertion. First, we consider the case $|\zeta| \leq 1$ in which $\mathcal{G}(\zeta) = 0$. In this case we have

$$
\frac{|\eta - \zeta|}{|\mathcal{G}(\eta) - \mathcal{G}(\zeta)|} = \frac{|\eta - \zeta|}{|\mathcal{G}(\eta)|} = \frac{|\eta - \zeta|}{|\eta| - 1} \leq \frac{|\eta| + 1}{|\eta| - 1} = 1 + \frac{2}{|\eta| - 1} \leq 1 + \frac{2}{\delta}.
$$

Next, we consider the case $|\zeta| > 1$. Recall that by assumption $|\eta| \geq 1 + \delta$. We start with the observation that $\mathcal{G} \colon \mathbb{R}^k \setminus \{|\eta| \leq 1\} \to \mathbb{R}^k \setminus \{0\}$ is a one to one mapping whose inverse mapping is given by $\mathcal{G}^{-1}(\tilde\xi) = \frac{|\tilde\xi| + 1}{|\tilde\xi|} \tilde\xi$. We let $\tilde\eta := \mathcal{G}(\eta)$ and $\tilde\zeta := \mathcal{G}(\zeta)$ and estimate with Lemma 2.1

$$
\frac{|\mathcal{G}^{-1}(\tilde\eta) - \mathcal{G}^{-1}(\tilde\zeta)|}{|\tilde\eta - \tilde\zeta|} = \frac{\left| \frac{|\tilde\eta| + 1}{|\tilde\eta|} \tilde\eta - \frac{|\tilde\zeta| + 1}{|\tilde\zeta|} \tilde\zeta \right|}{|\tilde\eta - \tilde\zeta|} = \frac{\left| \tilde\eta + \frac{\tilde\eta}{|\tilde\eta|} - \tilde\zeta - \frac{\tilde\zeta}{|\tilde\zeta|} \right|}{|\tilde\eta - \tilde\zeta|} \leq 1 + \frac{2}{|\tilde\eta|} \leq 1 + \frac{2}{\delta}.
$$

In the second to last estimate we used $|\tilde\eta| > \delta$. Using in the previous inequality the definition of $\tilde\eta$ and $\tilde\zeta$ the claim immediately follows. □

Lemma 2.4 *There exists a constant $c = c(p)$ such that for any $a > 1$ and $b \geq 0$ we have*

$$
|h(b) - h(a)| b \leq c(p) \frac{[a - 1 + (b - 1)_+]^{p-1}}{a - 1} |b - a|.
$$

Proof We apply Lemma 2.2 with $\alpha = p - 1 > 0$ to obtain

$$
\begin{aligned}
|h(a) - h(b)|b &= \left| h(a)(b - a) + (a - 1)^{p-1} - (b - 1)_+^{p-1} \right| \\
&\leq h(a)|b - a| + \left| (a - 1)^{p-1} - (b - 1)_+^{p-1} \right| \\
&\leq h(a)|b - a| + c \left[(a - 1) + (b - 1)_+ \right]^{p-2} |b - a| \\
&\leq c(p) \frac{[a - 1 + (b - 1)_+]^{p-1}}{a - 1} |b - a|.
\end{aligned}
$$

This proves the claim. □

Lemma 2.5 *For $a > 1$ we have*

$$
|h'(a)| \leq \frac{p(a - 1)^{p-2}}{a}.
$$

Moreover, for $a, b > 1$ there holds

$$
\left| h'(b)b - h'(a)a \right| \leq c(p) \left[(a - 1)^{p-3} + (b - 1)^{p-3} \right] |b - a|.
$$

Proof By direct computation we have for $a > 1$ that

$$
h'(a) = \frac{(p - 1)(a - 1)^{p-2}a - (a - 1)^{p-1}}{a^2} = \frac{(a - 1)^{p-2}[p - 2 + \frac{1}{a}]}{a}, \quad (2.4)
$$

from which the first claim immediately follows. We now turn our attention to the second claim. We may assume that $1 < a < b$; otherwise we interchange the role of a and b. In view of (2.4) we find

$$
|h'(b)b - h'(a)a| \leq (b - 1)^{p-2} \left| \tfrac{1}{a} - \tfrac{1}{b} \right| + \left| (p - 2) + \tfrac{1}{a} \right| \left| (b - 1)^{p-2} - (a - 1)^{p-2} \right|.
$$

For the first term we have

$$
(b - 1)^{p-2} \left| \tfrac{1}{a} - \tfrac{1}{b} \right| = \frac{(b - 1)^{p-2}}{ab} |a - b| \leq (b - 1)^{p-3} |a - b|.
$$

For the second term we estimate

$$
\begin{aligned}
&\left| (p - 2) + \tfrac{1}{a} \right| \left| (b - 1)^{p-2} - (a - 1)^{p-2} \right| \\
&\leq p \left| (b - 1)^{p-2} - (a - 1)^{p-2} \right| \\
&\leq p|p - 2| \max_{t \in [a,b]} (t - 1)^{p-3} |b - a| \\
&\leq p|p - 2| \left[(a - 1)^{p-3} + (b - 1)^{p-3} \right] |b - a|.
\end{aligned}
$$

This completes the proof of the lemma. □

Lemma 2.6 *For any $t \in \mathbb{R}_+$ we have*

$$g(t)^2 \leq h(t)(t-1)_+^p$$

and

$$g(t)^2 + g'(t)^2 t^2 \leq \frac{p^2}{p-1}\big[h(t) + h'(t)t\big](t-1)_+^p.$$

Proof The first assertion can be achieved, since for $t \in \mathbb{R}_+$ we have

$$g(t)^2 = \frac{(t-1)_+^{2p}}{t^2} \leq \frac{(t-1)_+^{2p-1}}{t} = h(t)(t-1)_+^p.$$

For the second claim, we first observe that both sides are zero for $t \leq 1$. Therefore, it remains to consider $t > 1$. Recalling (2.4) we compute

$$h(t) + h'(t)t = (p-1)(t-1)^{p-2}. \tag{2.5}$$

Moreover, we have

$$g'(t) = \frac{(t-1)^{p-1}[(p-1)t+1]}{t^2},$$

so that

$$\begin{aligned}
g(t)^2 + g'(t)^2 t^2 &= \frac{(t-1)^{2p}}{t^2} + \frac{(t-1)^{2p-2}[(p-1)t+1]^2}{t^2}\\
&= \frac{(t-1)^{2p-2}}{t^2}\Big[(t-1)^2 + \big[(p-1)t+1\big]^2\Big]\\
&\leq \frac{(t-1)^{2p-2}}{t^2}\Big[(t-1) + \big[(p-1)t+1\big]\Big]^2\\
&= p^2(t-1)^{2p-2}.
\end{aligned}$$

Dividing the previous inequality by $h(t) + h'(t)t$ we infer that

$$\frac{g(t)^2 + g'(t)^2 t^2}{h(t) + h'(t)t} \leq \frac{p^2(t-1)^{2p-2}}{(p-1)(t-1)^{p-2}} = \frac{p^2}{p-1}(t-1)^p$$

holds, proving the second claimed inequality. $\qquad\square$

2.3 Bilinear forms

For $\varepsilon \in [0,1]$ we define

$$h_\varepsilon(t) := h(t) + \varepsilon, \qquad \text{for } t \in \mathbb{R}_{\geq 0}.$$

We note that $h_\varepsilon \in C^1(\mathbb{R}_{\geq 0})$ for $p > 2$, while for $p \leq 2$, we have $h_\varepsilon \in W^{1,1}(\mathbb{R}_{>0}) \cap C^1([0, 1) \cap (1, \infty))$. For $p = 2$ we additionally have $h_\varepsilon \in W^{1,\infty}(\mathbb{R}_{\geq 0})$. In any case, $h_\varepsilon \equiv \varepsilon$ on the interval $[0, 1]$. Moreover, we let

$$\mathbf{A}_\varepsilon(\xi) := h_\varepsilon(|\xi|)\xi, \qquad \text{for } \xi \in \mathbb{R}^{Nn}. \tag{2.6}$$

For $\xi \in \mathbb{R}^{Nn} \setminus \{0\}$ with $|\xi| \neq 1$ if $1 < p < 2$ we define the bilinear forms

$$\boldsymbol{\mathcal{A}}_\varepsilon(\xi)(\eta, \zeta) := h_\varepsilon(|\xi|)\eta \cdot \zeta + h_\varepsilon'(|\xi|)|\xi| \sum_{i,j=1}^{N} \sum_{\alpha,\beta,\gamma=1}^{n} \frac{\xi_\alpha^i \eta_{\alpha\gamma}^i \xi_\beta^j \zeta_{\beta\gamma}^j}{|\xi|^2} \quad \text{for } \eta, \zeta \in \mathbb{R}^{Nn^2}$$

and

$$\boldsymbol{\mathcal{B}}_\varepsilon(\xi)(\eta, \zeta) := h_\varepsilon(|\xi|)\eta \cdot \zeta + h_\varepsilon'(|\xi|)|\xi| \sum_{i,j=1}^{N} \sum_{\alpha,\beta=1}^{n} \frac{\xi_\alpha^i \eta_\alpha^i \xi_\beta^j \zeta_\beta^j}{|\xi|^2} \quad \text{for } \eta, \zeta \in \mathbb{R}^{Nn} \tag{2.7}$$

and

$$\boldsymbol{\mathcal{C}}_\varepsilon(\xi)(\eta, \zeta) := h_\varepsilon(|\xi|)\eta \cdot \zeta + h_\varepsilon'(|\xi|)|\xi| \sum_{i=1}^{N_1} \sum_{\alpha,\beta=1}^{n} \frac{\xi_\alpha^i \eta_\alpha \xi_\beta^i \zeta_\beta}{|\xi|^2} \quad \text{for } \eta, \zeta \in \mathbb{R}^n. \tag{2.8}$$

Observe that all forms are symmetric in the arguments η and ζ. Due to the special structure of h_ε and h_ε' the compositions $\boldsymbol{\mathcal{A}}_\varepsilon(Dv)$, $\boldsymbol{\mathcal{B}}_\varepsilon(Dv)$ and $\boldsymbol{\mathcal{C}}_\varepsilon(Dv)$ are well defined for $v \in W^{1,p}$ as integrable functions. Therefore integral calculations involving these quantities make sense.

The next Lemma provides the relevant ellipticity and boundedness properties of the bilinear forms $\boldsymbol{\mathcal{A}}_\varepsilon(\xi)$, $\boldsymbol{\mathcal{B}}_\varepsilon(\xi)$ and $\boldsymbol{\mathcal{C}}_\varepsilon(\xi)$. The following abbreviations

$$\boldsymbol{\lambda}(t) := \min\{h(t), (p-1)(t-1)^{p-2}\} \qquad \text{for } t > 1,$$

and

$$\boldsymbol{\Lambda}(t) := \max\{h(t), (p-1)(t-1)^{p-2}\} \qquad \text{for } t > 1,$$

and $\boldsymbol{\lambda}(t) = 0 = \boldsymbol{\Lambda}(t)$ for $0 \leq t \leq 1$ prove to be useful in the formulation of the Lemma.

Lemma 2.7 *Let $\varepsilon \in [0, 1]$ and $\xi \in \mathbb{R}^{Nn} \setminus \{0\}$. The bilinear form $\boldsymbol{\mathcal{A}}_\varepsilon(\xi)$ defined above satisfies*

$$[\varepsilon + \boldsymbol{\lambda}(|\xi|)]|\zeta|^2 \leq \boldsymbol{\mathcal{A}}_\varepsilon(\xi)(\zeta, \zeta) \leq [\varepsilon + \boldsymbol{\Lambda}(|\xi|)]|\zeta|^2 \quad \text{for any } \zeta \in \mathbb{R}^{Nn^2}.$$

The analogous estimates hold for the bilinear form $\boldsymbol{\mathcal{B}}_\varepsilon$ and any $\eta, \zeta \in \mathbb{R}^{Nn}$, as well as for $\boldsymbol{\mathcal{C}}_\varepsilon$ and any $\eta, \zeta \in \mathbb{R}^n$.

Proof If $|\xi| \leq 1$ the inequality holds trivially. Therefore, it remains to consider the case $|\xi| > 1$. We first establish the lower bound. If $h'_\varepsilon(|\xi|) \geq 0$ we omit the second term in the definition of $\mathcal{A}_\varepsilon$ and obtain

$$\mathcal{A}_\varepsilon(\xi)(\zeta, \zeta) \geq h_\varepsilon(|\xi|)|\zeta|^2 \geq \big[\varepsilon + \boldsymbol{\lambda}(|\xi|)\big]|\zeta|^2,$$

while for $h'_\varepsilon(|\xi|) < 0$ we use Cauchy-Schwarz inequality and (2.5) to conclude

$$
\begin{aligned}
\mathcal{A}_\varepsilon(\xi)(\zeta, \zeta) &\geq \big[\varepsilon + h(|\xi|) + h'(|\xi|)|\xi|\big]|\zeta|^2 \\
&= \big[\varepsilon + (p-1)(|\xi|-1)^{p-2}\big]|\zeta|^2 \geq \big[\varepsilon + \boldsymbol{\lambda}(|\xi|)\big]|\zeta|^2.
\end{aligned}
$$

Now, we turn our attention to the upper bound. If $h'_\varepsilon(|\xi|) \geq 0$ we have

$$
\begin{aligned}
\mathcal{A}_\varepsilon(\xi)(\zeta, \zeta) &\leq \big[h_\varepsilon(|\xi|) + h'_\varepsilon(|\xi|)|\xi|\big]|\zeta|^2 \\
&= \big[\varepsilon + (p-1)(|\xi|-1)^{p-2}\big]|\zeta|^2 \leq \big[\varepsilon + \boldsymbol{\Lambda}(|\xi|)\big]|\zeta|^2,
\end{aligned}
$$

while for $h'_\varepsilon(|\xi|) < 0$ we obtain

$$\mathcal{A}_\varepsilon(\xi)(\zeta, \zeta) \leq h_\varepsilon(|\xi|)|\zeta|^2 \leq \big[\varepsilon + \boldsymbol{\Lambda}(|\xi|)\big]|\zeta|^2.$$

This proves the claim for the bilinear form $\mathcal{A}_\varepsilon$. The corresponding estimates for $\mathcal{B}_\varepsilon$ and $\mathcal{C}_\varepsilon$ follow in the same way. $\qquad\square$

It should also be mentioned that the coercive symmetric bilinear forms fulfill Cauchy-Schwarz inequality. In particular, we have

$$\big|\mathcal{C}_\varepsilon(\xi)(\eta, \zeta)\big| \leq \sqrt{\mathcal{C}_\varepsilon(\xi)(\eta, \eta)}\sqrt{\mathcal{C}_\varepsilon(\xi)(\zeta, \zeta)} \qquad \text{for any } \eta, \zeta \in \mathbb{R}^n.$$

In the next Lemma we put together the monotonicity and growth properties of the vector field $\mathbf{A}_\varepsilon$.

Lemma 2.8 *Let $\varepsilon \in [0, 1]$ and $\xi, \tilde{\xi} \in \mathbb{R}^k$ with $|\xi| > 1$. Then, we have*

$$\big|\mathbf{A}_\varepsilon(\tilde{\xi}) - \mathbf{A}_\varepsilon(\xi)\big| \leq c(p)\left[\varepsilon + \frac{[(|\xi|-1) + (|\tilde{\xi}|-1)_+]^{p-1}}{|\xi|-1}\right]|\tilde{\xi} - \xi|$$

and

$$\big(\mathbf{A}_\varepsilon(\tilde{\xi}) - \mathbf{A}_\varepsilon(\xi)\big) \cdot (\tilde{\xi} - \xi) \geq \left[\varepsilon + \frac{\min\{1, p-1\}}{2^{p+1}} \frac{(|\xi|-1)^p}{|\xi|(|\xi|+|\tilde{\xi}|)}\right]|\tilde{\xi} - \xi|^2.$$

Proof The first inequality results from the following chain of inequalities

$$|\mathbf{A}_\varepsilon(\tilde{\xi}) - \mathbf{A}_\varepsilon(\xi)| \le h_\varepsilon(|\xi|)|\tilde{\xi} - \xi| + \left|h(|\tilde{\xi}|) - h(|\xi|)\right||\tilde{\xi}|$$

$$\le h_\varepsilon(|\xi|)|\tilde{\xi} - \xi| + c\, \frac{[(|\xi| - 1) + (|\tilde{\xi}| - 1)_+]^{p-1}}{|\xi| - 1} \left||\tilde{\xi}| - |\xi|\right|$$

$$\le c(p)\left[\varepsilon + \frac{[(|\xi| - 1) + (|\tilde{\xi}| - 1)_+]^{p-1}}{|\xi| - 1}\right]|\tilde{\xi} - \xi|.$$

From the third last line to the second last we used Lemma 2.4. For the proof of the second inequality we abbreviate

$$\xi_s := \xi + s(\tilde{\xi} - \xi), \quad \text{for } s \in [0, 1].$$

Keeping this in mind we compute

$$\begin{aligned}
&\left(\mathbf{A}_\varepsilon(\tilde{\xi}) - \mathbf{A}_\varepsilon(\xi)\right) \cdot (\tilde{\xi} - \xi) \\
&= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s} \mathbf{A}_\varepsilon(\xi_s)\, \mathrm{d}s \cdot (\tilde{\xi} - \xi) \\
&= \int_0^1 \left[h_\varepsilon(|\xi_s|)(\tilde{\xi} - \xi) + \frac{h'_\varepsilon(|\xi_s|)}{|\xi_s|} \xi_s \cdot (\tilde{\xi} - \xi)\xi_s\right] \mathrm{d}s \cdot (\tilde{\xi} - \xi) \\
&= \int_0^1 \mathcal{B}_\varepsilon(\xi_s)(\tilde{\xi} - \xi, \tilde{\xi} - \xi)\, \mathrm{d}s \\
&\ge \int_0^1 \left[\varepsilon + \lambda(|\xi_s|)\right] \mathrm{d}s\, |\tilde{\xi} - \xi|^2 \\
&\ge \left[\varepsilon + \frac{\min\{1, p-1\}}{|\xi| + |\tilde{\xi}|} \int_0^1 (|\xi_s| - 1)_+^{p-1}\, \mathrm{d}s\right]|\tilde{\xi} - \xi|^2.
\end{aligned}$$

In turn we used Lemma 2.7 and the elementary inequality $(|\xi_s| - 1)_+ \le |\xi_s| \le |\xi| + |\tilde{\xi}|$. Now, we distinguish whether or not $|\tilde{\xi}| \le |\xi|$. If $|\tilde{\xi}| \le |\xi|$, then

$$|\xi_s| \ge (1 - s)|\xi| - s|\tilde{\xi}| \ge (1 - 2s)|\xi| > 1 \quad \forall s \in \left[0, \tfrac{|\xi|-1}{2|\xi|}\right).$$

For $s \in \left[0, \tfrac{|\xi|-1}{4|\xi|}\right]$ this implies a bound from below in the form

$$(|\xi_s| - 1)_+ = |\xi_s| - 1 \ge (1 - 2s)|\xi| - 1 \ge \left[1 - \frac{|\xi| - 1}{2|\xi|}\right]|\xi| - 1 = \tfrac{1}{2}(|\xi| - 1).$$

So it follows that

$$\int_0^1 (|\xi_s| - 1)_+^{p-1}\, \mathrm{d}s \ge \int_0^{\frac{|\xi|-1}{4|\xi|}} (|\xi_s| - 1)_+^{p-1}\, \mathrm{d}s = \frac{1}{2^{p+1}} \frac{(|\xi| - 1)^p}{|\xi|}$$

holds true. In the case that $|\tilde{\xi}| > |\xi|$, we estimate $|\xi_s|$ from below by

$$|\xi_s| \geq s|\tilde{\xi}| - (1-s)|\xi| > (2s-1)|\xi| > 1 \qquad \forall\, s \in \left(\tfrac{|\xi|+1}{2|\xi|}, 1\right].$$

Therefore, for $s \in \left[\tfrac{3|\xi|+1}{4|\xi|}, 1\right]$ we obtain

$$(|\xi_s| - 1)_+ = |\xi_s| - 1 \geq (2s-1)|\xi| - 1 \geq \left[\frac{3|\xi|+1}{2|\xi|} - 1\right]|\xi| - 1 = \tfrac{1}{2}(|\xi| - 1).$$

This yields

$$\int_0^1 (|\xi_s| - 1)_+^{p-1} \, ds \geq \int_{\frac{3|\xi|+1}{4|\xi|}}^1 (|\xi_s| - 1)_+^{p-1} \, ds \geq \frac{1}{2^{p+1}} \frac{(|\xi| - 1)^p}{|\xi|}.$$

Inserting this above, we obtain the second claim of the Lemma. □

Lemma 2.9 *Let $\varepsilon, \delta \in (0, 1]$ and $\xi, \tilde{\xi} \in \mathbb{R}^{Nn}$. Then, we have*

$$\varepsilon^{\frac{1}{2}}|\xi - \tilde{\xi}|^2 + \left|\mathcal{G}_\delta(\xi) - \mathcal{G}_\delta(\tilde{\xi})\right|^p \leq \varepsilon^{\frac{1}{2}}|\xi|^2 + c\varepsilon^{-\frac{1}{2}}\left(\mathbf{A}_\varepsilon(\tilde{\xi}) - \mathbf{A}_\varepsilon(\xi)\right) \cdot (\tilde{\xi} - \xi)$$

for a constant $c = c(p, \delta)$.

Proof If $|\xi|, |\tilde{\xi}| \leq 1+\delta$, we have $\mathcal{G}_\delta(\xi) = 0 = \mathcal{G}_\delta(\tilde{\xi})$. Therefore, the desired inequality follows from the second inequality in Lemma 2.8 after omitting the positive second term in the bracket on the right-hand side.

Therefore, it remains to consider the case where either $|\xi| > 1+\delta$ or $|\tilde{\xi}| > 1+\delta$. We again distinguish two cases and start with $|\xi| \geq |\tilde{\xi}|$. Note that this implies $|\xi| > 1+\delta$. If $p \geq 2$ we use Lemma 2.3 to conclude

$$\left|\mathcal{G}_\delta(\tilde{\xi}) - \mathcal{G}_\delta(\xi)\right|^p \leq 3^p|\tilde{\xi} - \xi|^p \leq 2^{p-2}3^p|\xi|^{p-2}|\tilde{\xi} - \xi|^2,$$

while in the case $p < 2$ we use Young's inequality and Lemma 2.3 to obtain

$$\begin{aligned}
\left|\mathcal{G}_\delta(\tilde{\xi}) - \mathcal{G}_\delta(\xi)\right|^p &= |\xi|^{\frac{p(2-p)}{2}}|\xi|^{\frac{p(p-2)}{2}}\left|\mathcal{G}_\delta(\tilde{\xi}) - \mathcal{G}_\delta(\xi)\right|^p \\
&\leq \tfrac{1}{2}\varepsilon^{\frac{1}{2}}|\xi|^p + c\varepsilon^{-\frac{2-p}{2p}}|\xi|^{p-2}\left|\mathcal{G}_\delta(\tilde{\xi}) - \mathcal{G}_\delta(\xi)\right|^2 \\
&\leq \tfrac{1}{2}\varepsilon^{\frac{1}{2}}|\xi|^2 + c\varepsilon^{-\frac{1}{2}}|\xi|^{p-2}|\tilde{\xi} - \xi|^2,
\end{aligned}$$

for a constant $c = c(p)$. Combining both cases, taking into account the elementary inequalities $\frac{1}{|\xi|} \leq \frac{2}{|\xi|+|\tilde{\xi}|}$ and $|\xi| \leq (1 + \frac{1}{\delta})(|\xi| - 1)$, and finally applying Lemma 2.8, we obtain

$$\begin{aligned}
\varepsilon^{\frac{1}{2}}|\tilde{\xi} - \xi|^2 + \left|\mathcal{G}_\delta(\xi) - \mathcal{G}_\delta(\tilde{\xi})\right|^p &\leq \varepsilon^{\frac{1}{2}}|\tilde{\xi} - \xi|^2 + \tfrac{1}{2}\varepsilon^{\frac{1}{2}}|\xi|^2 + \frac{c\,(|\xi| - 1)_+^p}{\varepsilon^{\frac{1}{2}}|\xi|(|\tilde{\xi}| + |\xi|)}\left|\tilde{\xi} - \xi\right|^2 \\
&\leq \tfrac{1}{2}\varepsilon^{\frac{1}{2}}|\xi|^2 + c\,\varepsilon^{-\frac{1}{2}}\left(\mathbf{A}_\varepsilon(\tilde{\xi}) - \mathbf{A}_\varepsilon(\xi)\right) \cdot (\tilde{\xi} - \xi),
\end{aligned}$$

where $c = c(p, \delta)$. This proves the claimed inequality in this case. In the remaning case, i.e. $|\tilde{\xi}| > |\xi|$, we obtain a similar estimate. We only have to replace on the right-hand side $|\xi|^2$ by $|\tilde{\xi}|^2$. Then, we use the estimate $|\tilde{\xi}|^2 \leq 2(|\tilde{\xi} - \xi|^2 + |\xi|^2)$ and absorb $\varepsilon^{\frac{1}{2}}|\tilde{\xi} - \xi|^2$ by Lemma 2.8 into the second term on the right-hand side. In this way, we obtain

$$\varepsilon^{\frac{1}{2}}|\tilde{\xi} - \xi|^2 + \left|\mathcal{G}_\delta(\xi) - \mathcal{G}_\delta(\tilde{\xi})\right|^p \leq \tfrac{1}{2}\varepsilon^{\frac{1}{2}}|\tilde{\xi}|^2 + c\,\varepsilon^{-\frac{1}{2}}\left(\mathbf{A}_\varepsilon(\tilde{\xi}) - \mathbf{A}_\varepsilon(\xi)\right) \cdot (\tilde{\xi} - \xi)$$
$$\leq \varepsilon^{\frac{1}{2}}|\xi|^2 + c\,\varepsilon^{-\frac{1}{2}}\left(\mathbf{A}_\varepsilon(\tilde{\xi}) - \mathbf{A}_\varepsilon(\xi)\right) \cdot (\tilde{\xi} - \xi),$$

proving the claim also in this case. $\qquad\square$

In the following Lemma we quantify the remainder term in the linearization of $\mathbf{A}_\varepsilon$. In the application, it can be assumed that the linearization only takes place in points ξ with $|\xi| > 1$ in a quantifiable way. The precise statement is

Lemma 2.10 *Let $\varepsilon \in [0, 1]$, and $\xi, \tilde{\xi} \in \mathbb{R}^{Nn}$ with $|\xi| \geq 1 + \frac{1}{4}\mu$ and $|\xi|, |\tilde{\xi}| \leq 1 + 2\mu$ for some $\mu > 0$. Then, we have*

$$\left|\mathcal{B}_\varepsilon(\xi)(\tilde{\xi} - \xi, \zeta) - \left(\mathbf{A}_\varepsilon(\tilde{\xi}) - \mathbf{A}_\varepsilon(\xi)\right) \cdot \zeta\right| \leq c(p)\,\mu^{p-3}|\tilde{\xi} - \xi|^2|\zeta|, \qquad \forall \zeta \in \mathbb{R}^{Nn}.$$

Proof We distinguish two cases. We start with the case $|\xi - \tilde{\xi}| \leq \frac{1}{8}\mu$. For $s \in [0, 1]$ we write $\xi_s := \xi + s(\tilde{\xi} - \xi)$. Note that

$$|\xi_s| \geq |\xi| - s|\tilde{\xi} - \xi| \geq 1 + \tfrac{1}{4}\mu - \tfrac{1}{8}\mu = 1 + \tfrac{1}{8}\mu \qquad \forall s \in [0, 1]. \qquad (2.9)$$

Similarly to the computations in the proof of Lemma 2.8 we have

$$\left(\mathbf{A}_\varepsilon(\tilde{\xi}) - \mathbf{A}_\varepsilon(\xi)\right) \cdot \zeta = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\mathbf{A}_\varepsilon\left(\xi + s(\tilde{\xi} - \xi)\right) \cdot \zeta \, \mathrm{d}s$$
$$= \int_0^1 \left[h_\varepsilon(|\xi_s|)(\tilde{\xi} - \xi) + \frac{h_\varepsilon'(|\xi_s|)}{|\xi_s|}\xi_s \cdot (\tilde{\xi} - \xi)\xi_s\right] \cdot \zeta \, \mathrm{d}s$$
$$= \int_0^1 \mathcal{B}_\varepsilon(\xi_s)(\tilde{\xi} - \xi, \zeta) \, \mathrm{d}s.$$

This allows us to re-write

$$\left|\mathcal{B}_\varepsilon(\xi)(\tilde{\xi} - \xi, \zeta) - \left(\mathbf{A}_\varepsilon(\tilde{\xi}) - \mathbf{A}_\varepsilon(\xi)\right) \cdot \zeta\right| = \left|\int_0^1 \left[\mathcal{B}_\varepsilon(\xi) - \mathcal{B}_\varepsilon(\xi_s)\right](\tilde{\xi} - \xi, \zeta) \, \mathrm{d}s\right|.$$

We decompose and estimate the integrand appearing on the right-hand side and obtain in this way

$$\left|\left[\mathcal{B}_\varepsilon(\xi) - \mathcal{B}_\varepsilon(\xi_s)\right](\tilde{\xi} - \xi, \zeta)\right| \leq \mathbf{I} + \mathbf{II},$$

where

$$\mathbf{I} := \left| h(|\xi|) - h(|\xi_s|) \right| |\tilde{\xi} - \xi| |\zeta|$$

$$\mathbf{II} := \left| \left(h'(|\xi|)|\xi| \frac{\xi_\alpha^i \xi_\beta^j}{|\xi|^2} - h'(|\xi_s|)|\xi_s| \frac{(\xi_s)_\alpha^i (\xi_s)_\beta^j}{|\xi_s|^2} \right) (\tilde{\xi} - \xi)_\alpha^i \zeta_\beta^j \right|$$

In view of Lemma 2.4 we have

$$\left| h_\varepsilon(|\xi|) - h_\varepsilon(|\xi_s|) \right| \le c(p) \frac{\left[|\xi| - 1 + (|\xi_s| - 1)_+ \right]^{p-1}}{|\xi_s|(|\xi| - 1)} |\xi - \xi_s|$$
$$\le c(p) \mu^{p-3} |\xi - \tilde{\xi}|,$$

which immediately implies

$$\mathbf{I} \le c(p) \mu^{p-3} |\xi - \tilde{\xi}|^2 |\zeta|.$$

For the second term we use Lemma 2.1, Lemma 2.5, the assumptions on ξ, $\tilde{\xi}$ and (2.9) to obtain

$$\mathbf{II} \le \left[\sum_{\alpha,\beta=1}^{n} \sum_{i,j=1}^{N} \left(h'(|\xi|)|\xi| \frac{\xi_\alpha^i \xi_\beta^j}{|\xi|^2} - h'(|\xi_s|)|\xi_s| \frac{(\xi_s)_\alpha^i (\xi_s)_\beta^j}{|\xi_s|^2} \right)^2 \right]^{\frac{1}{2}} |\tilde{\xi} - \xi| |\zeta|$$

$$= \left| h'(|\xi|)|\xi| \frac{\xi \otimes \xi}{|\xi|^2} - h'(|\xi_s|)|\xi_s| \frac{\xi_s \otimes \xi_s}{|\xi_s|^2} \right| |\tilde{\xi} - \xi| |\zeta|$$

$$\le \left[|h'(|\xi|)||\xi| \left| \frac{\xi \otimes \xi}{|\xi|^2} - \frac{\xi_s \otimes \xi_s}{|\xi_s|^2} \right| + \left| h'(|\xi|)|\xi| - h'(|\xi_s|)|\xi_s| \right| \right] |\tilde{\xi} - \xi| |\zeta|$$

$$\le \left[2|h'(|\xi|)| \frac{|\xi| + |\xi_s|}{|\xi|} |\xi - \xi_s| + \left| h'(|\xi|)|\xi| - h'(|\xi_s|)|\xi_s| \right| \right] |\tilde{\xi} - \xi| |\zeta|$$

$$\le c(p) \left[(|\xi| - 1)^{p-3} \frac{|\xi| + |\xi_s|}{|\xi|} + (|\xi| - 1)^{p-3} + (|\xi_s| - 1)^{p-3} \right] |\tilde{\xi} - \xi|^2 |\zeta|$$

$$\le c(p) \mu^{p-3} |\tilde{\xi} - \xi|^2 |\zeta|.$$

Inserting the preceding estimates above, we find that

$$\left| \mathcal{B}_\varepsilon(\xi)(\tilde{\xi} - \xi, \zeta) - \left(\mathbf{A}_\varepsilon(\tilde{\xi}) - \mathbf{A}_\varepsilon(\xi) \right) \cdot \zeta \right| \le c(p) \mu^{p-3} |\tilde{\xi} - \xi|^2 |\zeta|.$$

At this stage it remains to consider the case $|\xi - \tilde{\xi}| > \frac{1}{8}\mu$. Note that

$$\mathcal{B}_\varepsilon(\xi)(\tilde{\xi} - \xi, \zeta) = \varepsilon(\tilde{\xi} - \xi) \cdot \zeta + \mathcal{B}_0(\xi)(\tilde{\xi} - \xi, \zeta).$$

Therefore, by Lemma 2.7 and the assumption $\mu < 8|\tilde{\xi} - \xi|$ we obtain

$$
\begin{aligned}
\big|\mathcal{B}_\varepsilon(\xi)(\tilde{\xi} - \xi, \zeta) &- \big(\mathbf{A}_\varepsilon(\tilde{\xi}) - \mathbf{A}_\varepsilon(\xi)\big) \cdot \zeta\big| \\
&= \big|\mathcal{B}_0(\xi)(\tilde{\xi} - \xi, \zeta) - \big(h(|\tilde{\xi}|)\tilde{\xi} - h(|\xi|)\xi\big) \cdot \zeta\big| \\
&\leq \Big[\mathbf{\Lambda}(|\xi|)|\tilde{\xi} - \xi| + (|\tilde{\xi}| - 1)_+^{p-1} + (|\xi| - 1)^{p-1}\Big]|\zeta| \\
&\leq c(p)\big[\mu^{p-2}|\tilde{\xi} - \xi| + \mu^{p-1}\big]|\zeta| \\
&\leq c(p)\,\mu^{p-3}|\tilde{\xi} - \xi|^2|\zeta|.
\end{aligned}
$$

Joining both cases yields the claim. □

Lemma 2.11 *For any $\varepsilon \in [0, 1]$, any ball $B_R \subset \mathbb{R}^n$ and any $v \in W^{2,2}_{\mathrm{loc}}(B_R, \mathbb{R}^N)$ we have*

$$
\big|D\big[g(|Dv|)Dv\big]\big|^2 \leq \frac{2p^2}{p-1}\,\mathcal{A}_\varepsilon(Dv)\big(D^2v, D^2v\big)\,(|Dv| - 1)_+^p \quad a.e.\ in\ B_R,
$$

where g is defined in (2.1).

Proof For $\alpha \in \{1, \ldots, n\}$ we compute

$$
\begin{aligned}
\big|D_\alpha\big[g(|Dv|)Dv\big]\big|^2 &= \big|g(|Dv|)D_\alpha Dv + g'(|Dv|)D_\alpha|Dv|Dv\big|^2 \\
&\leq 2\Big[g(|Dv|)^2\big|D_\alpha Dv\big|^2 + g'(|Dv|)^2|Dv|^2\big|D_\alpha|Dv|\big|^2\Big].
\end{aligned}
$$

Summing with respect to $\alpha \in \{1, \ldots, n\}$, we obtain

$$
\begin{aligned}
\big|D\big[g(|Dv|)Dv\big]\big|^2 &= \sum_{\alpha=1}^n \big|D_\alpha\big[g(|Dv|)Dv\big]\big|^2 \\
&\leq 2\Big[g(|Dv|)^2|D^2v|^2 + g'(|Dv|)^2|Dv|^2\big|\nabla|Dv|\big|^2\Big].
\end{aligned}
$$

In the preceding inequality we replace the term $g(|Dv|)^2$ by $h(|Dv|)(|Dv| - 1)_+^p$, which is possible by an application of the first inequality in Lemma 2.6. Moreover, we would also like to replace $g'(|Dv|)^2|Dv|^2$ by $h'(|Dv|)|Dv|(|Dv| - 1)_+^p$. To this aim we have to distinguish two cases. If $h'(|Dv|) \geq 0$, we use the second inequality in Lemma 2.6 to replace $g'(|Dv|)^2|Dv|^2$ by $[h(|Dv|) + h'(|Dv|)|Dv|](|Dv| - 1)_+^p$. Thereby, we may omit the positive term $g(|Dv|)^2$ on the left-hand side. Inserting this

above and using also Kato's inequality in the form $|\nabla|Dv|| \le |D^2v|$, we obtain

$$
\begin{aligned}
\big|D\big[g(|Dv|)Dv\big]\big|^2 \\
\le \frac{p^2}{p-1}(|Dv|-1)_+^p\Big[h(|Dv|)|D^2v|^2 + \big[h(|Dv|)+h'(|Dv|)|Dv|\big]\big|\nabla|Dv|\big|^2\Big] \\
\le \frac{2p^2}{p-1}(|Dv|-1)_+^p\Big[h(|Dv|)|D^2v|^2 + h'(|Dv|)|Dv|\big|\nabla|Dv|\big|^2\Big] \\
\le \frac{2p^2}{p-1}\,\mathcal{A}_\varepsilon(Dv)\big(D^2v, D^2v\big)\,(|Dv|-1)_+^p.
\end{aligned}
$$

Otherwise, if $h'(|Dv|) < 0$ we use Kato's inequality twice and again the second inequality in Lemma 2.6 to obtain

$$
\begin{aligned}
\big|D\big[g(|Dv|)Dv\big]\big|^2 &\le 2\big[g(|Dv|)^2 + g'(|Dv|)^2|Dv|^2\big]|D^2v|^2 \\
&\le \frac{2p^2}{p-1}(|Dv|-1)_+^p\big[h(|Dv|)+h'(|Dv|)|Dv|\big]|D^2v|^2 \\
&\le \frac{2p^2}{p-1}(|Dv|-1)_+^p\Big[h(|Dv|)|D^2v|^2 + h'(|Dv|)|Dv|\big|\nabla|Dv|\big|^2\Big] \\
&\le \frac{2p^2}{p-1}\,\mathcal{A}_\varepsilon(Dv)\big(D^2v, D^2v\big)\,(|Dv|-1)_+^p.
\end{aligned}
$$

This proves the asserted inequality also in the second case. $\qquad\square$

3 Proof of Theorem 1.1

In this section we will prove Theorem 1.1 under the hypothesis that Propositions 3.4 and 3.5 below are true. The remainder of the paper is then devoted to the proof of those two propositions.

Here and in the following we denote by $u \in W^{1,p}(\Omega, \mathbb{R}^N)$ a weak solution of (1.5). We first observe that $u \in W^{1,\infty}_{\mathrm{loc}}(\Omega, \mathbb{R}^N)$; cf [2, 4] for the scalar case and [9] for the vectorial case. Therefore, we may always assume that Du is locally bounded in Ω.

3.1 Regularization

The first step in the proof consists in the construction of more regular approximating solutions. To this aim we consider a fixed ball $B_R \equiv B_R(y_o) \Subset \Omega$. We let $\varepsilon \in (0, 1]$ and $\mathfrak{p} := \max\{p, 2\}$ and recall the definition of the regularized coefficients $\mathbf{A}_\varepsilon$ from (2.6). By $u_\varepsilon \in u + W^{1,\mathfrak{p}}_0(B_R, \mathbb{R}^N)$ we denote the unique weak solution of the regularized elliptic system

$$
\begin{cases}
\operatorname{div}\mathbf{A}_\varepsilon(Du_\varepsilon) = f, & \text{in } B_R, \\
u_\varepsilon = u, & \text{on } \partial B_R.
\end{cases}
\tag{3.1}
$$

The weak formulation of (3.1) is

$$\int_{B_R} \mathbf{A}_\varepsilon(Du_\varepsilon) \cdot D\varphi \, dx = -\int_{B_R} f \cdot \varphi \, dx \qquad \forall \varphi \in W_0^{1,\mathfrak{p}}(B_R, \mathbb{R}^N). \qquad (3.2)$$

Note that $u_\varepsilon \in W_{\mathrm{loc}}^{1,\infty}(B_R, \mathbb{R}^N) \cap W_{\mathrm{loc}}^{2,2}(B_R, \mathbb{R}^N)$. This can be retrieved similarly as in [33].

Lemma 3.1 *For any $\varepsilon \in (0, 1]$ we have $u_\varepsilon \in W_{\mathrm{loc}}^{1,\infty}(B_R, \mathbb{R}^N) \cap W_{\mathrm{loc}}^{2,2}(B_R, \mathbb{R}^N)$. Moreover, for any ball $B_{2\varrho}(x_o) \Subset B_R$ the (uniform with respect to ε) quantitative L^∞-gradient bound*

$$\sup_{B_\varrho(x_o)} |Du_\varepsilon| \le c \left[\fint_{B_{2\varrho}(x_o)} \left(|Du_\varepsilon|^p + \varepsilon |Du_\varepsilon|^2 + 1 \right) dx \right]^{\frac{1}{p}}$$

with $c = c(n, N, p, \sigma, \|f\|_{L^{n+\sigma}(B_R)})$ and the quantitative $W^{2,2}$-estimate

$$\int_{B_\varrho(x_o)} |D^2 u_\varepsilon|^2 \, dx \le \frac{c}{\varepsilon \varrho^2} \left[\int_{B_{2\varrho}(x_o)} \left(|Du_\varepsilon|^p + \varepsilon |Du_\varepsilon|^2 \right) dx + \frac{\varrho^2}{\varepsilon} \int_{B_{2\varrho}(x_o)} |f|^2 \, dx \right]$$

with $c = c(n, N, p)$ hold true.

Our first observation is a uniform energy bound for Du_ε.

Lemma 3.2 *There exists a constant $c = c(n, p)$ such that for any $\varepsilon \in (0, 1]$ we have*

$$\int_{B_R} \left(|Du_\varepsilon|^p + \varepsilon |Du_\varepsilon|^2 \right) dx \le c \int_{B_R} \left(|Du|^p + \varepsilon |Du|^2 + 1 \right) dx + c \, R^n \|f\|_{L^n(B_R)}^{\frac{p}{p-1}}.$$

Proof The desired estimate can be deduced with a standard argument by testing the weak formulation (3.2) with the test-function $\varphi := u_\varepsilon - u$. Indeed, we have

$$\int_{B_R} \left[(|Du_\varepsilon| - 1)_+^p + \varepsilon |Du_\varepsilon|^2 \right] dx$$

$$\le \int_{B_R} h_\varepsilon(|Du_\varepsilon|) Du_\varepsilon \cdot Du_\varepsilon \, dx$$

$$= \int_{B_R} h_\varepsilon(|Du_\varepsilon|) Du_\varepsilon \cdot Du \, dx - \int_{B_R} f \cdot (u_\varepsilon - u) \, dx.$$

By Young's inequality we obtain for the first integral on the right-hand side

$$\int_{B_R} h_\varepsilon(|Du_\varepsilon|) Du_\varepsilon \cdot Du \, dx \le \int_{B_R} \left[(|Du_\varepsilon| - 1)_+^{p-1} |Du| + \varepsilon |Du_\varepsilon||Du| \right] dx$$

$$\le \tfrac{1}{4} \int_{B_R} (|Du_\varepsilon| - 1)_+^p \, dx + c \int_{B_R} |Du|^p \, dx$$

$$+ \tfrac{\varepsilon}{2} \int_{B_R} \left(|Du_\varepsilon|^2 + |Du|^2 \right) dx.$$

The second integral on the right-hand side is estimated with Hölder's and Sobolev's inequality, so that

$$\left| \int_{B_R} f \cdot (u_\varepsilon - u) \, dx \right|$$

$$\le \left[\int_{B_R} |f|^n \, dx \right]^{\frac{1}{n}} \left[\int_{B_R} |u_\varepsilon - u|^{\frac{n}{n-1}} \, dx \right]^{\frac{n-1}{n}}$$

$$\le c \left[\int_{B_R} |f|^n \, dx \right]^{\frac{1}{n}} \int_{B_R} |Du_\varepsilon - Du| \, dx$$

$$\le c \, |B_R|^{\frac{p-1}{p}} \left[\int_{B_R} |f|^n \, dx \right]^{\frac{1}{n}} \left[\int_{B_R} \left[(|Du_\varepsilon| - 1)_+^p + |Du|^p + 1 \right] dx \right]^{\frac{1}{p}}$$

$$\le c \, R^n \|f\|_{L^n(B_R)}^{\frac{p-1}{p}} + \tfrac{1}{4} \int_{B_R} \left[(|Du_\varepsilon| - 1)_+^p + |Du|^p + 1 \right] dx,$$

where $c = c(n, p)$. We insert these inequalities above and reabsorb the terms containing Du_ε from the right-hand side into the left. In this way, we get

$$\int_{B_R} \left[(|Du_\varepsilon| - 1)_+^p + \varepsilon |Du_\varepsilon|^2 \right] dx$$

$$\le c \int_{B_R} \left[|Du|^p + \varepsilon |Du|^2 + 1 \right] dx + c \, R^n \|f\|_{L^n(B_R)}^{\frac{p}{p-1}},$$

with a constant $c = c(n, p)$. The desired uniform energy bound can easily be deduced from the preceding inequality. $\qquad \square$

The next lemma ensures strong convergence of the approximating solutions, in the sense that $\mathcal{G}_\delta(Du_\varepsilon)$ strongly converges to $\mathcal{G}_\delta(Du)$ in L^p.

Lemma 3.3 *Let $\delta \in (0, 1]$ and u_ε with $\varepsilon \in (0, 1]$ be the unique solution of the Dirichlet problem* (3.1). *Then, we have*

$$\mathcal{G}_\delta(Du_\varepsilon) \to \mathcal{G}_\delta(Du) \quad \text{in } L^p\big(B_R, \mathbb{R}^{Nn}\big) \text{ as } \varepsilon \downarrow 0.$$

Proof Testing (3.2) and the weak formulation of (1.5) with $\varphi := u_\varepsilon - u$ we have

$$\int_{B_R} \big(\mathbf{A}_\varepsilon(Du_\varepsilon) - \mathbf{A}_\varepsilon(Du)\big) \cdot (Du_\varepsilon - Du)\, dx = \varepsilon \int_{B_R} Du \cdot (Du - Du_\varepsilon)\, dx.$$

Using Lemma 2.9 and Young's inequality we find

$$\int_{B_R} \Big[\varepsilon^{\frac{1}{2}} |Du_\varepsilon - Du|^2 + \big|\mathcal{G}_\delta(Du_\varepsilon) - \mathcal{G}_\delta(Du)\big|^p\Big]\, dx$$

$$\leq \varepsilon^{\frac{1}{2}} \int_{B_R} |Du|^2\, dx + c\varepsilon^{-\frac{1}{2}} \int_{B_R} \big(\mathbf{A}_\varepsilon(Du_\varepsilon) - \mathbf{A}_\varepsilon(Du)\big) \cdot (Du_\varepsilon - Du)\, dx$$

$$\leq \varepsilon^{\frac{1}{2}} \int_{B_R} |Du|^2\, dx + c\varepsilon^{\frac{1}{2}} \int_{B_R} |Du||Du_\varepsilon - Du|\, dx$$

$$\leq \varepsilon^{\frac{1}{2}} \int_{B_R} |Du_\varepsilon - Du|^2\, dx + c\varepsilon^{\frac{1}{2}} \int_{B_R} |Du|^2\, dx.$$

Re-absorbing the first integral from the right into the left-hand side, we obtain

$$\int_{B_R} \big|\mathcal{G}_\delta(Du_\varepsilon) - \mathcal{G}_\delta(Du)\big|^p\, dx \leq c\varepsilon^{\frac{1}{2}} \int_{B_R} |Du|^2\, dx.$$

The integral on the right-hand side is finite, since $Du \in L^\infty_{\mathrm{loc}}(\Omega, \mathbb{R}^{Nn})$. Therefore, the preceding inequality implies strong convergence $\mathcal{G}_\delta(Du_\varepsilon) \to \mathcal{G}_\delta(Du)$ in $L^p(B_R, \mathbb{R}^{Nn})$ as $\varepsilon \downarrow 0$. $\qquad\square$

3.2 Hölder-continuity of $\mathcal{G}_\delta(Du_\varepsilon)$

We recall the definition of $\mathcal{G}_\delta$ from (2.3). In this subsection we will prove that $\mathcal{G}_\delta(Du_\varepsilon)$ is locally Hölder continuous in B_R for any $\delta \in (0, 1]$. This will be achieved in Theorem 3.6. Thereby, it is essential that all constants are independent of ε.

The proof of Theorem 3.6 relies on the distinction between two different regimes—the *degenerate* and *non-degenerate regime*. In the non-degenerate regime we will prove an excess-decay estimate for $\mathcal{G}_\delta(Du_\varepsilon)$ (see Proposition 3.4), while in the degenerate regime we establish a reduction of the modulus of $\mathcal{G}_\delta(Du_\varepsilon)$ (see Proposition 3.5). The precise setup is as follows. We consider a ball $B_R \Subset \Omega$ and denote by u_ε the unique weak solution of the Cauchy–Dirichlet problem (3.1). For $0 < r_o < R$ we let $r_1 := \frac{1}{2}(R + r_o)$. Then, by Lemma 3.1 and Lemma 3.2 we have (uniform with respect to ε) boundedness of Du_ε on B_{r_1}. More precisely, there exists a constant

$$M = M\big(n, N, p, R - r_o, \|Du\|_{L^p(B_R)}, \ldots\big) \tag{3.3}$$

independent of $\varepsilon > 0$, such that $\|Du_\varepsilon\|_{L^\infty(B_{r_1})} \leq M$. We may assume that $M \geq 3$. Now, we consider a center $x_o \in B_{r_o}$ and a radius $\varrho \leq r_1$ such that $B_{2\varrho}(x_o) \subset B_{r_1}$. On

this ball we have

$$\sup_{B_{2\varrho}(x_o)} |Du_\varepsilon| \le 1 + \delta + \mu \tag{3.4}$$

for some $\mu > 0$ such that

$$1 + \delta + \mu \le M. \tag{3.5}$$

Next, for $\nu \in (0, 1)$ we define the super-level set of $|Du_\varepsilon|$ by

$$E_\varrho^\nu(x_o) := \big\{ x \in B_\varrho(x_o) : |Du_\varepsilon(x)| - 1 - \delta > (1 - \nu)\mu \big\}.$$

The definition of the super-level set allows us to distinguish between the degenerate regime which is characterized by the measure condition $|B_\varrho(x_o) \setminus E_\varrho^\nu(x_o)| \ge \nu |B_\varrho(x_o)|$ and the non-degenerate regime which is characterized by the reversed inequality. Roughly speaking, in the degenerate regime the set of points with $|Du_\varepsilon|$ small has large measure, while in the non-degenerate regime the set of points with $|Du_\varepsilon|$ small is small in measure. We start with the latter one. In the following we abbreviate $\beta := \frac{\sigma}{n+\sigma}$.

Proposition 3.4 *Let $\varepsilon, \delta \in (0, 1]$ and*

$$0 < \delta < \mu. \tag{3.6}$$

Then, there exist $\nu = \nu(n, N, p, \sigma, \|f\|_{L^{n+\sigma}(B_R)}, M, \delta) \in (0, \frac{1}{4}]$ and $\hat{\varrho} = \hat{\varrho}(n, N, p, \sigma, \|f\|_{L^{n+\sigma}(B_R)}, M, \delta) \in (0, 1]$ such that there holds: Whenever $B_{2\varrho}(x_o) \subset B_{r_1}$ is a ball with radius $\varrho \le \hat{\varrho}$ and center $x_o \in B_{r_o}$, and u_ε is the unique weak solution of the Dirichlet problem (3.1) *and hypothesis* (3.4) *and* (3.5) *and the measure condition*

$$\big| B_\varrho(x_o) \setminus E_\varrho^\nu(x_o) \big| < \nu \big| B_\varrho(x_o) \big| \tag{3.7}$$

are satisfied, then the limit

$$\Gamma_{x_o} := \lim_{r \downarrow 0} \big(\mathcal{G}_{2\delta}(Du_\varepsilon) \big)_{x_o, r} \tag{3.8}$$

exists, and the excess decay estimate

$$\fint_{B_r(x_o)} \big| \mathcal{G}_{2\delta}(Du_\varepsilon) - \Gamma_{x_o} \big|^2 \, dx \le \Big(\frac{r}{\varrho} \Big)^{2\beta} \mu^2 \quad \text{for any } 0 < r \le \varrho \tag{3.9}$$

holds true. Moreover, we have

$$|\Gamma_{x_o}| \le \mu.$$

The statement for the degenerate regime is as follows.

Proposition 3.5 *Let* $\varepsilon, \delta \in (0, 1]$, $\mu > 0$ *and* $v \in (0, \frac{1}{4}]$. *Then, there exist constants* $\kappa = \kappa(n, p, \sigma, \|f\|_{n+\sigma}, M, \delta, v) \in [2^{-\beta/2}, 1)$ *and* $c_o = c_o(n, p, \|f\|_{n+\sigma}, M, \delta, v) \geq$ 1 *such that the there holds: Whenever* $B_{2\varrho}(x_o) \subset B_{r_1}$ *is a ball with center* $x_o \in B_{r_o}$, *and* u_ε *is the unique weak solution of the Dirichlet problem* (3.1) *and hypothesis* (3.4) *and* (3.5) *and the measure condition*

$$\left| B_\varrho(x_o) \setminus E_\varrho^v(x_o) \right| \geq v \left| B_\varrho(x_o) \right| \tag{3.10}$$

are satisfied, then, either

$$\mu^2 < c_o \varrho^\beta,$$

or

$$\sup_{B_{\varrho/2}(x_o)} \left| \mathcal{G}_\delta(Du_\varepsilon) \right| \leq \kappa \mu$$

hold true.

We postpone the proofs of Proposition 3.4 and 3.5 to Chapters 4–6 and continue with formulating the main result of this subsection.

Theorem 3.6 *Let* $\varepsilon, \delta \in (0, 1]$ *and* u_ε *be the unique weak solution of the Dirichlet problem* (3.1) *in* B_R. *Then,* $\mathcal{G}_\delta(Du_\varepsilon)$ *is Hölder continuous in* B_{r_o} *for any* $0 < r_o <$ R *with Hölder-exponent* $\alpha_\delta \in (0, 1)$ *and a Hölder constant* c_δ *both depending on* $n, N, p, \sigma, \|f\|_{L^{n+\sigma}(B_R)}, M$ *and* δ.

Proof By

$$v = v(n, N, p, \sigma, \|f\|_{L^{n+\sigma}(B_R)}, M, \delta) \in (0, \tfrac{1}{4}]$$

and

$$\hat{\varrho} = \hat{\varrho}(n, N, p, \sigma, \|f\|_{L^{n+\sigma}(B_R)}, M, \delta) \in (0, 1]$$

we denote the constants from Proposition 3.4 and by

$$\kappa = \kappa(n, p, \sigma, \|f\|_{n+\sigma}, M, \delta, v) \in [2^{-\beta/2}, 1),$$
$$c_o = c_o(n, p, \|f\|_{n+\sigma}, M, \delta, v) \geq 1$$

we denote the ones from Proposition 3.5. The dependence of v on the structural parameters implies that κ depends on $n, N, p, \sigma, \|f\|_{n+\sigma}, M$ and δ. Finally, we let $\mu = M - 1 - \delta$ and

$$\varrho_* = \min \left\{ \hat{\varrho}, \left[\frac{(\kappa\mu)^2}{c_o} \right]^{\frac{1}{\beta}} \right\}.$$

We consider a ball $B_{2\varrho}(x_o) \subset B_{r_1}$ with center $x_o \in B_{r_o}$ and $\varrho \leq \varrho_*$ as described above. On this ball we have (3.3)–(3.5) satisfied. Our aim now is to prove that $\mathcal{G}_{2\delta}(Du_\varepsilon)$ is Hölder continuous in B_{r_o} with Hölder exponent

$$\alpha := -\frac{\log \kappa}{\log 2} \in \left(0, \tfrac{\beta}{2}\right].$$

In turn, substituting 2δ by δ, this proves the claim of the proposition. We proceed in two steps.

Step 1: We prove that the limit

$$\Gamma_{x_o} := \lim_{r \downarrow 0} \left(\mathcal{G}_{2\delta}(Du_\varepsilon)\right)_{x_o,r}$$

exists and that

$$\fint_{B_r(x_o)} \left|\mathcal{G}_{2\delta}(Du_\varepsilon) - \Gamma_{x_o}\right|^2 \mathrm{d}x \leq 8\left(\frac{r}{\varrho}\right)^{2\alpha} \mu^2 \quad \text{for any } 0 < r \leq \varrho \qquad (3.11)$$

holds true for a constant $c = c(n, p, \sigma, \|f\|_{n+\sigma}, M, \delta)$. To this aim, we define for $i \in \mathbb{N}_0$ radii

$$\varrho_i := 2^{-i}\varrho \quad \text{and} \quad \mu_i := \kappa^i \mu$$

and observe that

$$\mu_i = \kappa^i \mu \leq 2^{-\alpha i}\mu = \left(\frac{\varrho_i}{\varrho}\right)^{\alpha}\mu \quad \text{for any } i \in \mathbb{N}. \qquad (3.12)$$

Now, suppose that assumption (3.10) holds on $B_\varrho(x_o)$. Then, Proposition 3.5 yields that either $\mu^2 < c_o\varrho^\beta$ or

$$\sup_{B_{\varrho_1}(x_o)} |\mathcal{G}_\delta(Du_\varepsilon)| \leq \kappa\mu = \mu_1.$$

Note that the first alternative cannot happen, since it would imply

$$\mu^2 < c_o\varrho^\beta \leq c_o\varrho_*^\beta \leq \kappa^2\mu^2.$$

Hence, we conclude that (3.4) holds on $B_{\varrho_1}(x_o)$ with $\mu = \mu_1$. If the measure condition (3.10) is satisfied with $\varrho = \varrho_1$ and $\mu = \mu_1$, then a second application of Proposition 3.5 yields that either $\mu_1^2 < c_o\varrho_1^\beta$ or

$$\sup_{B_{\varrho_2}(x_o)} |\mathcal{G}_\delta(Du_\varepsilon)| \leq \kappa\mu_1 = \mu_2.$$

As before, the first alternative cannot happen, since it would imply

$$\mu_1^2 < c_o \varrho_1^\beta \le 2^{-\beta}(\kappa\mu)^2 \le \kappa^2\mu_1^2.$$

Assume now that (3.10) is satisfied for $i = 1, \ldots, i_o - 1$ up to some $i_o \in \mathbb{N}$, i.e. that (3.10) holds true on the balls $B_{\varrho_i}(x_o)$ with $\mu = \mu_i$. Then, we iteratively conclude that

$$\sup_{B_{\varrho_i}(x_o)} |\mathcal{G}_\delta(Du_\varepsilon)| \le \mu_i, \quad \text{for } i = 0, \ldots, i_o. \tag{3.13}$$

Now assume that (3.10) fails to hold for some $i_o \in \mathbb{N}_0$. If $\mu_{i_o} > \delta$, the hypothesis of Proposition 3.4 are satisfied on $B_{\varrho_{i_o}}(x_o)$ and we conclude that the limit

$$\Gamma_{x_o} := \lim_{r \downarrow 0} \left(\mathcal{G}_{2\delta}(Du_\varepsilon)\right)_{x_o,r}$$

exists and that

$$\fint_{B_r(x_o)} \left|\mathcal{G}_{2\delta}(Du_\varepsilon) - \Gamma_{x_o}\right|^2 dx \le \left(\frac{r}{\varrho_{i_o}}\right)^{2\beta} \mu_{i_o}^2 \quad \text{for any } 0 < r \le \varrho_{i_o}.$$

Moreover, we have

$$|\Gamma_{x_o}| \le \mu_{i_o}. \tag{3.14}$$

Therefore, we obtain from the preceding inequality and (3.12) that

$$\fint_{B_r(x_o)} \left|\mathcal{G}_{2\delta}(Du_\varepsilon) - \Gamma_{x_o}\right|^2 dx \le \left(\frac{\varrho_{i_o}}{\varrho}\right)^{2\alpha}\left(\frac{r}{\varrho_{i_o}}\right)^{2\beta}\mu^2 \le \left(\frac{r}{\varrho}\right)^{2\alpha}\mu^2$$

holds true for any $0 < r \le \varrho_{i_o}$. For a radius $r \in (\varrho_{i_o}, \varrho]$ there exists $i \in \{0, \ldots, i_o\}$ such that $\varrho_{i+1} < r \le \varrho_i$. Using (3.13), (3.14) and (3.12) we obtain

$$\fint_{B_r(x_o)} \left|\mathcal{G}_{2\delta}(Du_\varepsilon) - \Gamma_{x_o}\right|^2 dx \le 2 \sup_{B_{\varrho_i}(x_o)} |\mathcal{G}_\delta(Du_\varepsilon)|^2 + 2|\Gamma_{x_o}|^2$$

$$\le 4\mu_i^2 \le 4\left(\frac{\varrho_i}{\varrho}\right)^{2\alpha}\mu^2 \le 8\left(\frac{r}{\varrho}\right)^{2\alpha}\mu^2.$$

Combining the preceding two inequalities, we have shown (3.11) provided $\mu_{i_o} > \delta$.

In the case $\mu_{i_o} \le \delta$, we have $\mathcal{G}_{2\delta}(Du_\varepsilon) = 0$ on $B_{\varrho_{i_o}}(x_o)$. Combining this with (3.13) and keeping in mind that $\mathcal{G}_{2\delta}(Du_\varepsilon) \le \mathcal{G}_\delta(Du_\varepsilon)$, we obtain

$$\sup_{B_{\varrho_i}(x_o)} \left|\mathcal{G}_{2\delta}(Du_\varepsilon)\right| \le \mu_i, \quad \text{for any } i \in \{0, 1, \ldots\}. \tag{3.15}$$

In the final case when (3.10) holds for any $i \in \mathbb{N}$, then (3.13) is satisfied for any $i \in \mathbb{N}$ and hence we obtain (3.15) also in this case. (3.15), however, implies

$$\Gamma_{x_o} := \lim_{r \downarrow 0} \big(\mathcal{G}_{2\delta}(Du_\varepsilon)\big)_{x_o,r} = 0.$$

For $r \in (0, \varrho]$ we find $i \in \{0, 1, \dots\}$ such that $\varrho_{i+1} < r \leq \varrho_i$. Then, (3.15) and (3.12) imply

$$\fint_{B_r(x_o)} \big|\mathcal{G}_{2\delta}(Du_\varepsilon) - \Gamma_{x_o}\big|^2 \, dx \leq \sup_{B_{\varrho_i}(x_o)} \big|\mathcal{G}_{2\delta}(Du_\varepsilon)\big|^2$$

$$\leq \mu_i^2 \leq \Big(\frac{\varrho_i}{\varrho}\Big)^{2\alpha} \mu^2 \leq 2\Big(\frac{r}{\varrho}\Big)^{2\alpha} \mu^2.$$

This establishes (3.11) in the remaining cases.

Step 2: Here, we prove that the Lebesgue representative $x \mapsto \Gamma_x$ of $\mathcal{G}_{2\delta}(Du_\varepsilon)$ is Hölder continuous in B_{r_o}. The proof is standard once the excess decay (3.11) is established. For convenience of the reader we give the details. We consider $x_1, x_2 \in B_{r_o}$. If $r := |x_1 - x_2| \leq \varrho_*$ we define $\tilde{x} := \frac{1}{2}(x_1 + x_2)$ and obtain from Step 1 that

$$\big|\Gamma_{x_1} - \Gamma_{x_2}\big|^2 = \fint_{B_{r/2}(\tilde{x})} |\Gamma_{x_1} - \Gamma_{x_2}|^2 \, dx$$

$$\leq c(n)\bigg[\fint_{B_r(x_1)} |\mathcal{G}_{2\delta}(Du_\varepsilon) - \Gamma_{x_1}|^2 \, dx + \fint_{B_r(x_2)} |\mathcal{G}_{2\delta}(Du_\varepsilon) - \Gamma_{x_2}|^2 \, dx\bigg]$$

$$\leq c(n)\Big(\frac{|x_1 - x_2|}{\varrho_*}\Big)^{2\alpha} \mu^2.$$

This can be re-written in the form

$$\big|\Gamma_{x_1} - \Gamma_{x_2}\big| \leq c \left(\frac{|x_1 - x_2|}{\varrho_*}\right)^{\alpha} \mu.$$

Otherwise, if $r = |x_1 - x_2| > \varrho_*$, we trivially have

$$\big|\Gamma_{x_1} - \Gamma_{x_2}\big| \leq 2\mu \leq 2\left(\frac{|x_1 - x_2|}{\varrho_*}\right)^{\alpha} \mu.$$

Together, this establishes that the Lebesgue representative $x \mapsto \Gamma_x$ of $\mathcal{G}_{2\delta}(Du_\varepsilon)$ is Hölder continuous in B_{r_o} with Hölder exponent α. Note that α admits the same dependencies as κ, i.e. $\alpha = \alpha(n, N, p, \sigma, \|f\|_{n+\sigma}, M, \delta)$. This finishes the proof of the theorem. $\qquad\square$

3.3 Continuity of $\mathcal{G}(Du)$

In this subsection it is important that all estimates are independent of ε. More precisely, constants might depend on δ, but are independent of ε.

Proof of Theorem 1.1 We let $\varepsilon \in (0, 1]$ and consider a fixed ball $B_R \equiv B_R(y_o) \Subset \Omega$. By u_ε we denote the weak solution to (3.1) constructed in Sect. 3.1. Next, we fix $\delta \in (0, 1]$ and $r \in (0, R)$. From Theorem 3.6 we know that $\mathcal{G}_\delta(Du_\varepsilon)$ is Hölder continuous in $\overline{B}_r$ with Hölder-exponent $\alpha_\delta \in (0, 1)$ and constant c_δ, both depending at most on n, N, p, σ, $\|f\|_{n+\sigma}$, M, and δ. From Lemma 3.3 we know that $\mathcal{G}_\delta(Du_\varepsilon) \to \mathcal{G}_\delta(Du)$ in $L^p(B_R, \mathbb{R}^{Nn})$ as $\varepsilon \downarrow 0$. This implies that there exists a subsequence $\varepsilon_i \downarrow 0$ as $i \to \infty$ such that $\mathcal{G}_\delta(Du_{\varepsilon_i}) \to \mathcal{G}_\delta(Du)$ a.e. in B_R. On the other hand, by Ascoli-Arzelà's Theorem we conclude that $\mathcal{G}_\delta(Du_{\varepsilon_i})$ converges uniformly on compact subsets of B_R. Therefore the limit $\mathcal{G}_\delta(Du)$ is Hölder continuous in $\overline{B}_r$ with Hölder-exponent $\alpha_\delta \in (0, 1)$ and constant c_δ. In particular, $\mathcal{G}_\delta(Du)$ is continuous in $\overline{B}_r$ for any $\delta \in (0, 1]$. Moreover, we have $\mathcal{G}_\delta(Du) \to \mathcal{G}(Du)$ uniformly in $\overline{B}_r$ as $\delta \downarrow 0$. Indeed

$$\left|\mathcal{G}_\delta(Du) - \mathcal{G}(Du)\right| = \left|(|Du| - 1 - \delta)_+ \frac{Du}{|Du|} - (|Du| - 1)_+ \frac{Du}{|Du|}\right|$$
$$= \left|(|Du| - 1 - \delta)_+ - (|Du| - 1)_+\right| \le \delta$$

in $\overline{B}_r$. As the uniform limit of a sequence of continuous functions, $\mathcal{G}(Du)$ itself is continuous on $\overline{B}_r$. Observe that $\mathcal{G}(Du)$ is also uniformly continuous on $\overline{B}_r$.

Now, let $\mathcal{K} \colon \mathbb{R}^{Nn} \to \mathbb{R}$ be any continuous function vanishing on $\{\xi \in \mathbb{R}^{Nn} : |\xi| \le 1\}$. Since $u \in W^{1,\infty}_{\text{loc}}(\Omega, \mathbb{R}^N)$, we find $M > 0$ such that $|Du| \le M$ on $\overline{B}_r$. By $\omega \colon \mathbb{R}_+ \to \mathbb{R}_+$ we denote the modulus of continuity of $\mathcal{K}$ on $\{\xi \in \mathbb{R}^{Nn} : |\xi| \le M\}$, i.e. for any $\xi, \zeta \in \mathbb{R}^{Nn}$ with $|\xi|, |\zeta| \le M$ we have $|\mathcal{K}(\xi) - \mathcal{K}(\zeta)| \le \omega(|\xi - \zeta|)$. Next, given $\varepsilon \in (0, 1)$ we choose $\delta > 0$ such that

$$\left|\mathcal{G}(Du(x)) - \mathcal{G}(Du(y))\right| < \varepsilon \quad \text{for any } x, y \in \overline{B}_r \text{ with } |x - y| < \delta. \quad (3.16)$$

We now distinguish two cases. First, we assume $|Du(x)| \le 1 + \sqrt{\varepsilon}$. If $|Du(x)| \ge 1$ we use $\mathcal{K} = 0$ on $\{|\xi| \le 1\}$ to conclude

$$|\mathcal{K}(Du(x))| = \left|\mathcal{K}(Du(x)) - \mathcal{K}\left(\frac{Du(x)}{|Du(x)|}\right)\right| \le \omega\left(\left|Du(x) - \frac{Du(x)}{|Du(x)|}\right|\right) \le \omega(\sqrt{\varepsilon}).$$

The preceding estimate trivially holds if $|Du(x)| \le 1$. Moreover, by (3.16) we have

$$\left(|Du(y)| - 1\right)_+ = \left|\mathcal{G}(Du(y))\right| \le \left|\mathcal{G}(Du(y)) - \mathcal{G}(Du(x))\right| + \left|\mathcal{G}(Du(x))\right|$$
$$\le \varepsilon + \left(|Du(x)| - 1\right)_+ \le \varepsilon + \sqrt{\varepsilon} \le 2\sqrt{\varepsilon}.$$

This implies $|Du(y)| \le 1 + 2\sqrt{\varepsilon}$. Similarly as above we conclude

$$|\mathcal{K}(Du(y))| \le \omega(2\sqrt{\varepsilon}).$$

Combining the estimates from above, we end up with

$$\left| \mathcal{K}(Du(x)) - \mathcal{K}(Du(y)) \right| \leq \omega\left(\sqrt{\varepsilon}\right) + \omega\left(2\sqrt{\varepsilon}\right) \leq 2\omega\left(2\sqrt{\varepsilon}\right).$$

Now, we consider the case $|Du(x)| > 1 + \sqrt{\varepsilon}$. Here, Lemma 2.3 and (3.16) imply

$$|Du(x) - Du(y)| \leq \left(1 + \tfrac{2}{\sqrt{\varepsilon}}\right) \left| \mathcal{G}(Du(x)) - \mathcal{G}(Du(y)) \right| \leq \varepsilon + 2\sqrt{\varepsilon} \leq 3\sqrt{\varepsilon},$$

proving

$$\left| \mathcal{K}(Du(x)) - \mathcal{K}(Du(y)) \right| \leq \omega\left(3\sqrt{\varepsilon}\right).$$

Hence, $\mathcal{K}(Du)$ is continuous on $\overline{B}_r$. Since $B_r \Subset B_R \Subset \Omega$ were arbitrary, we have shown that $\mathcal{K}(Du)$ is continuous in Ω. This completes the proof of the theorem. $\qquad \square$

As mentioned above, we have now finished the proof of the main theorem Theorem 1.1 under the condition that Propositions 3.4 and 3.5 are true. The rest of the paper is now devoted to the proof of those two propositions.

4 Conclusions from the differentiated system

4.1 The main integral inequality for second derivatives

Throughout this subsection we assume as a *general requirement* that $u_\varepsilon \colon B_R \to \mathbb{R}^N$ is a weak solution to the regularized system (3.1). Instead of u_ε, we write u for the sake of simplicity. In contrast, we will continue to use the subscript ε in the notation for the coefficients h_ε and its associated bilinear forms, such as $\mathcal{C}_\varepsilon$. We recall that the bilinear forms have been defined in Sect. 2.3.

For some index $\beta = 1, \ldots, n$ we differentiate the regularized system (3.1) with respect to x_β and obtain

$$\int_{B_R} D_\beta\left[h_\varepsilon(|Du|)Du\right] \cdot D\varphi \, \mathrm{d}x = \int_{B_R} f \cdot D_\beta\varphi \, \mathrm{d}x, \qquad (4.1)$$

for any $\varphi \in W_0^{1,p}(B_R, \mathbb{R}^N)$. We have

$$
\begin{aligned}
D_\beta\left[h_\varepsilon(|Du|)D_\alpha u^i\right] &= h_\varepsilon(|Du|)D_\alpha D_\beta u^i + h'_\varepsilon(|Du|)|Du| \frac{D_\alpha u^i \, D_\gamma u^j \, u^j_{x_\gamma x_\beta}}{|Du|^2} \\
&= \left[h_\varepsilon(|Du|)\delta^{ij}\delta_{\gamma\alpha} + h'_\varepsilon(|Du|)|Du| \frac{D_\alpha u^i \, D_\gamma u^j}{|Du|^2}\right] u^j_{x_\gamma x_\beta}.
\end{aligned}
$$

In (4.1) we choose the testing function $\varphi = \zeta\phi(|Du|)D_\beta u$, where $\zeta \in C_0^1(B_R)$ is non-negative and $\phi \in W_{\text{loc}}^{1,\infty}(\mathbb{R}_{\geq 0}, \mathbb{R}_{\geq 0})$ is non-decreasing. Note that

$$D_\alpha\varphi = \zeta\phi(|Du|)D_\alpha D_\beta u + \zeta\phi'(|Du|)D_\alpha|Du|D_\beta u + D_\alpha\zeta\phi(|Du|)D_\beta u.$$

The resulting equations are then summed with respect to β from 1 to n. This leads to

$$\int_{B_R}\left[h_\varepsilon(|Du|)|D^2u|^2 + h'_\varepsilon(|Du|)|Du||\nabla|Du||^2\right]\phi(|Du|)\zeta\,dx$$

$$+ \int_{B_R}\left[h_\varepsilon(|Du|)|Du||\nabla|Du||^2 + h'_\varepsilon(|Du|)|Du\nabla|Du||^2\right]\phi'(|Du|)\zeta\,dx$$

$$+ \int_{B_R}\underbrace{\left[h_\varepsilon(|Du|)|Du|\nabla|Du| \cdot \nabla\zeta + h'_\varepsilon(|Du|)Du\nabla|Du| \cdot Du\nabla\zeta\right]}_{=\mathcal{C}_\varepsilon(Du)(\nabla|Du|,\nabla\zeta)|Du|}\phi(|Du|)\,dx$$

$$= \int_{B_R} f \cdot D_\beta\left[\zeta\phi(|Du|)D_\beta u\right]dx.$$

Now, we compute the right-hand side.

$$\int_{B_R} f \cdot D_\beta\left[\zeta\phi(|Du|)D_\beta u\right]dx$$

$$= \int_\Omega f \cdot D_\beta D_\beta u\,\phi(|Du|)\zeta\,dx + \int_\Omega f \cdot \phi'(|Du|)Du\nabla|Du|\zeta\,dx$$

$$+ \int_\Omega f \cdot \phi(|Du|)Du\nabla\zeta\,dx$$

$$= \mathbf{R}_1 + \mathbf{R}_2 + \mathbf{R}_3,$$

with the obvious meaning of $\mathbf{R}_i$. For the first term, we have by Young's inequality

$$\mathbf{R}_1 \leq \tau\int_{B_R} h_\varepsilon(|Du|)|D^2u|^2\phi(|Du|)\zeta\,dx + \frac{1}{\tau}\int_{B_R}\frac{|f|^2\phi(|Du|)}{h_\varepsilon(|Du|)}\zeta\,dx$$

for any $\tau \in (0, 1)$. Similarly, we get

$$\mathbf{R}_2 \leq \tau\int_{B_R} h_\varepsilon(|Du|)|Du||\nabla|Du||^2\phi'(|Du|)\zeta\,dx + \frac{1}{\tau}\int_{B_R}\frac{|f|^2\phi'(|Du|)|Du|}{h_\varepsilon(|Du|)}\zeta\,dx$$

and

$$\mathbf{R}_3 \leq \int_\Omega |f|\phi(|Du|)|Du||\nabla\zeta|\,dx.$$

Inserting this above and re-absorbing the terms containing second derivatives from the right-hand side into the left, we obtain

$$\int_{B_R} \left[(1-\tau)h_\varepsilon(|Du|)|D^2u|^2 + h'_\varepsilon(|Du|)|Du|\big|\nabla|Du|\big|^2\right]\phi(|Du|)\zeta\,dx$$

$$+ \int_{B_R} \left[(1-\tau)h_\varepsilon(|Du|)\big|\nabla|Du|\big|^2 + h'_\varepsilon(|Du|)|Du|\frac{\big|Du\nabla|Du|\big|^2}{|Du|^2}\right]\phi'(|Du|)|Du|\zeta\,dx$$

$$+ \int_{B_R} \mathcal{C}_\varepsilon(Du)\big(\nabla|Du|, \nabla\zeta\big)\phi(|Du|)|Du|\,dx$$

$$\leq \frac{1}{\tau}\int_{B_R} |f|^2\left[\frac{\phi(|Du|)}{h_\varepsilon(|Du|)} + \frac{\phi'(|Du|)|Du|}{h_\varepsilon(|Du|)}\right]\zeta\,dx + \int_{B_R} |f\phi(|Du|)|Du||D\zeta|\,dx$$

for any non-negative function $\zeta \in C^1_0(B_R)$. In the preceding inequality the parameter τ is at our disposal. We choose $\tau = \frac{1}{2}\min\{1, p-1\}$. For the first factor, i.e. the term $[\dots]$, in the integrand of the first integral on the left-hand side we have

$$(1-\tau)h_\varepsilon(|Du|)|D^2u|^2 + h'_\varepsilon(|Du|)|Du|\big|\nabla|Du|\big|^2 - \frac{1}{2}\mathcal{A}_\varepsilon(Du)\big(D^2u, D^2u\big)$$

$$= \frac{1}{2}\max\{0, 2-p\}h_\varepsilon(|Du|)|D^2u|^2 + \frac{1}{2}h'_\varepsilon(|Du|)|Du|\big|D|Du|\big|^2 \geq 0.$$

In fact, if $h'_\varepsilon(|Du|) \geq 0$ the inequality is obvious. If otherwise $h'_\varepsilon(|Du|) < 0$, which can only happen if $p < 2$ and $|Du| > 1$, the result follows by an application of Kato's inequality and (2.4). Indeed

$$\frac{1}{2}\max\{0, 2-p\}h_\varepsilon(|Du|)|D^2u|^2 + \frac{1}{2}h'_\varepsilon(|Du|)|Du|\big|\nabla|Du|\big|^2$$

$$\geq \frac{1}{2}\Big[(2-p)h(|Du|) + h'(|Du|)|Du|\Big]|D^2u|^2$$

$$= \frac{1}{2}\left[(2-p)\frac{(|Du|-1)^{p-1}_+}{|Du|} + \frac{(|Du|-1)^{p-2}_+[(p-2)|Du|+1]}{|Du|}\right]|D^2u|^2$$

$$= \frac{(|Du|-1)^{p-2}_+}{2|Du|}\Big[(2-p)(|Du|-1) + (p-2)|Du|+1\Big]|D^2u|^2$$

$$= (p-1)\frac{(|Du|-1)^{p-2}_+}{2|Du|}|D^2u|^2 \geq 0.$$

For the term in brackets of the second integral on the left-hand side a similar computation applies. The result of the calculation is

$$(1-\tau)h_\varepsilon(|Du|)\big|D|Du|\big|^2 + h'_\varepsilon(|Du|)|Du|\frac{\big|Du\nabla|Du|\big|^2}{|Du|^2}$$

$$- \frac{1}{2}\mathcal{C}_\varepsilon(Du)\big(\nabla|Du|, \nabla|Du|\big) \geq 0.$$

Using the last and second last inequality above, we obtain an inequality which can be interpreted in two ways. On the one hand it can be seen as an *energy inequality* for

the second derivatives of u. On the other hand—by discarding on the left-hand side the two non-negative terms containing second derivatives—the inequality implies that $|Du|$ is a *subsolution* of an elliptic equation with measurable coefficients.

Lemma 4.1 *Let $\varepsilon \in (0, 1]$ and $u = u_\varepsilon$ a weak solution to the regularized system* (3.1) *on B_R. Then, for any non-decreasing function $\phi \in W_{\mathrm{loc}}^{1,\infty}(\mathbb{R}_{\geq 0}, \mathbb{R}_{\geq 0})$ and any non-negative testing function $\zeta \in C_0^1(B_R)$ we have*

$$\int_{B_R} \left[\mathcal{A}_\varepsilon(Du)(D^2 u, D^2 u)\phi(|Du|) + \mathcal{C}_\varepsilon(Du)(\nabla|Du|, \nabla|Du|)\phi'(|Du|)|Du| \right] \zeta \, \mathrm{d}x$$

$$+ 2 \int_{B_R} \mathcal{C}_\varepsilon(Du)(\nabla|Du|, \nabla\zeta)\phi(|Du|)|Du| \, \mathrm{d}x$$

$$\leq c \int_{B_R} |f|^2 \left[\frac{\phi(|Du|)}{h_\varepsilon(|Du|)} + \frac{\phi'(|Du|)|Du|}{h_\varepsilon(|Du|)} \right] \zeta \, \mathrm{d}x + 2 \int_{B_R} |f|\phi(|Du|)|Du||\nabla\zeta| \, \mathrm{d}x,$$

where $c := \frac{4}{\min\{1, p-1\}}$. □

4.2 Subsolution to an elliptic equation

We start by showing that $(|Du_\varepsilon| - 1 - \delta)_+^2$ is a sub-solution of a certain elliptic equation. More precisely

Lemma 4.2 *Let $\varepsilon \in (0, 1]$ and $u_\varepsilon \in W^{1,p}(B_R, \mathbb{R}^N)$ be a weak solution of the regularized system* (3.1) *satisfying* (3.4) *and* (3.5) *on $B_\varrho(x_o) \subset B_{r_o} \Subset B_R$. Then, the function*

$$U_\varepsilon := \left(|Du_\varepsilon| - 1 - \delta \right)_+^2 \tag{4.2}$$

is a sub-solution of a linear elliptic equation on $B_\varrho(x_o)$, in the sense that

$$\int_{B_\varrho(x_o)} A_{\alpha\beta} D_\alpha U_\varepsilon D_\beta \zeta \, \mathrm{d}x \leq c \int_{B_\varrho(x_o)} |f|^2 \zeta \, \mathrm{d}x + c \mu \int_{B_\varrho(x_o)} |f||D\zeta| \, \mathrm{d}x, \tag{4.3}$$

holds true for any non-negative test function $\zeta \in C_0^1(B_\varrho(x_o))$ and with a universal constant $c = c(p, M, \delta)$. The coefficients $A_{\alpha\beta}$ are given by

$$A_{\alpha\beta}\eta_\alpha\zeta_\beta = |Du_\varepsilon|\,\mathcal{C}_\varepsilon(Du_\varepsilon)(\eta, \zeta) \quad \text{for } \eta, \zeta \in \mathbb{R}^n,$$

where $\mathcal{C}_\varepsilon$ is the bilinear form defined in (2.8).

Proof We apply Lemma 4.1 with $\phi(t) = (t - 1 - \delta)_+$. Due to Lemma 2.7 the first two integrals on the left-hand side are non-negative and therefore can be discarded. In

this way, we obtain with $c = \frac{4}{\min\{1, p-1\}}$ that

$$\mathbf{L} := \int_{B_\varrho(x_o)} |Du_\varepsilon| \, \mathcal{C}_\varepsilon(Du_\varepsilon)\big(\nabla|Du_\varepsilon|, \nabla\zeta\big)\phi(|Du_\varepsilon|) \, dx$$

$$\leq c \int_{B_\varrho(x_o)} |f|^2 \left[\frac{\phi(|Du_\varepsilon|)}{h_\varepsilon(|Du_\varepsilon|)} + \frac{\phi'(|Du_\varepsilon|)|Du_\varepsilon|}{h_\varepsilon(|Du_\varepsilon|)}\right]\zeta \, dx$$

$$+ \int_{B_\varrho(x_o)} |f|\phi(|Du_\varepsilon|)|Du_\varepsilon||D\zeta| \, dx =: \mathbf{R} \tag{4.4}$$

holds true for any non-negative $\zeta \in C_0^1(B_\varrho(x_o))$. To proceed further we compute

$$\frac{\phi(t)}{h_\varepsilon(t)} \leq \frac{\phi(t)}{h(t)} = \frac{t(t-1-\delta)_+}{(t-1)^{p-1}} \leq \frac{t^2}{\delta^{p-1}}$$

and

$$\frac{\phi'(t)t}{h_\varepsilon(t)} \leq \frac{\phi'(t)t}{h(t)} = \frac{\chi_{\{t>1+\delta\}}t^2}{(t-1)^{p-1}} \leq \frac{t^2}{\delta^{p-1}}.$$

The above inequalities together with the general assumptions (3.4), (3.5) allow to estimate

$$\frac{\phi(|Du_\varepsilon|)}{h_\varepsilon(|Du_\varepsilon|)} + \frac{\phi'(|Du_\varepsilon|)|Du_\varepsilon|}{h_\varepsilon(|Du_\varepsilon|)} \leq \frac{2M^2}{\delta^{p-1}}.$$

Moreover, we have $\phi(|Du_\varepsilon|)|Du_\varepsilon| \leq M\mu$. In this way, we obtain for the right-hand side in (4.4) the estimate

$$\mathbf{R} \leq c \int_{B_R} |f|^2\zeta \, dx + c\,\mu \int_{B_R} |f||D\zeta| \, dx,$$

where $c = c(p, M, \delta)$. Now, we consider the left-hand side in (4.4). Observe that $\nabla U_\varepsilon = 2\phi(|Du_\varepsilon|)\nabla|Du_\varepsilon|$, so that by the linearity of $\mathcal{C}_\varepsilon(Du_\varepsilon)$ with respect to the first variable we have

$$\mathcal{C}_\varepsilon(Du_\varepsilon)\big(\nabla|Du_\varepsilon|, \nabla\zeta\big)\phi(|Du_\varepsilon|) = \tfrac{1}{2}\mathcal{C}_\varepsilon(Du_\varepsilon)\big(\nabla U_\varepsilon, \nabla\zeta\big).$$

For the left-hand side this has the consequence that

$$\mathbf{L} = \tfrac{1}{2}\int_{B_\varrho(x_o)} |Du_\varepsilon|\mathcal{C}_\varepsilon(Du_\varepsilon)\big(\nabla U_\varepsilon, \nabla\zeta\big) \, dx = \tfrac{1}{2}\int_{B_\varrho(x_o)} A_{\alpha\beta}(x)D_\alpha U_\varepsilon D_\beta\zeta \, dx$$

holds true. Here, we have taken into account the definition of the coefficients $A_{\alpha\beta}$. Altogether we have shown the claim (4.3).   □

The coefficients $A_{\alpha\beta}$ in (4.3) can be explicitly written as

$$A_{\alpha\beta} := h_\varepsilon(|Du_\varepsilon|)|Du_\varepsilon| \left[\delta_{\alpha\beta} + \frac{h'_\varepsilon(|Du_\varepsilon|)D_\alpha u_\varepsilon \cdot D_\beta u_\varepsilon}{h_\varepsilon(|Du_\varepsilon|)|Du_\varepsilon|} \right].$$

They are only degenerate elliptic due to the factor $h_\varepsilon(|Du_\varepsilon|)|Du_\varepsilon|$ which, for $\varepsilon = 0$, vanishes on the set $\{|Du| \leq 1\}$. On the other hand U_ε has its support in the set $B_R \cap \{|Du_\varepsilon| \geq 1+\delta\}$. This allows us to modify the coefficients on $B_R \cap \{|Du_\varepsilon| \leq 1+\delta\}$. This idea will lead us to an energy estimate for U_ε in the next lemma.

Lemma 4.3 *Let $\varepsilon \in (0, 1]$ and $u_\varepsilon \in W^{1,p}(B_R, \mathbb{R}^N)$ be a weak solution of the regularized system* (3.1) *satisfying* (3.4) *and* (3.5) *on $B_\varrho(x_o) \subset B_{r_o} \Subset B_R$ and denote by U_ε the function defined in* (4.2). *Then, for any $k > 0$ and any $\tau \in (0, 1)$ we have*

$$\int_{B_{\tau\varrho}(x_o)} |D(U_\varepsilon - k)_+|^2 \, dx$$
$$\leq \frac{c}{(1-\tau)^2\varrho^2} \int_{B_\varrho(x_o)} (U_\varepsilon - k)_+^2 \, dx + c \, \|f\|_{n+\sigma}^2 \left| B_\varrho(x_o) \cap \{U_\varepsilon > k\} \right|^{1 - \frac{2}{n} + \frac{2\beta}{n}},$$

where $c = c(n, p, M, \delta)$ and $\beta = \frac{\sigma}{n+\sigma}$.

Proof We let $A_{\alpha\beta}$ be the coefficients defined in Lemma 4.2 and extend them from $B_R \cap \{|Du_\varepsilon| > 1+\delta\}$ to the complement $B_R \cap \{|Du_\varepsilon| \leq 1+\delta\}$ by letting $A_{\alpha\beta} \equiv \delta_{\alpha\beta}$. The new coefficients $\widetilde{A}_{\alpha\beta}$ are thus defined by

$$\widetilde{A}_{\alpha\beta}(x) := \begin{cases} \delta_{\alpha\beta}, & \text{on } \{x \in B_R : |Du_\varepsilon(x)| \leq 1 + \delta\}, \\ A_{\alpha\beta}(x), & \text{on } \{x \in B_R : |Du_\varepsilon(x)| > 1 + \delta\}. \end{cases}$$

From this definition and Lemma 4.2 we observe that U_ε is a weak sub-solution also with the modified coefficients. More precisely, we have that

$$\int_{B_\varrho(x_o)} \widetilde{A}_{\alpha\beta} D_\alpha U_\varepsilon D_\beta \zeta \, dx \leq c \int_{B_\varrho(x_o)} |f|^2 \zeta \, dx + c \, \mu \int_{B_\varrho(x_o)} |f||D\zeta| \, dx, \quad (4.5)$$

for any non-negative $\zeta \in C_0^1(B_R)$.

We now investigate the upper bound and ellipticity of the coefficients $\widetilde{A}_{\alpha\beta}$. We will show that there exist $0 < \lambda \leq \Lambda < \infty$ both depending at most on p, M and δ such that

$$\lambda|\zeta|^2 \leq \widetilde{A}_{\alpha\beta}(x)\zeta_\alpha\zeta_\beta \leq \Lambda|\zeta|^2$$

for any $x \in B_{r_o}$ and $\zeta \in \mathbb{R}^n$. We start with the former one. On the set where $|Du_\varepsilon| \leq 1 + \delta$ the upper bound holds with $\Lambda = 1$, while on the set where $|Du_\varepsilon| > 1 + \delta$ we have from Lemma 2.7 that

$$\begin{aligned}
\widetilde{A}_{\alpha\beta}\zeta_\alpha\zeta_\beta &= |Du_\varepsilon|\mathcal{C}_\varepsilon(Du_\varepsilon)(\zeta,\zeta) \\
&\leq |Du_\varepsilon|\big[\varepsilon + \Lambda(|Du_\varepsilon|)\big]|\zeta|^2 \\
&= \Big[\varepsilon|Du_\varepsilon| + \max\big\{(|Du_\varepsilon|-1)^{p-1}, (p-1)|Du_\varepsilon|(|Du_\varepsilon|-1)^{p-2}\big\}\Big]|\zeta|^2 \\
&\leq \Big[M + \frac{pM^p}{\delta}\Big]|\zeta|^2,
\end{aligned}$$

which proves the claim with $\Lambda = \Lambda(p,M,\delta) = M+\frac{pM^p}{\delta}$. Similarly, on the set where $|Du_\varepsilon| \leq 1+\delta$ the ellipticity holds with $\lambda = 1$, while on the set where $|Du_\varepsilon| > 1+\delta$ we have from Lemma 2.7 that

$$\begin{aligned}
\widetilde{A}_{\alpha\beta}\zeta_\alpha\zeta_\beta &= |Du_\varepsilon|\mathcal{C}_\varepsilon(Du_\varepsilon)(\zeta,\zeta) \\
&\geq |Du_\varepsilon|\big[\varepsilon + \lambda(|Du_\varepsilon|)\big]|\zeta|^2 \\
&\geq \min\big\{(|Du_\varepsilon|-1)^{p-1}, (p-1)|Du_\varepsilon|(|Du_\varepsilon|-1)^{p-2}\big\}|\zeta|^2 \\
&\geq \min\{1, p-1\}\delta^{p-1}|\zeta|^2,
\end{aligned}$$

which proves the claim with $\lambda = \lambda(p,\delta) = \min\{1, p-1\}\delta^{p-1}$.

Now, the claimed energy estimate follows in a standard way by choosing in (4.5) a test-function of the form $\zeta = \eta^2(U_\varepsilon - k)_+$ with a cut-off function $\eta \in C_0^1(B_\varrho(x_o))$ with $\eta \equiv 1$ on $B_{\tau\varrho}(x_o)$ and $|\nabla\eta| \leq \frac{2}{\tau\varrho}$, cf. [19, Chapter 10.1]. □

4.3 Energy estimates

Here, we assume that the hypothesis of Proposition 3.4 are in force. Our starting point is again Lemma 4.1. This time we keep the two non-negative terms containing the quadratic forms $\mathcal{A}_\varepsilon$ and $\mathcal{C}_\varepsilon$ on the left-hand side. For any non-decreasing function $\phi \in W_{\text{loc}}^{1,\infty}(\mathbb{R}_{\geq 0}, \mathbb{R}_{\geq 0})$ and any non-negative function $\zeta = \eta^2 \in C_0^1(B_\varrho(x_o))$ we have

$$\begin{aligned}
\int_{B_\varrho(x_o)} &\Big[\mathcal{A}_\varepsilon(Du_\varepsilon)(D^2u_\varepsilon, D^2u_\varepsilon)\phi(|Du_\varepsilon|) \\
&+ \mathcal{C}_\varepsilon(Du_\varepsilon)(\nabla|Du_\varepsilon|, \nabla|Du_\varepsilon|)\phi'(|Du_\varepsilon|)|Du_\varepsilon|\Big]\eta^2\,\mathrm{d}x \\
&\leq 4\Big|\int_{B_\varrho(x_o)} \mathcal{C}_\varepsilon(Du_\varepsilon)(\nabla|Du_\varepsilon|, \nabla\eta)\phi(|Du_\varepsilon|)|Du_\varepsilon|\eta\,\mathrm{d}x\Big| \\
&+ c(p)\int_{B_\varrho(x_o)} |f|^2\Big[\frac{\phi(|Du_\varepsilon|)}{h_\varepsilon(|Du_\varepsilon|)} + \frac{\phi'(|Du_\varepsilon|)|Du_\varepsilon|}{h_\varepsilon(|Du_\varepsilon|)}\Big]\eta^2\,\mathrm{d}x \\
&+ 4\int_{B_\varrho(x_o)} |f|\phi(|Du_\varepsilon|)|Du_\varepsilon||\nabla\eta|\eta\,\mathrm{d}x \\
&=: 4\mathbf{I} + c(p)\mathbf{II} + 4\mathbf{III}
\end{aligned}$$

with the obvious meaning of **I–III**. For **III** we have

$$4\mathbf{III} \leq 2 \int_{B_\varrho(x_o)} \left[|f|^2 \frac{\phi'(|Du_\varepsilon|)|Du_\varepsilon|}{h_\varepsilon(|Du_\varepsilon|)} \eta^2 + h_\varepsilon(|Du_\varepsilon|) \frac{\phi^2(|Du_\varepsilon|)|Du_\varepsilon|}{\phi'(|Du_\varepsilon|)} |\nabla\eta|^2 \right] dx.$$

Moreover, we estimate the integral **I** by Cauchy-Schwarz inequality and obtain

$$4\mathbf{I} \leq \int_{B_\varrho(x_o)} \mathcal{C}_\varepsilon(Du_\varepsilon)\big(\nabla|Du_\varepsilon|, \nabla|Du_\varepsilon|\big)\phi'(|Du_\varepsilon|)|Du_\varepsilon|\eta^2 \, dx$$

$$+ 4 \int_{B_\varrho(x_o)} \mathcal{C}_\varepsilon(Du_\varepsilon)\big(\nabla\eta, \nabla\eta\big) \frac{\phi^2(|Du_\varepsilon|)|Du_\varepsilon|}{\phi'(|Du_\varepsilon|)} \, dx$$

$$\leq \int_{B_\varrho(x_o)} \mathcal{C}_\varepsilon(Du_\varepsilon)\big(\nabla|Du_\varepsilon|, \nabla|Du_\varepsilon|\big)\phi'(|Du_\varepsilon|)|Du_\varepsilon|\eta^2 \, dx$$

$$+ 4 \int_{B_\varrho(x_o)} \big[\varepsilon + \mathbf{\Lambda}(|Du_\varepsilon|)\big] \frac{\phi^2(|Du_\varepsilon|)|Du_\varepsilon|}{\phi'(|Du_\varepsilon|)} |\nabla\eta|^2 \, dx.$$

From the second to last inequality we used the upper bound from Lemma 2.7 to estimate the second integral. Inserting the results above and re-absorbing the first term from the right into the left, we find that

$$\int_{B_\varrho(x_o)} \mathcal{A}_\varepsilon(Du_\varepsilon)\big(D^2u_\varepsilon, D^2u_\varepsilon\big)\phi(|Du_\varepsilon|)\eta^2 \, dx$$

$$\leq 4 \int_{B_\varrho(x_o)} \big[\varepsilon + \mathbf{\Lambda}(|Du_\varepsilon|) + h_\varepsilon(|Du_\varepsilon|)\big] \frac{\phi^2(|Du_\varepsilon|)|Du_\varepsilon|}{\phi'(|Du_\varepsilon|)} |\nabla\eta|^2 \, dx$$

$$+ c \int_{B_\varrho(x_o)} |f|^2 \frac{\phi(|Du_\varepsilon|) + \phi'(|Du_\varepsilon|)|Du_\varepsilon|}{h_\varepsilon(|Du_\varepsilon|)} \eta^2 \, dx \qquad (4.6)$$

holds true with a constant $c = c(p)$. We now choose $\phi(t) := (t-1)_+^p \tilde\phi(t)$, where $\tilde\phi \in W_{\mathrm{loc}}^{1,\infty}(\mathbb{R}_{\geq 0}, \mathbb{R}_{\geq 0})$ is non-decreasing. Note that

$$\phi'(t) = (t-1)_+^{p-1}\big[p\tilde\phi(t) + (t-1)_+\tilde\phi'(t)\big].$$

For $t \in [0, 1+2\mu]$ we compute

$$\big[\varepsilon + \mathbf{\Lambda}(t) + h_\varepsilon(t)\big] \frac{\phi^2(t)t}{\phi'(t)}$$

$$\leq 2\big[\varepsilon + \mathbf{\Lambda}(t)\big] \frac{(t-1)_+^{p+1} t\tilde\phi^2(t)}{p\tilde\phi(t) + (t-1)_+\tilde\phi'(t)}$$

$$\leq 2\Big[\varepsilon(t-1)_+^{p+1}t + \max\big\{(t-1)_+^{2p}, (p-1)(t-1)_+^{2p-1}t\big\}\Big] \frac{\tilde\phi^2(t)}{p\tilde\phi(t) + (t-1)_+\tilde\phi'(t)}$$

$$\leq c \big[\varepsilon\mu^{p+2} + \mu^{2p}\big] \frac{\tilde{\phi}^2(t)}{p\tilde{\phi}(t) + (t-1)_+\tilde{\phi}'(t)}$$

$$\leq c(p, M, \delta) \, \frac{\mu^{2p}\tilde{\phi}^2(t)}{p\tilde{\phi}(t) + (t-1)_+\tilde{\phi}'(t)}.$$

In turn we used $\delta \leq \mu \leq M$ from (3.5) and (3.6), which implies on the one hand $t \leq 1 + 2\mu \leq (2 + \frac{1}{\delta})\mu$, and on the other hand $\mu^2 \leq \max\{\delta^{2-p}, M^{2-p}\}\mu^p$. Next, we compute

$$\frac{\phi(t) + \phi'(t)t}{h_\varepsilon(t)} \leq \frac{\phi(t) + \phi'(t)t}{h(t)}$$
$$= \big[(t-1)_+\tilde{\phi}(t) + pt\tilde{\phi}(t) + (t-1)_+\tilde{\phi}'(t)t\big]t$$
$$\leq c(p, M, \delta)\,[\tilde{\phi}(t) + \tilde{\phi}'(t)t].$$

Due to assumptions (3.4) and (3.6) we know that $|Du_\varepsilon| \leq 1 + 2\mu$ on $B_\varrho(x_o)$. This allows us to use the preceding estimates in (4.6) to bound the right-hand side from above. Moreover, by Lemma 2.11 the left-hand side in (4.6) can be estimated from below. Proceeding in this way we obain

$$\int_{B_\varrho(x_o)} \big|D\big[g(|Du_\varepsilon|)Du_\varepsilon\big]\big|^2 \tilde{\phi}(|Du_\varepsilon|)\eta^2 \, dx$$

$$\leq c \int_{B_\varrho(x_o)} \frac{\mu^{2p}\tilde{\phi}^2(|Du_\varepsilon|)}{p\tilde{\phi}(|Du_\varepsilon|) + (|Du_\varepsilon| - 1)_+\tilde{\phi}'(|Du_\varepsilon|)} \, |\nabla\eta|^2 \, dx$$

$$+ c \int_{B_\varrho(x_o)} |f|^2 \big[\tilde{\phi}(|Du_\varepsilon|) + \tilde{\phi}'(|Du_\varepsilon|)|Du_\varepsilon|\big]\eta^2 \, dx, \qquad (4.7)$$

for any $\eta \in C_0^1(B_\varrho(x_o))$. The constant c depends only on p, M, and δ.

Different concrete choices of $\tilde{\phi}$ in (4.7) result in two important **energy inequalities**. The first one is

Lemma 4.4 *Let* $\varepsilon \in (0, 1]$ *and* $u_\varepsilon \in W^{1,p}(B_R, \mathbb{R}^N)$ *be a weak solution of the regularized system* (3.1) *such that hypotheses* (3.4), (3.5) *and* (3.6) *are in force on* $B_\varrho(x_o) \subset B_{r_o} \Subset B_R$. *Then, for any* $\tau \in (0, 1)$ *there holds*

$$\int_{B_{\tau\varrho}(x_o)} \big|D\big[g(|Du_\varepsilon|)Du_\varepsilon\big]\big|^2 dx \leq c\left[\frac{\mu^{2p}}{\varrho^2(1-\tau)^2} + \varrho^{-\frac{2n}{n+\sigma}}\|f\|_{L^{n+\sigma}(B_R)}^2\right]|B_\varrho|$$

for some universal constant $c = c(n, p, M, \delta)$.

Proof We apply inequality (4.7) with the choice $\tilde{\phi} \equiv 1$. The cut-off function $\eta \in C_0^1(B_\varrho(x_o))$ is chosen such that $\eta \equiv 1$ in $B_{\tau\varrho}(x_o)$, $0 \leq \eta \leq 1$, and $|\nabla\eta| \leq \frac{2}{(1-\tau)\varrho}$.

This leads us to

$$\int_{B_{\tau\varrho}(x_o)} \left|D\big[g(|Du_\varepsilon|)Du_\varepsilon\big]\right|^2 dx \leq c \int_{B_\varrho(x_o)} \left[\mu^{2p}|\nabla\eta|^2 + |f|^2\right] dx$$

$$\leq \frac{c\,\mu^{2p}}{(1-\tau)^2}\,\varrho^{n-2} + c\,\varrho^{n-\frac{2n}{n+\sigma}}\|f\|^2_{L^{n+\sigma}(B_R)}$$

with a constant $c = c(n, p, M, \delta)$, which is the claimed energy estimate. □

The second energy estimate is

Lemma 4.5 *Let $v \in (0, \frac{1}{4}]$, $\varepsilon \in (0, 1]$ and $u_\varepsilon \in W^{1,p}(B_R, \mathbb{R}^N)$ be a weak solution of the regularized system* (3.1) *such that hypotheses* (3.4), (3.5), (3.6) *and* (3.7) *are in force on $B_\varrho(x_o) \subset B_{r_o} \Subset B_R$. Then, for any $\tau \in (0, 1)$ we have*

$$\int_{E^v_{\tau\varrho}(x_o)} \left|D\big[g(|Du_\varepsilon|)Du_\varepsilon\big]\right|^2 dx \leq c\left[\frac{\mu^{2p}v}{\varrho^2(1-\tau)^2} + \frac{\varrho^{-\frac{2n}{n+\sigma}}}{v}\|f\|^2_{L^{n+\sigma}(B_R)}\right]|B_\varrho|,$$

for a constant $c = c(n, p, M, \delta)$.

Proof This time we choose

$$\tilde{\phi}(t) = (t - 1 - \delta - k)^2_+ \quad \text{with } k := (1 - 2v)\mu$$

in inequality (4.7), and obtain

$$\int_{B_\varrho(x_o)} \eta^2 \left|D\big[g(|Du_\varepsilon|)Du_\varepsilon\big]\right|^2 (|Du_\varepsilon| - 1 - \delta - k)^2_+ \, dx$$

$$\leq c \int_{B_\varrho(x_o)} \frac{\mu^{2p}(|Du_\varepsilon| - 1 - \delta - k)^3_+}{p(|Du_\varepsilon| - 1 - \delta - k) + 2(|Du_\varepsilon| - 1)}\,|\nabla\eta|^2 \, dx$$

$$+ c \int_{B_\varrho(x_o)} |f|^2 \, |Du_\varepsilon|(|Du_\varepsilon| - 1 - \delta - k)_+ \eta^2 \, dx.$$

On $B_\varrho(x_o) \cap \{|Du_\varepsilon| > 1 + \delta + k\}$ we have

$$(|Du_\varepsilon| - 1 - \delta - k)_+ \leq \mu - k = \mu - (1 - 2v)\mu = 2v\mu,$$

and

$$p(|Du_\varepsilon| - 1 - \delta - k) + 2(|Du_\varepsilon| - 1) \geq 2(\delta + k) \geq 2k \geq \mu,$$

since $v \leq \frac{1}{4}$. Again, we choose $\eta \in C^1_0(B_\varrho(x_o))$ to be a non-negative cut-off function with $\eta \equiv 1$ in $B_{\tau\varrho}(x_o)$, $0 \leq \eta \leq 1$, and $|\nabla\eta| \leq \frac{2}{(1-\tau)\varrho}$. This, together with the fact

that $|Du_\varepsilon| \leq 1 + 2\mu$ on $B_\varrho(x_o)$, allows us to estimate the right-hand side in the above inequality. Indeed, we have

$$\fint_{B_{\tau\varrho}(x_o)} \left|D\big[g(|Du_\varepsilon|)Du_\varepsilon\big]\right|^2 \big(|Du_\varepsilon| - 1 - \delta - k\big)_+^2 \, dx$$

$$\leq c\left[\frac{\nu^3\mu^{2p+2}}{(1-\tau)^2}\varrho^{n-2} + \nu\mu^2\int_{B_\varrho(x_o)}|f|^2\,dx\right].$$

Therefore it remains to estimate the left-hand side from below. The integral has to be taken only on the set of points $x \in B_\varrho(x_o)$ with $|Du_\varepsilon(x)| - 1 - \delta > k = (1-2\nu)\mu$. We shrink this set to those points satisfying the stronger condition $|Du_\varepsilon(x)| - 1 - \delta > (1-\nu)\mu > k$, i.e. to $E^\nu_{\tau\varrho}(x_o)$. On this set we have

$$|Du_\varepsilon| - 1 - \delta - k \geq (1-\nu)\mu - (1-2\nu)\mu = \nu\mu.$$

Inserting this above we conclude that

$$\nu^2\mu^2\int_{E^\nu_{\tau\varrho}(x_o)}\left|D\big[g(|Du_\varepsilon|)Du_\varepsilon\big]\right|^2\,dx \leq c\left[\frac{\nu^3\mu^{2p+2}}{\varrho^2(1-\tau)^2} + \nu\mu^2\varrho^{-\frac{2n}{n+\sigma}}\|f\|^2_{L^{n+\sigma}(B_R)}\right]|B_\varrho|$$

holds true. This proves the claim. $\qquad\square$

5 The non-degenerate regime

The aim of this section is to prove Proposition 3.4. Throughout this section we presume the following general assumptions. For given $\varepsilon \in (0,1]$ we denote by $u_\varepsilon \in W^{1,p}(B_R, \mathbb{R}^N)$ the unique weak solution of the regularized system (3.1). Moreover, we assume that for some $\delta \in (0,1]$ and $\mu > \delta$ and a ball $B_{2\varrho}(x_o) \subset B_{r_1}$ with $\varrho \leq 1$ assumptions (3.4)–(3.6) are in force. We denote by

$$\Phi(x_o, \varrho) := \fint_{B_\varrho(x_o)}\left|Du_\varepsilon - (Du_\varepsilon)_{x_o,\varrho}\right|^2\,dx \tag{5.1}$$

the L^2-excess of Du_ε on $B_\varrho(x_o)$, i.e. the L^2-mean square deviation of Du_ε from its mean value $(Du_\varepsilon)_{x_o,\varrho}$.

5.1 Higher integrability

An ingredient in the proof of Proposition 3.4 is the following higher integrability result.

Lemma 5.1 *Under the general assumptions of Sect. 5 there exist $\vartheta = \vartheta(n, p, \sigma, M, \delta) \in (0, \min\{\frac{1}{2}, \frac{n+\sigma}{2} - 1\}]$ and $c = c(n, p, M, \delta)$ such for any $\xi \in \mathbb{R}^{Nn}$ satisfying*

$$1 + \delta + \tfrac{1}{4}\mu \leq |\xi| \leq 1 + \delta + \mu,$$

we have

$$\fint_{B_{\varrho/2}(x_o)} |Du_\varepsilon - \xi|^{2(1+\vartheta)} \mathrm{d}x \le c \left[\fint_{B_\varrho(x_o)} |Du_\varepsilon - \xi|^2 \mathrm{d}x \right]^{1+\vartheta} + c \fint_{B_\varrho(x_o)} (\varrho|f|)^{2(1+\vartheta)} \mathrm{d}x.$$

Proof We consider a ball $B_s(z_o) \subset B_\varrho(x_o)$. We test the weak form (3.2) of the elliptic system by the testing function

$$\varphi := \eta^2 w, \quad \text{where} \quad w := u_\varepsilon - (u_\varepsilon)_{z_o,s} - \xi(x - z_o)$$

and $\eta \in C_0^1(B_s(z_o))$ is a standard cut-off function with $\eta \equiv 1$ in $B_{s/2}(z_o)$, $0 \le \eta \le 1$, and $|\nabla \eta| \le \frac{4}{s}$. We obtain

$$\begin{aligned}
0 &= \int_{B_s(z_o)} \Big[\mathbf{A}_\varepsilon(Du_\varepsilon) \cdot D\varphi + f \cdot \varphi \Big] \mathrm{d}x \\
&= \int_{B_s(z_o)} \Big[\big(\mathbf{A}_\varepsilon(Du_\varepsilon) - \mathbf{A}_\varepsilon(\xi) \big) \cdot D\varphi + f \cdot \varphi \Big] \mathrm{d}x \\
&= \int_{B_s(z_o)} \Big[\big(\mathbf{A}_\varepsilon(Du_\varepsilon) - \mathbf{A}_\varepsilon(\xi) \big) \cdot \big[\eta^2 Dw + 2\eta \nabla \eta \otimes w \big] + f \cdot \varphi \Big] \mathrm{d}x.
\end{aligned}$$

We use the monotonicity of $\mathbf{A}_\varepsilon$ from Lemma 2.8 in order to estimate the first term from below. Due to our assumption on ξ and (3.4) we have $|\xi| + |Du_\varepsilon| \le 5|\xi|$ and therefore obtain

$$\int_{B_s(z_o)} \eta^2 \big(\mathbf{A}_\varepsilon(Du_\varepsilon) - \mathbf{A}_\varepsilon(\xi) \big) \cdot Dw \, \mathrm{d}x \ge \left[\varepsilon + \lambda \frac{(|\xi| - 1)^p}{|\xi|^2} \right] \int_{B_s(z_o)} \eta^2 |Dw|^2 \, \mathrm{d}x,$$

where $\lambda = \frac{1}{5 \cdot 2^{p+1}} \min\{1, p-1\}$. To bound the second integral from above we use the structural upper bound from Lemma 2.8. Moreover, we observe that $(|Du_\varepsilon| - 1)_+ \le 4(|\xi| - 1)$ due to our assumption on ξ. This allows us to estimate

$$\begin{aligned}
2 \left| \int_{B_s(z_o)} \eta \big(\mathbf{A}_\varepsilon(Du_\varepsilon) - \mathbf{A}_\varepsilon(\xi) \big) \cdot \nabla \eta \otimes w \, \mathrm{d}x \right| \\
\le c \big[\varepsilon + (|\xi| - 1)^{p-2} \big] \int_{B_s(z_o)} \eta |Dw| |w| |\nabla \eta| \, \mathrm{d}x \\
\le \left[\varepsilon + \tfrac{1}{2} \lambda \frac{(|\xi| - 1)^p}{|\xi|^2} \right] \int_{B_s(z_o)} \eta^2 |Dw|^2 \, \mathrm{d}x \\
+ c \big[\varepsilon + |\xi|^2 (|\xi| - 1)^{p-4} \big] \int_{B_s(z_o)} |w|^2 |\nabla \eta|^2 \, \mathrm{d}x.
\end{aligned}$$

for a constant $c = c(p)$. Re-absorbing terms on the left-hand side, we find that

$$\frac{1}{2}\lambda \frac{(|\xi|-1)^p}{|\xi|^2} \int_{B_s(z_o)} \eta^2 |Dw|^2 \, dx$$
$$\leq c\big[1 + |\xi|^2(|\xi|-1)^{p-4}\big] \int_{B_s(z_o)} |w|^2 |D\eta|^2 \, dx + \int_{B_s(z_o)} |f||w| \, dx,$$

where $c = c(p)$. Due to the assumption on ξ, the assumption (3.5), and the particular choice of η, we conclude with an application of Hölder's and Sobolev–Poincaré's inequality a *reverse Hölder inequality* of the form

$$\fint_{B_{s/2}(z_o)} |Dw|^2 \, dx \leq \frac{c}{s^2} \fint_{B_s(z_o)} |w|^2 \, dx + c\,s^2 \fint_{B_s(z_o)} |f|^2 \, dx$$
$$\leq c\bigg[\fint_{B_s(z_o)} |Dw|^{\frac{2n}{n+2}} \, dx\bigg]^{\frac{n+2}{n}} + c\,\varrho^2 \fint_{B_s(z_o)} |f|^2 \, dx,$$

with a constant $c = c(n, p, M, \delta)$. The dependence of c upon M only occurs in the sub-quadratic case of $p < 2$. The claim, i.e. the higher integrability, now follows with Gehring's lemma, since $Dw = Du_\varepsilon - \xi$, cf. [1, Theorem 3.22], [22, Theorem 2.4] and [42, Theorem 3.3]. Note that ϑ can always be diminished if necessary. □

5.2 Comparison with a linear system

In this section we will consider the weak solution $v \in u_\varepsilon + W_0^{1,2}(B_{\varrho/2}(x_o), \mathbb{R}^N)$ of the linear elliptic system

$$\int_{B_{\varrho/2}(x_o)} \mathcal{B}_\varepsilon\big((Du_\varepsilon)_{x_o,\varrho/2}\big)(Dv, D\varphi) \, dx = 0, \tag{5.2}$$

for any $\varphi \in W_0^{1,2}(B_{\varrho/2}(x_o), \mathbb{R}^N)$ as comparison function to our solution u_ε of the regularized elliptic system (3.1). Recall that $\mathcal{B}_\varepsilon$ has been defined in (2.7).

Lemma 5.2 *Let the general assumptions of Sect. 5 be in force and assume that*

$$1 + \delta + \tfrac{1}{4}\mu \leq \big|(Du_\varepsilon)_{x_o,\varrho}\big| \leq 1 + \delta + \mu. \tag{5.3}$$

Then, there exists $\vartheta = \vartheta(n, p, M, \sigma, \delta) \in (0, \min\{\frac{1}{2}, \frac{n+\sigma}{2} - 1\}]$ and $c = c(n, p, M, \delta, \|f\|_{L^{n+\sigma}(B_R)})$ such that

$$\fint_{B_{\varrho/2}(x_o)} |Du_\varepsilon - Dv|^2 \, dx \leq c\bigg[\frac{\Phi(x_o, \varrho)}{\mu^2}\bigg]^\vartheta \Phi(x_o, \varrho) + c\,\varrho^{2\beta}.$$

Here, v is the unique weak solution of the Dirichlet problem (5.2) and $\beta = \frac{\sigma}{n+\sigma}$.

Proof Throughout the proof we omit the reference to the center x_o and write B_ϱ instead of $B_\varrho(x_o)$. Moreover, we abbreviate $\xi := (Du_\varepsilon)_\varrho$. Using the weak form (3.2) of the elliptic system we obtain

$$
\begin{aligned}
0 &= \int_{B_{\varrho/2}} \mathbf{A}_\varepsilon(Du_\varepsilon) \cdot D\varphi \, dx + \int_{B_{\varrho/2}} f \cdot \varphi \, dx \\
&= \int_{B_{\varrho/2}} \left(\mathbf{A}_\varepsilon(Du_\varepsilon) - \mathbf{A}_\varepsilon(\xi) \right) \cdot D\varphi \, dx + \int_{B_{\varrho/2}} f \cdot \varphi \, dx,
\end{aligned}
$$

for any $\varphi \in W_0^{1,2}(B_{\varrho/2}, \mathbb{R}^N)$. Using also the fact that v is a weak solution of the linear elliptic system (5.2), we find that

$$
\begin{aligned}
& \int_{B_{\varrho/2}} \mathcal{B}_\varepsilon(\xi)(Du_\varepsilon - Dv, D\varphi) \, dx \\
&= \int_{B_{\varrho/2}} \mathcal{B}_\varepsilon(\xi)(Du_\varepsilon, D\varphi) \, dx \\
&= \int_{B_{\varrho/2}} \mathcal{B}_\varepsilon(\xi)(Du_\varepsilon - \xi, D\varphi) \, dx \\
&= \int_{B_{\varrho/2}} \left[\mathcal{B}_\varepsilon(\xi)(Du_\varepsilon - \xi, D\varphi) - \left(\mathbf{A}_\varepsilon(Du_\varepsilon) - \mathbf{A}_\varepsilon(\xi) \right) \cdot D\varphi \right] dx - \int_{B_{\varrho/2}} f \cdot \varphi \, dx \\
&\leq c(p)\, \mu^{p-3} \int_{B_{\varrho/2}} |Du_\varepsilon - \xi|^2 |D\varphi| \, dx + \left(\int_{B_{\varrho/2}} |f|^2 \, dx \right)^{\frac{1}{2}} \left(\int_{B_{\varrho/2}} |\varphi|^2 \, dx \right)^{\frac{1}{2}}.
\end{aligned}
$$

Here we used from the second to last line Lemma 2.10. This is possible since (3.4) and (5.3) are in force. Since $u_\varepsilon - v \in W_0^{1,2}(B_{\varrho/2}, \mathbb{R}^N)$, we may choose the testing function $\varphi = u_\varepsilon - v$. Together with the bound from below from Lemma 2.7 and Hölder's and Poincaré's inequality this leads us to

$$
\begin{aligned}
& \gamma \mu^{p-2} \int_{B_{\varrho/2}} |Du_\varepsilon - Dv|^2 \, dx \\
&\leq \int_{B_{\varrho/2}} \mathcal{B}_\varepsilon(\xi)(Du_\varepsilon - Dv, Du_\varepsilon - Dv) \, dx \\
&\leq c(p)\, \mu^{p-3} \int_{B_{\varrho/2}} |Du_\varepsilon - \xi|^2 |Du_\varepsilon - Dv| \, dx \\
&\quad + \left(\int_{B_{\varrho/2}} |f|^2 \, dx \right)^{\frac{1}{2}} \left(\int_{B_{\varrho/2}} |u_\varepsilon - v|^2 \, dx \right)^{\frac{1}{2}} \\
&\leq c(p)\, \mu^{p-3} \left(\int_{B_{\varrho/2}} |Du_\varepsilon - \xi|^4 \, dx \right)^{\frac{1}{2}} \left(\int_{B_{\varrho/2}} |Du_\varepsilon - Dv|^2 \, dx \right)^{\frac{1}{2}} \\
&\quad + c(n)\, \varrho \left(\int_{B_{\varrho/2}} |f|^2 \, dx \right)^{\frac{1}{2}} \left(\int_{B_{\varrho/2}} |Du_\varepsilon - Dv|^2 \, dx \right)^{\frac{1}{2}}
\end{aligned}
$$

for a constant $\gamma = \gamma(p, \delta) > 0$. We divide both sides by

$$\gamma \mu^{p-2} \Big[\int_{B_{\varrho/2}} |Du_\varepsilon - Dv|^2 \, dx \Big]^{\frac{1}{2}},$$

square the result and finally take means. This implies

$$\fint_{B_{\varrho/2}} |Du_\varepsilon - Dv|^2 \, dx \leq \frac{c}{\mu^2} \fint_{B_{\varrho/2}} |Du_\varepsilon - \xi|^4 \, dx + c \, \varrho^{\frac{2\sigma}{n+\sigma}} \|f\|^2_{L^{n+\sigma}(B_R)}$$

with a constant $c = c(n, p, M, \delta)$. Here we have also used $\delta < \mu \leq M$. At this stage arrived, we want to reduce the integrability exponent on the right-hand side from 4 to $2(1 + \vartheta)$, where $\vartheta = \vartheta(n, p, M, \delta) \in (0, \min\{\frac{1}{2}, \frac{n+\sigma}{2} - 1\}]$ is the integrability exponent from the higher integrability Lemma 5.1. This is possible since $|Du_\varepsilon|$ and $|\xi|$ are bounded by $(2 + \frac{1}{\delta})\mu$ on account of (3.4) and (3.6). Then, the application of the higher integrability lemma yields

$$\fint_{B_{\varrho/2}} |Du_\varepsilon - \xi|^4 \, dx \leq \frac{c}{\mu^{2\vartheta}} \fint_{B_{\varrho/2}} |Du_\varepsilon - \xi|^{2(1+\vartheta)} \, dx$$

$$\leq \frac{c}{\mu^{2\vartheta}} \Big[\fint_{B_\varrho} |Du_\varepsilon - \xi|^2 \, dx \Big]^{1+\vartheta} + c \fint_{B_\varrho} (\varrho|f|)^{2(1+\vartheta)} \, dx$$

$$\leq c \Big[\frac{\Phi(\varrho)}{\mu^2} \Big]^{\vartheta} \Phi(\varrho) + c \, \varrho^{\frac{2\sigma(1+\vartheta)}{n+\sigma}} \|f\|^{2(1+\vartheta)}_{L^{n+\sigma}(B_R)},$$

where $c = c(n, p, M, \delta)$. Inserting this inequality above and noting that $\varrho \leq 1$ finishes the proof of the lemma.   □

The following a priori estimate for solutions to linear elliptic systems can be inferred from [5] once the ellipticity conditions for the quadratic form $\mathcal{B}_\varepsilon\big((Du_\varepsilon)_{x_o,\varrho}\big)$ are established; see also [21, Theorem 2.3].

Lemma 5.3 *Let the general assumptions of Sect. 5 be in force and assume that (5.3) holds true. Then, the weak solution $v \in W^{1,2}(B_{\varrho/2}(x_o), \mathbb{R}^N)$ of the linear elliptic system (5.2) satisfies $v \in W^{2,2}_{loc}(B_{\varrho/2}(x_o), \mathbb{R}^N)$ and there exists a constant $c_o = c_o(n, N, p, \delta)$ such that for any $\tau \in (0, \frac{1}{2}]$ we have*

$$\fint_{B_{\tau\varrho}(x_o)} |Dv - (Dv)_{x_o,\tau\varrho}|^2 \, dx \leq c_o \tau^2 \fint_{B_{\varrho/2}(x_o)} |Dv - (Dv)_{x_o,\varrho/2}|^2 \, dx.$$

Proof As mentioned before, the a priori estimate is standard. The constant c_o depends on the dimensions n, N and the ellipticity constant and the upper bound of the quadratic form $\mathcal{B}_\varepsilon\big((Du_\varepsilon)_{x_o,\varrho}\big)$. Due to assumption (5.3) and Lemma 2.7 these quantities only depend on p and δ.   □

5.3 Exploiting the measure theoretic information

The aim of this subsection is to convert the measure theoretic information (3.7) into a lower bound for the mean value of Du_ε and smallness of the excess.

Lemma 5.4 *Let the general assumptions of Sect. 5 be in force. Furthermore, assume that* (3.7) *holds for some* $v \in (0, \frac{1}{4}]$. *Then there exists a constant* $c = c(n, p, M, \delta)$ *such that for any* $\tau \in [\frac{1}{2}, 1)$ *there holds*

$$\Phi(x_o, \tau\varrho) \leq c\,\mu^2 \left[\frac{v^{\frac{2}{n}}}{(1-\tau)^2} + \frac{\varrho^{2\beta}}{v} \|f\|_{L^{n+\sigma}(B_R)}^2 \right].$$

Proof Throughout the proof we omit the reference to the center x_o and write B_ϱ instead of $B_\varrho(x_o)$. We define $\zeta \in \mathbb{R}^{Nn}$ by

$$|\zeta|^{p-1}\zeta = \left(|\mathcal{G}(Du_\varepsilon)|^{p-1}\mathcal{G}(Du_\varepsilon) \right)_{\tau\varrho}$$

and let

$$\tilde{\zeta} := \mathcal{G}^{-1}(\zeta).$$

Note that $|\zeta| \leq \delta + \mu$ by (3.4) and $|\tilde{\zeta}| \leq 1 + \delta + \mu$. Due to the minimality of the integral average $(Du_\varepsilon)_{\tau\varrho}$ with respect to the mapping $\xi \mapsto \int_{B_{\tau\varrho}} |Du_\varepsilon - \xi|^2 \, dx$, we have

$$\Phi(\tau\varrho) \leq \fint_{B_{\tau\varrho}} |Du_\varepsilon - \tilde{\zeta}|^2 \, dx$$

$$= \frac{1}{|B_{\tau\varrho}|} \int_{E_{\tau\varrho}^v} |Du_\varepsilon - \tilde{\zeta}|^2 \, dx + \frac{1}{|B_{\tau\varrho}|} \int_{B_{\tau\varrho} \setminus E_{\tau\varrho}^v} |Du_\varepsilon - \tilde{\zeta}|^2 \, dx =: \mathbf{I} + \mathbf{II}.$$

We recall that $|Du_\varepsilon|, |\tilde{\zeta}| \leq 1+\delta+\mu$ and hence by (3.6) we have $|Du_\varepsilon|, |\tilde{\zeta}| \leq (2+\frac{1}{\delta})\mu$. Due to assumption (3.7) we therefore obtain for the second integral

$$\mathbf{II} \leq \frac{c(\delta)\,\mu^2}{|B_{\tau\varrho}|} \left| B_\varrho \setminus E_\varrho^v \right| \leq \frac{c(\delta)\,v\mu^2}{\tau^n}.$$

For the estimate of $\mathbf{I}$ we first note that $|Du_\varepsilon| \geq 1 + \frac{3}{2}\delta$ on $E_{\tau\varrho}^v$ since $v \in (0, \frac{1}{4}]$ and $\mu \geq \delta$. Therefore, the application of Lemma 2.3 yields

$$\mathbf{I} \leq \frac{c(\delta)}{|B_{\tau\varrho}|} \int_{E_{\tau\varrho}^v} |\mathcal{G}(Du_\varepsilon) - \zeta|^2 \, dx.$$

Next, we note that

$$|\mathcal{G}(Du_\varepsilon)| + |\zeta| \geq |\mathcal{G}(Du_\varepsilon)| = (|Du_\varepsilon| - 1)_+ > \delta + (1-v)\mu > \tfrac{1}{2}\mu \quad \text{on } E_{\tau\varrho}^v.$$

Using this information, Lemma 2.2, the choice of ζ, and Poincaré's inequality we obtain

$$
\begin{aligned}
\mathbf{I} &\leq \frac{c(p,\delta)}{\mu^{2p-2}|B_{\tau\varrho}|} \int_{E_{\tau\varrho}^{\nu}} \big(|\mathcal{G}(Du_\varepsilon)| + |\zeta|\big)^{2p-2} \big|\mathcal{G}(Du_\varepsilon) - \zeta\big|^2 \, dx \\
&\leq \frac{c(p,\delta)}{\mu^{2p-2}} \fint_{B_{\tau\varrho}} \big||\mathcal{G}(Du_\varepsilon)|^{p-1} \mathcal{G}(Du_\varepsilon) - |\zeta|^{p-1}\zeta\big|^2 \, dx \\
&\leq \frac{c(p,\delta)\varrho^2}{\mu^{2p-2}} \left[\fint_{B_{\tau\varrho}} \big|D\big(|\mathcal{G}(Du_\varepsilon)|^{p-1}\mathcal{G}(Du_\varepsilon)\big)\big|^{\frac{2n}{n+2}} \, dx \right]^{\frac{n+2}{n}} \\
&= \frac{c(p,\delta)\varrho^2}{\mu^{2p-2}} \left[\fint_{B_{\tau\varrho}} \big|D\big[g(Du_\varepsilon)Du_\varepsilon\big]\big|^{\frac{2n}{n+2}} \, dx \right]^{\frac{n+2}{n}}.
\end{aligned}
$$

We once again decompose the domain of integration into $E_{\tau\varrho}^{\nu}$ and $B_{\tau\varrho} \setminus E_{\tau\varrho}^{\nu}$. Subsequently applying Hölder's inequality and taking into account assumption (3.7) leads us to

$$
\mathbf{I} \leq \frac{c(p,\delta)\varrho^2}{\mu^{2p-2}|B_{\tau\varrho}|} \left[\int_{E_{\tau\varrho}^{\nu}} \big|D\big[g(Du_\varepsilon)Du_\varepsilon\big]\big|^2 \, dx + \nu^{\frac{2}{n}} \int_{B_{\tau\varrho}\setminus E_{\tau\varrho}^{\nu}} \big|D\big[g(Du_\varepsilon)Du_\varepsilon\big]\big|^2 \, dx \right].
$$

We note that $\tau \geq \frac{1}{2}$ and hence $|B_{\tau\varrho}| \geq c(n)\varrho^n$. For the first integral we use Lemma 4.5, while for the second one we use Lemma 4.4 and the assumption $\mu \geq \delta$. In this way we obtain

$$
\begin{aligned}
\mathbf{I} &\leq \frac{c\,\mu^2}{(1-\tau)^2}\big[\nu + \nu^{\frac{2}{n}}\big] + \frac{c\,\varrho^{2-\frac{2n}{n+\sigma}}}{\nu\mu^{2p-2}} \|f\|_{L^{n+\sigma}(B_R)}^2 \\
&\leq c\,\mu^2 \left[\frac{\nu^{\frac{2}{n}}}{(1-\tau)^2} + \frac{\varrho^{\frac{2\sigma}{n+\sigma}}}{\nu} \|f\|_{L^{n+\sigma}(B_R)}^2 \right],
\end{aligned}
$$

for a constant $c = c(n,p,M,\delta)$. Inserting this above yields the desired estimate. □

Lemma 5.5 *Let the general assumptions of Sect. 5 be in force. Then, for any $\theta \in (0, \frac{1}{64}]$ there exist $\nu = \nu(n,p,M,\delta,\theta) \in (0,\frac{1}{4}]$ and $\varrho_o = \varrho_o(n,p,\sigma,\|f\|_{L^{n+\sigma}(B_R)}, M,\delta,\theta) \in (0,1]$ such that the smallness assumption $\varrho \leq \varrho_o$ and the measure theoretic hypothesis (3.7) imply*

$$
\big|(Du_\varepsilon)_{x_o,\varrho}\big| \geq 1 + \delta + \tfrac{1}{2}\mu \quad \text{and} \quad \Phi(x_o,\varrho) \leq \theta\mu^2. \tag{5.4}
$$

Proof We let $\tau \in [\frac{1}{2}, 1)$, $\nu \in (0,\frac{1}{4}]$ and $\varrho_o \in (0,1]$. Consider $B_\varrho(x_o) \subset B_R$ with $\varrho \leq \varrho_o$. For convenience in notation we omit the reference to the center x_o. Using the minimality of $(Du)_\varrho$ with respect to the mapping $\xi \mapsto \fint_{B_\varrho} |Du - \xi|^2 \, dx$ and

decomposing the domain of integration into $B_{\tau\varrho}$ and $B_\varrho \setminus B_{\tau\varrho}$, we obtain

$$
\begin{aligned}
\Phi(\varrho) &\leq \fint_{B_\varrho} |Du_\varepsilon - (Du_\varepsilon)_{\tau\varrho}|^2 \, dx \\
&= \frac{1}{|B_\varrho|} \left[\int_{B_{\tau\varrho}} |Du_\varepsilon - (Du_\varepsilon)_{\tau\varrho}|^2 \, dx + \int_{B_\varrho \setminus B_{\tau\varrho}} |Du_\varepsilon - (Du_\varepsilon)_{\tau\varrho}|^2 \, dx \right] \\
&= \mathbf{I} + \mathbf{II}.
\end{aligned}
$$

For the first integral we use Lemma 5.4 and obtain

$$
\mathbf{I} = \tau^n \Phi(\tau\varrho) \leq c\,\mu^2 \left[\frac{\nu^{\frac{2}{n}}}{(1-\tau)^2} + \frac{\varrho^{2\beta}}{\nu} \|f\|^2_{L^{n+\sigma}(B_R)} \right],
$$

where $c = c(n, p, M, \delta)$. For the second integral we use $|Du_\varepsilon| \leq 1 + \delta + \mu \leq c(\delta)\mu$ and get

$$
\mathbf{II} \leq \frac{4(1 + \delta + \mu)^2 |B_\varrho \setminus B_{\tau\varrho}|}{|B_\varrho|} \leq c(\delta)\,\mu^2(1 - \tau^n) \leq c(n, \delta)\,\mu^2(1 - \tau),
$$

so that

$$
\Phi(\varrho) \leq c\,\mu^2 \left[\frac{\nu^{\frac{2}{n}}}{(1-\tau)^2} + (1 - \tau) + \frac{\varrho^{2\beta}}{\nu} \|f\|^2_{L^{n+\sigma}(B_R)} \right],
$$

for a constant $c = c(n, p, M, \delta)$. Now, we first choose $\tau \in [\frac{1}{2}, 1)$ in dependence on n, p, M, δ and θ in such a way that $c(1 - \tau) \leq \frac{1}{3}\theta$. Subsequently, we choose $\nu \in (0, \frac{1}{4}]$ in dependence on n, p, M, δ and θ such that

$$
\nu \leq \min \left\{ \left(\frac{\theta(1 - \tau)^2}{3c} \right)^{\frac{n}{2}}, \frac{\delta}{4(1 + \delta)} \right\}.
$$

Finally, we choose $\varrho_o \in (0, 1]$ such that

$$
\varrho_o^{2\beta} \leq \frac{\nu\theta}{3c\left(1 + \|f\|^2_{L^{n+\sigma}(B_R)}\right)}.
$$

In this way we obtain (5.4)$_2$.

To prove (5.4)$_1$, we first observe that the measure theoretic assumption (3.7) implies

$$
|E_\varrho^\nu| > (1 - \nu)|B_\varrho|.
$$

Hence, due to the definition of the set E_ϱ^v, we obtain

$$\int_{B_\varrho} |Du_\varepsilon| \, dx \geq \int_{E_\varrho^v} |Du_\varepsilon| \, dx \geq \left(1 + \delta + (1-v)\mu\right)|E_\varrho^v|$$
$$\geq (1-v)\left(1 + \delta + (1-v)\mu\right)|B_\varrho|.$$

On the other hand, due to $(5.4)_2$, we have

$$\left| \fint_{B_\varrho} |Du_\varepsilon| \, dx - \left|(Du_\varepsilon)_\varrho\right| \right| \leq \fint_{B_\varrho} |Du_\varepsilon - (Du_\varepsilon)_\varrho| \, dx \leq \sqrt{\Phi(\varrho)} \leq \sqrt{\theta}\,\mu,$$

so that

$$\left|(Du_\varepsilon)_\varrho\right| \geq (1-v)\left(1 + \delta + (1-v)\mu\right) - \sqrt{\theta}\mu.$$

Due to the choice of v and the fact that $n \geq 2$ we have $v \leq (\frac{1}{3}\theta)^{\frac{n}{2}} \leq \frac{1}{3}\theta \leq \frac{1}{2}\sqrt{\theta}$ and $v \leq \frac{\delta}{4(1+\delta)}$. Together with the assumptions $\delta \leq \mu$ and $\theta \leq \frac{1}{64}$ we obtain

$$- v(1+\delta) + (1-v)^2\mu - \sqrt{\theta}\mu - \tfrac{1}{2}\mu \geq -\tfrac{1}{4}\delta + \left(\tfrac{1}{2} - 2v - \sqrt{\theta}\right)\mu$$
$$\geq \left[\tfrac{1}{4} - 2\sqrt{\theta}\right]\mu \geq 0.$$

Inserting this above yields the claim $(5.4)_1$ and finishes the proof of the lemma. □

5.4 Proof of Proposition 3.4

Our aim in this subsection is to prove Proposition 3.4. We start with an excess-decay estimate for the excess $\Phi(x_o, \varrho)$ of Du_ε.

Lemma 5.6 *Assume that the general hypotheses of Sect. 5 are in force. Let $\tau \in (0, \frac{1}{2}]$ and $\vartheta = \vartheta(n, p, M, \delta) \in (0, \frac{1}{2}]$ be the exponent from Lemma 5.2. If*

$$|(Du_\varepsilon)_{x_o,\varrho}| \geq 1 + \delta + \tfrac{1}{4}\mu \quad and \quad \Phi(x_o, \varrho) \leq \tau^{\frac{n+2}{\vartheta}}\mu^2, \tag{5.5}$$

hold true, then we have the quantitative excess decay estimate

$$\Phi(x_o, \tau\varrho) \leq c_*\left[\tau^2\Phi(x_o, \varrho) + \tau^{-n}\varrho^{2\beta}\mu^2\right]$$

with a constant $c_* = c_*(n, N, p, \|f\|_{L^{n+\sigma}(B_R)}, M, \delta)$.

Proof Throughout the proof we omit the reference to the center x_o and write B_ϱ instead of $B_\varrho(x_o)$. By $v \in u_\varepsilon + W_0^{1,2}(B_\varrho, \mathbb{R}^N)$ we denote the unique weak solution of the

linear elliptic system (5.2). For $\tau \in (0, \frac{1}{2}]$, we have

$$\Phi(\tau \varrho) \leq \fint_{B_{\tau \varrho}} |Du_\varepsilon - (Dv)_{\tau \varrho}|^2 \, dx$$

$$\leq 2 \fint_{B_{\tau \varrho}} |Du_\varepsilon - Dv|^2 \, dx + 2 \fint_{B_{\tau \varrho}} |Dv - (Dv)_{\tau \varrho}|^2 \, dx.$$

In view of Lemma 5.3 we deduce

$$\fint_{B_{\tau \varrho}} |Dv - (Dv)_{\tau \varrho}|^2 \, dx$$

$$\leq c \, \tau^2 \fint_{B_{\varrho/2}} |Dv - (Dv)_{\varrho/2}|^2 \, dx$$

$$\leq c \, \tau^2 \fint_{B_{\varrho/2}} |Dv - (Du_\varepsilon)_{\varrho/2}|^2 \, dx$$

$$\leq c \, \tau^2 \fint_{B_{\varrho/2}} |Du_\varepsilon - Dv|^2 \, dx + c \, \tau^2 \fint_{B_{\varrho/2}} |Du_\varepsilon - (Du_\varepsilon)_{\varrho/2}|^2 \, dx$$

$$\leq c \, \tau^2 \fint_{B_{\varrho/2}} |Du_\varepsilon - Dv|^2 \, dx + c \, \tau^2 \Phi(\varrho),$$

where $c = c(n, N, p, \delta)$. Inserting this above and applying Lemma 5.2 and assumption (5.5)$_2$, we end up with

$$\Phi(\tau \varrho) \leq \frac{c}{\tau^n} \fint_{B_{\varrho/2}} |Du_\varepsilon - Dv|^2 \, dx + c \, \tau^2 \Phi(\varrho)$$

$$\leq c \left[\frac{1}{\tau^n} \left[\frac{\Phi(\varrho)}{\mu^2} \right]^\vartheta + \tau^2 \right] \Phi(\varrho) + c \, \tau^{-n} \varrho^{2\beta}$$

$$\leq c_* \left[\tau^2 \Phi(\varrho) + \tau^{-n} \varrho^{2\beta} \mu^2 \right].$$

Note that the constant c_* depends on $n, N, p, \|f\|_{L^{n+\sigma}(B_R)}, M$ and δ. $\qquad \square$

Proof of Proposition 3.4 By $\vartheta = \vartheta(n, p, \sigma, M, \delta) \in (0, \min\{\frac{1}{2}, \frac{n+\sigma}{2} - 1\}]$ we denote the constant from Lemma 5.2 and by $c_* = c_*(n, N, p, \|f\|_{L^{n+\sigma}(B_R)}, M, \delta)$ the one from Lemma 5.6. For $\beta = \frac{\sigma}{n+\sigma} \in (0, 1)$ we define $\tau \in (0, \frac{1}{8}]$ by

$$\tau := \min \left\{ \frac{1}{8}, 2^{-\frac{1}{\beta}}, (10c_*)^{-\frac{1}{2(1-\beta)}} \right\}.$$

For the particular choice $\theta = \tau^{\frac{n+2}{\vartheta}}$ we let $\varrho_o = \varrho_o(n, p, \sigma, \|f\|_{L^{n+\sigma}(B_R)}, M, \delta) \in (0, 1]$ be the radius from Lemma 5.5. Finally, we define

$$\hat{\varrho} := \min \left\{ \varrho_o, (2c_*)^{-\frac{1}{2\beta}} \tau^{1+\frac{2n+2}{\beta \vartheta}} \right\},$$

so that $\hat{\varrho}$ depends on n, N, p, σ, $\|f\|_{L^{n+\sigma}(B_R)}$, M, δ. In the following we consider a ball $B_{2\varrho}(x_o) \subset B_{r_1}$ with $\varrho \leq \hat{\varrho}$. As before, we omit the reference to the center x_o and write B_ϱ instead of $B_\varrho(x_o)$. By $\nu = \nu(n, p, M, \delta, \theta = \frac{1}{2}\tau^{\frac{n+2}{\vartheta}}) \in (0, \frac{1}{4}]$ we denote the constant from Lemma 5.5 and assume that (3.7) is satisfied for this particular choice of ν. Note that by our choice of τ the parameter ν depends on n, N, p, σ, $\|f\|_{L^{n+\sigma}(B_R)}$, M and δ. From Lemma 5.5 applied with $\theta = \tau^{\frac{n+2}{\vartheta}}$ we infer that

$$\left|(Du)_\varrho\right| \geq 1 + \delta + \tfrac{1}{2}\mu \quad \text{and} \quad \Phi(\varrho) \leq \tau^{\frac{n+2}{\vartheta}}\mu^2. \tag{5.6}$$

By induction we shall prove that for any $i \in \mathbb{N}$ we have

$$\Phi(\tau^i \varrho) \leq \tau^{\frac{n+2}{\vartheta}}\tau^{2\beta i}\mu^2 \tag{I$_i$}$$

and

$$\left|(Du)_{\tau^i \varrho}\right| \geq 1 + \delta + \left[\frac{1}{2} - \frac{1}{8}\sum_{j=0}^{i-1}2^{-j}\right]\mu. \tag{II$_i$}$$

For $i = 1$ we can apply Lemma 5.6, since (5.6) ensures that the assumptions of the lemma are satisfied. Then, (I)$_1$ follows from Lemma 5.6, (5.6)$_2$ and our choices of τ and $\hat{\varrho}$, since

$$\begin{aligned}
\Phi(\tau\varrho) &\leq c_*\left[\tau^2\Phi(\varrho) + \tau^{-n}\varrho^{2\beta}\mu^2\right] \\
&\leq \tfrac{1}{2}\tau^{2\beta}\Phi(\varrho) + \frac{c_*}{\tau^n}\varrho^{2\beta}\mu^2 \\
&\leq \tau^{\frac{n+2}{\vartheta}}\tau^{2\beta}\left[\frac{1}{2} + \frac{c_*\hat{\varrho}^{2\beta}}{\tau^{\frac{n+2}{\vartheta}+n+2\beta}}\right]\mu^2 \\
&\leq \tau^{\frac{n+2}{\vartheta}}\tau^{2\beta i}\mu^2.
\end{aligned}$$

For the proof of (II)$_1$ we use (5.6)$_2$ and $\tau^{\frac{n+2}{\vartheta}} \leq \tau^{n+2}$ to obtain

$$\left|(Du)_{\tau\varrho} - (Du)_\varrho\right|^2 \leq \fint_{B_{\tau\varrho}}\left|Du - (Du)_\varrho\right|^2 \mathrm{d}x \leq \tau^{-n}\Phi(\varrho) \leq \tau^{\frac{n+2}{\vartheta}-n}\mu^2 \leq \tau^2\mu^2,$$

so that

$$\left|(Du)_{\tau\varrho} - (Du)_\varrho\right| \leq \tau\mu \leq \tfrac{1}{8}\mu.$$

Together with (5.6)$_1$ this implies (II)$_1$.

Now, we consider $i > 1$ and prove (I)$_i$ and (II)$_i$ assuming that (I)$_{i-1}$ and (II)$_{i-1}$ hold. From (I)$_{i-1}$ and (II)$_{i-1}$ we observe that the assumptions of Lemma 5.6 as formulated

in (5.5) are satisfied on $B_{\tau^{i-1}\varrho}$. Therefore, applying the lemma with $\tau^{i-1}\varrho$ instead of ϱ, recalling the choices of τ and $\hat{\varrho}$ and joining the result with $(I)_{i-1}$ yields

$$
\begin{aligned}
\Phi(\tau^i\varrho) &\leq c_* \left[\tau^2 \Phi(\tau^{i-1}\varrho) + \tau^{-n}(\tau^{i-1}\varrho)^{2\beta}\mu^2 \right] \\
&\leq \tfrac{1}{2}\tau^{2\beta}\Phi(\tau^{i-1}\varrho) + \frac{c_*}{\tau^n}(\tau^{i-1}\varrho)^{2\beta}\mu^2 \\
&\leq \tau^{\frac{n+2}{\vartheta}}\tau^{2\beta i}\left[\tfrac{1}{2} + \frac{c_*\hat{\varrho}^{2\beta}}{\tau^{\frac{n+2}{\vartheta}+n+2\beta}} \right]\mu^2 \\
&\leq \tau^{\frac{n+2}{\vartheta}}\tau^{2\beta i}\mu^2.
\end{aligned}
$$

This proves $(I)_i$. Moreover, from $(I)_{i-1}$ and $\tau^{\frac{n+2}{\vartheta}} \leq \tau^{n+2}$ we obtain

$$
\begin{aligned}
\left| (Du)_{\tau^i\varrho} - (Du)_{\tau^{i-1}\varrho} \right|^2 &\leq \fint_{B_{\tau^i\varrho}} \left| Du - (Du)_{\tau^{i-1}\varrho} \right|^2 dx \leq \tau^{-n}\Phi(\tau^{i-1}\varrho) \\
&\leq \tau^{\frac{n+2}{\vartheta}-n}\tau^{2\beta(i-1)}\mu^2 \leq \tau^{2\beta(i-1)}\tau^2\mu^2,
\end{aligned}
$$

so that

$$
\left| (Du)_{\tau^i\varrho} - (Du)_{\tau^{i-1}\varrho} \right| \leq \tau^{\beta(i-1)}\tau\mu \leq \tfrac{1}{8}2^{-(i-1)}\mu,
$$

by our choice of τ. Together with $(II)_{i-1}$, this proves $(II)_i$.

We now come to the proof of (3.8) and (3.9). For $i \in \mathbb{N}$ we obtain from the minimizing property of the mean value, Lemma 2.3, $(I)_i$, (5.6) and our choice of τ that

$$
\begin{aligned}
\Psi_{2\delta}(\tau^i\varrho) &:= \fint_{B_{\tau^i\varrho}} \left| \mathcal{G}_{2\delta}(Du) - \left(\mathcal{G}_{2\delta}(Du) \right)_{\tau^i\varrho} \right|^2 dx \\
&\leq \fint_{B_{\tau^i\varrho}} \left| \mathcal{G}_{2\delta}(Du) - \mathcal{G}_{2\delta}\left((Du)_{\tau^i\varrho} \right) \right|^2 dx \\
&\leq 9\,\Phi(\tau^i\varrho) \leq 9\,\tau^{\frac{n+2}{\vartheta}}\tau^{2\beta i}\mu^2 \leq \tau^{2n+2}\tau^{2\beta i}\mu^2.
\end{aligned}
\tag{5.7}
$$

This allows us to compute

$$
\begin{aligned}
\left| \left(\mathcal{G}_{2\delta}(Du) \right)_{\tau^i\varrho} - \left(\mathcal{G}_{2\delta}(Du) \right)_{\tau^{i-1}\varrho} \right|^2 &\leq \fint_{B_{\tau^i\varrho}} \left| \mathcal{G}_{2\delta}(Du) - \left(\mathcal{G}_{2\delta}(Du) \right)_{\tau^{i-1}\varrho} \right|^2 dx \\
&\leq \tau^{-n}\Psi_{2\delta}\left(\tau^{i-1}\varrho \right) \leq \tau^{n+2}\tau^{2\beta(i-1)}\mu^2.
\end{aligned}
$$

Given $j < k$, we use the preceding inequality to conclude that

$$\left|\left(\mathcal{G}_{2\delta}(Du)\right)_{\tau^j \varrho} - \left(\mathcal{G}_{2\delta}(Du)\right)_{\tau^k \varrho}\right| \leq \sum_{i=j+1}^{k} \left|\left(\mathcal{G}_{2\delta}(Du)\right)_{\tau^i \varrho} - \left(\mathcal{G}_{2\delta}(Du)\right)_{\tau^{i-1} \varrho}\right|$$

$$\leq \tau^{\frac{n}{2}+1} \sum_{i=j+1}^{k} \tau^{\beta(i-1)} \, \mu \leq \tau^{\frac{n}{2}+1} \frac{\tau^{\beta j}}{1 - \tau^\beta} \, \mu$$

$$\leq 2\tau^{\frac{n}{2}+1} \tau^{\beta j} \, \mu. \tag{5.8}$$

This shows that $((\mathcal{G}_{2\delta}(Du))_{\tau^i \varrho})_{i=1}^{\infty}$ is a Cauchy sequence and therefore the limit

$$\Gamma_{x_o} := \lim_{i \to \infty} \left(\mathcal{G}_{2\delta}(Du)\right)_{\tau^i \varrho}$$

exists. Passing to the limit $k \to \infty$ in (5.8) yields

$$\left|\left(\mathcal{G}_{2\delta}(Du)\right)_{\tau^j \varrho} - \Gamma_{x_o}\right| \leq 2\tau^{\frac{n}{2}+1} \tau^{\beta j} \mu \qquad \text{for any } j \in \mathbb{N}.$$

Joining this with (5.7), we find

$$\fint_{B_{\tau^j \varrho}} \left|\mathcal{G}_{2\delta}(Du) - \Gamma_{x_o}\right|^2 dx \leq 2\Psi_{2\delta}(\tau^j \varrho) + 2\left|\left(\mathcal{G}_{2\delta}(Du)\right)_{\tau^j \varrho} - \Gamma_{x_o}\right|^2$$

$$\leq 10\tau^{n+2} \tau^{2\beta j} \mu^2.$$

For $r \in (0, \varrho]$ there exists $j \in \mathbb{N}_0$ such that $\tau^{j+1}\varrho < r \leq \tau^j \varrho$. Then, we obtain from the last inequality

$$\fint_{B_r} \left|\mathcal{G}_{2\delta}(Du) - \Gamma_{x_o}\right|^2 dx \leq \tau^{-n} \fint_{B_{\tau^j \varrho}} \left|\mathcal{G}_{2\delta}(Du) - \Gamma_{x_o}\right|^2 dx$$

$$\leq 10\tau^2 \tau^{2\beta j} \mu^2 \leq \left(\frac{r}{\varrho}\right)^{2\beta} \mu^2.$$

This implies

$$\left|\left(\mathcal{G}_{2\delta}(Du)\right)_r - \Gamma_{x_o}\right|^2 \leq \fint_{B_r} \left|\mathcal{G}_{2\delta}(Du) - \Gamma_{x_o}\right|^2 dx \leq \left(\frac{r}{\varrho}\right)^{2\beta} \mu^2,$$

so that also

$$\Gamma_{x_o} = \lim_{r \downarrow 0} \left(\mathcal{G}_{2\delta}(Du)\right)_r.$$

Finally, due to assumption (3.4) we have $|(\mathcal{G}_{2\delta}(Du))_r| \leq \mu$ for any $0 < r \leq \varrho$, which implies $|\Gamma_{x_o}| \leq \mu$. This finishes the proof of Proposition 3.4. $\square$

6 The degenerate regime

Our aim in this section is to prove Proposition 3.5, which treats the degenerate regime. The proof relies on a De Giorgi type reduction argument reducing the supremum of $U_\varepsilon = (|Du_\varepsilon| - 1 - \delta)_+^2$ under the measure theoretic assumption (3.10). The starting point is the energy estimate for U_ε from Lemma 4.3.

As in Section 5, we first formulate the general assumptions. For $\varepsilon \in (0, 1]$ we denote by $u_\varepsilon \in W^{1,p}(B_R, \mathbb{R}^N)$ the unique weak solution to the Dirichlet problem (3.1) associated to the regularized system. We assume that (3.4) is in force for some $\mu, \delta > 0$ on some ball $B_{2\varrho}(x_o) \subset B_{r_1} \Subset B_R$. Let $U_\varepsilon := (|Du_\varepsilon| - 1 - \delta)_+^2$ denote the function defined in (4.2). Note that (3.4) implies

$$\sup_{B_{2\varrho}(x_o)} U_\varepsilon \le \mu^2.$$

Moreover, we set $\beta := \frac{\sigma}{n+\sigma} \in (0, 1)$.

We start by a De Giorgi type lemma for U_ε, which can for instance be deduced as in [19, Chap. 10, Proposition 4.1] by the use of the energy estimate from Lemma 4.3. For the readers convenience we provide the proof in the appendix Sect. 7.

Lemma 6.1 (Reducing the supremum) *Assume that the general assumptions of Sect. 6 are in force and let $\theta \in (0, 1)$. Then, there exists $\tilde{\nu} = \tilde{\nu}(n, p, \|f\|_{n+\sigma}, M, \delta) \in (0, 1)$ such that the measure theoretic assumption*

$$\left|\left\{x \in B_\varrho(x_o) : U_\varepsilon(x) > (1-\theta)\mu^2\right\}\right| < \tilde{\nu}\left|B_\varrho(x_o)\right|,$$

implies that either

$$\mu^2 < \frac{\varrho^\beta}{\theta},$$

or

$$U_\varepsilon \le \left(1 - \tfrac{1}{2}\theta\right)\mu^2 \quad \text{in } B_{\varrho/2}(x_o)$$

hold true. □

The proof of the next Lemma can be deduced as in [19, Chap. 10, Proposition 5.1] utilizing the energy estimate from Lemma 4.3; see also Sect. 7.

Lemma 6.2 *Assume that the general assumptions of Sect. 6 are in force and assume that (3.10) is satisfied for some $\nu \in (0, 1)$. Then, for any $i_* \in \mathbb{N}$ we either have*

$$\mu^2 < 2^{i_*}\varrho^\beta/\nu$$

or

$$\left|\left\{x \in B_\varrho(x_o) : U_\varepsilon(x) > (1 - 2^{-i_*}\nu)\mu^2\right\}\right| < \frac{c_*}{\nu\sqrt{i_*}}\left|B_\varrho(x_o)\right|$$

for a constant $c_* = c_*(n, p, \|f\|_{n+\sigma}, M, \delta)$.

Now, we have all the prerequisites at hand to provide the

Proof of Proposition 3.5 Let $\tilde{\nu} \in (0, 1)$ and c_* be the constants from Lemmas 6.1 and 6.2. Note that both depend on n, p, $\|f\|_{n+\sigma}$, M and δ. Choose $i_* \in \mathbb{N}$ such that

$$i_* \geq \left(\frac{c_*}{\nu\tilde{\nu}}\right)^2.$$

Then i_* depends n, p, $\|f\|_{n+\sigma}$, M, δ and ν. Lemma 6.2 implies that either $\mu^2 < 2^{i_*}\varrho^\beta/\nu$, or

$$\left|\left\{x \in B_\varrho(x_o) : U_\varepsilon(x) > (1 - 2^{-i_*}\nu)\mu^2\right\}\right| < \frac{c_*}{\nu\sqrt{i_*}}\left|B_\varrho(x_o)\right| \leq \tilde{\nu}\left|B_\varrho(x_o)\right|.$$

In the first case the proposition is proved with $c_o = 2^{i_*}/\nu$, while in the second case we may apply Lemma 6.1 with $\theta = 2^{-i_*}\nu$. Therefore either $\mu^2 < 2^{i_*}\varrho^\beta/\nu$ or

$$U_\varepsilon \leq \left(1 - 2^{-(i_*+1)}\nu\right)\mu^2 \quad \text{in } B_{\varrho/2}(x_o).$$

The first alternative coincides with the first alternative above, while the second one implies the sup-bound for $\mathcal{G}_\delta(Du_\varepsilon)$ for any $\kappa \geq \sqrt{1 - 2^{-(i_*+1)}\nu}$ since $U_\varepsilon = |\mathcal{G}_\delta(Du_\varepsilon)|^2$. Therefore we my choose $\kappa \in [2^{-\beta/2}, 1)$ as required. $\qquad\square$

7 Appendix

Here, we provide the proofs of Lemmas 6.1 and 6.2. We first state a well known iteration lemma which can be found for instance in [19, Chap. 9.15.1, Lemma 15.1].

Lemma 7.1 *Let $(Y_i)_{i\in\mathbb{N}_0}$ be a sequence of non-negative numbers satisfying*

$$Y_{i+1} \leq \kappa\, b^i Y_i^{1+\gamma} \quad \text{for all } i \in \mathbb{N}_0$$

with some positive constants κ, γ and $b > 1$. If

$$Y_0 \leq \kappa^{-\frac{1}{\gamma}} b^{-\frac{1}{\gamma^2}},$$

then $Y_i \to 0$ as $i \to \infty$.

The next lemma can be seen as a discrete version of the isoperimetric inequality, cf. [19, Chap. 10.5.1, inequality (5.4)].

Lemma 7.2 *Let $v \in W^{1,1}(B_\varrho(x_o))$ and $k, \ell \in \mathbb{R}$ with $k < \ell$. Then, there exists a constant c depending on n such that*

$$(\ell - k)|B_\varrho(x_o) \cap \{v < k\}| \leq \frac{c\varrho^{n+1}}{|B_\varrho(x_o) \cap \{v > \ell\}|} \int_{B_\varrho(x_o) \cap \{k < v < \ell\}} |Dv| \, dx.$$

Proof of Lemma 6.1 For $i \in \mathbb{N}$ we let

$$B_i := B_{\varrho_i}(x_o), \quad \text{where } \varrho_i := \tfrac{1}{2}\varrho(1 + 2^{-i}),$$

and consider for the levels

$$k_i := \left(1 - \tfrac{1}{2}\theta - 2^{-(i+1)}\theta\right)\mu^2$$

the normalized measure of the associated super-level sets

$$Y_i := \frac{|A_i|}{|B_i|}, \quad \text{where } A_i := \{x \in B_i : U_\varepsilon(x) > k_i\}.$$

In view of Hölder's, Sobolev's inequality and Lemma 4.3 we now estimate[1]

$$\int_{B_{i+1}} (U_\varepsilon - k_i)_+^2 \, dx \leq |A_i|^{\frac{2}{n}} \left[\int_{B_{i+1}} (U_\varepsilon - k_i)_+^{\frac{2n}{n-2}} \, dx \right]^{\frac{n-2}{n}}$$

$$\leq c \, |A_i|^{\frac{2}{n}} \int_{B_{i+1}} \left[\left| D(U_\varepsilon - k_i)_+ \right|^2 + \frac{(U_\varepsilon - k_i)_+^2}{\varrho_{i+1}^2} \right] dx$$

$$\leq \frac{c \, 4^i}{\varrho^2} |A_i|^{\frac{2}{n}} \int_{B_i} (U_\varepsilon - k_i)_+^2 \, dx + c \, \|f\|_{n+\sigma}^2 |A_i|^{1+\frac{2\beta}{n}}$$

$$\leq \frac{c \, 4^i \theta^2 \mu^4}{\varrho^2} |A_i|^{1+\frac{2}{n}} + c \, |A_i|^{1+\frac{2\beta}{n}},$$

where $c = c(n, p, \|f\|_{n+\sigma}, M, \delta)$. On the other hand, by the definition of k_i we have

$$|A_{i+1}| \leq \frac{4^{i+2}}{\theta^2 \mu^4} \int_{B_{i+1}} (U_\varepsilon - k_i)_+^2 \, dx.$$

---

[1] If $n = 2$ we may choose any exponent > 2 instead of $\frac{2n}{n-2}$.

Here we used that $U_\varepsilon - k_i \geq k_{i+1} - k_i$ on A_{i+1}. Joining the preceding two inequalities and recalling the definition of Y_i shows

$$
\begin{aligned}
Y_{i+1} &= \frac{|A_{i+1}|}{|B_{i+1}|} \\
&\leq \frac{c\, 4^{2i} |A_i|^{1+\frac{2\beta}{n}}}{|B_i|^{1+\frac{2\beta}{n}}} \left[\frac{|A_i|^{\frac{2(1-\beta)}{n}} |B_i|^{\frac{2\beta}{n}}}{\varrho^2} + \frac{|B_i|^{\frac{2\beta}{n}}}{\theta^2 \mu^4} \right] \\
&\leq c\, 4^{2i} Y_i^{1+\frac{2\beta}{n}} \left[1 + \frac{\varrho^{2\beta}}{\theta^2 \mu^4} \right].
\end{aligned}
$$

If $\mu^2 < \varrho^\beta / \theta$, the lemma is proved. Otherwise, the term in brackets on the right-hand side is bounded by 2, so that

$$
Y_{i+1} \leq c\, 4^{2i} Y_i^{1+\frac{2}{n}},
$$

where $c = c(n, p, \|f\|_{n+\sigma}, M, \delta)$. Therefore, Lemma 7.1 ensures that $Y_i \to 0$ in the limit $i \to \infty$, provided that $Y_0 \leq 2^{-n^2} c^{-\frac{n}{2}} =: \tilde{\nu}$ and hence $U_\varepsilon \leq (1 - \frac{1}{2}\theta)\mu^2$ on $B_{\varrho/2}(x_o)$. $\qquad \square$

Proof of Lemma 6.2 For convenience in notation we omit the reference to the center x_o. For $i \in \{0, \ldots, i_*\}$ we consider the super-level sets

$$
A_i := \left\{ x \in B_\varrho : U_\varepsilon(x) > k_i \right\}, \qquad \text{where } k_i := (1 - 2^{-i}\nu)\mu^2.
$$

Then, due to assumption (3.10) we have $|B_\varrho \setminus A_i| \geq \nu |B_\varrho|$ for any $i \in \{0, \ldots, i_*\}$. Applying Lemma 7.2 to U_ε on B_ϱ with $k = k_i$ and $\ell = k_{i+1}$, we obtain with a constant $c = c(n)$ that

$$
\begin{aligned}
\frac{\nu \mu^2}{2^{i+1}} |A_{i+1}| &\leq \frac{c\, \varrho^{n+1}}{|B_\varrho \setminus A_i|} \int_{A_i \setminus A_{i+1}} |DU_\varepsilon| \, dx \\
&\leq \frac{c\, \varrho}{\nu} \left[\int_{B_\varrho} |D(U_\varepsilon - k_i)_+|^2 \, dx \right]^{\frac{1}{2}} \left(|A_i| - |A_{i+1}| \right)^{\frac{1}{2}}.
\end{aligned}
$$

In view of the energy estimate from Lemma 4.3 and the definition of k_i we have

$$
\begin{aligned}
\int_{B_\varrho} |D(U_\varepsilon - k_i)_+|^2 \, dx &\leq \frac{c}{\varrho^2} \int_{B_{2\varrho}} |(U_\varepsilon - k_i)_+|^2 \, dx + c \|f\|_{n+\sigma}^2 \varrho^{n-2+2\beta} \\
&\leq \frac{c\, \nu^2 \mu^4}{2^{2i} \varrho^2} \left[1 + \left(\frac{2^i \varrho^\beta}{\nu \mu^2} \right)^2 \right] |B_\varrho|
\end{aligned}
$$

with a constant $c = c(n, p, \|f\|_{n+\sigma}, M, \delta)$. If $\mu^2 < 2^i \varrho^\beta / \nu$, the lemma is proved. Otherwise, the term in brackets on the right-hand side is bounded by 2. Inserting this above yields

$$|A_{i+1}|^2 \leq \frac{c}{\nu^2} \left(|A_i| - |A_{i+1}| \right) |B_\varrho|,$$

for $c = c(n, p, \|f\|_{n+\sigma}, M, \delta)$. Now, we add these inequalities for $i = 0, 1, \ldots, i_* - 1$ and obtain

$$i_* |A_{i_*}|^2 \leq \frac{c}{\nu^2} |B_\varrho| \sum_{i=0}^{i_*-1} \left(|A_i| - |A_{i+1}| \right) \leq \frac{c}{\nu^2} |A_0| |B_\varrho| \leq \frac{c}{\nu^2} |B_\varrho|^2.$$

Therefore, we have

$$|A_{i_*}| \leq \frac{c}{\nu \sqrt{i_*}} |B_\varrho|.$$

This proves the assertion of the lemma. □

References

1. Björn, A., Björn, J.: Nonlinear Potential Theory on Metric Spaces. EMS Tracts in Mathematics, vol. 17. European Mathematical Society (EMS), Zürich (2011)
2. Brasco, L.: Global L^∞ gradient estimates for solutions to a certain degenerate elliptic equation. Nonlinear Anal. **74**(2), 516–531 (2011)
3. Brasco, L., Carlier, G.: On certain anisotropic elliptic equations arising in congested optimal transport: local gradient bounds. Adv. Calc. Var. **7**, 379–407 (2014)
4. Brasco, L., Carlier, G., Santambrogio, F.: Congested traffic dynamics, weak flows and very degenerate elliptic equations. J. Math. Pures Appl. (9) **93**(6), 652–671 (2010)
5. Campanato, S.: Equazioni ellittiche del II ordine e spazi $\mathcal{L}^{2,\lambda}$. Ann. Mat. Pura Appl. **69**, 321–381 (1965)
6. Carlier, G., Jimenez, C., Santambrogio, F.: Optimal transportation with traffic congestion and Wardrop equilibria. SIAM J. Control Optim. **47**(3), 1330–1350 (2008)
7. Carlier, G., Santambrogio, F.: A continuous theory of traffic congestion and Wardrop equilibria. Jo. Math. Sci. **181**(6), 792–804 (2012)
8. Chipot, M., Evans, L.C.: Linearization at infinity and Lipschitz estimates for certain problems in calculus of variations. Proc. R. Soc. Edinburgh Sect. A **102**, 291–303 (1986)
9. Clop, A., Giova, R., Hatami, F., Passarelli di Napoli, A.: Very degenerate elliptic equations under almost critical Sobolev regularity. Forum Math. **32**(6), 1515–1537 (2020)
10. Colombo, M., Figalli, A.: An excess-decay result for a class of degenerate elliptic equations. Discret. Contin. Dyn. Syst. Ser. S **7**(4), 631–652 (2014)
11. Colombo, M., Figalli, A.: Regularity results for very degenerate elliptic equations. J. Math. Pures Appl. (9) **101**(1), 94–117 (2014)
12. Cupini, G., Giannetti, F., Giova, R., Passarelli di Napoli, A.: Higher integrability for minimizers of asymptotically convex integrals with discontinuous coefficients. Nonlinear Anal. **154**, 7–24 (2017)
13. Cupini, G., Giannetti, F., Giova, R., Passarelli di Napoli, A.: Regularity results for vectorial minimizers of a class of degenerate convex integrals. J. Differ. Equ. **265**(9), 4375–4416 (2018)
14. Cupini, G., Guidorzi, M., Mascolo, E.: Regularity of minimizers of vectorial integrals with $p - q$ growth. Nonlinear Anal. **54**(4), 591–616 (2003)
15. De Giorgi, E.: Un esempio di estremali discontinue per un problema variazionale di tipo ellitico. Boll. UMI **4**, 135–137 (1968)
16. DiBenedetto, D.: Continuity of weak solutions to certain singular parabolic equations. Ann. Mat. Pura Appl. **130**(4), 131–176 (1982)
17. DiBenedetto, D.: $C^{1+\alpha}$ local regularity of weak solutions of degenerate elliptic equations. Nonlinear Anal. **7**(8), 827–850 (1983)
18. DiBenedetto, E.: Degenerate Parabolic Equation. Universitext. Springer, New York (1993)

19. DiBenedetto, E.: Partial Differential Equations. Cornerstones, 2nd edn. Birkhäuser Boston Ltd, Boston (2010)
20. DiBenedetto, E., Friedman, A.: Hölder estimates for nonlinear degenerate parabolic systems. J. Reine Angew. Math. **357**, 1–22 (1985)
21. Duzaar, F., Grotowski, J.F.: Optimal interior partial regularity for nonlinear elliptic systems: the method of A-harmonic approximation. Manuscripta Math. **103**(3), 267–298 (2000)
22. Duzaar, F., Grotowski, J.F., Kronz, M.: Partial and full boundary regularity for minimizers of functionals with nonquadratic growth. J. Convex Anal. **11**(2), 437–476 (2004)
23. Eleuteri, M., Marcellini, P., Mascolo, E.: Lipschitz estimates for systems with ellipticity conditions at infinity. Ann. Mat. Pura Appl. **195**, 1575–1603 (2016)
24. Fonseca, I., Fusco, N., Marcellini, P.: An existence result for a nonconvex variational problem via regularity. ESAIM Control Optim. Calc. Var. **7**, 69–95 (2002)
25. Foss, M.: Global regularity for almost minimizers of nonconvex variational problems. Ann. Mat. Pura Appl. (4) **187**(2), 263–321 (2008)
26. Foss, M., Passarelli di Napoli, A., Verde, A.: Global Morrey regularity results for asymptotically convex variational problems. Forum Math. **20**, 921–953 (2008)
27. Foss, M., Passarelli di Napoli, A., Verde, A.: Global Lipschitz regularity for almost minimizers of asymptotically convex variational problems. Ann. Mat. Pura Appl. (4) **189**, 127–162 (2010)
28. Giaquinta, M., Modica, G.: Remarks on the regularity of the minimizers of certain degenerate functionals. Manuscripta Math. **57**, 55–99 (1986)
29. Giusti, E.: Direct Methods in the Calculus of Variations. World Scientific Publishing Company, Tuck Link (2003)
30. Leone, C., Passarelli di Napoli, A., Verde, A.: Lipschitz regularity for some asymptotically subquadratic problems. Nonlinear Anal. **67**, 1532–1539 (2007)
31. Liao, N.: On the logarithmic type boundary modulus of continuity for the Stefan problem. (preprint) (2021)
32. Manfredi, J.J.: Regularity for minima of functionals with p-growth. J. Differ. Equ. **76**(2), 203–212 (1988)
33. Marcellini, P.: Everywhere regularity for a class of elliptic systems without growth conditions. Ann. Scuola Norm. Sup. Pisa Cl. Sci.(4) **23**(1), 1–25 (1996)
34. Passarelli di Napoli, A., Verde, A.: A regularity result for asymptotically convex problems with lower order terms. J. Convex Anal. **15**, 131–148 (2008)
35. Raymond, J.P.: Lipschitz regularity for some asymptotically convex problems. Proc. R. Soc. Edinburgh Sect. A **117**, 59–73 (1991)
36. Santambrogio, F., Vespri, V.: Continuity in two dimensions for a very degenerate elliptic equation. Nonlinear Anal. **73**, 3832–3841 (2010)
37. Scheven, C., Schmidt, T.: Asymptotically regular problems. I. Higher integrability. J. Differ. Equ. **248**, 745–791 (2010)
38. Sverák, V., Yan, X.: Non-Lipschitz minimizers of smooth uniformly convex variational integrals. Proc. Natl. Acad. Sci. USA **99**(24), 15269–15276 (2002)
39. Tolksdorf, P.: Everywhere-regularity for some quasilinear systems with a lack of ellipticity. Ann. Mat. Pura Appl. **134**(4), 241–266 (1983)
40. Uhlenbeck, K.: Regularity for a class of nonlinear elliptic systems. Acta Math. **138**, 219–240 (1977)
41. Ural'ceva, N.N.: Degenerate quasilinear elliptic systems. Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **7**, 184–222 (1968)
42. Zatorska-Goldstein, A.: Very weak solutions of nonlinear subelliptic equations. Ann. Acad. Sci. Fenn. Math. **30**(2), 407–436 (2005)