Calc. Var. (2016) 55:154
DOI 10.1007/s00526-016-1085-5

Calculus of Variations

Regularity for free interface variational problems in a general class of gradients

Adolfo Arroyo-Rabasa[1]

Received: 28 March 2016 / Accepted: 5 October 2016 / Published online: 30 November 2016

Abstract We present a way to study a wide class of optimal design problems with a perimeter penalization. More precisely, we address existence and regularity properties of saddle points of energies of the form

$$(u, A) \quad \mapsto \quad \int_\Omega 2fu\,\mathrm{d}x \;-\int_{\Omega\cap A} \sigma_1\mathscr{A}u\cdot\mathscr{A}u\;\mathrm{d}x \;-\int_{\Omega\setminus A} \sigma_2\mathscr{A}u\cdot\mathscr{A}u\;\mathrm{d}x \;+\;\mathrm{Per}\,(A;\overline{\Omega}),$$

where Ω is a bounded Lipschitz domain, $A\subset\mathbb{R}^N$ is a Borel set, $u:\Omega\subset\mathbb{R}^N\to\mathbb{R}^d$, $\mathscr{A}$ is an operator of gradient form, and σ_1,σ_2 are two not necessarily well-ordered symmetric tensors. The class of operators of gradient form includes scalar- and vector-valued gradients, symmetrized gradients, and higher order gradients. Therefore, our results may be applied to a wide range of problems in elasticity, conductivity or plasticity models. In this context and under mild assumptions on f, we show for a solution (w, A), that the topological boundary of $A\cap\Omega$ is locally a C^1-hypersurface up to a closed set of zero $\mathscr{H}^{N-1}$-measure.

Mathematics Subject Classification Primary 49J20 · 49J35 · 49N60 · 49Q20; Secondary 49J45

1 Introduction

The problem of finding optimal designs involving two materials goes back to the work of Hashin and Shtrikman. In [1], the authors made the first successful attempt to derive the optimal bounds of a composite material. It was later on, in the series of papers [2–4], that Kohn and Strang described the connection between composite materials, the method of relaxation, and the homogenization theory developed by Murat and Tartar [5,6]. In the context

Communicated by L. Ambrosio.

✉ Adolfo Arroyo-Rabasa
adolfo.arroyo.rabasa@hcm.uni-bonn.de

1 Institut für Angewandte Mathematik, Endenicher Allee 60, 53115 Bonn, Germany

of homogenization, *better* designs tend to develop finer and finer geometries; a process which results in the creation of non-classical designs. One way to avoid the mathematical abstract of infinitely fine mixtures is to add a cost on the interfacial energy. In this regard, there is a large amount of optimal design problems that involve an interfacial energy and a Dirichlet energy. The study of regularity properties in this setting has been mostly devoted to problems where the Dirichlet energy is related to a scalar elliptic equation; see [7–12], where partial C^1-regularity on the interface is shown for an optimization problem oriented to find dielectric materials of maximal conductivity. We shall study regularity properties of similar problems in a rather general framework. Our results extend the aforementioned results to linear elasticity and linear plate theory models.

Before turning to a precise mathematical statement of the problem let us first present the model in linear plate theory that motivated our results. Let $\Omega = \omega \times [-h, h]$ be the reference configuration of a plate of thickness $2h$ and cross section $\omega \subset \mathbb{R}^2$. The linear equations governing a clamped plate Ω as h tends to zero for the Kirchhoff model are

$$\begin{cases} \nabla \cdot (\nabla \cdot (\sigma \nabla^2 u)) = f & \text{in } \omega, \\ \partial_\nu u = u = 0 & \text{in } \partial\omega, \end{cases} \tag{1}$$

where $u : \omega \to \mathbb{R}$ represents the displacement of the plate with respect to a vertical load $f \in L^\infty(\omega)$, and the design of the plate is described by a symmetric positive definite fourth-order tensor σ (up to a cubic dependence on the constant h). Here, we denote the second gradient by

$$\nabla^2 u := \left(\frac{\partial^2 u}{\partial x_i \partial x_j} \right)_{ij}, \quad i, j = 1, 2.$$

Consider the physical problem of a thin plate Ω made-up of two elastic materials. More precisely, for a given set $A \subset \omega \subset \mathbb{R}^2$ we define the symmetric positive tensor

$$\sigma_A(x) := \mathbb{1}_A \sigma_1 + (1 - \mathbb{1}_A)\sigma_2,$$

where $\sigma_1, \sigma_2 \in \operatorname{Sym}(\mathbb{R}^{2\times 2}, \mathbb{R}^{2\times 2})$. In this way, to each Borel subset $A \subset \omega$, there corresponds a displacement $u_A : \omega \to \mathbb{R}$ solving Eq. (1) with $\sigma = \sigma_A$. One measure of the rigidity of the plate is the so-called compliance, i.e., the work done by the loading. The smaller the compliance, the stiffer the plate is. A reasonable optimal design model consists in finding the most rigid design A under the aforementioned costs. One seeks to minimize an energy of the form

$$A \mapsto \int_\omega \sigma_A \nabla^2 u_A \cdot \nabla^2 u_A \, dx \; + \; \operatorname{Per}(A; \omega), \quad \text{among Borel subsets } A \text{ of } \mathbb{R}^2.$$

Optimality conditions for a stiffest plate can be derived by taking local variations on the design. For such analysis to be meaningful, one has to ensure first that the variational equations of optimality have a suitable meaning *in* the interface. Hence, it is natural to ask for the maximal possible regularity of ∂A and $\nabla^2 u_A$.

We will introduce a more general setting where one can replace the second gradient ∇^2 by an operator $\mathscr{A}$ of *gradient type* (see Definition 2.1 and the subsequent examples in the next section for a precise description of this class).

1.1 Statement of the problem

Let $N \geq 2$, and let d, k be positive integers. We shall work in $\Omega \subset \mathbb{R}^N$; a nonempty, open, and bounded Lipschitz domain. We also fix a function $f \in L^\infty(\Omega; \mathbb{R}^d)$ and let σ_1 and σ_2 be two positive definite tensors in $\mathrm{Sym}(\mathbb{R}^{dN^k} \otimes \mathbb{R}^{dN^k})$ satisfying a strong pointwise Gårding inequality: there exists a positive constant M such that

$$\frac{1}{M}|P|^2 \leq \sigma_i\, P \cdot P \leq M|P|^2 \quad \text{for all } P \in \mathbb{R}^{dN^k}, \quad i \in \{1, 2\}. \tag{2}$$

For a fixed Borel set $A \subset \mathbb{R}^N$, define the two-point valued tensor

$$\sigma_A(x) := \mathbb{1}_A \sigma_1 + \mathbb{1}_{(\mathbb{R}^N \setminus A)} \sigma_2. \tag{3}$$

We consider a k'th-order homogeneous linear differential operator $\mathscr{A} : \mathrm{L}^2\left(\Omega; \mathbb{R}^d\right) \to \mathrm{W}^{-k,2}(\Omega; \mathbb{R}^{dN^k})$ of gradient form (see Definition 2.1 in Sect. 2). As a consequence of the definition of operators of gradient form, the following equation

$$\mathscr{A}^*(\sigma_A \mathscr{A} u) = f \quad \text{in } \mathscr{D}'(\Omega; \mathbb{R}^d), \quad u \in \mathrm{W}_0^{\mathscr{A}}(\Omega) \subset \mathrm{W}_0^{k,2}(\Omega; \mathbb{R}^d), \tag{4}$$

has a unique solution (cf. Theorem 1.1). We will refer to Eq. (4) as the *state constraint* and we will denote by w_A its unique solution.

It is a physically relevant question to ask which designs have the least dissipated energy. To this end, consider the energy defined as[1]

$$A \;\mapsto\; E(A) := \int_\Omega f w_A \,\mathrm{d}x \;+\; \mathrm{Per}\,(A; \overline{\Omega}) \quad \text{among Borel subsets } A \text{ of } \mathbb{R}^N.$$

We will be interested in the optimal design problem with Dirichlet boundary conditions on sets:

$$\text{minimize} \quad \left\{ E(A) \;:\; A \subset \mathbb{R}^N \;\text{ is a Borel set, }\; A \cap \Omega^c \equiv A_0 \cap \Omega^c \right\}, \tag{5}$$

where $A_0 \subset \mathbb{R}^N$ is a set of locally finite perimeter.[2]

Most attention has been drawn to the case where designs are mixtures of two well-ordered materials. The presentation given here places no comparability hypotheses on σ_1 and σ_2. Instead, we introduce a weaker condition on the decay of generalized minimizers of a double-well problem. Our technique also holds under various constraints other than Dirichlet boundary conditions; in particular, any additional cost that scales as $O(r^{N-1+\varepsilon})$. For example, a constraint on the volume occupied by a particular material (cf. [8,12,13]). Lastly, we remark that our technique is robust enough to treat models involving the *maximization* of dissipated energy.

1.2 Main results and background of the problem

Existence of a minimizer of (5) can be established by standard methods. We are interested in proving that a solution pair (w_A, A) enjoys better *regularity* properties than the ones needed for existence. The notion of regularity for a set A will be understood as the local regularity of ∂A seen as a submanifold of $\mathbb{R}^N$, whereas the notion of regularity for w_A will refer to its differentiability and integrability properties.

It can be seen from the energy, that the deviation from being a perimeter minimizer for a solution A of problem (5) is bounded by the dissipated energy. Therefore, one may not expect better regularity properties for A than the ones for perimeter minimizers; and thus,

[1] Here, $\mathrm{Per}(A; \overline{\Omega}) = |\mu_A|(\overline{\Omega})$, where μ_A is the Gauss–Green measure of A; see Sect. 2.4.

[2] Due to the nature of the problem, we cannot replace $\mathrm{Per}(A; \overline{\Omega})$ with $\mathrm{Per}(A; \Omega)$ in $E(A)$ because it possible that minimizing sequences tend to accumulate perimeter in $\partial\Omega$.

one may only expect regularity up to singular set (we refer the reader to [14,15] for classic results, see also [12] for a partial regularity result in a similar setting to ours).

Since a constrained problem may be difficult to treat, we will instead consider an equivalent variational unconstrained problem by introducing a multiplier as follows. Consider the saddle point problem

$$\inf_{A\subset\Omega}\ \sup_{u\in \mathrm{W}_0^{\mathscr{A}}(\Omega)}\ I_\Omega(u,A), \tag{P}$$

where

$$I_\Omega(u,A) := \int_\Omega 2fu\,\mathrm{d}x \ -\ \int_\Omega \sigma_A \mathscr{A}u\cdot\mathscr{A}u\,\mathrm{d}x \ +\ \operatorname{Per}(A;\overline{\Omega}).$$

Our first result shows the equivalence between problem (P) and the minimization problem (5) under the state constraint (4):

Theorem 1.1 (Existence) *There exists a solution* (w, A) *of problem* (P)*. Furthermore, there is a one to one correspondence*

$$(w,A)\mapsto (w_A,A)$$

between solutions of the problem (P) *and solutions of the minimization problem* (5) *under the state constraint* (4).

We now turn to the question of regularity. Let us depict an outline of the key steps and results obtained in this regard. The Morrey space $\mathrm{L}^{p,\lambda}(\Omega;\mathbb{R}^d)$ is the subspace of $\mathrm{L}^p(\Omega;\mathbb{R}^d)$ for which the semi-norm

$$[u]^p_{\mathrm{L}^{p,\lambda}(\Omega)} := \sup\left\{\frac{1}{r^\lambda}\int_{B_r(x)}|u|^p\,\mathrm{d}y : B_r(x)\subset\Omega\right\},\quad 0<\lambda\le N,$$

is finite.

The *first step* in proving regularity for solutions (w, A) consists in proving a critical $\mathrm{L}^{2,\,N-1}$ local estimate for $\mathscr{A}w$. This estimate arises naturally since we expect a kind of balance between $\int_{B_r(x)}\sigma_A\mathscr{A}w\cdot\mathscr{A}w\,\mathrm{d}y$ and the perimeter part $\operatorname{Per}(A;B_r(x))$ that scales as r^{N-1} in balls of radius r.

To do so, let us recall a related relaxed problem. As part of the assumptions on $\mathscr{A}$ there must exist an m'th-order diferential operator $\mathscr{B}:\mathrm{L}^2(\Omega;Z)\to \mathrm{W}^{-m,2}(\Omega;\mathbb{R}^n)$ with constant rank and $\operatorname{Ker}(\mathscr{B}) = \mathscr{A}[\mathrm{W}^{\mathscr{A}}(\Omega)]$.[3] It has been shown by Fonseca and Müller [16], that a necessary and sufficient condition for the lower semi-continuity of integral energies with superlinear growth under a constant rank differential constraint $\mathscr{B}v = 0$ is the $\mathscr{B}$-quasiconvexity of the integrand. In this context, the $\mathscr{B}$-free quasiconvex envelope of the double-well $W(P) := \min\{\sigma_1 P\cdot P,\ \sigma_2 P\cdot P\}$, at a point $P\in Z\subset\mathbb{R}^{dN^k}$, is given by

$$\begin{aligned} Q_{\mathscr{B}}W(P) := \inf\Bigg\{&\int_{[0,1]^N} W(P+v(y))\,\mathrm{d}y :\\ &v\in \mathrm{C}^\infty_{\mathrm{per}}\big([0,1]^N;Z\big),\ \mathscr{B}v=0 \quad\text{and}\quad \int_{[0,1]^N} v(y)\,\mathrm{d}y = 0\Bigg\}.\end{aligned}$$

The idea is to get an $\mathrm{L}^{2,\,N-1}$ estimate by transferring the regularizing effects from generalized minimizers of the energy $u\mapsto\int_{B_1}W(\mathscr{A}u)$ onto our original problem. In order to achieve

[3] Here, $\mathrm{W}^{\mathscr{A}}(\Omega) = \{u\in\mathrm{L}^2(\Omega;\mathbb{R}^d) : \mathscr{A}u\in\mathrm{L}^2(\Omega;\mathbb{R}^{dN^k})\}$ is the $\mathscr{A}$-Sobolev space of Ω.

this, we use a Γ-convergence argument with respect to a perturbation in the interfacial energy from which the next result follows:

Theorem 1.2 (Upper bound) *Let* (w, A) *be a variational solution of problem* (P). *Assume that, for some* $\delta \in [0, 1)$ *and some positive constant c, the higher integrability condition*

$$[\mathscr{A}\tilde{u}]^2_{\mathrm{L}^{2,N-\delta}(B_{1/2})} \leq c\|\mathscr{A}\tilde{u}\|^2_{\mathrm{L}^2(B_1)}, \tag{Reg}$$

holds for local minimizers of the energy $u \mapsto \int_{B_1} Q_{\mathscr{B}}W(\mathscr{A}u)$, *where* $u \in \mathrm{W}^{\mathscr{A}}(\Omega)$. *Then, for every compactly contained set* $K \subset\subset \Omega$, *there exists a positive constant* Λ_K *such that*

$$\begin{aligned}&\int_{B_r(x)} \sigma_A \mathscr{A}w \cdot \mathscr{A}w \;\mathrm{d}y \;+\; \mathrm{Per}(A; B_r(x)) \\ &\quad\leq \Lambda_K r^{N-1} \quad \textit{for all } x \in K \;\; \textit{and every } r \in (0, \mathrm{dist}(K, \partial\Omega)).\end{aligned} \tag{6}$$

Remark 1.3 (Well-ordering assumption) If σ_1, σ_2 are well-ordered, say $\sigma_2 - \sigma_1$ is positive definite, then $Q_{\mathscr{B}}W$ is precisely the quadratic form $\sigma_2\, P \cdot P$. Due to standard elliptic regularity results (cf. Lemma 2.6), estimate (Reg) holds for $\delta = 0$; therefore, assuming that the materials are well-ordered is a sufficient condition for the higher integrability assumption (Reg) to hold.

Remark 1.4 (Non-comparable materials) In dimensions $N = 2, 3$ and restricted to the setting $\mathscr{A} = \nabla$, $d = 1$, condition (Reg) is strictly weaker than assuming the materials to be well-ordered. Indeed, one can argue by a Moser type iteration as in [17] to lift the regularity of minimizers. For higher-order gradients or in the case of systems it is not clear to us whether assumption (Reg) is equivalent to the well-ordering of the materials.

The *second step*, consists of proving a *discrete monotonicity* for the excess of the Dirichlet energy on balls under a low perimeter density assumption. More precisely, on the function that assigns

$$r \quad\mapsto\quad \frac{1}{r^{N-1}} \int_{B_r(x)} |\mathscr{A}w|^2 \,\mathrm{d}x, \quad x \in \partial A,\; r > 0.$$

The discrete monotonicity of the map above, together with the upper bound estimate (6), will allow us to prove a local lower bound λ_K on the density of the perimeter:

$$\frac{\mathrm{Per}(A; B_r(x))}{r^{N-1}} \geq \lambda_K \quad \text{for every } x \in (K \cap \partial A), \text{ and every } 0 < r \leq r_K. \tag{LB}$$

As usual, the lower bound on the density of the perimeter is the cornerstone to prove regularity of almost perimeter minimizers. In fact, once the estimate (LB) is proved we simply apply the excess improvement results of [8, Sections 4 and 5] to obtain our main result:

Theorem 1.5 (Partial regularity) *Let* (w, A) *be a saddle point of problem* (P) *in* Ω. *Assume that the operator* $P_H u = \mathscr{A}^*(\sigma_H \mathscr{A}u)$ *is hypoelliptic and regularizing for the half-space problem* (see properties (60), (61)), *and that the higher integrability* (Reg) *holds. Then there exists a positive constant* $\eta \in (0, 1]$ *depending only on* N *such that*

$$\mathscr{H}^{N-1}((\partial A \backslash \partial^* A) \cap \Omega) = 0, \quad \textit{and} \quad \partial^* A \quad \textit{is an open } \mathrm{C}^{1,\eta/2}\text{- hypersurface in } \Omega.$$

Moreover if $\mathscr{A}$ *is a first-order partial differential operator, then* $\mathscr{A}w \in \mathrm{C}^{0,\eta/8}_{\mathrm{loc}}(\Omega \backslash (\partial A \backslash \partial^* A))$; *and hence, the trace of* $\mathscr{A}w$ *exists on either side of* $\partial^* A$.

Let us make a quick account of previous results. To our knowledge, only optimal design problems modeling the maximal dissipation of energy have been treated.

In [7] Ambrosio and Buttazzo considered the case where $\mathscr{A} = \nabla$ is the gradient operator for scalar-valued ($d = 1$) functions and where $\sigma_2 \geq \sigma_1$ in the sense of quadratic forms. The authors proved existence of solutions and showed that, up to choosing a good representative, the topological boundary is the closure of the reduced boundary and $\mathscr{H}^{N-1}(\partial A \backslash \partial^* A) = 0$. Soon after, Lin [8], and Kohn and Lin [9] proved, in the same case, that $\partial^* A$ is an open C^1-hypersurface. From this point on, there have been several contributions aiming to discuss the optimal regularity of the interface for this particular case. In this regard and in dimension $N = 2$, Larsen [10] proved that connected components of A are C^1 away from the boundary. In arbitrary dimensions, Larsen's argument cannot be further generalized because it relies on the fact that convexity and positive curvature are equivalent in dimension $N = 2$. During the time this project was developed, we have learned that Fusco and Julin [11] found a different proof for the same results as stated in [8]; besides this, De Philippis and Figalli [12] recently obtained an improvement on the dimension of the singular set $(\partial^* A \backslash \partial A)$.

The paper is organized as follows. In the beginning of Sect. 2 we fix notation and discuss some facts of linear operators, Young measures and sets of finite perimeter. We also give the precise definition of gradient type operators and include a compensated compactness result that will be employed throughout the paper. In Sect. 3 we show the equivalence of the constrained problem (4), (5) and the unconstrained problem (P) (Theorem 1.1). In the first part of Sect. 4 we shortly discuss how the higher integrability assumption (Reg) holds for various operators of gradient form. The rest of the section is devoted to the proof of the Upper bound (6). Section 5 is devoted to the proof of the Lower bound estimate (LB). Finally, in Sect. 6 we recall the flatness excess improvement [8] from which Theorem 1.5 easily follows.

2 Notation and preliminaries

We will write Ω to represent a non-empty, open, bounded subset of $\mathbb{R}^N$ with Lipschitz boundary $\partial\Omega$. The use of capital letters $A, B, \ldots,$ will be reserved to denote Borel subsets of $\mathbb{R}^N$ and we will write $\mathfrak{B}(\mathbb{R}^N)$ to denote the Borel σ-algebra of $\mathbb{R}^N$.

The letters x, y will denote points in Ω; while $z \in \mathbb{R}^d$ and $P \in \mathbb{R}^{dN^k}$ will be reserved for vectors and arrays in Euclidean space. The Greek letters $\varepsilon, \delta, \rho$ and γ shall be used for general smallness or scaling constants. We follow Lin's convention in [8], bounding constants will be generally denoted by $c_1 \geq c_2 \geq \cdots$, while smallness and decay constants will be usually denoted by $\varepsilon_1 \geq \varepsilon_2 \geq \cdots$, and $\theta_1 \geq \theta_2 \geq \cdots$, respectively. Let us mention that in proving regularity results one may often find it impractical to keep track of numerical constants due to the large amount of parameters; to illustrate better their uses and dependencies we have included a glossary of constants at the end of the paper.

It will often be useful to write a point $x \in \mathbb{R}^N = \mathbb{R}^{N-1} \times \mathbb{R}$ as $x = (x', x_N)$, in the same fashion we will also write $\nabla = (\nabla', \partial_N)$ to decompose the gradient operator. The bilinear form $\mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R} : (x, y) \mapsto x \cdot y$ will stand for the standard inner product between two points while we will use the notation $|x| := \sqrt{x \cdot x}$ to represent the standard p-dimensional Euclidean norm. To denote open balls in $\mathbb{R}^N$ centered at a point x with radius r we will simply write $B_r(x)$. Similarly, $B_r' := \{x' \in \mathbb{R}^{N-1} : (x', 0) \in B_r\}$.

We keep the standard notation for L^p and $\mathrm{W}^{l,p}$ spaces. We write $\mathrm{C}^l(\Omega; Z)$, and $\mathrm{C}^l_\mathrm{c}(\Omega; \mathbb{R}^d)$ to denote the spaces of functions with values in $\mathbb{R}^d$ and with continuous l-th derivative, and its subspace of functions compact support respectively. Similar notation stands for $\mathscr{M}(\Omega; \mathbb{R}^d)$

the space of bounded Radon measures in Ω, and $\mathscr{D}(\Omega;\mathbb{R}^d)$ the space of smooth functions in Ω with compact support. For X and Y Banach spaces, the standard pairing between X and Y will be denoted by $\langle\cdot,\cdot\rangle : X\times Y\to\mathbb{R} : (u,v)\mapsto\langle u,v\rangle$.

2.1 Operators of gradient form

We introduce an abstract class of linear differential operators $\mathscr{A} : \mathrm{L}^2(\Omega;\mathbb{R}^d)\to \mathrm{W}^{-k,2}(\Omega;\mathbb{R}^{dN^k})$.

This class contains scalar- and vector-valued gradients, higher gradients, and symmetrized gradients among its elements. The motivation behind it is that we may treat different models by employing a general and neat abstract setting. At a first glance this framework may appear too sterile, however, this definition is only meant to capture some of the essential regularity and rigidity properties of gradients.

Let $\mathscr{A} : \mathrm{L}^2(\Omega;\mathbb{R}^p)\to \mathrm{W}^{-k,2}(\Omega;\mathbb{R}^q)$ be a k'-th order homogeneous partial differential operator of the form

$$\mathscr{A}=\sum_{|\alpha|=k} A_\alpha\partial^\alpha, \tag{7}$$

where $A_\alpha\in\mathrm{Lin}\,(\mathbb{R}^p;\mathbb{R}^q)$, and $\partial^\alpha=\partial_1^{\alpha_1}\dots\partial_N^{\alpha_N}$ for every multi-index $\alpha=(\alpha_1,\dots,\alpha_N)\in(\mathbb{N}\cup\{0\})^N$ with $|\alpha|:=|\alpha_1|+\cdots|\alpha_N|$. We define the $\mathscr{A}$-Sobolev space of Ω as

$$\mathrm{W}^{\mathscr{A}}(\Omega):=\left\{u\in\mathrm{L}^2(\Omega;\mathbb{R}^p)\ :\ \mathscr{A}u\in\mathrm{L}^2(\Omega;\mathbb{R}^q)\right\}$$

endowed with the norm $\|u\|^2_{\mathrm{W}^{\mathscr{A}}(\Omega)}:=\|u\|^2_{\mathrm{L}^2(\Omega)}+\|\mathscr{A}u\|^2_{\mathrm{L}^2(\Omega)}$. We also define the $\mathscr{A}$-Sobolev space with zero boundary values in $\partial\Omega$ by letting

$$\mathrm{W}_0^{\mathscr{A}}(\Omega):=\mathrm{cl}\left\{\mathrm{C}_\mathrm{c}^\infty(\Omega;\mathbb{R}^p),\|\cdot\|_{\mathrm{W}^{\mathscr{A}}(\Omega)}\right\}.$$

The principal symbol of $\mathscr{A}$ is the positively k-homogeneous map defined as

$$\xi\mapsto\mathbb{A}(\xi):=\sum_{|\alpha|=k}\xi^\alpha A_\alpha\in\mathrm{Lin}\,(\mathbb{R}^p,\mathbb{R}^q),\quad \xi\ \in\mathbb{R}^N,$$

where $\xi^\alpha=\xi_1^{\alpha_1}\cdots\xi_N^{\alpha_N}$. One says that $\mathscr{A}$ has the constant rank property if there exists a positive integer r such that

$$\mathrm{rank}\,(\mathbb{A}(\xi))=r\quad\text{for all }\xi\in\mathbb{R}^N\setminus\{0\}. \tag{†}$$

Definition 2.1 (Operators of gradient form) Let $\mathscr{A}$ a homogeneous partial differential operator as in (7) with $p=d$ and $q=dN^k$. We say that $\mathscr{A}$ is an operator of gradient form if the following properties hold:

1. *Compactness.* There exists a positive constant $C(\Omega)$ for which

$$\|\varphi\|^2_{\mathrm{W}^{k,2}(\Omega)}\le C(\Omega)\left(\|\varphi\|^2_{\mathrm{L}^2(\Omega)}+\|\mathscr{A}\varphi\|^2_{\mathrm{L}^2(\Omega)}\right) \tag{8}$$

for all $\varphi\in\mathrm{C}^\infty(\overline{\Omega};\mathbb{R}^d)$. Even more, for every $u\in W^{\mathscr{A}}(\Omega)$ the following Poincaré inequality holds:

$$\inf\left\{\ \|u-v\|^2_{\mathrm{W}^{k,2}(\Omega)}\ :\ v\in\mathrm{W}^{\mathscr{A}}(\Omega),\mathscr{A}v=0\ \right\}\le C(\Omega)\|\mathscr{A}u\|^2_{\mathrm{L}^2(\Omega)}. \tag{9}$$

2. *Exactness.* There exists an m'-th homogeneous partial differential operator

$$\mathscr{B} := \sum_{|\alpha|=m} B_\alpha \partial^\alpha, \tag{10}$$

with coefficients $B_\alpha \in \mathrm{Lin}\,(Z; \mathbb{R}^n)$ for some positive integer n and a subspace Z of $\mathbb{R}^{dN^k}$, such that for every open and simply connected subset $\omega \subset \Omega$ we have the property

$$\left\{\mathscr{A}u : u \in \mathrm{W}^{\mathscr{A}}(\omega)\right\} = \left\{v \in \mathrm{L}^2(\omega; Z) : \mathscr{B}v = 0 \text{ in } \mathscr{D}'(\omega; \mathbb{R}^n)\right\}.$$

We write $\mathscr{A}^*$ to denote the L^2-adjoint of $\mathscr{A}$, which is given by

$$\mathscr{A}^* := (-1)^k \sum_{|\alpha|=k} A_\alpha^T \partial^\alpha.$$

Remark 2.2 (Constant rank) Let $\mathscr{A}$ and $\mathscr{B}$ be two linear differential operators satisfying an exactness property as in Definition 2.1. Then both operators $\mathscr{A}$ and $\mathscr{B}$ have the constant rank property (†). This follows from the lower semi-continuity of the rank in any subspace of matrices.

Remark 2.3 (Rigidity) The wave cone of an operator $\mathscr{A}$ of the form (7) which is defined as

$$\Lambda_{\mathscr{A}} := \bigcup_{|\xi|=1} \mathrm{Ker}(\mathbb{A}(\xi)) \subset \mathbb{R}^p,$$

contains the admissible amplitudes in Fourier space for which concentration and oscillation behavior is allowed under the constraint $\mathscr{A}u = 0$. As in the case of gradients, it can be seen from the compactness assumption in Definition 2.1 that the wave cone $\Lambda_{\mathscr{A}}$ of a gradient operator $\mathscr{A}$ is the zero space. In particular, there exists a positive constant λ (depending only on the coefficients of $\mathscr{A}$) such that

$$|\mathbb{A}(\xi)z|^2 \geq \lambda|\xi|^{2k}|z|^2 \quad \text{for all } \xi \in \mathbb{R}^N \setminus \{0\} \text{ and all } z \in \mathbb{R}^d. \tag{11}$$

Remark 2.4 (Poincaré inequality II) It follows from the definition of $\mathrm{W}_0^{\mathscr{A}}(\Omega)$ and the compactness assumption of $\mathscr{A}$ that $\mathrm{W}_0^{\mathscr{A}}(\Omega) \subset \mathrm{W}_0^{k,2}(\Omega; \mathbb{R}^d)$. In particular, $\mathrm{Ker}(\mathscr{A}) \cap \mathrm{W}_0^{\mathscr{A}}(\Omega) = \{0\} \subset \mathrm{L}^2(\Omega; \mathbb{R}^d)$ and $\mathscr{A}[\mathrm{W}_0^{\mathscr{A}}(\Omega)]$ is closed in the L^2 norm. Thus, by [18, Theorem 2.21], there exists a constant[4] $C(\Omega)$ such that

$$\|u\|_{\mathrm{L}^2(\Omega)}^2 \leq C(\Omega)\|\mathscr{A}u\|_{\mathrm{L}^2(\Omega)}^2 \quad \text{for all } u \in \mathrm{W}_0^{\mathscr{A}}(\Omega). \tag{12}$$

2.1.1 *Elliptic regularity*

Let $\mathscr{A}$ be an operator of gradient form as in Definition 2.1 and let $\sigma \in \mathrm{L}^\infty(\Omega; \mathbb{R}^{dN^k})$ be a tensor of variable coefficients satisfying the strong pointwise Gårding inequality (see (2))

$$\frac{1}{M}|P|^2 \leq \sigma(x)\, P \cdot P \leq M|P|^2 \quad \text{for almost every } x \in \Omega \text{ and every } \; P \in \mathbb{R}^{dN^k}. \tag{13}$$

If we define

$$\mathbf{A}_{\beta\alpha}^{ij} := (A_\alpha)_{i\beta,j} \quad \text{for } |\alpha| = |\beta| = k, \text{ and } 1 \leq i, j \leq d,$$

[4] Possibly abusing the notation, we will denote by $C(\Omega)$ the Poincaré constants from Definition 2.1 and Remark 2.4.

then we may write

$$\mathscr{A}\varphi = \mathbf{A}\nabla^k\varphi \quad \text{for every } \varphi \in \mathrm{C}^k(\overline{\Omega}; \mathbb{R}^d). \tag{14}$$

It is easy to verify, using the compactness assumption of $\mathscr{A}$, that $\mathbf{C} := (\mathbf{A}^T\sigma\ \mathbf{A})$ satisfies the weak Gårding inequality

$$\langle \mathbf{C}\ \nabla^k\varphi, \nabla^k\varphi\rangle \geq \left(\frac{1}{MC}\right)\|\nabla^k\varphi\|^2_{\mathrm{L}^2(\Omega)} - \left(\frac{1}{M}\right)\|\varphi\|^2_{\mathrm{L}^2(\Omega)}, \tag{15}$$

where $C = C(\Omega)$ is the constant in the compactness assumption of Definition 2.1; for all smooth, $\mathbb{R}^d$-valued functions φ in $\overline{\Omega}$.

Lemma 2.5 (Caccioppoli inequality) *Let $\sigma \in \mathrm{L}^\infty(\Omega; \mathbb{R}^{dN^k})$ satisfy the strong pointwise Gårding inequality (13) and let $w \in \mathrm{W}^{\mathscr{A}}(\Omega)$ be a solution of the state equation*

$$\mathscr{A}^*(\sigma\mathscr{A}u) = 0 \quad \textit{in } \mathscr{D}'(\Omega; \mathbb{R}^d).$$

Then there exists a positive constant C depending only on M, N, σ and $\mathscr{A}$ such that

$$\int_{B_r(x)} |\nabla^k w|^2\,\mathrm{d}x \ \leq\ \frac{C}{(R-r)^{2k}}\int_{B_R(x)} |w|^2\,\mathrm{d}x \quad \textit{for every } B_r(x) \subset B_R(x) \subset \Omega.$$

Proof We may re-write $\mathscr{A}^*(\sigma\mathscr{A}u)$ as the elliptic operator in divergence form

$$(-1)^k \sum \partial^\beta(\mathbf{C}^{ij}_{\beta\alpha}\partial^\alpha u^j),$$

for coefficients $\mathbf{C} = (\mathbf{A}^T\sigma\mathbf{A})$ satisfying a weak Gårding inequality as in (15). The assertion then follows from Corollary 22 in [19]. □

Using Lemma 2.5 one can show, by classical methods, the following lemma on the regularizing properties of elliptic operators with constant coefficients:

Lemma 2.6 (Constant coefficients) *Let $\mathscr{A}$ be an operator of gradient form and let $\sigma_0 \in$ $\mathrm{Lin}(\mathbb{R}^{dN^k}; \mathbb{R}^{dN^k})$ be a tensor satisfying the strong Gårding inequality (13). Then the operator*

$$L_{\sigma_0}u := \mathscr{A}^*(\sigma_0\mathscr{A}u)$$

is hypoelliptic in the sense that if Ω is open and connected, and $w \in \mathrm{L}^2(\Omega; \mathbb{R}^d)$, then

$$L_{\sigma_0}w = 0 \quad \Rightarrow \quad w \in \mathrm{C}^\infty_{\mathrm{loc}}(\Omega; \mathbb{R}^d).$$

Furthermore, there exists a constant $c = c(M, N) \geq 2^N$ such that

$$\frac{1}{\rho^N}\int_{B_\rho(x)} |\nabla^k u|^2\,\mathrm{d}x \leq \frac{c}{r^N}\int_{B_r(x)} |\nabla^k u|^2\,\mathrm{d}x \quad \textit{for all} \quad 0 < \rho \leq \frac{r}{2},$$

$$\frac{1}{\rho^N}\int_{B_\rho(x)} |\mathscr{A}u|^2\,\mathrm{d}x \leq \frac{c}{r^N}\int_{B_r(x)} |\mathscr{A}u|^2\,\mathrm{d}x \quad \textit{for all} \quad 0 < \rho \leq \frac{r}{2},$$

for every $B_r(x) \subset \Omega$.

2.1.2 Examples

Next, we gather some well-known differential structures that fit into the definition of operators of gradient form.

(i) *Gradients.* Let $\mathscr{A} : \mathrm{L}^2(\Omega; \mathbb{R}^d) \to \mathrm{W}^{-1,2}(\Omega; \mathbb{R}^{dN}) : u \mapsto (\partial_j u^i)$ for $1 \le i \le d$ and $1 \le j \le N$. In this case

$$A_j z = z \otimes \mathbf{e}_j \quad \text{for every } z \in \mathbb{R}^d.$$

Hence, $\mathrm{W}^{\mathscr{A}}(\Omega) = \mathrm{W}^{1,2}(\Omega; \mathbb{R}^d)$ and the compactness property is a consequence of the classical Poincaré inequality on Ω.
The exactness assumption is the result of the characterization of gradients via curl-free vector fields.
Let $\mathscr{B} : \mathrm{L}^2(\Omega; \mathbb{R}^{dN}) \to \mathrm{W}^{-1,2}(\Omega; \mathbb{R}^{dN^2})$ be the curl operator

$$\mathscr{B}v = (\mathrm{curl}(v^i))_i := (\partial_l v_{ir} - \partial_r v_{il})_{ilr} \quad 1 \le i \le d, \quad 1 \le l, r \le N,$$

then condition (10) is fulfilled for $\mathscr{B} = \sum_{j=1}^N B_j \partial_i$ with coefficients

$$(B_j)_{ilr,pq} = \delta_{ip}(\delta_{jl}\delta_{rq} - \delta_{jr}\delta_{lq}) \quad 1 \le l, r, q \le N, \quad 1 \le i, p \le d.$$

Observe that $\mathscr{B}v = 0$ if and only if curl $v^i = 0$, for every $1 \le i \le d$; or equivalently, $v^i = \nabla u^i$ for some function $u^i : \Omega \subset \mathbb{R}^N \to \mathbb{R}$, for every $1 \le p \le d$ (as long as Ω is simply connected). Hence,

$$\big\{\nabla u : u \in \mathrm{W}^{1,2}(\omega; \mathbb{R}^d)\big\} = \big\{v \in \mathrm{L}^2(\omega; \mathbb{R}^{dN}) : \mathscr{B}v = 0\big\},$$

for all Lipschitz, and simply connected $\omega \subset\subset \Omega$.

(ii) *Higher gradients.* Let $\mathscr{A} : \mathrm{L}^2(\Omega) \to \mathrm{W}^{-k,2}(\Omega; \mathbb{R}^{N^k})$ be the linear operator given by

$$u \mapsto \partial^\alpha u, \quad \text{where } |\alpha| = k.$$

Compactness is similar to the case of gradients.
We focus on the exactness condition: Let $\mathscr{B}^k : \mathrm{L}^2(\Omega; \mathrm{Sym}(\mathbb{R}^{N^k})) \to \mathrm{W}^{-1,2}(\Omega; \mathbb{R}^{N^{k+1}})$ be the curl operator on symmetric functions defined by the coefficients

$$(B_j^k)_{pq\beta_2\ldots\beta_k,\alpha_1\ldots\alpha_k} := \left(\delta_{jp}\delta_{\alpha_1 q}\prod_{h=2}^k \delta_{\alpha_h\beta_h} - \delta_{jq}\delta_{\alpha_1 p}\prod_{h=2}^k \delta_{\alpha_h\beta_h}\right),$$
$$1 \le p, q, \beta_h, \alpha_h \le N, \quad h \in \{2, \cdots, k\}.$$

We write

$$\mathscr{B}^k v := \sum_{i=1}^N B_j^k \, \partial_j v, \quad v : \Omega \subset \mathbb{R}^N \to \mathrm{Sym}\,(\mathbb{R}^{N^k}).$$

It easy to verify that $\mathscr{B}^k v = 0$ if and only if

$$\mathrm{curl}((v_{p\alpha'})_p) = 0 \quad \text{for all } |\alpha'| = k - 1.$$

If Ω is simply connected, then there exists a function $u^{\alpha'} : \Omega \to \mathbb{R}$ such that $v_{p\alpha'} = \partial_p u^{\alpha'}$ for every $|\alpha'| = k - 1$. Using the symmetry of v under the permutation of its coordinates one can further deduce the existence of a function $u_k : \Omega \to \mathrm{Sym}\,(\mathbb{R}^{N^{k-1}})$ with

$$v = \nabla u_k \quad \text{and} \quad (u_k)_{\alpha'} = u^{\alpha'}.$$

Moreover, $\mathscr{B}^{k-1} u_k = 0$. By induction one obtains that

$$v = \nabla^k u_0 \quad \text{for some function } u_0 : \Omega \subset \mathbb{R}^N \to \mathbb{R}.$$

(iii) *Symmetrized gradients.* Let $\mathscr{E} : \mathrm{L}^2(\Omega; \mathbb{R}^N) \to \mathrm{W}^{-1,2}(\Omega; \mathrm{Sym}\,(\mathbb{R}^{N^2}))$ be the linear operator given by

$$u \mapsto \mathscr{E}u := \frac{1}{2}(\partial_j u^i + \partial_i u^j)_{ij}, \quad \text{for } 1 \le i, j \le N.$$

The compactness property is a direct consequence of Korn's inequality. Consider the second-order homogeneous differential operator $\mathscr{B} : \mathrm{L}^2(\Omega; \mathrm{Sym}\,(\mathbb{R}^{N^2})) \to \mathrm{W}^{-2,2}(\Omega; \mathbb{R}^{N^3})$ defined in the following way

$$\mathscr{B}v = \mathrm{curl}\,(\mathrm{curl}(v)) = \left(\frac{\partial^2 v_{ij}}{\partial x_i \partial x_l} + \frac{\partial^2 v_{il}}{\partial x_i \partial x_j} - \frac{\partial^2 v_{ii}}{\partial x_j \partial x_l} - \frac{\partial^2 v_{jl}}{\partial x_i \partial x_i} \right)_{1 \le i,j,l \le N}.$$

Then $\mathscr{B}v = 0$, if and only if $v = \mathscr{E}u$ for some $u \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^N) = \mathrm{W}^{\mathscr{E}}(\Omega)$.[5]

Remark 2.7 In the previous examples, we have omitted the characterization of higher gradients of vector-valued functions; however, the ideas remain the same as in the examples (i) and (ii).

Remark 2.8 (Two-dimensional elasticity) In dimension $N = 2$ and provided that Ω is simply connected, the fourth-order equation for pure bending of a thin plate given by

$$\nabla \cdot (\nabla \cdot (\mathbf{D}(x) \nabla^2 u(x))) = 0 \quad \text{for } u \in \mathrm{W}^{2,2}(\Omega)$$

is equivalent to the in-plane elasticity equation

$$\nabla \cdot (\mathbf{S}(x) \mathscr{E} w(x)) = 0 \quad \text{where } w \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^2),$$

for some tensor $\mathbf{S}$ such that $\mathbf{D} = (\mathbf{R}_\perp \mathbf{S}^{-1} \mathbf{R}_\perp)$, and where $\mathbf{R}_\perp$ is the fourth-order tensor whose action is to rotate a second-order tensor by 90° (see, e.g., [20, Chapter 2.3]). Furthermore,

$$\mathbf{S}(x) \mathscr{E} w(x) = \mathbf{R}_\perp \nabla^2 u(x) \quad \text{and} \quad \nabla \cdot (\nabla \cdot (\mathbf{R}^\perp \mathscr{E} w(x))) = 0.$$

For this reason, when working with the linear equations for pure bending of a thin plate we may indistinctly use regularizing properties of any of the equations above in the portions where $\mathbf{D}$ is regular.

2.2 Compensated compactness

The following theorem is a generalized version of the well-known div-curl Lemma.

Lemma 2.9 *Let $\mathscr{A}$ be a k'-th order operator of gradient form and let $\{\sigma_h\} \subset \mathrm{L}^2(\Omega; \mathrm{Sym}(\mathbb{R}^{dN^k} \otimes \mathbb{R}^{dN^k}))$ be a sequence of strongly elliptic tensors as in (13). Assume also that $\{u_h\} \subset \mathrm{W}^{\mathscr{A}}(\Omega)$ and $\{f_h\} \subset \mathrm{W}^{-k,2}(\Omega; \mathbb{R}^d)$ are sequences for which*

$$\mathscr{A}^*(\sigma_h \mathscr{A} u_h) = f_h \quad \text{in } \mathscr{D}'(\Omega; \mathbb{R}^d), \quad \text{for every } h \in \mathbb{N}.$$

[5] Here, $\mathscr{B}$ is a second order operator expressing the Saint-Venant compatibility conditions.

Further assume there exist $\sigma \in \mathrm{L}^2(\Omega; \mathrm{Sym}(\mathbb{R}^{dN^k} \otimes \mathbb{R}^{dN^k}))$*,* $u \in \mathrm{W}^{\mathscr{A}}(\Omega)$*, and* $f \in \mathrm{W}^{-k,2}(\Omega; \mathbb{R}^d)$ *for which*

$$\mathscr{A}u_h \rightharpoonup \mathscr{A}u \quad \text{in } \mathrm{L}^2(\Omega; \mathbb{R}^{dN^k}), \quad f_h \to f \quad \text{in } \mathrm{W}^{-k,2}(\Omega; \mathbb{R}^d), \quad \text{and} \quad \sigma_h \to \sigma \quad \text{in } \mathrm{L}^2(\Omega; \mathbb{R}^{dN^k} \otimes \mathbb{R}^{dN^k}).$$

Then,

$$\mathscr{A}^*(\sigma \mathscr{A}u) = f \quad \text{in } \mathscr{D}'(\Omega; \mathbb{R}^d), \quad \sigma_h \mathscr{A}u_h \cdot \mathscr{A}u_h \to \sigma \mathscr{A}u \cdot \mathscr{A}u \quad \text{in } \mathscr{D}'(\Omega).$$

In particular,

$$\mathscr{A}u_h \to \mathscr{A}u \quad \text{in } \mathrm{L}^2_{\mathrm{loc}}(\Omega; \mathbb{R}^{dN^k}).$$

Proof For simplicity we denote $\tau_h := \sigma_h \mathscr{A}u_h$, $\tau := \sigma \mathscr{A}u$. It suffices to observe that $\tau_h \rightharpoonup \tau$ in L^2 to prove that

$$\mathscr{A}^*\tau = f \quad \text{in } \mathscr{D}'(\Omega; \mathbb{R}^d).$$

The strong convergence on compact subsets of Ω requires a little bit more effort. Considering that $\mathscr{A}$ is a k'-th order linear differential operator, we may find constants $c_{\alpha\beta}$ with $|\alpha| + |\beta| \leq k$, $|\beta| \geq 1$ such that

$$\mathscr{A}(u_h\varphi) = (\mathscr{A}u_h)\varphi + \sum_{\alpha,\beta} c_{\alpha\beta}\partial^\alpha u_h \partial^\beta \varphi \in \mathrm{L}^2(\Omega; \mathbb{R}^d) \quad \forall\, \varphi \in \mathscr{D}(\Omega), \forall\, h \in \mathbb{N}.$$

Hence,

$$\langle \tau_h \cdot \mathscr{A}u_h, \varphi \rangle = \langle f_h, u_h\varphi \rangle - \left\langle \tau_h, \sum_{\alpha,\beta} c_{\alpha\beta}\partial^\alpha u_h \partial^\beta \varphi \right\rangle.$$

By the compactness assumption on $\mathscr{A}$ we may assume without loss of generality that $u_h \rightharpoonup u$ in $\mathrm{W}^{k,2}(\Omega; \mathbb{R}^d)$. Thus, passing to the limit we obtain

$$\lim_{h\to\infty} \langle \tau_h \cdot \mathscr{A}u_h, \varphi \rangle = \langle f, u\varphi \rangle - \left\langle \tau, \sum_{\alpha,\beta} c_{\alpha\beta}\partial^\alpha u \, \partial^\beta \varphi \right\rangle = \langle \tau \cdot \mathscr{A}u, \varphi \rangle,$$

for every $\varphi \in \mathscr{D}(\Omega)$. One concludes that

$$\sigma_h \mathscr{A}u_h \cdot \mathscr{A}u_h \to \sigma \mathscr{A}u \cdot \mathscr{A}u \quad \text{in } \mathscr{D}'(\Omega). \tag{16}$$

Fix $\omega \subset\subset \Omega$ and let $0 \leq \varphi \in \mathscr{D}(\Omega)$ with $\varphi \equiv 1$ on ω. Using the convergence in (16), the uniform ellipticity (2) and the symmetry of $\{\sigma_h\}$, one gets

$$\begin{aligned}
\lim_{h\to\infty} \|\mathscr{A}u_h - \mathscr{A}u\|_{\mathrm{L}^2(\omega)} &\leq M \cdot \lim_{h\to\infty} \langle \sigma_h(\mathscr{A}(u_h - u)) \cdot \mathscr{A}(u_h - u), \varphi \rangle \\
&\leq M \cdot \Bigg(\lim_{h\to\infty} \langle \sigma_h \mathscr{A}u_h \cdot \mathscr{A}u_h, \varphi \rangle \\
&\quad - \lim_{h\to\infty} 2\langle \sigma_h \mathscr{A}u_h \cdot \mathscr{A}u, \varphi \rangle + \langle \sigma_h \mathscr{A}u \cdot \mathscr{A}u, \varphi \rangle \Bigg) \\
&= 0.
\end{aligned}$$

□

2.3 Young measures and lower semi-continuity of integral energies

In this section $\mathscr{B} : \mathrm{L}^2(\Omega; Z) \to \mathrm{W}^{-m,2}(\Omega; \mathbb{R}^n)$ is assumed to be a an m'-th order homogeneous partial differential operator of the form

$$\sum_{\alpha} B_\alpha \partial^\alpha, \quad B_\alpha \in \mathrm{Lin}\,(Z; \mathbb{R}^n), \text{ with Z a linear subspace of } \mathbb{R}^{dN^k},$$

satisfying the constant rank condition (†).

Next, we recall some facts about $\mathscr{B}$-quasiconvexity, lower semi-continuity and Young measures. The results in this section hold for differential operators with coefficients B_α in arbitrary spaces Lin $(\mathbb{R}^p; \mathbb{R}^q)$ for p, q a pair of positive integers; however, we only present versions where the dimensions match our current setting. We start by stating a version of the Fundamental theorem for Young measures due to Ball [21].

Theorem 2.10 (Fundamental theorem for Young measures) *Let $\Omega \subset \mathbb{R}^N$ be a measurable set with finite measure and let $\{v_j\}$ be a sequence of measurable functions $v_j : \Omega \to Z$. Then there exists a subsequence $\{v_{h(j)}\}$ and a weak* measurable map $\mu : \Omega \to \mathscr{M}(Z)$ with the following properties:*

1. *We denote $\mu_x := \mu(x)$ for simplicity, then $\mu_x \geq 0$ in the sense of measures and $|\mu_x|(Z) \leq 1$ for a.e. $x \in \Omega$.*
2. *If one additionally assumes that $\{v_{h(j)}\}$ is uniformly bounded in $\mathrm{L}^1(\Omega; Z)$, then $|\mu_x|(Z) = 1$ for a.e. $x \in \Omega$.*
3. *If $F : \mathbb{R}^{dN^k} \to \mathbb{R}$ is a Borel and lower semi-continuous function, and is also bounded from below, then*
$$\int_\Omega \langle \mu_x, F \rangle \,\mathrm{d}x \leq \liminf_{j\to\infty} \int_\Omega F(v_{h(j)}) \,\mathrm{d}x.$$
4. *If $\{v_{h(j)}\}$ is uniformly bounded in $\mathrm{L}^1(\Omega; Z)$ and $F : \mathbb{R}^{dN^k} \to \mathbb{R}$ is a continuous function, and bounded from below, then*
$$\int_\Omega \langle \mu_x, F \rangle \,\mathrm{d}x = \liminf_{j\to\infty} \int_\Omega F(v_{h(j)}) \,\mathrm{d}x$$
if and only if $\{F \circ v_{h(j)}\}$ is equi-integrable. In this case,
$$F \circ v_{h(j)} \rightharpoonup \langle \mu_x, F \rangle \quad \textit{in } \mathrm{L}^1(\Omega).$$

In the sense of Theorem 2.10, we say that the sequence $\{v_{h(j)}\}$ generates the Young measure μ.

The following proposition tells us that a uniformly bounded sequence in the L^p norm, which is also sufficiently close to Ker($\mathscr{B}$), may be approximated by a p-equi-integrable sequence in Ker($\mathscr{B}$) in a weaker L^q norm. We remark that this rigidity result is the only one where Murat's constant rank condition (†) is used.

Proposition 2.11 [16, Lemma 2.15] *Let $1 < p < \infty$. Let $\{v_h\}$ be a bounded sequence in $\mathrm{L}^p(\Omega; Z)$ generating a Young measure μ, with $v_h \rightharpoonup v$ in $\mathrm{L}^p(\Omega; Z)$ and $\mathscr{B}v_h \to 0$ in $\mathrm{W}^{-m,p}(\Omega; \mathbb{R}^n)$. Then there exists a p-equi-integrable sequence $\{u_h\}$ in $\mathrm{L}^p(\Omega; Z) \cap$ Ker($\mathscr{B}$) that generates the same Young measure μ and is such that*

$$\int_\Omega v_h \,\mathrm{d}x = \int_\Omega u_h \,\mathrm{d}x, \quad \|v_h - u_h\|_{\mathrm{L}^q(\Omega)} \to 0, \quad \textit{for all } 1 \leq q < p.$$

□

Let $F : \mathbb{R}^{dN^k} \to \mathbb{R}$ be a lower semi-continuous function with $0 \leq F(P) \leq C(1 + |P|^p)$ for some positive constant C. The $\mathscr{B}$-quasiconvex envelope of F at $P \in Z \subset \mathbb{R}^{dN^k}$ is defined as

$$Q_{\mathscr{B}} F(P) := \inf \left\{ \int_{[0,1]^N} F(P + v(y)) \, dy : \right. \\ \left. v \in C^\infty_{\text{per}}\left([0,1]^N; Z\right), \quad \mathscr{B}v = 0 \text{ and } \int_{[0,1]^N} v \, dy = 0 \right\}. \tag{17}$$

The most relevant feature of $Q_{\mathscr{B}} F$ is that, for $p > 1$, the lower semi-continuous envelope with respect to the weak-L^p topology of the functional

$$v \mapsto \int_\Omega F(v) \, dx, \quad \text{where } v \in L^p(\Omega; Z) \text{ and } \mathscr{B}v = 0, \tag{18}$$

is given by the functional

$$v \mapsto \int_\Omega Q_{\mathscr{B}} F(v) \, dx, \quad \text{where } v \in L^p(\Omega; Z) \text{ and } \mathscr{B}v = 0.$$

If μ is a Young measure generated by a sequence $\{v_h\}$ in $L^p(\Omega; Z)$ such that $\mathscr{B}v_h = 0$ for every $h \in \mathbb{N}$, then we say that μ is a $\mathscr{B}$-free Young measure.

We recall the following Jensen inequality for $\mathscr{B}$-free Young measures [16, Theorem 4.1]:

Theorem 2.12 *Let $1 < p < \infty$. Let μ be a $\mathscr{B}$-free Young measure in Ω. Then for a.e. $x \in \Omega$ and all lower semi-continuous functions that satisfy $|F(P)| \leq C(1+|P|^p)$ for some positive constant C and all $P \in \mathbb{R}^{dN^k}$, one has that*

$$\langle \mu_x, F \rangle \geq Q_{\mathscr{B}} F(\langle \mu_x, \text{id} \rangle).$$

2.4 Geometric measure theory and sets of finite perimeter

Most of the facts collected in this section can be found in [22] and [23]; however, some notions as the slicing of sets of finite perimeter are presented there only in a formal way. For a better understanding of such topics we refer the reader to [24].

Let $A \subset \mathbb{R}^N$ be a Borel set. The Gauss-Green measure μ_A of A is the derivative of the characteristic function of A in the sense of distributions, i.e., $\mu_A := D(\mathbb{1}_A)$. We say that A is a set of locally finite perimeter if and only if $|\mu_A|$ is a vector-valued Radon measure in $\mathbb{R}^N$. We write $A \in BV_{\text{loc}}(\mathbb{R}^N)$ to express that A is a set of locally finite perimeter in $\mathbb{R}^N$.

Let $\omega \subset\subset \mathbb{R}^N$ be a Borel set. The perimeter in ω of a set A with locally finite perimeter is defined as

$$\text{Per}(A, \omega) := |\mu_A|(\omega).$$

The Radon–Nikodým differentiation theorem states that the set of points

$$\partial^* A := \left\{ x \in \mathbb{R}^N : \lim_{r \downarrow 0} \frac{\text{Per}(A; B_r(x))}{\text{vol}\,(B'_1) \cdot r^{N-1}} = 1, \right. \\ \left. \text{and } \frac{d\mu_A}{d|\mu_A|}(x) \text{ exists and belongs to } \mathbb{S}^{N-1} \right\}$$

has full $|\mu_A|$-measure in $\mathbb{R}^N$; this set is commonly known as the *reduced boundary* of A. We will also use the notation

$$\nu_A(x) := \frac{\mathrm{d}\mu_A}{\mathrm{d}|\mu_A|}(x) \quad x \in \partial^* A;$$

the *measure theoretic normal* of A.

In general, for $s \geq 0$, we will denote by $\mathscr{H}^s$ the *s-dimensional Hausdorff measure in* $\mathbb{R}^N$. The following well-known theorem captures the structure of sets with finite perimeter in terms of the measure $\mathscr{H}^{N-1}$:

Theorem 2.13 (De Giorgi's Structure Theorem) *Let A be a set of locally finite perimeter. Then*

$$\partial^* A = \bigcup_{j=1}^{\infty} K_j \cup N,$$

where

$$|\mu_A|(N) = 0,$$

and each K_j is a compact subset of a C^1-hypersurface S_j for every $j \in \mathbb{N}$. Furthermore, $\nu_A|_{S_j}$ is normal to S_j and

$$\mu_A = \nu_A \, \mathscr{H}^{N-1} \llcorner \partial^* A.$$

From De Giorgi's Structure Theorem it is clear that spt $\mu_A = \overline{\partial^* A}$. Actually, up to modifying A on a set of zero measure, one has that $\partial A = \overline{\partial^* A}$ (see [22, Proposition 12.19]). From this point on, each time we deal with a set A of finite perimeter, we will assume without loss of generality that

$$\partial A = \text{spt } \mu_A = \overline{\partial^* A}. \tag{19}$$

For a set of locally finite perimeter A, the deviation from being a *perimeter minimizer* in Ω, at a given scale r, is quantified by the monotone function

$$\mathrm{Dev}_\Omega(A, r) := \sup\{\mathrm{Per}(A; B_r(x)) - \mathrm{Per}(E; B_r(x)) : E \Delta A \subset\subset B_r(x) \subset \Omega\}.$$

The next result, due to Tamanini [25], states that a set of locally finite perimeter with small deviation Dev_Ω at every scale is actually a C^1-hypersurface up to a lower dimensional set.

Theorem 2.14 *Let $A \subset \mathbb{R}^N$ be a set of locally finite perimeter and let $c(x)$ be a locally bounded function for which*

$$Dev_\Omega(A, r) \leq c(x) r^{N-1+2\eta} \quad \text{for some } \eta \in (0, 1/2\,].$$

Then the reduced boundary in Ω, $(\partial^* A \cap \Omega)$, is an open $\mathrm{C}^{1,\eta}$-hypersurface and the singular set $\Omega \cap (\partial A \backslash \partial^* A)$ has at most Hausdorff dimension $(N - 8)$.

2.4.1 Slicing sets of finite perimeter

Given a Borel set $E \subset \mathbb{R}^N$ and a Lipschitz function $g : \mathbb{R}^N \to \mathbb{R}$, we shall consider the level set slices

$$E_t := E \cap \{g = t\}, \quad t \in \mathbb{R}.$$

For a set $A \subset \mathbb{R}^N$ of finite perimeter in Ω, the level set slice of the reduced boundary $(\partial^* A)_t$ is $\mathscr{H}^{N-2}$-rectifiable for almost every $t \in \mathbb{R}$. Furthermore, by the co-area formula, $t \mapsto \mathscr{H}^{N-2}((\partial^* A)_t) \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$.

If the set $\{g = t\}$ is a C^1-manifold and t is such that $\mathscr{H}^{N-2}((\partial^* A)_t) < \infty$, we shall define the *slice of* A in $g^{-1}\{t\}$ as

$$\langle A, g, t\rangle := \mathscr{H}^{N-2} \llcorner (\partial^* A)_t.$$

It turns out that, for $g(x) = |x|$, the level set slice A_t is locally diffeomorphic to a set of finite perimeter in $\mathbb{R}^{N-1}$. Even more,

$$\mathscr{H}^{N-2} \llcorner \partial^* A_t = \langle A, g, t\rangle \qquad \text{for a.e. } t > 0, \text{ and} \tag{20}$$

$$\pi_g \nu_A := (\mathrm{id}_{\mathbb{R}^N} - \nabla g \otimes \nabla g)\nu_A \neq 0 \qquad \text{for } \mathscr{H}^{N-2}\text{- a.e. } x \in (\partial^* A)_t. \tag{21}$$

Here, $\partial^* A_t$ is understood as the image, under local diffeomorphisms, of the reduced boundary of a set of finite perimeter. These properties can be inferred from the classical slicing by hyperplanes, see e.g., [22, Chapter 18.3].

We also define the cone extension of a set $E \subset \mathbb{R}$ containing $\{0\}$ by letting

$$D_E := \{\lambda x \in \mathbb{R}^N : \lambda > 0,\ x \in E\}.$$

For a.e. $t > 0$ and $g(x) = |x|$, the cone extension of A_t is a set of locally finite perimeter in $\mathbb{R}^N$ with

$$\partial^* D_{A_t} = D_{(\partial^* A)_t} \quad \text{and} \quad \mathrm{Per}(D_{A_t}; B_\rho) = \left(\frac{1}{N-1}\right)\frac{\rho^{N-1}}{t^{N-2}} \cdot \mathscr{H}^{N-2}((\partial^* A)_t). \tag{22}$$

In order to attend different variational problems involving the minimization of perimeter, a well-known technique is to modify a set A within balls B_t without modifying its Gauss-Green measure in $(B_t)^c$.

For almost every $t > 0$, where $\langle A, g, t\rangle$ is well-defined and (20), (21) hold, we construct a cone-like comparison set of A by setting

$$\tilde{A} := \mathbb{1}_{B_t} D_{A_t} + \mathbb{1}_{\Omega \setminus B_t} A. \tag{23}$$

Exploiting the basic properties of reduced boundaries, it follows by (20) that

$$\mu_{\tilde{A}} = \mu_{D_{A_t}} \llcorner B_t + \mu_A \llcorner (B_t)^c; \tag{24}$$

and, in particular,

$$\mathrm{Per}(\tilde{A}; B_r) = \mathrm{Per}(D_{\partial^* A_t}; B_t) + \mathrm{Per}(A; (B_t)^c \cap B_r) \quad \text{for all } r > t.$$

On the other hand, again by the co-area formula,

$$\mathscr{H}^{N-1}((\partial^* A)_t \cap \{g = t\}) = 0 \quad \text{for almost every } t > 0.$$

Using the monotonicity of $r \mapsto \mathrm{Per}(A; B_r)$ and the general version of the co-area formula (see [24, Theorem 3.2.22]) one can show that the derivative of $r \mapsto \mathrm{Per}(A; B_r)$ exists at almost every $t > 0$; even more, one has that

$$\left.\frac{\mathrm{d}}{\mathrm{d}r}\right|_{r=t} \mathrm{Per}(A; B_r) \geq |\pi_t \nu_A|^{-1} \mathscr{H}^{N-2}((\partial^* A)_t) \geq \langle A, g, t\rangle(\mathbb{R}^N). \tag{25}$$

The previous estimate will play a crucial role in proving the Lower bound (LB).

3 Existence of solutions: proof of Theorem 1.1

We show an equivalence between the constrained problem (5) and the unconstrained problem (P) for which existence of solutions and regularity properties for minimizers are discussed in the present and subsequent sections. We fix $\mathscr{A} : \mathrm{L}^2(\Omega\ \mathbb{R}^d) \to \mathrm{W}^{-k,2}(\Omega; \mathbb{R}^{dN^k})$ an operator of gradient from as in Definition 2.1. We also fix $A_0 \subset \mathbb{R}^N$, a set of locally finite perimeter.

Recall that, the minimization problem (5) under the state constraint (4) reads:

$$\text{minimize} \quad \left\{ \int_\Omega f w_A \ + \ \mathrm{Per}(A; \overline{\Omega}) \ : \ A \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^N), \ A \cap \Omega^c \equiv A_0 \cap \Omega^c \right\},$$

where w_A is the unique distributional solution to the state equation

$$\mathscr{A}^*(\sigma_A \mathscr{A} u) = f, \quad u \in \mathrm{W}_0^{\mathscr{A}}(\Omega).$$

On the other hand, the associated saddle point problem (P) reads[6]:

$$\inf \left\{ \sup_{u \in \mathrm{W}_0^{\mathscr{A}}(\Omega)} I_\Omega(u, A) : A \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^N), \ A \cap \Omega^c \equiv A_0 \cap \Omega^c \right\}, \tag{P}$$

where

$$I_\Omega(u, A) := \int_\Omega 2 f u \,\mathrm{d}x - \int_\Omega \sigma_A \mathscr{A} u \cdot \mathscr{A} u \,\mathrm{d}x \ + \ \mathrm{Per}(A; \overline{\Omega}).$$

Theorem 1.1 (Existence) *There exists a solution* (w, A) *of problem* (P). *Furthermore, there is a one to one correspondence*

$$(w, A) \mapsto (w_A, A)$$

between solutions of the problem (P) *and solutions of the minimization problem* (5) *under the constraint* (4).

Proof We employ the direct method. We begin by proving existence of solutions to problem (P). To do so, we will first prove the following:

Claim 1. *For any set* $A \subset \mathbb{R}^N$ *as in the assumptions, there exists* $w_A \in \mathrm{W}_0^{\mathscr{A}}(\Omega)$ *such that*

$$0 \leq I_\Omega(w_A, A) = \sup_{u \in \mathrm{W}_0^{\mathscr{A}}(\Omega)} I_\Omega(u, A) < \infty.$$

The tensor σ_A is a positive definite tensor and therefore the mapping

$$u \mapsto I_\Omega(u, A) = \int_\Omega 2 f u - \sigma_A \mathscr{A} u \cdot \mathscr{A} u \,\mathrm{d}x + \mathrm{Per}\,(A; \overline{\Omega})$$

is strictly concave. Observe that $\sup_{u \in \mathrm{W}_0^{\mathscr{A}}(\Omega)} I_\Omega(u, A) \geq \mathrm{Per}\,(A; \overline{\Omega})$; indeed, we may take $u \equiv 0 \in \mathrm{W}_0^{\mathscr{A}}(\Omega)$. Hence,

$$\sup_{u \in \mathrm{W}_0^{\mathscr{A}}(\Omega)} I_\Omega(u, A) \geq \mathrm{Per}\,(A; \overline{\Omega}) \geq 0. \tag{26}$$

[6] As stated in Sect. 2.4, we write $A \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^N)$ to express that A is a Borel set of locally finite perimeter in $\mathbb{R}^N$.

Because of this, we may find a maximizing sequence $\{w_h\}$ in $\mathrm{W}_0^{\mathscr{A}}(\Omega)$, i.e.,

$$I_\Omega(w_h, A) \to \sup_{u \in \mathrm{W}_0^{\mathscr{A}}(\Omega)} I_\Omega(u, A), \quad \text{as } h \text{ tends to infinity.}$$

Even more, one has from (2) that

$$-\frac{1}{M}\|\mathscr{A} w_h\|^2_{\mathrm{L}^2(\Omega)} \geq -\int_\Omega \sigma_A \mathscr{A} w_h \cdot \mathscr{A} w_h \,\mathrm{d}x$$

and consequently from (26) and (12) one infers that

$$C(\Omega)^{-1} \cdot \limsup_{h\to\infty} \frac{1}{M}\|w_h\|^2_{\mathrm{L}^2(\Omega)} \leq \limsup_{h\to\infty} \frac{1}{M}\|\mathscr{A} w_h\|^2_{\mathrm{L}^2(\Omega)} \leq 2\|f\|_{\mathrm{L}^2(\Omega)} \cdot \limsup_{h\to\infty} \|w_h\|_{\mathrm{L}^2(\Omega)}. \tag{27}$$

A fast calculation shows that $\|w_h\|_{\mathrm{L}^2(\Omega)} \leq 2MC(\Omega)\|f\|_{\mathrm{L}^2(\Omega)}$; in return, (27) also implies that

$$\limsup_{h\to\infty} \|\mathscr{A} w_h\|^2_{\mathrm{L}^2(\Omega)} \leq 4C(\Omega)M^2\|f\|^2_{\mathrm{L}^2(\Omega)}.$$

Hence, using again the compactness property of $\mathscr{A}$, we may pass to a subsequence (which we will not relabel) and find $w_A \in \mathrm{W}_0^{\mathscr{A}}(\Omega)$ with

$$w_h \to w_A \quad \text{in } \mathrm{L}^2(\Omega; \mathbb{R}^d), \quad \mathscr{A} w_h \rightharpoonup \mathscr{A} w_A \quad \text{in } \mathrm{L}^2(\Omega; \mathbb{R}^{dN^k}).$$

The concavity of $-\sigma_A z \cdot z$ is a well-known sufficient condition for the upper semi-continuity of the functional $\mathscr{A} u \mapsto -\int_\Omega \sigma_A \mathscr{A} u \cdot \mathscr{A} u$. Therefore,

$$\sup_{u \in \mathrm{W}_0^{\mathscr{A}}(\Omega)} I_\Omega(u, A) = \lim_{h\to\infty} I_\Omega(w_h, A) \leq I_\Omega(w_A, A).$$

This proves the claim.

Now, we use **Claim 1** to find a minimizing sequence $\{A_h\}$ for $A \mapsto I_\Omega(w_A, A)$. Since the uniform bound (27) does not depend on A, we may again assume (up to a subsequence) that there exists $\tilde{w} \in \mathrm{W}_0^{\mathscr{A}}(\Omega)$ such that

$$w_{A_h} \to \tilde{w} \quad \text{in } \mathrm{L}^2(\Omega; \mathbb{R}^d), \quad \mathscr{A} w_{A_h} \rightharpoonup \mathscr{A} \tilde{w} \quad \text{in } \mathrm{L}^2(\Omega; \mathbb{R}^{dN^k}), \quad \text{and} \quad \mathscr{A}^*(\sigma_{A_h} \mathscr{A} w_{A_h}) = f.$$

Even more, since $\{A_h\}$ is minimizing, it must be that $\sup_h\{\mathrm{Per}(A_h; B_R)\} < \infty$, for some ball B_R properly containing Ω, and thus (for a further subsequence) there exists a set $\tilde{A} \subset \mathbb{R}^N$ of locally finite perimeter with $\tilde{A} \cap \Omega^c \equiv A_0 \cap \Omega^c$ and such that

$$\mathbb{1}_{A_h} \to \mathbb{1}_{\tilde{A}} \text{ in } \mathrm{L}^1(B_R), \quad |\mu_{\tilde{A}}|(B_R) \leq \liminf_{h\to\infty} |\mu_{\tilde{A}_h}|(B_R).$$

Therefore

$$\begin{aligned}\mathrm{Per}(\tilde{A}; \overline{\Omega}) &= |\mu_{\tilde{A}}|(B_R) - |\mu_{A_0}|(B_R \backslash \overline{\Omega}) \\ &\leq \liminf_{h\to\infty} |\mu_{A_h}|(B_R) - |\mu_{A_0}|(B_R \backslash \overline{\Omega}) = \liminf_{h\to\infty} \mathrm{Per}(A_h; \overline{\Omega})\end{aligned} \tag{28}$$

A consequence of Lemma 2.9 is that[7]

$$\mathscr{A}^*(\sigma_{\tilde{A}} \mathscr{A} \tilde{w}) = f \quad \text{in } \mathscr{D}'(\Omega; \mathbb{R}^d), \quad \text{and} \quad \int_\Omega \sigma_{A_h} \mathscr{A} w_{A_h} \cdot \mathscr{A} w_{A_h} \to \int_\Omega \sigma_{\tilde{A}} \mathscr{A} \tilde{w} \cdot \mathscr{A} \tilde{w}. \tag{29}$$

[7] The convergence of the total energy is not covered by Lemma 2.9; however, this can be deduced using integration by parts and the fact that w_h has zero boundary values for every $h \in \mathbb{N}$.

By taking the limit as h goes to infinity we get from (28) and the convergence above that

$$\min_{A} \sup_{u \in \mathrm{W}_0^{\mathscr{A}}(\Omega)} I_\Omega(u, A) = \lim_{h \to \infty} I_\Omega(w_{A_h}, A_h) \geq I_\Omega(\tilde{w}, \tilde{A}) = I_\Omega(w_{\tilde{A}}, \tilde{A}),$$

where the last equality is a consequence of the identity $\tilde{w} = w_{\tilde{A}}$ which can be easily derived by using the equation and the strict concavity of I_Ω in the first variable. Thus, the pair $(w_{\tilde{A}}, \tilde{A})$ is a solution to problem (P).

The equivalence of problem (P) and problem (5) under the state constraint (4) follows easily from (29), the strict concavity of $I_\Omega(\cdot, A)$, and a simple integration by parts argument. □

4 The energy bound: proof of Theorem 1.2

Throughout this section and for the rest of the manuscript we fix $\mathscr{A} : \mathrm{L}^2(\Omega; \mathbb{R}^d) \to \mathrm{W}^{-k,2}(\Omega; \mathbb{R}^{dN^k})$ in the class of operators of gradient form. Accordingly, the notations Z and $\mathscr{B}$ shall denote the subspace of $\mathbb{R}^{dN^k}$ and the homogeneous operator associated to $\mathscr{A}$ (see Definition 2.1). We will also write (w, A) to denote a particular solution of problem (P).

Consider the energy $J_\omega : \mathrm{L}^2(\Omega; Z) \times \mathfrak{B}(\mathbb{R}^N) \to \mathbb{R}$ defined as

$$J_\omega(v, E) := \int_\omega \sigma_E v \cdot v \,\mathrm{d}y \; + \; \mathrm{Per}(E; \omega), \quad \text{for } \omega \subset \mathbb{R}^N \text{ an open set.}$$

The goal of this section is to prove a local bound for the map $x \mapsto J_{B_r(x)}(\mathscr{A}w, A)$. More precisely, we aim to prove that for every compactly contained subset K of Ω there exists a positive number Λ_K such that

$$J_{B_r(x)}(\mathscr{A}w, A) \leq \Lambda_K r^{N-1} \quad \text{for all } x \in K \text{ and every } r \in (0, \mathrm{dist}(K, \partial\Omega)). \tag{30}$$

Our strategy will be the following. We first define a one-parameter family J^ε of perturbations of J_{B_1} in the perimeter term. In Theorem 4.2 we show that, as the perimeter term vanishes, these perturbations Γ-converge (with respect to the L^2-weak topology) to the relaxation of the energy

$$w \mapsto \int_{B_1} W(\mathscr{A}w) \,\mathrm{d}x,$$

for which we will assume certain regularity properties (cf. property (Reg)). Then, using a compensated compactness argument, we prove Theorem 1.2 (Upper bound) by transferring the regularity properties of the relaxed problem to our original problem.

Before moving forward, let us shortly discuss how the higher integrability property (Reg) stands next to the standard assumption that the materials σ_1 and σ_2 are well-ordered.

4.1 A digression on the regularization assumption

As commented beforehand in the introduction, a key assumption in the proof of the upper bound (30) is that *generalized* local minimizers of the energy

$$u \mapsto \int_{B_1} W(\mathscr{A}u) \,\mathrm{d}y, \quad \text{where } u \in \mathrm{W}^{\mathscr{A}}(B_1),$$

possess improved decay estimates. More precisely, we require that *local* minimizers $\tilde{u}$ of the functional

$$u \mapsto \int_{B_1} Q_{\mathscr{B}} W(\mathscr{A} u)\, \mathrm{d}y, \quad \text{where } u \in \mathrm{W}^{\mathscr{A}}(B_1), \tag{31}$$

possess a higher integrability estimate of the form

$$[\mathscr{A}\tilde{u}]^2_{\mathrm{L}^{2,N-\delta}(B_{1/2})} \leq c\|\mathscr{A}\tilde{u}\|^2_{\mathrm{L}^2(B_1)} \quad \text{for some } \delta \in [0,1). \tag{Reg}$$

Only then, we will be able to transfer a decay estimate of order ρ^{N-1} to solutions of our original problem.

Remark 4.1 (The case of gradients) In the case $\mathscr{A} = \nabla$, condition (Reg) boils down to regularity above the critical $\mathrm{C}^{0,1/2}$ local regularity. More specifically,

$$\frac{1}{r^{N-\delta+2}} \int_{B_r(x)} |w - (w)_{r,x}|^2\, \mathrm{d}y \leq [\nabla w]^2_{\mathrm{L}^{2,N-\delta}(B_{1/2})} \leq c\|\nabla w\|^2_{\mathrm{L}^2(B_1)}$$

for all $B_r(x) \subset B_{1/2}$.

By Poincaré's inequality and Campanato's Theorem one can easily deduce that $w \in \mathrm{C}^{0,\frac{1}{2}+\varepsilon}_{\mathrm{loc}}(B_{1/2})$ (cf. [9]).

Let us give a short account of some cases where one may find (Reg) to be a *natural assumption*.

4.1.1 The well-ordered case

The notion of well-ordering in Materials Science is not only justified as the comparability of two materials, one being at least *better* than the other. It has also been a consistent assumption when dealing with optimization problems because it allows explicit calculations. See for example [1,26,27], where the authors discuss how the well-ordering assumption plays a role in proving the optimal lower bounds of an effective tensor made-up by two materials. If σ_1 and σ_2 are well-ordered, say $\sigma_2 \geq \sigma_1$ as quadratic forms, then $W(P) = \sigma_2 P \cdot P$. Hence, by Lemma 2.6, the desired higher integrability (Reg) holds with $\delta = 0$.

4.1.2 The non-ordered case

Applications for this setting are mostly reserved for gradients of scalar valued functions. In this particular case one can ensure that $Q_{\mathscr{B}} W = W^{**}$, where W^{**} is the convex envelope of W. For example, one may consider an optimal design problem involving the linear conductivity equations for two dielectric materials which happen to be incomparable as quadratic forms. In this setting, it is not hard to see that indeed $QW = W^{**}$ and even that $W^{**} \in \mathrm{C}^{1,1}(\mathbb{R}^{dN^k}, \mathbb{R})$. In dimensions $N = 2, 3$, one can employ a Moser-iteration technique for the dual problem as the one developed in [17] to show better regularity of minimizers of (31).

Regarding the case of systems, if no well-ordering of the materials is assumed, it is not clear to us that (Reg) necessary holds (compare to [28,29]).

4.2 Proof of Theorem 1.2

We define an ε-perturbation of $v \mapsto \int_{B_1} \sigma_A v \cdot v$ as follows. Consider the functional

$$(v, A) \mapsto J^{\varepsilon}(v, A) := \int_{B_1} \sigma_A v \cdot v \, \mathrm{d}y \; + \; \varepsilon^2 \mathrm{Per}(A; B_1), \quad \text{for } \varepsilon \in [0, 1]; \quad J := J^1. \tag{32}$$

By a scaling argument one can easily check that

$$\varepsilon^2 J(v, A) = J^{\varepsilon}(\varepsilon v, A). \tag{33}$$

Furthermore,

$$v \text{ is a local minimizer of } J(\,\cdot\,, A) \text{ if and only if } \varepsilon v \text{ is a local minimizer of } J^{\varepsilon}(\,\cdot\,, A). \tag{34}$$

We also consider the following one-parameter family of functionals:

$$v \mapsto G^{\varepsilon}(v) := \begin{cases} \min\limits_{A \in \mathfrak{B}(\mathbb{R}^N)} J^{\varepsilon}(v, A) & \text{if } v \in \mathrm{L}^2(\Omega; Z) \text{ and } \mathscr{B}v = 0, \\ \infty & \text{otherwise.} \end{cases} \tag{35}$$

The next result characterizes the Γ-limit of these functionals as ε tends to zero.

Theorem 4.2 *The Γ-limit of the functionals G^{ε}, as ε tends to zero, and with respect to the weak-L^2 topology is given by the functional*

$$G(v) := \begin{cases} \int_{B_1} Q_{\mathscr{B}} W(v) \, \mathrm{d}y & \text{if } v \in \mathrm{L}^2(\Omega; Z) \text{ and } \mathscr{B}v = 0, \\ \infty & \text{else}. \end{cases} \tag{36}$$

Proof We divide the proof into three steps. First, we will prove the following auxiliary lemma.

Lemma 4.3 *Let $\omega \subset \mathbb{R}^N$ be an open and bounded domain. Let $p > 1$ and let $F : \mathbb{R}^{dN^k} \to [0, \infty)$ be a continuous integrand with p-growth, i.e.,*

$$0 \le F(P) \le C(1 + |P|^p), \quad P \in \mathbb{R}^{dN^k}.$$

If $v \in \mathrm{L}^p(\omega; Z)$ and $\mathscr{B}v = 0$, then there exists a p-equi-integrable recovery sequence $\{v_h\} \subset \mathrm{L}^p(\omega; Z)$ for v such that

$$\mathscr{B}v_h = 0 \quad \text{and} \quad F(v_h) \rightharpoonup Q_{\mathscr{B}} F(v) \quad \text{in } \mathrm{L}^1(\omega).$$

Proof Since $v \mapsto \int_{\omega} Q_{\mathscr{B}} F(v)$ is the lower semi-continuous envelope of $v \mapsto \int_{\omega} F(v)$ (see (17), (18)) with respect to the weak-L^p topology, we may find a sequence $\{v_h\}$ with the following properties:

$$\mathscr{B}v_h = 0, \quad v_h \overset{\mathrm{L}^p}{\rightharpoonup} v,$$

and

$$\int_{\omega} Q_{\mathscr{B}} F(v) \, \mathrm{d}x \ge \int_{\omega} F(v_h) \, \mathrm{d}x - \frac{1}{h}.$$

Passing to a subsequence if necessary, we may assume that the sequence $\{v_h\}$ generates a $\mathscr{B}$-free Young measure which we denote by μ. We then apply [16, Lemma 2.15] to find a p-equi-integrable sequence $\{v_h'\}$ (with $\mathscr{B}v_h = 0$) generating the same Young measure μ. On the one hand, the Fundamental Theorem for Young measures (Theorem 2.10) and the fact that $\{v_h\}$ generates μ yield

$$\liminf_{h\to\infty} \int_\omega F(v_h)\,\mathrm{d}x \geq \int_\omega \langle \mu_x, F\rangle\,\mathrm{d}x.$$

Even more, due to the same theorem and the equi-integrability of the sequence $\{|v_h'|^p\}$ one gets the convergence $F(v_h') \rightharpoonup \langle \mu_x, F\rangle \in \mathrm{L}^1$. In other words,

$$\lim_{h\to\infty} \int_\omega F(v_h')\,\mathrm{d}x = \int_\omega \langle \mu_x, F\rangle\,\mathrm{d}x.$$

The three relations above yield

$$\int_\omega Q_{\mathscr{B}}F(v)\,\mathrm{d}x \geq \limsup_{h\to\infty} \int_\omega F(v_h) \geq \int_\omega \langle \mu_x, F\rangle\,\mathrm{d}x = \lim_{h\to\infty} \int_\omega F(v_h')\,\mathrm{d}x \geq \int_\omega Q_{\mathscr{B}}F(v)\,\mathrm{d}x. \tag{37}$$

We summon the characterization for $\mathscr{B}$-free Young measures from Theorem 2.12 to observe that

$$\langle \mu_x, F\rangle \geq Q_{\mathscr{B}}F(\langle \mu_x, \mathrm{id}\rangle) = Q_{\mathscr{B}}F(v(x)) \quad \text{a.e. } x \in \omega.$$

This inequality and (37) imply

$$\langle \mu_x, F\rangle = Q_{\mathscr{B}}F(v(x)) \quad \text{a.e. } x \in \omega.$$

We conclude by recalling that $F(v_h') \rightharpoonup \langle \mu_x, F\rangle$ in $\mathrm{L}^1(\omega)$. □

The lower bound. Let $v \in \mathrm{L}^2(B_1; Z)$ and let $\{v_\varepsilon\}$ be a sequence in $\mathrm{L}^2(B_1; Z)$ such that $v_\varepsilon \rightharpoonup v$ in $\mathrm{L}^2(B_1; Z)$. We want to prove that

$$\liminf_{\varepsilon\downarrow 0} G^\varepsilon(v_\varepsilon) \geq G(v).$$

Notice that, we may reduce the proof to the case where $\mathscr{B}v_\varepsilon = 0$ for every ε. From the inequality $\sigma_A \geq W \geq Q_{\mathscr{B}}W$ (as quadratic forms), we infer that

$$J^\varepsilon(v_\varepsilon) \geq \int_{B_1} Q_{\mathscr{B}}W(v_\varepsilon)\,\mathrm{d}y.$$

Next, we recall that $v \mapsto \int_{B_1} Q_{\mathscr{B}}W(v)$ is lower semi-continuous in $\{v \in \mathrm{L}^2(\Omega; Z) : \mathscr{B}v = 0\}$ with respect to the weak-L^2 topology. Hence,

$$\liminf_{\varepsilon\downarrow 0} G^\varepsilon(v_\varepsilon) \geq \int_{B_1} Q_{\mathscr{B}}W(v)\,\mathrm{d}y.$$

This proves the lower bound inequality.

The upper bound. We fix $v \in \mathrm{L}^2(B_1; Z)$, we want to show that there exists a sequence $\{v_\varepsilon\}$ in $\mathrm{L}^2(B_1; Z)$ with $v_\varepsilon \rightharpoonup v$ in $\mathrm{L}^2(B_1; Z)$ and such that

$$\limsup_{\varepsilon\downarrow 0} G^\varepsilon(v_\varepsilon) \leq G(v).$$

We may assume that $\mathscr{B}v = 0$, for otherwise the inequality occurs trivially. Lemma 4.3 guarantees the existence of a 2-equi-integrable sequence $\{v_h\}_{h=1}^\infty$ for which

$$\mathscr{B}v_h = 0, \quad v_h \rightharpoonup v \quad \text{in } \mathrm{L}^2(B_1; Z), \quad \text{and} \quad W(v_h) \rightharpoonup Q_{\mathscr{B}}W(v) \quad \text{in } \mathrm{L}^1(B_1). \tag{38}$$

Next, we define an h-parametrized sequence of subsets of B_1 in the following way:

$$A_h := \{x \in B_1 : (\sigma_1 - \sigma_2)v_h \cdot v_h \le 0\}.$$

Using the fact that smooth sets are dense in the broader class of subsets with respect to measure convergence, we may take a smooth set $A'_h \subset B_1$ such that the following estimates hold for some strictly monotone function $L : \mathbb{N} \to \mathbb{N}$ (with $\lim_{h\to\infty} L(h) = \infty$):

$$|(A'_h \Delta A_h) \cap B_1| = O(h^{-1}), \quad \mathrm{Per}(A'_h; B_1) \le L(h). \tag{39}$$

Observe that, by the 2-equi-integrability of $\{v_h\}$, one gets that

$$\|(\sigma_{A_h} - \sigma_{A'_h})v_h \cdot v_h\|_{\mathrm{L}^2(B_1)} \le M\|v_h\|^2_{\mathrm{L}^2(S_h)} = O(h^{-1}), \quad \text{where } S_h := A'_h \Delta A_h. \tag{40}$$

The next step relies, essentially, on stretching the sequence $\{v_h\}$. Define the ε-sequence

$$\overline{v}_\varepsilon := v_{K(\varepsilon)}, \quad \varepsilon \le \frac{1}{L(1)},$$

where $K : \mathbb{R}_+ \to \mathbb{N}$ is the piecewise constant decreasing function defined as

$$K := \sum_{h=1}^{\infty} h \cdot \mathbb{1}_{R_h}, \quad R_h := \left(\frac{1}{L(h+1)}, \frac{1}{L(h)} \right].$$

Claim 1. *$L \circ K(\varepsilon) \le \varepsilon^{-1}$, if $\varepsilon \in (0, L(1)^{-1}]$.*
2. *$K(\varepsilon) = h$, where h is such that $\varepsilon \in R_h$.*

Proof To prove 1, observe from the strict monotonicity of L that $\cup_{h=1}^{\infty} R_h = (0, L(1)^{-1}]$. A simple calculation gives

$$L(K(\varepsilon)) = L\left(\sum_{h=1}^{\infty} h \cdot \mathbb{1}_{R_h}(\varepsilon)\right) = \sum_{h=1}^{\infty} L(h) \cdot \mathbb{1}_{R_h}(\varepsilon) = L(h_0) \cdot \mathbb{1}_{R_{h_0}}(\varepsilon) \le \frac{1}{\varepsilon}, \tag{41}$$

where h_0 is such that $\varepsilon \in R_{h_0}$. The proof of 2 is an easy consequence of the definition of K and the fact that $\{R_h\}$ is a disjoint family of sets. Indeed, if $\varepsilon \in R_h$ then $K(\varepsilon) = h \cdot \mathbb{1}_{R_h}(\varepsilon) = h$. □

Since K is a piecewise decreasing function and $K(\mathbb{R}_+) = \mathbb{N} \cup \{0\}$, it remains true that

$$\overline{v}_{K(\varepsilon)} \rightharpoonup v \text{ in } \mathrm{L}^2(B_1; \mathbb{R}^{dN^k}), \quad \text{as } \varepsilon \to 0.$$

We are now in position to calculate the lim sup inequality:

$$\begin{aligned} G^\varepsilon(v_{K(\varepsilon)}) &= \min_{A \in \mathfrak{B}(B_1)} \int_{B_1} \sigma_A v_{K(\varepsilon)} \cdot v_{K(\varepsilon)} + \varepsilon^2 \mathrm{Per}(A; B_1) \\ &\le \int_{B_1} \sigma_{A'_{K(\varepsilon)}} v_{K(\varepsilon)} \cdot v_{K(\varepsilon)} + \varepsilon^2 \mathrm{Per}(A'_{K(\varepsilon)}; B_1) \\ &\le \int_{B_1} \sigma_{A_{K(\varepsilon)}} v_{K(\varepsilon)} \cdot v_{K(\varepsilon)} + O(K(\varepsilon)^{-1}) + \varepsilon^2 L(K(\varepsilon)) \\ &\le \int_{B_1} W(v_{K(\varepsilon)}) + O(\varepsilon) + \varepsilon. \end{aligned}$$

Hence, by (38)

$$\limsup_{\varepsilon\downarrow 0} G^\varepsilon(\overline{v}_\varepsilon) \le \limsup_{\varepsilon\downarrow 0} \int_{B_1} W(v_{K(\varepsilon)}) = \lim_{h\to\infty} \int_{B_1} W(v_h) = \int_{B_1} Q_{\mathscr{B}} W(v).$$

This proves the upper bound inequality. □

Corollary 4.4 *Let $\{w_\varepsilon\} \subset \mathrm{W}^{\mathscr{A}}(B_1)$ be a sequence of almost local minimizers of the sequence of functionals*

$$\{u \mapsto G^\varepsilon(\mathscr{A}u)\}.$$

Assume that $\{\mathscr{A}w_\varepsilon\}$ is 2-equi-integrable in B_s for every $s < 1$. Assume also that there exists $w \in \mathrm{W}^{\mathscr{A}}(B_1)$ such that

$$\mathscr{A}w_\varepsilon \rightharpoonup \mathscr{A}w \quad \text{in } \mathrm{L}^2(B_1; \mathbb{R}^{dN^k}).$$

Then,

$$Q_{\mathscr{B}}W(\mathscr{A}w_\varepsilon) \rightharpoonup Q_{\mathscr{B}}(\mathscr{A}w) \quad \text{in } \mathrm{L}^1_{\mathrm{loc}}(B_1).$$

Moreover, w is a local minimizer of $u \mapsto G(\mathscr{A}u)$.

Proof The first step is to check that

$$Q_{\mathscr{B}}W(\mathscr{A}w_\varepsilon) \rightharpoonup Q_{\mathscr{B}}(\mathscr{A}w) \quad \text{in } \mathrm{L}^1(B_s), \text{ for every } s < 1. \tag{42}$$

The sequence $\mathscr{A}w_\varepsilon$ generates (up to taking a subsequence) a $\mathscr{B}$-free Young measure $\mu : B_1 \to \mathscr{M}(Z)$ so that by Theorem 2.10, Theorem 2.12 and the local 2-equi-integrability assumption,

$$W(\mathscr{A}w'_\varepsilon) \rightharpoonup \langle \mu_x, W\rangle \geq Q_{\mathscr{B}}W(\mathscr{A}w) \quad \text{in } \mathrm{L}^1_{\mathrm{loc}}(B_1). \tag{43}$$

Fix $s \in (0, 1)$ and consider the rescaled functions

$$w^s_\varepsilon := \frac{w_\varepsilon(sy)}{s^{k-\frac{1}{2}}}, \quad w^s := \frac{w(sy)}{s^{k-\frac{1}{2}}}.$$

It is not hard to see that, because of the (almost) minimization properties of $\{w_\varepsilon\}$, the rescaled sequence $\{w^s_\varepsilon\}$ is also a sequence of almost local minimizers of the sequence of functionals $\{u \mapsto G(\mathscr{A}u)\}$.[8] Moreover, $\mathscr{A}w^s_\varepsilon \rightharpoonup \mathscr{A}w^s$ in $\mathrm{L}^2(B_1; Z)$.

From the proof of the lower bound in Theorem 4.2, we may find a 2-equi-integrable recovery sequence $\{v'_\varepsilon\}$ for v, i.e., such that $v'_\varepsilon \rightharpoonup \mathscr{A}w^s$ and

$$\lim_{\varepsilon\downarrow 0} G^\varepsilon(v'_\varepsilon) = G(\mathscr{A}w^s).$$

Recall that, by the exactness assumption of $\mathscr{A}$ and $\mathscr{B}$, there are functions $w'_\varepsilon \in \mathrm{W}^{\mathscr{A}}(B_1)$ such that

$$v'_\varepsilon = \mathscr{A}w'_\varepsilon \quad \text{for every } \varepsilon > 0.$$

A recovery sequence with the same boundary values. The next step is to show that one may assume, without loss of generality, that $\mathrm{spt}(w'_\varepsilon - w^s_\varepsilon) \subset\subset B_1$.

We may further assume (without loss of generality) that $\{w^s_\varepsilon\}$ and $\{w'_\varepsilon\}$ are $\mathrm{W}^{k,2}$-uniformly bounded, and that $w^s_\varepsilon - w'_\varepsilon \rightharpoonup 0$ in $\mathrm{W}^{k,2}(B_1; \mathbb{R}^d)$.

Define

$$\tilde{v}_{h,\varepsilon} := \mathscr{A}(\varphi_h w'_\varepsilon + (1-\varphi_h)w^s_\varepsilon) = \varphi_h\mathscr{A}w'_\varepsilon + (1-\varphi_h)\mathscr{A}w^s_\varepsilon + \overbrace{\sum_{\substack{|\beta|\geq 1\\ |\alpha|+|\beta|=k}} c_{\alpha\beta}\partial^\alpha(w'_\varepsilon - w^s_\varepsilon)\partial^\beta\varphi_h}^{g(h)};$$

[8] This scaling has the property that $s^{N-1}J_{B_1}(\mathscr{A}w^s, A^s) = J_{B_s}(\mathscr{A}w, A)$.

where, for every $h \in \mathbb{N}$, $\varphi_h \in C^\infty(B_1; [0,1])$ with $\varphi_h \equiv 1$ in $B_{1-1/h}$. Since $\|g(h)\|_{L^2(B_1)} \to 0$ as $\varepsilon \to 0$, we infer that

$$\limsup_{\varepsilon\downarrow 0} \|\tilde{v}_{h,\varepsilon} - \mathscr{A} w'_\varepsilon\|_{L^2(B_1)} \le \limsup_{\varepsilon\downarrow 0} \|\mathscr{A} w'_\varepsilon\|_{L^2(B_1\setminus B_{1-1/h})} + \limsup_{\varepsilon\downarrow 0} \|\mathscr{A} w_\varepsilon\|_{L^2(B_1\setminus B_{1-1/h})}.$$

We now let $h \to \infty$ and use the 2-equi-integrability of $\{\mathscr{A} w^s_\varepsilon\}$ and $\{\mathscr{A} w'_\varepsilon\}$ to get

$$\limsup_{h\to\infty} \limsup_{\varepsilon\downarrow 0} \|\tilde{v}_{h,\varepsilon} - \mathscr{A} w'_\varepsilon\|_{L^2(B_1)} = 0.$$

Thus, we may find a diagonal sequence $\tilde{v}_\varepsilon = \tilde{v}_{h(\varepsilon),\varepsilon} = \mathscr{A}\tilde{w}^s_\varepsilon$ which is 2-equi-integrable, $\operatorname{spt}(w^s_\varepsilon - \tilde{w}_\varepsilon) \subset\subset B_1$, and such that

$$\lim_{\varepsilon\downarrow 0} \|\mathscr{A} w'_\varepsilon - \mathscr{A}\tilde{w}_\varepsilon\|_{L^2(B_s)} = O(\varepsilon).$$

In particular, the (almost) local minimizing property of $\{\mathscr{A} w^s_\varepsilon\}$ gives

$$\begin{aligned}\limsup_{\varepsilon\downarrow 0} \int_{B_1} W(\mathscr{A} w^s_\varepsilon) &\le \limsup_{\varepsilon\downarrow 0} G^\varepsilon(\mathscr{A} w^s_\varepsilon) \le \limsup_{\varepsilon\downarrow 0} G^\varepsilon(\mathscr{A}\tilde{w}_\varepsilon) \le \lim_{\varepsilon\downarrow 0} G^\varepsilon(\mathscr{A} w'_\varepsilon)\\ &= G(\mathscr{A} w^s).\end{aligned}$$

Rescaling back, the inequality above yields

$$\limsup_{\varepsilon\downarrow 0} \int_{B_s} W(\mathscr{A} w_\varepsilon) \le \int_{B_s} Q_{\mathscr{B}} W(\mathscr{A} w),$$

which together with (43) proves (42).

Local minimizer of G. The second step is to show that w is a local minimizer of $u \mapsto G(\mathscr{A} u)$. We argue by contradiction: assume that w is not a local minimizer of $u \mapsto G(\mathscr{A} u)$, then we would find $s \in (0,1)$ and $\eta \in C^\infty_c(B_s; \mathbb{R}^{dN^k})$ for which

$$G(\mathscr{A} w) > G(\mathscr{A} w + \mathscr{A}\eta).$$

Again, using a re-scaling argument, this would imply that

$$G(\mathscr{A} w^s) > G(\mathscr{A} w^s + \mathscr{A}\eta^s).$$

Similarly to the previous step, we can find a 2-equi-integrable recovery sequence $\{\mathscr{A}(\phi^s_\varepsilon + \eta^s)\}$ of $(\mathscr{A} w^s + \mathscr{A}\eta^s)$ with the property that $\operatorname{spt}(\phi^s_\varepsilon - w^s_\varepsilon) \subset\subset B_1$, for every $\varepsilon > 0$. On the other hand, the (almost) minimizing property of $\mathscr{A} w^s_\varepsilon$ and (42) yield

$$G(\mathscr{A} w^s + \mathscr{A}\eta^s) < G(\mathscr{A} w^s) = \lim_{\varepsilon\downarrow 0} G^\varepsilon(\mathscr{A} w^s_\varepsilon) \le \lim_{\varepsilon\downarrow 0} G^\varepsilon(\mathscr{A}\phi^s_\varepsilon + \mathscr{A}\eta^s) = G(\mathscr{A} w^s + \mathscr{A}\eta^s),$$

which is a contradiction. This shows that w is a local minimizer of $u \mapsto G(\mathscr{A} u)$. □

Let us recall, for the proof of the next proposition, that the higher integrability assumption (Reg) on local minimizers $\tilde{u}$ of $u \mapsto G(\mathscr{A} u)$ reads:

$$[\mathscr{A}\tilde{u}]^2_{L^{2,N-\delta}(B_{1/2})} \le c\|\mathscr{A}\tilde{u}\|^2_{L^2}(B_1), \quad \text{for some } \delta \in [0,1). \tag{Reg}$$

Proposition 4.5 *Let* (w, A) *be a saddle-point of problem* (P). *Assume that the higher integrability condition* (Reg) *holds for local minimizers of* $u \mapsto G(\mathscr{A} u)$. *Then, for every* $K \subset\subset \Omega$ *there exists a positive constant* $C(K) > 1$ *and a smallness constant* $\rho \in (0, 1/2)$ *such that at least one of the following properties*

1. $J_{B_r(x)}(\mathscr{A}w, A) \leq C(K)r^{N-1}$,
2. $J_{B_{\rho r}(x)}(\mathscr{A}w, A) \leq \rho^{N-(1+\delta)/2} J_{B_r(x)}(\mathscr{A}w, A)$,

holds for all $x \in K$ *and every* $r \in (0, \operatorname{dist}(K, \partial\Omega))$. Here,

$$J_{B_r(x)}(\mathscr{A}u, A) = \int_{B_r(x)} \sigma_A \mathscr{A}u \cdot \mathscr{A}u \ \mathrm{d}y \ + \ \operatorname{Per}(A; B_r(x)).$$

Proof Let (w, A) be a saddle-point of (P) and fix $\rho \in (0, 1)$ (to be specified later in the proof). We argue by contradiction through a blow-up technique: Negation of the statement would allow us to find a sequence $\{(x_h, r_h)\}$ of points $x_h \in K$ and positive radii $r_h \downarrow 0$ for which

$$J_{B_{r_h}(x_h)}(\mathscr{A}w, A) > hr_h^{N-1}, \quad \text{and} \tag{44}$$

$$J_{B_{\rho r_h}(x_h)}(\mathscr{A}w, A) > \rho^{N-(1+\delta)/2} J_{B_{r_h}(x_h)}(\mathscr{A}w, A). \tag{45}$$

An equivalent variational problem. It will be convenient to work with a similar variational problem: Consider the saddle-point problem

$$\inf \left\{ \sup_{u \in \mathrm{W}_0^{\mathscr{A}}(\Omega)} \tilde{I}_\Omega(\mathscr{A}u, A) : A \subset \mathbb{R}^N \text{ Borel set}, \ A \cap \Omega^c \equiv A_0 \cap \Omega^c \right\}, \tag{$\tilde{\mathrm{P}}$}$$

where

$$\tilde{I}_\Omega(\mathscr{A}u, A) := \int_\Omega 2\tau_A \cdot \mathscr{A}u \, \mathrm{d}x - \int_\Omega \sigma_A \mathscr{A}u \cdot \mathscr{A}u \, \mathrm{d}x + \operatorname{Per}(A; \overline{\Omega}).$$

Here we recall the notation $\tau_A := \sigma_A \mathscr{A} w_A$, where $w_A \in \mathrm{W}_0^{\mathscr{A}}(\Omega)$ is the unique maximizer of $u \mapsto I_\Omega(u, A)$. It follows immediately from the identity

$$\int_\Omega \tau_A \cdot \mathscr{A}u \, \mathrm{d}x = \int_\Omega fu \, \mathrm{d}x \quad u \in \mathrm{W}_0^{\mathscr{A}}(\Omega),$$

that saddle-points (w, A) of problem (P) are also saddle-points of ($\tilde{\mathrm{P}}$) and vice versa; hence, in the following we will make no distinction between saddle-points of (P) and ($\tilde{\mathrm{P}}$). A special property of $\tilde{I}$ is that, its density is always positive on saddle-points (w, A) of (P). Indeed, in this case $w = w_A$ and therefore

$$\tilde{I}_{B_r(x)}(\mathscr{A}w, A) = \int_{B_r(x)} \sigma_A \mathscr{A}w_A \cdot \mathscr{A}w_A + \operatorname{Per}(A; B_r(x)) = J_{B_r(x)}(\mathscr{A}w, A), \quad B_r(x) \subset \Omega. \tag{46}$$

A re-scaling argument. We re-scale and translate $B_r(x)$ into B_1 by letting

$$A^{r,x} := \frac{A}{r} - x, \quad f^{r,x}(y) := r^{k+\frac{1}{2}} f(ry+x) \to 0 \text{ in } \mathrm{L}^\infty(B_1), \quad \text{and} \quad w^{r,x}(y) := \frac{w(ry+x)}{r^{k-\frac{1}{2}}}. \tag{47}$$

A further normalization on the sequence takes place by setting

$$\varepsilon(h)^2 := r_h^{N-1} \cdot J_{B_{r_h}(x_h)}(\mathscr{A}w, A)^{-1} = O(h^{-1}),$$

and defining

$$A_{\varepsilon(h)} := A^{r_h, x_h}, \quad f_{\varepsilon(h)} := \varepsilon(h) \cdot f^{r_h, x_h}, \quad w_{\varepsilon(h)} := \varepsilon(h) \cdot w^{r_h, x_h},$$
$$\text{and} \quad \tau_{\varepsilon(h)} := \sigma_{A_{\varepsilon(h)}} \mathscr{A} w_{\varepsilon(h)}.$$

It is easy to check that the scaling rule (33), and the relations (45) and (46) imply

$$J^{\varepsilon(h)}(\mathscr{A} w_{\varepsilon(h)}, A_{\varepsilon(h)}) = 1, \quad \text{and} \tag{48}$$

$$\int_{B_\rho} \sigma_{A_{\varepsilon(h)}} \mathscr{A} w_{\varepsilon(h)} \cdot \mathscr{A} w_{\varepsilon(h)} + \varepsilon(h)^2 \mathrm{Per}(A_{\varepsilon(h)}; B_\rho) > \rho^{N-(1+\delta)/2}. \tag{49}$$

In particular, due to the coercivity of σ_1 and σ_2, the norms $\|\mathscr{A} w_{\varepsilon(h)}\|^2_{\mathrm{L}^2(B_1)}$ are h-uniformly bounded by M.

Local almost-minimizers of $G^{\varepsilon(h)}$. The next step is to show that $\{w_{\varepsilon(h)}\}$ is $O(\varepsilon)$-close in L^2 to a sequence $\{\tilde{w}_\varepsilon\}$ of almost minimizers of $\{u \mapsto G^{\varepsilon(h)}(\mathscr{A} u)\}$. Observe that $w_{\varepsilon(h)}$ is the unique solution to the equation

$$\mathscr{A}^*(\sigma_{A_\varepsilon} \mathscr{A} u) = f_{\varepsilon(h)}, \quad u \in \mathrm{W}^{\mathscr{A}}_{w_{\varepsilon(h)}}(B_1).$$

Let $\tilde{w}_{\varepsilon(h)}$ be the unique minimizer of $u \mapsto J^{\varepsilon(h)}(\mathscr{A} u, A_{\varepsilon(h)})$ – see (32) – in the affine space $W^{\mathscr{A}}_{w_{\varepsilon(h)}}(B_1)$. Thus, in particular, $\tilde{w}_{\varepsilon(h)}$ is the unique solution of the equation

$$\mathscr{A}^*(\sigma_{A_{\varepsilon(h)}} \mathscr{A} u) = 0, \quad u \in \mathrm{W}^{\mathscr{A}}_{w_{\varepsilon(h)}}(B_1).$$

A simple integration by parts, considering that $\tilde{w}_{\varepsilon(h)} - w_{\varepsilon(h)} \in \mathrm{W}^{\mathscr{A}}_0(B_1)$, gives the estimate

$$\|\mathscr{A} w_{\varepsilon(h)} - \mathscr{A} \tilde{w}_{\varepsilon(h)}\|^2_{\mathrm{L}^2(B_1)} \le C(B_1) \cdot M^2 \|f_{\varepsilon(h)}\|^2_{\mathrm{L}^2(B_1)} = O(h^{-1}), \tag{50}$$

where $C(B_1)$ is the Poincaré constant from (12); and therefore $\|w_{\varepsilon(h)} - \tilde{w}_{\varepsilon(h)}\|_{\mathrm{W}^{k,2}_0(B_1)} = O(h^{-1})$.

Lastly, we use strongly the fact that (w, A) is a saddle-point of (P) to see that $\{(w_{\varepsilon(h)}, A_{\varepsilon(h)})\}$ is also a *local* saddle-point of the energy

$$(u, E) \mapsto \tilde{I}^{\varepsilon(h)}(\mathscr{A} u, E) := \int_{B_1} 2\tau_E \cdot \mathscr{A} u \,\mathrm{d}y - \int_{B_1} \sigma_E \mathscr{A} u \cdot \mathscr{A} u \,\mathrm{d}y \ + \ \varepsilon(h)^2 \mathrm{Per}(E; B_1).$$

Moreover, by (33), (46) and (50) one has that

$$\tilde{I}^{\varepsilon(h)}(\mathscr{A} w_{\varepsilon(h)}, A_{\varepsilon(h)}) = J^{\varepsilon(h)}(\mathscr{A} w_{\varepsilon(h)}, A_{\varepsilon(h)}) = J^{\varepsilon(h)}(\mathscr{A} \tilde{w}_{\varepsilon(h)}, A_{\varepsilon(h)}) + O(h^{-1}). \tag{51}$$

An immediate consequence of the two facts above is that $\{\tilde{w}_{\varepsilon(h)}\}$ is a sequence of *local* almost minimizers of the sequence of functionals $\{u \mapsto G^{\varepsilon(h)}(\mathscr{A} u)\}$. The local (almost) minimizing properties of the sequence $\{\tilde{w}_{\varepsilon(h)}\}$ – with respect to the functionals $\{u \mapsto G^{\varepsilon(h)}(\mathscr{A} u)\}$ – are not affected by subtracting $\mathscr{A}$-free fields; hence, using the compactness assumption of $\mathscr{A}$ once more, we may assume without loss of generality that $\sup_h \|\tilde{w}_{\varepsilon(h)}\|_{\mathrm{W}^{k,2}(B_1)} < \infty$. Upon passing to a further subsequence, we may also assume that there exists $\tilde{w} \in \mathrm{W}^{k,2}(B_1)$ such that

$$\tilde{w}_{\varepsilon(h)} \rightharpoonup \tilde{w} \quad \text{in } \mathrm{W}^{k,2}(B_1; \mathbb{R}^d).$$

Equi-integrability of $\{\mathscr{A} \tilde{w}_{\varepsilon(h)}\}$. The last but one step is to show that $\{\mathscr{A} \tilde{w}_\varepsilon\}$ is a 2-equi-integrable sequence in B_s, for every $s < 1$.

Since σ_{A_ε} is uniformly bounded, there exists $\tilde{\tau} \in \mathrm{L}^2(B_1; \mathbb{R}^{dN^k})$ such that (upon passing to a further subsequence)

$$\sigma_{A_{\varepsilon(h)}} \mathscr{A} \tilde{w}_{\varepsilon(h)} =: \tilde{\tau}_{\varepsilon(h)} \rightharpoonup \tilde{\tau} \quad \text{in } \mathrm{L}^2(B_1; \mathbb{R}^{dN^k}), \quad \mathscr{A}^* \tilde{\tau}_{\varepsilon(h)} = \mathscr{A}^* \tilde{\tau} = 0. \tag{52}$$

Let $\varphi \in \mathscr{D}(B_1)$ and fix $\varepsilon > 0$, integration by parts yields

$$\langle \tilde{\tau}_{\varepsilon(h)} \cdot \mathscr{A}\tilde{w}_{\varepsilon(h)}, \varphi \rangle = - \sum_{\substack{|\beta|\geq 1 \\ |\alpha|+|\beta|=k}} c_{\alpha\beta} \langle \tilde{\tau}_{\varepsilon(h)}, \partial^\alpha \tilde{w}_{\varepsilon(h)} \partial^\beta \varphi \rangle \qquad c_{\alpha,\beta} \in \mathbb{R}.$$

Since the term in the right hand side of the equality depends only on $\nabla^{k-1}\tilde{w}_{\varepsilon(h)}$, the strong convergence $\tilde{w}_\varepsilon \to \tilde{w}$ in $\mathrm{W}^{k-1,2}(B_1; \mathbb{R}^d)$ gives

$$\lim_{\varepsilon \to 0} \langle \tilde{\tau}_{\varepsilon(h)} \cdot \mathscr{A}\tilde{w}_{\varepsilon(h)}, \varphi \rangle = - \sum_{\substack{|\beta|\geq 1 \\ |\alpha|+|\beta|=k}} c_{\alpha\beta} \langle \tilde{\tau}, \partial^\alpha \tilde{w} \partial^\beta \varphi \rangle = \langle \tilde{\tau} \cdot \mathscr{A}\tilde{w}, \varphi \rangle.$$

Therefore,

$$\sigma_{A_{\varepsilon(h)}} \mathscr{A}\tilde{w}_{\varepsilon(h)} \cdot \mathscr{A}\tilde{w}_{\varepsilon(h)} = \tilde{\tau}_{\varepsilon(h)} \cdot \mathscr{A}\tilde{w}_{\varepsilon(h)} \overset{*}{\rightharpoonup} \tilde{\tau} \cdot \mathscr{A}\tilde{w} \in \mathrm{L}^1(B_1) \quad \text{weakly* in } \mathscr{M}^+(B_1).$$

The positivity of $\sigma_{A_\varepsilon} \mathscr{A}\tilde{w}_\varepsilon \cdot \mathscr{A}\tilde{w}_\varepsilon$, the Dunford-Pettis Theorem and the convergence above imply that the sequence

$$\{\sigma_{A_\varepsilon} \mathscr{A}\tilde{w}_\varepsilon \cdot \mathscr{A}\tilde{w}_\varepsilon\} \quad \text{is equi-integrable in } B_s; \ \text{for every } s < 1.$$

In turn, due to the uniform coerciveness and boundedness of $\{\sigma_{A_\varepsilon}\}$, both sequences $\{\mathscr{A}\tilde{w}_\varepsilon\}$ and $\{\tilde{\tau}_\varepsilon\}$ are 2-equi-integrable in B_s; for every $s < 1$.

The contradiction. We are in position to apply Proposition 4.4 to the sequence $\{\tilde{w}_\varepsilon\}$, which in particular implies

$$\begin{aligned} &\varepsilon(h)^2 \mathrm{Per}(A_{\varepsilon(h)}; B_\rho) \to 0, \\ &\sigma_{A_{\varepsilon(h)}} \mathscr{A}\tilde{w}_{\varepsilon(h)} \cdot \mathscr{A}\tilde{w}_{\varepsilon(h)} \rightharpoonup Q_{\mathscr{B}} W(\mathscr{A}\tilde{w}) \leq M|\mathscr{A}\tilde{w}|^2 \quad \text{in } \mathrm{L}^1_{\mathrm{loc}}(B_1), \end{aligned} \tag{53}$$

and that w is a local minimizer of $u \mapsto G(\mathscr{A}u)$. On the other hand, the higher integrability assumption (Reg) tells us that

$$[\mathscr{A}\tilde{w}]^2_{\mathrm{L}^{2,N-\delta}(B_{1/2})} \leq c\|\mathscr{A}\tilde{w}\|^2_{\mathrm{L}^2(B_1)}. \tag{54}$$

We set the value of $\rho \in (0, 1/2)$ to be such that $2cM^2\rho^{(1-\delta)/2} \leq 1$. Taking the limit in (48) and (49), using Fatou's Lemma, (50), (51), (53) and (54), we get

$$\begin{aligned} \frac{1}{M}\|\mathscr{A}\tilde{w}\|^2_{\mathrm{L}^2(B_1)} &\leq \lim_{h\to\infty} J^{\varepsilon(h)}(\mathscr{A}\tilde{w}_{\varepsilon(h)}, A_{\varepsilon(h)}) = 1 \\ &\leq \left(\frac{1}{\rho^{N-(1+\delta)/2}}\right)\|Q_{\mathscr{B}}W(\mathscr{A}\tilde{w})\|_{\mathrm{L}^1(B_\rho)} \leq \left(\frac{M\rho^{(1-\delta)/2}}{\rho^{N-\delta}}\right)\|\mathscr{A}\tilde{w}\|^2_{\mathrm{L}^2(B_\rho)} \\ &\leq M\rho^{(1-\delta)/2}[\mathscr{A}\tilde{w}]^2_{\mathrm{L}^{2,N-\delta}(B_{1/2})} \leq cM\rho^{(1-\delta)/2}\|\mathscr{A}\tilde{w}\|^2_{\mathrm{L}^2(B_1)} \\ &\leq \frac{1}{2M}\|\mathscr{A}\tilde{w}\|^2_{\mathrm{L}^2(B_1)}; \end{aligned}$$

a contradiction. □

Theorem 1.2 (Upper bound) *Let (w, A) be a variational solution of problem* (P)*. Assume that the higher integrability condition*

$$[\mathscr{A}\tilde{u}]^2_{\mathrm{L}^{2,N-\delta}(B_{1/2})} \leq c\|\mathscr{A}\tilde{u}\|^2_{\mathrm{L}^2(B_1)}, \quad \text{for some } \delta \in [0,1) \text{ and some positive constant } c,$$

holds for local minimizers of the energy $u \mapsto \int_{B_1} Q_{\mathscr{B}} W(\mathscr{A}u)$, *where* $u \in \mathrm{W}^{\mathscr{A}}(B_1)$. *Then, for every compactly contained set* $K \subset\subset \Omega$, *there exists a positive constant* Λ_K *such that*

$$\int_{B_r(x)} \sigma_A \mathscr{A}w \cdot \mathscr{A}w \; \mathrm{d}y \; + \; \mathrm{Per}(A; B_r(x)) \leq \Lambda_K r^{N-1} \quad \forall\, x \in K, \forall\, r \in (0, \mathrm{dist}(K, \partial\Omega)). \tag{55}$$

Proof Let $x \in K$, and set

$$\varphi(r, x) := J_{B_r(x)}(\mathscr{A}w, A),$$

where we recall that

$$J_{B_r(x)}(\mathscr{A}w, A) = \int_{B_r(x)} \sigma_A \mathscr{A}w \cdot \mathscr{A}w \; \mathrm{d}y \; + \; \mathrm{Per}(A; B_r(x))$$

Proposition 4.5 tells us that there exists a positive constant $\rho \in (0, 1/2)$ such that if $B_r(x) \subset \Omega$, then

$$\varphi(\rho r, x) \leq \rho^{N-(1+\delta)/2} \varphi(r, x) + C(K) r^{N-1}.$$

An application of the Iteration Lemma (stated below) to $r \in (0, \min\{1, \mathrm{dist}(K, \partial\Omega\})$, and $\alpha_1 := N - (1+\delta)/2 > \alpha_2 := N - 1$ yields the existence of positive constants $c = c(x)$, and $r = r(K)$ such that

$$\varphi(s, x) \leq c s^{N-1} \quad \forall\, s \in (0, R(K)).$$

Notice that the constants c and r depend continuously on $x \in \Omega$. Hence, for any $K \subset\subset \Omega$ we may find $\Lambda_K > 0$ for which

$$J_{B_r(x)}(\mathscr{A}w, A) \leq \Lambda_K r^{N-1} \quad \forall\, x \in K, \;\; \forall\, r \in (0, \mathrm{dist}(K, \partial\Omega)).$$

□

Lemma 4.6 (Iteration Lemma [30, Lemma 2.1, Chapter III]) *Assume that* $\varphi(\rho)$ *is a non-negative, real-valued, non-decreasing function defined on the* (0, 1) *interval. Assume further that there exists a number* $\tau \in (0, 1)$ *such that for all* $r < 1$ *we have*

$$\varphi(\tau r) \leq \tau^{\alpha_1} \varphi(r) + C r^{\alpha_2}$$

for some non-negative constant C, *and positive exponents* $\alpha_1 > \alpha_2$. *Then there exists a positive constant* $c = c(\tau, \alpha_1, \alpha_2)$ *such that for all* $0 \leq \rho \leq r \leq R$ *we have*

$$\varphi(\rho) \leq c \left(\frac{\rho}{r}\right)^{\alpha_2} \varphi(r) + C \rho^{\alpha_2}.$$

Corollary 4.7 (Compactness of blow-up sequences) *Let* (w, A) *be a variational solution of problem* (P). *Under the assumptions of the Upper bound Theorem* 1.2, *there exists a positive constant* C_K *such that*

$$[\mathscr{A}w]^2_{\mathrm{L}^{2,N-1}(K)} \leq C_K. \tag{56}$$

Proof The assertion follows directly from the Upper bound Theorem and the coercivity of σ_1 and σ_2. □

5 The Lower bound: proof of estimate (LB)

During this section we will write (w, A) to denote a solution of problem (P) under the assumptions of Theorem 1.2. In light of the results obtained in the previous section we will assume, throughout the rest of the paper, that for every compact set $K \subset\subset \Omega$ there exist positive constants C_K, and Λ_K such that

$$\begin{aligned}\mathrm{Per}(A; B_r(x)) \leq \Lambda_K r^{N-1},\\ \|\mathscr{A} w^{x,r}\|^2_{\mathrm{L}^2(B_1)} \leq [\mathscr{A} w]^2_{\mathrm{L}^{2,N-1}(K)} \leq C_K,\end{aligned}$$

for all $x \in K$ and every $r \in (0, \mathrm{dist}(K, \partial\Omega))$. Here, $w^{x,r} := w(x + ry)/r^{k-\frac{1}{2}}$.

The main result of this section is a lower bound on the density of the perimeter in $\partial^* A$. In other words, there exists a positive constant $\lambda_K = \lambda_K(N, M)$ such that

$$\mathrm{Per}(A; B_r(x)) \geq \lambda_K r^{N-1} \quad \text{for every } 0 < r < \mathrm{dist}(x, \partial\Omega). \tag{LB}$$

There are two major consequences from estimate (LB). The first one (cf. Corollary 5.8) is that the difference between the topological boundary of A and the reduced boundary of A is at most a set of zero $\mathscr{H}^{N-1}$-measure. In other words, $(\partial A \backslash \partial^* A) = \Sigma$ where $\mathscr{H}^{N-1}(\Sigma) = 0$ (cf. [7, Theorem 2.2]). The second is that (LB) is a necessary assumption for the Height bound Lemma and the Lipschitz approximation Lemma, which are essential tools to prove the flatness excess improvement in the next section.

Throughout this section and the rest of the manuscript we will constantly use the following notations:

The scaled Dirichlet energy

$$D(w; x, r) := \frac{1}{r^{N-1}} \int_{B_r(x)} |\mathscr{A} w|^2 \, \mathrm{d}y,$$

and the excess for γ-weighted energy

$$E_\gamma(w, A; x, r) := D(w; x, r) + \frac{\gamma}{r^{N-1}} \mathrm{Per}(A, B_r(x)).$$

Granted that the spatial-, radius-, or (w, A)- dependence is clear, we will shorten the notations to the only relevant variables, e.g., $D(r)$ and $E_\gamma(r)$. Recall that, up to translation and re-scaling, we may assume

$$0 \in \partial^* A \cap K, \quad \text{and} \quad B_1 \subset K + B_9 \subset \Omega.$$

Bear also in mind that all the constants in this section are universal up to their dependence on Λ_K and C_K.

We will proceed as follows. First we prove in Lemma 5.1 that if the density of the perimeter is sufficiently small, one may regard the regularity properties of solutions as those ones for an elliptic equation with constant coefficients. Then, in Lemma 5.2, we prove a lower bound on the decay of the density of the perimeter in terms of D. Combining these results, we are able to show a discrete monotonicity formula on the decay of E_γ.

The proof of the Lower density bound (LB) follows easily from this discrete monotonicity formula, De Giorgi's Structure Theorem, and the Upper bound Theorem of the previous section. Finally, we prove that the difference between ∂A and $\partial^* A$ is $\mathscr{H}^{N-1}$-negligible (Theorem 5.8) as a corollary of the estimate (LB).

Lemma 5.1 (Approximative solutions of the constant coefficient problem) *For every* $\theta_1 \in (0, 1/2)$*, there exist positive constants*[9] $c_1(\theta_1, N, M)$ *and* $\varepsilon_1(\theta_1, N, M)$ *such that either*

$$\int_{B_\rho} |\mathscr{A} w|^2 \,\mathrm{d}y \le c_1 \rho^N \|f\|^2_{\mathrm{L}^\infty(B_1)},$$

or

$$\int_{B_\rho} |\mathscr{A} w|^2 \,\mathrm{d}y \le 2c\rho^N \int_{B_1} |\mathscr{A} w|^2 \,\mathrm{d}y \quad \textit{for every } \rho \in [\theta_1, 1),$$

where $c = c(N, M)$ *is the constant from Lemma 2.6; whenever*

$$\mathrm{Per}(A; B_1) \le \varepsilon_1.$$

Proof Since $c \ge 2^N$, the result holds if we assume $\rho \ge 1/2$, therefore we focus only on the case where $\rho \in (\theta_1, 1/2]$. Fix $\theta_1 \in (0, 1/2)$. We argue by contradiction: We would find a sequence of pairs (w_h, A_h) (locally solving (P) in B_1 for a source function f_h) and constants $\rho_h \in [\theta_1, 1/2]$, such that

$$\delta_h^2 := \int_{B_{\rho_h}} |\mathscr{A} w_h|^2 \,\mathrm{d}y > 2\, c\rho_h^N \int_{B_1} |\mathscr{A} w_h|^2 \,\mathrm{d}y, \tag{57}$$

and simultaneously

$$\rho_h^N \cdot \frac{\|f_h\|^2_{\mathrm{L}^\infty(B_1)}}{\delta_h^2} \le \frac{1}{h}, \quad \text{and} \quad \mathrm{Per}(A_h; B_1) \le \frac{1}{h}.$$

The estimate above yields $\delta_h^{-1} f_h \to 0$ in $\mathrm{L}^2(B_1; \mathbb{R}^d)$. Also, since $\mathrm{Per}(A_h; B_1) \to 0$, the isoperimetric inequality yields that either $\sigma_{A_h} \to \sigma_1$ or $\sigma_{A_h} \to \sigma_2$ in L^2 as h tends to infinity. Let us assume that the former convergence $\sigma_{A_h} \to \sigma_1$ holds.

Let $u_h := \delta_h^{-1} w_h$, and observe that

$$\sup_h \|\mathscr{A} u_h\|_{\mathrm{L}^2(B_1)} < \infty.$$

We use that w_h is a (local) solution to (P) for A_h as indicator set and f_h as source term, to see that

$$\mathscr{A}^*(\sigma_{A_h} \mathscr{A} u_h) = \delta_h^{-1} f_h \quad \text{in } B_1.$$

Up to modifying the sequence by $\mathscr{A}$-free fields and passing to a further subsequence, we may assume that $u_h \rightharpoonup u$ in $\mathrm{W}^{k,2}(B_1; \mathbb{R}^{dN^k})$. We may then apply the compensated compactness result from Lemma 2.9 to obtain that

$$\mathscr{A}^*(\sigma_1 \mathscr{A} u) = 0 \quad \text{in } B_1,$$

and

$$D(u_h; s) \to D(u; s) \quad \text{where } \rho_h \to s \in [\theta_1, 1/2].$$

Hence, by (57) and Fatou's Lemma one gets

$$2\, cs^N D(u; 1) \le \lim_{h\to\infty} c\rho_h^N D(u_h; 1) \le 1 = \lim_{h\to\infty} D(u_h; \rho_h) = \lim_{h\to\infty} D(u_h; s) = D(u; s).$$

This is a contradiction to Lemma 2.6 because u is a solution for the problem with constant coefficients σ_1. The case when $\sigma_{A_h} \to \sigma_2$ can be solved by similar arguments. □

[9] As it can be seen from the proof of Lemma 5.1, the constant c_1 does not depend on K.

The next lemma is the principal ingredient in proving the (LB) estimate. It relies on a cone-like comparison to show that the decay of the perimeter density is controlled by $D(r)/r$: The perimeter density cannot blow-up at smaller scales, while for a fixed scale, the perimeter density is small.

Lemma 5.2 (Universal comparison decay) *There exists a positive constant*[10] $c_2 = c_2(N, M)$ *such that*

$$\frac{\mathrm{d}}{\mathrm{d}r}\bigg|_{\rho=r} \left(\frac{\mathrm{Per}(A;\, B_\rho)}{\rho^{N-1}}\right) \geq -c_2 \frac{D(r)}{r} \quad \textit{for a.e. } r \in (0, 1].$$

Proof For a.e. $r \in (0, 1)$ the slice $\langle A, g, r\rangle$, where $g(x) = |x|$, is well defined (see Sect. 2.4). Fix one such r and let $\tilde{A}$ be the cone-like comparison set to A as in (23). By minimality of (w, A) and a duality argument, we get

$$\int_{B_r} \sigma_A^{-1} \tau_A \cdot \tau_A \,\mathrm{d}y + \mathrm{Per}(A;\, B_r) \leq \int_{B_r} \sigma_{\tilde{A}}^{-1} \tau_A \cdot \tau_A \,\mathrm{d}y + \mathrm{Per}(\tilde{A};\, B_r)$$

for $\tau_A = \sigma_A \mathscr{A} w$. Hence,

$$\begin{aligned} \mathrm{Per}(A;\, B_r) &\leq \mathrm{Per}(\tilde{A};\, B_r) + M^3 \int_{B_r} |\mathscr{A} w_A|^2 \,\mathrm{d}y \\ &\leq \frac{r}{N-1} \langle A, g, r\rangle(\mathbb{R}^N) + M^3 r^{N-1} D(r). \end{aligned} \tag{58}$$

To reach the inequality in the last row we have used that the cone extension $\tilde{A}$ is precisely built (cf. (24)) so that the Green-Gauss measures $\mu_{\tilde{A}}$ and μ_A agree in $(B_r)^c$; where, by (22),

$$\begin{aligned} \mathrm{Per}(\tilde{A};\, B_\rho) &= \frac{1}{(N-1)} \left(\frac{\rho^{N-1}}{r^{N-2}}\right) \mathscr{H}^{N-2}(\partial^* A \cap \{g = r\}) \\ &\leq \frac{1}{(N-1)} \left(\frac{\rho^{N-1}}{r^{N-2}}\right) \langle A, g, r\rangle(\mathbb{R}^N) \quad \forall\, 0 < \rho \leq r. \end{aligned}$$

We know from (25) that $\frac{\mathrm{d}}{\mathrm{d}\rho}\big|_r \mathrm{Per}(A;\, B_\rho) \geq \langle A, g, r\rangle(\mathbb{R}^N)$ for a.e. $r > 0$. Since (58) and the previous inequality are valid almost everywhere in $(0, 1)$, a combination of these arguments yields

$$\frac{\mathrm{d}}{\mathrm{d}r}\bigg|_{\rho=r} \left(\frac{\mathrm{Per}(A;\, B_\rho)}{\rho^{N-1}}\right) \geq -M^3(N-1) \frac{D(r)}{r} \quad \text{for a.e. } r \in (0, 1).$$

The result follows for $c_2 := M^3(N-1)$. □

The following result is a discrete monotonicity for the weighted excess energy E_γ. We remark that, in general, a monotonicity formula may not be expected in the case of systems.

Theorem 5.3 (Discrete monotonicity) *There exist positive constants* $\gamma = \gamma(N, M)$, $\varepsilon_2 = \varepsilon_2(\gamma, N) \leq \mathrm{vol}(B_1') \cdot \gamma/2$, *and* $\theta_2 = \theta_2(N, M) \in (0, 1/2)$ *such that*

$$E_\gamma(\theta_2) \leq E_\gamma(1) + c_1(\theta_2) \|f\|^2_{\mathrm{L}^\infty(B_1)}, \quad \textit{whenever} \quad E_\gamma(1) \leq \varepsilon_2. \tag{59}$$

[10] The constant c_2 is independent of the compact set K; indeed, this is the result of universal comparison estimates in Ω.

Proof We fix γ and θ_1 such that

$$\gamma c_2 \max\{c, c_1(\theta_1)\} \le \frac{1}{4}, \quad \text{where} \quad 2\theta_1 c \le \frac{1}{2}.$$

Set $\theta_2 := \theta_1$. Recall that c_2 is the constant from Lemma 5.2, and c is the constant of Lemma 2.6.

Let also $\varepsilon_2 = \varepsilon_2(\gamma, \varepsilon_1)$ be a positive constant with $\varepsilon_2 \le \min\{\gamma\varepsilon_1(\theta_2), \gamma \cdot \mathrm{vol}\,(B_1')/2\}$. This implies

$$\mathrm{Per}(A; B_1) \le \varepsilon_1(\theta_2),$$

which in turn gives, for $c_1 = c_1(\theta_2)$,

$$E_\gamma(\theta_2) \le \frac{\gamma}{\theta_2^{N-1}} \mathrm{Per}(A; B_{\theta_2}) + 2c\theta_2 D(1) + c_1\theta_2 \|f\|^2_{\mathrm{L}^\infty(B_1)}.$$

Now, we apply Lemmas 5.1 and 5.2 to $s \in (\theta_2, 1)$ to get

$$\begin{aligned}
E_\gamma(\theta_2) &\le \frac{\gamma}{\theta_2^{N-1}} \mathrm{Per}(A; B_{\theta_2}) + 2c\theta_2 D(1) + c_1\theta_2 \|f\|^2_{\mathrm{L}^\infty(B_1)} \\
&\le \gamma\, \mathrm{Per}(A; B_1) + \gamma \int_{\theta_2}^1 -\frac{\mathrm{d}}{\mathrm{d}r}\Big|_{r=s} \left(\frac{\mathrm{Per}(A, B_r)}{r^{N-1}} \right) \mathrm{d}s + \frac{1}{2} D(1) + c_1\theta_2 \|f\|^2_{\mathrm{L}^\infty(B_1)} \\
&\le \gamma\, \mathrm{Per}(A; B_1) + \gamma c_2 \int_{\theta_1}^1 \frac{D(s)}{s}\, \mathrm{d}s + \frac{1}{2} D(1) + c_1\theta_2 \|f\|^2_{\mathrm{L}^\infty(B_1)} \\
&\le \gamma\, \mathrm{Per}(A; B_1) + 2\gamma c c_2 D(1) + \gamma c_2 c_1 \|f\|^2_{\mathrm{L}^\infty(B_1)} + \frac{1}{2} D(1) + c_1\theta_2 \|f\|^2_{\mathrm{L}^\infty(B_1)} \\
&\le \gamma\, \mathrm{Per}(A; B_1) + D(1) + c_1 \|f\|^2_{\mathrm{L}^\infty(B_1)} \\
&= E_\gamma(1) + c_1 \|f\|^2_{\mathrm{L}^\infty(B_1)}.
\end{aligned}$$

This proves the desired result. □

Lemma 5.4 *For every $\varepsilon > 0$, there exist positive constants $\theta_0(N, M, K, \varepsilon) \in (0, 1/2)$ and $\kappa(N, M, K, \varepsilon) > 0$ such that*

$$E_\gamma(\theta_0) \le \varepsilon + c_1 \|f\|^2_{\mathrm{L}^\infty(B_1)};$$

whenever

$$\mathrm{Per}(A; B_1) \le \kappa.$$

Proof The result follows by taking θ_0 such that $2c\theta_0 C_K \le \varepsilon/2$ (recall that, $D(s) \le C_K$ for every $s \in (0, 1)$) and $\kappa \le \min\left\{\frac{\varepsilon\theta_0^{N-1}}{2\gamma}, \varepsilon_1(\theta_0)\right\}$ and then simply applying Lemma 5.1. □

Lemma 5.5 *Let (w, A) be a saddle-point of (P) and let $x \in K \subset\subset \Omega$. Then, for every $\varepsilon > 0$ there exists a positive radius $r_0 = r_0(N, M, K, \|f\|_{\mathrm{L}^\infty(B_1)}, \varepsilon)$ for which*

$$E_\gamma(w, A; x, r) \le 2\varepsilon;$$

whenever $r \le r_0$ and $\mathrm{Per}\big(A; B_{\theta_0^{-1} r}\big) \le \kappa(\varepsilon) \cdot \left(\frac{r}{\theta}\right)^{N-1}$.

Proof Let r_0 be a positive constant such that $c_1 r_0^{2k+1}\|f\|^2_{L^\infty(B_1)} \le \theta_0^{2k+1}\varepsilon$ and let us set $s := \theta_0^{-1} r$. Since

$$\mathrm{Per}(A^{x,s}; B_1) = s^{-(N-1)}\mathrm{Per}(A; B_s) \le \kappa(\varepsilon),$$

it follows from the previous lemma and a rescaling argument that

$$E_\gamma(w, A; r) = E_\gamma(w, A; \theta_0 s) \le \varepsilon + c_1\|f^s\|^2_{L^\infty(B_1)} = \varepsilon + c_1 s^{2k+1}\|f\|^2_{L^\infty(B_1)} \le 2\varepsilon.$$

□

Theorem 5.6 (Lower bound) *Let* (w, A) *be a solution of problem* (P) *in* Ω*. Let* $K \subset\subset \Omega$ *be a compact subset. Then, there exist positive constants* λ_K *and* r_K *depending only on* K*, the dimension* N*, the constant* M *in the assumption* (2)*, and* f *such that*

$$\mathrm{Per}(A; B_r(x)) \ge \lambda_K r^{N-1}, \tag{LB}$$

for every $r \in (0, r_K)$ *and every* $x \in \partial^* A \cap K$*.*

Proof Let $p(\theta_2) := \sum_{h=0}^{\infty} \theta_2^{(2k+1)h} \in \mathbb{R}$ and let $r_1 \in (0, 1)$ be a positive constant for which

$$r_1^{2k+1} c_1(\theta_2) p(\theta_2)\|f\|^2_{L^\infty(B_1)} \le \frac{\varepsilon_2}{4}.$$

We argue by contradiction. If the assertion does not hold, we would be able to find a point $x \in \partial^* A$ and a radius $r \le \min\{r_0, r_1\}$ for which

$$\mathrm{Per}\big(A; B_{\frac{r}{\theta_0}}(x)\big) \le \left(\frac{r}{\theta_0}\right)^{N-1} \kappa(\varepsilon), \quad \varepsilon := \frac{\varepsilon_2}{4}.$$

After translation, we may assume that $x = 0$. The fact that $r \le r_0$ and Lemma 5.5 yield the estimate

$$E_\gamma(w, A; r) \le 2\varepsilon \le \frac{\varepsilon_2}{2};$$

in return, Lemma 5.3 and a rescaling argument give (recall that $f^r(y) = r^{k+\frac{1}{2}} f(ry)$)

$$E_\gamma(w, A; \theta_2 r) \le E_\gamma(w^r, A^r; 1) + c_1\|f^r\|^2_{L^\infty(B_1)} \le \frac{\varepsilon_2}{2} + c_1 r^{2k+1}\|f\|^2_{L^\infty(B_1)} \le \varepsilon_2.$$

A recursion of the same argument gives the estimate

$$E_\gamma(w, A; \theta_2^j r) \le E_\gamma(w, A; r) + c_1 r^{2k+1}\|f\|^2_{L^\infty(B_1)}\left(\sum_{h=0}^{j} \theta_2^{(2k+1)h}\right) \le \varepsilon_2.$$

Taking the limit as $j \to \infty$ we get

$$\limsup_{j\to\infty} \frac{\mathrm{Per}(A; B_{\theta_2^j r})}{\mathrm{vol}\,(B_1') \cdot (\theta_2^j r)^{N-1}} \le \limsup_{j\to\infty} \frac{E_\gamma(w, A; \theta_2^j r)}{\mathrm{vol}\,(B_1') \cdot \gamma} \le \frac{\varepsilon_2}{\mathrm{vol}\,(B_1') \cdot \gamma} \le \frac{1}{2}.$$

This a contradiction to the fact that $x = 0 \in \partial^* A$ (cf. Sect. 2.4). □

Corollary 5.7 *Let* (w, A) *be a solution for problem* (P) *in* Ω*. Let* $K \subset\subset \Omega$ *be a compact subset. Then, there exist positive constants* λ_K *and* r_K *depending only on* K*, the dimension* N*, and* f *such that*

$$\mathrm{Per}(A; B_r(x)) \ge \lambda_K r^{N-1},$$

for every $r \in (0, r_K)$ *and for every* $x \in \partial A \cap K$*.*

Proof The property (LB) from the Lower bound theorem is a topologically closed property, i.e., it extends to $\overline{\partial^* A} = \operatorname{spt} \mu_A = \partial A$ (cf. (19)). □

Corollary 5.8 *Under the same assumptions of Theorem 5.6, the following characterization for the topological boundary of A holds:*

$$\partial A = \partial^* A \cup \Sigma, \quad \textit{where } \mathscr{H}^{N-1}(\Sigma) = 0.$$

Proof An immediate consequence of the previous corollary is that $\mathscr{H}^{N-1} \llcorner \partial A \ll |\mu_A|$ as measures in Ω. The assertion follows by De Giorgi's Structure Theorem. □

6 Proof of Theorem 1.5

As we have established in the past section, we will assume that for every $K \subset\subset \Omega$ there exist positive constants λ_K, C_K such that $D(w; x, r) \le C_K$ and

$$\operatorname{Per}(A, B_r(x)) \ge \lambda_K r^{N-1} \qquad \forall\, x \in (\partial A \cap K), \forall\, r \in (0, \operatorname{dist}(K, \partial\Omega)). \tag{LB}$$

Half-space regularity. Throughout this section we shall work with the additional assumption for solutions of the half-space problem: let $H := \{\, x \in \mathbb{R}^N \;:\; x_N > 0 \,\}$ and let σ_H be the two-point valued tensor defined in (3) for $\Omega = B_1$ (so that $\sigma_H = \sigma_1$ in $H \cap B_1$), then the operator

$$P_H u := \mathscr{A}^*(\sigma_H \mathscr{A} u)$$

is hypoelliptic in $B_1 \backslash \partial H$ in the sense that, if $w \in \mathrm{L}^2(B_1; \mathbb{R}^d)$, then[11]

$$P_H w = 0 \quad \Rightarrow \quad w \in \mathrm{C}^\infty(\overline{B_r^+}; \mathbb{R}^d) \cup \mathrm{C}^\infty(\overline{B_r^-}; \mathbb{R}^d) \quad \text{for every } 0 < r < 1. \tag{60}$$

Furthermore, there exists a positive constant $c^* = c^*(N, M, \mathscr{A})$ such that

$$\begin{aligned}
\frac{1}{\rho^N} \int_{B_\rho} |\nabla^k w|^2 \,\mathrm{d}x &\le c^* \int_{B_1} |\nabla^k w|^2 \,\mathrm{d}x && \text{for all } 0 < \rho \le \frac{1}{2},\\
\frac{1}{\rho^N} \int_{B_\rho} |\mathscr{A} w|^2 \,\mathrm{d}x &\le c^* \int_{B_1} |\mathscr{A} w|^2 \,\mathrm{d}x && \text{for all } 0 < \rho \le \frac{1}{2},\\
\sup_{B_\rho^+ \cup B_\rho^-} |\nabla^{k+1} w|^2 &\le c^* \int_{B_1} |w|^2 \,\mathrm{d}x && \text{for all } 0 < \rho \le \frac{1}{2}.
\end{aligned} \tag{61}$$

Remark 6.1 (Half-space regularity in applications) For 1-st order operators of gradient form it is relatively simple to show that such estimates as in (61) hold. This case includes gradients and symmetrized gradients; while the linear plate equations may be also reduced to this case (cf. Remark 2.8).

A sketch of the proof is as follows: the first step is to observe that the tangential derivatives $(i \ne N)$ $\partial_i w$ of a solution w of $P_H u = 0$ are also solutions of $P_H u = 0$. The second step is to repeat recursively the previous step and use the Caccioppoli inequality from Lemma 2.5 to estimate

$$\int_{B_{1/2}} |\partial^\alpha w|^2 \,\mathrm{d}x \le C(|\alpha|) \int_{B_1} |w|^2 \,\mathrm{d}x \quad \text{for arbitrary } \alpha \text{ with } \alpha_N \le 1. \tag{62}$$

[11] The notation $B_r^\pm$ stands for the upper and lower half ball of radius r: $B_r \cap H$ and $B_r \cap -H$ respectively.

The third step consists in using the ellipticity of $A_N = \mathbb{A}(\mathbf{e}_N)$ (cf. Remark 2.3) and the equation to express $\partial_{NN} w$ in terms of the rest of derivatives[12]: The tensor $(A_N^T \sigma A_N)$ is invertible, this can be seen from the inequality $|\mathbb{A}(\mathbf{e}_N)z|^2 \geq \lambda(\mathscr{A})|z|^2$ for every $z \in \mathbb{R}^d$ (cf. 2.3) and the fact that σ_H satisfies Gårding's strong inequality (2) with M^{-1}. Hence, using that $P_H w = 0$, we may write

$$\partial_{NN} w = -(A_N^T \sigma_H A_N)^{-1} \sum_{ij \neq NN} (A_i^T \sigma_1 A_j)\partial_{ij} w \quad \text{in } B_1^+, \tag{63}$$

from which estimates for $\partial_{NN} w$ of the form (62) in the upper half ball easily follow (similarly for the lower half ball). Further ∂_N differentiation of the equation in $B_1^{\pm}$ and iteration of this procedure together with the Sobolev embedding yield bounds as in (61).

For arbitrary higher-order gradients and other general elliptic systems one cannot rely on the same method. However, the Schauder and L^p boundary regularity of such systems has been systematically developed in [31,32] through the so called *complementing condition*. In the case of strongly elliptic systems (cf. (2) and (11)) this complementing condition is fulfilled, see [32, pp 43-44]; see also [33] where a closely related *natural notion* of hypoellipticity of the half-space problem is assumed.

Flatness excess. Given a set $A \subset \mathbb{R}^N$ of locally finite perimeter, the *flatness excess* of A at x for scale r and with respect to the direction $\nu \in \mathbb{S}^{n-1}$, is defined as

$$e(A; x, r, \nu) := \frac{1}{r^{N-1}} \int_{C(x,r,\nu)\cap\partial^* A} \frac{|\nu_E(y) - \nu|^2}{2} \, \mathrm{d}\mathscr{H}^{n-1}(y).$$

Here, $C(x, r, \nu)$ denotes for the cylinder centered at x with height 2, that is parallel to ν, of radius r.

Intuitively, the flatness excess expresses (for a set A) the deviation from being a hyperplane at a given scale r. Again, up to re-scaling, translating and rotating, it will be enough to work the case $x = 0$, $\nu = \mathbf{e}_N$, and $r = 1$. In this case, we will simply write $e(\mathrm{A})$. The hyper-plane energy excess is defined as

$$H_{\mathrm{ex}}(w, A; x, r, \nu) := e(A; x, r, \nu) + D(w, A; x, r),$$

and as long as its dependencies are understood we will simply write $H_{\mathrm{ex}}(r) = e(r) + D(r)$.

The following result relies on the (LB) property, a proof can be found in [24, Section 5.3] or [22, Theorem 22.8].

Lemma 6.2 (Height bound) *There exist positive constants $c_1^* = c_1^*(N)$ and $\varepsilon_1^* = \varepsilon_1^*(N)$ with the following property. If $A \subset \mathbb{R}^N$ is a set of locally finite perimeter with the* (LB) *property,*

$$0 \in \partial A \quad \textit{and} \quad \mathrm{e}(9) \leq \varepsilon_1^*,$$

then

$$\sup\{|y_N| : y \in B_1' \times [-1, 1] \cap \partial A\} \leq c_1^* \cdot e(4)^{\frac{1}{2N-2}}. \tag{HB}$$

The next decay lemma is the half-space problem analog of Lemma 5.1. The proof is similar except that it relies on the half-space regularity assumptions (60), (61) (instead of the ones given by Lemma 2.6), and the Height bound Lemma stated above.

[12] Recall that, for a 1-st order operator as in (7), the coefficients A_α can be simply denoted by A_i with $i = 1, \ldots, N$.

Lemma 6.3 (Approximative solutions of the half-space problem) *Let* (w, A) *be a solution of problem* (P) *in* B_1. *Then, for every* $\theta_1^* \in (0, 1/2)$ *there exist positive constants* $c_2^*(\theta_1^*, N, M)$ *and* $\varepsilon_2^*(\theta_1^*, N, M)$ *such that either*

$$\int_{B_\rho} |\mathscr{A} w|^2 \, \mathrm{d}x \le c_2^* \rho^N \|f\|^2_{\mathrm{L}^\infty(B_1)},$$

or

$$\int_{B_\rho} |\mathscr{A} w|^2 \, \mathrm{d}x \le 2c^* \rho^N \int_{B_1} |\mathscr{A} w|^2 \, \mathrm{d}x \quad \textit{for every } \rho \in [\theta_1, 1),$$

where $c^* = c^*(N, M)$ *is the constant from the regularity condition* (61)*; whenever*

$$\mathrm{Per}(A; B_1) \le \varepsilon_2^*.$$

□

Remark 6.4 Let $\delta \in (0, 1)$. Then there exists $\kappa^* = \kappa^*(N, M, \delta)$ such that if $e(1) \le \kappa^*$, and if one further assumes that the excess function $r \mapsto e(r)$ is monotone increasing, then the scaling $w(ry)/r^{(k-\frac{\delta}{2})}$ and the Iteration Lemma 4.6 imply that

$$\frac{1}{r^{N-\delta}} \int_{B_r} |\mathscr{A} w|^2 \le C_\delta \big(\|\mathscr{A} w\|^2_{\mathrm{L}^2(B_1)} + c_2^* \|f\|^2_{\mathrm{L}^\infty(B_1)} \cdot r^{2k+\delta} \big) \quad \text{for every } r \in (0, 1/2),$$

for some positive constant $C_\delta = C_\delta(N, M)$. □

The next crucial result can be found in [8, Section 5]. We have decided not to include a proof because the ideas remain the same. The ingredients for the proof are: the estimate (LB), the Height bound Lemma, the Lipschitz approximation Theorem, the estimates from Lemma 6.3 and the higher integrability for solutions to elliptic equations.[13]

Lemma 6.5 (Flatness excess improvement) *Let* (w, A) *be a saddle point of problem* (P) *in* Ω. *There exist positive constants* $\eta \in (0, 1]$, c_3^*, *and* ε_3 *depending only on* K, *the dimension* N, *the constant* M *in* (2)*, and* $\|f\|_{\mathrm{L}^\infty}$ *with the following properties: If* (w, A) *is a saddle point of problem* (P) *in* B_9, *and*

$$H_{\mathrm{ex}}(9) \le \varepsilon_3^*,$$

then, for every $r \in (0, 9)$*, there exists a direction* $\nu(r) \in \mathbb{S}^{N-1}$ *for which*

$$|\nu(r) - \mathbf{e}_N| \le c_3^* H_{\mathrm{ex}}(9) \quad \textit{and} \quad H_{\mathrm{ex}}(r, \nu(r)) \le c_3^* r^\eta H_{\mathrm{ex}}(9).$$

□

Theorem 1.5 (Partial regularity) *Let* (w, A) *be a saddle point of problem* (P) *in* Ω. *Assume that the operator* $P_H u = \mathscr{A}^*(\sigma \mathscr{A} u)$ *is hypoelliptic and regularizing as in* (60), (61)*, and that the higher integrability condition*

$$[\mathscr{A} \tilde{u}]^2_{\mathrm{L}^{2,N-\delta}(B_{1/2})} \le c \|\mathscr{A} \tilde{u}\|^2_{\mathrm{L}^2(B_1)}, \quad \text{for some } \delta \in [0, 1),$$

holds for every local minimizer $\tilde{u}$ *of the energy* $u \mapsto \int_{B_1} Q_{\mathscr{B}} W(\mathscr{A} u)$, *where* $u \in \mathrm{W}^{\mathscr{A}}(B_1)$. *Then there exists a positive constant* $\eta \in (0, 1]$ *depending only on* N *such that*

$$\mathscr{H}^{N-1}((\partial A \backslash \partial^* A) \cap \Omega) = 0, \quad \text{and} \quad \partial^* A \quad \text{is an open } \mathrm{C}^{1,\eta/2}\text{- hypersurface in } \Omega.$$

13 $\mathrm{L}^{2^*}(\Omega)$-integrability of $\mathscr{A} w$, for some exponent $2^* > 2$, can be established by standard methods through the use of the Caccioppoli inequality in Lemma 2.5.

Moreover if $\mathscr{A}$ is a first-order differential operator, then $\mathscr{A}w \in \mathrm{C}^{0,\eta/8}_{\mathrm{loc}}(\Omega\setminus(\partial A\setminus\partial^* A))$; *and hence, the trace of $\mathscr{A}w$ exists on either side of $\partial^* A$.*

Proof **The reduced boundary is an open hypersurface.** The first assertion $\mathscr{H}^{N-1}((\partial A\setminus \partial^* A)\cap\Omega) = 0$ is a direct consequence of Corollary 5.8.

To see that $\partial^* A$ is relatively open in ∂A we argue as follows: De Giorgi's Structure Theorem guarantees that for every $x \in \partial^* A$ there exist $r > 0$ (sufficiently small) and $\nu \in \mathbb{S}^{N-1}$ such that

$$H_{\mathrm{ex}}(w, A; r, x, \nu) \leq \frac{1}{2}\varepsilon_3^*, \quad \text{and} \quad \mu_A(\partial B_r(x)) = 0.$$

The map $y \mapsto \mu_A(B_r(y)) = 0$ is continuous at x, therefore we may find $\delta(x) \in (0,1)$ such that

$$H_{\mathrm{ex}}(w, A; r, y, \nu) \leq \varepsilon_3^* \quad \text{for every } y \in B_\delta(x)\cap\partial A.$$

We may then apply Lemma 6.5 to get an estimate of the form

$$\inf_{\xi\in\mathbb{S}^{N-1}} H_{\mathrm{ex}}(w, A; y, \rho, \xi) \leq c_3^*\rho^\eta H_{\mathrm{ex}}(w, A; y, r, \nu) \quad \text{for all } y \in B_\delta(x), \text{and all } \rho \in (0, r).$$

This and the first assertion of Lemma 6.5 imply that $y \in \partial^* A$ for every $y \in B_\delta(x)\cap\partial A$. Therefore, the reduced boundary $\partial^* A$ is a relatively open subset of the topological boundary ∂A.

We proceed to prove the regularity for $\partial^* A$. It follows from the last equation that

$$\begin{aligned} D(w; y, \rho) &\leq \inf_{\xi\in\mathbb{S}^{N-1}} H_{\mathrm{ex}}(w, A; y, \rho, \xi) \\ &\leq c_3^*\varepsilon_3^*\rho^\eta \leq C\rho^\eta \quad \text{for every } y \in B_\delta(x), \text{and every} \rho \in (0, r), \end{aligned} \tag{64}$$

for some constant $C = C(C_{B_\delta(x)}, \Lambda_{B_\delta(x)}, N, M)$.

Through a simple comparison, we observe from (64) and the property that (w, A) is a local saddle point of problem (P) in $B_\delta(x)$, that

$$\begin{aligned} &\mathrm{Dev}_{B_\delta(x)}(A, \rho) \leq 2M\rho^{N-1} D(w; y, \rho) \\ &\quad\leq 2MC\rho^{N-1+\eta}, \quad \text{for all } \rho \in (0, r) \text{ and every } y \in B_\delta(x). \end{aligned}$$

We conclude with an application of Tamanini's Theorem 2.14:

$$\partial A = \partial^* A \ \text{ is a } \mathrm{C}^{1,\eta/2}\text{- hypersurface in } B_\delta(x).$$

The assertion follows by observing that the regularity of $\partial^* A$ is a local property.

Jump conditions for the hyper-space problem. Let $\tau \in \mathrm{L}^2_{\mathrm{loc}}(B_1; Z) \cap (\mathrm{C}^\infty(\overline{B_\rho^+}; Z) \cup \mathrm{C}^\infty(\overline{B_\rho^-}; Z))$ for every $\rho \in (0,1)$, assume furthermore that τ is a solution of the equation

$$\mathscr{A}^*\tau = 0 \quad \text{in } B_1.$$

Let $\eta \in \mathrm{C}_c^\infty(B_1'; \mathbb{R}^d)$ be an arbitrary test function and choose a function $\varphi \in \mathrm{C}_c^\infty(B_1; \mathbb{R}^d)$ with the following property:

$$\varphi(y', y_N) = \frac{y_N^{k-1}}{(k-1)!}\eta(y') \quad \text{in a neighborhood of } B_1'.$$

Then, integration by parts and Green's Theorem yield that

$$0 = \int_{B_1} \tau\cdot\mathscr{A}\varphi \,\mathrm{d}y = \int_{\partial H\cap B_1} [\mathbb{A}(\mathbf{e}_N)^T\cdot\tau]\cdot\eta\,\mathrm{d}y',$$

where $[\mathbb{A}(\mathbf{e}_N)^T \cdot \tau] = \mathbb{A}(\mathbf{e}_N)^T \cdot (\tau^+ - \tau^-)$. Here, τ^+ and τ^- are the traces of τ in ∂H from B_1^+ and B_1^- respectively. Since η is arbitrary, a density argument shows that

$$[\mathbb{A}(\mathbf{e}_N)^T \cdot \tau] = 0 \quad \text{in } \partial H \cap B_1, \quad \text{and hence } \mathbb{A}(\mathbf{e}_N)^T \cdot \tau \ \in \mathrm{W}^{1,2}_{\mathrm{loc}}(B_1; \mathbb{R}^d). \tag{65}$$

Regularity of $\mathscr{A} w$. From this point and until the end of the proof we will assume that $\mathscr{A}$ is a first-order differential operator of gradient form; we may as well assume that $\partial^* A$ is locally parametrized by $\mathrm{C}^{1,\eta/2}$ functions.

Due to Campanato's Theorem ($\mathrm{C}^{0,\eta/8} \simeq \mathrm{L}^{2,\,N+(\eta/4)}$ on Lipschitz domains), our goal is to show local boundedness of the map

$$x \mapsto \sup_{r \le 1} \left\{ \frac{1}{r^{N+(\eta/4)}} \int_{B_r(x)\cap A} |\mathscr{A} w - (\mathscr{A} w)_{B_r(x)\cap A}|^2 \,\mathrm{d}y \right\} \quad x \in (\Omega \backslash (\partial A \backslash \partial^* A)); \tag{66}$$

and a similar result for A^c instead of A. □

Also, since Campanato estimates in the interior are a simple consequence of Lemma 2.6, we may restrict our analysis to show only local boundedness at points $x \in \partial^* A$. We first prove the following decay for solutions of the half-space:

Lemma 6.6 *Let $\tilde{w} \in \mathrm{W}^{\mathscr{A}}(B_1)$ be such that*

$$\mathscr{A}^*(\sigma_H \mathscr{A} \tilde{w}) = 0 \quad \textit{in } B_1. \tag{67}$$

Then $\tilde{w}$ satisfies an estimate of the form

$$\frac{1}{\rho^{N+2}} \int_{B_\rho} |R_H \tilde{w} - (R_H \tilde{w})_\rho|^2 \,\mathrm{d}y \le c(N, \sigma_1, \sigma_2) \int_{B_1} |R_H \tilde{w} - (R_H \tilde{w})_1|^2 \,\mathrm{d}y \tag{68}$$

for all $0 < \rho \le 1$, where we have defined

$$R_A u := \big(\nabla' u, A_N^T(\sigma_A \mathscr{A} u)\big), \quad A \subset B_1 \textit{ Borel.}$$

Proof Since for $\rho \ge 1/2$ one can use $c := 2^{(N+2)}$, we only focus on proving the estimate for $\rho \in (0, 1/2)$. It is easy to verify that $\mathscr{A}^*(\sigma_H \mathscr{A}(\partial_i \tilde{w} - \lambda)) = 0$ in $\mathscr{D}'(B_1; \mathbb{R}^d)$ for all $\lambda \in \mathbb{R}^d$, and every $i = 1, \dots, N-1$. In particular, by (61) we know that

$$\frac{1}{\rho^{N+2}} \int_{B_\rho} |\partial_i \tilde{w} - (\partial_i \tilde{w})_\rho|^2 \,\mathrm{d}y \le \frac{C}{\rho^N} \int_{B_\rho} |\nabla \partial_i \tilde{w}|^2 \,\mathrm{d}y \le c^* C \int_{B_1} |\partial_i \tilde{w} - (\partial_i \tilde{w})_1|^2 \,\mathrm{d}y, \tag{69}$$

for every $\rho \in (0, 1/2)$, and every $i = 1, \dots, N-1$. Here, $C = C(N)$ is the standard scaled Poincaré constant for balls. Summation over $i \in \{1, \dots, N-1\}$ yields an estimate of the form (68) for $\nabla' \tilde{w}$.

We are left to calculate the decay estimate for $g_H(\tilde{w}) := A_N^T(\sigma_H \mathscr{A} \tilde{w}) = \mathbb{A}(\mathbf{e}_N) \cdot (\sigma_H \mathscr{A} \tilde{w})$. By the hypoellipticity assumption (60) and the jump condition (65), we infer that $g_H(\tilde{w}) \in \mathrm{W}^{1,2}_{\mathrm{loc}}(B_1; \mathbb{R}^d)$.

Even more, by the classical Poincaré's inequality

$$\frac{1}{\rho^{N+2}} \int_{B_\rho} |g(\tilde{w}) - (g(\tilde{w}))_\rho|^2 \,\mathrm{d}y \le \frac{C}{\rho^N} \int_{B_\rho \backslash \partial H} |\nabla (g(\tilde{w}))|^2 \,\mathrm{d}y \tag{70}$$

for every $\rho \in (0, 1/2)$. On the other hand, it follows from the equation in $(B_1 \backslash \partial H)$ and (63) that one may write $\nabla g(\tilde{w})$ in terms of $\nabla(\nabla' \tilde{w})$ for almost every $x \in (B_r \backslash \partial H)$. We may then find a constant $C' = C'(\sigma_1, \sigma_2, \mathscr{A})$ such that

$$|\nabla g(\tilde{w}(x))|^2 \le C' |\nabla(\nabla' \tilde{w})(x)|^2 \quad \text{for every } x \in (B_\rho \backslash \partial H).$$

Using the same calculation as in the derivation of (69), it follows from (70) that

$$\frac{1}{\rho^{N+2}}\int_{B_\rho}|g(\tilde{w})-(g(\tilde{w}))_\rho|^2\,\mathrm{d}y \le c^*\,C\,C'\int_{B_1}|\nabla'\tilde{w}-(\nabla'\tilde{w})_1|^2\,\mathrm{d}y$$
$$\le c^*\,C\,C'\int_{B_1}|R_H\tilde{w}-(R_H\tilde{w})_1|^2\,\mathrm{d}y,$$

for every $\rho \in (0,1/2)$. The assertion follows by letting $c(N,\sigma_1,\sigma_2) := c^*C\max\{1, C'\}$. □

The next corollary can be inferred from (68) by following the strategy of Lin in [8, pp 166–167]:

Corollary 6.7 *Let $\tilde{w} \in \mathrm{W}^{\mathscr{A}}(B_2)$ solve the equation*

$$\mathscr{A}^*(\sigma_A\mathscr{A}u) = f \quad \text{in } B_2, \qquad \text{and assume furthermore that}$$
$$\|\tilde{w}\|_{\mathrm{L}^2(B_2)} \le 1 \quad \text{and} \quad \|f\|_{\mathrm{L}^\infty(B_2)} \le 1, \tag{71}$$

where $A := \{\, x \in B_2' \times \mathbb{R} \;:\; x_N > \varphi(x') \,\}$ for some function $\varphi \in \mathrm{C}^{1,\eta/2}(B_2')$ with $\varphi(0) = |\nabla\varphi|(0) = 0$, and $\|\varphi\|_{\mathrm{C}^{1,\eta/2}(B_2')} \le 1$. Then there exist positive constants $\theta(N,\sigma_1,\sigma_2) \in (0,1/2)$, and $C(N,\sigma_1,\sigma_2)$ such that either

$$\frac{1}{\theta^{N+1}}\int_{B_\theta}|R_A\tilde{w}-(R_A\tilde{w})_\theta|^2\,\mathrm{d}y \le \int_{B_1}|R_A\tilde{w}-(R_A\tilde{w})_1|^2\,\mathrm{d}y, \tag{72}$$

or

$$\int_{B_\theta}|R_A\tilde{w}-(R_A\tilde{w})_\theta|^2\,\mathrm{d}y \le C\Big(\|\varphi\|_{\mathrm{C}^{1,\eta/2}(B_1')} + \|f\|^2_{\mathrm{L}^\infty(B_1)}\Big). \tag{73}$$

□

We are now in the position to prove (66). Let $\delta \in (0,\eta/2)$ and let (w, A) be solution of problem (P). Since local regularity properties of the pair (w, A) are inherited to any (possibly rotated and translated) re-scaled pair $(w^{x,r}, A^{x,r})$ – as defined in (47), where in particular the source $f^{x,r}$ tends to zero – with $r \le \operatorname{dist}(x,\partial\Omega)$, we may do the following assumptions without any loss of generality: $B_4 \subset \Omega$ and $x = 0 \in \partial^*A$, ∂A^* is parametrized in B_2 by a function $\varphi \in \mathrm{C}^{1,\eta/2}(B_2')$ such that $\varphi(0) = |\nabla\varphi(0)| = 0$, and $\|\varphi\|_{\mathrm{C}^{1,\eta/2}(B_2')}, \|f\|_{\mathrm{L}^\infty(B_2;\mathbb{R}^d)} \le \min\{1,\kappa^*\}$ where $\kappa^* = \kappa^*(\delta, N, M)$ is the constant of Remark 6.4. Additionally, since (w, A) is a solution of problem (P), we know that

$$\mathscr{A}^*(\sigma_A\mathscr{A}w) = f \quad \text{in } B_2, \tag{74}$$

and

$$\frac{1}{r^{N-\delta}}\int_{B_r}|\mathscr{A}w|^2\,\mathrm{d}y \le C_\delta\big(\|\mathscr{A}w\|^2_{\mathrm{L}^2(B_2)} + \|f\|^2_{\mathrm{L}^\infty(B_1)}\big) \quad \text{for every } r \in (0,1), \tag{75}$$

where $C_\delta(N, M)$ is the constant from Remark 6.4.

Notice that the rescaled functions[14] $w^r(y) := (w(ry) - v_r(ry))/r^{1-(\delta/2)}$ and $\varphi^r(y) := \varphi(ry)/r$ still solve (74) for $f^r(y) := r^{1+(\delta/2)}f(ry)$ and $A^r := A/r$ with $\|\varphi^r\|_{\mathrm{C}^{1,\eta/2}(B_2')}$, $\|f^r\|_{\mathrm{L}^\infty(B_2;\mathbb{R}^d)} \le \min\{1,\kappa^*\}$. In particular, by (75) and Poincaré's inequality

$$\|w^r\|^2_{\mathrm{L}^2(B_1)} \le C(B_1)\|\mathscr{A}w^r\|^2_{\mathrm{L}^2(B_1)} < \overline{C} := C(B_1)C_\delta\big(\|\mathscr{A}w\|^2_{\mathrm{L}^2(B_2)} + 1\big).$$

14 Here, v_r is the $\mathscr{A}$-free corrector function for w in B_r, see Definition 2.1.

Recall that $\|\varphi^r\|_{C^{1,\eta/2}(B_1')}$ scales as $r^{\eta/2}\|\varphi\|_{C^{1,\eta/2}(B_r')}$ and, in view of its definition, $\|f^r\|^2_{L^\infty(B_1)}$ scales as $r^{2+\delta}$. In view of these properties, we are in position to apply Corollary 6.7 to $w^r / \max\{1, \overline{C}^{1/2}\}$: We infer that either

$$\frac{1}{\theta^{N+1}} \int_{B_\theta} |R_{A^r} w^r - (R_{A^r} w^r)_\theta|^2 \, dy \le \int_{B_1} |R_{A^r} w^r - (R_{A^r} w^r)_1|^2 \, dy, \tag{76}$$

or

$$\int_{B_\theta} |R_{A^r} w^r - (R_{A^r} w^r)_\theta|^2 \, dy \le \max\{1, \overline{C}\} \cdot C(N, \sigma_1, \sigma_2) \Big(\|\varphi^r\|_{C^{1,\eta/2}(B_1')} + r^{2+\delta} \Big), \tag{77}$$

where $\theta = \theta(N, \sigma_1, \sigma_2) \in (0, 1/2)$ is the constant from Corollary 6.7.

It is not difficult to verify, with the aid of the Iteration Lemma 4.6, that re-scaling in (76) and (77) conveys a decay of the form

$$\frac{1}{r^{N+\eta/2-\delta}} \int_{B_r} |R_A(w - \nu_r) - (R_A(w - \nu_r))_r|^2 \, dy \le c' \quad \text{for all } r \in (0, 1), \tag{78}$$

and some constant $c' = c'(\delta, N, \sigma_1, \sigma_2, \|\mathscr{A} w\|_{L^2(B_2)})$.

The last step of the proof consists in showing that $R_A(w - \nu_r)$ dominates $\nabla(w - \nu_r)$. By the definition of R_A, it is clear that $|\nabla'(w - \nu_r)(x) - (\nabla'(w - \nu_r))_{B_r \cap A}|^2 \le |R_A(w - \nu_r)(x) - (R_A(w - \nu_r))_{B_r \cap A}|^2$ for all $x \in B_1$ and every $r \in (0, 1)$. We show a similar estimate for $\partial_N(w - \nu_r)$:

The pointwise Gårding inequality (2) and (11) imply, in particular, that the tensor $(\mathbb{A}(\mathbf{e}_N)^T \sigma_1 \; \mathbb{A}(\mathbf{e}_N)) = (A_N^T \sigma_1 A_N) \in \mathrm{Lin}\,(\mathbb{R}^d; \mathbb{R}^d)$ is invertible (use, e.g., Lax-Milgram in $\mathbb{R}^d$). Hence,

$$\partial_N(w - \nu_r) = (A_N^T \sigma_1 A_N)^{-1} \Big(g(w - \nu_r) \; - \sum_{j \ne N} (A_N^T \sigma_1 \; A_j) \partial_j (w - \nu_r) \Big) \quad \text{in } B_1 \cap A, \tag{79}$$

from where we deduce that

$$\begin{aligned} &\frac{1}{r^{N+(\eta/2)-\delta}} \int_{B_r \cap A} |\partial_N(w - \nu_r) - (\partial_N(w - \nu_r))_{B_r \cap A}|^2 \, dy \\ &\le \frac{c''}{r^{N+(\eta/2)-\delta}} \int_{B_r \cap A} |R_A(w - \nu_r) - (R_A(w - \nu_r))_{B_r \cap A}|^2 \, dy \end{aligned}$$

for some constant $c'' = c''(\sigma_1, \mathscr{A}) \ge 1$ bounding the right hand side of (79) in terms of $\nabla'(w - v_r)$ and $g(w - v_r)$.

By (78) and the estimate above we obtain

$$\begin{aligned} &\frac{1}{r^{N+(\eta/2)-\delta}} \int_{B_r \cap A} |\mathscr{A} w - (\mathscr{A}(w))_{B_r \cap A}|^2 \, dy \\ &= \frac{1}{r^{N+(\eta/2)-\delta}} \int_{B_r \cap A} |\mathscr{A}(w - \nu_r) - (\mathscr{A}(w - \nu_r))_{B_r \cap A}|^2 \, dy \\ &\le \frac{C(\mathscr{A})}{r^{N+(\eta/2)-\delta}} \int_{B_r \cap A} |\nabla(w - \nu_r) - (\nabla(w - \nu_r))_{B_r \cap A}|^2 \, dy \\ &\le \overline{c}(N, \sigma_1, \sigma_2, \|\mathscr{A} w\|_{L^2(B_2)}) := C(\mathscr{A}) \cdot c' \cdot c'', \end{aligned}$$

for every $r \in (0, 1)$. The assertion follows by taking $\delta = \eta/4$.

Notice that the dependence on $\|\mathscr{A} w\|_{\mathrm{L}^2(B_2)}$ is local since we assumed $B_4 \subset \Omega$; this means that in general we may not expect a uniform boundedness of the decay. Similar bounds for A replaced by A^c can be derived by the same method. □

Remark 6.8 (Regularity I) In general, for a k'-th order operator $\mathscr{A}$ of gradient form, the only feature required to prove the regularity of $\nabla^k w$ up to the boundary $\partial^* A$ by the same methods as for first-order operators of gradient form is to obtain an analog of Lemma 6.6 (and its Corollary 6.7) for higher-order operators.

More specifically, if $\tilde{w} \in \mathrm{W}^{\mathscr{A}}(B_1)$ is a solution of the equation

$$\mathscr{A}^*(\sigma_H \mathscr{A} u) = 0 \quad \text{in } B_1,$$

then $\tilde{w}$ satisfies an estimate of the form

$$\frac{1}{\rho^{N+2}} \int_{B_\rho} |R_H \tilde{w} - (R_H \tilde{w})_\rho|^2 \,\mathrm{d}y \leq c(N, \sigma_1, \sigma_2) \int_{B_1} |R_H \tilde{w} - (R_H \tilde{w})_1|^2 \,\mathrm{d}y \qquad (80)$$

for all $0 < \rho \leq 1$,

where

$$R_A u := \big(\nabla' u, \mathbb{A}(\mathbf{e}_N)^T (\sigma_A \mathscr{A} u)\big), \quad A \subset B_1.$$

Unfortunately, for $2k$'-th order systems of elliptic equations (with $k > 1$) it is not clear to us whether one can prove such decay estimates by standard methods. While a decay estimate for $\nabla^{k-1}(\nabla' u)$ can be shown by the very same method as the one in the proof of Theorem 1.5, the main problem centers in proving a decay estimate for the term $\mathbb{A}(\mathbf{e}_N)^T(\sigma \mathscr{A} u) \in \mathrm{W}^{1,2}(B_1)$ – cf. (65). Technically, the issue is that one cannot use the equation on half-balls to describe $\partial^{(0,\ldots,0,k)} u$ in terms of $\nabla^{k-1}(\nabla' u)$.

Remark 6.9 (Regularity II: linear plate theory) In the particular case of models in linear plate theory ($\mathscr{A} = \nabla^2$, $N = 2$, and $d = 1$) it is possible to show a decay estimate as in (80) for solutions $w \in \mathrm{W}^{2,2}_0(B_2)$ of the equation

$$\nabla \cdot (\nabla \cdot (\sigma_H \nabla^2 u)) = 0.$$

By Remark 2.8, there exists a field $w \in W^{1,2}(B_2; \mathbb{R}^2)$ which turns out to be a solution of the equation

$$\nabla \cdot (\mathbf{S}_H \, \mathscr{E} w) = 0,$$

where $\mathbf{S}$ is a positive fourth-order symmetric tensor such that $\sigma_H(x) = \mathbf{R}_\perp \mathbf{S}_H^{-1}(x) \mathbf{R}_\perp$; furthermore, $\mathbf{R}_\perp \mathscr{E} w = \sigma_H \nabla^2 u$. Since $\mathscr{A} = \nabla^2$, it is easy to verify that $A_\alpha = A_{(i,j)} = \mathbf{e}_i \otimes \mathbf{e}_j$ for $i, j \in \{1, 2\}$, a simple calculation shows that

$$g_H(u) := \mathbb{A}(\mathbf{e}_N)^T(\sigma_H \mathscr{A} u) = (\sigma_H \nabla^2 u)_{22} = (\mathbf{R}_\perp \mathscr{E} w)_{22} = \partial_1 w^1;$$

and thus, since $\mathscr{E}$ is an operator of gradient form of order one, it follows form the proof of Theorem 1.5 that an estimate of the form (80) indeed holds for $g_H(u)$.

Acknowledgements I wish to extend many thanks to Prof. Stefan Müller for his advice and fruitful discussions in this beautiful subject. I would also like to thank the reviewer and the editor for their patience and care which derived in the correct formulation of Theorem 1.5. The support of the University of Bonn and the Hausdorff Institute for Mathematics is gratefully acknowledged. The research conducted in this paper forms part of the author's Ph.D. thesis at the University of Bonn.

7 Glossary of constants

N spatial dimension
M coercivity and bounding constant for the tensors σ_1 and σ_2 (as quadratic forms)
K an arbitrary compact set in Ω
λ_K local upper bound constant

Other constants Groups of constants are numbered in non-increasing order, e.g., $c_1^* \geq c_2^* \geq c_3^*$. The following constants play an important role in our calculations:

Constant	Dependence	Description
θ_1	arbitrary in $(0, 1/2)$	Ratio constant
c_1	θ_1, N, M	Universal constant
ε_1	θ_1, N, M	Smallness of perimeter density
c_2	N, M	Universal constant
γ	N, M	Universal constant
θ_2	N, M	Universal constant
ε_2	N, M	Smallness of excess energy
$\theta_0(\varepsilon)$	N, M, K	Smallness of perimeter density
c_1^*	λ_K, N	Constant in the Height bound Lemma
θ_1^*	arbitrary in $(0, 1/2)$	Ratio constant
c_2^*	θ_1^*, N, M	Universal constant
ϵ_2^*	θ_1^*, N, M	Smallness of flatness excess
c_3^*	K, N, M, f	Flatness excess improvement scaling constant
ε_3^*	K, N, M, f	Smallness of flatness excess

References

1. Hashin, Z., Shtrikman, S.: A variational approach to the theory of the elastic behaviour of multiphase materials. J. Mech. Phys. Solids **11**, 127–140 (1963)
2. Kohn, R.V., Strang, G.: Optimal design and relaxation of variational problems. I. Commun. Pure Appl. Math. **39**(1), 113–137 (1986)
3. Kohn, R.V., Strang, G.: Optimal design and relaxation of variational problems. II. Commun. Pure Appl. Math. **39**(2), 139–182 (1986)
4. Kohn, R.V., Strang, G.: Optimal design and relaxation of variational problems. III. Commun. Pure Appl. Math. **39**(3), 353–377 (1986)
5. Murat, F., Tartar, L.: Optimality conditions and homogenization. In: Nonlinear variational problems (Isola d'Elba, 1983), Res. Notes in Math., vol. 127, pp. 1–8. Pitman, Boston (1985)
6. Murat, F., Tartar, L.: Calcul des variations et homogénéisation. In: Homogenization methods: theory and applications in physics (Bréau-sans-Nappe, 1983), Collect. Dir. Études Rech. Élec. France, vol. 57, pp. 319–369. Eyrolles, Paris (1985)
7. Ambrosio, L., Buttazzo, G.: An optimal design problem with perimeter penalization. Calc. Var. Partial Differ. Equ. **1**(1), 55–69 (1993)
8. Lin, F.H.: Variational problems with free interfaces. Calc. Var. Partial Differ. Equ. **1**(2), 149–168 (1993)
9. Kohn, R.V., Lin, F.H.: Partial regularity for optimal design problems involving both bulk and surface energies. Chin. Ann. Math. Ser. B **20**(2), 137–158 (1999)
10. Larsen, C.J.: Regularity of components in optimal design problems with perimeter penalization. Calc. Var. Partial Differ. Equ. **16**(1), 17–29 (2003)
11. Fusco, N., Julin, V.: On the regularity of critical and minimal sets of a free interface problem. Interfaces Free Bound **17**(1), 117–142 (2015)
12. De Philippis, G., Figalli, A.: A note on the dimension of the singular set in free interface problems. Differ. Integr. Equ. **28**(5–6), 523–536 (2015)

13. Carozza, M., Fonseca, I., Di Napoli, A.P.: Regularity results for an optimal design problem with a volume constraint. ESAIM Control Optim. Calc. Var. **20**(2), 460–487 (2014)
14. De Giorgi, E.: Frontiere orientate di misura minima. Seminario di Matematica della Scuola Normale Superiore di Pisa, 1960–61. Editrice Tecnico Scientifica, Pisa (1961)
15. Almgren, F.: Existence and regularity almost everywhere of solutions to elliptic variational problems among surfaces of varying topological type and singularity structure. Ann. Math. **2**(87), 321–391 (1968)
16. Fonseca, I., Müller, S.: $\mathscr{A}$-quasiconvexity, lower semicontinuity, and Young measures. SIAM J. Math. Anal. **30**(6), 1355–1390 (1999)
17. Carstensen, C., Müller, S.: Local stress regularity in scalar nonconvex variational problems. SIAM J. Math. Anal. **34**(2), 495–509 (2002)
18. Brézis, H.: Analyse fonctionnelle. Collection Mathématiques Appliquées pour la Maîtrise, Masson, Paris
19. Barton, A.: Gradient estimates and the fundamental solution for higher-order elliptic systems with rough coefficients. Manuscripta Mathematica **151**(3–4), 375–418 (2016)
20. Milton, G.W.: The Theory of Composites, Cambridge Monographs on Applied and Computational Mathematics, vol. 6. Cambridge University Press, Cambridge (2002)
21. Ball, J.: A version of the fundamental theorem for Young measures. In: PDEs and continuum models of phase transitions. Proceedings of an NSF-CNRS joint seminar held in Nice, France, January 18–22, 1988, pp. 207–215. Springer, Berlin (1989)
22. Maggi, F.: Sets of Finite Perimeter and Geometric Variational Problems, Cambridge Studies in Advanced Mathematics, vol. 135. Cambridge University Press, Cambridge (2012). (An introduction to geometric measure theory)
23. 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)
24. Federer, H.: Geometric measure theory. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York (1969)
25. Tamanini, I.: Regularity Results for Almost Minimal Oriented Hypersurfaces in $\mathbb{R}^n$. Dipartimento di Matematica dell'Università di Lecce, Lecce (1984)
26. Allaire, G., Kohn, R.V.: Optimal bounds on the effective behavior of a mixture of two well-ordered elastic materials. Q. Appl. Math. **51**(4), 643–674 (1993)
27. Allaire, G., Kohn, R.V.: Optimal lower bounds on the elastic energy of a composite made from two non-well-ordered isotropic materials. Q. Appl. Math. **52**(2), 311–333 (1994)
28. Evans, L.C.: Quasiconvexity and partial regularity in the calculus of variations. Arch. Rational Mech. Anal. **95**(3), 227–252 (1986)
29. Sverák, V., Yan, X.: Non-Lipschitz minimizers of smooth uniformly convex functionals. Proc. Natl. Acad. Sci. USA **99**(24), 15269–15276 (2002)
30. Giaquinta, M.: Multiple Integrals In the Calculus of Variations and Nonlinear Elliptic Systems, Annals of Mathematics Studies, vol. 105. Princeton University Press, Princeton (1983)
31. Agmon, S., Douglis, A., Nirenberg, L.: Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. I. Commun. Pure Appl. Math. **12**, 623–727 (1959)
32. Agmon, S., Douglis, A., Nirenberg, L.: Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. II. Commun. Pure Appl. Math. **17**, 35–92 (1964)
33. Simon, L.: Schauder estimates by scaling. Calc. Var. Partial Differ. Equ. **5**(5), 391–407 (1997)