REGULARITY FOR CONVEX VARIATIONAL PROBLEMS ON BD

FRANZ XAVER GMEINEDER

ABSTRACT. In this work we study the regularity properties of minima of convex variational problems on BD, a space which is larger than BV by Ornstein’s Non-Inequality. In essence, we prove that from a Sobolev regularity and partial regularity perspective, such problems lead to the same regularity of minima as the corresponding full gradient functionals on BV. As to the former, we establish the existence of $W^{1,1}_0$-minima under similar conditions known for the BV-theory, covering the limiting case of the area-integrand. As to the latter, we establish a novel Poincaré-type inequality that yields the claimed partial regularity by adapting an approach due to Anzellotti & Giaquinta.

1. INTRODUCTION

A variety of physically relevant convex problems that describe the displacements of bodies subject to external forces are posed in the space BD of functions of bounded deformation, cf. [8, 34] for instance. For a given open set $\Omega \subset \mathbb{R}^n$, this space consists of all $u \in L^1(\Omega; \mathbb{R}^n)$ such that the distributional symmetric gradient $\text{Eu} := \frac{1}{2}(\text{Du} + \text{Du}^T)$ is a finite, matrix-valued Radon measure on $\Omega$. By Ornstein’s Non-Inequality [59, 23, 47, 46] this space is in fact larger than BV, and the full distributional gradients of BD-maps in general do not need to exist as (local) $\mathbb{R}^{n \times n}$-valued Radon measures. As a consequence, the regularity theory for variational problems on BD – cf. (1.1) below – requires a more subtle analysis than those posed on BV. The overarching question here in particular is as to which extent the regularity available for generic full gradient problems on BV can be proven to hold for the respective problems on BD, too. In this sense, the aim of the present paper is to establish that for convex variational problems on BD to be described in more detail below, the regularity of generalised minima is in fact the same provided measured in the right space scale. As we hope, the methods employed to arrive at this aim might also prove useful for other related convex variational problems on BD.

1.1. Aims and scope. Let $\Omega$ be an open and bounded Lipschitz subset of $\mathbb{R}^n$. The aim of this paper is to establish a regularity theory for convex variational problems of the form

\begin{equation}
(1.1) \quad \text{to minimise } F[v] := \int_\Omega f(\varepsilon(v)) \, dx \text{ over a Dirichlet class } \mathcal{D}_{u_0},
\end{equation}

where $\mathcal{D}_{u_0}$ is a suitable Dirichlet class and $\varepsilon(v) := \frac{1}{2}(Dv + Dv^T)$ denotes the symmetric gradient of a map $v: \mathbb{R}^n \to \mathbb{R}^n$. Here, $f: \mathbb{R}^{n \times n} \to \mathbb{R}$ is assumed to be a continuous integrand of linear growth. By the latter, we understand that there exists a constant $L > 0$ such that

\begin{equation}
(\text{LG}) \quad |f(z)| \leq L(1 + |z|) \quad \text{for all } z \in \mathbb{R}^{n \times n}.\n\end{equation}

In view of coerciveness of $F$ this is too bad a bound and one needs to assume, e.g., that there exist constants $c_1, c_2, \gamma > 0$ such that there holds

\begin{equation}
(\text{LG'}) \quad c_1|z| - \gamma \leq f(z) \leq c_2(1 + |z|) \quad \text{for all } z \in \mathbb{R}^{n \times n}.\n\end{equation}

By the linear growth assumption – and as shall be recalled in detail below – such functionals are usually considered on the space BD($\Omega$) of functions of bounded deformation. As a routine consequence, for $F$ to be defined for BD-maps, it must be suitably relaxed. In this respect, we let $u_0 \in \text{LD}(\Omega) := \{v \in L^1(\Omega; \mathbb{R}^n) : \varepsilon(v) \in L^1(\Omega; \mathbb{R}^{n \times n})\}$ and work with the weak*-relaxation (cf. Section 2 for more detail) subject to the Dirichlet constraint $u|_{\partial \Omega} = u_0$ and record the representation (cf. [44, 42], [40, Sec. 5])

\begin{equation}
(1.2) \quad F[u] = \int_{\Omega} f(\varepsilon(u)) \, d\mathcal{L}^n + \int_{\Omega} f^\infty\left(\frac{\text{d}E^u}{\text{d}E^u}\right) \, dE^u u + \int_{\partial \Omega} f^\infty(\text{Tr}(u - u_0) \circ \nu_{\partial \Omega}) \, d\mathcal{H}^{n-1},\n\end{equation}

arXiv:1804.11123v1 [math.AP] 30 Apr 2018
where, for \( u \in \text{BD}(\Omega) \) we denote the Lebesgue-Radon-Nikodym decomposition \( E_u = E^a u + E^s u = \mathcal{E} u \mathcal{L}^n + E^a u \) of \( E_u \) into its absolutely continuous and singular parts for \( \mathcal{L}^n \). Moreover, \( f^\infty(z) := \lim_{t \to 0} f(z/t) \) denotes the recession function of \( f \), capturing the integrand’s behaviour at infinity. Consequently, we call a map \( u \in \text{BD}(\Omega) \) a generalised minimiser if \( F[u] \leq F[v] \) for all \( v \in \text{BD}(\Omega) \), and we denote the set of all generalised minima by \( \text{GM}(F; u_0) \). As a consequence of [40, Sec. 5], generalised minimisers always exist in this framework.

When dealing problems with linear growth, the integrand of primary interest is that of the area integrand and is roughly depicted in Figure 1. By the combination of a lack of Korn inequalities, the typical degeneracies of the integrands and the very weak compactness properties of BD, this task requires a more delicate analysis than corresponding problems on \( W^{1,p} \) or BV and shall be carefully explained now.

1.2. Existence and \( W^{1,1} \)-Regularity of Minima. One of the overall challenges is to prove to which extent the regularity results from the BV-theory are also shared by generalised minima of the relaxed functional. A first step in this direction has been taken by Kristensen and the author in [40] (also cf. the author’s thesis [36]), but was restricted to a fairly small range of strong ellipticity. The first aim of this paper is to show that from a Sobolev and partial regularity perspective, the passage from BV to BD comes with the same regularity as is presently known for full gradient functionals. Here, as pointed out in [40], the essential obstruction is the derivation of estimates for the full second derivatives by ellipticity of \( f \). By Ornstein’s Non-Inequality, we are not allowed to a priori utilise \( L^1 \)-estimates on the full difference quotients. This obstruction led to fractional estimates in [40]. Briefly worded, the first aim of this paper is to give suitable estimates on the full second derivatives of suitable viscosity approximations. This requires a somewhat finer analysis of the Euler-Lagrange equations satisfied by suitable viscosity approximations and is inspired by the foundational work of Seregin [65, 66, 67]. Motivated by Bernstein’s genre [14, 42, 69] and the conditions considered by Ladyzhenskaya & Ural’ceva [53], a natural scale of \( C^2 \)-integrands is given by those \( f : \mathbb{R}^{n \times n} \to \mathbb{R} \) that satisfy for some \( 0 < \lambda \leq \Lambda < \infty \)

\[
(1.4) \quad \lambda \frac{[\xi]^2}{(1 + |z|^2)^\frac{n}{2}} \leq \langle f''(z)\xi, \xi \rangle \leq \Lambda \frac{[\xi]^2}{(1 + |z|^2)^\frac{n}{2}} \quad \text{for all } \xi \in \mathbb{R}^{n \times n}.
\]

If the symmetric gradient of a locally integrable map is a measure, we write \( E_{ua} \) instead of \( e(u) \).
For such integrands, (1.4) precisely describes the degeneration of the ellipticity ratio of $f''$ and has been rediscovered by Bildhauer & Fuchs [16, 17, 18] in the BV-context under the name of $\mu$-ellipticity, where $\mu = a$ in our terminology. Let us remark that the area integrand (1.3) satisfies (1.4) with $a = 3$, cf. [42, Eq. (3.6)]. In turn, we only allow for $a > 1$ as $a = 1$ corresponds to integrands of $L \log L$-growth for which the regularity theory traces back to SerEgin and collaborateurs [33, 34].

Now, in the BV-case, it is known from [17] that if $a \leq 3$ and a suitable viscosity approximation sequence remains locally bounded, then there exists one generalised minimiser which has full gradient in $L \log L_{\text{loc}}$. By possible non-uniqueness of generalised minima, this only applies to one particular candidate. Note that the positively 1-homogeneous recession function $f \infty$ in (1.2) acting on the singular part of the relevant measure might lead to non-uniqueness of generalised minima. However, due to the approach of Beck & Schmidt [19] based on Ekeland’s variational principle, it is known that $a \leq 3$ implies that $D^1 u \equiv 0$ in the interior for all generalised minimisers. As mentioned above, this strategy was extended to convex functionals on BD in [40] and led to $\text{GM}(F ; u_0) \subset W^{1,1}_{\text{loc}}(\Omega; \mathbb{R}^n)$ for if $1 < a < 1 + \frac{2}{n}$ by virtue of weighted Nikolskii-estimates. As a main drawback, these estimates rely on perturbed Euler-Lagrange equations – essentially differential inequalities – which make it difficult to utilise the little but nevertheless available crucial identities satisfied by generalised minimisers. Here we do not focus so much on uniqueness but on the existence of suitably regular generalised minima. In this respect, our first main result reads as follows:

**Theorem 1.1** (Existence of $W^{1,1}_{\text{loc}}$-regular minimisers). If $f \in C^2(\mathbb{R}^{n \times n})$ satisfies ($LG'$) and (1.4) with

(i) $1 < a < 1 + \frac{2}{n}$, then there exists a generalised minimiser of class $\text{LD}(\Omega) \cap W^{1,p}_{\text{loc}}(\Omega; \mathbb{R}^n)$ for some $1 < p < \infty$.

(ii) $1 + \frac{2}{n} \leq a \leq 3$ and a suitable sequence of viscosity approximations remains locally bounded, then there exists a generalised minimiser of class $\text{LD}(\Omega) \cap W^{1,L \log L}_{\text{loc}}(\Omega; \mathbb{R}^n)$.

Let us note that the condition appearing in (ii) of the previous theorem – which will be made precise in Section 3 below, cf. (3.13) – is also required for the corresponding result to hold for the respective problem on BV, cf. [17, 18, 19]. Moreover, let us note that even for the Dirichlet problem on BV presently no Sobolev regularity result is available if $a > 3$; the only unconditional $W^{1,1}_{\text{loc}}$-result concerns the Neumann problem on BV and has been established by Beck, Buliček and the author [12]. For further reference we remark, however, that in the context of $x$-dependent $C^2$-integrands, $a > 3$ is not sufficient in the BV-Dirichlet case to yield generalised minima with vanishing singular part; cf. [18, Chpt. 4.4].

We wish to stress that the preceding theorem is in fact independent from the prequel [40] and thereby complements the results obtained therein; indeed, Theorem 1.1 only addresses the existence of suitably regular generalised minima, whereas [40, Thm. 1.2] covers all generalised minimisers for if $1 < a < 1 + \frac{2}{n}$. Moreover, the second main result in [40] which appeals to the existence of one generalised minimiser subject to a so-called uniform local $\text{BMO}$ condition – a substitute for (3.13) below – turns out to work unconditionally by Theorem 1.1, but the resulting higher integrability in [40] is slightly better than the one that shall follow from Theorem 1.1.

### 1.3. Partial Regularity

The second part of this paper is devoted to the partial (Hölder) regularity of generalised minima of $F$. We note that, essentially because of the fact that the minimisation of $F$ constitutes a vectorial problem, full Hölder continuity in general is not to be expected; see [35, 43, 11, 56, 57] and the references therein. To streamline terminology, in this paper we say that a map $v \in L^1_{\text{loc}}(\Omega; \mathbb{R}^n)$ is partially regular if there exists a relatively open subset $\Omega_\varepsilon \subset \Omega$ such that $v$ is of class $C^{1,\alpha}_{\text{loc}}$ in a neighbourhood of any of the elements of $\Omega_\varepsilon$ for any $0 < \alpha < 1$. There is an abundant literature on the topic of partial regularity and proof strategies, most notably the (indirect) blow-up method with roots in De Giorgi’s work [25] and the $A$-harmonic approximation method with roots in Almgren’s and Allard’s work in geometric measure theory [4, 3]. These proof strategies have been adapted to the setting of functionals of the type (1.1) with $\varepsilon$ replaced by the full gradient, see [1, 2, 31, 28, 29, 30, 55, 27] for an incomplete list. For instance, even in the convex full-gradient linear case, indirect methods such as blow-up are difficult to implement by the relatively weak compactness properties of BV as long as no additional Sobolev regularity is available. On the other hand, from an application perspective, a plenty of variational integrands of linear growth degenerates...
and coincides with an affine-linear function for large arguments. Then ellipticity of \( f \) is lost for large arguments – and hence higher Sobolev regularity – but partial regularity still survives provided the gradients of minima are assumed small. In this sense, a (local) boost from \( W^{1,\infty} \) to \( C^{1, \alpha} \)-regularity might take place.

Overall, up to date the only partial regularity result available for convex linear growth functionals is due to Anzellotti & Giaquinta \cite{AG} (also see the related result of Schmidt \cite{S} for the model integrands \( f(\cdot) = (1 + |\cdot|^p)^{1/p} \)). In the terminology of \cite{GS}, this is a *local-in-phase-space* partial regularity result for a BV-generalised minimiser \( u \in BV \): If \( x_0 \in \Omega \) is a Lebesgue point for \( Du \) (with Lebesgue value \( z \)) and \( f''(z) \) exists and is positive definite, then \( u \) is of class \( C^{1, \alpha} \) in a neighbourhood of \( x_0 \). Relating to the discussion from above, it is thus natural to ask for a generalisation of the above result to BD and convex variational problems \((1.1)\). Here, the chief obstruction stems from the strategy of \cite{AG} which is based on comparing generalised minima with of suitable mollifications, cf. Section 4 for more detail. As a main tool, we utilise a new convolution inequality for BD-maps (cf. Proposition 4.3) to establish the following result which, in turn, seems to be the first of its kind:

**Theorem 1.2 (Local-in-phase-space regularity).** Let \( f \in C^2(\mathbb{R}^n) \) be convex and of linear growth in the sense of \((LG)\). Let \( u \in GM(F; u_0) \). If \((x, z) \in \Omega \times \mathbb{R}^n\) is such that

\[
\lim_{R \to 0} \int_{B(x, R)} |\delta^2 u - z| \, dx + \frac{E^u_\Omega(B(x, R))}{Z^n(B(x, R))} = 0
\]

and \( f''(z) \) is positive definite, then there holds \( u \in C^{1, \alpha}(U; \mathbb{R}^n) \) for a suitable neighbourhood \( U \) of \( x \) for all \( 0 < \alpha < 1 \).

In consequence, we exemplarily obtain the conclusion that if \( f''(z) > 0 \) in the sense of positivity of bilinear forms for all \( |z| \leq M \) and \( u \in GM(F; u_0) \) satisfies \( |\varepsilon(u)| \leq M \), then automatically \( Du \in C^{0, \alpha}(\Omega_u; \mathbb{R}^{n \times n}) \) on an open set \( \Omega_u \subset \Omega \) of full Lebesgue measure. On the other hand, referring to Figure 1, if \( f''(z) > 0 \) globally, then any generalised minimiser is partially is partially regular. Previously, such results were only known *under quantified versions of global strong convexity*, cf. \cite{H}. Again, let us emphasize that the key here is that *no* sort of uniform convexity is required. If, for instance, minima were known to have higher order derivatives, then the blow-up method would certainly do, but our overarching aim is to cope with the degenerate cases also.

To keep our exposition at a reasonable length and to focus on the chief difficulties, let us note that the preceding theorem is established to \cite{AG} once the key estimate of Proposition 4.1 below is established (also see the discussion afterwards). In Section 4, where Theorem 1.2 is proved, we hence confine to the proof of Proposition 4.1 and refer the interested reader to \cite{AG, H} for the easy modification of the remaining steps. Also, the approach is robust enough to apply to non-autonomous functionals as well, and this is discussed in Section 5 together with other generalisations.

### 1.4. Organisation of the Paper

Let us finally explain the structure of the paper. In Section 2 we fix notation, record basic definitions – with emphasis on functions of measures and function spaces – and auxiliary estimates. In Section 3 we provide the proof of Theorem 1.1 and discuss selected implications, in particular, the merger of Theorem 1.1 with the uniqueness results obtained in the prequel \cite{G} of this paper. In Section 4, we provide the proof of Theorem 1.2 and then conclude the paper with extensions of the results gathered so far in Section 5.

**Acknowledgments.** The results presented in this paper are partially obtained (and partially extend those) in the author’s thesis \cite{H}. As such, he is indebted to his former advisor Jan Kristensen for proposing the partial regularity for convex functionals on BD, and to Gianni Dal Maso and Gregory Seregin for acting as referees for the author’s thesis and thereby making valuable suggestions, which eventually lead to the Sobolev regularity improvement compared to \cite{H, G}. Financial support through the HCM Bonn is gratefully acknowledged.

## 2. Preliminaries

### 2.1. General Notation and Background

We briefly comment on the notation used throughout. All finite dimensional vector spaces are equipped with the euclidean (or, in the matrix case, Frobenius) norm, and the inner product on such spaces is denoted \( \langle \cdot, \cdot \rangle \). Given \( a, b \in \mathbb{R}^n \), the symmetric tensor
product is given by $a \otimes b := \frac{1}{2}(ab^T + ba^T)$. Given $x_0 \in \mathbb{R}^n$ and $r > 0$, the open ball of radius $r > 0$ centered at $x_0$ in $\mathbb{R}^n$ is denoted $B(x_0, r) := \{ x \in \mathbb{R}^n : |x - x_0| < r \}$. To distinguish from balls in matrix space, we write $B(z, r) := \{ y \in \mathbb{R}_{\text{sym}}^n : |y - z| < r \}$ for $z \in \mathbb{R}_{\text{sym}}^n$ and $r > 0$. The $n$-dimensional Lebesgue and $(n - 1)$-dimensional Hausdorff measure are denoted $\mathcal{L}^n$ and $\mathcal{H}^{n-1}$, respectively. Accordingly, the Hausdorff dimension of a Borel set $A \in \mathcal{B}(\mathbb{R}^n)$ is denoted $\dim_h(A)$.

For $u \in L_{\text{loc}}^1$ and an open set $U \subset \mathbb{R}^n$, we use the shorthand $(u)_U := \int_U u \, dx := \mathcal{L}^n(U)^{-1} \int_U u \, dx$ whereas, if $U = B(x, r)$ is ball, we abbreviate $(u)_{x,r} := (u)_{B(x,r)}$. Cubes in $\mathbb{R}^n$ are tacitly assumed to be non-degenerate, and we denote by $\ell(Q)$ their sidelengths. Moreover, for a given finite dimensional real vector space $V$, we denote $\mathcal{M}(\Omega; V)$ the $V$-valued finite Radon measures on $\Omega$. For $\mu \in \mathcal{M}(\Omega; V)$, its Lebesgue-Radon-Nikodým decomposition is given by $\mu = \mu^\alpha + \frac{du}{\mu^\alpha} |\mu^*|$, where $\mu^* \ll \mathcal{L}^n$ and $\mu^* \perp \mathcal{L}^n$. Given a function $f : V \to \mathbb{R}$ of linear growth (i.e., $f$ satisfies $(\text{LG})$), its recession function $f^\infty$ is defined as after $(1.2)$ with the obvious modifications. The corresponding measure $f[\mu](A)$ for Borel sets $A \in \mathcal{B}(\Omega)$ then is given by

\begin{equation}
(2.1) \quad f[\mu](A) := \int_A f(\mu) := \int_A f\left(\frac{\mu}{\mathcal{L}^n}\right) \, d\mathcal{L}^n + \int_A f^\infty\left(\frac{\mu}{\mu^*}\right) \, d\mu^*,
\end{equation}

and we refer the reader to [60, 26, 7, 6] for more background information on applying convex functions to measures. By $c, C > 0$ we denote generic constants whose value might change from line to line, and shall only be specified if their precise value is required. Finally, the symbols $A \lesssim B$ means that $A \leq cB$ with $c > 0$ not depending on $A$ or $B$, and we write $A \approx B$ if $A \lesssim B \lesssim A$.

2.2. Function Spaces and Integral Operators. Let $\Omega \subset \mathbb{R}^n$ be open and bounded. We then define $\text{BD}(\Omega)$ as the space of all $u \in L^1(\Omega; \mathbb{R}^n)$ for which their total deformation

\begin{equation}
(2.2) \quad |Eu|(\Omega) := \sup \left\{ \int_{\Omega} \langle \nu, \text{div}(\varphi) \rangle \, dx : \varphi \in C_c^1(\Omega; \mathbb{R}_{\text{sym}}^n), \|\varphi\|_{L^\infty(\Omega; \mathbb{R}_{\text{sym}}^n)} \leq 1 \right\}
\end{equation}

is finite; note that by writing $Eu$ we indicate that the symmetric gradient of $u$ is a measure whereas by $\varepsilon(u)$ we tacitly understand that it is representable by an $L^1$-map. This space has been introduced in [24, 73] and studied from various perspectives in [8, 5, 71, 10]; until otherwise stated, all of the following can be traced back to these references. Let $u, u_1, u_2, \ldots \in \text{BD}(\Omega)$. We say that $u_k \overset{\ast}{\rightharpoonup} u$ if and only if $u_k \to u$ in $L^1(\Omega; \mathbb{R}^n)$ and $\text{En}_u \overset{\ast}{\rightharpoonup} Eu$ in $\mathcal{M}(\Omega; \mathbb{R}_{\text{sym}}^n)$. If $u_k \overset{\ast}{\rightharpoonup} u$ as just defined and $|\text{En}_{u_k}|(\Omega) \to |\text{En}_u|(\Omega)$, then we say that $(u_k)$ converges area-strictly to $u$. If, moreover, $\sqrt{1 + |\text{En}_u|^2}(\Omega) \to \sqrt{1 + |\text{En}_u|^2}(\Omega)$ in the sense explained above (cf. $(2.1)$), then we say that $(u_k)$ converges area-strictly to $u$. By smooth approximation, $C^\infty(\Omega; \mathbb{R}^n) \cap \text{LD}(\Omega)$ is dense in $\text{BD}(\Omega)$ with respect to strict convergence.

Given $u \in \text{BD}(\Omega)$, the Lebesgue-Radon-Nikodým decomposition of $Eu$ reads as $Eu = E^\alpha u + E^\varepsilon u = \varepsilon u \mathcal{L}^n|\Omega + |(\nabla u)|^2 + |E^\alpha u|. Here, $\varepsilon u$ takes the rôle of the symmetric part of the approximate gradient (cf. [6]) for this terminology.

Now let $\Omega$ have Lipschitz boundary. Both $\text{LD}(\Omega)$ and $\text{BD}(\Omega)$ then have trace space $L^1(\partial\Omega; \mathbb{R}^n)$. Note that the trace operator onto $L^1(\partial\Omega; \mathbb{R}^n)$ is continuous with respect to weak convergence on $\text{LD}(\Omega)$ whereas it is not with respect to weak*-convergence in $\text{BD}(\Omega)$. In the latter case, continuity can only be achieved when $\text{BD}(\Omega)$ is equipped with strict convergence. Moreover, as $\Omega$ has Lipschitz boundary, any $u \in \text{BD}(\Omega)$ can be extended by zero to the entire $\mathbb{R}^n$; this extension $\overline{u}$ again belongs to $\text{BD}(\mathbb{R}^n)$ and we have

\begin{equation}
\overline{Eu} = Eu \mathcal{L}^n|\Omega + \text{Tr}(u) \otimes \nu_{\partial\Omega} \mathcal{H}^{n-1}|\partial\Omega,
\end{equation}

where $\nu_{\partial\Omega}$ is the outward unit norm to $\partial\Omega$. Also, we have the Gauß-Green formula

\begin{equation}
(2.3) \quad \int_{\Omega} \langle \varphi, Eu \rangle + \int_{\Omega} \langle \text{div}(\varphi), u \rangle \, dx = \int_{\partial\Omega} \langle \varphi, \text{Tr}(u) \otimes \nu_{\partial\Omega} \rangle \, d\mathcal{H}^{n-1}
\end{equation}

for all $u \in \text{BD}(\Omega)$ and all $\varphi \in C(\overline{\Omega}; \mathbb{R}_{\text{sym}}^n) \cap C^1(\Omega; \mathbb{R}_{\text{sym}}^n)$; here, $\text{div}$ denotes the row-wise divergence.
Next, let us recall the SMITH representation formula, cf. [70, 45]. If $u \in C_0^\infty(\mathbb{R}^n; \mathbb{R}^n)$, then $u = (u^1, \ldots, u^n)$ can be retrieved from $\varepsilon(u) = (\varepsilon_{ij}(u))_{i,j=1}^n$ via

\begin{equation}
\varepsilon(u)^k = \frac{2}{n \omega_n} \sum_{i,j \in \{1, \ldots, n\}} \varepsilon_{ik}(u) \varepsilon_{jk}(u) \partial_i K_{ij} - \varepsilon_{ij}(u) \partial_k K_{ij} + \varepsilon_{ki} \partial_j K_{ij}, \quad k \in \{1, \ldots, n\},
\end{equation}

where $K_{ij}(x) := x_i x_j / |x|^n$ for $x \in \mathbb{R}^n \setminus \{0\}$. This representation readily implies that the map $\Phi: \varepsilon(u) \mapsto Du$ is given by a singular integral operator of convolution type (satisfying the usual Hörmander condition). This, in turn, indicates the failure of Korn’s inequality in the $L^1$-setup. Also, the following lemma is based on (2.4):

**Lemma 2.1.** Let $\Omega \subset \mathbb{R}^n$ be open and let $v, u_1, u_2, \ldots \in L^1_{\text{loc}}(\Omega; \mathbb{R}^n)$ be given.

(a) If $(u_k) \subset \text{LD}(\Omega)$ and there exists $\beta > 0$ such that for every relatively compact $K \Subset \Omega$ there holds $\sup_{k\in \mathbb{N}} \|\varepsilon(u_k)\|_{L^{1+\alpha}(\Omega; \mathbb{R}^{n \times n})} < \infty$, then there exists $u \in W^{1,1}_{\text{loc}}(\Omega; \mathbb{R}^n)$ such that for every relatively compact Lipschitz subset $K' \Subset \Omega$ there holds $u_k |_{K'} \to u |_{K'}$ weakly in $W^{1,1}(K'; \mathbb{R}^n)$ as $k \to \infty$.

(b) If $u$ satisfies $\varepsilon(u) \in L^p_{\text{loc}}(\Omega; \mathbb{R}^{n \times n})$ for some $1 < p < \infty$, then there exists a constant $c > 0$ which only depends on $p$ such that for all balls $B(x_0, r) \Subset \Omega$ there holds

\[
\int_{B(x_0, r)} |Du - (Du)_{B(x_0, r)}|^p \, dx \leq c \int_{B(x_0, r)} |\varepsilon(u) - (\varepsilon(u))_{B(x_0, r)}|^p \, dx.
\]

**Proof.** Ad (a). In the following, we will utilise some terminology that is presented in the appendix, Section 6.3. Given a Young function $A: \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, $\|\cdot\|_{L^A}$ denotes the corresponding Luxembourg norm, and this notation particularly applies to $A(t) := t^{1+\alpha}(1 + t)$ for $\alpha \geq 0$. We then denote $E_0^\prime L^{1+\alpha}_A(\Omega; \mathbb{R}^n)$ the closure of $C_0^\infty(\mathbb{R}^n; \mathbb{R}^n)$ with respect to $\|\varepsilon(u)\|_{L^{1+\alpha}_A(\Omega; \mathbb{R}^{n \times n})}$. Relying on (2.4), CIANCHE proved in the recent work [22, Ex. 3.8] that if $\alpha \geq 0$, then there exists a constant $c = c(\alpha, n) > 0$ such that for all $u \in E_0^\prime L^{1+\alpha}(\Omega; \mathbb{R}^n)$ there holds $\|Du\|_{L^{1+\alpha}_A(\Omega; \mathbb{R}^{n \times n})} \leq c\|\varepsilon(u)\|_{L^{1+\alpha}_A(\Omega; \mathbb{R}^{n \times n})}$. Now let $B \subset \Omega$ be a ball with $\text{dist}(B; \partial\Omega) > 0$ and choose a cut-off function $\rho \in C_c^\infty(\Omega; [0, 1])$ with $\rho_B \leq \rho$. Put $u_k := \rho_{uk}$; then CIANCHE’s inequality implies that $(Du_k)$ is uniformly bounded in $L^{1+\alpha}_A$, and since $\alpha > 0$, it must possess a weakly converging subsequence as a consequence of the DeVallée-Poussin criterion. Upon localisation, (a) now follows. Ad (b). This can be conveniently traced back to [20].

In the sequel, we augment (2.4) by a decomposition result due to RESHETNYAK [61]. To this end, let us remind the reader of the space of rigid deformations

\begin{equation}
\mathcal{R}(\Omega) := \{u: \Omega \to \mathbb{R}^n: u(x) = Ax + b, \ A \in \mathbb{R}^{n \times n}, b \in \mathbb{R}^n\}
\end{equation}

which, for connected $\Omega$, is precisely the nullspace of $\varepsilon$. Also note that, since elements of $\mathcal{R}(\Omega)$ are polynomials, we shall often identify $\mathcal{R}(\Omega)$ with $\mathcal{R}(\mathbb{R}^n)$. In the original version, this next lemma is stated in slightly higher generality, but we confine to the case of cubes.

**Lemma 2.2 (RESHETNYAK, [61]).** For each $n \in \mathbb{N}$ there exists a constant $c = c(n) > 0$ with the following property: If $Q$ is an open, non-empty cube in $\mathbb{R}^n$, there exists an projection $\Pi_Q: C^\infty(\mathbb{Q}; \mathbb{R}^n) \cap \text{LD}(Q) \to \mathcal{R}(Q)$ and an integral operator $T_Q: (C^\infty \cap L^1)(Q; \mathbb{R}^{n \times n}) \to L^1(Q; \mathbb{R}^n)$ such that for any $v \in \text{LD}(Q)$ there holds

\begin{equation}
v(x) = (\Pi_Q(v))(x) + T_Q[\varepsilon(v)](x) \quad \text{for all } x \in Q.
\end{equation}

Moreover, the operator $T_Q$ is of the form

\begin{equation}
T_Q[\varepsilon(v)](x) = \int_Q R_Q(x, y) \varepsilon(v)(y) \, dy,
\end{equation}

where $R_Q: Q \times Q \to \mathcal{R}(\mathbb{R}^{n \times n}; \mathbb{R}^n)$ satisfies $|R_Q(x, y)| \leq C_R/|x - y|^{n-1}$ for all $x, y \in Q$ with $C_R > 0$ independent of $Q$. Moreover, the integral kernel $R$ can be obtained by shifting and rescaling that for the unit cube.
Proof. In [61] the argument is carried out for sets \( U \subset \mathbb{R}^n \) being starshaped with respect to a ball. Here we consider all cubes \( Q \) to be starshaped with respect to a ball \( B_Q \) having the same center as the cube and with radius \( \frac{\text{dist}(Q, \partial D)}{100} \); in this way, the corresponding estimates will scale correctly. By [61, Eqs. (2.38), (2.39) ff.], we can write \( v = (v^1, \ldots, v^n) \in (C^\infty \cap W^{1,p})(Q; \mathbb{R}^n) \) as (2.6) and we have the explicit representation
\[
(T_Q[e(v)])^i(x) = \int_Q \sum_{k,l=1}^n \varepsilon_{kl}(v) R^{(i)}_{kl,Q}(x,y) \, dy \quad x \in Q,
\]
and \( R^{(i)}_{kl,Q} = \sum_{k,l=1}^n \varepsilon_{kl}(v) R^{(i)}_{kl,Q} \) satisfies for a constant \( C = C(Q, B_Q) \geq 0, |R_Q(x,y)| \leq C/|x-y|^{n-1} \) for all \( x,y \in Q \). By our choice of the ball \( B_Q \) with respect to which \( Q \) is starshaped, it is seen that \( C > 0 \) actually can be chosen independently of \( Q \), and even though (2.6) is stated for \( C^\infty \cap W^{1,1} \)-maps, the above bound on \( R_Q \) shows that it equally remains valid for \( v \in C^\infty(Q; \mathbb{R}^n) \cap LD(Q) \).

2.3. Auxiliary Estimates. We now collect some auxiliary estimates on the reference integrands to be dealt with later. To this end, we define for \( t \in \mathbb{R} \) the function \( e(t) := (1 + t^2)^{-\frac{1}{2}} - 1 \). If \( z \in \mathbb{R}^m \) for some \( m \in \mathbb{N} \), then we tacitly understand \( e(z) := (1 + |z|^2)^{-\frac{1}{2}} - 1 \), but no ambiguities will arise from this. The functions \( e \) will help to define our excess quantity later on, and we record from [9] the following

Lemma 2.3 ([9, Prop. 2.5]). The auxiliary integrand \( e: \mathbb{R}^m \to \mathbb{R} \) satisfies the following properties:

(i) There exists a positive constant \( c = c(p) \) such that for all \( z_1, z_2 \in \mathbb{R}^m \) and all \( t > 0 \) there holds \( e(z_1) \leq c \max\{t, t^2\} e(z_1) \) and \( e(z_1 + z_2) \leq (c e(z_1) + e(z_2)) \).

(ii) Given \( L > 0 \), there exist positive constants \( C_1 = C_1(L) \) and \( C_2 = C_2(L) \) such that for all \( z \in \mathbb{R}^m \) with \( |z| \leq L \) there holds \( C_1 |z|^2 \leq e(z) \leq C_2 |z|^2 \).

(iii) There exists a constant \( \ell = \ell(m) > 0 \) such that for all \( z \in \mathbb{R}^m \) there holds \( \ell \min\{|z|, |z|^2\} \leq e(z) \leq \ell^{-1} \min\{|z|, |z|^2\} \).

On the other hand, raising the inequality \( |t| \leq (1 + |t|^2)^{\frac{1}{2}} \) to the \( 2\frac{m}{m-2}p \)-th power yields for \( 1 < a < 2 \):

Lemma 2.4. For \( 1 < a < 2 \), define \( V_a(t) := (1 + t^2)^{\frac{a}{2}} \) for \( t \in \mathbb{R} \). Let \( U \subset \mathbb{R}^n \) be open and bounded. If \( u \in L^1(U; \mathbb{R}^m) \) satisfies \( V_a(u) \in L^p(U; \mathbb{R}^m) \) for some \( 1 \leq p < \infty \), then there holds \( u \in L^\infty(U; \mathbb{R}^m) \) with \( \|u\|_{L^\infty(U; \mathbb{R}^m)} \leq \max\{1, 2\frac{m}{m-2}p\} \).

Next, some implications of \( a \)-ellipticity; albeit stated in [17] for the gradient case, these inherit in a straightforward manner to the case of symmetric matrices.

Lemma 2.5 ([42, 17]). Suppose that \( f \in C^2(\mathbb{R}^{n \times n}_\text{sym}) \) satisfies (LG) and (1.4). Then the following hold:

(a) There exist \( \vartheta, \theta > 0 \) such that \( \langle f'(\xi), \xi \rangle \geq \vartheta E(\xi) - \theta \) holds for all \( \xi \in \mathbb{R}^{n \times n}_\text{sym} \).

(b) There exists \( L > 0 \) such that \( |f'| \leq L \).

The reader will notice that by convexity and linear growth, (b) of the preceding lemma remains valid for the respective \( C^1 \)-integrands.

3. Local \( W^{1,1} \)-Regularity

The purpose of this section is to prove the \( W^{1,1} \)-regularity results asserted by Theorem 1.1. Here we adapt the vanishing viscosity approach as employed first by Seregin [66, 68]. This will lead us to weighted second order estimates, and in the final subsection 3.3 we shall draw some further conclusions of Theorem 1.1 that will be of independent interest in the subsequent Section 4.

3.1. Viscosity Approximations. In what follows, let \( u_0 \in LD(\Omega) \). We employ a two-layer approximation scheme to establish the regularity assertions of Theorem 1.1. In a first step, we utilise the Lipschitz regularity of \( \partial D \) to find a sequence \( (u_l^{(0)}) \subset W^{1,2}(\Omega; \mathbb{R}^n) \) such that \( \|u_0 - u_l^{(0)}\|_{LD(\Omega)} \to 0 \), \( l \to \infty \). This can be arranged in a way such that \( \|u_l^{(0)}\|_{W^{1,2}(\Omega; \mathbb{R}^n)} \leq \alpha(l) \) as \( l \to \infty \) but is independent of \( u_0 \). In a second step, we then fix \( l \in \mathbb{N} \) and consider for \( j \in \mathbb{N} \) the
auxiliary variational problem of finding $v_{l,j} \in \mathcal{D}_l := u_l^{\beta_0} + W^{1,2}_0(\Omega; \mathbb{R}^n)$ that minimises the stabilised functional

$$
F_{l,j}[v] := F_{l,j}[v; \Omega] := \int_{\Omega} f(\varepsilon(v)) \, dx + \frac{1}{2\alpha(l,j)} \int_{\Omega} |\varepsilon(v)|^2 \, dx \quad \text{over } v \in \mathcal{D}_l.
$$

It is clear that by convexity, each $F_{l,j}$ possesses a unique minimiser $v_{l,j} \in \mathcal{D}_l$. We next record some easy properties and a priori-estimates for the double sequence $(v_{l,j})$. To this end, it is customary to introduce the notation

$$
f_{l,j} := f + \frac{1}{2\alpha(l,j)} |\cdot|^2 \quad \text{and} \quad \sigma_{l,j} := f'_{l,j}(\varepsilon(v_{l,j})).
$$

**Lemma 3.1.** Let $f \in C^2(\mathbb{R}^{n \times n})$ satisfy (LG') and have positive definite and bounded Hessian everywhere. Then the double sequence $(v_{l,j})$ satisfies the following:

(a) $\sup_{l \in \mathbb{N}} \sup_{j \in \mathbb{N}} \|v_{l,j}\|_{L^1(\Omega)} < \infty$ and $\sup_{l \in \mathbb{N}} \sup_{j \in \mathbb{N}} \|\alpha(l,j)^{-1}\|\|\varepsilon(v_{l,j})\|^{L^2(\Omega; \mathbb{R}^{n \times n})} < \infty$.

(b) For every $(l, j) \in \mathbb{N}$ we have $\text{div}(\sigma_{l,j}) \equiv 0$ in the sense of distributions on $\Omega$.

(c) For every $(l, j) \in \mathbb{N}$ we have $v_{l,j} \in W^{2,2}_0(\Omega; \mathbb{R}^n)$.

(d) For every $l \in \mathbb{N}$, a suitable subsequence of $(v_{l,j})_j$ converges weakly to some $v_l \in \text{BD}(\Omega)$.

Also, a suitable subsequence $(v_l)$ converges to some $u \in \text{BD}(\Omega)$ in the weak*-sense which is a generalised minimiser: $u \in \text{GM}(F; u_0)$.

**Proof.** Ad (a). For each $l \in \mathbb{N}$, $u_l^{\beta_0}$ is an admissible for the minimisation of $F_l$ over $\mathcal{D}_l$. Hence, we obtain by use of (LG') and $\|u_l^{\beta_0} - u_0\|_{L^1(\Omega)} \to 0$ as $l \to \infty$

$$
F_{l,j}[v_{l,j}] \leq F_{l,j}[u_l^{\beta_0}] = \int_{\Omega} f(\varepsilon(u_l^{\beta_0})) \, dx + \frac{1}{2j} \|u_0\|_{L^2(\Omega)}^2 
\leq c_2 \int_{\Omega} |\varepsilon(u_l^{\beta_0})| \, dx + (c_2 \mathcal{L}^n(\Omega) + 1) + C \|u_0\|_{L^2(\Omega)}^2 
\leq c_2 \|u_l\|_{L^1(\Omega)} + \|u_0\|_{L^2(\Omega)}^2 + c_3 \leq C,
$$

and now the lower bound provided by (LG') yields that $(\varepsilon(v_{l,j}))$ is uniformly bounded in $L^1(\Omega; \mathbb{R}^{n \times n})$. On the other hand, as $v_{l,j} - u_l^{\beta_0} \in L^1(\Omega)$, we obtain by Poincaré’s inequality on $L^1(\Omega)$

$$
\|v_{l,j}\|_{L^1(\Omega; \mathbb{R}^n)} \leq \|u_l^{\beta_0}\|_{L^1(\Omega; \mathbb{R}^n)} + C \|\varepsilon(v_{l,j} - u_l^{\beta_0})\|_{L^1(\Omega; \mathbb{R}^{n \times n})} \leq C \|u_l^{\beta_0}\|_{L^2(\Omega)} + C
$$

by the first part, and since $\sup_{l \in \mathbb{N}} \|u_l^{\beta_0}\|_{L^2(\Omega)} < \infty$, the proof of (a) is complete. Ad (b). We only sketch the argument: Let $B \subset \Omega$ be an arbitrary but fixed ball and let $\rho \in C^2_B(\Omega; [0,1])$ satisfy $\rho|_{\partial B} \equiv 1$. Note that for each $(l, j) \in \mathbb{N}^2$, $v_{l,j}$ satisfies the Euler-Lagrange equation

$$
\int_{\Omega} \langle \sigma_{l,j}, \varepsilon(\varphi) \rangle \, dx = 0 \quad \text{for all } \varphi \in W^{1,2}_0(\Omega; \mathbb{R}^n).
$$

Then choosing $\varphi := \Delta_{s,h}^{-} (\rho^2 \Delta_{s,h} v_j)$, where $\Delta_{s,h}$ and $\Delta_{s,h}^{+}$ are the forward and backward difference quotients in direction $s \in \{1, \ldots, n\}$ of stepsize $h$, respectively, $\varphi$ is admissible in (3.3). Discrete integration by parts for $\Delta_{s,h}$, positive definiteness of $f''$ then yields by Young’s inequality and absorbing suitably in a standard manner

$$
\frac{1}{\alpha(l,j)} \int_{\Omega} |\rho \varepsilon(v_{l,j})| \, dx \leq C \int_{\Omega} \langle f_{l,j}'(\varepsilon(v_{l,j}))(D \varphi \ast v_{l,j}), (D \varphi \ast v_{l,j}) \rangle \, dx 
\leq C \int_{\Omega} \|f_{l,j}'(\varepsilon(v_{l,j}))(D \varphi \ast v_{l,j})\|_{L^1(\Omega)} \, dx \leq C(\varphi, l, j)
$$

as $f_{l,j}'$ is bounded independently of $l, j$. In consequence, $(\varepsilon(\varphi(v_{l,j})))$ is locally uniformly bounded in $L^2$ and so the claim of (c) is a consequence of Korn’s inequality. Ad (d). The claimed convergence to some $u \in \text{BD}(\Omega)$ is an immediate consequence of (a), and hence it suffices to prove minimality. Relabelling if necessary, we can assume without loss of generality that for each $l \in \mathbb{N}$, $v_{l,j} \rightharpoonup v_l$ as $j \to \infty$ in $\text{BD}(\Omega)$. Then $v_l \in \text{BD}(\Omega)$ is a generalised minimiser of $F$ over $\mathcal{D}_l$, that is, we have $\mathcal{F}_l[u_l] \leq \mathcal{F}_l[w]$ for all $w \in \text{BD}(\Omega)$, where $\mathcal{F}_l := \mathcal{F}_l^{\beta_0}$ as adapted from (1.2). This follows, e.g., by adapting the approach of SEREGIN [65] as is done in the author’s thesis [38, Thm. 5.10] to which the interested
reader is referred to. To conclude, let again \( w \in BD(\Omega) \) be arbitrary. Then, by weak*-convergence of \((v_l)\) to \( u \in BD(\Omega)\), we have by Reshetnyak’s lower semicontinuity theorem in the first, minimality of \( v_l \) for \( \mathcal{F} \) over \( \mathcal{G}_l \) in the second and Lipschitz continuity of \( f^\infty \) in the third step

\[
\mathcal{F}[u] - \mathcal{F}[w] \leq \liminf_{l \to \infty} \int_{\Omega} f(Ev_l) - f(Ew) + \int_{\partial \Omega} f^\infty(\text{Tr}(u_l - v_l) \otimes \nu_{\partial \Omega}) \, d\mathcal{H}^{n-1} \\
\leq \liminf_{l \to \infty} \left( \int_{\Omega} f(Ev_l) + \int_{\partial \Omega} f^\infty(\text{Tr}(u_l^{\partial \Omega} - v_l) \otimes \nu_{\partial \Omega}) \, d\mathcal{H}^{n-1} \right) \\
- \int_{\Omega} f(Ew) - \int_{\partial \Omega} f^\infty(\text{Tr}(u_l^{\partial \Omega} - w) \otimes \nu_{\partial \Omega}) \, d\mathcal{H}^{n-1} \\
+ \int_{\Omega} \|f^\infty(\text{Tr}(u_l^{\partial \Omega} - v_l) \otimes \nu_{\partial \Omega}) - f^\infty(\text{Tr}(u_l^{\partial \Omega} - v_l) \otimes \nu_{\partial \Omega})\| \, d\mathcal{H}^{n-1} \\
+ \int_{\partial \Omega} \|f^\infty(\text{Tr}(u_l^{\partial \Omega} - w) \otimes \nu_{\partial \Omega}) - f^\infty(\text{Tr}(u_l^{\partial \Omega} - w) \otimes \nu_{\partial \Omega})\| \, d\mathcal{H}^{n-1} \right) \\
\leq \liminf_{l \to \infty} L \int_{\partial \Omega} |\text{Tr}(u_l - u_l^{\partial \Omega})| \, d\mathcal{H}^{n-1} \leq \liminf_{l \to \infty} CL\|u_l - u_l^{\partial \Omega}\|_{L^1D(\Omega)} = 0.
\]

In conclusion, \( u \in GM(F; u_0) \) and the proof is complete. \( \square \)

**Theorem 3.2.** Let \( f \in C^2(\mathbb{R}^{n \times n}_{\text{sym}}) \) be such that \( f''(z) \) is positive definite on \( \mathbb{R}^{n \times n}_{\text{sym}} \) for every \( z \in \mathbb{R}^{n \times n}_{\text{sym}} \) and, for some \( \lambda > 0 \) there holds

\[
0 < \langle f''(z)\xi, \xi \rangle \leq \lambda \frac{|\xi|^2}{(1 + |z|^2)^2} \quad \text{for all } z, \xi \in \mathbb{R}^{n \times n}_{\text{sym}}.
\]

Then for every ball \( B \subset \Omega \) with \( \text{dist}(B; \partial \Omega) > 0 \) there exists a finite constant \( C > 0 \) such that

\[
\sup_{B} \sup_{j \in \mathbb{N}} \sum_{k=1}^{n} \int_{B} (f''(\varepsilon(v_{l,j}))) \partial_k \varepsilon(v_{l,j}) \partial_k \varepsilon(v_{l,j}) \, dx \leq C.
\]

**Proof.** Step 1. In a first step, for \( j \in \mathbb{N} \) and \( \nu \in \mathbb{R}^{n \times n}_{\text{sym}} \), we define bilinear forms \( A[\nu; \cdot, \cdot] \) and \( A_{ij}[\nu; \cdot, \cdot] \) by

\[
A[\nu; \eta, \xi] := \langle f''(\nu)\eta, \xi \rangle, \\
A_{ij}[\nu; \eta, \xi] := A[\nu; \eta, \xi] + \frac{1}{j} \langle \xi, \eta \rangle, \quad \nu, \eta, \xi \in \mathbb{R}^{n \times n}_{\text{sym}}.
\]

Note that, by positive definiteness of \( f''(\nu) \) for every \( \nu \in \mathbb{R}^{n \times n}_{\text{sym}} \), both \( A[\nu; \cdot, \cdot] \) and \( A_{ij}[\nu; \cdot, \cdot] \) for all \( j \in \mathbb{N} \) define symmetric positive definite bilinear forms and thus we have a suitable Cauchy–Schwarz inequality for \( A_{ij} \) at our disposal. For notational brevity, let us further assume \( l \in \mathbb{N} \) fixed and write \( v_{l,j}, A_{ij} \) or \( \sigma_{l,j} \) instead of \( v_{l,j}, A_{ij} \) and \( \sigma_{l,j} \), respectively.

By Lemma 3.1 (c), we have \( (v_{l,j}) \in W^{2,2}_{\text{loc}}(\Omega; \mathbb{R}^n) \). Hence, integrating by parts, we obtain for all \( \varphi \in C^\infty_c(\Omega; \mathbb{R}^n) \) and all \( l \in \{1, \ldots, n\} \) that

\[
\int_{\Omega} \langle \partial_l \sigma_{l,j}, \varepsilon(\varphi) \rangle \, dx = 0.
\]

By smooth approximation, (3.7) then is equally seen to hold for competitor maps \( \varphi \in W^{1,2}_{0}(\Omega; \mathbb{R}^n) \).

Step 2. To establish (3.5), we shall use the weak Euler–Lagrange equation (3.7) satisfied by \( \sigma_{l,j}, \) \( j \in \mathbb{N} \). Let \( k \in \{1, \ldots, n\} \) and let \( B \subset \Omega \) be an arbitrary open ball with \( \text{dist}(B; \partial \Omega) > 0 \). We choose a cut-off function \( \rho \in C^1_c(\Omega; [0, 1]) \) such that \( \rho \equiv 1 \) on \( B \). Then, since \( v_{l,j} \in W^{2,2}_{\text{loc}}(\Omega; \mathbb{R}^n) \) by Lemma 3.1 (c), we obtain that \( \varphi := \rho^2 \partial_k (v_{l,j} - a_j) =: \rho^2 \partial_k w_j \) belongs to \( W^{1,2}_{0}(\Omega; \mathbb{R}^n) \) and hence qualifies as a competitor map in (3.7). Here, \( a_j \in \mathcal{A}(\Omega) \) is an arbitrary rigid deformation to be specified later on, and \( w_j \) is defined in the obvious manner. Writing \( A = (A_{im})_{i,m=1}^{n} \) for an \( (n \times n) \)-matrix \( A \) and denoting the \( l \)-th component of a vector \( u \in \mathbb{R}^n \) by \( u^{(l)} \), applying (3.7) to \( l = k \) and summing over \( k \in \{1, \ldots, n\} \) yields

\[
\sum_{k,i,m} \int_{\Omega} (\partial_k \sigma_{l,j}^{(m)}) \varepsilon^{im}(\rho^2 \partial_k w_j) \, dx = 0,
\]
where the sum is taken over all indices $k, i, m \in \{1, \ldots, n\}$. Moreover, note that because of

$$
\sum_{k=1}^{n} \int_{\Omega} \langle f''(v_j) \rangle \partial_k h_e(v_j), \partial_k e(v_j) \rangle \, dx \leq \sum_{k=1}^{n} \int_{\Omega} A_j \langle e(v_j); \rho \partial_k e(v_j), \rho \partial_k e(v_j) \rangle \, dx \\
\leq \sum_{k,i,m} \int_{\Omega} (\partial_h \sigma^m_{j}) \rho^2 \partial_k e \langle v_j \rangle \, dx,
$$

(3.9)

it suffices to estimate the right hand side in view of (3.8). We first rewrite (3.8) as

$$
\sum_{k,i,m} \int_{\Omega} \langle \partial_k \sigma^m_{j} \rangle (\partial_k (\rho^2 \partial_k w_j^{(m)}) + \partial_m (\rho^2 \partial_k w_j^{(k)})) \, dx = 0.
$$

(3.10)

Thus we find by expanding and regrouping terms

$$
2 \sum_{k,i,m} \int_{\Omega} \langle \partial_k \sigma^m_{j} \rangle (\rho^2 \partial_k e \langle v_j \rangle) \, dx = - \sum_{k,i,m} \int_{\Omega} \langle \partial_k \sigma^m_{j} \rangle ((\partial_k \rho^2) \partial_k w_j^{(m)} + (\partial_m \rho^2) \partial_k w_j^{(i)}) \, dx \\
= - \sum_{k,i,m} \int_{\Omega} \langle \partial_k \sigma^m_{j} \rangle ((\partial_k \rho^2) \partial_k w_j^{(m)} + (\partial_m \rho^2) \partial_m w_j^{(k)}) \, dx \\
+ \sum_{k,i,m} \int_{\Omega} \langle \partial_k \sigma^m_{j} \rangle ((\partial_k \rho^2) \partial_m w_j^{(k)} + (\partial_m \rho^2) \partial_k w_j^{(k)}) \, dx \\
- \sum_{k,i,m} \int_{\Omega} \langle \partial_k \sigma^m_{j} \rangle ((\partial_m \rho^2) \partial_k w_j^{(i)} + (\partial_m \rho^2) \partial_k w_j^{(k)}) \, dx \\
=: \text{I + II + III}.
$$

Ad I and III. Let us note that since the indices $i, m$ run over all numbers $1, \ldots, n$ and $\sigma(x) \in \mathbb{R}^{n \times n}$, we have $I = III$. Moreover, we note that the artificial terms leading to the appearance of $II$ are just introduced to have the symmetric gradient appearing, that is, terms which are conveniently controllable. In consequence, defining $j \Theta_k := (j \Theta^m_{k})_{i,m=1}^{n}$ and $j \hat{\Theta}_k := (j \hat{\Theta}^m_{k})_{i,m=1}^{n}$ with

$$
\text{I} + \text{III} \leq 2|\text{III}| \leq 4 \sum_{k,i,m} \int_{\Omega} \langle \partial_k \sigma^m_{j} \rangle (\partial_k \rho^2) \partial_k e \langle w_j \rangle \, dx = 4 \sum_{k=1}^{n} \int_{\Omega} \langle \partial_h \sigma_{j}, j \Theta_k \rangle \, dx =: \text{IV}.
$$

We now employ the definition of $\sigma_j$ in conjunction with the bilinear forms $A[e(v_j); \cdot, \cdot]$ and $A_j[e(v_j); \cdot, \cdot]$. Then we obtain, applying the Cauchy-Schwarz inequality to the bilinear forms $A_j$, for some $\theta > 0$ to be fixed later,

$$
\text{IV} \leq \theta \sum_{k,i,m} \int_{\Omega} \langle A_j \rangle \langle e(v_j); \rho \partial_k e(v_j), \rho \partial_k e(v_j) \rangle \, dx + 4C(\theta) \sum_{k=1}^{n} \int_{\Omega} A_j \langle e(v_j); j \hat{\Theta}_k, j \hat{\Theta}_k \rangle \, dx =: \text{IV}'.
$$

Appealing to (3.4), we then further estimate

$$
\text{IV}' \leq \theta \sum_{k=1}^{n} \int_{\Omega} A_j \langle e(v_j); \rho \partial_k e(v_j), \rho \partial_k e(v_j) \rangle \, dx \\
+ 16 A \theta \| \nabla \rho \|_{L^2} \sum_{k=1}^{n} \int_{spt(\rho)} (1 + |e(v_j)|^2)^{\frac{1}{2}} \, dx + \frac{16 \theta \| \nabla \rho \|_{L^2}}{\alpha(l_j)} \int_{spt(\rho)} |e(v_{t,j})|^2 \, dx \\
=: \text{V}_1 + \text{V}_2 + \text{V}_3.
$$
By virtue of Lemma 3.1 (d), $V_2$ and $V_3$ are uniformly bounded in $j \in \mathbb{N}$ (and, as is implicit here, in $l \in \mathbb{N}$). Consequently, we find for some constant $c(\rho, \vartheta) > 0$ independent of $j, l \in \mathbb{N}$,

$$\text{(3.11)} \quad |I + III| \leq 4\theta \sum_{k=1}^{n} \int_{\Omega} A_j[\varepsilon(v_j); \rho \partial_k \varepsilon(v_j), \rho \partial_k \varepsilon(v_j)] \, dx + c(\rho, \vartheta).$$

Ad II. By symmetry of $\sigma_j$, i.e., $\sigma_j^{im} = \sigma_j^{mi}$ for all $i, m \in \{1, \ldots, n\}$ and all $j \in \mathbb{N}$, and a permutation of indices, it suffices to estimate the term

$$2|IV| := 2 \sum_{k, i, m} \int_{\Omega} (\partial_k \sigma_j^{im})(\partial_i \rho^2)(\partial_m w_j^{(k)}) \, dx$$

with an obvious definition of IV. Integrating by parts twice yields

$$\begin{align*}
V_1 &= \sum_{k, i, m} \int_{\Omega} \partial_k \sigma_j^{im} \partial_i \rho^2 \partial_m w_j^{(k)} \, dx = -\sum_{k, i, m} \int_{\Omega} \sigma_j^{im} \partial_k (\partial_i \rho^2 \partial_m w_j^{(k)}) \, dx \\
&= -\sum_{k, i, m} \int_{\Omega} \sigma_j^{im} (\partial_{ik} \rho^2 \partial_m w_j^{(k)}) + \partial_i \rho^2 \partial_{mk} w_j^{(k)} \, dx \\
&= \sum_{k, i, m} \int_{\Omega} (\partial_m \sigma_j^{im}) (\partial_{ik} \rho^2) w_j^{(k)} + \sigma_j^{im} (\partial_{ikm} \rho^2) w_j^{(k)} \, dx \\
&= \sum_{k, i, m} \int_{\Omega} (\partial_m \sigma_j^{im}) (\partial_{ik} \rho^2) w_j^{(k)} + \sigma_j^{im} (\partial_{ikm} \rho^2) w_j^{(k)} \, dx \\
&=: V_1 + \ldots + V_4,
\end{align*}$$

where $V_1, \ldots, V_4$ are defined in the obvious manner. Note that by the $W^{2,2}_{\text{loc}}$–regularity of $v_j$, this is a valid computation. The crucial point in this calculation is that the only derivatives that apply to $w_j$ appear in the form $\partial_k w_j^{(k)}$ (and are decoupled from the $(i, m)$–components), and summation over $k \in \{1, \ldots, n\}$ corresponds to taking the divergence of $w_j$. Now, denoting $A^{(i)}$ the $i$-th row of a matrix $A \in \mathbb{R}^{n \times n}$, we employ the row-wise solenoidality of $\sigma$ (cf. Lemma 3.1 (b)) to find by

$$\psi_{j,k} := (\psi_{j,k})_{n}^{i} := ((\partial_{ik} \rho^2) w_j^{(k)})_{n}^{i} \in W^{1,2}_{0}(\Omega; \mathbb{R}^n)$$

$$V_1 = \sum_{k, i} \int_{\Omega} \text{div}(\sigma_j^{(i)})(\partial_{ik} \rho^2) w_j^{(k)} \, dx = -\sum_{k=1}^{n} \int_{\Omega} \langle \sigma_j, D\psi_{j,k} \rangle \, dx = 0.$$

Similarly, we obtain $V_3 = 0$. Next recall that we still have the freedom to choose the rigid deformations $a_j$ as they appear in the definition of $w_j$. As we can assume without loss of generality that $\text{spt}(\rho)$ is connected, we find a constant $C > 0$ such that for all $v \in LD(\text{spt}(\rho)) \cap W^{1,2}(\text{spt}(\rho); \mathbb{R}^n)$ there exists $a \in \mathcal{S}(\mathbb{R}^n)$ such that

$$\int_{\text{spt}(\rho)} |v - a| \, dx \leq C(\rho) \int_{\text{spt}(\rho)} |\varepsilon(v)| \, dx \quad \text{and} \quad \int_{\text{spt}(\rho)} |v - a|^2 \, dx \leq C(\rho) \int_{\text{spt}(\rho)} |\varepsilon(v)|^2 \, dx.\)$$

It is important that for each such $v$ we can choose one rigid deformation $a$ to make both inequalities work, and by [40, Lem. 4.6], this is in fact possible. Accordingly, we choose for each $j \in \mathbb{N}$ some $a_j \in \mathcal{S}(\mathbb{R}^n)$ such that the ultimate inequality holds with $v$ being replaced by $v_j$ and with $a$ being replaced by $a_j$. Turning to $V_2$, we go back to the definition of $\sigma_j$ and thereby obtain by virtue of Young’s inequality and the above Poincaré inequalities that

$$\begin{align*}
|V_2| &\leq \sum_{k, i, m} \int_{\Omega} (|f'(|\varepsilon(v_j)|)| + \frac{1}{\alpha(l)} |\varepsilon(v_j)|) |(\partial_{ikm} \rho^2)| w_j^{(k)} \, dx \\
&\leq LC(\rho) \int_{\Omega} |w_j| \, dx + \frac{C(\rho)}{2\alpha(l)} \left( \int_{\Omega} |w_j|^2 \, dx + \int_{\Omega} |\varepsilon(v_j)|^2 \, dx \right).
\end{align*}$$
\[ \leq LC(\rho) \int_{\Omega} |\varepsilon(v_{ij})| \, dx + \frac{C(\rho)}{2\alpha(l)j} \int_{\Omega} |\varepsilon(w_{ij})|^2 \, dx \]
\[ \leq C(\rho), \]

where again \( C(\rho) > 0 \) is independent of \( j, l \in \mathbb{N} \). As to VI₄, we note that since \( (\varepsilon(v_{ij})) \) is uniformly bounded in \( L^1(\Omega; \mathbb{R}^{n \times n}) \) by Lemma 3.1, so is \( (\text{div}(v_{ij})) \) in \( L^1(\Omega) \). We then estimate as follows:

\[ \text{VI}_4 \leq C(\rho) \int_{\Omega} |\sigma| |\text{div}(w_{ij})| \, dx \leq C(\rho) \int_{\Omega} L|\varepsilon(v_{ij})| + \frac{1}{\alpha(l)j} |\varepsilon(v_{ij})|^2 \, dx \leq C(\rho), \]

and again, \( C(\rho) > 0 \) is independent of \( j, l \in \mathbb{N} \). In particular, we infer that there holds

\[ (3.12) \quad \text{II} \leq 2|\text{VI}_1 + \ldots + \text{VI}_4| \leq C(\rho). \]

**Step 3.** We now gather the estimates obtained so far; precisely, we go back to (3.9) and estimate

\[ \sum_{k=1}^{n} \int_{\Omega} A_j[\varepsilon(v_{ij}) ; \rho \partial_{k} \varepsilon(v_{ij}), \rho \partial_{k} \varepsilon(v_{ij})] \, dx \leq (3.9) \sum_{k,i,m} (\partial_{k} \sigma_{j}^{im}) \rho^2 |\varepsilon(v_{ij})|^2 \, dx \leq |\text{I} + \text{II} + \text{III}| \]

\[ \leq 4\vartheta \sum_{k=1}^{n} \int_{\Omega} A_j[\varepsilon(v_{ij}) ; \rho \partial_{k} \varepsilon(v_{ij}), \rho \partial_{k} \varepsilon(v_{ij})] \, dx \]
\[ + c(\rho, \vartheta), \]

where we have tacitly synthesised constants in the ultimate step. At this stage, we choose \( \vartheta > 0 \) so small such that the first side on the very right hand side of the previous inequality can be absorbed into its left hand side. In view of arbitrariness of \( B \) and (3.9), this implies (3.5) and the proof is hereby complete. \( \square \)

### 3.2. Proof of Theorem 1.1.

Based on Theorem 3.2, we can proceed to the proof of Theorem 1.1. It needs to be noted that the second order estimate given in (3.5) is the decisive ingredient which we lacked in [40], and in the following we demonstrate how (3.5) leads to a Sobolev regularity improvement. Here, we are lead by the ideas exposed in [17, 19] for the gradient case.

**Proof of Theorem 1.1(i).** In the ellipticity regime \( 1 < a < 2 \), we introduce the auxiliary function

\[ V_a(\xi) := (1 + |\xi|^2)^{\frac{2-a}{2}}, \quad \xi \in \mathbb{R}^{n \times n}, \]

cf. Lemma 2.4. Now let \((v_{ij}) \subset W^{2,2}_{loc}(\Omega; \mathbb{R}^n)\) be the double sequence of viscosity approximations from the previous paragraph. Now, differentiating \( V_a(\varepsilon(v_{ij})) \) (with the corresponding viscosity approximations \( v_{ij} \) of the last section), we obtain for all \( k \in \{1, ..., n\} \)

\[ |\partial_{k} V_a(\varepsilon(v_{ij}))|^2 \leq \left( \frac{2-a}{2} \right)^2 |\partial_{k} \varepsilon(v_{ij})|^2 |\varepsilon(v_{ij})|^2(1 + |\varepsilon(v_{ij})|^2)^{-\frac{2-a}{2}} \leq a(1 + |\varepsilon(v_{ij})|^2) \]

By virtue of the lower bound in (1.4), we deduce from (3.5) that for any open ball \( B' \subset \Omega \) the double sequence \( (DV_a(\varepsilon(v_{ij})))_{B'} \) is uniformly bounded in \( L^2(B'; \mathbb{R}^{n \times n}) \). Standard localisation then yields that for every open ball \( B \subset B' \) the double sequence \( (V_a(\varepsilon(v_{ij})))_{B} \) is uniformly bounded in \( W^{1,2}(B) \). Since \( W^{1,2}(B) \hookrightarrow \text{BMO}(B) \) for \( n = 2 \), we deduce from the classical John-Nirenberg lemma (BMO \( \hookrightarrow \text{L}^p \) for any \( 1 \leq p < \infty \) and Lemma 2.4 that if \( n = 2 \) and \( 1 < a < 2 \), then for any \( 1 \leq p < \infty \) the double sequence \( (\varepsilon(v_{ij})) \) is locally uniformly bounded in any \( L^p \), \( 1 \leq r < \infty \). For \( \Phi : \varepsilon(u) \mapsto Du \) is a singular integral operator of convolution type, another localisation yields that \((v_{ij})\) is locally uniformly bounded in \( W^{1,r}(\Omega; \mathbb{R}^n) \) for any \( 1 \leq r < \infty \). From here, it is then routine to show that the generalised minimiser \( u \) from Lemma 3.1 equally belongs to \( W^{1,r}_{loc}(\Omega; \mathbb{R}^n) \); in particular, we obtain for the singular part \( E^u u \equiv 0 \).

If \( n \geq 3 \), then we can only use the embedding \( W^{1,2}(B) \hookrightarrow L^{2n/(n-2)}(B) \). Similarly as above, this yields local uniform boundedness of \( (\varepsilon(v_{ij})) \) in \( L^q \) for \( q = \frac{2n}{n-2} \), and the latter number satisfies \( q > 1 \) if and only if \( 1 < a < 1 + \frac{2}{n} \). The proof of Theorem 1.1(i) is hereby complete. \( \square \)
We come to the second part of Theorem 1.1. As we briefly argued in the introduction, in this context it is customary\(^2\) to require the local boundedness condition (cf. [17, Ass. 4.11]) which here translates to: For every relatively compact subset \(U \subseteq \Omega\) we have
\[
\sup_{l \in \mathbb{N}} \sup_{j \in \mathbb{N}} \|v_{l,j}\|_{L^\infty(U; \mathbb{R}^n)} < \infty.
\]
In the full gradient case, such a condition can be justified by use of a maximum principle (e.g., cf. [19, Appendix D]) but by dependence of the integrands on the symmetric gradient, such arguments are not at our disposal here. However, as the above proof and alongside discussion shows, there are ellipticity regimes for which such \(L^\infty\)-bounds can be expected but, as shall be addressed in a future publication, can also be achieved provided certain smallness and smoothness assumption are imposed on the Dirichlet data.

**Proof of Theorem 1.1(ii).** Again, let \(B \subseteq \Omega\) be an open ball with \(\text{dist}(B; \partial \Omega) > 0\) and choose a cut-off function \(\rho \in C^1_c(\Omega; [0, 1])\) with \(\rho \equiv 1\) on \(B\). Then we put, for \(l, j \in \mathbb{N}\) arbitrary but fixed, as a preliminary test function (with the shorthand \(v_{l,j}\) instead of \(v_{i,j}\))
\[
\varphi := \rho^2 \log^2(1 + |\epsilon(v_{l,j})|^2)v_{j},
\]
which, because of \((v_{j}) \subset W^{1,2}_{\text{loc}}(\Omega; \mathbb{R}^n)\), has weak symmetric gradient
\[
\epsilon(\varphi) = 2\rho \log^2(1 + |\epsilon(v_{j})|^2) D\rho \circ v_{j} + \rho^2 D(\log^2(1 + |\epsilon(v_{j})|^2)) \circ v_{j}
\]
\[
+ \rho^2 \log^2(1 + |\epsilon(v_{j})|^2)\epsilon(v_{j})
\]
\(\mathcal{L}^n\text{-a.e.}\). By virtue of assumption (3.13) and Korn’s inequality, we conclude that \(\varphi \in W^{1,2}(\Omega; \mathbb{R}^n)\): It is clear that the first term on the right hand side belongs to \(L^2(\Omega; \mathbb{R}^{n \times n})\). On the other hand, we have
\[
|\rho^2 D(\log^2(1 + |\epsilon(v_{j})|^2)) \circ v_{j}| \leq 4\rho \frac{\log(1 + |\epsilon(v_{j})|^2)|D^2 v_j||\epsilon(v_{j})|}{1 + |\epsilon(v_{j})|^2} \leq C\rho^2 |D^2 v_j|,
\]
where we have crucially utilised (3.13), and then it is seen that the corresponding term in the decomposition of \(\epsilon(\varphi)\) also is square-integrable. Since \(\epsilon(v_{j}) \in L^{2n/(n-2)}_{\text{loc}}(\Omega; \mathbb{R}^{n \times n})\) for if \(n \geq 3\) and \(\epsilon(v_{j}) \in \text{BMO}_{\text{loc}}(\Omega; \mathbb{R}^{n \times n})\) for if \(n = 2\) (recall \(v_{j} \in W^{1,2}_{\text{loc}}(\Omega; \mathbb{R}^n)\)), we conclude the admissibility of \(\varphi\). We now insert \(\varphi\) into the Euler-Lagrange equation (3.3). This consequently gives
\[
\int_{\Omega} (f'_{l,j}(\epsilon(v_{j})), \rho^2 \log^2(1 + |\epsilon(v_{j})|^2)\epsilon(v_{j})) \, dx = -\int_{\Omega} (f'_{l,j}(\epsilon(v_{j})), 2\rho \log^2(1 + |\epsilon(v_{j})|^2) D\rho \circ v_{j}) \, dx
\]
\[
- \int_{\Omega} (f'_{l,j}(\epsilon(v_{j})), \rho^2 D(\log^2(1 + |\epsilon(v_{j})|^2)) \circ v_{j}) \, dx
\]
\[
\Leftrightarrow: \mathbf{I} = \mathbf{II} + \mathbf{III}.
\]

**Ad I.** In the following, we use the auxiliary function \(E := \sqrt{1 + |\eta|^2}\). Here we estimate, using Lemma 2.5, the trivial bound \(\log^2(1 + |\eta|^2) \leq c(1 + |\eta|)\) for \(\eta \in \mathbb{R}^{n \times n}\) together with the uniform bound \(\sup_{l \in \mathbb{N}} \sup_{j \in \mathbb{N}} \|u_{l,j}\|_{L^2(\Omega)} < \infty\) of Lemma 3.1 (a),
\[
\mathbf{I} = \int_{\Omega} (f'_{l,j}(\epsilon(v_{j})), \rho^2 \log^2(E^2(\epsilon(v_{j})))\epsilon(v_{j})) \, dx \geq \theta \int_{\Omega} \rho^2 E(\epsilon(v_{j})) \log^2(E^2(\epsilon(v_{j}))) \, dx
\]
\[
- \theta \int_{\Omega} \rho^2 \log^2(E^2(\epsilon(v_{j}))) \, dx
\]
\[
+ \frac{1}{\theta} \int_{\Omega} |\rho \log(1 + |\epsilon(v_{j})|^2)|\epsilon(v_{j})| \, dx
\]
\[
\geq \left( \theta \int_{\Omega} \rho^2 E(\epsilon(v_{j})) \log^2(E^2(\epsilon(v_{j}))) \, dx - C \right)
\]
\(^2\)In their entirety, all Sobolev regularity results in the BV-Dirichlet case without radial symmetry of the integrands for the ellipticity regime considered here make use of this assumption; cf. [17, 18, 19]. In particular, toward our aim of showing essentially that the same regularity is available for BV and BD, this hypothesis is hereby justified in our setup, too.
where \( C > 0 \) does not depend on \( l, j \in \mathbb{N} \). Turning to \( \Pi \), we then invoke hypothesis (3.13) to find, again by the simple estimate for \( \log^2 \) as in the estimation of \( I \) and Lipschitz continuity of \( f \) (cf. Lemma 2.5),

\[
\Pi \leq LC(\rho) \int_{B(x_0, R)} \log^2(E^2(\varepsilon(v_j))) \, dx + \frac{1}{\alpha(l)} \int_\Omega E(\varepsilon(v_j)) \log^2(E^2(\varepsilon(v_j))) \, dx
\]

\[
\leq LC(\rho) \int_\Omega E(\varepsilon(v_j)) \, dx + \frac{C}{\alpha(l)} \int_\Omega E^2(\varepsilon(v_j)) \, dx \leq C,
\]

where \( C > 0 \) is now also independent of \( l, j \in \mathbb{N} \). Ad \( \Pi \). Here we recall (3.15) and therefore, again by Lipschitz continuity of \( f \), for \( \ell > 0 \) to be fixed later

\[
\Pi \leq 4L \int_\Omega \rho^2 \log(E^2(\varepsilon(v_j))) \frac{D_\varepsilon(\varepsilon(v_j))}{1 + |\varepsilon(v_j)|^2} |v_j| \, dx + \frac{C}{\alpha(l)} \int_\Omega \rho^2 |D_\varepsilon(\varepsilon(v_j))|^2 \, dx + \frac{1}{\alpha(l)} \int_\Omega \rho^2 |D_\varepsilon(\varepsilon(v_j))|^2 1 + |\varepsilon(v_j)|^2 \, dx
\]

\[
\leq 4L \int_\Omega \rho^2 \log(E^2(\varepsilon(v_j))) \frac{D_\varepsilon(\varepsilon(v_j))}{1 + |\varepsilon(v_j)|^2} |v_j| \, dx + \frac{C}{\alpha(l)} \int_\Omega \rho^2 |D_\varepsilon(\varepsilon(v_j))|^2 \, dx + 4L\ell \int_\Omega \rho^2 \log^2(E^2(\varepsilon(v_j))) E(\varepsilon(v_j)) \, dx
\]

\[
= : \Pi_1 + \Pi_2 + \Pi_3 + \Pi_4.
\]

Choosing \( \ell > 0 \) small enough, we may absorb \( \Pi_2 \) into the term \( \Pi_1 \) in the overall inequality. The terms \( \Pi_1 \) and \( \Pi_2 \) are bounded by Theorem 3.2, and the term \( \Pi_4 \) is bounded by the simple estimate for \( \log^2 \) from above and \( \sup_{l \in \mathbb{N}} \|v_{l, j}\|_{L^2(\Omega)} < \infty \). Now we gather estimates and absorb as indicated to find (recall that \( v_j \) tacitly abbreviates \( v_{l, j} \))

\[
(3.16) \quad \sup_{l \in \mathbb{N}} \sup_{j \in \mathbb{N}} \int_B \log^2(E^2(\varepsilon(v_{l, j})) E(\varepsilon(v_{l, j}))) \, dx < \infty.
\]

By passing to a suitable subsequence, this estimate implies in a straightforward manner that the weak*-limit \( u \in BD(\Omega) \) as found in Lemma 3.1 satisfies \( E^* u = 0 \) (so \( Eu = \varepsilon(u) \) in our notation) in \( \Omega \) and, moreover, \( \varepsilon(u) \in L \log^2 L_{loc}(\Omega; \mathbb{R}^n) \). Now we use Lemma 2.1 to find by standard localisation that actually \( \partial u \in L \log L_{loc}(\Omega; \mathbb{R}^n) \) and hence \( u \in W^{1,1}_{loc}(\Omega; \mathbb{R}^n) \). The proof is complete. \( \square \)

We conclude this subsection with two easy remarks on the case of intermediate ellipticities \( 1 + \frac{2}{n} < a < 3 \) subject to the local boundedness hypothesis (3.13) and on crucial differences to the BV-case.

\textbf{Remark 3.3.} If \( 1 + \frac{2}{n} < a < 3 \) and (3.13) is in action, then we can – following [17] – utilise the test maps \( \varphi = \rho^2 E^p(\varepsilon(v_j)) v_j \) for some \( b > 0 \) instead of the (3.14) to obtain the existence of a generalised minimiser that actually belongs to \( W^{1,1}_{loc}(\Omega; \mathbb{R}^n) \) for every \( p < 4 - \max\{2, a\} \); this is similar to the derivation of the corresponding BV-bound, cf. [37].

\textbf{Remark 3.4.} To obtain a result for all generalised minimisers, the only fruitful method known to the author is an application of the Ekeland variational principle (see [19] for the BV- and [40] for the BD-case). Here one also works with viscosity approximations; however, these do not satisfy Euler-Lagrange equations but – by the very method of proof – differential inequalities (in the terminology of [19], \textit{perturbed Euler-Lagrange equations}). Unfortunately, it is not clear to us how to fruitfully employ this approach in the setting of Theorem 3.2 as here the approximate solenoidality of the relevant terms \( \sigma_{l, j} \) does not seem tantamount to control the arising critical terms. Hence we only obtain statements for
one generalised minimiser, but we will merge the available improvements with [40] in the next section to obtain selected improved statements for all generalised minimisers.

3.3. Selected Implications. We now collect some consequences of the results established above and particularly improve the results from [40].

**Corollary 3.5** (Existence of second derivatives). Suppose that $f \in C^2(\mathbb{R}^{n \times n})$ satisfies (LG') and (1.4) for some $a > 1$. Then the following hold:

(a) If $n = 2$ and $1 < a < 2$, then there exists $u \in \text{GM}(F; u_0)$ with $u \in W^{2,\#}_{\text{loc}}(\Omega; \mathbb{R}^n)$ for any $1 \leq q < 2$.

(b) Let $n \geq 3$. If $1 \leq q < \frac{n-2}{n-1}$ and $1 < a \leq \frac{2n}{n-2}q$, there exists $u \in \text{GM}(F; u_0)$ such that $u \in W^{2,\#}_{\text{loc}}(\Omega; \mathbb{R}^n)$.

**Proof.** Let $B \subset \Omega$ be a ball. Then estimate, using Young’s inequality with $p = \frac{n}{q} > 1$ by $q < 2$,

$$
\int_B |D\varepsilon(v_{l,j})|^q \, dx \leq \frac{\tilde{q}}{2} \int_B |D\varepsilon(v_{l,j})|^2 \frac{q}{(1 + |\varepsilon(v_{l,j})|^2)^{\frac{q}{2}}} \, dx + \frac{2 - q}{2} \int_B (1 + |\varepsilon(v_{l,j})|^2)^{\frac{q}{2}} x^{a-\frac{n}{2}} \, dx.
$$

The first term is uniformly controlled by Theorem 3.2. If $n = 2$, then the second term is uniformly bounded in $l, j \in \mathbb{N}$ regardless of $1 < q < 2$. Appealing to the proof of Theorem 1.1(a), the second term is uniformly bounded in $l, j \in \mathbb{N}$ provided

$$
(3.17) \quad 2 - q \leq \frac{2 - a}{n - 2} \quad \text{that is,} \quad a \leq \frac{2 - q}{n - q} =: \tilde{q}(n).
$$

Now note that $\tilde{q}(n) > 1$ if and only if $1 \leq q < \frac{n-2}{n-1}$, and hence the statement follows. 

Obviously, a similar corollary can be established for $1 + \frac{1}{2n} < a < 3$, now invoking Remark 3.3. Also, merging the above results with the universal vanishing of $E^s_u$ for all $u \in \text{GM}(F; u_0)$ for if $1 < a < \frac{n+1}{n}$ (cf. [40, Thm. 1.2]), we obtain the following result:

**Corollary 3.6.** Let $\Omega$ be an open and bounded Lipschitz subset of $\mathbb{R}^n$. Suppose further that $f \in C^2(\mathbb{R}^{n \times n})$ satisfies (LG') and (1.4) with $1 < a < 1 + \frac{1}{n}$. Then

(a) if $n = 2$ and $1 < a \leq \frac{3}{2}$, any $u \in \text{GM}(F; u_0)$ satisfies $u \in W^{2,\#}_{\text{loc}}(\Omega; \mathbb{R}^n)$ for all $1 \leq q < 2$.

(b) if $n \geq 3$ and $1 < a \leq \frac{n+1}{n}$, any $u \in \text{GM}(F; u_0)$ satisfies $u \in W^{2,\#}_{\text{loc}}(\Omega; \mathbb{R}^n)$ for all $1 \leq q < \frac{n^2-n}{n^2-2n-1}$.

**Proof.** By [40, Thm. 1.2] whose proof is based on the Ekeland variational principle, $E^s_u \equiv 0$ in $\Omega$ for every $u \in \text{GM}(F; u_0)$ provided $1 < a < 1 + \frac{1}{n}$. We then deduce in a standard way that on any open connected component of $\Omega$, any two generalised minima only differ by a constant and so, in particular, necessarily share the same interior regularity. Now (a) immediately follows from Corollary 3.5. For part (b), note that (3.17) with $a = 1 + \frac{1}{n} - \varepsilon$ and consequently sending $\varepsilon \searrow 0$ yields $q < \frac{n^2-n}{n^2-2n-1}$ and we conclude.

Compared to [40], we have now established that for the ellipticity regime $1 < a < \frac{n+1}{n}$, all generalised minima possess second derivatives in some $L^q_{\text{loc}}, q > 1$. Finally, an easy application of the measure density lemma [43, Prop. 2.7] yields the following

**Corollary 3.7.** Let $\Omega$ be an open and bounded Lipschitz subset of $\mathbb{R}^n$ and that $f \in C^2(\mathbb{R}^{n \times n})$ satisfies (LG') with $1 < a \leq 3$. For a given map $v \in \text{BD}(\Omega)$, put

$$
\Sigma_v := \left\{ x \in \Omega : \limsup_{R \searrow 0} \left[ \frac{\int_{B(x,R)} |\mathcal{E}v - z| \, dx + |E^sv|_{B(x,R)} \, dx}{R^n} \right] > 0 \text{ for all } z \in \mathbb{R}^{n \times n} \right\}.
$$

Then the following holds:

(a) If $n = 2$ and $1 < a < \frac{3}{2}$, then any $u \in \text{GM}(F; u_0)$ satisfies $\text{dim}_{\mathcal{H}}(\Sigma_u) = 0$.

(b) If $n \geq 3$ and $1 < a < \frac{n+1}{n}$, then any $u \in \text{GM}(F; u_0)$ satisfies $\text{dim}_{\mathcal{H}}(\Sigma_u) \leq \frac{n^2-n^2+n}{n^2-2n-1}$.
4. Partial Regularity and the Proof of Theorem 1.2

4.1. Outline of the proof and setup. Compared to the standard apparatus of partial regularity proofs, the method as employed in [9] is fairly uncommon and we briefly pause to sketch the main ideas. In order to establish the partial regularity for $u \in GM(F; u_0)$, $u$ is compared with suitable mollifications. Based on the minimality of $u$, these are – in some specified sense – close to solving a linear elliptic system. Hence the mollifications satisfy good decay estimates, and these need to be shown to be inherited by $u$. This will eventually yield the desired partial regularity and alongside reveals two major issues: As mollifications and so convolutions with smooth bumps work well in conjunction with convex functions by Jensen’s inequality, the method is somewhat designed for the convex case. On the other hand, to fruitfully accomplish the comparison, we shall be lead to measure the distance of $u$ to its mollifications in terms of a suitable auxiliary function, and here we need a novel Poincaré-type inequality, see Section 4.2. Let us note in advance that, by Ornstein’s Non-Inequality, the latter cannot be obtained by reducing to the full gradient and entails proper modification of the partial regularity proof. More precisely, we proceed as follows:

(i) Section 4.3.1: Estimates for comparison maps. Here we establish, by linearisation, that if a $C^{1,\alpha}$-Hölder continuous function satisfies a certain smallness condition and has symmetric gradient close to some carefully chosen reference point, then it almost enjoys the typical decay for linear systems. Crucially, this typical decay is only perturbed by its $L^2$-deviation from the aforementioned reference point and by a quantity that measures how far the considered function is away from minimising the variational integral at our disposal.

(ii) Section 4.3.2: Smoothing and selection of good radii. To construct the $C^{1,\alpha}$-comparison maps required for step (i), we carefully mollify the given generalised minimiser and demonstrate that – under the conditions of Proposition 4.1 – the mollification parameters can be chosen so that the comparison estimates from (i) become available, cf. Lemma 4.6. Once these preparations are accomplished, we prove that there exist good radii in the following sense: On the balls corresponding to good radii, the difference of the functionals applied to the generalised minima and the functionals applied to their mollified variants can be sharply controlled, cf. Lemma 4.8. This step – as in [9] – is key to the approach and seems to require convexity.

(iii) Section 4.3.3: Comparison estimates and decay. Here we carry out the aforementioned comparison argument and employ minimality to deduce validity of Proposition 4.1 below. As it stands to reason, such an argument shall require – besides the sharp estimates on balls with good radii from step (ii) – a Poincaré-type inequality for locally measuring the distance of the generalised minimiser to its mollified version in terms of the auxiliary function $e$, cf. Section 2.3, and its symmetric gradient. This inequality, which is one of the main tools in the proof, is carefully derived in Section 4.2.

We now introduce the requisite terminology for the proof below: Given $z \in \Omega$ and $R > 0$ such that $B(z, R) \Subset \Omega$, we define two excess quantities by

$$E(u; z, R) := \int_{B(z,R)} e(Eu - (Eu)_z) \quad \text{and} \quad \hat{E}(u; z, R) := \frac{E(u; z, R)}{R^n},$$

where the mean value in the definition of $E$ is taken with respect to $\mathcal{L}^n$. Theorem 1.2 will follow from the next proposition.

Proposition 4.1. Let $f : \mathbb{R}_+^{n \times n} \to \mathbb{R}$ be a convex $C^2$-function that satisfies the assumptions of Theorem 1.2. Suppose that $m \in \mathbb{R}_+^{n \times n}$ and $\sigma > 0$ are such that

$$|f''(m') - f''(m)| \leq \omega(|m - m'|) \quad \text{for all} \ m' \in B(m, \sigma)$$

with a non-decreasing function $\omega : \mathbb{R}_+^n \to \mathbb{R}_+$ which, in addition, satisfies $|\omega(t)| \leq M = M(|m|)$ for all $t \in \mathbb{R}_+$ and $\lim_{t \to 0} \omega(t) = 0$, such that we have

$$\lambda |z|^2 \leq (f''(m)z, z) \leq \Lambda |z|^2$$
for two constants $0 < \lambda \leq \Lambda < \infty$ and all $z \in \mathbb{R}^{n \times n}_{\text{sym}}$. There exists $\tau > 0$ with the following property: If for $z \in \Omega$ and $R > 0$ with $B(z, R) \subset \Omega$ a map $u \in \text{BD}(\Omega)$ minimises

$$F_R[u] := \int_{B(z, R)} f(Eu)$$

with its own boundary values, satisfies $(Eu)_z = m$ and the excess satisfies $\hat{E}(u; z, R) < \tau$, then for any $0 < r < R/4$ we have the estimate

$$(4.4) \quad E(u; z, r) \lesssim \left((r/R)^{n+2} + \left(\rho\left(\frac{1}{r}\vert\nabla u\vert\right) + (R/r)^{n+1}\right)E(u; z, R)\right)$$

with a bounded non-decreasing function $\rho: \mathbb{R}^+ \to \mathbb{R}^+$ such that $\lim_{t \to 0} \rho(t) = 0$. Here, both $\tau$ and the constants implicit in '$\lesssim$' only depend on $n$, $m$, $\lambda$, $\Lambda$ and the linear growth parameters from (LG), but not on $u$, $r$ or $R$.

The way in which Theorem 1.2 then is analogous to the reasoning in [9] and is omitted here. Instead, let us note that from Proposition 4.1 it is evident that Theorem 1.2 also generalised to local generalised minimisers. These are somewhat defined canonically, namely, we say that $u \in \text{BD}_{\text{loc}}(\Omega)$ is a local generalised minimiser if and only if for every Lipschitz subset $\omega \Subset \Omega$ there holds $F_u[u; \omega] \leq F_u[v; \omega]$ for all $v \in \text{BD}(\omega)$, where

$$F_u[u; \omega] := \int_{\omega} f(Eu) + \int_{\partial \omega} f^\infty(\text{Tr}(u - v) \circ \nu_{\partial \omega}) \, dH^{n-1}, \quad v \in \text{BD}(\omega).$$

That is to say, on every Lipschitz subset $\omega$ of $\Omega$ the map $u$ minimises $F\left[-; \omega\right]$ with respect to its own boundary values. As in the usual $W^{1, p}$-case, generalised minima are local generalised minima but not necessarily vice versa. We conclude with a remark on notation.

**Notation 4.2.** In the following, we will use the single symbol $\varepsilon$ for both a mollification parameter and $\varepsilon(u)$ for the symmetric gradient of an LD-map $u$, but no ambiguities will arise from this.

### 4.2. Generalised Poincaré-type inequalities

In this section we discuss a family of convolution inequlities, and we start off with the precise version as required in the partial regularity proof later on. Here we provide the construction in detail and then comment on related inequalities that can be obtained similarly. To fix terminology, any radially symmetric $\rho: \mathbb{R}^n \to [0, 1]$ with $\int_{\mathbb{R}^n} \rho \, dx = 1$ which is $C^\infty$ in the interior of its support shall be referred to as a *standard mollifier*, and we denote for $\varepsilon > 0$ by $\rho_\varepsilon(x) := \varepsilon^{-n} \rho(x/\varepsilon)$ its $\varepsilon$-rescaled variant. The main result of this section now reads as follows.

**Proposition 4.3.** Let $\lambda > 1$ and $n \in \mathbb{N}$. Then there exists a constant $C = C(n, \lambda) > 0$ such that the following holds: For every open and bounded Lipschitz domain $U \Subset \mathbb{R}^n$, $u \in \text{BD}(\mathbb{R}^n)$ and numbers $\varepsilon > 0$ and $L > 0$ there holds

$$(4.5) \quad \int_U e\left(|L(u - \rho_\varepsilon * u)|\right) \, dx \leq C \max\{\langle L\varepsilon \rangle, \langle L\varepsilon \rangle^2\} \int_{U + B(0, \lambda \sqrt{\varepsilon})} e\left(|Eu|\right).$$

Here, $\rho: \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is an arbitrary standard mollifier in the above sense and $\rho_\varepsilon$ its $\varepsilon$-rescaled variant.

Before we pass to the proof of the preceding proposition, a remark is in order.

**Remark 4.4 (Obstructions).** If in Proposition 4.3 the symmetric gradient is replaced by the full gradient, then the statement follows by the fundamental theorem of calculus in a straightforward way. In fact, the difference $u - \rho_\varepsilon * u$ can be controlled for $L^\infty$-a.e. $x$ by virtue of the inequality

$$|u(x) - \rho_\varepsilon * u(x)| \leq \int_{\mathbb{R}^n} \rho(y)\left|u(x - \varepsilon y) - u(x)\right| \, dy \leq \varepsilon \int_{\mathbb{R}^n} \rho(y) \left|\langle \nabla u\rangle(x - t\varepsilon y)\right| \, dy \, dt$$

in conjunction with Jensen’s inequality as is done in [9]. In the symmetric gradient case, we thus must directly avoid passing to $\nabla u$ as this term cannot be controlled unless we assume some strong $\alpha$-ellipticity (cf. (1.4)) and appeal to the results of Section 3. However, it is exactly at this point that the partial regularity result is not subject to a strong convexity assumption. Also note that the prefactor $\max\{\varepsilon, \varepsilon^2\}$ is crucial in the proof below; if it were replaced by $\max\{\varepsilon^b, \varepsilon^{2b}\}$ for some $0 < b < 1$, then the proposition would follow potential theoretic considerations by virtue of (2.4) and the implied embedding $\text{BD} \hookrightarrow \text{W}^{2, 1}_{\text{loc}}$ for $0 < \alpha < 1$, but this shall not be sufficient to conclude the decay below.
**Proof of Proposition 4.3.** We divide the proof into four main steps; first we record some background facts and fix terminology, then provide a prove a Poincaré-type inequality for the function \( e \) and then conclude the claim for LD-maps. Finally, we briefly argue how to obtain the statement for BD-maps in the fourth step.

**Step 1. Preliminaries.** Let \( \lambda > 1 \) be given. Let \( U \) be as in the lemma and denote, for \( t > 0 \), \( N_t(U) := \{ x \in \mathbb{R}^n : \text{dist}(x, U) < t \} \) the \( t \)-neighbourhood of \( U \). We then first determine and consequently fix a number \( \ell \in \mathbb{N} \) such that \( \frac{1}{\ell} < t \). In a next step, we put \( \varepsilon_{\lambda, n} := \frac{1}{\ell} \).

Now we consider the lattice \( \Gamma_{\varepsilon_{\lambda, n}} := \{ \lambda, n \mathbb{Z}^n \} \) and denote \( Q_{\varepsilon_{\lambda, n}} \) the collection of all open cubes of sidelength \( \varepsilon_{\lambda, n} \) and edge points contained in \( \Gamma_{\varepsilon_{\lambda, n}} \). Given \( Q \in Q_{\varepsilon_{\lambda, n}} \), we denote \( \tilde{Q} \) the cube which has the same center as \( Q \) and sides parallel to those of \( Q \) but sidelength \( (2\ell + 1)\varepsilon_{\lambda, n} \), see Figure 2.

Then \( \tilde{Q} \) has all its edge points equally contained in \( \Gamma_{\varepsilon_{\lambda, n}} \) and \( N_\varepsilon(Q) = Q + B(0, \varepsilon) \subset \tilde{Q} \), and can be written as the union of \( \mathcal{N}(\lambda, n) \) cubes from \( Q_{\varepsilon_{\lambda, n}} \); for notational convenience, we denote these cubes \( Q^{(i)} \), \( i = 1, \ldots, \mathcal{N}(\lambda, n) \), and arrange that for all \( Q \in Q_{\varepsilon_{\lambda, n}} \), the relative positioning of \( Q^{(i)} \) to \( Q \) is the same. Moreover, note that if \( Q \in Q_{\varepsilon_{\lambda, n}} \) satisfies \( Q \cap U \neq \emptyset \), then we have \( \tilde{Q} \subset N_{\lambda \sqrt{\varepsilon}}(U) \). In fact, in this case there exists \( x_0 \in Q \cap U \) and thus for any \( z \in Q \) we have \( \text{dist}(z, \partial U) \leq |x_0 - z| \). By the geometry of \( \tilde{Q} \) (see Figure 2), it is clear that \( |x_0 - z| \) is at most

\[
\sqrt{n \varepsilon_{\lambda, n}} + \sqrt{n \varepsilon} = \sqrt{n \varepsilon_{\lambda, n}} (1 + \frac{1}{\ell}) < \lambda \sqrt{n \varepsilon}
\]

and hence \( \text{dist}(z, \partial U) < \lambda \sqrt{n \varepsilon} \) so that \( z \in N_{\lambda \sqrt{n \varepsilon}}(U) \). Summarising, for every \( Q \in Q_{\varepsilon_{\lambda, n}} \) with \( Q \cap U \neq \emptyset \), we have \( \tilde{Q} = \bigcup_{i=1}^{\mathcal{N}(\lambda, n)} Q^{(i)} \subset N_{\lambda \sqrt{n \varepsilon}}(U) \).

As a main feature of the symmetric gradient operator, let us note that as first order polynomials, all elements \( \alpha \in \mathcal{A}(\mathbb{R}^n) \) of its nullspace are harmonic. Thus they satisfy the mean value property and, as a consequence, convolution with standard mollifiers locally turns out to be the identity on the rigid deformations, cf. [32, Chpt. 2.2.3, Thm. 6].

**Step 2. A Poincaré-type inequality for the function e.** In a second step, we claim that for every open cube \( Q \subset \mathbb{R}^n \), every \( L > 0 \) and every \( u \in \text{LD}(\mathbb{R}^n) \), there exists \( r_Q \in \mathcal{A}(\mathbb{R}^n) \) such that

\[
\int_Q e(|L(u - r_Q)|) \, dx \leq C \max \{ L \ell(Q), (L \ell(Q))^2 \} \int_{\mathbb{R}^n} e(|\varepsilon(u)|) \, dx,
\]

and here \( C > 0 \) does neither depend on \( Q \) nor on \( u \). It is crucial for this inequality to be available in this form, and so we provide the details. We thus assume \( u \in C^\infty(\mathbb{R}^n; \mathbb{R}^n) \cap \text{LD}(\mathbb{R}^n) \) first and employ the representation from Lemma 2.2: \( u(x) - r_Q(x) = T_Q[e(v)](x) \) for \( x \in Q \), where \( r_Q = \Pi_Q v \). As \( T_Q \) is linear, \( L(u(x) - r_Q(x)) = T_Q[L(e(u))](x) \) for all \( x \in Q \). Now recall the representation (2.7) together with the bounds on its integral kernel \( R_Q \) afterwards: \( |R_Q(x, y)| \leq C_R/|x - y|^{n-1} \) for \( \mathcal{L}^n \)-a.e. \( x, y \in Q \), where \( C_R > 0 \) is independent of \( Q \). Let \( x \in Q \). We define a measure \( \mu_x : \mathcal{B}(Q) \rightarrow \mathbb{R} \)

**Figure 2.** Neighbouring cube notation.
so \( \ell Q \) for all cubes \( (4.7) \)
\[ x, y \]
sequence, as \(|x - y| < \ell(Q)\) and so \( \ell(Q)^{n-1} \leq |x - y|^n - 1 \). We also need a remark on the upper bound. Namely, if \( x \in Q \), then \( Q \subset B(x, \sqrt{\ell(Q)}) \) independently of \( x \). Thus
\[ \mu_x(Q) \leq C R \int_{B(x, \sqrt{\ell(Q)})} \frac{dy}{|x - y|^{n-1}} = C R \int_{B(0, \sqrt{\ell(Q)})} \frac{dy}{|y|^{n-1}} \leq C \ell(Q), \]
and \( C > 0 \) here neither depends on \( x \) nor \( Q \). In conclusion, we have
\[ (4.7) \]
for all cubes \( Q \) and \( x \) in \( Q \) with two constants \( 0 < C_1 \leq C_2 < \infty \) independent of \( Q \) and \( x \). Now, \( \mu_x/\mu_x(Q) \) is a probability measure for every (non-degenerate) cube \( Q \) and every \( x \in Q \). In consequence, as \( |u - r_Q| \leq |T_Q[e(u)]| \) pointwisely and \( e \) is monotone, we estimate by Jensen’s inequality
\[
\int_Q e(|L(u - r_Q)|) \, dx \leq \int_Q e \left( L \int_Q R_Q(x, y) e(u(y)) \, dy \right) \, dx \\
\leq \int_Q e \left( L e \left( L \int_Q e(u(y)) \, dy \right) \frac{dy}{\mu_x(Q)} \right) \, dx \\
\leq C \max \{ (Le(Q)), (Le(Q))^2 \} \times \int_Q e \left( \int_Q \frac{|e(u(y)|}{\mu_x(Q)} \, dy \right) \, dx \\
\overset{\text{Lemma 2.3(1) \& (4.7)}}{\leq} C \max \{ (Le(Q)), (Le(Q))^2 \} \times \int_Q e \left( \int_Q \frac{|e(u(y)|}{\mu_x(Q)} \, dy \right) \, dx \\
\overset{\text{Jensen}}{\leq} C \max \{ (Le(Q)), (Le(Q))^2 \} \times \int_Q e \left( \int_Q \frac{|e(u(y)|}{\mu_x(Q)} \, dy \right) \, dx 
\]
where the ultimate inequality follows from Lemma \(\hat{\epsilon}\). Consequently, going back to \(\tilde{\epsilon}\), we have used at 

\[(4.8)\]

C \max\{(\ell(Q)), (\ell(Q))^2\} \frac{1}{\ell(Q)} \int_Q e(|\epsilon(u)(y)|) |dy| dx.

Fubini

\[\leq C \max\{(\ell(Q)), (\ell(Q))^2\} \frac{1}{\ell(Q)} \int_Q e(|\epsilon(u)(y)|) |dy| dx\]

\[(4.7)\]

C \max\{(\ell(Q)), (\ell(Q))^2\} \int_Q e(|\epsilon(u)(y)|) |dy| dx.

This establishes \((4.6)\) for \(u \in C^\infty(\mathbb{R}^n;\mathbb{R}^n) \cap LD(\mathbb{R}^n);\) note that \(C > 0\) here does not depend on \(Q\) and \(u\). Now it suffices to approximate \(u \in LD(\mathbb{R}^n)\) in the \(L^\infty\)-norm topology by maps \(u_k \in C^\infty(\mathbb{R}^n;\mathbb{R}^n) \cap LD(\mathbb{R}^n)\) to conclude \((4.6)\) by virtue of Lebesgue’s theorem on dominated convergence.

**Step 3. Conclusion for \(L^\infty\)-maps.** For any \(Q \in \mathcal{Q}_{\varepsilon, \lambda, n}\), we recall the definition of the cube \(Q\) from step 1 and denote \(r_Q \in \mathcal{B}(\mathbb{R}^n)\) the rigid deformation appearing in \((4.6)\) with \(Q\) being systematically replaced by \(\tilde{Q}\). We then obtain, using Lemma 2.3(i) in the third step

\[
\int_{U} e(|L(u - \rho \ast u)|) \, dx \leq \sum_{Q \in \mathcal{Q}_{\varepsilon, \lambda, n}} \int_{Q} e(|L(u - \rho \ast u)|) \, dx
\]

\[
\leq C \sum_{Q \in \mathcal{Q}_{\varepsilon, \lambda, n}} \int_{Q} e(|L(u - r_{\tilde{Q}}) - \rho \ast L(u - r_{\tilde{Q}})|) \, dx \quad \text{(as} \ (\rho \ast r_{\tilde{Q}})|Q = r_Q|Q)
\]

\[
\leq C \sum_{Q \in \mathcal{Q}_{\varepsilon, \lambda, n}} \int_{Q} e(|L(u - r_{\tilde{Q}})|) + e(|\rho \ast L(u - r_{\tilde{Q}})|) \, dx
\]

=: (*),

where the ultimate inequality follows from Lemma 2.3(i). At this stage, we use Jensen’s and Young’s convolution inequality to conclude that for any \(Q \in \mathcal{Q}_{\varepsilon, \lambda, n}\)

\[
\int_{Q} e(|\rho \ast L(u - r_{\tilde{Q}})|) \, dx \leq \int_{Q} \rho \ast e(|u - r_{\tilde{Q}}|) \, dx
\]

\[
\leq \int_{N_{\varepsilon}(Q)} e(|L(u - r_{\tilde{Q}})|) \, dx
\]

\[
\leq \int_{Q} e(|L(u - r_{\tilde{Q}})|) \, dx \quad \text{(as} \ N_{\varepsilon}(Q) \subset \tilde{Q})
\]

\[(4.6)\]

\[
\leq C \max\{\ell(\tilde{Q}), (\ell(\tilde{Q}))^2\} \int_{\tilde{Q}} e(|\epsilon(u)|) \, dx
\]

\[\leq C \max\{\ell(\tilde{Q}), (\ell(\tilde{Q}))^2\} \sum_{j=1}^{N(\lambda, n)} \int_{Q^{(j)}} e(|\epsilon(u)|) \, dx.
\]

Here we have used at (**) that \(\ell(\tilde{Q}) = \frac{\ell}{2}\) and so \(C\) consequently depends on \(\lambda\) and \(n\) only. Consequently, going back to (*), we obtain with \(Q \subset \tilde{Q}\) in the first step

\[
\int_{U} e(|L(u - \rho \ast u)|) \, dx \leq C \left( \sum_{Q \in \mathcal{Q}_{\varepsilon, \lambda, n}} \int_{Q} e(|L(u - r_{\tilde{Q}})|) \right) + \left( \sum_{Q \in \mathcal{Q}_{\varepsilon, \lambda, n}} \int_{Q} e(|\rho \ast L(u - r_{\tilde{Q}})|) \right)
\]

\[
\leq C \max\{\ell(\tilde{Q}), (\ell(\tilde{Q}))^2\} \sum_{Q \in \mathcal{Q}_{\varepsilon, \lambda, n}} \sum_{j=1}^{N(\lambda, n)} \int_{Q^{(j)}} e(|\epsilon(u)|) \, dx
\]
\[ C \max\{ (L\varepsilon), (L\varepsilon)^2 \} \sum_{j=1}^{\mathcal{N}(\lambda, n)} \sum_{Q \in \mathcal{S}(\lambda, n)_{\mathbb{Q}}} \int_{Q} \varepsilon(\varepsilon(u)) \, dx \leq C \max\{ (L\varepsilon), (L\varepsilon)^2 \} \int_{\mathcal{N}(\lambda, n + u)} \varepsilon(\varepsilon(u)) \, dx. \]

from where (4.5) follows in the LD-case. Here we have used that, by step 1, every \( Q^{(j)} \) appears at most \( \mathcal{N}(\lambda, n) \)-times in the sum of the penultimate line and each \( Q^{(j)} \) is contained in \( N_{\lambda, \sqrt{\varepsilon}}(U) \); hence, \( C = C(\lambda, n) \) again. It therefore remains to give the corresponding bound in the slightly more general BD-case.

**Step 4. Passage to the general case.** For the general case, we smoothly approximate \( u \in \text{BD}(\mathbb{R}^n) \) in the strict topology by \( (u_k) \subset C^\infty(\mathbb{R}^n; \mathbb{R}^n) \cap \text{LD}(\mathbb{R}^n) \) obtained by setting \( u_k := \eta_{1/k} \ast u \), where \( \eta \) is a possibly different smooth standard mollifier. This yields by Fatou’s lemma for all \( \varepsilon > 0 \)

\[
\int_{U} e((L(u - \rho_{\varepsilon} \ast u))) \, dx \leq \liminf_{k \to \infty} \int_{Q} e((L(u_k - \rho_{\varepsilon} \ast u_k))) \, dx \leq C \max\{ (L\varepsilon), (L\varepsilon)^2 \} \liminf_{k \to \infty} \int_{\mathcal{N}(\lambda, n + u_k)} e((\varepsilon u_k)(y)) \, dy \leq C \max\{ (L\varepsilon), (L\varepsilon)^2 \} \int_{\mathcal{N}(\lambda, n + u)} e((\varepsilon u))
\]

This is the inequality claimed in the proposition and the proof is hereby complete. \( \square \)

For consistency, let us moreover note that if the right hand side of (4.5) is zero, then \( u \) must coincide with a rigid deformation on each of the connected components of \( U + \text{B}(0, \sqrt{n}) \) and so on those of \( U \); in consequence, it must coincide with its mollification on each of these connected components and hence the left hand side is zero indeed.

Other inequalities of this sort are also possible, some of which we enumerate now: First, if one uses STRAUSS’ embedding [72] \( \text{BD}(\mathbb{R}^n) \hookrightarrow L^{\frac{n}{n-1}}(\mathbb{R}^n; \mathbb{R}^n) \), then an analogous approach yields the inequality

\[
\left( \int_{\mathbb{R}^n} |u - \rho_{\varepsilon} \ast u|^{\frac{n}{n-1}} \, dx \right)^{\frac{n-1}{n}} \leq C \varepsilon |\text{Eu}|(\mathbb{R}^n), \quad \text{for } u \in \text{BD}(\mathbb{R}^n)
\]

with \( C = C(n) > 0 \) being a constant.

For the upcoming subsection, we shall work with two mollification parameters. We therefore put

\[
\rho^{(1)}(x) := \mathcal{L}^{-1}(B(0, 1))^{-1} \mathbb{B}_{B(0, 1)}(x) \text{ and } \rho^{(2)}(x) := \gamma_n \mathbb{B}_{B(0, 1)}(x) \exp \left( \frac{1}{1 - |x|^2} \right), \quad x \in \mathbb{R}^n,
\]

\( \gamma_n \) being adjusted in a way such that \( \| \rho^{(2)} \|_{L^1} = 1 \). For \( u \in L^1_{\text{loc}}(\mathbb{R}^n; \mathbb{R}^n) \), we arrange that

\[
u_\delta := \rho^{(1)} \ast u \text{ and } u_{\delta, \varepsilon} := \rho^{(2)} \ast u_{\delta}. \]

Finally, if any of the estimates presented so far is invoked to estimate the differences \( u - (u_\delta)_\varepsilon \), then we simply write \( u - (u_\delta)_\varepsilon = u - u_\delta + u_\delta - (u_\delta)_\varepsilon \) and apply the gathered estimates to \( u - u_\delta \) and \( u_\delta - (u_\delta)_\varepsilon \) separately. The resulting integrals involving \( u_\delta \) are then treated by Jensen’s inequality by

\[
\int_{U} e((L(u - \rho_{\varepsilon} \ast u))) \, dx \leq C \max\{ (L\varepsilon), (L\varepsilon)^2 \} \int_{U + \text{B}(0, \sqrt{n} + \delta)} e((\varepsilon u))
\]

4.3. **Proof of Theorem 1.2.** After the preparations of the previous section, we now carry out the steps (i), (ii) and (iii) as outlined in Section 4.1 above.
4.3.1. Estimates for comparison maps. Throughout this paragraph, we fix \( m \in \mathbb{R}^{n\times n}_\text{sym} \), \( \sigma > 0 \) and assume that \( f \in C^2(\mathbb{R}^{n\times n}_\text{sym}) \) satisfies \( |\lambda|^2 \leq \langle f''(m), z \rangle \leq |\lambda|^2 \) for some \( 0 < \lambda \leq \Lambda < \infty \) and all \( z \in \mathbb{R}^{n\times n}_\text{sym} \). Moreover, we assume that there exists a bounded and non-decreasing function \( \omega: \mathbb{R}_0^+ \to \mathbb{R}_0^+ \) with \( \lim_{t \to 0} \omega(t) = 0 \) such that for all \( m' \in \mathcal{B}(m, \sigma) \) we have

\[
[f''(m) - f''(m')] \leq \omega(|m - m'|).
\]

Finally, for \( 0 < r < R \) and \( v \in C^{1,\alpha}(\mathcal{B}(z, r); \mathbb{R}^n) \) we put

\[
\text{dev}(v; z, r) := \int_{\mathcal{B}(z, r)} f(\varepsilon(v)) \, dx - \inf \left\{ \int_{\mathcal{B}(z, r)} f(\varepsilon(w)) \, dx : w \in C^{1,\alpha}(\mathcal{B}(z, r); \mathbb{R}^n) \cup \{0\}, w = v \text{ on } \partial \mathcal{B}(z, r) \right\},
\]

\[
\mathfrak{t}(v; z, r) := \sup_{\mathcal{B}(z, r)} |\varepsilon(v) - m| + 2^\alpha r^\alpha |\varepsilon(v)|_{C^{\alpha, \alpha}(\mathcal{B}(z, r); \mathbb{R}^{n\times n})}.
\]

Notice that \( \text{dev} \) is an indicator of how far \( v \) is away from minimising the variational integral \( F \) restricted to \( \mathcal{B}(z, r) \). Besides, the function \( \mathfrak{t} \) will prove useful to find the mentioned smallness condition which is necessary to infer the decay estimate of the Hölder continuous comparison maps. Now we have the following result.

**Proposition 4.5.** Let \( 0 < \alpha < 1 \). Then there exists \( c_1 > 0 \) such that the following holds: If \( v \in C^{1,\alpha}(\mathcal{B}(z, R/2); \mathbb{R}^n) \) satisfies \( \mathfrak{t}(v; z, R/2) < \min\{\sigma/c_1, 1\} \), then there exists a bounded, non-decreasing function \( \vartheta: \mathbb{R}_0^+ \to \mathbb{R}_0^+ \) with \( \lim_{t \to 0} \vartheta(t) = 0 \) such that for all \( 0 < r < R/2 \) we have

\[
\int_{\mathcal{B}(z, r)} |\varepsilon(v) - (\varepsilon(v))_{z, r}|^2 \, dx \lesssim \left( \frac{r}{R} \right)^{n+2} \int_{\mathcal{B}(z, R/2)} |\varepsilon(v) - (\varepsilon(v))_{z, R/2}|^2 \, dx
\]

\[
+ \vartheta(\mathfrak{t}(v; z, R/2)) \int_{\mathcal{B}(z, R/2)} |\varepsilon(v) - m|^2 \, dx + \text{dev}(v; z, R/2).
\]

Here, the constants implicit in \( \lesssim \) only depend on \( \lambda, \Lambda \) and \( n \), and \( \vartheta \) only depends on \( \lambda, \Lambda, n \) and \( \omega \).

**Proof.** The strategy of the proof is as follows: Aiming to linearise, we pass to the second order Taylor polynomial of the integrand \( f \) to obtain an integrand \( g \) of quadratic growth. Then we apply the classical decay estimates for linear elliptic systems to its minimisers and prove that the conditions of the present proposition are tantamount for these decay estimates inherit to \( v \) in a way such that (4.10) follows.

We therefore begin by defining the auxiliary integrand \( g: \mathcal{B}(m, \sigma) \to \mathbb{R} \) through

\[
g(z) := f(m) + \langle f'(m), (z - m) \rangle + \frac{1}{2} \langle f''(m)(z - m), (z - m) \rangle, \quad z \in \mathcal{B}(m, \sigma).
\]

Using a Taylor expansion of \( f \) up to order two around \( m \), we deduce by (4.9) that

\[
|f(z) - g(z)| \lesssim \omega(|z - m|)|z - m|^2, \quad z \in \mathcal{B}(m, \sigma).
\]

By Lemma 6.2, the solution \( h \) of the auxiliary minimisation problem

\[
\int_{\mathcal{B}(z, R/2)} \varepsilon(h) \, dx \quad \text{among all } w \in W^{1,2}_v(B(z, R/2); \mathbb{R}^n),
\]

where \( W^{1,2}_v(B(z, R/2); \mathbb{R}^n) = v + W^{1,2}_v(B(z, R/2); \mathbb{R}^n) \), belongs to \( C^{1,\alpha}(B(z, R/2); \mathbb{R}^n) \). In consequence, we have

\[
\int_{\mathcal{B}(z, r)} |\varepsilon(h) - (\varepsilon(h))_{z, r}|^2 \, dx \lesssim \left( \frac{r}{R} \right)^{n+2} \int_{\mathcal{B}(z, R/2)} |\varepsilon(h) - (\varepsilon(h))_{z, R/2}|^2 \, dx
\]

for all \( 0 < r < R/2 \). Moreover, Lemma 6.2 (b) gives

\[
|\varepsilon(h)|_{C^{\alpha, \alpha}(B(z, R/2); \mathbb{R}^{n\times n})} \lesssim |\varepsilon(v)|_{C^{\alpha, \alpha}(B(z, R/2); \mathbb{R}^{n\times n})},
\]

and in both (4.13) and (4.14) the constants implicit in \( \lesssim \) only depend on \( n, \lambda, \Lambda \). We will now compare \( v \) with \( h \). To this end, we first notice that

\[
\int_{\mathcal{B}(z, r)} |\varepsilon(v) - (\varepsilon(v))_{z, r}|^2 \, dx \lesssim \int_{\mathcal{B}(z, r)} |\varepsilon(v) - \varepsilon(h)|^2 \, dx
\]
\[
\begin{align*}
&+ \int_{B(z,r)} |\varepsilon(h) - (\varepsilon(h))_{z,r}|^2 \, dz \\
&+ \int_{B(z,r)} |(\varepsilon(h))_{z,r} - (\varepsilon(v))_{z,r}|^2 \, dx =: \text{I} + \text{II} + \text{III}.
\end{align*}
\]
We shall estimate \( \text{II} \) by means of (4.13). Keeping this in mind, we turn to the remaining two terms \( \text{I} \) and \( \text{III} \). By Jensen’s inequality, we firstly find
\[
\begin{align*}
\text{I} + \text{III} & \lesssim \int_{B(z,r)} |\varepsilon(v) - \varepsilon(h)|^2 \, dx + \int_{B(z,r)} \int_{B(z,r)} |\varepsilon(v)(y) - \varepsilon(h)(y)|^2 \, dy \, dx \\
& \lesssim \int_{B(z,r)} |\varepsilon(v) - \varepsilon(h)|^2 \, dx,
\end{align*}
\]
and by the minimality property of \( h \) we deduce through (4.13) that
\[
\text{II} \lesssim \left( \frac{n}{R} \right)^{n+2} \int_{B(z,R/2)} |\varepsilon(h) - (\varepsilon(h))_{z,R/2}|^2 \, dx \lesssim \left( \frac{n}{R} \right)^{n+2} \int_{B(z,R/2)} |\varepsilon(v) - (\varepsilon(v))_{z,R/2}|^2 \, dx.
\]
We then combine the estimates for \( \text{I}, \text{II} \) and \( \text{III} \), we need to estimate
\[
\int_{B(z,R/2)} |\varepsilon(v) - \varepsilon(h)|^2 \, dx \lesssim \int_{B(z,R/2)} \langle f''(m)(\varepsilon(v) - \varepsilon(h)), (\varepsilon(v) - \varepsilon(h)) \rangle \, dx
\]
\[
\lesssim \int_{B(z,R/2)} g(\varepsilon(v)) - g(\varepsilon(h)) \, dx
\]
\[
\lesssim \int_{B(z,R/2)} g(\varepsilon(v)) - f(\varepsilon(v)) \, dx
\]
\[
+ \int_{B(z,R/2)} f(\varepsilon(v)) - f(\varepsilon(h)) \, dx
\]
\[
+ \int_{B(z,R/2)} f(\varepsilon(h)) - g(\varepsilon(h)) \, dx =: \text{I}_1 + \text{I}_2 + \text{I}_3.
\]
To justify (\( \star \)), we obtain (recall that \( f''(m) \in \mathbb{R}^{n \times n}_{\text{sym}} \))
\[
\langle f''(m)(\varepsilon(v) - m), (\varepsilon(v) - m) \rangle - \langle f''(m)(\varepsilon(h) - m), (\varepsilon(h) - m) \rangle
\]
\[
= \langle f''(m)\varepsilon(v), \varepsilon(v) \rangle - \langle f''(m)\varepsilon(h), \varepsilon(h) \rangle - 2\langle f''(m)m, \varepsilon(v - h) \rangle.
\]
Hence we obtain
\[
\int_{B(z,R/2)} g(\varepsilon(v)) - g(\varepsilon(h)) \, dx = \int_{B(z,R/2)} \langle f'(m), \varepsilon(v - h) \rangle \, dx
\]
\[
+ \frac{1}{2} \int_{B(z,R/2)} \langle f''(m)\varepsilon(v), \varepsilon(v) \rangle - \langle f''(m)\varepsilon(h), \varepsilon(h) \rangle \, dx
\]
\[
- \int_{B(z,R/2)} \langle f''(m)m, \varepsilon(v - h) \rangle \, dx =: \text{J}_1 + \text{J}_2 + \text{J}_3.
\]
On the other hand, expanding terms yields
\[
\int_{B(z,R/2)} \langle f''(m)(\varepsilon(v) - \varepsilon(h)), (\varepsilon(v) - \varepsilon(h)) \rangle \, dx
\]
\[
= \int_{B(z,R/2)} \langle f''(m)\varepsilon(v), \varepsilon(v) \rangle \, dx - 2 \int_{B(z,R/2)} \langle f''(m)\varepsilon(h), \varepsilon(v) \rangle + \int_{B(z,R/2)} \langle f''(m)\varepsilon(h), \varepsilon(h) \rangle \, dx
\]
\[
= \int_{B(z,R/2)} \langle f''(m)\varepsilon(v), \varepsilon(v) \rangle \, dx - 2 \int_{B(z,R/2)} \langle f''(m)\varepsilon(h), \varepsilon(v - h) \rangle \, dx
\]
\[
- \int_{B(z,R/2)} \langle f''(m)\varepsilon(h), \varepsilon(h) \rangle \, dx = 2\text{J}_2 - 2 \int_{B(z,R/2)} \langle f''(m)\varepsilon(h), \varepsilon(v - h) \rangle \, dx =: 2\text{J}_2 - \text{J}_4.
\]
Then, working from the Euler-Lagrange equation associated with (4.12), we conclude by \( v - h \in W^{1,2}_0(B(z, \frac{R}{2}); \mathbb{R}^n) \) that \( \text{J}_1 = \text{J}_3 = \text{J}_4 = 0 \). This yields (\( \star \)), and we can turn to the estimation of
I₁, I₂, I₃. Among them, we readily obtain through the definition of dev and h that I₂ ≤ dev(v; z, R/2).
Now notice that h solves the auxiliary problem (4.12) and hence a routine estimation yields ∥e(h) − m∥₂(B(z, R/2); ℝⁿ × n) ≤ ∥e(v) − m∥₂(B(z, R/2); ℝⁿ × n) with the constants implicit in ‘≤’ only depending on λ, Λ. Therefore we deduce by virtue of (4.14)
\[
\sup_{B(z, R/2)} |e(h) − m| ≤ \sup_{B(z, R/2)} |e(h) − (e(h))_{z, R/2}| + \sup_{B(z, R/2)} |(e(h))_{z, R/2} − m|
\leq R^\alpha|e(h)|_{C^{\alpha, \sigma}(B(z, R/2); \mathbb{R}^n \times n)} + \left(\int_{B(z, R/2)} |e(h) − m|^2 \, dx\right)^{\frac{1}{2}}
\leq R^\alpha|e(v)|_{C^{\alpha, \sigma}(B(z, R/2); \mathbb{R}^n \times n)} + \left(\int_{B(z, R/2)} |e(v) − m|^2 \, dx\right)^{\frac{1}{2}}
\leq R^\alpha|e(v)|_{C^{\alpha, \sigma}(B(z, R/2); \mathbb{R}^n \times n)} + \sup_{B(z, R/2)} |e(v) − m|
\]
where c₀ = c₀(λ, Λ, n) > 0 is a constant. At this point we make our definition of c₁ > 0 as it appears in the assumptions of the present proposition by putting c₁ := c₀. Then, by assumption, we have that e(h)(x) ∈ B(m, σ) provided x ∈ B(z, R/2) and so \(|f(e(h)(x)) − g(e(h)(x))| \leq \omega(|e(h)(x) − m|)|e(h)(x) − m|^2\) for all x ∈ B(z, R/2) by (4.11). On the other hand, \(|e(h) − m|\|_{L^2(B(z, R/2); \mathbb{R}^n \times n)} \leq \|e(v) − m\|_{L^2(B(z, R/2); \mathbb{R}^n \times n)}\) and thus, in consequence, we obtain for I₃
\[
\int_{B(z, R/2)} f(e(h)) − g(e(h)) \, dx \leq \omega(c₀t(v; z, R/2)) \int_{B(z, R/2)} |e(h) − m|^2 \, dx
\leq \omega(c₀t(v; z, R/2)) \int_{B(z, R/2)} |e(v) − m|^2 \, dx.
\]
The remaining integral I₁ is estimated as follows: By assumption, we have the estimate t(v; z, R/2) < min{σ/c₁, 1} so that it holds e(h)(x) ∈ B(m, σ) for all x ∈ B(z, R/2). Referring to (4.11) we then conclude
\[
\int_{B(z, R/2)} g(e(v)) − f(e(v)) \, dx \leq \omega(t(v; z, R/2)) \int_{B(z, R/2)} |e(v) − m|^2 \, dx
\]
so that the claim follows for \(δ\) being a suitable multiple of \(ω(\cdot) + ω(c₀t)\). The proof is complete. □

4.3.2. Smoothing and selection of good radii. In this section we concentrate on step (ii) and establish the required adjusting of the smoothing parameters. We begin with the following lemma and its corollary whose proofs closely follow [9, Lem. 4.2] but with a slight change in the relevant constants. Here and in all what follows, we choose and fix a constant λ > 1; for instance, λ := 1 + \(\frac{1}{100}\) will do.

**Lemma 4.6.** Let u ∈ BD(ℝⁿ), x₀ ∈ ℝⁿ, r > 0 and put m := (Eu)ₓ₀,r. Fix 0 < α < 1 and suppose that \(\hat{E}(u; x₀, r) < 1\). Then there exist 0 < γ < \(\frac{1}{n + 2α}\) and β > 0 such that if
(4.15)
\[
δ = ε = \frac{1}{48\sqrt{nΛ}}\hat{E}(u; x₀, r)^γ,
\]
then t(uδ,ε; x₀, r/2) ≤ \(\hat{E}(u; x₀, r)^β\); note that the appearance of m is implicit in the definition of t. Here, the constants implicit in ‘≤’ only depend on n, γ and α.

**Proof.** First observe that, as a consequence of the elementary estimates for convolutions, we obtain
(4.16)
\[
t(uδ,ε; x₀, r/2) \leq ε \left(1 + \left(\frac{r}{δ}\right)^{α}\right) \sup_{x ∈ B(x₀, r/2 + δ)} |e(uε) − m|.
\]
In fact, we have
(4.17)
\[
|e(uδ,ε)(x) − m| = |(e(uε) − m)_δ(x) |
\]
and hence $\sup_{B(x, r/2)} |\varepsilon(u_{\delta, \varepsilon})(x) - m| \leq \sup_{B(x_0, r/2 + \delta)} |\varepsilon(u_{\varepsilon})(x) - m|$. On the other hand, for any radially symmetric mollifier $\eta$ there exists a constant $c_\eta > 0$ such that for all $g \in L^1(\mathbb{R}^n; \mathbb{R}^{n \times n})$ and $\delta > 0$ there holds

$$[\hat{g} * \eta]_{C^{0, \alpha}}(B(x_0, r/2 + \delta)) \leq \frac{c_\eta}{\delta^\alpha} \sup_{B(x_0, r + \delta)} |g - \xi| \quad \text{for all } \xi \in \mathbb{R}^{n \times n}$$

which can be established by straightforward computation. Therefore,

$$r^\alpha [\varepsilon(u_{\delta, \varepsilon})]_{C^{0, \alpha}}(B(x_0, r/2)) \leq C \left( \frac{r}{\delta} \right)^\alpha \sup_{B(x_0, r/2 + \delta)} |\varepsilon(u_{\varepsilon}) - m|.$$  

In consequence, adding (4.17) and (4.19) yields (4.16), and in order to arrive at the claimed estimate, we must give an estimate for $\sup_{B(x_0, r/2 + \delta)} |\varepsilon(u_{\varepsilon}) - m|.$

Now recall that $\varepsilon$ and $\delta$ are adjusted according to (4.15). Then, by Jensen’s inequality and since $\varepsilon < r$,

$$e(\varepsilon(u_{\varepsilon})(x) - m) \leq \int_{B(x, r)} e(\varepsilon(u) - m) \, dy \leq \left( \frac{r}{\delta} \right)^n \hat{E}(u; x_0, r) \leq K \hat{E}(u; x_0, r)^{1 - n\gamma} \leq K,$$

where $K = (48\sqrt{m\lambda})^n$. Here, the last estimate is valid due to our assumption $\hat{E}(u; x_0, r) < 1$ and $1 - n\gamma > 1 - \frac{n}{n + d} > 0$ as $\alpha \in (0, 1)$. For any $K > 0$ there exists a constant $c = c(K)$ such that if $|z| \leq K$, then $|z|^2 \leq c \varepsilon(z)$, cf. Lemma 2.3. Applying this to our choice of $K$ and using (4.20), we obtain for all $x \in B(x_0, \frac{2}{3} + r)$

$$|\varepsilon(u_{\varepsilon})(x) - m|^2 \leq c e(\varepsilon(u_{\varepsilon})(x) - m) \leq \hat{E}(u; x_0, r)^{1 - n\gamma}.$$  

Now observe that by (4.16) we have by the specific choice of $\delta$ by (4.15) and $\hat{E}(u; x_0, r) < 1$ that

$$t(u_{\delta, \varepsilon}; x_0, r/2) \leq c \left( 1 + (\hat{E}(u; x_0, r))^{\gamma \alpha} \right)^\alpha \hat{E}(u; x_0, r)^{\frac{\gamma}{\alpha}} \leq \hat{E}(u; x_0, r)^\beta$$

with $\beta := \frac{1}{\gamma}(1 - n\gamma) - \gamma \alpha > 0$ (recall that $0 < \gamma < 1/(n + 2\alpha)$). This proves the claim. 

**Corollary 4.7.** *In the situation of Lemma 4.6, we have*

$$\int_{B(x_0, r/2)} |\varepsilon(u_{\delta, \varepsilon}) - m|^2 \, dx \lesssim \hat{E}(u; x_0, r), \quad \text{and}$$

$$\int_{B(x_0, r/2)} |\varepsilon(u_{\delta, \varepsilon}) - (\varepsilon(u_{\delta, \varepsilon}))_{B(x_0, r/2)}|^2 \, dx \lesssim \hat{E}(u; x_0, r).$$

*Again, the constants implicit in “$\lesssim$” only depend on $n$, $\alpha$ and $\gamma$ as specified and fixed in Lemma 4.6.*

**Proof.** We recall from Lemma 4.6 that $t(u_{\delta, \varepsilon}; x_0, r/2) \leq C \hat{E}(u; x_0, r)^\beta \leq C, C > 0$ only depending on $n$, $\alpha$ and $\gamma$. In particular, there holds $\sup_{x \in B(x_0, r/2)} |\varepsilon(u_{\delta, \varepsilon}) - m| \leq C$. Therefore we have $|\varepsilon(u_{\delta, \varepsilon})(x) - m|^2 \leq c e(\varepsilon(u_{\delta, \varepsilon}) - m)$ for all $x \in B(x_0, r/2)$ with a constant $c > 0$ that only depends on $C$ and $e$. We conclude by Jensen’s inequality that

$$\int_{B(x_0, r/2)} |\varepsilon(u_{\delta, \varepsilon}) - m|^2 \, dx \lesssim \int_{B(x_0, r/2)} e(\varepsilon(u_{\delta, \varepsilon}) - m) \, dx \lesssim \int_{B(x_0, r/2)} (e(\varepsilon(u) - m))_{\delta, \varepsilon} \, dx = \hat{E}(u; x_0, r),$$

the last estimate being valid due to $\frac{2}{3} + \delta + \varepsilon < r$ by assumption. The second inequality directly follows from this and hence the complete statement of the corollary.

The previous lemma and corollary establish that, once the smoothing parameters are chosen suitably, then the $\varepsilon$-deviation will be small provided the excess $\hat{E}$ is. This assertion works without the assumption of minimality, and so does the next lemma, the proof of which is exactly along the lines of [9] and thus is omitted. Still, it is convexity that crucially enters by means of Jensen’s inequality here:

**Lemma 4.8.** *Let $u \in BD(\Omega)$ and let $\varepsilon, \delta, R > 0$ and $z \in \Omega$ be such that $B(z, R) \subseteq \Omega$. Then the following holds for any $f \in C^2(\mathbb{R}^{n \times n})$ with (LG’):*
Let \( f \in \mathcal{F} \). Moreover, since \( f \approx e \), we also have \( \tilde{f} \approx e \). Given \( w : B(z, R) \to \mathbb{R}^n \) and \( m' \in \mathbb{R}^{n \times n}_{\text{sym}} \), we define \( \tilde{w} : B(z, R) \to \mathbb{R}^n \) by
\[
\tilde{w}(x) := w(x) - m'(x - z).
\]
and record the connection of the variational integrals corresponding to \( f \) and \( \tilde{f} \) first:

**Lemma 4.9.** Let \( u, v \in \text{BD}(B(z, R)) \) such that \( u = v \mathcal{H}^{n-1} \text{-a.e. on } \partial B(z, R) \) in the sense of traces and define \( \tilde{f} \) and \( \tilde{u}, \tilde{v} \) by (4.23) and (4.24), respectively. Then there holds
\[
\int_{B(z, R)} (f(Eu) - f(Eu)) = \int_{B(z, R)} (\tilde{f}(E\tilde{u}) - \tilde{f}(E\tilde{v})).
\]

**Proof.** Writing out the definitions, we obtain for \( u, v \in \text{LD}(B(z, R)) \)
\[
\int_{B(z, R)} \tilde{f}(\varepsilon(\tilde{u})) - \tilde{f}(\varepsilon(\tilde{v})) \, dx = \int_{B(z, R)} f(m' + \varepsilon(\tilde{u})) - f(m' + \varepsilon(\tilde{v})) \, dx
\]
\[- \int_{B(z, R)} f'(m') \varepsilon(\tilde{u}) - f'(m') \varepsilon(\tilde{v}) \, dx \]
\[
= \int_{B(z, R)} f(\varepsilon(u)) - f(\varepsilon(v)) \, dx - \int_{B(z, R)} f'(m') \varepsilon(\tilde{u}) - f'(m') \varepsilon(\tilde{v}) \, dx.
\]

Now notice that \( \varepsilon(\tilde{u}) - \varepsilon(\tilde{v}) = \varepsilon(u) - \varepsilon(v) \) and so, by the Gauß-Green theorem on \( \text{BD} \) (cf. (2.3)),
\[
\int_{B(z, R)} f'(m') (\varepsilon(\tilde{u}) - \varepsilon(\tilde{v})) \, dx = \int_{\partial B(z, R)} (f'(m'), \text{Tr}(u - v) \otimes v) \, d\mathcal{H}^{n-1}
\]
\[- \int_{B(z, R)} \text{div}(f'(m')) (u - v) \, dx = 0.
\]

because \( \text{div}(f'(m')) = 0 \) and, by our assumptions on \( u \) and \( v \), \( \text{Tr}(u) = \text{Tr}(v) \mathcal{H}^{n-1} \text{-a.e. on } \partial B(z, R) \). The claim follows. \( \square \)

**Proposition 4.10.** Let \( f : \mathbb{R}^{n \times n}_{\text{sym}} \to \mathbb{R} \) be a convex function with (LG'). Then there exists \( \alpha > 0 \) such that for any generalised minimiser \( u \in \text{BD}(\Omega) \) of \( F \) the following holds: Suppose that for a ball \( B(z, R) \subseteq \Omega \) there holds
\[(a) \ (Eu)_{z, R} = m, \]
\[(b) \ (E(u; z, R) < a, \]
\[(c) \ there exists \( \sigma > 0 \) such that \( f \) is of class \( C^2 \) in \( \mathbb{B}(m, \sigma) \) verifying
\[
|f''(m) - f''(m')| \leq \omega(|m - m'|) \quad \text{for all } m' \in \mathbb{B}(m, \sigma)
\]
with a non-decreasing function \( \omega : \mathbb{R}_+^0 \to \mathbb{R}_+^+ \) which, in addition, satisfies \( |\omega(t)| \leq M = M(|m|) \) for all \( t \in \mathbb{R} \geq 0 \) and \( \lim_{t \to 0} \omega(t) = 0 \), such that we have
\[
\lambda |Z|^2 \leq (f''(m)Z, Z) \leq \Lambda |Z|^2
\]
Then there exists a constant \( c > 0 \) and a bounded and non-decreasing function \( h: \mathbb{R}^+ \to \mathbb{R}^+ \) with \( \lim_{\lambda \searrow 0} h(t) = 0 \) such that

\[
E(u; z, r) \leq cE(v; z, 2r) + ch(\tilde{E}(u; z, R)) \left( 1 + \left( \frac{R}{r} \right)^{n+1} \right) E(u; z, R)
\]

holds for all \( 0 < r < R/4 \). Here we have set \( v := u_{\delta, \varepsilon} \) with \( \delta, \varepsilon \) adjusted according to Lemma 4.6 applied to the radius \( R > 0 \).

**Proof.** We split the proof into three steps.

**Step 1. Preliminaries.** Let \( 0 < r < R/4 \) and put \( m' := (\varepsilon(v))_{z, r} \). Due to our choice of \( \delta \) and \( \varepsilon \), Lemma 4.6 applied to the radius \( R > 0 \) yields

\[
\sup_{B(z, r/4)} |\varepsilon(v) - m| + R^\alpha |\varepsilon(v)|_{C^{0, \alpha}(B(z, r))} \lesssim c\tilde{E}(u; z, R)^{\beta}
\]

for \( \beta > 0 \) as in Lemma 4.6 (note that both \( c > 0 \) and \( \beta > 0 \) in the previous inequality depend on \( n, \alpha \) and \( \gamma \) as in Lemma 4.6, however, the latter are thought as fixed throughout). In consequence, we have because of \( B(z, r) \subset B(z, \frac{r}{4}) \)

\[
|m' - m| \leq \int_{B(z, r)} |\varepsilon(v) - m| \, dx \lesssim \sup_{B(z, \frac{r}{4})} |\varepsilon(v) - m| \leq c\tilde{E}(u; z, R)^{\beta}.
\]

Thus we may choose

\[
0 < a < \min \left\{ \frac{1}{2}, \left( \frac{3}{25} \right)^{\frac{1}{\gamma}} \right\}
\]

sufficiently small such that \( \tilde{E}(u; z, R) < a \) yields \( |m - m'| < \sigma/2 \); here, \( \gamma \) is as in Lemma 4.6. Since \( B(m', \frac{r}{2}) \subset B(m, \sigma) \), (4.25) and (4.26) continue to hold in \( B(m', \frac{r}{2}) \). Let us go back to the definition of \( \tilde{f} \) as given in (4.23), where we have the appearance of \( m' \). Here we have the estimate \( \tilde{f} \approx e \), that is, there exist two constants \( c^{(1)}, c^{(2)} > 0 \) such that \( c^{(1)}e(\xi) \leq \tilde{f}(\xi) \leq c^{(2)}e(\xi) \), and it is easily seen that these constants do not depend on the specific point \( m' \).

**Step 2. Comparison Estimates.** We observe that, as a consequence of Lemma 2.3, Jensen’s inequality and \( \tilde{f} \approx e \), for all \( 0 < r < \frac{R}{4} \) there holds

\[
E(u; z, r) \lesssim \int_{B(z, r)} e(\tilde{E}u - m') + \int_{B(z, r)} e((\tilde{E}u)_{z, r} - m') \, dx
\]

\[
\lesssim \int_{B(z, r)} e(\tilde{E}u - m') \, dx \lesssim \int_{B(z, r)} \tilde{f}(\tilde{E}u).
\]

Now fix \( R/2 < t < s < \frac{3R}{4} \) and define \( \rho \in C_c(B(z, R); [0, 1]) \) by

\[
\rho(x) := \frac{2}{s-t}(\delta = \frac{R}{2}) \mathbb{1}_{(\delta \leq |x| \leq (s+t)/2)}(x) + \mathbb{1}_{(|x| > (s+t)/2)}(x), \quad x \in B(z, R).
\]

Then we have \( \tilde{v} + \rho(\tilde{u} - \tilde{v}) \in BD(\Omega) \) and, in particular, \( \tilde{v}|_{\partial B(z, R)} = \tilde{u}|_{\partial B(z, R)} \). By Lemma 4.9, \( \tilde{u} \) minimises the functional

\[
w \mapsto \int_{B(z, R)} \tilde{f}(\varepsilon(w)) \, dx
\]

with respect to its own boundary values. Using this in the second step, we deduce

\[
\int_{B(z, s)} \tilde{f}(\varepsilon(\tilde{u})) \, dx \leq \left( \int_{B(z, r)} + \int_{B(z, t) \setminus B(z, r)} + \int_{B(z, s) \setminus B(z, t)} \right) \tilde{f}(\varepsilon(\tilde{u})) \, dx
\]

\[
\leq \int_{B(z, r)} \tilde{f}(\varepsilon(\tilde{v})) \, dx + \int_{B(z, t) \setminus B(z, r)} \tilde{f}(\varepsilon(\tilde{v})) \, dx
\]

\[
+ \int_{B(z, s) \setminus B(z, t)} \tilde{f}(\varepsilon(\tilde{v} + \rho(\tilde{u} - \tilde{v}))) \, dx.
\]
We recall estimate (4.29) and regroup terms to obtain
\[
\mathbf{E}(u; z, r) \lesssim \int_{B(z, r)} \tilde{f}(\varepsilon(v)) \, dx + \int_{B(z, t) \setminus B(z, r)} \tilde{f}(\varepsilon(v)) - \tilde{f}(\varepsilon(\tilde{u})) \, dx
\]
\[
+ \int_{B(z, s) \setminus B(z, r)} \tilde{f}(\varepsilon(v + \rho(\tilde{u} - \tilde{v}))) - \tilde{f}(\varepsilon(\tilde{u})) \, dx =: \mathbf{I} + \mathbf{II} + \mathbf{III}.
\]
We now estimate the single terms. In view of \(\mathbf{I}\), we directly work from the definition of \(\tilde{f}\) and \(\tilde{v}\). Then it follows by \(\tilde{f} \approx e\) and the definition of \(m'\) that
\[
(4.32) \quad \mathbf{I} = \int_{B(z, r)} \tilde{f}(\varepsilon(v) - m') \, dx \lesssim \mathbf{E}(v; z, r).
\]
The term \(\mathbf{II}\) is already in a convenient form and shall be dealt with later. Let us put \(A(z; t, s) := B(z, s) \setminus B(z, t)\). We turn to \(\mathbf{III}\); recalling the product rule (i.e., \(e(\varphi u) = \varphi e(u) + \nabla \varphi \otimes u\) for \(\varphi: \mathbb{R}^n \to \mathbb{R}\) and \(u: \mathbb{R}^n \to \mathbb{R}^n\), monotonicity of \(e\) together with Lemma 2.3(i), we then have
\[
\int_{A(z; t, s)} \tilde{f}(\varepsilon(v + \rho(\tilde{u} - \tilde{v}))) \, dx \lesssim \int_{A(z; t, s)} e(\varepsilon(v)) + e(\varepsilon(\tilde{u})) + e\left(\frac{\tilde{u} - \tilde{v}}{s - t}\right) \, dx.
\]
Keeping in mind \(s < \frac{1}{2} R\) and therefore \(s + \varepsilon + \delta < s + \frac{1}{8} R \leq R\), we see that \(A(z; t - 2\varepsilon, s + 2\varepsilon) \subset \Omega\) and so it is admissible to estimate
\[
\int_{A(z; t, s)} e(\varepsilon(v)) \, dx \leq \int_{A(z; t - 2\varepsilon, s + 2\varepsilon)} e(\varepsilon(\tilde{u})) \, dx \leq \int_{A(z; t - 2\varepsilon, s + 2\varepsilon)} \tilde{f}(\varepsilon(\tilde{u})) \, dx.
\]
Here we have used Jensen’s inequality and implicitly the fact that \(\delta = \varepsilon\). Combining the last inequality with Proposition 4.3 and \(\delta = \varepsilon\), it follows that
\[
\int_{A(z; t, s)} \tilde{f}(\varepsilon(v + \rho(\tilde{u} - \tilde{v}))) \, dx \lesssim \int_{A(z; t - 2\varepsilon, s + 2\varepsilon)} \tilde{f}(\varepsilon(\tilde{u})) \, dx
\]
\[
+ \max\left\{\frac{\varepsilon^2}{(s - t)^2}, \frac{\varepsilon^p}{(s - t)^p}\right\} \int_{A(z; t - 2\varepsilon, s + 2\varepsilon + 2\varepsilon)} \tilde{f}(\varepsilon(\tilde{u})) \, dx.
\]
Overall, we get the intermediate estimate
\[
\mathbf{III} \leq \int_{B(z, r) \setminus B(z, t)} \tilde{f}(\varepsilon(v + \rho(\tilde{u} - \tilde{v}))) + \tilde{f}(\varepsilon(\tilde{u})) \, dx
\]
\[
\lesssim \int_{A(z; t - 2\varepsilon, s + 2\varepsilon)} \tilde{f}(\varepsilon(\tilde{u})) \, dx + \max\left\{\left(\frac{\varepsilon}{s - t}\right), \left(\frac{\varepsilon}{s - t}\right)^2\right\} \int_{A(z; t - 2\varepsilon, s + 2\varepsilon + 2\varepsilon)} \tilde{f}(\varepsilon(\tilde{u})) \, dx
\]
and hence, putting together our preliminary estimates for \(\mathbf{I}\), \(\mathbf{II}\) and \(\mathbf{III}\), we end up with
\[
\mathbf{E}(u; z, r) \lesssim \mathbf{E}(v; z, r) + \int_{B(z, t) \setminus B(z, r)} \tilde{f}(\varepsilon(v)) - \tilde{f}(\varepsilon(\tilde{u})) \, dx
\]
\[
+ \int_{A(z; t - 2\varepsilon, s + 2\varepsilon)} \tilde{f}(\varepsilon(\tilde{u})) \, dx
\]
\[
+ \max\left\{\left(\frac{\varepsilon}{s - t}\right), \left(\frac{\varepsilon}{s - t}\right)^2\right\} \int_{A(z; t - 2\varepsilon, s + 2\varepsilon + 2\varepsilon)} \tilde{f}(\varepsilon(\tilde{u})) \, dx
\]
being valid for any \(0 < r < R/2\), and the constants implicit in ’\(\lesssim\)’ are independent of \(u, \varepsilon, \delta, s\) and \(t\).

**Step 3. Conclusion.** The aim of the remaining proof is to utilise the last inequality for certain choices of \(s, t\). For \(l \in \mathbb{N}_0^+\) denote the integer part of \(l\) by \([l]\) and define \(N := [25(2\hat{E}^{\gamma/2}(u; z, R))^{-1}]\) with \(0 < \gamma < 1/(n + 2\alpha)\) given by Lemma 4.6. For \(k \in \{1, \ldots, N\}\), put
\[
a_k := \frac{5}{8} R + k \frac{R}{400} \hat{E}^{\gamma/2}(u; z, R)
\]
so that \( a_k \in [\frac{3}{4}R, \frac{1}{4}R] \). Due to Lemma 4.8 (ii), the choice \( r < R/4 \) implies that for each \( k = 1, \ldots, N \) there exists \( t_k \in (a_{sk-1}, a_{sk}) \) and \( r_k \in (r, 2r) \) with
\[
\int_{A(z, r_k, t_k)} \tilde{f}(\varepsilon(\tilde{u})) \, dx \leq 2 \left( \frac{\varepsilon + \delta}{a_{sk} - a_{sk-1}} + \frac{\varepsilon + \delta}{r} \right) \int_{B(z, r)} \tilde{f}(E\tilde{u})
\leq 8\varepsilon \max \left\{ \frac{400}{R^\gamma/2(u; z, R)}, \frac{1}{r} \right\} \int_{B(z, r)} \tilde{f}(E\tilde{u})
\leq \frac{R\tilde{E}^\gamma(u; z, R)}{6\lambda n^2} \max \left\{ \frac{400}{R^\gamma/2(u; z, R)}, \frac{1}{r} \right\} \int_{B(z, r)} \tilde{f}(E\tilde{u})
\leq \max \left\{ \frac{70R\tilde{E}^\gamma(u; z, R)}{6\lambda n^2}, \frac{1}{6\lambda n} \right\} \tilde{E}^\gamma(u; z, R) \times \int_{B(z, R)} \tilde{f}(E\tilde{u}).
\]
We can choose \( L > 0 \) so small such that for \( k = 1, \ldots, N \) the annuli
\[
A_k := B(z, s_k + 2\lambda\sqrt{\varepsilon}) \setminus B(z, t_k + 2\lambda\sqrt{\varepsilon}), \quad s_k := t_k + LR\tilde{E}^\gamma/2(u; z, R)
\]
are pairwise disjoint and subsets of \( B(z, R) \); for instance, the choice \( L = \frac{1}{400} \) will do. Let us address this point in detail: Disjointness of \( A_k \) and \( A_{k+1} \) is equivalent to
\[
A_{k+1} \cap A_k = \emptyset \iff t_{k+1} - 2\lambda\sqrt{n} > s_k + 2\lambda\sqrt{n}
\iff t_{k+1} - t_k > \frac{R}{400} \tilde{E}^\gamma/2(u; z, R) + 2\lambda\sqrt{n}
\iff t_{k+1} - t_k > \frac{R}{400} \tilde{E}^\gamma/2(u; z, R) + \frac{R}{12} \tilde{E}^\gamma(u; z, R).
\]
Now note that by construction, \( t_{k+1} - t_k > a_{sk+7} - a_{sk} = \frac{7}{400}R\tilde{E}^\gamma/2(u; z, R) \), and so the last inequality of (4.35) is certainly satisfied provided
\[
\frac{6}{400} > \frac{1}{12} \tilde{E}^\gamma/2(u; z, R), \quad \text{i.e.,} \quad \tilde{E}(u; z, R) < \left( \frac{3}{25} \right)^{\frac{1}{\gamma}},
\]
and the last inequality is valid by virtue of (4.28). Now, \( A_k \subseteq B(z, R) \) as a consequence of
\[
s_N + 2\lambda\sqrt{n} = a_{sN} + \frac{R}{400} \tilde{E}^\gamma/2(u; z, R) + 2\lambda\sqrt{n}
\leq \frac{5}{8}R + \frac{8NR}{400} \tilde{E}^\gamma/2(u; z, R) + \frac{R}{400} \tilde{E}^\gamma/2(u; z, R) + \frac{1}{24}R\tilde{E}^\gamma(u; z, R)
\leq \frac{5}{8}R + \frac{104}{400}R + \frac{R}{400} + \frac{1}{24}R \leq R.
\]
Here we have used the definition of \( s_N \) in terms of \( t_N, t_N \leq a_{sN} \) in the first step, the definition of \( a_{sN} \) in the second and \( \tilde{E}(u; z, R) < 1 \) together with the definition of \( N \) in the third step.

We can therefore conclude that
\[
\int_{A_1} \tilde{f}(\varepsilon(\tilde{u})) \, dx + \ldots + \int_{A_N} \tilde{f}(\varepsilon(\tilde{u})) \, dx \leq \int_{B(z, R)} \tilde{f}(\varepsilon(\tilde{u})) \, dx
\]
and so we find \( k' \in \{1, \ldots, N\} \) such that
\[
\int_{A_{k'}} \tilde{f}(\varepsilon(\tilde{u})) \, dx \leq N^{-1} \int_{B(z, R)} \tilde{f}(\varepsilon(\tilde{u})) \, dx \leq \frac{1}{10} \tilde{E}^\gamma/2(u; z, R) \int_{B(z, R)} \tilde{f}(\varepsilon(\tilde{u})) \, dx,
\]
where (*) holds because of our choice of \( N \).
We go back to \((4.33)\) and insert the choices \(r = r\varepsilon, t = t\varepsilon\) and \(s = s\varepsilon\). For these choices, we obtain by
\[
\frac{\varepsilon}{s - t} = \frac{400}{R\tilde{E}^{\gamma/2}(u; z, R)} \frac{R\tilde{E}^{\gamma}(u; z, R)}{48\lambda \sqrt{n}} \leq \frac{10}{\lambda \sqrt{n}} \frac{3}{25} < 1
\]
that
\[
\max \left\{ \left( \frac{\varepsilon}{s - t} \right), \left( \frac{\varepsilon}{s - t} \right)^2 \right\} = \frac{10}{\lambda \sqrt{n}} \tilde{E}^{\gamma/2}(u; z, R).
\]
Going back to \((4.33)\) and employing \((4.34), (4.36)\), we consequently obtain
\[
\mathbf{E}(u; z, r) \lesssim \mathbf{E}(u; z, 2r) + \max \left\{ \tilde{\mathbf{E}}^{\gamma/2}(u; z, R), \left( \frac{R}{r} \right) \tilde{\mathbf{E}}^{\gamma}(u; z, R) \right\} \int_{B(z,R)} \tilde{f}(\varepsilon(\tilde{u})) \, dx
\]
\[(4.37)\]
\[
\lesssim \mathbf{E}(u; z, 2r) + \left( 1 + \frac{R}{r} \right) \tilde{\mathbf{E}}^{\gamma/2}(u; z, R) \int_{B(z,R)} \tilde{f}(\varepsilon(\tilde{u})) \, dx
\]
for all \(r < R/2\). This is not yet the decay estimate \((4.27)\), so we shall provide a suitable estimate on the last factor on the right side of \((4.37)\). By Lemma 2.3 we conclude
\[
\int_{B(z,R)} \tilde{f}(\varepsilon(\tilde{u})) \lesssim \int_{B(z,R)} e(\varepsilon(u) - m')
\]
\[
\lesssim \left( \int_{B(z,R)} e(\varepsilon(u) - m) \right) + R^n e(m - m')
\]
\[
\lesssim \left( \int_{B(z,R)} e(\varepsilon(u) - m) \right) + R^n e \left( \int_{B(z,r)} \varepsilon(u) - m \, dx \right)
\]
\[
\lesssim \left( \int_{B(z,R)} e(\varepsilon(u) - m) \right) + \left( \frac{R}{r} \right)^n \int_{B(z,r)} e((u - m)_{\lambda,\varepsilon}) \, dx
\]
\[
\lesssim \mathbf{E}(u; z, R) + (R/r)^n \int_{B(z,R)} e(\varepsilon(u) - m) \lesssim (1 + (R/r)^n)\mathbf{E}(u; z, R),
\]
the penultimate estimate being valid due to \(r + \varepsilon + \delta < R\). Combining this estimate with \((4.37)\), we obtain
\[
\mathbf{E}(u; z, r) \lesssim \mathbf{E}(u; z, 2r) + h(\tilde{\mathbf{E}}(u; z, R)) \left( 1 + (R/r)^n \right) \left( 1 + R/r \right) \mathbf{E}(u; z, R)
\]
\[(4.38)\]
we have set \(h(t) := t^{\gamma/2}\); note that the constants implicit in \(\lesssim\) do not depend on \(u, z, r, R, \varepsilon\) or \(\delta\). Since \(h\) obviously is non-decreasing and moreover satisfies \(\lim_{t \to 0} h(t) = 0\), \((4.38)\) implies the claim.

**Proposition 4.11.** In the situation of Proposition 4.10 we have
\[
\text{dev}(v; z, R/2) \lesssim h(\tilde{\mathbf{E}}(u; z, R))\mathbf{E}(u; z, R).
\]
Again, the constants implicit in \(\lesssim\) do not depend on \(\varepsilon, \delta, u, z\) or \(R\).

**Proof.** We fix \(R/2 < t < s < R\) and let \(\theta > 0\) be arbitrary. We then put
\[
\mathcal{A} := \{ w \in W^{1,\infty}(B(z, t); \mathbb{R}^n) : w = \tilde{v} \text{ on } \partial B(z, t) \}
\]
\[
\mathcal{B} := \{ w \in W^{1,\infty}(B(z, s) \setminus B(z, t); \mathbb{R}^n) : w = \tilde{v} \text{ on } \partial B(z, t) \cup \partial B(z, s) \}
\]
and find \(w_1 \in \mathcal{A}\) and \(w_2 \in \mathcal{B}\) such that
\[
\int_{B(z,t)} \tilde{f}(\varepsilon(w_1)) \, dx \leq \inf_{w \in \mathcal{A}} \int_{B(z,t)} \tilde{f}(\varepsilon(w)) \, dx + \frac{\theta}{2}
\]
\[
\int_{B(z,s) \setminus B(z,t)} \tilde{f}(\varepsilon(w_2)) \, dx \leq \inf_{w \in B} \int_{B(z,s) \setminus B(z,t)} \tilde{f}(\varepsilon(w)) \, dx + \frac{\theta}{2}.
\]

By Lemma 4.9, we have
\[
\text{dev} \left( v; z, \frac{R}{2} \right) := \int_{B(z, \frac{R}{2})} \tilde{f}(\varepsilon(\tilde{v})) \, dx - \inf \left\{ \int_{B(z, R/2)} \tilde{f}(\varepsilon(w)) \, dx : w \in C^{1,\alpha}(B(z, \frac{R}{2}); \mathbb{R}^n) \right\}.
\]

Since \( w_1, w_2 \) are Lipschitz and coincide on \( \partial B(z, t) \), we deduce that \( w_3 := \| w_1 \|_{B(z,s)} \setminus B(z,t) + w_2 \) belongs to \( W^{1,\infty}(B(z, s); \mathbb{R}^n) \). We then obtain
\[
\text{dev}(\tilde{v}; z, t) \leq \int_{B(z,t)} \tilde{f}(\varepsilon(\tilde{v})) \, dx - \int_{B(z,s)} \tilde{f}(\varepsilon(\tilde{v})) \, dx = \int_{A(z,s)} \tilde{f}(\varepsilon(\tilde{v})) \, dx - \int_{A(z,t)} \tilde{f}(\varepsilon(\tilde{v})) \, dx \leq \text{dev}(\tilde{v}, \lambda).
\]

Next, if we let \( \rho \) be the function given by (4.30), we similarly obtain by exploiting minimality of \( \tilde{u} \) that
\[
\int_{B(z,s)} \tilde{f}(\varepsilon(\tilde{u})) - \tilde{f}(\varepsilon(\tilde{w}_3)) \, dx \leq \int_{B(z,s)} \tilde{f}(\varepsilon(\tilde{w}_3) + \eta(\tilde{u} - \tilde{v})) \, dx \leq \int_{A(z,t,s)} \epsilon(\varepsilon(\tilde{w}_3) + \epsilon(\tilde{u}) + \epsilon(\tilde{v})) + \epsilon \left( \frac{\tilde{u} - \tilde{v}}{s - t} \right) \, dx.
\]

Now note that \( w_3 = w_2 \) on \( B(z, s) \setminus B(z, t) \) and therefore
\[
\int_{A(z,t,s)} \epsilon(\varepsilon(\tilde{w}_3)) \, dx \leq \int_{A(z,t,s)} \tilde{f}(\varepsilon(\tilde{w}_3)) \, dx \leq \int_{A(z,t,s)} \tilde{f}(\varepsilon(\tilde{v})) \, dx + \frac{\theta}{2}.
\]

Going back to the estimate of III in the proof of the preceding Proposition 4.10, we see that the same estimates used there also apply to the present problem, yielding
\[
\int_{B(z,s)} \tilde{f}(\varepsilon(\tilde{u})) - \tilde{f}(\varepsilon(\tilde{w}_3)) \, dx \leq \int_{A(z,t-2\lambda \sqrt{\epsilon}; s+2\lambda \sqrt{\epsilon})} \tilde{f}(\varepsilon(\tilde{u})) + \epsilon \left( \frac{\tilde{u} - \tilde{v}}{s - t} \right) \, dx.
\]

Because \( \theta > 0 \) was assumed arbitrary and the constants which implicitly appear in the above estimates do not depend on \( \theta \), we may pass \( \theta \to 0 \) in these estimates. In combination with our above preliminary estimate on \( \text{dev}(v; z, R/2) \), we then obtain
\[
\text{dev}(v; z, R/2) \leq \int_{B(z,s)} \tilde{f}(\varepsilon(\tilde{v})) - \tilde{f}(\varepsilon(\tilde{u})) + \int_{A(z,t-2\lambda \sqrt{\epsilon}; s+2\lambda \sqrt{\epsilon})} \tilde{f}(\varepsilon(\tilde{v})) + \frac{\theta}{2}.
\]

In the same way as it is done in [9], we now view this last estimate as the substitute of (4.33) in the proof of Proposition 4.10. Working from this observation and employing the same strategy to find suitable choices of \( s \) and \( t \) as outlined in the proof of Proposition 4.10, the statement of the present proposition follows.

We eventually come to the

**Proof of Proposition 4.1.** Let \( z \in \Omega, \, R > 0 \) be as in Proposition 4.1 and put \( m := (E u)_{z, R} \). Moreover, let \( 0 < r < R/4 \). Then \( 2r < R/2 \) and hence, due to

\[
\int_{B(z,r)} \tilde{f}(\varepsilon(\tilde{v})) \, dx \leq \int_{B(z,s) \setminus B(z,t)} \tilde{f}(\varepsilon(w_2)) \, dx \leq \int_{B(z,s) \setminus B(z,t)} \tilde{f}(\varepsilon(w)) \, dx + \frac{\theta}{2}.
\]
Lemma 4.6, \(\tau_1 > 0\) so small such that (with \(\varepsilon, \delta\) adjusted as indicated therein) \(\tilde{E}(u; z, R) < \tau_1\) implies \(t(u_{\delta, z}; z, R/2) < \min\{\sigma/c_1, 1\}\), where \(c_1 > 0\) is as in Proposition 4.5. Then the latter yields, using Lemma 2.3 (ii), \(0 < 2r < R/2\) and Corollary 4.7,
\[
E(v; z, 2r) \lesssim \left( \left( \frac{r}{R} \right)^{n+2} + \vartheta(\tilde{E}^\beta(v; z, R)) \right) E(u; z, R) + \text{dev}(v; z, R/2).
\]

- Proposition 4.5, we choose \(\tau_2 := a\). Adopting the terminology in this statement, we see that \(\tilde{E}(u; z, R) < \tau_2\) implies
\[
E(u; z, r) \lesssim E(v; z, 2r) + h(\tilde{E}(u; z, R))(1 + \left( \frac{R}{r} \right)^{n+1}) E(u; z, R).
\]

On the other hand, Proposition 4.10 yields that \(h(u_{\delta, z}; z, R/2) \lesssim h(\tilde{E}(u; z, R))E(u; z, R)\). In conclusion, we find with \(\tau := \min\{\tau_1, \tau_2\}\)
\[
E(u; z, r) \lesssim \left( \left( \frac{r}{R} \right)^{n+2} + \vartheta(\tilde{E}(u; z, R))(1 + \left( \frac{R}{r} \right)^{n+1}) \right) E(u; z, R),
\]
where \(\vartheta(t) := \vartheta(|t|^\beta) + 2h(t)\) obviously matches the conditions required in Proposition 4.1.

The proof is complete. \(\square\)

## 5. Remarks and Extensions

We conclude the paper with some remarks on possible generalisations of Theorems 1.1 and 1.2; we begin with the latter. Let us first emphasize, in analogy with [9, Sec. 6], that under suitable smoothness assumptions on \(f: \Omega \times R^{n \times n} \to R\) and \(g: \Omega \times R^n \to R\), variational integrals of the form
\[
F[u; \Omega] := \int_\Omega f(x, \varepsilon(u)) \, dx + \int_\Omega g(x, u) \, dx
\]
can be handled analogously by the method presented in Section 4 under the following conditions (cf. [9, (6.2)–6.5]):

\[
\begin{align*}
&\{c_1|z| - c_2 \leq f(x, z) \leq c_3(1 + |z|)\} \quad \text{for all } x \in \Omega, \ z \in R^{n \times n}, \\
&\{|f(x_1, z) - f(x_2, z)| \leq c_4|x_1 - x_2|^{\gamma}(1 + |z|)\} \quad \text{for all } x_1, x_2 \in \Omega, \ z \in R^{n \times n}, \\
&\{|g(x_1, y_1) - g(x_2, y_2)| \leq c_5||x_1 - x_2| + |y_1 - y_2||^{\gamma}\} \quad \text{for all } x_1, x_2 \in \Omega, \ y_1, y_2 \in R^n,
\end{align*}
\]

where \(c_1, \ldots, c_5 > 0\) and \(0 < \gamma < 1\) are constants. In this case, if \(f\) moreover is of class \(C^2\) and for each \(x_0 \in \Omega\) and \(z \in R^{n \times n}\) there holds for some \(\lambda > 0\)
\[
|\lambda|\xi|^2 \leq \langle f''(x, z)\xi, \xi \rangle \quad \text{for all } \xi \in R^{n \times n}
\]
uniformly for \(x\) in a neighbourhood of \(x_0\), then generalised minima of (5.1) are equally \(C^1,\alpha\)-partially regular. Here, the dashes appearing in \(f''\) are taken with respect to the symmetric gradient variable exclusively.

However, let us note that it is not clear to us how an analogous result should be proved if the overall variational integrand \(f(x, \xi) + g(x, z)\) does not possess the splitting structure but is of the form \(f(x, z, \xi)\). In this situation, a common device is to employ a Caccioppoli inequality in conjunction with Gehring’s lemma to arrive at a higher integrability result, but this proof scheme is somewhat ruled out in the linear growth case:

**Remark 5.1** (Caccioppoli and Gehring). To cope with fully non-autonomous integrands \(f: (x, y, z) \mapsto f(x, y, z)\), one usually invokes Caccioppoli’s inequality in conjunction with Gehring’s lemma on higher integrability to conclude that minima of elliptic problems belong to some \(W^{1,p}_{\text{loc}}\), \(p > r\), where \(p\) is the Lebesgue exponent of the natural energy space \(W^{1,p}\). In the linear growth situation, this is in general not possible: As will be shown in [21], there exist linear growth integrands and generalised minimisers \(u \in BV \setminus W^{1,1}\) which do satisfy a Caccioppoli type inequality in one space dimension. This easily carries over to the BD-situation, and hereby rules out any integrability boost by virtue of Gehring. On the other hand, even for semiautonomous integrands \((x, z) \mapsto f(x, z)\), a well-known counterexample due to BILDHAUER [17, Thm. 4.39] asserts that if \(f \in C^2(\bar{\Omega}; R^{N \times n})\) satisfies a uniform variant of (1.4).
for $a > 3$, then generalised minima might in fact belong to $BV \setminus W^{1,1}$. In particular, the Caccioppoli inequality itself cannot yield higher integrability results in the linear growth setting.

On the other hand, the approach of Section 4 is robust enough can cope with integrands $(x, y, z) \mapsto f(x, y, z)$ if suitable superlinear growth in the last variable is imposed and thus the Gehring-type improvement is available. The reader will notice that this furthermore generalises to the case of convex integrands that have Orlicz growth in their gradient variable – provided the corresponding Orlicz function has $\Delta_2 \cap \nabla_2$-growth. Indeed, based on the results of [27] which can be adapted to the fully non-autonomous setting too, we can in fact derive higher integrability of minima. Furthermore, based on the results of [20], the required Korn-type inequalities are at our disposal. It is not so clear how the method of Section 4 can be adapted if the underlying Orlicz function does not have $\Delta_2 \cap \nabla_2$-growth. This case will be tackled in a future work. For the situation of primary interest, however, we make

**Remark 5.2** ($p$-growth functionals: Partial regularity). The reader will notice that our above reasoning to establish partial regularity for linear growth functionals generalises without further efforts to the superlinear growth regime. First, if $1 < p < \infty$, and $f : \Omega \times \mathbb{R}^n \times \mathbb{R}^{n \times n} \to \mathbb{R}$ and $g : \Omega \times \mathbb{R}^n \to \mathbb{R}$ are integrands that satisfies for constants $c_1, \ldots, c_6 > 0$ and $0 < \gamma < 1$,

$$
\begin{align*}
&c_1 |z|^p - c_2 \leq f(x, y, z) \leq c_3 (1 + |z|^p) \\
&|f(x_1, y_1, z) - f(x_2, y_2, z)| \leq c_4 (|x_1 - x_2| + |y_1 - y_2|)^\gamma (1 + |z|^p) \\
&|g(x_1, y_1) - g(x_2, y_2)| \leq c_5 |x_1 - x_2| + |y_1 - y_2| \gamma \\
&|f'(x, y, z)| \leq c_6 |z|^{p-1}
\end{align*}
$$

for all $x, x_1, x_2 \in \Omega$, $y, y_1, y_2 \in \mathbb{R}^n$ and $z \in \mathbb{R}^{n \times n}$ then a straightforward adaptation of Proposition 4.3 together with the modifications outlined in [9, Sec. 6] yield that minima of the functional (5.1) are $C^{1,\alpha}$-partially regular.

As to partial regularity, we have omitted quasiconvex or – in the setting of the main part – symmetric quasiconvex functionals throughout. In fact, at present it is not known how to modify the method exposed in Section 4 even in the full gradient case (also see the discussion in [9, 63]). The only result available in the BV-full gradient, quasiconvex case is due to KRISTENSEN and the author [41], and a generalisation to the BD-case will appear in [39]; also cf. the author’s thesis [38]. However, compared to the very degenerate setting of Theorem 1.2, the corresponding partial regularity result will require a strong quasiconvexity condition. If this condition pro forma is introduced for convex $C^{2,\alpha}$-integrands, then it translates to $3$-elliptic integrands in the sense of (1.4). In this sense, the results of the forthcoming work [39] and Theorem 1.2 are independent.

Let us now briefly comment on the Sobolev regularity. As we mentioned, the case of autonomous $a$-elliptic functionals of linear growth subject to (3.13) requires a bound as given in [18, Lem. 4.19(i)], but we believe that this can be accomplished similarly as is done for Theorem 3.2. This will possibly be done in a future work. The case of non-autonomous integrands $(x, z) \mapsto f(x, z)$, which satisfy the obvious modification of (1.4) uniformly in $x$, however, is more intricate. Even if $f$ is of class $C^2$ in the joint variable and satisfies the estimates corresponding to [18, Ass. 4.22], it is not fully clear to arrive at the decoupling estimates that eliminate the divergence as done in the proof of Theorem 3.2.

Whereas for partial regularity $C^{0,\alpha}$-Hölder continuous $x$-dependence still suffices, the corresponding Sobolev regularity theory is far from clear when aiming at an ellipticity regime beyond $1 < a < 1 + \frac{1}{p}$. Namely, in this case the Euler-Lagrange equations satisfied by (generalised) minima cannot be differentiated. In the full gradient, superlinear growth regime, this setting has been extensively studied by MINGIONE [54, 56, 57] and KRISTENSEN & MINGIONE [50, 51, 52]; the BV-case will be dealt with in [21]. Here, Nikol’skii estimates are employed but – as a matter of fact – do not use any information apart from the Euler-Lagrange equation itself and the continuity properties of $f$ with respect to its first variable. Such a strategy has been pursued in [40] for autonomous functionals (in the regime $1 < a < 1 + \frac{1}{p}$), and in light of [21] is expected to work in the symmetric gradient case as well.

However, if we wish to amplify the ellipticity regime as is done in Theorem 1.1, then we ought to use the instrumental identities for the minimisers that come out as a byproduct of second order estimates, cf. Theorem 3.2. As the latter are obtained by differentiating the first variation-style Euler-Lagrange...
equation, it seems difficult to make the approach presented in Section 3 work in the non-differentiable case.

6. APPENDIX

6.1. Reshetnyak (Semi-)Continuity Theory. At various stages in the main body of the paper, we have used (semi-)continuity properties of convex multiple integrals. These can be conveniently traced back to the following theorem — originally due to RESHETNYAK [60] — which subsumes these properties in the quite elaborate form as given by BECK & SCHMIDT [19, Thm. 2.4]:

**Theorem 6.1** (Reshetnyak (Lower Semi-)Continuity Theorem). Let \( m \in \mathbb{N} \) and suppose that \( \Omega \) is an open and bounded subset of \( \mathbb{R}^n \). Let \( (\mu_k) \) be a sequence in \( \mathcal{M}(\Omega; \mathbb{R}^m) \) that converges in the weak*-sense to some \( \mu \in \mathcal{M}(\Omega; \mathbb{R}^m) \). Moreover, assume that all of \( \mu, \mu_1, \mu_2, \ldots \) take values in some closed convex cone \( K \subset \mathbb{R}^m \). Then the following holds:

(a) If \( f : K \to [0, \infty) \) is a lower semicontinuous function which is convex and 1-homogeneous, then there holds

\[
\int_{\Omega} f\left(\frac{d\mu}{|d\mu|}\right) d|\mu| \leq \liminf_{k \to \infty} \int_{\Omega} f\left(\frac{d\mu_k}{|d\mu_k|}\right) d|\mu_k|.
\]

(b) If \( f : K \to [0, \infty) \) is a continuous function which is 1-homogeneous and if \( (\mu_k) \) converges strictly to \( \mu \), then there holds

\[
\int_{\Omega} f\left(\frac{d\mu}{|d\mu|}\right) d|\mu| = \lim_{k \to \infty} \int_{\Omega} f\left(\frac{d\mu_k}{|d\mu_k|}\right) d|\mu_k|.
\]

As an important special case, let us remark that this theorem entails continuity of convex variational integrals

\[
F[u; \Omega] := \int_{\Omega} f(\mathcal{E}u) \, dx + \int_{\Omega} F^\infty(\frac{d\mathcal{E}^s u}{|d\mathcal{E}^s u|}) \, d|\mathcal{E}^s u|, \quad u \in \text{BD}(\Omega)
\]

with respect to area-strict convergence; again, we implicitly assume that \( \Omega \subset \mathbb{R}^n \) is open, bounded and Lipschitz and that \( f : \mathbb{R}^{n\times n}_{\text{sym}} \to \mathbb{R} \) is continuous, convex and satisfies (LG'). In fact, here we apply the continuity part (b) of the preceding theorem to \( K := [0, \infty) \times \mathbb{R}^{n\times n}_{\text{sym}} \) and \( f : K \to [0, \infty) \) given by

\[
f(t, \xi) := \begin{cases} tf\left(\frac{\xi}{t}\right), & t > 0, \ \xi \in \mathbb{R}^{n\times n}_{\text{sym}}, \\ F^\infty(\xi), & t = 0, \ \xi \in \mathbb{R}^{n\times n}_{\text{sym}}. \end{cases}
\]

By continuity and linear growth of \( f \), \( f \) is in fact finite, continuous and 1-homogeneous on \( K \). Then, if \( u \in \text{BD}(\Omega) \), we put \( \mu := (\mathcal{L}^n, \mathcal{E}u) \); this is a measure on \( \Omega \) with values in \( K \). Directly inserting the definitions, we obtain

\[
F[u; \Omega] = \int_{\Omega} f\left(\frac{d\mu}{|d\mu|}\right) d|\mu|.
\]

Consequently, we find that if \( (u_k) \subset \text{BD}(\Omega) \) is a sequence with \( u_k \to u \) area-strictly in \( \text{BD}(\Omega) \), then \( (\mathcal{L}^n, \mathcal{E}u_k))(\Omega) \to (\mathcal{L}^n, \mathcal{E}u)(\Omega) \). Thus part (b) of Theorem 6.1 immediately yields \( F[u_k; \Omega] \to F[u; \Omega] \) as \( k \to \infty \).

6.2. Linear Comparison Estimates. Let \( \Omega \) be an open and bounded Lipschitz subset of \( \mathbb{R}^n \) and consider for \( w \in W^{1,2}(\Omega; \mathbb{R}^n) \) the variational principle

\[
(6.1) \quad \text{to minimise } \mathcal{F}[v] := \int_{B} g(\varepsilon(v)) \, dx \quad \text{over } v \in w + W^{1,2}_0(\Omega; \mathbb{R}^n),
\]

where \( g(\xi) := \langle A\xi, \xi \rangle + \langle b, \xi \rangle + c \) is a polynomial of degree two on \( \mathbb{R}^{n\times n}_{\text{sym}} \) with arbitrary but fixed \( b \in \mathbb{R}^{n\times n} \) and \( c \in \mathbb{R} \), whereas we assume that \( A : \mathbb{R}^{n\times n}_{\text{sym}} \times \mathbb{R}^{n\times n}_{\text{sym}} \to \mathbb{R} \) is bounded and, moreover, elliptic in the sense that there exists \( \ell > 0 \) such that \( \langle A\xi, \xi \rangle \geq \ell |\xi|^2 \) holds for all \( \xi \in \mathbb{R}^{n\times n}_{\text{sym}} \). As a small variation of [34, Lem. 3.0.5], we obtain:

**Lemma 6.2.** There exists a unique solution \( u \in w + W^{1,2}_0(\Omega; \mathbb{R}^n) \) of (6.1). Moreover, this solution satisfies the following:
(a) There exists a constant \( c = c(n, \ell) > 0 \) such that if \( B(z, R) \subseteq \Omega \), then for all \( 0 < r < R/2 \) there holds
\[
\int_{B(z,r)} |u(x) - (\varepsilon(u))_{z,r}|^2 \, dx \leq c \left( \frac{r}{R} \right)^{n+2} \int_{B(z,R/2)} |u(x) - (\varepsilon(u))_{z,R/2}|^2 \, dx.
\]

(b) There exists a constant \( c = c(n, \ell) > 0 \) such that if \( B(z, R) \subseteq \Omega \) and \( w \in C^{1,\alpha}(\overline{\Omega};\mathbb{R}^n) \), then
\[
\|\varepsilon(u)\|_{C^{0,\alpha}(\overline{B(z,R/2)};\mathbb{R}^{n \times n})} \leq c\|\varepsilon(w)\|_{C^{0,\alpha}(\overline{B(z,R/2)};\mathbb{R}^{n \times n})}.
\]

The proof follows as a direct merger of the arguments exposed in [34, Lem. 3.0.5] and classical Schauder theory (cf. [35, 11]) and is left to the interested reader.

6.3. Some Remarks on Young functions and Orlicz spaces. For the reader’s convenience, we record here some elementary facts and notational conventions regarding Young functions and Orlicz spaces and refer to [20, 22] for more detailed information. We say that \( A: [0, \infty) \to [0, \infty) \) is a Young function provided \( A(t) = \int_0^t a(s) \, ds \) for all \( t \geq 0 \), where \( a: [0, \infty) \to [0, \infty) \) is non-decreasing, left-continuous and \( a \not\equiv 0 \), \( a \not\equiv \infty \). For an open set \( \Omega \subset \mathbb{R}^n \), the Orlicz space \( L^A(\Omega; \mathbb{R}^m) \) consists of all measurable maps \( u: \Omega \to \mathbb{R}^m \) such that the Luxembourg norm
\[
\|u\|_{L^A(\Omega; \mathbb{R}^m)} := \inf \left\{ \lambda > 0 : \int_{\Omega} A\left( \frac{|u(x)|}{\lambda} \right) \, dx \leq 1 \right\}
\]
is finite. As tacitly utilised in Section 2, we put
\[
E^1 L^A(\Omega; \mathbb{R}^m) := \{ u \in L^A(\Omega; \mathbb{R}^m) : \varepsilon(u) \in L^A(\Omega; \mathbb{R}^{n \times n}) \},
\]
where as usual \( \varepsilon \) denotes the distributional symmetric gradient and
\[
E_0 L^A(\Omega; \mathbb{R}^m) := \{ u : \Omega \to \mathbb{R}^m : \text{ the trivial extension of } u \text{ to } \mathbb{R}^n \text{ belongs to } E^1 L^A(\mathbb{R}^m; \mathbb{R}^n) \}.
\]
Both spaces are endowed with the corresponding canonical norm \( \|u\|_{E^1 L^A(\Omega; \mathbb{R}^m)} := \|u\|_{L^A(\Omega; \mathbb{R}^m)} + \|\varepsilon(u)\|_{L^A(\Omega; \mathbb{R}^{n \times n})} \). Note that, when \( A(t) = |t| \log^{1+\alpha}(1 + |t|) \), we also write \( L \log^{1+\alpha} L := L^A \).

REFERENCES

[1] Acerbi, E.; Fusco, N.: A regularity theorem for minimizers of quasiconvex integrals. Arch. Rational Mech. Anal. 99 (1987), no. 3, 261–281.
[2] Acerbi, E.; Fusco, N.: Regularity for minimizers of nonquadratic functionals: the case \( 1 < p < 2 \). J. Math. Anal. Appl. 140 (1989), no. 1, 115–135.
[3] Allard, W.K.: On the first variation of a varifold. Ann. of Math. (2) 95 (1972), 417–491.
[4] Ambrosio, L.; Fusco, N.: Functions of bounded variation and free discontinuity problems. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000. xviii+434 pp.
[5] Ambrosio, L.; Coscia, A.; Dal Maso, G.: Fine properties of functions with bounded deformation. (English summary) Arch. Rational Mech. Anal. 139 (1997), no. 3, 201–238.
[6] Ambrosio, L.; Fusco, N.; Pallara, D.: Functions of bounded variation and free discontinuity problems. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000. xviii+434 pp.
[18] Bildhauer, M.: Convex variational problems. Linear, nearly linear and anisotropic growth conditions. Lecture Notes in Mathematics, 1818. Springer-Verlag, Berlin, 2003. x+217 pp.

[19] Beck, L.; Schmidt, T.: On the Dirichlet problem for variational integrals in BV. J. Reine Angew. Math. 674 (2013), 113-194.

[20] Breit, D.; Diening, L.: Sharp conditions for Korn inequalities in Orlicz spaces. J. Math. Fluid Mech. 14 (2012), no. 3, 565–573.

[21] Bulíček, M.; Gmeineder, F.: On Non-Autonomous Functionals of Linear Growth. In preparation.

[22] Cianchi, A.: Korn type inequalities in Orlicz spaces. Journal of Functional Analysis 267 (2014), 2313–2352.

[23] Conti, S., Faraco, D., Maggi, F.: A new approach to counterexamples to L1-estimates: Korn’s inequality, geometric rigidity, and regularity for gradients of separately convex functions. Arch. Rat. Mech. Anal. 175, 287–300 (2005).

[24] Matthies, H.; Strang, G.; Christiansen, E.: The saddle point of a differential program. Energy methods in finite element analysis, pp. 309-318. Wiley, Chichester, 1979.

[25] De Giorgi, E.: Frontiere orientate di misura minima. (Italian) Seminario di Matematica della Scuola Normale Superiore di Pisa, 1960-61 Editrice Tecnico Scientifica, Pisa 1961 57 pp.

[26] Demengel, F.; Temam, R.: Convex functions of a measure and applications. Indiana Univ. Math. J. 33 (1984), no. 5, 673-709.

[27] Diening, L.; Lengeler, D.; Stroffolini, B.; Verde, A.: Partial regularity for minimizers of quasi-convex functionals with general growth. SIAM J. Math. Anal. 44 (2012), no. 5, 3594–3616.

[28] Duzaar, F.; Grotowski, J.F.: Optimal interior partial regularity for nonlinear elliptic systems: the method of A-harmonic approximation. Manuscripta Math. 103 (2000), no. 3, 267–298.

[29] Duzaar, F.; Steffen, K.: Optimal interior and boundary regularity for almost minimizers to elliptic variational integrals. J. Reine Angew. Math. 546 (2002), 73–138.

[30] Duzaar, F.; Grotowski, J.F.; Kronz, M.: Regularity of almost minimizers of quasi-convex variational integrals with sub-quadratic growth. Ann. Mat. Pura Appl. (4) 184 (2005), no. 4, 421–448.

[31] Evans, L.C.: Quasiconvexity and partial regularity in the calculus of variations. Arch. Rational Mech. Anal. 95 (1986), no. 3, 227–252.

[32] Evans, L.C.: Partial differential equations. Second edition. Graduate Studies in Mathematics, 19. American Mathematical Society, Providence, RI, 2010.

[33] Fuchs, M.; Seregin, G.: A regularity theory for variational integrals with LlnL-growth. Calc. Var. Partial Differential Equations 6 (1998), no. 2, 171–187.

[34] Fuchs, M.; Seregin, G.: Variational methods for problems from plasticity theory and for generalized Newtonian fluids. Lecture Notes in Mathematics, 1749. Springer-Verlag, Berlin, 2000. vi+269 pp.

[35] Giaquinta, M.: Multiple integrals in the calculus of variations and nonlinear elliptic systems. Annals of Mathematics Studies, 105. Princeton University Press, Princeton, NJ, 1983.

[36] Gmeineder, F.: Symmetric-Convex Functionals of Linear Growth. Journal of Elliptic and Parabolic Problems, 2016.

[37] Gmeineder, F.: On the singular set of generalised minima in BV. OxPDE Technical Report.

[38] Gmeineder, F.: Regularity Theory for Variational Problems on BD. DPhil Thesis, University of Oxford, 2017.

[39] Gmeineder, F.: Regularity Theory for Semiconvex Functionals on BV. In Preparation.

[40] Gmeineder, F.; Kristensen, J.: Partial regularity for BV minimizers. Preprint.

[41] Gmeineder, F.; Modica, G.; Soucek, J.: Functionals with linear growth in the calculus of variations. I. II. Comment. Math. Univ. Carolin. 20 (1979), no. 1, 143-156, 157-172.

[42] Giusti, E.: Direct methods in the calculus of variations. World Scientific Publishing Co., Inc., River Edge, NJ, 2003.

[43] Goffman, C.; Serrin, J.: Sublinear functions of measures and variational integrals, Duke Math. J. 31 (1964), 159–178.

[44] Hajlasz, P.: On approximate differentiability of functions with bounded deformation. Manuscripta Math. 91 (1996), no. 1, 61–72.

[45] Kazaniecki, K.; Stolyarov, D.M.; Wojciechowski, M.: Anisotropic Ornstein non inequalities, Anal. PDE 10 (2017) 351–366.

[46] Kirchheim, B.; Kristensen, J.: On rank one convex functions that are homogeneous of degree one. Arch. Ration. Mech. Anal. 221 (2016), no. 1, 527–558.

[47] Kohn, R.V.: New Estimates for Deformations in Terms of their Strains. Ph.D. Thesis, Princeton Univ., 1979.

[48] Kohn, R.V.: New integral estimates for deformations in terms of their nonlinear strains. Arch. Rational Mech. Anal. 78 (1982), no. 2, 131–172.

[49] Kristensen, J.; Mingione, G.: The singular set of minima of integral functionals. Arch. Ration. Mech. Anal. 180 (2006), no. 3, 331–398.

[50] Kristensen, J.; Mingione, G.: The singular set of ω-minima. Arch. Ration. Mech. Anal. 177 (2005), no. 1, 93–114.

[51] Kristensen, J.; Mingione, G.: Non-differentiable functionals and singular sets of minima. C. R. Math. Acad. Sci. Paris 340 (2005), no. 1, 93–98.

[52] Ladyzhenskaya, O.A.: Ural’ceva, N.N.: Local estimates for gradients of solutions of non-uniformly elliptic and parabolic equations. Comm. on Pure and Appl. Math. 83(1970), 677-703.

[53] Mingione, G.: The singular set of solutions to non-differentiable elliptic systems. Arch. Ration. Mech. Anal. 166 (2003), no. 4, 287-301.

[54] Mingione, G.: Regularity of minima: an invitation to the dark side of the calculus of variations. Appl. Math. 51 (2006), no. 4, 355–426.
[56] Mingione, G.: Singularities of minima: a walk on the wild side of the calculus of variations. (English summary) J. Global Optim. 40 (2008), no. 1-3, 209-223.

[57] Mingione, G.: Bounds for the singular set of solutions to non linear elliptic systems. Calc. Var. Partial Differential Equations 18 (2003), no. 4, 373–400.

[58] Morrey Jr., C.B.: Multiple integrals in the calculus of variations.

[59] Ornstein, D.: A non-equality for differential operators in the $L_1$-norm. Arch. Rational Mech. Anal. 11 1962 40–49.

[60] Reshetnyak, Yu. G.: Weak convergence of completely additive vector functions on a set. Siberian Mathematical Journal (1968), Volume 9, Issue 6, 1039-1045.

[61] Reshetnyak, Yu. G.: Estimates for certain differential operators with finite-dimensional kernel. (Russian) Sibirsk. Mat. Z. 11 1970 414–428.

[62] Schmidt, T.: A simple partial regularity proof for minimizers of variational integrals. NoDEA Nonlinear Differential Equations Appl. 16 (2009), no. 1, 109–129.

[63] Schmidt, T.: Partial regularity for degenerate variational problems and image restoration models in BV. Indiana Univ. Math. J. 63, 213–279 (2014).

[64] Schmidt, T.: BV Minimizers of Variational Integrals: Existence, Uniqueness, Regularity, Habilitationsschrift (postdoctoral thesis), Friedrich-Alexander-Universität Erlangen-Nürnberg, 86 pp. (2015).

[65] Seregin, G.: Variation-difference schemes for problems in the mechanics of ideally elastoplastic media, U.S.S.R. Comput. Math. Phys. 25 (1985), 153–165; translation from Zh. Vychisl. Mat. Mat. Fiz. 25 (1985), 237–253.

[66] Seregin, G.: Differential properties of solutions of variational problems for functionals of linear growth, J. Sov. Math. 64 (1993), 1256–1277; translation from Probl. Mat. Anal. 11 (1990), 51–79.

[67] Seregin, G.: Two-dimensional variational problems of the theory of plasticity, Izv. Math. 60 (1996), 179–216; translation from Izv. Ross. Akad. Nauk, Ser. Mat. 60 (1996), 175–210.

[68] Seregin, G.: Differentiability properties of weak solutions of certain variational problems in the theory of perfect elastoplastic plates. Appl. Math. Optim. 28 (1993), 307–335.

[69] Serrin, J.: The problem of Dirichlet for quasilinear elliptic differential equations with many independent variables. Phil. Trans. of the Royal Soc. of London. Series A, n. 1153 vol. 264(1969), 413-496.

[70] Smith, K. T.: Formulas to represent functions by their derivatives. Math. Ann. 188 1970 53–77.

[71] Strang, G.; Temam, R.: Functions of bounded deformation. Arch. Rational Mech. Anal. 75 (1980/81), no. 1, 7–21.

[72] Strauss, M.I.: Variations of Korn’s and Sobolev’s equalities. Partial differential equations (Proc. Sympos. Pure Math., Vol. XXIII, Univ. California, Berkeley, Calif., 1971), pp. 207–214. Amer. Math. Soc., Providence, R.I., 1973.

[73] Suquet, P.-M.: Sur un nouveau cadre fonctionnel pour les Équations de la plasticité. (French) C. R. Acad. Sci. Paris Sér. A-B 286 (1978), no. 23, A1129–A1132.

**Author’s Address:** Mathematical Institute, University of Bonn, Endenicher Allee 60, 53116 Bonn, Germany