ENERGY SCALING LAWS FOR GEOMETRICALLY LINEAR ELASTICITY MODELS FOR MICROSTRUCTURES IN SHAPE MEMORY ALLOYS*

SERGIO CONTI[1], JOHANNES DIERMEIER[1], DAVID MELCHING[2]
AND BARBARA ZWICKNAGL[3],[**]

Abstract. We consider a singularly-perturbed two-well problem in the context of planar geometrically linear elasticity to model a rectangular martensitic nucleus in an austenitic matrix. We derive the scaling regimes for the minimal energy in terms of the problem parameters, which represent the shape of the nucleus, the quotient of the elastic moduli of the two phases, the surface energy constant, and the volume fraction of the two martensitic variants. We identify several different scaling regimes, which are distinguished either by the exponents in the parameters, or by logarithmic corrections, for which we have matching upper and lower bounds.

1. INTRODUCTION

Solid–solid phase transitions are a classical model problem in the variational study of pattern formation in solids, both in the context of the theory of relaxation and in the study of singularly perturbed problems. Their study has led on the one side to many important abstract developments in the calculus of variations, on the other side to a mathematical explanation of the physical behavior of shape-memory alloys and other materials with peculiar properties [3–5, 9, 34, 42, 48]. The basic model is a vectorial, nonconvex variational problem, where the integrand depends on the gradient of the deformation field. The study of the macroscopic material behavior is strongly coupled to the development of the theory of quasiconvexity and relaxation [29, 48], and focuses on average properties of the microstructures without resolving the geometric details and the microscopic length scales.

A finer analysis requires the introduction of a length scale, typically in the form of a small parameter times a convex function of a second gradient, which penalizes interfaces. The resulting singularly-perturbed nonconvex problem contains a scale dependence and is much more difficult to study in detail, a numerical treatment is in

[1] Institut für Angewandte Mathematik, Universität Bonn, 53115 Bonn, Germany.
[2] Fakultät für Mathematik, Universität Wien, 1090 Wien, Austria.
[3] Institut für Mathematik, Humboldt-Universität zu Berlin, 10117 Berlin, Germany.

[**] Corresponding author: barbara.zwicknagl@math.hu-berlin.de

most cases not feasible either. Starting with the papers by Kohn and Müller [43, 44] it has become clear that the key property is the scaling of the optimal energy in terms of the parameters present in the problem, and that it is appropriate to start by focusing on the exponents and ignoring the prefactor. One obtains mesoscopic phase diagrams which characterize the different regimes of material behavior and the qualitative properties of the microstructure [10, 38, 39, 42]. The techniques developed for singularly-perturbed functionals modeling martensitic microstructures have proven useful also in the study of a variety of other physical problems, such as for example magnetic microstructures [17, 19, 40], flux tubes in superconductors [18, 23, 25], diblock copolymers [16], wrinkling in thin elastic films [7, 8, 36], and compliance minimization [45].

One aspect which is very important for practical applications of materials with solid–solid phase transitions is the detailed study of the transformation path from austenite to martensite and the corresponding hysteresis. It is known that the amplitude of the hysteresis cycle crucially depends on the microstructures that emerge during nucleation [28, 52, 53]. Specifically, transition-state theory explains that the transformation from austenite to martensite is strongly influenced by the energetics of the critical nucleus, which is a small inclusion of martensite in an austenitic matrix. It is known that stress-free inclusions with interfaces of finite total area (or length, in two dimensions) are possible only for special material parameters [12, 24, 33, 37–39, 49, 50].

We investigate here a variational model for the formation of microstructures in a martensitic nucleus embedded in an austenitic matrix. The mechanical framework is the theory of geometrically linear elasticity, the mathematical framework is a singularly perturbed nonconvex vectorial functional. Before discussing the large body of mathematical literature that has been devoted to variants of this problem in the last decades, let us briefly introduce the setting. For simplicity we work in two spatial dimensions and consider a large body, identified with $\mathbb{R}^2$, which is mostly austenitic with a bounded martensitic inclusion $\omega \subset\subset \mathbb{R}^2$. The inclusion ω is selected by a process slower than elastic equilibration and therefore, for the present purposes, fixed.

We take the austenite state as reference configuration and denote by $u : \mathbb{R}^2 \to \mathbb{R}^2$ the elastic displacement. We assume that two variants of martensite are relevant, which are characterized by strains $A, B \in \mathbb{R}^{2 \times 2}_{\mathrm{sym}}$. Experimentally it is known that martensitic transformations are to a very good approximation volume preserving [9], therefore we assume $\mathrm{Tr}\, A = \mathrm{Tr}\, B = 0$. Relaxation theory predicts zero macroscopic energy if A and B are compatible and austenite can be realized as a weighted average of the two martensitic variants. This means that there are matrices $\hat{A}, \hat{B} \in \mathbb{R}^{2 \times 2}$ with $\hat{A} - A$ and $\hat{B} - B$ skew-symmetric such that $\mathrm{rank}(\hat{A} - \hat{B}) = 1$ and $(1 - \theta)\hat{A} + \theta\hat{B} = 0$ for some $\theta \in (0, 1)$. In this situation, a finer analysis, which includes a singular perturbation regularizing the microstructure, is necessary in order to understand the detailed material behavior. Since austenite/martensite interfaces which are not aligned with the rank-one direction have very large energy, one expects the nucleus to be elongated in the rank-one direction. For mathematical simplicity it is convenient to further restrict the geometry. Since $\hat{A}$ and $\hat{B}$ are rank-one connected, we have $\hat{A} - \hat{B} = c \otimes n$ for some $c, n \in \mathbb{R}^2$. By scaling we can assume $|c| = |n| = 1$. From $\mathrm{Tr}\, A = \mathrm{Tr}\, B = 0$ one obtains $\mathrm{Tr}\, \hat{A} = \mathrm{Tr}\, \hat{B} = 0$ and therefore $c \cdot n = 0$, and from $(1 - \theta)\hat{A} + \theta\hat{B} = 0$ one obtains $\hat{A} = \theta c \otimes n$, $\hat{B} = (\theta - 1)c \otimes n$. By a change of variables one can reduce to the case that $\theta \leq \frac{1}{2}$, $c = e_1$ and $n = e_2$. We shall then assume that the martensitic domain is a rectangle elongated along e_1, and by scaling it suffices to consider

$$\Omega_{2L} := (0, 2L) \times (0, 1) \subset \mathbb{R}^2. \tag{1.1}$$

The same pair of matrices allows for a second rank-one connection, rotated by 90 degrees. Therefore we can assume without loss of generality that

$$L \geq \frac{1}{2}.$$

In particular, both edges of Ω_{2L} are aligned with the habit planes of exact austenite/martensite interfaces. In the austenite the elastic energy vanishes if the strain $e(u)$, defined by

$$e(u) := (\nabla u)_{\mathrm{sym}} := \frac{1}{2}(\nabla u + \nabla^T u), \tag{1.2}$$

vanishes; in the martensite if $e(u) \in \{A, B\}$, and (assuming sufficient regularity) grows quadratically close to these minima. The relevant constructions have strains which are not larger than a multiple of the order parameter, hence we do not expect the behavior of the energy at infinity to be important for the scaling results we shall derive, provided sufficient coercivity is present. For simplicity we restrict to quadratic energies, characterized as the squared distance from the energy wells. We use $|a| := (\mathrm{Tr}\, a^T a)^{1/2} = (\sum a_{ij}^2)^{1/2}$ for the Euclidean norm of a matrix and $\mathrm{dist}(a, M) := \inf\{|a - m| : m \in M\}$ for the distance of a matrix to a set and consider the functional $J : W_{\mathrm{loc}}^{1,2}(\mathbb{R}^2; \mathbb{R}^2) \to \mathbb{R} \cup \{\infty\}$ given by

$$J(u) := \mu \int_{\mathbb{R}^2 \setminus \Omega_{2L}} |e(u)|^2 \, \mathrm{d}\mathcal{L}^2 + \int_{\Omega_{2L}} \mathrm{dist}^2(e(u), K) \, \mathrm{d}\mathcal{L}^2 + \varepsilon |D^2 u|(\Omega_{2L}). \tag{1.3}$$

Let us briefly explain the terms in the functional. The first term in $J(u)$ represents the elastic energy of the surrounding austenite, where μ stands for the ratio of typical elastic moduli of austenite and martensite. This term favors configurations whose gradients are approximately skew symmetric. The second term measures the elastic energy inside the martensitic nucleus, which vanishes on

$$K := \left\{ \frac{1}{2} \begin{pmatrix} 0 & \theta \\ \theta & 0 \end{pmatrix}, \ \frac{1}{2} \begin{pmatrix} 0 & -1 + \theta \\ -1 + \theta & 0 \end{pmatrix} \right\}. \tag{1.4}$$

The parameter $\theta \in (0, 1/2]$ measures the compatibility between this majority martensitic variant and the surrounding austenite. The so-measured compatibility has been found to play an important role in the control of the thermal hysteresis of the phase transition (see *e.g.* [28, 35, 52] and the references therein). Of particular interest is the almost compatible case $\theta \ll 1$ which corresponds to particularly low hysteresis [28, 52, 53]. The third term in (1.3) is a singular perturbation that regularizes the nonconvex part of the functional. It prevents too fine oscillations between the martensitic variants in Ω_{2L}, and can be related to an interfacial energy, $\varepsilon > 0$ being a typical surface energy constant per unit length. Whereas one could physically imagine that similar terms are present also in the austenitic phase, they are normally not included since the convex austenitic energy does not need regularization. Although we expect most of our results to carry over to a setting in which this term is extended to $\mathbb{R}^2$, for brevity we do not pursue this investigation here. The energy of the austenite/martensite interface depends only on the shape of the inclusion, which is fixed here, and is hence irrelevant for the present purposes. We denote by $D^2 u$ the second distributional derivative of u. If it is a measure, then we denote by $|D^2 u|(\Omega_{2L})$ the total variation of $D^2 u$, otherwise we set $|D^2 u|(\Omega_{2L}) := \infty$. Existence of minimizers can be readily established by the direct method of the calculus of variations and will not be discussed explicitly, as it is not important for the study of the scaling of the energy.

To determine exact minimizers of functionals like (1.3) is generally not possible, and we follow the strategy to determine the scaling regimes of the minima in terms of the problem parameters L, μ, θ and ε. We remark that $L \to \infty$ corresponds to the long-inclusion limit, which is relevant due to the compatibility condition; $\theta \to 0$ is the almost-compatible limit, which is the one of low-hysteresis materials; $\varepsilon \to 0$ is the large-body limit, in which complex structures arise.

Our main result is the following scaling law for the minimal energy. For an overview over the individual regimes we refer to Section 2.2 below. Explicit constructions are given in Section 2.1. We remark that the same result holds for nonlinear energy densities $W_A(e(u))$, $W_M(e(u))$ with $\frac{1}{c}|a|^2 \leq W_A(a) \leq c|a|^2$ and $\frac{1}{c}\mathrm{dist}^2(a, K) \leq W_M(a) \leq c\,\mathrm{dist}^2(a, K)$ for all $a \in \mathbb{R}^{2 \times 2}$.

Theorem 1.1. *There exists a constant $c > 0$ such that for all $\mu > 0$, $\varepsilon > 0$, $\theta \in (0, 1/2]$, and $L \geq 1/2$*

$$\frac{1}{c} \mathcal{I}(\mu, \varepsilon, \theta, L) \leq \min_u J(u) \leq c \mathcal{I}(\mu, \varepsilon, \theta, L),$$

where J was defined in (1.1), (1.2), (1.3), (1.4) and

$$
\begin{aligned}
\mathcal{I}(\mu, \varepsilon, \theta, L) := \min \Big\{ & \theta^2 L, && \text{(constant)} \\
& \mu \theta^2 \ln(3 + L), && \text{(affine)} \\
& \mu \theta^2 \ln\left(3 + \frac{L}{\mu}\right) + \varepsilon \theta, && \text{(linear interpolation)} \\
& \mu \theta^2 \ln\left(3 + \frac{\varepsilon L}{\mu \theta^2}\right) + \mu \theta^2 \ln\left(3 + \frac{\varepsilon}{\mu^2 \theta^2}\right) + \varepsilon^{1/2} \theta^{3/2}, && \text{(single truncated branching)} \\
& \mu \theta^2 \ln\left(3 + \frac{\varepsilon L}{\mu \theta^2}\right) + \mu \theta^2 \ln\left(3 + \frac{\theta}{\mu}\right), && \text{(corner laminate)} \\
& \varepsilon^{2/3} \theta^{2/3} L^{1/3} + \varepsilon L, && \text{(branching)} \\
& \mu^{1/2} \varepsilon^{1/2} \theta L^{1/2} \left(\ln(3 + \frac{1}{\theta^2})\right)^{1/2} + \varepsilon L, && \text{(laminate)} \\
& \mu^{1/2} \varepsilon^{1/2} \theta L^{1/2} \left(\ln(3 + \frac{\varepsilon}{\mu^3 \theta^2 L})\right)^{1/2} + \varepsilon L \Big\} && \text{(two-scale branching)}.
\end{aligned}
$$

Proof. The upper bound follows from Theorem 2.7, and the lower bound follows from Theorem 3.13. $\square$

The proof of Theorem 1.1 is split into two main steps. In Section 2 we combine constructions from the literature with some new ones for the upper bound. The ansatz-free lower bound is proven in Section 3.

We remark that the constant 3 inside the terms of the form $\ln(3 + x)$ is, to a certain degree, arbitrary. We choose 3 so that $\ln(3 + x) \geq 1$ for all $x \geq 0$. This simplifies some estimates in the proofs. The constant 3 could be replaced by 2 or by any number larger than 1, changing correspondingly the constant c in the statement.

We point out that in contrast to previous works on a single austenite/martensite interface (see *e.g.* [21, 27, 53]) there is no relevant regime which corresponds to a construction with a single laminate near the left and right boundaries of the nucleus.

1.1. Comparison to the literature and new contributions

We point out that the study of microstructures in shape memory alloys by means of the Calculus of Variations has a long history, and we generalize and build on several earlier works that we will briefly discuss to point out our new contributions. Our proof uses techniques developed in the study of scalar-valued models taking into account all problem parameters on the one hand, and of vectorial models in specific parameter regimes on the other hand, combined with several new arguments. As will be outlined in more detail below, the latter include in particular

– the careful treatment of the elastic energy in the austenite part which completely surrounds a martensitic inclusion. This introduces new difficulties in both the upper and the lower bound of Theorem 1.1; and
– the explicit use of the full geometrically linearized energy instead of the scalar simplification, which requires *e.g.* a BD-type slicing argument in the proof of the lower bound.

In the 1990s, Kohn and Müller proposed a reduced scalar-valued model for the formation of microstructures near interfaces between austenite and twinned martensite [43, 44]. These models are by now well understood in terms of scaling of the minimal energy and more quantitative properties of minimizers in specific regimes (see *e.g.* [20–22, 27, 30, 32, 53]). Roughly speaking, depending on the problem parameters, minimizers are expected to be uniform, or show laminated structures, or branched patterns, where the different martensitic variants

finely mix close to the interface. Compared to the setting we consider here, in these earlier works there were two main simplifications:

First, only one component of the displacement has been taken into account, *i.e.*, only functions $u = (u_1, u_2)$ with $u_2 = 0$ are considered, which makes the problem scalar (similarly for $u_1 = 0$). On a more technical level, the first two terms of the functional (1.3) then in particular provide control on the full gradient of the displacements. In our more general setting, only the symmetric part of the gradient is directly controlled by the functional, and this introduces several additional difficulties in the proof of the lower bound (see also the discussion of vectorial models below). Nevertheless, the constructions we use to prove the upper bound here, are in fact scalar valued. Some of them build upon constructions introduced in the above mentioned references, but others are new, as for example the one for the corner laminate and the single truncated branching, see the proof of Theorem 2.7 in Section 2.

Second, in the above references, only one austenite/martensite interface is considered. That is, in the elastic energy of the austenite part (the first term in (1.3)), only the contribution from $(-\infty, 0) \times (0, 1)$ is taken into account. If we restricted the energy in (1.3) to this strip, there would be configurations with vanishing total energy, *e.g.*,

$$u(x) = \begin{cases} 0, & \text{if } x \in (-\infty, 0] \times (0, 1), \\ (0, \theta x_1), & \text{if } x \in (0, 2L) \times (0, 1). \end{cases}$$

The functional (1.3) is more nonlocal than the scalar valued models in the sense that interactions between the traces at the upper and lower boundaries of Ω_{2L} (captured by the elastic energy of the austenite part) make it sometimes more favorable to pay elastic energy inside Ω_{2L} to release elastic energy in the austenite part. Let us consider a typical example. Deep in Ω_{2L}, in the simplified setting one expects the affine configuration $u_1(x) = \theta x_2$, which has zero energy in Ω_{2L}. In our setting instead, by Rellich's trace theorem, this configuration bears elastic energy in the austenite part. It therefore competes with a single laminate, which requires only surface energy in the interior of Ω_{2L}. The proof of the lower bound correspondingly needs a treatment of the interplay of the energy in the austenite and the martensite on many different scales. This is done by line integrals, inspired by the arguments used for proving Korn-Poincaré inequalities in BD, see for example Lemmas 3.5 and 3.6 below. One important ingredient is a separate treatment of the austenite part, where one controls three of the four components of ∇u (but with a coefficient μ), and the martensite part, where only the two diagonal entries are controlled independently of the variant.

On the other hand, we extend techniques developed in [13–15]. In these works, the geometrically nonlinear analogue to (1.3) has been considered for the case $\theta = \frac{1}{2}$ and hard austenite $\mu = \infty$. Some of their techniques, in particular related to localization in the proof of the lower bound, have been adopted to the geometrically linear setting and refined in two of the authors' Master's theses [31, 47] on which we build here. A main difficulty in our setting compared to those works (in addition to the 'non-locality' due to the elastic energy in the austenite part discussed above) lies in the treatment of small θ. As pointed out in [27], such localization techniques are not sufficient to obtain the precise scaling of laminated structures since the logarithmic corrections require a rather precise understanding of the geometry of the set in which the minority variant is active. Here, a careful BD-type slicing argument for almost diagonal slices allows us to combine the techniques from the studies of vectorial models with techniques developed to treat small volume fractions in the scalar valued case. In particular, test functions need to be obtained that reproduce the fine-scale structure of the martensite and have a controlled behavior at the boundaries, see Lemma 3.7 where for example separate test functions on the top and bottom boundaries need to be constructed for $\theta \geq \mu$ (the parameter range with corner laminates) and $\theta \leq \mu$ (without corner laminates), and also Lemma 3.11 where the corner logarithm is treated by a test function on the boundary. At the same time the "horizontal" interpolation between different variants needs to be localized in order to capture the optimal power of θ (Lem. 3.10).

Analytical results on microstructures for related three-dimensional models based on geometrically linearized elasticity functionals were obtained in [10, 11, 51] for a cubic-to-tetragonal phase transition and in [49, 50] for a

cubic-to-orthorhombic transition. As in our case, the focus there lies on planar austenite/martensite interfaces. Some results that take into account also the volume dependence of the energy of a martensitic inclusion (by penalizing the area of the austenite/martensite interfaces) and the resulting optimal shapes of nuclei (which typically differ significantly from a rectangle) were given in [38] for a two-well potential and in [6, 39, 41] for the cubic-to-tetragonal transition in whole space and domains with generic corners, respectively. We note that these works predict a different scaling behavior. We hope that a precise understanding of microstructures in a fixed domain as derived here provides also a step towards a better understanding of the full nucleation problem.

1.2. Notation

Throughout the text, we denote by c positive constants that may change from expression to expression, we use $x \lesssim y$ to state that there is $c > 0$ such that $x \leq cy$. We use capitalized letters and c_i with indices $i \in \mathbb{N}$ to denote specific fixed constants that will not be changed throughout the text.

For a measurable set $A \subset \mathbb{R}^d$ with $\mathcal{L}^d(A) \neq 0$ and a function $w \in L^1(A)$, we denote the average by $\langle w \rangle_A := \mathcal{L}^d(A)^{-1} \int_A w \, d\mathcal{L}^d$.

Energy. Let us first fix a notation for the function space

$$\mathcal{X} := \left\{ u \in W^{1,2}_{\mathrm{loc}}(\mathbb{R}^2, \mathbb{R}^2) \, : \, \partial_i u_j \in BV(\Omega_{2L}) \text{ for } i,j \in \{1,2\}, \nabla u \in L^2(\mathbb{R}^2, \mathbb{R}^{2\times 2}) \right\}$$

on which the energy is finite. In the proofs it will be convenient to consider a slightly modified energy functional where the symmetrized gradient in the elastic energy of the austenite part is replaced by the full gradient. This does not change the scaling regimes of the minimal energy due to Korn's inequality since the constant in Korn's inequality in $\mathbb{R}^2 \setminus \Omega_{2L}$ can be chosen independently of L, *i.e.*, there is a constant $C_K > 0$ independent of L such that

$$\min_{A \in \mathrm{Skew}(2)} \|\nabla u - A\|_{L^2(\mathbb{R}^2 \setminus \Omega_{2L})} \leq C_K \|e(u)\|_{L^2(\mathbb{R}^2 \setminus \Omega_{2L})} \text{ for all } u \in W^{1,2}_{\mathrm{loc}}(\mathbb{R}^2, \mathbb{R}^2). \tag{1.5}$$

To see this, we use the decomposition

$$\mathbb{R}^2 \setminus \Omega_{2L} = [(-\infty, 0) \times \mathbb{R}] \cup [\mathbb{R} \times (1, \infty)] \cup [(2L, \infty) \times \mathbb{R}] \cup [\mathbb{R} \times (-\infty, 0)].$$

Each one of the sets on the right-hand side is a half-space, and therefore on each of them, a Korn's inequality holds with a constant independent of L (see [46]). Further, every set intersects another one on a set of infinite measure. Hence, for any function u, the Korn's inequality in each one of the four parts holds with the same skew symmetric matrix A, and therefore, (1.5) holds. We may therefore without changing the qualitative scaling behavior replace the symmetrized gradient in the first term in (1.3) by the full gradient and define

$$I(u) := \mu \int_{\mathbb{R}^2 \setminus \Omega_{2L}} |\nabla u|^2 \, d\mathcal{L}^2 + \int_{\Omega_{2L}} \min \left\{ |e(u) - \theta e_1 \odot e_2|^2, |e(u) + (1-\theta)e_1 \odot e_2|^2 \right\} \, d\mathcal{L}^2 + \varepsilon |D^2 u|(\Omega_{2L})$$

where

$$e_1 \odot e_2 := (e_1 \otimes e_2)_{\mathrm{sym}}.$$

Sometimes it will be useful to consider the energy only on parts of the domain. For any Borel set $A \subset \mathbb{R}^2$ we define

$$I_A(u) := \mu \int_{A \setminus \Omega_{2L}} |\nabla u|^2 \, \mathrm{d}\mathcal{L}^2$$
$$+ \int_{A \cap \Omega_{2L}} \min\left\{|e(u) - \theta e_1 \odot e_2|^2, |e(u) + (1-\theta)e_1 \odot e_2|^2\right\} \, \mathrm{d}\mathcal{L}^2 + \varepsilon |D^2 u|(A \cap \Omega_{2L}),$$
$$I^{\mathrm{int}}(u) := I_{\Omega_{2L}}(u) \qquad \text{and} \qquad I^{\mathrm{ext}}(u) := I_{\mathbb{R}^2 \setminus \Omega_{2L}}(u).$$

The $H^{1/2}$-norm. It has proven useful to interpret the energetic contribution in the austenite region as a trace norm at the austenite/martensite interface. For $\rho > 0$ and $u_0 \in L^2((0,\rho))$ we define the $H^{1/2}$-seminorm by

$$[u_0]^2_{H^{1/2}((0,\rho))} := \inf\left\{ \int_{-\infty}^0 \int_0^\rho |\nabla v(x_1, x_2)|^2 \, \mathrm{d}x_2 \, \mathrm{d}x_1 \,:\, v(0, x_2) = u_0(x_2), v \in W^{1,2}_{\mathrm{loc}}((-\infty, 0) \times (0, \rho)) \right\}.$$

The subspace of $L^2((0,\rho))$ on which this seminorm is finite is called $H^{1/2}((0,\rho))$. We state a variant of Lemma 4.1 from [27].

Lemma 1.2. *Let $\omega \subset\subset (0,1)$. Then there is $c = c(\omega) > 0$ such that for all $v \in H^{1/2}((0,1))$ and $\psi \in H^{1/2}((0,1)) \cap H^1((0,1))$ with $\mathrm{supp}\,\psi \subset \omega$ one has*

$$\int_0^1 v(t)\psi'(t)\mathrm{d}t \le c[v]_{H^{1/2}((0,1))}[\psi]_{H^{1/2}((0,1))}.$$

Remark. Lemma 4.1 from [27] incorrectly does not state that c depends on ω (or on $\mathrm{supp}\,\psi$). This is not relevant for the usage in [27], since the test function ψ can be constructed to be supported in $(1/12, 11/12)$.

Proof. If $\mathrm{supp}\,v \subset\subset (0,1)$, then the assertion is readily proven by Fourier series,

$$\left| \sum_k \hat{v}_k^* ik\psi_k \right| \le \left(\sum_k |k| \, |v_k|^2\right)^{1/2} \left(\sum_k |k| \, |\psi_k|^2\right)^{1/2}.$$

Otherwise, one fixes $\varphi \in C_c^\infty((0,1))$ with $\varphi = 1$ on ω, applies the Poincaré estimate for $H^{1/2}$ to obtain $\|v - v_0\|_{L^2((0,1))} \le [v]_{H^{1/2}((0,1))}$ for some $v_0 \in \mathbb{R}$, and then applies the previous assertion to $\tilde{v} := \varphi(v - v_0)$ and ψ. $\square$

In our proof of the lower bound, we shall use a related estimate given in the next lemma.

Lemma 1.3. *Let $\omega \subset \mathbb{R}^2$ be a bounded Lipschitz set, $\Psi \in \mathrm{Lip}(\bar{\omega})$, $v \in W^{1,2}(\omega)$. Then*

$$\left| \int_{\partial\omega} v \partial_\tau \Psi \, \mathrm{d}\mathcal{H}^1 \right| \le \|\nabla\Psi\|_{L^2(\omega)} \|\nabla v\|_{L^2(\omega)},$$

where ∂_τ denotes the tangential derivative and v is identified with its trace on $\partial\omega$.

Proof. Assume first that $\Psi \in C^2(\bar{\omega})$. We define $f \in W^{1,2}(\omega; \mathbb{R}^2)$ by $f := v(\nabla\Psi)^\perp = (-v\partial_2\Psi, v\partial_1\Psi)$ and observe that $\mathrm{div}\, f = \nabla\Psi \times \nabla v = \partial_2 v \partial_1 \Psi - \partial_1 v \partial_2 \Psi$. Then

$$\left| \int_{\partial\omega} v \, \partial_\tau \Psi \, \mathrm{d}\mathcal{H}^1 \right| = \left| \int_{\partial\omega} f \cdot \nu \, \mathrm{d}\mathcal{H}^1 \right| = \left| \int_\omega \mathrm{div}\, f \, \mathrm{d}x \right| = \left| \int_\omega \nabla\Psi \times \nabla v \, \mathrm{d}x \right|$$

implies the result. In the general case, we choose a sequence of smooth functions $\Psi_\varepsilon \in C^\infty(\mathbb{R}^2)$ such that $\|\nabla \Psi_\varepsilon\|_{L^\infty(\omega)} \le \|\nabla \Psi\|_{L^\infty(\omega)}$, with Ψ_ε converging uniformly and strongly in $W^{1,2}(\omega)$ to Ψ. Then $\partial_\tau \Psi_\varepsilon$ converges weakly-$*$ to $\partial_\tau \Psi$ in $L^\infty(\partial\omega)$, therefore

$$\left| \int_{\partial\omega} v\, \partial_\tau \Psi \, d\mathcal{H}^1 \right| = \left| \lim_{\varepsilon \to 0} \int_{\partial\omega} v\, \partial_\tau \Psi_\varepsilon \, d\mathcal{H}^1 \right| = \left| \lim_{\varepsilon \to 0} \int_\omega \nabla \Psi_\varepsilon \times \nabla v \, dx \right| = \left| \int_\omega \nabla \Psi \times \nabla v \, dx \right|.$$

$\square$

We will moreover use the following variant of Lemma 2 in [21] that has also been used in the proof of Theorem 1 in [53]:

Lemma 1.4. *There is $c > 0$ such that for all $\rho > 0$, $v \in H^{1/2}((0,\rho))$, $a \in \mathbb{R}$, and $b \in \mathbb{R}$ there holds*

$$c\rho^2 a^2 \le |a| \int_0^\rho |v(y) - ay - b|\, dy + [v]^2_{H^{1/2}((0,\rho))}.$$

BD-type slicing. For any $u : \mathbb{R}^2 \to \mathbb{R}^2$, $x_1 \in \mathbb{R}$, and $\xi \in \mathbb{R}^2$ with $\xi_1 \ne 0 \ne \xi_2$, we define the one-dimensional, scalar-valued function on the slice $(x_1, 0) + \mathbb{R}\xi$ as

$$u^\xi_{x_1}(s) := \frac{1}{\xi_1 \xi_2} u((x_1, 0) + s\xi) \cdot \xi. \tag{1.6}$$

Notice that this definition is motivated by the characterization of BD via suitable one-dimensional sections as introduced by Ambrosio, Coscia and Dal Maso ([1], Prop. 3.2). For convenience, our definition differs from the one given in the above reference by the prefactor $\frac{1}{\xi_1 \xi_2}$.

If $u \in W^{1,2}_{\text{loc}}(\mathbb{R}^2, \mathbb{R}^2)$ we have $u^\xi_{x_1} \in W^{1,2}_{\text{loc}}(\mathbb{R})$ for almost every $x_1 \in \mathbb{R}$, and

$$u^\xi_{x_1}{}'(s) = \left(\frac{\xi_1}{\xi_2} \partial_1 u_1 + \partial_2 u_1 + \partial_1 u_2 + \frac{\xi_2}{\xi_1} \partial_2 u_2 \right)((x_1, 0) + s\xi). \tag{1.7}$$

Throughout this work we will fix the direction $\xi := (\frac{1}{4}, 1)$ and define for some $x_1 \in (0, 2L - \xi_1)$ the almost diagonal segment with base point $(x_1, 0)$ as

$$\Delta^\xi_{x_1} := \{(x_1, 0) + s\xi \,:\, s \in (0,1)\}. \tag{1.8}$$

With a small abuse of notation we shall write, for $f : \Delta^\xi_{x_1} \to \mathbb{R}$,

$$\|f\|^2_{L^2(\Delta^\xi_{x_1})} := \int_0^1 |f|^2((x_1, 0) + s\xi)\, ds, \tag{1.9}$$

and the same for L^1. This definition differs from the usual one, in which one integrates with respect to $\mathcal{H}^1$, by a factor of $\sqrt{17/16}$, which is irrelevant for our argument but would make notation cumbersome.

For any $a \in \mathbb{R}$, almost every $x_1 \in (0, 2L - \xi_1)$ and almost every $s \in (0,1)$ we compute

$$|u^\xi_{x_1}{}'(s) - a| = \left| \frac{1}{\xi_1 \xi_2} \sum_{ij} \xi_i \xi_j (e(u) - ae_1 \odot e_2)_{ij} \right|((x_1, 0) + s\xi)$$

$$\le \frac{|\xi|^2}{|\xi_1 \xi_2|} |e(u) - ae_1 \odot e_2|((x_1, 0) + s\xi) \le 5|e(u) - ae_1 \odot e_2|((x_1, 0) + s\xi) \tag{1.10}$$

where in the last step we used the specific choice $\xi = (1/4, 1)$. This implies in particular

$$\min\left\{|u_{x_1}^{\xi}{}' - \theta|^2, |u_{x_1}^{\xi}{}' + (1-\theta)|^2\right\}(s) \leq 25 \min\left\{|e(u) - \theta e_1 \odot e_2|^2, |e(u) + (1-\theta)e_1 \odot e_2|^2\right\}((x_1, 0) + s\xi). \tag{1.11}$$

We remark that vertical slices cannot be used to obtain similar estimates, since $\partial_2 u_2$ does not distinguish between the two variants and $\partial_2 u_1$ cannot be controlled without an independent estimate on $\partial_1 u_2$.

2. Upper bound

In this section we will prove the upper bound, *i.e.*, the second inequality in Theorem 1.1. For that, we provide explicit constructions for the different energy scaling regimes in Section 2.1. In Section 2.2 we give an overview over typical parameter ranges to illustrate that indeed all scalings are attained.

2.1. Explicit constructions

Let us point out that all our constructions will be scalar valued. We define for every $u \in W_{\text{loc}}^{1,2}(\mathbb{R}^2, \mathbb{R})$ with $\nabla u \in BV(\Omega, \mathbb{R}^2)$

$$E(u) := I((u, 0)) \leq \int_{\Omega_{2L}} \min\left\{|\nabla u - \theta e_2|^2, |\nabla u + (1-\theta)e_2|^2\right\} \, \mathrm{d}\mathcal{L}^2 + \mu \int_{\mathbb{R}^2 \setminus \Omega_{2L}} |\nabla u|^2 \, \mathrm{d}\mathcal{L}^2 + \varepsilon |D^2 u|(\Omega_{2L})$$

and correspondingly $E_A(u) := I_A((u, 0))$, E^{int}, and E^{ext}. Some of the test functions we consider below are taken from the literature, some constructions have to be modified, and some are new. We shall use only constructions that are symmetric with respect to the axis $\{x_1 = L\}$, working explicitly in $(-\infty, L] \times \mathbb{R}$ and then extending each construction by symmetry. This introduces an additional term $|D^2 u|(\{L\} \times (0, 1))$. In some cases, it vanishes (constant, affine, linear interpolation). In the other cases, we use that the relevant gradients are bounded, and hence the term $\varepsilon|D^2 u|(\{L\} \times (0, 1))$ is bounded by ε. Then this term can be incorporated in the regimes using $\varepsilon \leq \mu\theta^2$, see the proof of Theorem 2.7. We will therefore not explicitly mention this term in the discussion of the constructions below. We shall use the following short-hand notation: For $a < b$, we set

$$\iota_{a,b} : [a, b] \to \mathbb{R}, \quad \iota_{a,b}(t) := \frac{t - a}{b - a}, \tag{2.1}$$

i.e., $\iota_{a,b}$ is the affine function with $\iota_{a,b}(a) = 0$, $\iota_{a,b}(b) = 1$.
We start with auxiliary lemmata to estimate the energy contribution from $\mathbb{R}^2 \setminus \Omega_{2L}$.

Lemma 2.1. *Let $\bar{L} \in [1/2, \infty)$, $1 \leq \alpha < \beta \leq \bar{L}$. Then there exists $u_{\alpha,\beta,\bar{L}} \in W_{\text{loc}}^{1,2}\left(\left((-\infty, \bar{L}] \times \mathbb{R}\right) \setminus \Omega_{\bar{L}}\right) \cap C^0\left(\left((-\infty, \bar{L}] \times \mathbb{R}\right) \setminus \Omega_{\bar{L}}\right)$ such that*

(i) *on the lower boundary $u_{\alpha,\beta,\bar{L}}(x_1, 0) = 0$ for all $x_1 \in [0, \bar{L}]$;*
(ii) *on the upper boundary*

$$u_{\alpha,\beta,\bar{L}}(x_1, 1) = \begin{cases} 0, & \text{if } x_1 \in [0, \alpha], \\ \theta\iota_{\alpha,\beta}(x_1), & \text{if } x_1 \in (\alpha, \beta), \\ \theta, & \text{if } x_1 \in (\beta, \bar{L}]; \end{cases}$$

(iii) *on the interface $u_{\alpha,\beta,\bar{L}}(0, x_2) = 0$ for all $x_2 \in [0, 1]$;*

(iv) and there is a constant $c > 0$ independent of α, β, $\bar{L}$ and θ such that

$$\int_{\mathbb{R}^2 \setminus \Omega_{2\bar{L}}} |\nabla u_{\alpha,\beta,\bar{L}}|^2 \, d\mathcal{L}^2 \leq c\theta^2 \left(\ln \left(3 + \frac{\bar{L}}{\alpha} \right) + \frac{\beta + \alpha}{\beta - \alpha} \right).$$

Proof. We use polar coordinates, denoting by $\phi(x) \in (0, 2\pi)$ and $r(x) \in [0, \infty)$ the coordinates of $x = (x_1, x_2) \in \mathbb{R}^2 \setminus ((0, \infty) \times \{0\})$ so that

$$x_1 = r(x) \cos(\phi(x)), \qquad x_2 = r(x) \sin(\phi(x)).$$

We define $\hat{f}_{\alpha,\beta} : (0, 2\pi) \times [0, \infty) \to \mathbb{R}$ by

$$\hat{f}_{\alpha,\beta}(\phi, r) := \begin{cases} 0, & \text{if } r \in [0, \alpha], \\ (1 - \frac{\phi}{2\pi}) \iota_{\alpha,\beta}(r)\theta, & \text{if } r \in (\alpha, \beta], \\ (1 - \frac{\phi}{2\pi})\theta, & \text{if } r \in (\beta, \bar{L}], \\ \frac{\bar{L}^{1/2}}{r^{1/2}} (1 - \frac{\phi}{2\pi})\theta, & \text{if } r \in (\bar{L}, \infty), \end{cases}$$

and consider the transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$T(x_1, x_2) := \begin{cases} (x_1, x_2), & \text{if } x_2 \leq 0, \\ (x_1, 0), & \text{if } 0 < x_2 \leq 1, \\ (x_1, x_2 - 1), & \text{if } x_2 > 1. \end{cases}$$

We set for $x \in \left((-\infty, \bar{L}] \times \mathbb{R} \right) \setminus \overline{\Omega_{2\bar{L}}}$

$$u_{\alpha,\beta,\bar{L}}(x) := \hat{f}_{\alpha,\beta} \left(\phi(T(x)), r(T(x)) \right).$$

Note that for $x_1 > 0$ and $x_2 \searrow 1$, we have $r(T(x_1, x_2)) \to x_1$, and $\phi(T(x_1, x_2)) \searrow 0$. Similarly, for $x_1 > 0$ and $x_2 \nearrow 0$, we have $r(T(x_1, x_2)) \to x_1$ and $\phi(T(x_1, x_2)) \to 2\pi$. Finally, for $x_2 \in (0, 1)$ and $x_1 \nearrow 0$, we have $r(T(x_1, x_2)) \to 0$. Therefore $u_{\alpha,\beta,\bar{L}}$ has a continuous extension to $\left((-\infty, \bar{L}] \times \mathbb{R} \right) \setminus \Omega_{\bar{L}}$ and it satisfies (i), (ii) and (iii). It remains to verify (iv). From the definition of T we obtain, with $f_{\alpha,\beta}(x) := \hat{f}(\phi(x), r(x))$,
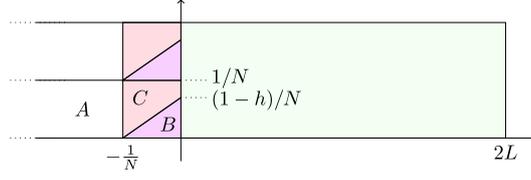
$$\|\nabla u_{\alpha,\beta,\bar{L}}\|^2_{L^2(((-\infty,\bar{L}) \times \mathbb{R}) \setminus \Omega_{\bar{L}})} = \|\nabla f_{\alpha,\beta}\|^2_{L^2(((-\infty,\bar{L}) \times \mathbb{R}) \setminus ([0,\bar{L}] \times \{0\}))} + \|\partial_1 f_{\alpha,\beta}(\cdot, 0)\|^2_{L^2((-\infty,0))}$$

where ∂_1 refers to the usual derivative in direction e_1. Using polar coordinates we compute

$$\|\nabla f_{\alpha,\beta}\|^2_{L^2(B_{\bar{L}}(0) \setminus ([0,\bar{L}] \times \{0\}))} = \int_0^{\bar{L}} \int_0^{2\pi} \frac{1}{r} |\partial_\phi \hat{f}_{\alpha,\beta}|^2 + r |\partial_r \hat{f}_{\alpha,\beta}|^2 \, d\phi \, dr \leq c\theta^2 \left(\int_\alpha^{\bar{L}} \frac{1}{r} \, dr + \int_\alpha^\beta \frac{r}{(\beta - \alpha)^2} dr \right)$$

$$\leq c\theta^2 \left(\ln \frac{\bar{L}}{\alpha} + \frac{\beta + \alpha}{\beta - \alpha} \right)$$

and

$$\|\nabla f_{\alpha,\beta}\|^2_{L^2(((-\infty,\bar{L}) \times \mathbb{R}) \setminus B_{\bar{L}}(0))} \leq \int_{\bar{L}}^\infty \int_0^{2\pi} \frac{1}{r} |\partial_\phi \hat{f}_{\alpha,\beta}|^2 + r |\partial_r \hat{f}_{\alpha,\beta}|^2 \, d\phi \, dr \leq c\theta^2 \int_{\bar{L}}^\infty \frac{\bar{L}}{r^2} \, dr = c\theta^2.$$

FIGURE 1. Construction in the proof of Lemma 2.3 for $N = 2$.

We also estimate,

$$\|\partial_1 f_{\alpha,\beta}(\cdot,0)\|^2_{L^2((-\infty,0))} = \frac{\theta^2}{4} \int_\alpha^\beta \frac{1}{(\beta-\alpha)^2} \, dx_1 + \frac{\theta^2}{16} \int_{\bar{L}}^\infty \frac{\bar{L}}{x_1^3} \, dx_1 = \frac{\theta^2/4}{\beta-\alpha} + \frac{\theta^2/32}{\bar{L}} \leq \frac{\theta^2}{\beta-\alpha}.$$

Recalling that $\alpha \geq 1$, we conclude

$$\|\nabla u_{\alpha,\beta,\bar{L}}\|^2_{L^2(((-\infty,\bar{L})\times\mathbb{R})\setminus((0,\bar{L})\times(0,1)))} \leq c\theta^2 \left(\ln\left(3 + \frac{\bar{L}}{\alpha}\right) + \frac{\beta+\alpha}{\beta-\alpha} \right).$$

$\square$

Remark 2.2. We will often use that the upper bound in Lemma 2.1(iv) is monotonically increasing in $\bar{L}$ and decreasing in β.

The following lemma has been used in [27] without being explicitly stated. We refer to Figure 1 for a sketch.

Lemma 2.3. *Let* $N \in \mathbb{N}$, $N \geq 1$, $h \in (0, \frac{1}{2}]$, $\theta \in [0, \frac{1}{2}]$. *Then there exists* $v_{N,h} \in W^{1,2}_{\mathrm{loc}}((-\infty,0)\times(0,1)) \cap C^0((-\infty,0]\times[0,1])$ *such that*

(i) $v_{N,h}(x_1,x_2) = v_{N,h}(x_1, x_2 + \frac{1}{N})$ *for all* $x_2 \in [0, 1-\frac{1}{N}]$ *and all* $x_1 \in (-\infty,0]$;
(ii) for $x_1 = 0$ *we have*

$$v_{N,h}(0,x_2) = \begin{cases} \theta x_2, & \text{if } x_2 \in [0, \frac{1-h}{N}], \\ \theta \frac{1-h}{h}(\frac{1}{N} - x_2), & \text{if } x_2 \in (\frac{1-h}{N}, \frac{1}{N}]; \end{cases}$$

(iii) for all $x_1 \in (-\infty, 0]$, *we have* $v_{N,h}(x_1, 0) = v_{N,h}(x_1, 1) = 0$;
(iv) for all $x_1 \in (-\infty, -\frac{1}{N}]$ *and all* x_2, *we have* $v_{N,h}(x_1,x_2) = 0$;
(v) and there exists $c > 0$ *independent of* N, θ *and* h *such that*

$$\int_{(-\infty,0)\times(0,1)} |\nabla v_{N,h}|^2 \, d\mathcal{L}^2 \leq c\frac{\theta^2}{N} \ln\left(3 + \frac{1}{h}\right).$$

Proof. We write v for $v_{N,h}$ and set

$$v(x) := \begin{cases} 0, & \text{if } x \in A := (-\infty, -\frac{1}{N}] \times [0, \frac{1}{N}), \\ \theta x_2, & \text{if } x \in B := \{x_1 \in (-\frac{1}{N}, 0], 0 \leq x_2 \leq (1-h)(x_1 + \frac{1}{N})\}, \\ \frac{(1-Nx_2)\theta(1-h)(Nx_1+1)}{N(h-(1-h)Nx_1)}, & \text{if } x \in C := \{x_1 \in (-\frac{1}{N}, 0], (1-h)(x_1 + \frac{1}{N}) < x_2 < \frac{1}{N}\}, \end{cases}$$

and extend it periodically in x_2 to $(-\infty, 0] \times [0, 1]$ as stated in (i) (see Fig. 1). One easily checks that v is continuous and satisfies (ii), (iii) and (iv). To show (v), we first work in $(-\frac{1}{N}, 0] \times [0, \frac{1}{N}) = B \cup C$. In region C
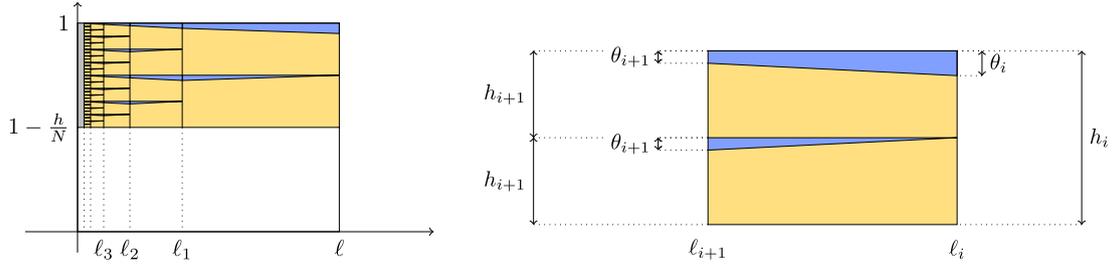
FIGURE 2. Sketch of the branched construction in Lemma 2.4. *Left*: global construction, with $h = 1/2$, $N = 1$, and four refinement steps. The gray region $(0, \ell_{k+1}) \times (1 - \frac{h}{N}, 1)$ is the one where linear interpolation is used. *Right*: enlargement of a unit cell $(\ell_{i+1}, \ell_i) \times (1 - h_i, 1)$ in the refinement. The colors mark the regions $\{\partial_2 u = \theta\}$ (yellow) and $\{\partial_2 u = \theta - 1\}$ (blue).

we have $(1 - h)(Nx_1 + 1) \le Nx_2 < 1$, and therefore $0 < 1 - Nx_2 \le h - (1 - h)Nx_1$, so that

$$|\nabla v|(x) \le \theta(1 - h)\left(\frac{2}{h - (1 - h)Nx_1} + \frac{1 - Nx_2}{(h - (1 - h)Nx_1)^2}\right) \le \frac{3\theta(1 - h)}{h - (1 - h)Nx_1} \text{ for all } x \in C.$$

Since $|\nabla v| = \theta$ in B we obtain

$$\int_{-1/N}^0 \int_0^{1/N} |\nabla v|^2 \, \mathrm{d}x_2 \, \mathrm{d}x_1 \le \frac{\theta^2}{N^2} + 9\theta^2 \int_{-1/N}^0 \int_{(1-h)(x_1 + \frac{1}{N})}^{1/N} \frac{(1 - h)^2}{(h - (1 - h)Nx_1)^2} \, \mathrm{d}x_2 \, \mathrm{d}x_1$$

$$= \frac{\theta^2}{N^2} + \frac{9\theta^2}{N} \int_{-1/N}^0 \frac{(1 - h)^2}{h - (1 - h)Nx_1} \, \mathrm{d}x_1 \le \frac{\theta^2}{N^2} + \frac{9\theta^2}{N^2} \ln \frac{1}{h} \le c \frac{\theta^2}{N^2} \ln\left(3 + \frac{1}{h}\right).$$

By periodicity the proof is concluded.   □

We now recall the basic branching construction from [27], which refines the one in [44] (see Fig. 2).

Lemma 2.4. *Suppose that* $\theta \in (0, 1/2]$, $N \in \mathbb{N}$, $N \ge 1$, $h \in [\theta, 1]$, $\ell \in [\theta, \infty)$. *Then there exists* $u := u_{h,\ell,N} \in W^{1,\infty}\left((0, \ell) \times (1 - \frac{h}{N}, 1)\right) \cap C^0\left([0, \ell] \times [1 - \frac{h}{N}, 1]\right)$ *with the following properties:*

(i) $u(0, x_2) = \frac{\theta}{h}(1 - h)(1 - x_2)$ *for all* $x_2 \in [1 - \frac{h}{N}, 1]$;
(ii)

$$u(\ell, x_2) = \begin{cases} \theta x_2 - \theta(1 - \frac{1}{N}), & \text{if } x_2 \in [1 - \frac{h}{N}, 1 - \frac{\theta}{N}], \\ (1 - \theta)(1 - x_2), & \text{if } x_2 \in (1 - \frac{\theta}{N}, 1]; \end{cases}$$

(iii) $u(x_1, 1 - \frac{h}{N}) = \frac{1}{N}\theta(1 - h)$ *for all* $x_1 \in [0, \ell]$;
(iv) $u(x_1, 1) = 0$ *for all* $x_1 \in [0, \ell]$;
(v) $|D^2 u|\left((0, \ell) \times (1 - \frac{h}{N}, 1)\right) \le c(\ell + \frac{h}{N})$ *and* $\|\nabla u\|_{L^\infty} \le c$;
(vi)

$$\int_{(0, \ell) \times (1 - \frac{h}{N}, 1)} (\partial_1 u)^2 + \min\left\{(\partial_2 u - \theta)^2, (\partial_2 u + 1 - \theta)^2\right\} \, \mathrm{d}\mathcal{L}^2 \le c \frac{\theta^2 h}{N^3 \ell}.$$

Proof. One can use the finite branching construction given in [27] building on Lemma ([27], Lem. 5.2) and truncation parameter $I \in \mathbb{N}$ such that $(3/2)^I \sim \ell/\theta$. The estimates then follow from the considerations in the proof of Proposition 6.1 from [27].

For the convenience of the reader we sketch the main steps of the construction, referring to Figure 2 for an illustration. For $i \in \mathbb{N}$ we set $h_i := 2^{-i}h/N$, $\theta_i := 2^{-i}\theta/N$, $\ell_i := 3^{-i}\ell$, and let k be the largest integer such that $\theta_k \leq \ell_k$. For $0 \leq i \leq k$, the function $x_2 \mapsto \partial_2 u(\ell_i, x_2)$ is h_i-periodic, with $\partial_2 u(\ell_i, x_2) = \theta$ for $x_2 \in (1 - h_i, 1 - \theta_i)$ and $\partial_2 u(\ell_i, x_2) = \theta - 1$ for $x_2 \in (1 - \theta_i, 1)$. In $(\ell_{i+1}, \ell_i) \times (1 - \frac{h}{N}, 1)$, $0 \leq i < k$, the function $x_2 \mapsto u(x_1, x_2) - \frac{\theta}{h}(1 - h)(1 - x_2)$ is h_i-periodic in the x_2 direction, u obeys (iii) and (iv), and is defined interpolating the boundary values as sketched in Figure 2. By construction $\partial_2 u \in \{\theta, \theta - 1\}$ almost everywhere in this set. One checks that $|\partial_1 u| \leq c\theta_i/\ell_i$, leading to $|D^2 u|([\ell_{i+1}, \ell_i] \times (1 - \frac{h}{N}, 1)) \leq c2^i \ell_i + c2^i \frac{h_i \theta_i}{\ell_i}$ and $\|\partial_1 u\|^2_{L^2([\ell_{i+1}, \ell_i] \times (1 - \frac{h}{N}, 1))} \leq c(\frac{\theta_i}{\ell_i})^2 \ell_i \frac{h}{N}$. Summing the two geometric series, this leads to the bounds in (v) and (vi) on $(\ell_k, \ell) \times (1 - \frac{h}{N}, 1)$. In $(0, \ell_k) \times (1 - \frac{h}{N}, 1)$ we use an affine interpolation. The condition $\theta_k \leq \ell_k$ gives $|\nabla u| \leq c$ in this region, so that the elastic energy is bounded by $c\frac{h}{N}\ell_k$. Since $\ell_{k+1} < \theta_{k+1}$ we have $\ell \leq \frac{\theta}{N}(\frac{3}{2})^{k+1} \leq 2\frac{\theta}{N}3^{k/2}$, which implies $\ell_k \ell \leq 4\frac{\theta^2}{N^2}$ and hence (v). In turn, $|D^2 u|((0, \ell_k) \times (1 - \frac{h}{N}, 1)) \leq c(2^k \ell_k + \frac{h}{N}) \leq c(\ell + \frac{h}{N})$. This concludes the proof. $\qquad \square$

Finally, we shall frequently use a function that interpolates between a single laminate and an affine function.

Lemma 2.5. *Let $\theta \in (0, 1/2]$ and $\tilde{\beta} > 0$. There exists a function $\tilde{v}_{\tilde{\beta}} = \tilde{v} \in C^0([0, \tilde{\beta}] \times [0, \theta])$ such that*

(i) $\tilde{v}(0, x_2) = (-1 + \theta)x_2$ and $\tilde{v}(\tilde{\beta}, x_2) = \theta x_2$ for all $x_2 \in [0, \theta]$;

(ii) $\tilde{v}(x_1, 0) = 0$ for all $x_1 \in [0, \tilde{\beta}]$;

(iii) $\tilde{v}(x_1, \theta) = \theta \iota_{0, \tilde{\beta}}(x_1) + \theta(-1 + \theta)$, with $\iota_{0, \tilde{\beta}}$ as in (2.1);

(iv) $\|\nabla \tilde{v}\|_{L^\infty} \leq 1 + \frac{\theta}{\tilde{\beta}}$ and the energy is estimated by

$$\int_{(0, \tilde{\beta}) \times (0, \theta)} \min \left\{ |\nabla \tilde{v} - \theta e_2|^2, |\nabla \tilde{v} + (1 - \theta)e_2|^2 \right\} \, d\mathcal{L}^2 + \varepsilon |D^2 \tilde{v}|((0, \tilde{\beta}) \times (0, \theta)) \leq c \left(\frac{\theta^3}{\tilde{\beta}} + \varepsilon(\tilde{\beta} + \frac{\theta^2}{\tilde{\beta}}) \right).$$

Proof. The standard interpolation

$$\tilde{v}(x) := \begin{cases} \theta x_2, & \text{if } x_2 \leq \frac{\theta}{\tilde{\beta}} x_1, \\ -(1 - \theta)x_2 + \frac{\theta}{\tilde{\beta}} x_1, & \text{if } x_2 > \frac{\theta}{\tilde{\beta}} x_1 \end{cases} \tag{2.2}$$

satisfies all required properties. We simplified the estimate using $\theta \leq \tilde{\beta} + \frac{\theta^2}{\tilde{\beta}}$. $\qquad \square$

We shall now provide the proof of the upper bound in Theorem 1.1. We proceed in two steps: in the first step, we adapt a result from the literature (Prop. 2.6), and in the second step, we provide test functions for the remaining regimes (Thm. 2.7).

Proposition 2.6 (Upper bound: constructions from the literature). *There exists $c > 0$ such that for all $\mu > 0$, $\varepsilon > 0$, $\theta \in (0, 1/2]$, and $L \geq 1/2$ there holds*

$$\min_u E(u) \leq c \min \left\{ \varepsilon^{2/3}\theta^{2/3}L^{1/3} + \varepsilon L, \; \mu^{1/2}\varepsilon^{1/2}\theta L^{1/2} \left(\ln \left(3 + \frac{1}{\theta^2} \right) \right)^{1/2} + \varepsilon L, \right.$$

$$\left. \mu^{1/2}\varepsilon^{1/2}\theta L^{1/2} \left(\ln \left(3 + \frac{\varepsilon}{\mu^3 \theta^2 L} \right) \right)^{1/2} + \varepsilon L \right\}.$$
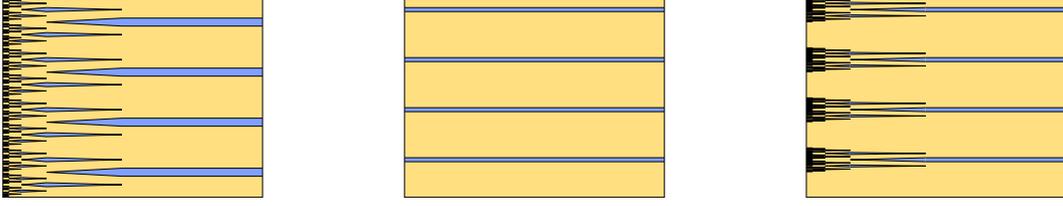
FIGURE 3. Sketch of the three regimes from Proposition 2.6. Only the left half of the martensite, in Ω_L, is plotted, for parameters for which the εL term is not relevant. *Left*: branching construction from Proposition 2.6(a). *Middle*: laminate from Proposition 2.6(b). *Right*: two-scale branching from Proposition 2.6(c).

Proof. The assertion follows from the proof of Proposition 5.1 in [27], checking carefully that the differences between the two functionals are not relevant. For clarity we provide a short self-contained argument, based on Lemmas 2.3 and 2.4 (both taken from [27]). The three constructions are illustrated in Figure 3.

(a) Let $N \geq 1$, $h := 1$, $\ell := L$, and let $u \in C^0([0, L] \times [1 - \frac{1}{N}, 1])$ be as in Lemma 2.4, extended periodically in x_2 to $[0, L] \times [0, 1]$, symmetrically for $x_1 \in (L, 2L]$ and then by zero outside $\overline{\Omega_{2L}}$. We have

$$E(u) \leq cN\varepsilon L + c\frac{\theta^2}{N^2 L} + \varepsilon.$$

Choosing N to be the smallest integer above $\theta^{2/3}\varepsilon^{-1/3}L^{-2/3}$, we obtain
$E(u) \leq c(\varepsilon^{2/3}\theta^{2/3}L^{1/3} + \varepsilon L)$.

(b) Let $N \geq 1$, $h := \theta$, and let $v_{N,\theta}$ be as in Lemma 2.3. We extend it setting $v_{N,\theta}(x) = v_{N,\theta}(0, x_2)$ for $x_1 \in (0, L]$, symmetrically for $x_1 \geq L$, and by zero on $\mathbb{R} \times (\mathbb{R} \setminus [0, 1])$. By Lemma 2.3(ii) we have $\partial_1 v_{N,\theta} = 0$ and $\partial_2 v_{N,\theta} \in \{\theta, \theta - 1\}$ in Ω_{2L}, so that $I_{\Omega_{2L}}(v_{N,\theta}) \leq 4N\varepsilon L$. By Lemma 2.3(v),

$$E(v_{N,\theta}) \leq c\mu\theta^2 N^{-1}\ln\left(3 + \frac{1}{\theta^2}\right) + cN\varepsilon L.$$

Choosing N as the smallest integer above $(\varepsilon^{-1}L^{-1}\mu\theta^2\ln(3 + \frac{1}{\theta^2}))^{1/2}$, we obtain
$E(v_{N,\theta}) \leq c(\mu^{1/2}\varepsilon^{1/2}\theta L^{1/2}(\ln(3 + \frac{1}{\theta^2}))^{1/2} + \varepsilon L)$.

(c) It remains to show that $I(c) \leq c\left(\mu^{1/2}\varepsilon^{1/2}\theta L^{1/2}(\ln(3 + \frac{\varepsilon}{\mu^3\theta^2 L}))^{1/2} + \varepsilon L\right)$. If $\mu^{3/2}\varepsilon^{-1/2}\theta L^{1/2} \leq \theta$, this follows from (b). Otherwise, we fix again $N \geq 1$, $h \in [\theta, 1]$, $\ell := L$ and use Lemma 2.4 and Lemma 2.3 to obtain a function u with

$$E(u) \leq c\mu\theta^2 N^{-1}\ln\left(3 + \frac{1}{h^2}\right) + cN\varepsilon L + c\frac{\theta^2 h}{N^2 L} + \varepsilon,$$

where we used that $\ln(3 + \frac{1}{h}) \leq \ln(3 + \frac{1}{h^2})$ since $h \leq 1$. We choose $h := \min\{1, \mu L(\mu\theta^2/(\varepsilon L))^{1/2}\} = \min\{1, \mu^{3/2}\varepsilon^{-1/2}\theta L^{1/2}\} \in [\theta, 1]$ and N to be the smallest integer above $(\mu\theta^2\ln(3 + \frac{\varepsilon}{\mu^3\theta^2 L})/(\varepsilon L))^{1/2}$. Using $h \leq \mu^{3/2}\varepsilon^{-1/2}\theta L^{1/2}$ and $\ln(3 + \frac{\varepsilon}{\mu^3\theta^2 L}) \geq 1$, the proof is concluded.
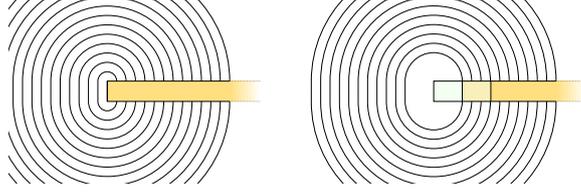
$\square$

FIGURE 4. *Left*: Sketch of the affine regime (see proof of Thm. 2.7(ii)). The martensite in Ω_{2L} has an affine deformation $u = \theta x_2$; in the austenite the field lines of ∇u are sketched. As usual, we only plot the region with $x_1 \leq L$. *Right*: Sketch of the linear interpolation regime (see proof of Thm. 2.7(iii)). The martensite has $u = \theta x_2$ only for $x_1 \in [2\mu, L]$, it has $u = 0$ for $x_1 \in [0, \mu]$, and the affine interpolation in between. Correspondingly the field lines start at $x_1 = \mu$. This construction is relevant for $\mu \gg 1$, for simplicity $\mu = 1.4$ is plotted.

Theorem 2.7 (Upper bound: conclusion). *There is a constant $c > 0$ such that for all $\mu > 0$, $\varepsilon > 0$, $\theta \in (0, 1/2]$, and $L \geq 1/2$*

$$\min_u J(u) \leq c\mathcal{I}(\mu, \varepsilon, \theta, L),$$

where $\mathcal{I}(\mu, \varepsilon, \theta, L)$ is given in Theorem 1.1.

Proof. By the estimate $|e(u)| \leq |\nabla u|$, it suffices to show an upper bound for I, which in turn follows from an upper bound for E. We provide test functions for the respective regimes separately. Some constructions are used for several test functions. We will describe them in detail the first time we use them and refer to the arguments in the path of the proof.

(0) Branching, laminate, two-scale branching: By Proposition 2.6, we have

$$\min_u E(u) \leq c \min \left\{ \varepsilon^{2/3}\theta^{2/3}L^{1/3} + \varepsilon L, \mu^{1/2}\varepsilon^{1/2}\theta L^{1/2} \left(\ln \left(3 + \frac{1}{\theta^2} \right) \right)^{1/2} + \varepsilon L, \right.$$

$$\left. \mu^{1/2}\varepsilon^{1/2}\theta L^{1/2} \left(\ln \left(3 + \frac{\varepsilon}{\mu^3\theta^2 L} \right) \right)^{1/2} + \varepsilon L \right\}.$$

(i) Constant: Set $u^{(1)} := 0$ in $\mathbb{R}^2$. This shows $\min_u E(u) \leq E(u^{(1)}) \leq 2\theta^2 L$.

(ii) Affine: we aim to show that $\min_u E(u) \leq c\mu\theta^2 \ln(3 + L)$. Choose $\bar{L} := L + 2$, $\beta := 2$ and $\alpha := 1$. We note that the assumptions of Lemma 2.1 are satisfied, and we shall use the function $u_{\alpha,\beta,\bar{L}}$. Precisely, we set

$$u^{(2)}(x) := \begin{cases} \theta x_2, & \text{if } x \in \Omega_L, \\ u_{1,2,L+2}(x_1 + 2, x_2), & \text{if } x \in ((-\infty, L] \times \mathbb{R}) \setminus ((-2, L] \times (0, 1)), \\ 0, & \text{if } x \in (-2, -1] \times (0, 1), \\ (1 + x_1)\theta x_2, & \text{if } x \in (-1, 0] \times (0, 1), \\ u^{(2)}(2L - x_1, x_2), & \text{if } x_1 > L \end{cases}$$

see Figure 4, left panel. We obtain, by Lemma 2.1 (using that $\frac{\beta+\alpha}{\beta-\alpha} = 3 \leq c\ln(3 + L)$) and an explicit computation in $(-1, 0) \times (0, 1)$,

$$\min_u E(u) \leq E(u^{(2)}) = E^{\text{ext}}(u^{(2)}) \leq c\mu\theta^2 \ln(3 + L).$$

(iii) Linear interpolation: We aim to show that $\min_u E(u) \leq c \left(\mu\theta^2 \ln(3 + L/\mu) + \varepsilon\theta \right)$.
We distinguish some cases.
a) If $\mu \leq 1$, this holds by (ii).
b) If $\mu \geq L/3$, this holds by (i).
c) If $\mu \in (1, L/3)$, we choose $\alpha := \mu$, $\beta := 2\mu$ and $\overline{L} := L$. Note that these choices are admissible for Lemma 2.1 since $\alpha = \mu > 1$ and $L \geq 3\mu \geq \beta$. We set (with ι as defined in (2.1))

$$
u^{(3)}(x) := \begin{cases}
0, & \text{if } x \in [0, \mu] \times [0, 1], \\
\iota_{\mu, 2\mu}(x_1)\theta x_2, & \text{if } x \in (\mu, 2\mu] \times [0, 1], \\
\theta x_2, & \text{if } x \in (2\mu, L] \times [0, 1], \\
u_{\mu, 2\mu, L}(x), & \text{if } x \in ((-\infty, L] \times \mathbb{R}) \setminus \overline{\Omega_L}, \\
u^{(3)}(2L - x_1, x_2), & \text{if } x_1 > L
\end{cases}
$$

see Figure 4, right panel. Then $E^{\text{ext}}(u^{(3)}) \leq c\mu\theta^2 \ln(3 + \frac{L}{\mu})$ by Lemma 2.1, and hence by an explicit computation in Ω_L using that $\mu < 1$

$$
E(u^{(3)}) \leq c \left(\mu\theta^2 \ln\left(3 + \frac{L}{\mu}\right) + \varepsilon\theta + \frac{\varepsilon\theta}{\mu} + \mu\theta^2 \right) \leq c \left(\mu\theta^2 \ln\left(3 + \frac{L}{\mu}\right) + \varepsilon\theta \right).
$$

(iv) Next we aim to show an auxiliary result, namely that

$$
\min_u E(u) \leq c \left(\mu\theta^2 \ln\left(3 + \frac{\varepsilon L}{\mu\theta^2}\right) + \mu\theta^2 \ln\left(3 + \frac{1}{\theta^2}\right) + \varepsilon^{1/2}\theta^{3/2} \right). \tag{2.3}
$$

This bound is not needed for the proof of the theorem, but it introduces a new construction method that will be used for cases (v) and (vi) below. The idea behind the construction is a single laminate close to the left and right boundaries of the nucleus, interpolated to an affine function deep in the bulk. Related estimates play also a role in the proof of the lower bound, see *e.g.* the assumptions of Proposition 3.3.
We distinguish three cases:
a) If $\mu\theta^2\varepsilon^{-1} \leq 1$, then (2.3) follows from (ii).
b) If $\mu\theta^2\varepsilon^{-1} \geq L$, we use the function $v_{1,\theta}$ from Lemma 2.3 with $N = 1$ and $h = \theta$, and set

$$
u^{(4)}(x) := \begin{cases}
\theta x_2, & \text{if } x \in [0, L] \times [0, 1 - \theta], \\
(1 - \theta)(1 - x_2), & \text{if } x \in [0, L] \times (1 - \theta, 1], \\
v_{1,\theta}(x), & \text{if } x \in (-\infty, 0) \times [0, 1], \\
0, & \text{if } x \in (-\infty, L] \times (\mathbb{R} \setminus [0, 1]), \\
u^{(4)}(2L - x_1, x_2), & \text{if } x_1 > L.
\end{cases}
$$

Then by Lemma 2.3,

$$
E(u^{(4)}) \leq c \left(\mu\theta^2 \ln\left(3 + \frac{1}{\theta}\right) + \varepsilon L \right) \leq c\mu\theta^2 \ln\left(3 + \frac{1}{\theta^2}\right)
$$

which concludes the proof of (2.3).
c) If $\mu\theta^2\varepsilon^{-1} \in (1, L)$, we use $v_{1,\theta}$ as above and Lemma 2.1 with $\alpha := \mu\theta^2\varepsilon^{-1}$, $\beta := \alpha + \tilde{\beta}$ with $\tilde{\beta} \geq \alpha$ chosen

below, and $\bar{L} := \max\{L+1, \beta+1\}$. Precisely, we set with $\tilde{v}_{\tilde{\beta}}$ from Lemma 2.5 and $v_{1,\theta}$ from Lemma 2.3

$$
U^{(4)}(x) :=
\begin{cases}
\theta x_2, & \text{if } x \in [0, L] \times [0, 1-\theta], \\
(1-\theta)(1-x_2), & \text{if } x \in [0, \alpha] \times (1-\theta, 1], \\
\tilde{v}_{\tilde{\beta}}(x_1 - \alpha, x_2 - (1-\theta)) + \theta(1-\theta), & \text{if } x \in (\alpha, \alpha + \tilde{\beta}] \times (1-\theta, 1], \\
\theta x_2, & \text{if } x \in (\alpha + \tilde{\beta}, L] \times (1-\theta, 1], \\
v_{1,\theta}(x), & \text{if } x \in (-1, 0) \times [0, 1], \\
u_{\alpha+1, \beta+1, \bar{L}}(x_1 + 1, x_2), & \text{if } x \in ((-\infty, L] \times \mathbb{R}) \setminus ((-1, L] \times [0, 1]), \\
U^{(4)}(2L - x_1, x_2), & \text{if } x_1 > L.
\end{cases}
$$

One readily checks that $U^{(4)}$ is continuous. By Lemmas 2.1 and 2.3, we have

$$
E^{\text{ext}}(U^{(4)}) \le c\mu\theta^2 \left(\ln\left(3 + \frac{L + \alpha + \tilde{\beta} + 1}{\alpha}\right) + \frac{\beta + \alpha}{\beta - \alpha} + \ln\left(3 + \frac{1}{\theta^2}\right) \right).
$$

Altogether, we obtain using Lemma 2.5 and $\beta \ge 2\alpha$

$$
\begin{aligned}
E(U^{(4)}) &\le c\left(\frac{\theta^3}{\tilde{\beta}} + \varepsilon\tilde{\beta} + \frac{\varepsilon\theta^2}{\tilde{\beta}}\right) + c\mu\theta^2 \ln\left(3 + \frac{L}{\alpha} + \frac{\tilde{\beta}}{\alpha}\right) + c\mu\theta^2 \ln\left(3 + \frac{1}{\theta^2}\right) + \varepsilon \\
&\le c\left(\frac{\theta^3}{\tilde{\beta}} + \varepsilon\tilde{\beta}\right) + c\mu\theta^2 \ln\left(3 + \frac{\varepsilon L}{\mu\theta^2} + \frac{\tilde{\beta}}{\alpha}\right) + c\mu\theta^2 \ln\left(3 + \frac{1}{\theta^2}\right)
\end{aligned}
$$

where in the second step we used that $\tilde{\beta} \ge 1$ implies $\varepsilon\theta^2/\tilde{\beta} \le \varepsilon\tilde{\beta}$, and the assumption $\varepsilon \le \mu\theta^2$. At this point we distinguish two further subcases. If $\mu\theta^2 \le \varepsilon^{1/2}\theta^{3/2}$ then we set $\tilde{\beta} := \varepsilon^{-1/2}\theta^{3/2} \ge \mu\theta^2\varepsilon^{-1} = \alpha$ and obtain

$$
E(U^{(4)}) \le c\left(\varepsilon^{1/2}\theta^{3/2} + \mu\theta^2 \ln\left(3 + \frac{\varepsilon L}{\mu\theta^2} + \frac{\varepsilon^{1/2}\theta^{3/2}}{\mu\theta^2}\right)\right) + c\mu\theta^2 \ln\left(3 + \frac{1}{\theta^2}\right).
$$

We treat the first logarithm using $\ln(3 + x + y) \le \ln(3 + x) + \ln(1 + y) \le \ln(3 + x) + y$ for $x, y \ge 0$, leading to

$$
E(U^{(4)}) \le c\left(\varepsilon^{1/2}\theta^{3/2} + \mu\theta^2 \ln\left(3 + \frac{\varepsilon L}{\mu\theta^2}\right)\right) + c\mu\theta^2 \ln\left(3 + \frac{1}{\theta^2}\right),
$$

which concludes the proof of (2.3). If instead $\mu\theta^2 > \varepsilon^{1/2}\theta^{3/2}$ then we set $\tilde{\beta} := \alpha \ge \varepsilon^{-1/2}\theta^{3/2}$ and obtain

$$
E(U^{(4)}) \le c\left(\varepsilon^{1/2}\theta^{3/2} + \mu\theta^2 \ln\left(3 + \frac{\varepsilon L}{\mu\theta^2}\right)\right) + c\mu\theta^2 \ln\left(3 + \frac{1}{\theta^2}\right).
$$

(v) Single truncated branching: we aim to show that $\min_u E(u) \le c\left(\mu\theta^2 \ln(3 + \frac{\varepsilon L}{\mu\theta^2}) + \mu\theta^2 \ln(3 + \frac{\varepsilon}{\mu^2\theta^2}) + \varepsilon^{1/2}\theta^{3/2}\right)$. Again, we distinguish several cases.

a) If $\mu\theta^2\varepsilon^{-1} \le 1$, this follows from (ii).

b) If $\mu^2\theta^2\varepsilon^{-1} < \theta$, then this follows from (iv).

c) If $\mu^2\theta^2\varepsilon^{-1} \ge \theta$ and $\mu\theta^2\varepsilon^{-1} > 1$, we use the truncated branching construction (see Fig. 5) from the proofs of Theorems 3.1, 3.2 in [53], in the version of Lemma 2.4 and $v_{1,h}$ from Lemma 2.3. Precisely, we choose
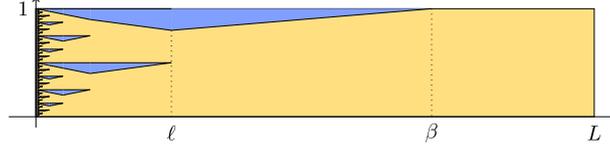
FIGURE 5. Sketch of the single truncated branching construction. We refer to Figure 2 for details of the branching construction on the left, to Figure 4 for the field lines in the austenite.

$h := \min\{1, \mu^2\theta^2\varepsilon^{-1}\} \in [\theta, 1]$, $N := 1$, and $\ell := \mu\theta^2\varepsilon^{-1} \geq 1 \geq h$ and set

$$u^{(5)}(x) := \begin{cases} \theta x_2, & \text{if } x \in [0, \min\{\ell, L\}] \times [0, 1-h], \\ u_{h,\ell,1}(x), & \text{if } x \in [0, \min\{\ell, L\}] \times (1-h, 1], \\ v_{1,h}(x), & \text{if } x \in [-1, 0) \times [0, 1], \end{cases} \tag{2.4}$$

which satisfies

$$E_{[-1,\ell)\times(0,1)}(w) \leq c\frac{\theta^2 h}{\ell} + c\varepsilon(\ell + h) + c\mu\theta^2 \ln\left(3 + \frac{1}{h}\right) \leq c\mu\theta^2 \ln(3 + \frac{\varepsilon}{\mu^2\theta^2}). \tag{2.5}$$

In the second inequality we used that $\frac{\theta^2 h}{\ell} = \frac{\varepsilon h}{\mu} \leq \mu\theta^2$ and $\ell \geq h$.

If $\ell \geq L/2$, we extend $u^{(5)}$ inside Ω_L by a simple laminate, *i.e.*,

$$u^{(5)}(x) := \begin{cases} \theta x_2, & \text{if } x \in (\ell, L] \times [0, 1-\theta], \\ (1-\theta)(1-x_2), & \text{if } x \in (\ell, L] \times (1-\theta, 1], \\ 0, & \text{if } x \in ((-\infty, L] \times \mathbb{R}) \setminus ([-1, L] \times [0,1]), \\ u^{(5)}(2L - x_1, x_2), & \text{if } x_1 > L. \end{cases}$$

Note that $(\ell, L] = \emptyset$ if $L < \ell$. We have $E(u^{(5)}) \leq c\left(E_{[-1,\ell)\times(0,1)}(u^{(5)}) + \varepsilon\ell\right)$, and the assertion follows. Otherwise, if $\ell < L/2$, we proceed as in (iv)c) using Lemma 2.5 and Lemma 2.1 with $\tilde{\beta} := \alpha := \ell$, $\beta := \ell + \tilde{\beta}$, and $\bar{L} := \max\{\beta + 1, L + 1\}$ and set

$$u^{(5)}(x) := \begin{cases} \theta x_2, & \text{if } x \in (\ell, L] \times [0, 1-\theta], \\ \tilde{v}_{\tilde{\beta}}(x_1 - \ell, x_2 - (1-\theta)) + \theta(1-\theta), & \text{if } x \in (\ell, \ell+\tilde{\beta}] \times (1-\theta, 1], \\ \theta x_2, & \text{if } x \in (\ell+\tilde{\beta}, L] \times (1-\theta, 1], \\ u_{\ell+1,\ell+\tilde{\beta}+1,\bar{L}}(x_1+1, x_2), & \text{if } x \in ((-\infty, L] \times \mathbb{R}) \setminus ((-1, L] \times [0,1]), \\ u^{(5)}(2L - x_1, x_2), & \text{if } x_1 > L. \end{cases}$$

This leads to

$$E(u^{(5)}) \leq c\left(\mu\theta^2 \ln\left(3 + \frac{\varepsilon}{\mu^2\theta^2}\right) + \varepsilon\ell + \frac{\theta^3}{\tilde{\beta}} + \varepsilon\tilde{\beta} + \frac{\varepsilon\theta^2}{\tilde{\beta}} + \mu\theta^2 \ln\left(3 + \frac{L+\ell+1}{\ell}\right) + \varepsilon\right),$$

and the assertion follows as in (iv)c).

(vi) Corner laminate: we show that

$$\min_u E(u) \le c\left(\mu\theta^2 \ln(3 + \frac{\varepsilon L}{\mu\theta^2}) + \mu\theta^2 \ln(3 + \frac{\theta}{\mu})\right).$$

Again, we distinguish several cases.

a) If $\varepsilon \ge \mu\theta^2$, then the assertion follows from (ii).

b) If $\theta \le \varepsilon < \mu\theta^2$, then $\frac{\varepsilon L}{\mu\theta^2} \ge \frac{L}{\mu\theta} > \frac{L}{\mu}$, and $\mu\theta^2 > \varepsilon > \varepsilon\theta$, and the assertion follows from (iii).

c) If $\theta/\mu \ge L$, this follows from (ii).

d) If $\mu^2\theta \le \varepsilon$, then $\frac{\varepsilon L}{\mu\theta^2}\frac{\theta}{\mu} \ge L$. The assertion follows from (ii) using that $\ln(3 + a) + \ln(3 + b) \ge \ln(3 + ab)$ for any $a, b \ge 0$.

e) It remains to consider the case that $\varepsilon < \min\{\mu\theta^2, \theta, \mu^2\theta\}$ and $\theta/\mu < L$. We choose $\gamma := \max\{\frac{1}{4}, \frac{\theta}{\mu \ln(3 + \frac{\theta}{\mu})}\}$ and note that $\gamma < L$ since $\theta/\mu < L$ and $L \ge \frac{1}{2}$. We first construct a function $\tilde{u}^{(6)}$ in $R_\gamma := [-\gamma, \gamma] \times [-\gamma, \gamma + 1]$,

$$\tilde{u}^{(6)}(x) := \begin{cases} \theta x_2, & \text{if } x_1 \in [0, \gamma],\ 0 \le x_2 \le 1 - \theta\frac{x_1}{\gamma}, \\[2mm] (1-\theta)(1-x_2) + \theta - \theta\dfrac{x_1}{\gamma}, & \text{if } x_1 \in [0, \gamma],\ 1 - \theta\frac{x_1}{\gamma} < x_2 \le 1, \\[2mm] \theta\left(1 - \dfrac{x_1 + x_2 - 1}{\gamma}\right), & \text{if } x_1 \in [0, \gamma],\ 1 < x_2 \le 1 + \gamma - x_1, & (2.6a) \\[2mm] \theta\left(1 - \dfrac{x_2 - 1}{\gamma}\right), & \text{if } x_1 \in [-\gamma, 0],\ 1 - x_1 \le x_2 \le \gamma + 1, & (2.6b) \\[2mm] \theta\left(1 + \dfrac{x_1}{\gamma}\right)\dfrac{x_2}{1 - x_1}, & \text{if } x_1 \in [-\gamma, 0] \text{ and } 0 \le x_2 < 1 - x_1, & (2.6c) \\[2mm] 0, & \text{elsewhere in } R_\gamma, \end{cases}$$

see Figure 6. One easily checks that $\tilde{u}^{(6)}$ is continuous and that

$$E_{(0,\gamma)\times(0,1)}(\tilde{u}^{(6)}) \le c\left(\frac{\theta^3}{\gamma} + \varepsilon(\gamma + \theta) + \frac{\varepsilon\theta^2}{\gamma}\right) \le c\left(\frac{\theta^3}{\gamma} + \varepsilon\gamma + \varepsilon\theta\right),$$

where we used in the last estimate that $\varepsilon \le \theta$. To estimate the energy outside Ω_{2L}, we observe that $|\nabla\tilde{u}^{(6)}| \le c\theta/\gamma$ in the parts given in (2.6a) and (2.6b), and $|\nabla\tilde{u}^{(6)}(x)| \le c\frac{\theta}{|1-x_1|}$ in the part given in (2.6c) (recall that $\gamma \ge 1/4$), which yields

$$E_{R_\gamma\setminus(0,\gamma)\times(0,1)}(\tilde{u}^{(6)}) \le c\mu\theta^2 \ln(3 + \gamma).$$

We then proceed as in (iv)c), using $\tilde{v}_{\tilde{\beta}}$ from Lemma 2.5 with $\tilde{\beta} := 3\mu\theta^2\varepsilon^{-1}$. Since $\varepsilon < \mu\theta^2$ we have $1 \le \tilde{\beta}$. We use Lemma 2.1 with $\alpha := \tilde{\beta}$, $\beta := \alpha + \tilde{\beta} = 2\alpha$ and $\bar{L} := \max\{\beta, L\}$ and define $u^{(6)}$ by

FIGURE 6. Construction for $\tilde{u}^{(6)}$. The left panel shows the construction in the martensite, the right panel the subdivision of the domain in the austenite phase.
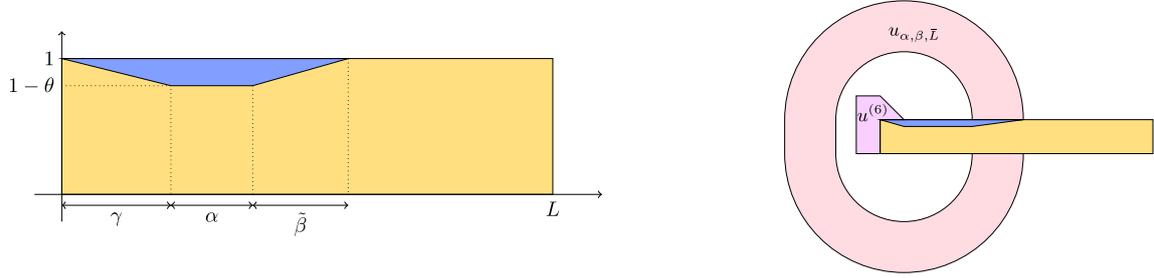

FIGURE 7. Corner-laminate construction for $u^{(6)}$. The left panel shows the construction in the martensitic region Ω_L, the right panel (on a different scale, and with different parameters) the construction in the austenite. The shaded regions are those where $\tilde{u}^{(6)} \neq 0$ and $u_{\alpha,\beta,\bar{L}}(x_1 - \gamma, x_2) \neq 0$, respectively (see also Fig. 6).

$$u^{(6)}(x) := \begin{cases} \tilde{u}^{(6)}(x), & \text{if } x \in R_\gamma, \\ \theta x_2, & \text{if } x \in (\gamma, \gamma + \alpha + \tilde{\beta}] \times [0, 1 - \theta], \\ (1 - \theta)(1 - x_2), & \text{if } x \in (\gamma, \gamma + \alpha] \times (1 - \theta, 1], \\ \tilde{v}_{\tilde{\beta}}(x_1 - (\gamma + \alpha), x_2 - (1 - \theta)) + \theta(1 - \theta), & \text{if } x \in (\gamma + \alpha, \gamma + \alpha + \tilde{\beta}] \times (1 - \theta, 1], \\ \theta x_2, & \text{if } x \in (\gamma + \alpha + \tilde{\beta}, L] \times [0, 1], \\ u_{\alpha,\beta,\bar{L}}(x_1 - \gamma, x_2), & \text{otherwise in } (-\infty, L] \times \mathbb{R}, \\ u^{(6)}(2L - x_1, x_2), & \text{if } x_1 > L, \end{cases}$$

see Figure 7. The condition $\varepsilon \leq \min\{\mu^2\theta, \mu\theta^2\}$ implies $3\gamma \leq \alpha$, so that the construction for $\tilde{u}^{(6)}$ and the one for $u_{\alpha,\beta,\bar{L}}(x_1 - \gamma, x_2)$ match continuously. The function $u^{(6)}$ is continuous and

$$E(u^{(6)}) \leq c \left(\frac{\theta^3}{\gamma} + \varepsilon\gamma + \varepsilon\theta + \mu\theta^2 \ln(3 + \gamma) + \varepsilon\alpha + \frac{\theta^3}{\tilde{\beta}} + \varepsilon \left(\tilde{\beta} + \frac{\theta^2}{\tilde{\beta}} \right) + \mu\theta^2 \ln \left(3 + \frac{\bar{L}}{\alpha} \right) + \mu\theta^2 \frac{\beta + \alpha}{\beta - \alpha} + \varepsilon \right)$$

$$\leq \mu\theta^2 \ln \left(3 + \frac{\theta}{\mu} \right) + \mu\theta^2 \ln \left(3 + \frac{\varepsilon L}{\mu\theta^2} \right).$$

We used here $\varepsilon\gamma \leq \max\{\varepsilon, \frac{\varepsilon\theta}{\mu}\} \leq \mu\theta^2$ since $\varepsilon \leq \min\{\mu\theta^2, \mu^2\theta\}$; and similarly $\varepsilon(1 + \theta + \alpha + \tilde{\beta}) \leq c\mu\theta^2$ and $\frac{\varepsilon\theta^2}{\tilde{\beta}} \leq \frac{\theta^3}{\tilde{\beta}} = \frac{\varepsilon\theta^3}{3\mu\theta^2} \leq \mu\theta^2$.

This concludes the proof of the upper bound. $\square$

2.2. Comments on the scaling law

The purpose of this subsection is two-fold: on the one hand, in Section 2.2.1 we shall prove that all terms in the definition of $\mathcal{I}(\mu, \varepsilon, \theta, L)$ are relevant in the sense that the statement is false if we remove one of them. Furthermore, we give some intuition on the constructions used in the proof of the upper bound. On the other hand, in Section 2.2.2, we shall explain the different parameter regimes and motivate why they are treated separately in the proof of the lower bound.

2.2.1. Do all regimes really exist?

We will use the following abbreviatory notation: We denote as *scaling* an expression like $\varepsilon^{1/2}\theta^{3/2}$, and as *regime* something like $\mu\theta^2 \ln(3 + \frac{L}{\mu}) + \varepsilon\theta$ (which is the sum of a few scalings). In particular, $\mathcal{I}$ is defined in Theorem 1.1 as the minimum of eight regimes.

We show below that no regime $R = R(\mu, \varepsilon, \theta, L)$ can be eliminated from the definition of $\mathcal{I}$ in Theorem 1.1. To do this, we shall exhibit a sequence of parameters $\mu_j, \varepsilon_j, \theta_j, L_j$ such that $\lim_{j\to\infty} \frac{\mathcal{I}_R(\mu_j, \varepsilon_j, \theta_j, L_j)}{R(\mu_j, \varepsilon_j, \theta_j, L_j)} \to \infty$, where $\mathcal{I}_R$ is the minimum in Theorem 1.1 without regime R.

Additionally, we show that no scaling can be eliminated in the regimes that consist of more than one scaling. Consider a regime R which consists of the scalings $S^{(k)}$, in the sense that $R = S^{(1)} + \cdots + S^{(K)}$ for $K \geq 2$. For any $i \in \{1, \ldots, K\}$ let $R_i := \sum_{k \neq i} S^{(k)}$ be the regime R without $S^{(i)}$. We shall provide a sequence $(\mu_j, \varepsilon_j, \theta_j, L_j)$ such that $\frac{R_i(\mu_j, \varepsilon_j, \theta_j, L_j)}{\mathcal{I}(\mu_j, \varepsilon_j, \theta_j, L_j)} \to 0$, proving that R cannot be replaced by R_i. In most cases, this will be done constructing a sequence with $\frac{\mathcal{I}_R}{R} \to \infty$, $\frac{S^{(i)}}{R} \to 1$ and $\frac{S^{(i)}}{S^{(k)}} \to 0$ if $k \neq i$ along that sequence, which additionally shows that the scaling $S^{(i)}$ dominates the regime R.

To briefly sketch the ideas behind the constructions in the proof of the upper bound, we describe them only inside the martensitic nucleus. They should be considered to be extended optimally (in the sense of trace) to the austenite part. The precise constructions and references to the literature are given in Section 2. We recall that we write $\ln^\alpha x$ for $(\ln x)^\alpha$, and the same for $\ln^\alpha \ln x$. We write $a_j \sim b_j$ if there is a constant $c > 0$ such that $\frac{1}{c} a_j \leq b_j \leq c a_j$ for all $j \in \mathbb{N}$.

i. $R = \theta^2 L$ (constant): this regime is attained by a constant test function, corresponding to austenite. This regime is the only one that does not depend on ε nor μ. We take $\theta_j = L_j = \frac{1}{2}$, $\varepsilon_j = \mu_j \to \infty$. Then all other regimes have diverging energy.

ii. $R = \mu\theta^2 \ln(3 + L)$ (affine): this regime is attained by using an affine function inside the nucleus corresponding to the majority variant of martensite (see Fig. 4 (left)). We take $\theta_j = \mu_j = \frac{1}{2}$, $L_j \to \infty$, $\varepsilon_j = e^{L_j}$. All regimes which contain one of the scalings εL, $\varepsilon\theta$, $\varepsilon^{1/2}\theta^{3/2}$, $\theta^2 L$ have energies which diverge at least as a power of L_j, and also $\mu_j\theta_j^2 \ln(3 + \frac{\varepsilon_j L_j}{\mu_j \theta_j^2}) \gg L_j$, and only $\mu\theta^2 \ln(3 + L)$ is logarithmic.

iii. $R = \mu\theta^2 \ln(3 + \frac{L}{\mu}) + \varepsilon\theta$ (linear interpolation). This regime is (when relevant) attained by a test function that is constant near the left and the right boundaries of Ω_{2L} (corresponding to austenite), and affine near the middle $\{x_1 = L\}$ of the nucleus (corresponding to the majority variant of martensite). There is a competition of the energy inside the nucleus, which favours the test function to be in the martensitic variant on a large part, and the energy contribution from the austenite part, which favours the function to be constant in a large neighbourhood of the left and right boundaries (see Fig. 4 (right)).

   (a) $S = \mu\theta^2 \ln(3 + \frac{L}{\mu})$: We take $L_j \to \infty$, $\theta_j = \frac{1}{2}$, $\mu_j = \frac{L_j}{\ln L_j}$, $\varepsilon_j = \mu_j\theta_j$. Then $\varepsilon_j\theta_j = \mu_j\theta_j^2$, $\frac{L_j}{\mu_j} = \ln L_j$, so that $R(\mu_j, \varepsilon_j, \theta_j, L_j) \sim S(\mu_j, \varepsilon_j, \theta_j, L_j) \sim L_j \frac{\ln\ln L_j}{\ln L_j}$, whereas $\theta_j^2 L_j \sim L_j$, $\mu_j\theta_j^2 \ln(3 + L_j) \sim L_j$, $\varepsilon_j L_j \sim L_j^2/\ln L_j$, and $\frac{\varepsilon_j L_j}{\mu_j\theta_j^2} = \frac{L_j}{\theta_j}$ implies $\mu_j\theta_j^2 \ln(3 + \frac{\varepsilon_j L_j}{\mu_j\theta_j^2}) \sim \mu_j \ln L_j \sim L_j$.

   (b) $S = \varepsilon\theta$: As above, we take $L_j \to \infty$, $\theta_j = \frac{1}{2}$, $\mu_j = \frac{L_j}{\ln L_j}$, but this time $\varepsilon_j = \mu_j\theta_j(\ln\ln L_j)^2$. Then $\frac{L_j}{\mu_j} = \ln L_j$, so that $R(\mu_j, \varepsilon_j, \theta_j, L_j) \sim S(\mu_j, \varepsilon_j, \theta_j, L_j) \sim \theta_j^2 L_j \frac{(\ln\ln L_j)^2}{\ln L_j}$, in the other terms the $\ln\ln L_j$ correction does not change the argument.

iv. $R = \mu\theta^2 \ln(3 + \frac{\varepsilon L}{\mu\theta^2}) + \mu\theta^2(3 + \frac{\varepsilon}{\mu^2\theta^2}) + \varepsilon^{1/2}\theta^{3/2}$ (single truncated branching). This regime is (when relevant) attained by a test function that consists of roughly three parts: close to the left and right boundaries of Ω_{2L}, a branching construction is used, which goes over to a single laminate, and then interpolates to an affine function (which corresponds to the majority variant of martensite) near the vertical middle $\{x_1 = L\}$ of the nucleus, see Figure 5.

(a) $S = \varepsilon^{1/2}\theta^{3/2}$: We take $\theta_j = \frac{1}{2}$, $L_j \to \infty$, $\mu_j = \frac{1}{L_j}$, $\varepsilon_j = \frac{\mu_j\theta_j^2}{L_j}\ln L_j = \frac{\ln L_j}{4L_j^2}$. Then $\mu_j\theta_j^2 = \frac{1}{4L_j}$, $\varepsilon_j L_j = \frac{\ln L_j}{4L_j}$. In particular, $\theta_j^2 L_j \sim L_j$, $\mu_j\theta_j^2 \ln L_j \sim \mu_j\theta_j^2 \ln\frac{L_j}{\mu_j} \sim \frac{\ln L_j}{L_j}$. We have $\frac{\varepsilon_j L_j}{\mu_j\theta_j^2} = \ln L_j$, $\frac{\varepsilon_j}{\mu_j^2\theta_j^2} = \ln L_j$, and $S(\mu_j,\varepsilon_j,\theta_j,L_j) = \varepsilon_j^{1/2}\theta_j^{3/2} \sim \frac{1}{L_j}\ln^{1/2}L_j$, so that $R(\mu_j,\varepsilon_j,\theta_j,L_j) = O(\frac{1}{L_j}\ln\ln L_j) + S(\mu_j,\varepsilon_j,\theta_j,L_j) = O(\frac{1}{L_j}\ln\ln L_j) + \frac{1}{L_j}\ln^{1/2}L_j$, and $S(\mu_j,\varepsilon_j,\theta_j,L_j)/R(\mu_j,\varepsilon_j,\theta_j,L_j) \to 1$. All regimes which contain the scaling $\varepsilon_j L_j \sim \frac{\ln L_j}{L_j}$ can be ignored. Finally, $\theta_j/\mu_j = \theta_j L$, hence $\mu_j\theta_j^2 \ln\frac{\theta_j}{\mu_j} \sim \frac{1}{L_j}\ln L_j \gg S(\mu_j,\varepsilon_j,\theta_j,L_j) \sim R(\mu_j,\varepsilon_j,\theta_j,L_j)$. This concludes the proof.

(b) $S = \mu\theta^2\ln(3 + \frac{\varepsilon L}{\mu\theta^2})$: We take $\theta_j = \frac{1}{2}$, $L_j \to \infty$, $\mu_j = \frac{1}{L_j^{1/2}}$, $\varepsilon_j = \frac{\mu_j\theta_j^2}{L_j}\ln L_j = \frac{1}{4L_j^{3/2}}\ln L_j$. Then $\mu_j\theta_j^2 = \frac{1}{4L_j^{1/2}}$, $\varepsilon_j L_j = \frac{1}{4L_j^{1/2}}\ln L_j$, $\frac{\varepsilon_j L_j}{\mu_j\theta_j^2} = \ln L_j$, $\frac{\varepsilon_j}{\mu_j^2\theta_j^2} = \frac{1}{L_j^{1/2}} \to 0$, and $\varepsilon_j^{1/2} \sim \frac{1}{L_j^{3/4}}\ln^{1/2}L_j$. Therefore $S(\mu_j,\varepsilon_j,\theta_j,L_j)$ dominates $R(\mu_j,\varepsilon_j,\theta_j,L_j)$, and $S(\mu_j,\varepsilon_j,\theta_j,L_j) \sim R(\mu_j,\varepsilon_j,\theta_j,L_j) \sim \frac{1}{L_j^{1/2}}\ln\ln L_j$. All regimes with $\varepsilon_j L_j$ are higher, as is obviously $\theta_j^2 L_j$. Further, $\frac{\theta_j}{\mu_j} \sim L_j^{1/2}$ implies $\mu_j\theta_j^2 \ln(3 + \frac{\theta_j}{\mu_j}) \sim \frac{1}{L_j^{1/2}}\ln L_j$, and finally, $\mu_j\theta_j^2 \ln L_j \sim \mu_j\theta_j^2 \ln\frac{L_j}{\mu_j} \sim \frac{1}{L_j^{1/2}}\ln L_j \gg R(\mu_j,\varepsilon_j,\theta_j,L_j))$.

(c) $S = \mu\theta^2\ln(3 + \frac{\varepsilon}{\mu^2\theta^2})$: We take $L_j \to \infty$, $\theta_j = \frac{1}{\ln^{2/5}L_j}$, $\mu_j = \frac{1}{L_j\ln^{1/5}L_j}$, $\varepsilon_j = \frac{\ln^5\ln L_j}{L_j^2\ln L_j}$. Then $\varepsilon_j L_j = \frac{\ln^5\ln L_j}{L_j\ln L_j}$, $\mu_j\theta_j^2 = \frac{\ln\ln L_j}{L_j\ln L_j}$, $(\varepsilon_j L_j\mu_j\theta_j^2)^{1/2} = \mu_j\theta_j^2\ln^2\ln L_j$. Further, $\frac{\varepsilon_j L_j}{\mu_j\theta_j^2} = \ln^4\ln L_j$, $\frac{\varepsilon_j}{\mu_j^2\theta_j^2} = \ln^{1/5}L_j\ln^3\ln L_j$, and $S(\mu_j,\varepsilon_j,\theta_j,L_j) \sim \mu_j\theta_j^2\ln\ln L_j \gg \mu_j\theta_j^2\ln\frac{\varepsilon_j L_j}{\mu_j\theta_j^2} \sim \mu_j\theta_j^2\ln\ln\ln L_j$. For the third scaling in this regime, $\varepsilon_j^{1/2}\theta_j^{3/2} = \frac{\ln^{5/2}\ln L_j}{L_j\ln^{1/2}L_j\ln^{3/5}L_j} = \mu_j\theta_j^2\frac{\ln^{3/2}\ln L_j}{\ln^{1/10}L_j} \ll S(\mu_j,\varepsilon_j,\theta_j,L_j)$. For the corner laminate regime, we estimate $\frac{\theta_j}{\mu_j} = \frac{L_j}{\ln^{1/5}L_j\ln\ln L_j}$ which gives $\mu_j\theta_j^2\ln(3 + \frac{\theta_j}{\mu_j}) \sim \mu_j\theta_j^2\ln L_j \gg S(\mu_j,\varepsilon_j,\theta_j,L_j)$. The other regimes are simpler. $\theta_j^2 L_j$ is linear in L_j, $\mu_j\theta_j^2\ln(3 + L_j)$ and $\mu\theta^2\ln(3 + \frac{L}{\mu})$ behave as $\mu_j\theta_j^2\ln L_j$, which is much larger than $S(\mu_j,\varepsilon_j,\theta_j,L_j) \sim \mu_j\theta_j^2\ln\ln L_j$, and the last ones (branching, laminate and two-scale-branching) are eliminated by $\varepsilon_j L_j$.

v. $R = \mu\theta^2\ln(3 + \frac{\theta}{\mu}) + \mu\theta^2\ln\left(3 + \frac{\varepsilon L}{\mu\theta^2}\right)$ (corner laminate): this scaling is (when relevant) attained by a construction sketched in Figure 7 (left). Note that this leads to two relevant contributions from the austenite part as sketched in Figure 7 (right).

(a) $S = \mu\theta^2\ln(3 + \frac{\theta}{\mu})$: We take $L_j \to \infty$, $\theta_j = \frac{1}{L_j^2}$, $\mu_j = \frac{1}{L_j^2\ln L_j}$, $\varepsilon_j = \frac{\ln^4\ln L_j}{L_j^7\ln L_j}$. Then $\frac{\theta_j}{\mu_j} = \ln L_j$, $\mu_j\theta_j^2 = \frac{1}{L_j^6\ln L_j}$, $\varepsilon_j L_j = \frac{\ln^4\ln L_j}{L_j^6\ln L_j}$. In particular, $\ln\frac{\varepsilon_j L_j}{\mu_j\theta_j^2} \sim \ln\ln\ln L_j \ll \ln\frac{\theta_j}{\mu_j} = \ln\ln L_j$, and $S(\mu_j,\varepsilon_j,\theta_j,L_j) \sim \frac{1}{L_j^6\ln L_j}\ln\ln L_j$. Therefore $S(\mu_j,\varepsilon_j,\theta_j,L_j)$ dominates $R(\mu_j,\varepsilon_j,\theta_j,L_j)$. To eliminate the other regimes, we observe that $\frac{\varepsilon_j}{\mu_j^2\theta_j^2} = L_j(\ln L_j)(\ln^4\ln L_j)$ shows that $\mu_j\theta_j^2\ln(3 + \frac{\varepsilon_j}{\mu_j^2\theta_j^2})/S(\mu_j,\varepsilon_j,\theta_j,L_j) \to \infty$. Since $\varepsilon_j L_j = \frac{\ln^4\ln L_j}{L_j^6\ln L_j} \gg S(\mu_j,\varepsilon_j,\theta_j,L_j)$, branching, laminates and two-scale branching are ruled out. Since $\mu_j\theta_j^2\ln L_j \sim \mu_j\theta_j^2 L_j\ln\frac{L_j}{\mu_j} \sim \frac{1}{L_j^6} \gg S(\mu_j,\varepsilon_j,\theta_j,L_j)$, and $\theta_j^2 L_j = \frac{1}{L_j^3}$, all remaining regimes are eliminated.

(b) $S = \mu\theta^2\ln(3 + \frac{\varepsilon L}{\mu\theta^2})$. We take $L_j \to \infty$, $\theta_j = \mu_j = \frac{1}{L_j^2}$, $\varepsilon_j = \frac{\ln L_j}{L_j^7}$. Then $\frac{\theta_j}{\mu_j} = 1$, $\mu_j\theta_j^2 = \frac{1}{L_j^6}$, $\varepsilon_j L_j = \frac{\ln L_j}{L_j^6}$. In particular, $\ln\frac{\varepsilon_j L_j}{\mu_j\theta_j^2} \sim \ln\ln L_j \gg \ln(3 + \frac{\theta_j}{\mu_j}) = \ln 4$, and $S(\mu_j,\varepsilon_j,\theta_j,L_j) \sim \frac{1}{L_j^6}\ln\ln L_j$. Therefore $S(\mu_j,\varepsilon_j,\theta_j,L_j)$ dominates $R(\mu_j,\varepsilon_j,\theta_j,L_j)$. To eliminate the other regimes, we observe that $\frac{\varepsilon_j}{\mu_j^2\theta_j^2} = $

$L_j \ln L_j$ shows that $\mu_j \theta_j^2 \ln(3 + \frac{\varepsilon_j}{\mu_j^2 \theta_j^2})/S(\mu_j, \varepsilon_j, \theta_j, L_j) \to \infty$. The bottom ones (branching, laminates and two-scale branching) are eliminated by $\varepsilon_j L_j/S(\mu_j, \varepsilon_j, \theta_j, L_j) \sim \frac{\ln L_j}{\ln \ln L_j} \to \infty$, the top ones (constant, affine and linear interpolation) by $\mu_j \leq 1$ which implies $\min\{\theta_j^2 L, \mu_j \theta_j^2 \ln(3 + \frac{L_j}{\mu_j})\} \geq \mu_j \theta_j^2 \ln L_j \sim \frac{\ln L_j}{L_j^6} \gg S(\mu_j, \varepsilon_j, \theta_j, L_j)$.

vi. $R = \varepsilon^{2/3} \theta^{2/3} L^{1/3} + \varepsilon L$ (branching): This regime is (when relevant) attained by a branching construction sketched in Figure 3 (left).

(a) $S = \varepsilon^{2/3} \theta^{2/3} L^{1/3}$. We take $\theta_j = \mu_j = L_j = \frac{1}{2}$, $\varepsilon_j = \frac{1}{j} \to 0$. Only the last three regimes (branching, laminates and two-scale branching) have infinitesimal energy. We have $\varepsilon_j^{2/3} \theta_j^{2/3} L_j^{1/3} \sim j^{-2/3}$, whereas $\mu_j^{1/2} \varepsilon_j^{1/2} \theta_j L_j^{1/2} \sim j^{-1/2}$. At the same time, $\varepsilon_j^{2/3} \gg \varepsilon_j$, and hence $S(\mu_j, \varepsilon_j, \theta_j, L_j)$ dominates $R(\mu_j, \varepsilon_j, \theta_j, L_j)$.

(b) $S = \varepsilon L$. We take $\theta_j = \mu_j = \frac{1}{2}$, $L_j = j^{2/3} \to \infty$, $\varepsilon_j = \frac{1}{j} \to 0$. Then $\varepsilon_j L_j = j^{-1/3} \to 0$, $\varepsilon_j^{2/3} L_j^{1/3} = j^{-4/9} \ll j^{-1/3}$, $\mu_j \theta_j^2 = \frac{1}{8}$, $\mu_j^{1/2} \varepsilon_j^{1/2} \theta_j L_j^{1/2} \sim j^{-1/6} \gg j^{-1/3}$.

vii. $R = \mu^{1/2} \varepsilon^{1/2} \theta L^{1/2} \ln^{1/2}(3 + \frac{1}{\theta^2}) + \varepsilon L$ (laminate): this regime is attained by a laminate construction as sketched in Figure 3 (middle).

(a) $S = \mu^{1/2} \varepsilon^{1/2} \theta L^{1/2} \ln^{1/2}(3 + \frac{1}{\theta^2})$: We take $L_j \to \infty$, $\theta_j = \frac{1}{L_j}$, $\mu_j = L_j^2 e^{-L_j}$, $\varepsilon_j = \frac{1}{L_j^2} e^{-L_j}$. Then $\theta_j^2 L_j = \frac{1}{L_j}$, $\mu_j \theta_j^2 = e^{-L_j}$, $\varepsilon_j L_j = \frac{1}{L_j} e^{-L_j}$, and $\mu_j^{1/2} \varepsilon_j^{1/2} \theta_j L_j^{1/2} = \frac{1}{L_j^{1/2}} e^{-L_j}$. We compute in detail the last two regimes. Since $\ln(3 + \frac{1}{\theta_j^2}) \sim \ln L_j$, we have $S(\mu_j, \varepsilon_j, \theta_j, L_j) \sim \frac{1}{L_j^{1/2}} e^{-L_j} \ln L_j$ and $\varepsilon_j L_j/S(\mu_j, \varepsilon_j, \theta_j, L_j) \to 0$. Since $\ln(3 + \frac{\varepsilon_j}{\mu_j^3 L_j \theta_j^2}) = \ln(3 + \frac{e^{2L_j}}{L_j^7}) \sim L_j$, we have $\mu_j^{1/2} \varepsilon_j^{1/2} \theta_j L_j^{1/2} \ln^{1/2}(3 + \frac{\varepsilon_j}{\mu_j^3 L_j \theta_j^2}) \sim e^{-L_j}$. With $\mu_j \theta_j^2/S(\mu_j, \varepsilon_j, \theta_j, L_j) \to \infty$ the proof is concluded.

(b) $S = \varepsilon L$: We take $L_j \to \infty$, $\theta_j = \frac{1}{\ln^{1/2} L_j}$, $\mu_j = \frac{1}{L_j^{3/2}}$, $\varepsilon_j = \frac{1}{L_j^{5/2} \ln^{1/2} L_j}$. Then $\mu_j \theta_j^2 = \frac{1}{L_j^{3/2} \ln L_j}$, $\varepsilon_j L_j = \frac{1}{L_j^{3/2} \ln^{1/2} L_j}$, and $\mu_j^{1/2} \varepsilon_j^{1/2} \theta_j L_j^{1/2} = \frac{1}{L_j^{3/2} \ln^{3/4} L_j}$. We compute in detail the last two regimes. Since $\ln(3 + \frac{1}{\theta_j^2}) \sim \ln \ln L_j$, we have $\mu_j^{1/2} \varepsilon_j^{1/2} \theta_j L_j^{1/2} \ln^{1/2}(3 + \frac{1}{\theta_j^2}) \sim \frac{1}{L_j^{3/2} \ln^{3/4} L_j} \ln^{1/2} \ln L_j \ll \varepsilon_j L_j$. Since $\ln(3 + \frac{\varepsilon_j}{\mu_j^3 L_j \theta_j^2}) = \ln(3 + L_j \ln^{1/2} L_j) \sim \ln L_j$, we have $\mu_j^{1/2} \varepsilon_j^{1/2} \theta_j L_j^{1/2} \ln^{1/2}(3 + \frac{\varepsilon_j}{\mu_j^3 L_j \theta_j^2}) \sim \frac{1}{L_j^{3/2} \ln^{1/4} L_j} \gg \varepsilon_j L_j$. Further, $\theta_j^2 L_j \sim L_j/\ln L_j$, $\mu \leq 1$, and the three terms $\ln(3 + L_j)$, $\ln(3 + \frac{\varepsilon_j}{\mu_j^2 \theta_j^2})$ and $\ln(3 + \frac{\theta_j}{\mu_j})$ behave as $\ln L_j$, eliminating the first five regimes (constant, affine, linear interpolation, single truncated branching and corner laminate).

viii. $R = \mu^{1/2} \varepsilon^{1/2} \theta L^{1/2} \ln^{1/2}(3 + \frac{\varepsilon}{\mu^3 \theta^2 L}) + \varepsilon L$ (two-scale branching): this regime is (when relevant) attained by a two-scale branching construction sketched in Figure 3 (right).

(a) $S = (\mu \theta^2 \varepsilon L \ln(3 + \frac{\varepsilon}{\mu^3 \theta^2 L}))^{1/2}$: We take $L_j \to \infty$, $\theta_j = \frac{1}{L_j}$, $\mu_j = \frac{1}{L_j^{5/2}}$, $\varepsilon_j = \mu_j^3 \theta_j^2 L_j \ln L_j = \frac{1}{L_j^{8 + \frac{1}{2}}} \ln L_j$. Then $\mu_j \theta_j^2 = \frac{1}{L_j^{4 + \frac{1}{2}}}$, $\varepsilon_j L_j = \frac{1}{L_j^{7 + \frac{1}{2}}} \ln L_j$, $\varepsilon_j^{2/3} \theta_j^{2/3} L_j^{1/3} = \frac{1}{L_j^6} \ln^{2/3} L_j$, $\mu_j^{1/2} \varepsilon_j^{1/2} \theta_j L_j^{1/2} = \frac{1}{L_j^6} \ln^{1/2} L_j$, $\frac{\varepsilon}{\mu_j^3 \theta_j^2 L_j} = \ln L_j$, so that $S(\mu_j, \varepsilon_j, \theta_j, L_j) \sim \frac{1}{L_j^6} \ln^{1/2} L_j \ln^{1/2} \ln L_j$ and $\varepsilon_j L_j/S(\mu_j, \varepsilon_j, \theta_j, L_j) \to 0$. The laminate is eliminated by $(\mu_j \theta_j^2 \varepsilon L_j)^{1/2} \ln^{1/2}(3 + \frac{1}{\theta_j^2}) \sim \frac{1}{L_j^6} \ln L_j \gg S(\mu_j, \varepsilon_j, \theta_j, L_j)$, those with $\mu_j \theta_j^2$ by $\mu_j \theta_j^2/S(\mu_j, \varepsilon_j, \theta_j, L_j) \to \infty$, and $\theta_j^2 L_j = \frac{1}{L_j}$.

(b) $S = \varepsilon L$: We take $L_j \to \infty$, $\theta_j = \frac{1}{\ln^{1/5} L_j}$, $\mu_j = \frac{1}{L_j \ln L_j}$, $\varepsilon_j = \frac{1}{L_j^2 \ln^{3/5} L_j}$. Then $\mu_j \theta_j^2 = \frac{1}{L_j \ln^{7/5} L_j}$, $\varepsilon_j L_j = \frac{1}{L_j \ln^{3/5} L_j} = \mu_j \theta_j^2 \ln^{4/5} L_j$, $\frac{\varepsilon_j}{\mu_j^3 \theta_j^2 L_j} = \ln^{14/5} L_j$, which implies that $S_{TSB} := \mu_j^{1/2} \varepsilon_j^{1/2} \theta_j L_j^{1/2} \ln^{1/2}(3 + \frac{\varepsilon}{\mu^3 \theta^2 L}) \sim (\mu_j \theta_j^2 \varepsilon_j L_j)^{1/2} \ln^{1/2} \ln L_j \sim \frac{\ln^{1/2} \ln L_j}{L_j \ln L_j} \sim \mu_j \theta_j^2 (\ln^{2/5} L_j)(\ln^{1/2} \ln L_j) \ll \varepsilon_j L_j$. To conclude, we need to check that for all regimes R entering $\mathcal{I}$ we have $S_{TSB}/R \to 0$. This is obvious for $\theta_j^2 L_j$, for

$\mu_j \theta_j^2 \ln L_j$, and for all regimes that contain an $\varepsilon_j L_j$ scaling. Since $\mu_j \leq 1$, the regime with $\ln(3 + L_j/\mu_j)$ is also irrelevant. Since $\varepsilon_j^{1/2} \theta_j^{3/2} = \varepsilon_j L_j$, this is also true for the regimes that contain the $\varepsilon_j^{1/2} \theta_j^{3/2}$ scaling, and finally $\frac{\theta_j}{\mu_j} = L_j \ln^{4/5} L_j$ shows that this also holds true for the corner laminate.

2.2.2. Rough overview over some parameter ranges

The proof of the lower bound in Section 3 is split into several parts that address different parameter ranges. We shall briefly motivate and sketch heuristically why different behaviours are expected in the considered ranges, and how this is reflected in our scaling law.

(i) We first consider the range in which ε is not so small, in the sense that $\varepsilon \geq \min\{\theta^2, \mu\theta^2\}$. This is the range considered in Section 3.1. Roughly speaking, interfacial energy is expensive, and one expects rather uniform structures. Note that in the scaling regimes, the 'uniform' constructions of constant functions (austenite, $L\theta^2$) and affine functions (majority variant of martensite, $\mu\theta^2 \ln(3 + L)$) scale differently in the size of the nucleus L. Hence, comparing these two regimes leads for large μ to a competition between μ and L.

(a) If $\mu < 1$ then elastic strain in the austenite part is more favourable than elastic strain in the martensite part. Further, $\varepsilon \geq \min\{\mu\theta^2, \theta^2\} = \mu\theta^2$ implies that also interfacial energy is expensive compared to elastic energy in the austenite part. Therefore, one would expect that low energy configurations behave roughly like affine functions (corresponding to the majority variant of martensite) inside the nucleus. This is reflected in our scaling law: Since $\mu\theta^2 \ln(3 + L) \leq \mu\theta^2 \ln(3 + \frac{L}{\mu}) \leq c\mu\theta^2(3 + \frac{L}{\mu}) \leq c(\mu\theta^2 + L\theta^2) \leq cL\theta^2$, $\frac{\varepsilon L}{\mu\theta^2} \geq L$, and $\varepsilon L \geq \mu\theta^2 L \geq c\mu\theta^2 \ln(3 + L)$, we obtain $\mathcal{I}(\mu, \varepsilon, \theta, L) \sim \mu\theta^2 \ln(3 + L)$, which corresponds to the affine test function.

(b) If $\mu \geq 1$, then the behaviour is different, and the above mentioned competition between L and μ becomes relevant. Note that $\mu \geq 1$ and $\varepsilon \geq \min\{\mu\theta^2, \theta^2\} = \theta^2$ implies that $\varepsilon L \geq \theta^2 L$, and hence all the branching and laminate regimes with a scaling εL are not relevant. To see that the regimes with three scalings are not relevant is more complicated: We always have $\frac{\varepsilon L}{\mu\theta^2} \geq \frac{L}{\mu}$. Hence, if $\mu\theta^2 \ln(3 + \frac{L}{\mu}) \gtrsim \varepsilon\theta$, we have $\mu\theta^2 \ln(3 + \frac{\varepsilon L}{\mu\theta^2}) \geq \mu\theta^2 \ln(3 + \frac{L}{\mu}) \gtrsim \mu\theta^2 \ln(3 + \frac{L}{\mu}) + \varepsilon\theta$, and we are done. Otherwise, $\varepsilon\theta \gtrsim \mu\theta^2 \ln(3 + \frac{L}{\mu})$ implies $\varepsilon \gtrsim \mu\theta^2$ and hence $\frac{\varepsilon L}{\mu\theta^2} \gtrsim L$, which yields $\mu\theta^2 \ln(3 + \frac{\varepsilon L}{\mu\theta^2}) \gtrsim \mu\theta^2 \ln(3 + L)$. Summarizing, we obtain

$$\mathcal{I}(\mu, \varepsilon, \theta, L) \sim \min\left\{\theta^2 L, \mu\theta^2 \ln(3 + L), \mu\theta^2 \ln\left(3 + \frac{L}{\mu}\right) + \varepsilon\theta\right\}.$$

The examples given in Section 2.2.1 show that all scalings are relevant in this parameter range.

(ii) The parameter range $\varepsilon \leq \min\{\theta^2, \mu\theta^2\}$ is more delicate since here many contributions compete. Note that in this range, the scaling $\theta^2 L$ is not relevant since $4\theta^2 L \geq \varepsilon^{2/3}\theta^{2/3}L^{1/3} + \varepsilon L$ (recall that $L \geq 1/2$). Also the regime $\mu\theta^2 \ln(3 + \frac{L}{\mu}) + \varepsilon\theta$ does not occur: Indeed, if $\mu \leq 1$ then $\mu\theta^2 \ln(3 + L) \leq \mu\theta^2 \ln(3 + \frac{L}{\mu})$. If $\mu > 1$, then $\varepsilon^{1/2}\theta^{3/2} \leq \theta^2 \leq \mu\theta^2$, $\frac{\varepsilon L}{\mu\theta^2} \leq \frac{L}{\mu}$ and $\frac{\varepsilon}{\mu^2\theta^2} \leq \frac{1}{\mu^2} \leq 1$. Summarizing, $4\mu\theta^2 \ln(3 + \frac{L}{\mu}) \geq \min\{\mu\theta^2 \ln(3 + L), \mu\theta^2 \ln(3 + \frac{\varepsilon L}{\mu\theta^2}) + \mu\theta^2 \ln(3 + \frac{\varepsilon}{\mu^2\theta^2}) + \varepsilon^{1/2}\theta^{3/2}\}$, and hence $\mu\theta^2 \ln(3 + \frac{L}{\mu}) + \varepsilon\theta$ does not occur.

In this parameter range, the main difficulty lies in the logarithmic corrections. Roughly speaking, complex patterns and rather uniform structures can occur, and the overall behaviour is mainly determined by the comparison of εL and several scalings with $\mu\theta^2$. The latter, however, contain logarithmic corrections that make the comparison rather involved and lead to mixtures of different constructions. There are mainly two qualitatively different reasons for the logarithmic terms: Some of them arise (rather locally) for laminated structures in the vicinity of the left and right boundaries of the nucleus (see Lem. 2.3). Others are due to the fact that in long nuclei affine structure deep inside the nucleus lead to non-periodic boundary conditions at the top and bottom boundaries of the nucleus and hence to elastic strain in the austenite (see Lem. 2.1). To indicate the different phenomena, we consider several subcases, corresponding to the competition between $\varepsilon^{2/3}\theta^{2/3}L^{1/3}$ and εL, and the size of μ.

(a) Assume $\varepsilon \geq \frac{\theta^2}{L^2}$. For these rather large values of ε, one expects that the relevant structures are rather uniform with few horizontal interfaces passing through the whole nucleus. This behaviour is reflected in our scaling law as follows: We have $\varepsilon L \geq \varepsilon^{2/3}\theta^{2/3}L^{1/3}$ which means that the branching regime behaves as εL and that the laminate and two-scale branching regimes are not relevant. Since $\left(\frac{\varepsilon}{\mu^2\theta^2}\right)^{1/2} \leq \frac{\varepsilon L}{\mu\theta^2}$, the single truncated branching regime reduces to $\mu\theta^2 \ln(3 + \frac{\varepsilon L}{\mu\theta^2}) + \varepsilon^{1/2}\theta^{3/2}$. The corner laminate regime can then be removed. Indeed, if $\varepsilon \leq \mu^2\theta$ then $\varepsilon^{1/2}\theta^{3/2} \leq \mu\theta^2$. If $\mu^2\theta \leq \varepsilon$, then $L \leq \frac{\theta}{\mu}\frac{\varepsilon L}{\mu\theta^2}$ implies $\ln(3 + L) \leq \ln(3 + \frac{\theta}{\mu}) + \ln(3 + \frac{\theta}{\mu}\frac{\varepsilon L}{\mu\theta^2})$, for details see the proof of Theorem 2.7 (vi)d). Therefore,

$$\mathcal{I}(\mu, \varepsilon, \theta, L) \sim \min\left\{\mu\theta^2 \ln(3 + L),\ \mu\theta^2 \ln\left(3 + \frac{\varepsilon L}{\mu\theta^2}\right) + \varepsilon^{1/2}\theta^{3/2},\ \varepsilon L\right\}.$$

The corresponding lower bound is the statement of Proposition 3.8 with the additional assumption $\varepsilon L^2 \geq \theta^2$. Using $\varepsilon L \geq \varepsilon^{1/2}\theta^{3/2}$, $\frac{\varepsilon L}{\mu\theta^2} \leq L$ and $\mu\theta^2 \ln(3 + \frac{\varepsilon L}{\mu\theta^2}) \leq \mu\theta^2 + \varepsilon L$, one can see that all the scalings are relevant.

(b) Assume $\varepsilon \leq \min\{\theta^2, \mu\theta^2, \theta^2/L^2\}$. Note that the condition $\varepsilon \leq \frac{\theta^2}{L^2}$ implies in particular $\varepsilon L \leq \theta^2/L$, i.e., roughly speaking, in the martensite part, a single laminate is cheaper than having a constant function or interpolating from a constant function at the left and right boundaries to an affine function deep inside the nucleus. The first two conditions indicate that interfacial energy is cheap compared to elastic energy in both, the austenite and the martensite part. However, there are various competitions between the interfacial energy, the elastic energies in the austenite and martensite parts, and the size of the nucleus. We shall outline the main points in these competitions by considering the cases $\mu \geq 1$ (i.e., elastic energy in the austenite is more expensive than in the martensite part), $\frac{1}{L} \leq \mu \leq 1$ (i.e., elastic energy in the austenite part is less expensive than in the martensite part but the nucleus is rather large), and the case $\mu \leq \frac{1}{L}$ (i.e., elastic energy in the austenite part is less expensive than in the martensite part and the nucleus is not so large).

– Assume $\mu \geq 1$. Then interfacial energy is rather cheap, the size of the nucleus is small in terms of ε, and elastic energy in the austenite part is rather expensive. Therefore, one expects that optimal configurations form complex microstructures inside the nucleus, with little strain the austenite part. This is reflected in our scaling law as follows: We have $\varepsilon L \leq \varepsilon^{2/3}\theta^{2/3}L^{1/3} \leq (\frac{\theta^2}{L^2})^{2/3}\theta^{2/3}L^{1/3} \lesssim \mu\theta^2$ (since $L \geq 1/2$ and $\mu \geq 1$), which shows that all regimes with a scaling $\mu\theta^2 \ln(3 + X)$ are irrelevant. Since also $\varepsilon^{2/3}\theta^{2/3}L^{1/3} \lesssim \mu^{1/2}\varepsilon^{1/2}\theta L^{1/2}$, also the laminate and two-scale branching regimes with a scaling $\mu^{1/2}\varepsilon^{1/2}\theta L^{1/2}\ln^{1/2}(3 + Y)$ are not relevant, and therefore we are in the branching regime of Figure 3(left),

$$\mathcal{I}(\mu, \varepsilon, \theta, L) \sim \varepsilon^{2/3}\theta^{2/3}L^{1/3}.$$

– Assume $\frac{1}{L} \leq \mu \leq 1$. The situation is similar to the case above. However, if the size L of the nucleus is large (in terms of μ), then one expects a competition between the formation of complex patterns inside the nucleus and elastic energy in the austenite part in the vicinity of the left and right boundaries: This is reflected in our scaling law as follows: Again, $\varepsilon L \leq \varepsilon^{2/3}\theta^{2/3}L^{1/3} \leq (\frac{\theta^2}{L^2})^{2/3}\theta^{2/3}L^{1/3} \lesssim \mu\theta^2$ implies that all regimes with a scaling $\mu\theta^2 \ln(3 + X)$ are not relevant. Furthermore, $\frac{\varepsilon}{\mu^3\theta^2 L} \leq \frac{1}{\theta^2}$ since $\mu^3 L \geq \frac{1}{L^2} \geq \frac{\theta^2}{L^2} \geq \varepsilon$, which means that two-scale branching is more favourable than laminates. Therefore, in this parameter range

$$\mathcal{I}(\mu, \varepsilon, \theta, L) \sim \min\left\{\varepsilon^{2/3}\theta^{2/3}L^{1/3}, \mu^{1/2}\varepsilon^{1/2}\theta L^{1/2}\ln^{1/2}(3 + \frac{\varepsilon}{\mu^3\theta^2 L}) + \varepsilon L\right\}.$$

Using that $\varepsilon^{2/3}\theta^{2/3}L^{1/3} \geq \mu^{1/2}\varepsilon^{1/2}\theta L^{1/2}\ln^{1/2}(3 + \frac{\varepsilon}{\mu^3\theta^2 L})$ is equivalent to $y\ln^{1/2}(3 + y) \geq 1$ for $y := \frac{\varepsilon}{\mu^3 L\theta^2}$, i.e., $y \geq c$, one easily checks that all three scalings are relevant.

– Assume finally $\mu \leq \frac{1}{L}$. This is the richest and most complex parameter range.

If ε is very small in the sense that additionally $\varepsilon \leq \frac{\mu^{3/2}\theta^2}{L^{1/2}} \leq \frac{\mu}{L}\theta^2$ then one expects the formation of complex patterns inside the nucleus. This is reflected in our scaling law as follows: On the one hand $\varepsilon L \leq \varepsilon^{2/3}\theta^{2/3}L^{1/3} \leq \mu\theta^2$, which shows that branching behaves as $\varepsilon^{2/3}\theta^{2/3}L^{1/3}$, and that all the regimes with a term $\mu\theta^2 \ln(3 + X)$ are not relevant. On the other hand, $\varepsilon L \leq \mu^{1/2}\varepsilon^{1/2}\theta L^{1/2}$. Summarizing, only branching, two-scale branching and laminates (see Fig. 3) are relevant, and

$$\mathcal{I}(\mu, \varepsilon, \theta, L) \sim \min\left\{ \varepsilon^{2/3}\theta^{2/3}L^{1/3}, \mu^{1/2}\varepsilon^{1/2}\theta L^{1/2}\ln^{1/2}\left(3 + \frac{\varepsilon}{\mu^3\theta^2 L}\right), \mu^{1/2}\varepsilon^{1/2}\theta L^{1/2}\ln^{1/2}\left(3 + \frac{1}{\theta^2}\right)\right\}.$$

As above, one easily checks that all the scalings are relevant. For the case of large $\theta \geq m_1$, the logarithms disappear since $\ln(3 + \frac{1}{\theta^2}) \leq c$, and the corresponding lower bound is proven in Lemma 3.4.

Let us finally address the remaining range $\frac{\mu^{3/2}\theta^2}{L^{1/2}} \leq \varepsilon \leq \min\{\mu\theta^2, \frac{\theta^2}{L^2}\}$ in which the overall behaviour is essentially determined by the logarithmic terms. Setting $y := \frac{\varepsilon}{L\mu^3\theta^2} \geq 1$, we have $y/\ln^3(3 + y) \geq c$ and hence $\varepsilon^{2/3}\theta^{2/3}L^{1/3} \geq c\mu^{1/2}\varepsilon^{1/2}\theta L^{1/2}\ln^{1/2}(3 + \frac{\varepsilon}{\mu^3\theta^2 L})$, which shows that branching is not relevant. We also note that $\frac{\varepsilon L}{\mu\theta^2} \leq \frac{\varepsilon}{\mu^2\theta^2}$. In this case,

$$\mathcal{I}(\theta, \varepsilon, L, \mu) = \min\left\{ \mu\theta^2\ln(3 + L), \mu\theta^2\ln(3 + \frac{\varepsilon}{\mu^2\theta^2}) + \varepsilon^{1/2}\theta^{3/2}, \mu\theta^2\ln(3 + \frac{\varepsilon L}{\mu\theta^2}) + \mu\theta^2\ln(3 + \frac{\theta}{\mu}),\right.$$

$$\left. \mu^{1/2}\varepsilon^{1/2}\theta L^{1/2}\ln^{1/2}(3 + \frac{1}{\theta^2}) + \varepsilon L, \mu^{1/2}\varepsilon^{1/2}\theta L^{1/2}\ln^{1/2}(3 + \frac{\varepsilon}{\mu^3\theta^2 L}) + \varepsilon L\right\}.$$

Here many competitions between the 'more local' (lower line) and 'global' (upper line) logarithms take place, and the different contributions are treated separately in the proof of the lower bound. Precisely, in Lemma 3.11, the 'global' logarithms are captured which always are in competition with a single laminate (εL). Combined with the energy required for an interpolation from a constant to an affine function (see Lem. 2.5), this leads to the lower bound in Proposition 3.8. The case of a single laminate requires additional care since there the incompatibility at the left and right boundaries lead to a competition between complex microstructures inside the nucleus and elastic strain in the austenite part. We point out that the situation here is (even in a scalar-valued setting) more complicated and the scaling behaviour is more complex than in the well-studied case of a vertical austenite/martensite interface with periodic boundary conditions at the top and bottom boundaries. This is in particular reflected in Proposition 3.3 by the additional regime $\mu\theta^2\ln(3 + \frac{\theta}{\mu})$.

3. LOWER BOUND

The proof of the lower bound will be divided in three main parts, addressing various regimes in which qualitatively different behavior is expected from the constructions in Section 2.1. We shall briefly outline the structure of the proof:

In Section 3.1, we deal with the case that ε is not very small. Specifically, we assume that one of θ^2 and $\mu\theta^2$ is below ε. Roughly speaking, in this regime, one expects rather uniform structures inside the nucleus. The lower bound in this regime is given in Proposition 3.1. The key competition is between the bulk energy in the martensite and the bulk energy in the austenite, and is made quantitative in Lemma 3.2.

In Section 3.2, we treat the case of small ε, in which we expect microstructure. This is the most interesting and richest regime, in which a variety of one- and two-scale branching patterns appear. The smallness of ε corresponds to two conditions: firstly, it should be such that there is at least a single interface over the entire length of the sample, as made quantitative by comparing εL with $\mu\theta^2$ (up to a logarithmic factor, see below for the precise condition). Secondly, it must be such that the cost of a branching pattern is not dominated by the cost of a single straight interface, in the sense that $\varepsilon L \leq \varepsilon^{2/3}\theta^{2/3}L^{1/3}$, which is equivalent to $\varepsilon L^2 \leq \theta^2$. Roughly speaking, in this regime, one expects complex patterns inside the whole martensitic nucleus, and contributions

from the austenite part only close to the left and right boundaries of Ω_{2L}. The lower bound is derived in Proposition 3.3, which builds upon a series of Lemmata for specific parts of the estimate.

In Section 3.3, we address the remaining part of the small-ε range which is not covered in Section 3.2, corresponding to the cases that μ is small (in the sense that $\mu\theta^2 \lesssim \varepsilon L$) or that L is large (in the sense that $\theta^2 < \varepsilon L^2$). In this case, one expects that there are parts inside the nucleus in which the displacement is affine or a single laminate. The relevant lower bound is obtained in Proposition 3.8.

Finally, in Section 3.4, we put together the above results and conclude the proof of the lower bound.

We start by making a few general observations and definitions that will be used all over the argument. The condition $\nabla u \in BV(\Omega_{2L}; \mathbb{R}^{2\times 2})$ implies that u has a representative which is continuous on $\overline{\Omega_{2L}}$ (see, for example, Lemma 9 in [26]). We work with this representative, and mainly work on slices in direction $\xi = (1/4, 1)$. One important quantity is the set $\mathcal{C}$ of slices which are almost affine with slope θ or $\theta - 1$ (recall (1.6))

$$\mathcal{C} := \left\{ x_1 \in (0, L - \xi_1) \, : \, \min\left\{ \|u_{x_1}^\xi(s) - u_{x_1}^\xi(0) - s\theta\|_{L^\infty((0,1))}, \|u_{x_1}^\xi(s) - u_{x_1}^\xi(0) - s(\theta-1)\|_{L^\infty((0,1))} \right\} < \frac{1}{16}\theta \right\}. \tag{3.1}$$

These are slices which have almost no energy in the martensitic nucleus. The boundary values on the top and bottom of the slice, however, differ by approximately θ (or $1 - \theta$). Therefore, these slices generate a large energy in the austenitic matrix.

Correspondingly, we shall consider the set $\mathcal{P}$ of slices where the boundary values are close,

$$\mathcal{P} := \left\{ x_1 \in (0, L - \xi_1) \, : \, |u((x_1, 0) + \xi) - u(x_1, 0)| \leq 2^{-7}\theta \right\}. \tag{3.2}$$

These slices generate very small energy in the austenitic matrix, but cannot be low energy inside the martensitic nucleus. They can be realized either by microstructure (with energy density at least ε) or by having a deformation which does not match the eigendeformation of the martensite (with energy density at least θ^2). The sets $\mathcal{C}$ and $\mathcal{P}$ are clearly disjoint. This competition and these energy contributions will be made precise in Lemma 3.2 below.

Over the entire lower bound we shall often focus on "typical" slices, and relate the one-dimensional integrals over slices to the energy via Fubini's theorem. For example, recalling the definition (1.9), one has
$\int_{(0,L)\times(0,1)} |f|^2 \, dx \geq \int_0^{L-\xi_1} \|f\|_{L^2(\Delta_{x_1}^\xi)}^2 \, dx_1 \geq \alpha^2 \mathcal{L}^1\left(\left\{ x_1 \in (0, L - \xi_1) : \|f\|_{L^2(\Delta_{x_1}^\xi)} \geq \alpha \right\} \right)$ for any $\alpha > 0$.

3.1. A lower bound in the parameter range $\theta^2 \leq \varepsilon$ or $\mu\theta^2 \leq \varepsilon$

In this case the structure inside the martensite is coarse, and the optimal bound is obtained considering a path that contains the segment $(x_1, 0)$ to $(x_1 + \xi_1, 1)$, and then goes back in the austenite phase, staying at a distance of order x_1 from the martensite (see Fig. 8 and Lem. 3.2).

Proposition 3.1. *There exists $c > 0$ such that for all $u \in \mathcal{X}$, $\mu > 0$, $\theta \in (0, 1/2]$, $\varepsilon > 0$, and $L \in [1/2, \infty)$ with*

$$\min\{\theta^2, \mu\theta^2\} \leq \varepsilon$$

there holds

$$I(u) \geq c \min\left\{ \mu\theta^2 \ln(3 + L), \theta^2 L, \mu\theta^2 \ln\left(3 + \frac{L}{\mu}\right) + \varepsilon\theta \right\}.$$

The key estimate, that will be useful also later in the proof, is the following. We recall the definition of the sets $\mathcal{P}$ and $\mathcal{C}$ in (3.1) and (3.2), the definition of $\Delta_{x_1}^\xi$ in (1.8) and of the $L^2(\Delta_{x_1}^\xi)$ norm in (1.9). The geometry is illustrated in Figure 8.

Lemma 3.2 (Bulk energy). *Suppose that $\theta \in (0, 1/2]$, $L \in [1/2, \infty)$, $u \in \mathcal{X}$. The following holds:*
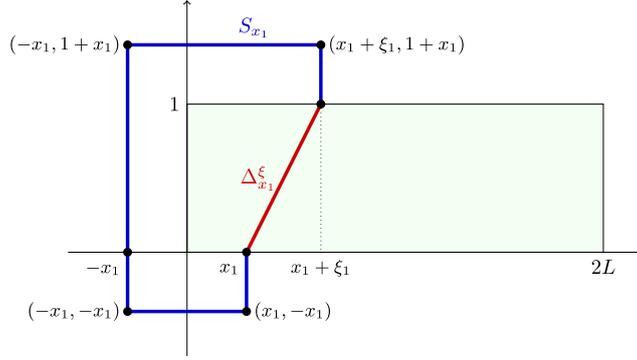
FIGURE 8. Sketch of the geometry in Lemma 3.2. The set S_{x_1} is marked blue, the segment $\Delta_{x_1}^\xi$ red. The path used here is analogous to the one in the upper bound construction, see Figure 4 and Lemma 2.1.

i. For $i \in \{1,2\}$ and for almost every $x_1 \in (0, L - \xi_1)$ one has

$$\mu \int_{S_{x_1}} |\nabla u_i|^2 \, d\mathcal{H}^1 \geq \frac{1}{8} \mu \frac{1}{1+x_1} |u_i(x_1 + \xi_1, 1) - u_i(x_1, 0)|^2,$$

where S_{x_1} is the polygonal arc in $\mathbb{R}^2 \setminus \Omega_{2L}$ joining the points

$$(x_1 + \xi_1, 1),\ (x_1 + \xi_1, 1 + x_1),\ (-x_1, 1 + x_1),\ (-x_1, -x_1),\ (x_1, -x_1),\ (x_1, 0).$$

ii. If $x_1 \in [0, L - \xi_1]$ obeys $\| \min\{|e(u) - \theta e_1 \odot e_2|, |e(u) + (1-\theta)e_1 \odot e_2|\} \|^2_{L^2(\Delta_{x_1}^\xi)} \leq \tilde{C}_1 \theta^2$ and $|\partial_s \partial_s u_{x_1}^\xi|((0,1)) \leq \tilde{C}_1$, where $\tilde{C}_1 := 2^{-13}$, then $x_1 \in \mathcal{C}$.

iii. For any $x_1 \in \mathcal{C}$ we have $|u_{x_1}^\xi(1) - u_{x_1}^\xi(0)| \geq \frac{3}{4}\theta$. In particular, the sets $\mathcal{P}$ and $\mathcal{C}$ are disjoint.

iv. One has

$$I(u) \geq c \min\{\varepsilon, \theta^2\} \mathcal{L}^1([0, L - \xi_1] \setminus \mathcal{C}),$$

and

$$I(u) \geq c\mu\theta^2 \ln \frac{L + 1 - \xi_1}{\mathcal{L}^1(P) + 1} + c \min\{\varepsilon, \theta^2\} \mathcal{L}^1(\mathcal{P}).$$

Proof. (i) The assertion follows by a direct computation, using that

$$\mathcal{H}^1(S_{x_1}) = 8x_1 + 1 + \xi_1 \leq 8(1 + x_1).$$

Indeed, for almost every $x_1 \in (0, L - \xi_1)$ we have that $u_i \in W^{1,2}(S_{x_1})$. It then follows that

$$\int_{S_{x_1}} |\nabla u_i|^2 \, d\mathcal{H}^1 \geq (\mathcal{H}^1(S_{x_1}))^{-1} \left(\int_{S_{x_1}} |\nabla u_i| \, d\mathcal{H}^1 \right)^2 \geq \frac{|u_i(x_1 + \xi_1, 1) - u_i(x_1, 0)|^2}{8(1 + x_1)}.$$

(ii) Fix such an x_1 and define $v(s) := u^\xi_{x_1}(s)$. Since $|\partial_s \partial_s v| \leq \frac{1}{8}$, there is $b \in \mathbb{R}$ such that $|v'(s) - b| \leq \frac{1}{8}$ for almost all $s \in (0,1)$. We observe that by Hölder's inequality and (1.11)

$$\min\{|b - \theta|, |b + (1 - \theta)|\} \leq \int_0^1 \min\{|v' - \theta|, |v' + (1 - \theta)|\} + |v' - b| \, ds \leq 5\tilde{C}_1^{1/2}\theta + \frac{1}{8} \leq \frac{1}{4}.$$

Therefore there is $\sigma \in \{0, 1\}$ such that $|b + \sigma - \theta| \leq \frac{1}{4}$, and correspondingly $|v'(s) + \sigma - \theta| \leq \frac{1}{2}$ for almost all s. Since $|\theta - (1 - \theta)| = 1$ it follows that $|v'(s) + \sigma - \theta| = \min\{|v' - \theta|, |v' + 1 - \theta|\}$ for almost every s, and thus

$$\int_0^1 |v' + \sigma - \theta| \, ds = \int_0^1 \min\{|v' - \theta|, |v' + 1 - \theta|\} \, ds \leq 5\tilde{C}_1^{1/2}\theta.$$

Integrating we obtain

$$|v(s) + (\sigma - \theta)s - v(0)| \leq 5\tilde{C}_1^{1/2}\theta \text{ for all } s \in (0, 1). \tag{3.3}$$

Since $\tilde{C}_1 = 2^{-13}$ this implies $x_1 \in \mathcal{C}$.

(iii) For $x_1 \in \mathcal{C}$ we have $|u^\xi_{x_1}(1) - u^\xi_{x_1}(0)| \geq \min_{\sigma \in \{0,1\}} |\sigma - \theta| - 2\frac{1}{16}\theta \geq \frac{3}{4}\theta$. The second assertion follows from $\frac{3}{4}\theta \leq |u^\xi_{x_1}(1) - u^\xi_{x_1}(0)| \leq 4|\xi| \, |u((x_1, 0) + \xi) - u(x_1, 0)|$ and $|\xi| \leq 2$.

(iv) The first assertion follows from (ii) with Fubini's theorem. Indeed, if $x_1 \in (0, L - \xi_1) \setminus \mathcal{C}$ then $\| \min\{|e(u) - \theta e_1 \odot e_2|, |e(u) + (1 - \theta)e_1 \odot e_2|\} \|^2_{L^2(\Delta^\xi_{x_1})} > \tilde{C}_1\theta^2$ or $\varepsilon|\partial_s \partial_s u^\xi_{x_1}|((0, 1)) > \varepsilon\tilde{C}_1$. Integrating over all such x_1 we obtain $I(u) \geq c \min\{\varepsilon, \theta^2\} \mathcal{L}^1([0, L - \xi_1] \setminus \mathcal{C})$.

To prove the other one, for almost every $x_1 \in (0, L - \xi_1) \setminus \mathcal{P}$ we obtain by (i) that

$$\mu \int_{S_{x_1}} |\nabla u|^2 \, d\mathcal{H}^1 \geq c\frac{\mu\theta^2}{1 + x_1}$$

so that, using Fubini and monotonicity of $1/(1 + x_1)$,

$$I^{\text{ext}}(u) \geq c\mu\theta^2 \int_{(0, L - \xi_1) \setminus \mathcal{P}} \frac{1}{x_1 + 1} \, dx_1 \geq c\mu\theta^2 \int_{\mathcal{L}^1(\mathcal{P})}^{L - \xi_1} \frac{1}{x_1 + 1} \, dx_1 = c\mu\theta^2 \ln \frac{L + 1 - \xi_1}{\mathcal{L}^1(\mathcal{P}) + 1}.$$

Finally, since $\mathcal{P}$ and $\mathcal{C}$ are disjoint, we have $\mathcal{P} \subset [0, L - \xi_1] \setminus \mathcal{C}$ and $I(u) \geq c \min\{\varepsilon, \theta^2\} \mathcal{L}^1([0, L - \xi_1] \setminus \mathcal{C}) \geq c \min\{\varepsilon, \theta^2\} \mathcal{L}^1(\mathcal{P})$.

$\square$

Proof of Proposition 3.1. By Lemma 3.2(iv) we have, with $p := \mathcal{L}^1(\mathcal{P})$,

$$I(u) \geq c\mu\theta^2 \ln \frac{L + 1 - \xi_1}{p + 1} + cp \min\{\varepsilon, \theta^2\}. \tag{3.4}$$

We distinguish two cases.

i. Assume $\mu \leq 1$. Then the assumption on ε implies $\mu\theta^2 \leq \varepsilon$, so that $\mu\theta^2 \leq \min\{\varepsilon, \theta^2\}$ and (3.4) gives

$$I(u) \geq c \min_{p' \in [0, L - \xi_1]} \left(\mu\theta^2 \ln \frac{L + \frac{3}{4}}{p' + 1} + p'\mu\theta^2 \right) = c\mu\theta^2 \ln \left(L + \frac{3}{4} \right) \geq c\mu\theta^2 \ln(3 + L),$$

which concludes the proof. We used here that the expression to be minimized is nondecreasing in p', hence the minimum is attained at $p' = 0$.

ii. Assume $1 < \mu$. Then the assumption on ε implies $\theta^2 \leq \varepsilon$, and (3.4) gives

$$I(u) \geq c\Big(\mu\theta^2 \ln \frac{L + \frac{3}{4}}{p + 1} + p\theta^2\Big).$$

If $p \geq \frac{1}{5}L$, then $I(u) \geq c\theta^2 L$ and we are done. If $p = 0$, then $I(u) \geq c\mu\theta^2 \ln(3 + L)$ as above and we are done. Assume now $0 < p < \frac{1}{5}L$. We observe that $p \leq \frac{1}{5}L$ and $\frac{1}{2} \leq L$ imply $\frac{10}{9}(p + 1) \leq L + \frac{3}{4}$ and therefore $I(u) \geq c\mu\theta^2 \ln \frac{10}{9} \geq c\mu\theta^2$. Further,

$$I(u) \geq c \min_{p' \in [0, L - \xi_1]} \Big(\mu\theta^2 \ln \frac{L + \frac{3}{4}}{p' + 1} + p'\theta^2\Big) \geq c \min \Big\{\mu\theta^2 \ln(3 + L), \mu\theta^2 \ln\Big(\frac{L + \frac{3}{4}}{\mu}\Big), \theta^2 L\Big\}$$

(we used here that the only zero of the derivative is at $p' + 1 = \mu$, hence the minimum over $[0, \infty)$ is at $p' = \mu - 1$). Since we had already proven $I(u) \geq c\mu\theta^2$, distinguishing the two cases $L/\mu \geq 3$ and $L/\mu \leq 3$ gives

$$I(u) \geq c \min \Big\{\mu\theta^2 \ln(3 + L), \mu\theta^2 \ln\Big(3 + \frac{L}{\mu}\Big), \theta^2 L\Big\}. \tag{3.5}$$

At this point it only remains to obtain the $\varepsilon\theta$ term. We are working under the assumption that $p > 0$. For almost any $x_p \in \mathcal{P}$ we have

$$\int_{(0,1)} (u_{x_p}^{\xi})'(s) \, ds = |u_{x_p}^{\xi}(1) - u_{x_p}^{\xi}(0)| \leq \frac{1}{4}\theta.$$

If $\mathcal{C} = \emptyset$ then by Lemma 3.2(iv) we have $I(u) \geq c \min\{\varepsilon, \theta^2\}(L - \xi_1) \geq c\theta^2 L$, and we are done. Otherwise, for any $x_c \in \mathcal{C}$ we have

$$\frac{3}{4}\theta \leq |u_{x_c}^{\xi}(1) - u_{x_c}^{\xi}(0)| \leq \int_{(0,1)} (u_{x_c}^{\xi})'(s) \, ds.$$

Therefore

$$\frac{1}{2}\theta \leq \int_{(0,1)} |(u_{x_c}^{\xi})' - (u_{x_p}^{\xi})'| \, ds \leq 4|D^2 u|((0, L) \times (0, 1))$$

and therefore $I(u) \geq c\varepsilon\theta$. Recalling (3.5) we have

$$I(u) \geq c \min \Big\{\mu\theta^2 \ln(3 + L), \mu\theta^2 \ln\Big(3 + \frac{L}{\mu}\Big) + \varepsilon\theta, \theta^2 L\Big\}$$

which concludes the proof. $\square$

3.2. A lower bound in the parameter range $\varepsilon L^2 \leq \theta^2$ and $\varepsilon L \lesssim \mu\theta^2$

We turn to the case in which $\varepsilon L^2 \leq \theta^2$, and formulate a lower bound on the energy restricted to the set

$$\tilde{\Omega}_\ell := (-\infty, \ell) \times \mathbb{R}. \tag{3.6}$$

We give these estimates for general $\ell \leq 2L$ since this does not require extra work. To prove the main theorem, however, we will only need the case $\ell = 2L$.

Proposition 3.3. *There exists $c > 0$ such that for all $u \in \mathcal{X}$, $\theta \in (0, 1/2]$, $\varepsilon > 0$, $\ell > 0$, $\mu > 0$ which obey*

$$1 \leq \ell \leq 2L \,, \qquad \varepsilon \ell^2 \leq \theta^2 \,, \qquad and \qquad \varepsilon \ell \leq \mu \theta^2 \min \left\{ \ln \left(3 + \frac{1}{\theta^2} \right), \ln \left(3 + \frac{\varepsilon}{\mu^3 \theta^2 \ell} \right) \right\}$$

there holds

$$I_{\tilde{\Omega}_\ell}(u) \geq c \min \left\{ \mu \theta^2 \ln(3 + \ell), \mu \theta^2 \ln \left(3 + \frac{\theta}{\mu} \right), \varepsilon^{2/3} \theta^{2/3} \ell^{1/3}, \right.$$
$$\left. \mu^{1/2} \varepsilon^{1/2} \theta \ell^{1/2} \left(\ln \left(3 + \frac{1}{\theta^2} \right) \right)^{1/2}, \mu^{1/2} \varepsilon^{1/2} \theta \ell^{1/2} \left(\ln \left(3 + \frac{\varepsilon}{\mu^3 \theta^2 \ell} \right) \right)^{1/2} \right\}.$$

The key estimate for the most difficult case, in which θ is small, is proven in Lemma 3.7 below. A proof of similar statements has been provided by two of the authors in [27] in the context of simplified scalar-valued models for austenite/martensite interfaces and crystal plasticity. Our proof follows the general strategy of the proof and builds on techniques from there with two main differences: First, the vectorial structure requires more refined arguments; and second, the isotropic elastic modulus allows for more flexibility which is treated differently than in the corresponding model for dislocation microstructures.

The vector-valued setting including a symmetrized gradient requires more careful slicing techniques than in [27]. We follow a BD-type approach and consider diagonal slices instead of nearly vertical slices. An intuitive choice of the direction of slices would be $(1, 1)$. However, this would lead to problems in the case $L = 1/2$. Indeed, in this case $\Omega_{2L} = (0, 1)^2$ contains only a single segment parallel to $(1, 1)$ which connects the bottom and the top boundaries, and it would not be possible to choose a 'typical' slice (in the sense of Fubini's theorem). Note that the energy controls only the symmetric part of the gradient of u, and hence we rely on BD-type slicing results which hold along diagonal slices but not on vertical ones. For these reasons, we fixed the direction of the slices as

$$\xi = \left(\frac{1}{4}, 1 \right).$$

We first treat the simpler case in which θ is bounded away from zero. In this case we use a different proof which uses ideas that were introduced in [15] in the geometrically nonlinear setting and were refined in the linear setting in [31, 47]. All these works treat only the case $\theta = \frac{1}{2}$, but the argument can easily be extended to the case $\theta \in [m_1, \frac{1}{2}]$. We start with this argument, which is simpler and does not require much preparation.

Lemma 3.4 (The case of large θ). *Let $m_1 \in (0, 1/2]$. There exists $c > 0$ depending only on m_1 such that for all $u \in \mathcal{X}$, $\varepsilon > 0$, $\ell > 0$, $\mu > 0$, $\theta > 0$ which obey*

$$1 \leq \ell \leq 2L \,, \qquad \varepsilon \ell^2 \leq 1 \,, \qquad \varepsilon \ell \leq \mu \,, \qquad and \qquad m_1 \leq \theta \leq \frac{1}{2}$$

we have

$$I_{\tilde{\Omega}_\ell}(u) \geq c \min \left\{ \varepsilon^{2/3} \ell^{1/3}, \mu^{1/2} \varepsilon^{1/2} \ell^{1/2} \right\}.$$

Proof. Let $\sigma : \Omega_\ell \to \{\theta, -1 + \theta\}$ be the function that indicates which variant is locally attained, *i.e.*, such that $\min\{|e(u) - \theta e_1 \odot e_2|, |e(u) + (1 - \theta) e_1 \odot e_2|\} = |e(u) - \sigma e_1 \odot e_2|$ almost everywhere. We fix $\rho \in (0, 1]$, and choose $\tilde{x}_1 \in [0, \ell - \rho]$, $\tilde{x}_2 \in [0, 1 - \rho]$ such that the infinite horizontal strip $S_h := (-\infty, \ell) \times (\tilde{x}_2, \tilde{x}_2 + \rho)$, and the infinite
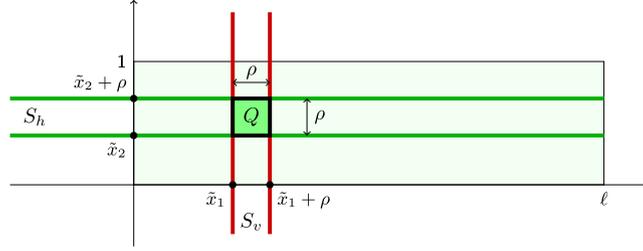
FIGURE 9. Sketch of the geometry in the proof of Lemma 3.4. The horizontal and vertical stripes S_h and S_v, as well as their intersection Q, are emphasized.

vertical strip $S_v := (\tilde{x}_1, \tilde{x}_1 + \rho) \times \mathbb{R}$ obey, recalling the definition of $\tilde{\Omega}_\ell$ in (3.6),

$$I_Q(u) \leq c\ell^{-1}\rho^2 I_{\tilde{\Omega}_\ell}(u), \quad I_{S_v}(u) \leq c\ell^{-1}\rho I_{\tilde{\Omega}_\ell}(u), \text{ and } I_{S_h}(u) \leq c\rho I_{\tilde{\Omega}_\ell}(u), \tag{3.7}$$

where $Q := S_h \cap S_v \subset \Omega_\ell$ (see Fig. 9). By Poincaré's inequality we have that

$$\|\partial_2 u_1 - \langle \partial_2 u_1 \rangle_Q\|_{L^1(Q)} \leq \rho|D^2 u_1|(Q) \quad \text{and} \quad \|\partial_1 u_2 - \langle \partial_1 u_2 \rangle_Q\|_{L^1(Q)} \leq \rho|D^2 u_2|(Q) \tag{3.8}$$

where $\langle f \rangle_Q$ denotes the average of f over Q. Combined with Hölder's inequality, we obtain

$$\begin{aligned}
\|\sigma - \langle \partial_2 u_1 \rangle_Q - \langle \partial_1 u_2 \rangle_Q\|_{L^1(Q)} &\leq \|\sigma - (\partial_2 u_1 + \partial_1 u_2)\|_{L^1(Q)} + \|\partial_2 u_1 - \langle \partial_2 u_1 \rangle_Q\|_{L^1(Q)} \\
&\quad + \|\partial_1 u_2 - \langle \partial_1 u_2 \rangle_Q\|_{L^1(Q)} \\
&\leq c\left(\rho I_Q^{1/2}(u) + \rho\varepsilon^{-1} I_Q(u)\right) \leq c\left(\rho^2 \ell^{-1/2} I_{\tilde{\Omega}_\ell}^{1/2}(u) + \rho^3 \varepsilon^{-1}\ell^{-1} I_{\tilde{\Omega}_\ell}(u)\right).
\end{aligned}$$

If $\mathcal{L}^2(\{\sigma = \theta\} \cap Q) \geq \frac{1}{2}\rho^2$, then

$$\frac{1}{2}\rho^2 |\theta - \langle \partial_2 u_1 \rangle_Q - \langle \partial_1 u_2 \rangle_Q| \leq \|\sigma - \langle \partial_2 u_1 \rangle_Q - \langle \partial_1 u_2 \rangle_Q\|_{L^1(Q)}.$$

Otherwise $\mathcal{L}^2(\{\sigma = \theta - 1\} \cap Q) \geq \frac{1}{2}\rho^2$ and hence

$$\frac{1}{2}\rho^2 |(\theta - 1) - \langle \partial_2 u_1 \rangle_Q - \langle \partial_1 u_2 \rangle_Q| \leq \|\sigma - \langle \partial_2 u_1 \rangle_Q - \langle \partial_1 u_2 \rangle_Q\|_{L^1(Q)}.$$

Hence if $|\langle \partial_2 u_1 \rangle_Q| \leq m_1/4$ and $|\langle \partial_1 u_2 \rangle_Q| \leq m_1/4$, then since $m_1 \leq \theta$, $|\theta - \langle \partial_2 u_1 \rangle_Q - \langle \partial_1 u_2 \rangle_Q| \geq \theta/2$ and $|\theta - 1 - \langle \partial_2 u_1 \rangle_Q - \langle \partial_1 u_2 \rangle_Q| \geq \theta/2$, we deduce

$$m_1 \rho^2 \leq c\left(\rho^2 \ell^{-1/2} I_{\tilde{\Omega}_\ell}^{1/2}(u) + \rho^3 \varepsilon^{-1}\ell^{-1} I_{\tilde{\Omega}_\ell}(u)\right),$$

which implies that

$$I_{\tilde{\Omega}_\ell}(u) \geq c \min\{\ell, \varepsilon\ell\rho^{-1}\}. \tag{3.9}$$

This estimate will be considered as one part of (3.11) below.

It remains to consider the other cases. If $|\langle \partial_2 u_1 \rangle_Q| > m_1/4$, we set $a := \langle \partial_2 u_1 \rangle_Q$ and $b := \langle u_1 - ax_2 \rangle_Q$. Then by Poincaré's and Hölder's inequalities (see also (3.8))

$$\|u_1 - ax_2 - b\|_{L^1(Q)} \leq \rho\|\nabla(u_1 - \langle \partial_2 u_1 \rangle_Q x_2)\|_{L^1(Q)} \leq \rho\|\partial_1 u_1\|_{L^1(Q)} + \rho\|\partial_2 u_1 - \langle \partial_2 u_1 \rangle_Q\|_{L^1(Q)}$$
$$\leq \rho^2 I_Q^{1/2}(u) + \rho^2 \varepsilon^{-1} I_Q(u).$$

By Fubini's theorem, there is $x_1^* \in (\tilde{x}_1, \tilde{x}_1 + \rho)$ such that

$$\int_{\tilde{x}_2}^{\tilde{x}_2 + \rho} |u_1(x_1^*, x_2) - ax_2 - b| \, dx_2 \leq \rho I_Q^{1/2}(u) + \rho\varepsilon^{-1} I_Q(u)$$

and $u_1(x_1^*, \cdot)$ is the trace of u_1 on $\{x_1 = x_1^*\}$. We use the fundamental theorem to transfer this information to a corresponding slice on the boundary. Precisely, we estimate

$$\int_{\tilde{x}_2}^{\tilde{x}_2 + \rho} |u_1(0, x_2) - ax_2 - b| \, dx_2$$
$$\leq \int_{\tilde{x}_2}^{\tilde{x}_2 + \rho} |u_1(x_1^*, x_2) - ax_2 - b| \, dx_2 + \int_{\tilde{x}_2}^{\tilde{x}_2 + \rho} \int_0^{x_1^*} |\partial_1 u_1|(x_1, x_2) \, dx_1 \, dx_2$$
$$\leq c\left(\rho I_Q^{1/2}(u) + \rho\varepsilon^{-1} I_Q(u) + \ell^{1/2}\rho^{1/2} I_{S_h}^{1/2}(u)\right)$$
$$\leq c\left(\rho^2 \ell^{-1/2} I_{\tilde{\Omega}_\ell}^{1/2}(u) + \rho^3 \varepsilon^{-1}\ell^{-1} I_{\tilde{\Omega}_\ell}(u) + \rho\ell^{1/2} I_{\tilde{\Omega}_\ell}^{1/2}(u)\right).$$

Since $|a| \geq m_1/4$, we get by Lemma 1.4 applied to $u_1(0, \tilde{x}_2 + \cdot)$ that

$$m_1\rho^2 \leq c\left(\rho^2 \ell^{-1/2} I_{\tilde{\Omega}_\ell}^{1/2}(u) + \rho^3 \varepsilon^{-1}\ell^{-1} I_{\tilde{\Omega}_\ell}(u) + \rho\ell^{1/2} I_{\tilde{\Omega}_\ell}^{1/2}(u) + m_1^{-1}\mu^{-1} I_{S_h}(u)\right)$$
$$\leq c\left(\rho^2 \ell^{-1/2} I_{\tilde{\Omega}_\ell}^{1/2}(u) + \rho^3 \varepsilon^{-1}\ell^{-1} I_{\tilde{\Omega}_\ell}(u) + \rho\ell^{1/2} I_{\tilde{\Omega}_\ell}^{1/2}(u) + \mu^{-1}\rho I_{\tilde{\Omega}_\ell}(u)\right), \tag{3.10}$$

where in the last step we subsumed m_1 in the constant c. Finally, if $|\langle \partial_1 u_2 \rangle_Q| > m_1/4$, we proceed analogously, interchanging the indices 1 and 2, and obtain the estimate

$$m_1\rho^2 \leq c\left(\int_{\tilde{x}_1}^{\tilde{x}_1 + \rho} |u_2(x_1, x_2^*) - \tilde{a}x_1 - \tilde{b}| \, dx_1 + \int_{\tilde{x}_1}^{\tilde{x}_1 + \rho} \int_0^{x_2^*} |\partial_2 u_2| \, dx + \mu^{-1}[u(\cdot, 0)]^2_{H^{1/2}(\tilde{x}_1, \tilde{x}_1 + \rho)}\right)$$
$$\leq c\left(\rho^2 \ell^{-1/2} I_{\tilde{\Omega}_\ell}^{1/2}(u) + \rho^3 \varepsilon^{-1}\ell^{-1} I_{\tilde{\Omega}_\ell}(u) + \rho I_{\tilde{\Omega}_\ell}^{1/2}(u) + \mu^{-1}\ell^{-1}\rho I_{\tilde{\Omega}_\ell}(u)\right)$$

with $\tilde{a} := \langle \partial_1 u_2 \rangle_Q$ and $\tilde{b} := \langle u_2 - \tilde{a}x_1 \rangle_Q$ and an appropriately chosen $x_2^* \in (\tilde{x}_2, \tilde{x}_2 + \rho)$.

Putting this together with (3.9) and (3.10), we see that for any $\rho \in (0, 1)$

$$I_{\tilde{\Omega}_\ell}(u) \geq c\min\{\ell, \varepsilon\ell\rho^{-1}, \ell^{-1}\rho^2, \mu\rho\} = c\min\{\varepsilon\ell\rho^{-1}, \ell^{-1}\rho^2, \mu\rho\}, \tag{3.11}$$

where we used that $\rho \leq \ell$ implies $\ell^{-1}\rho^2 \leq \ell$.

If $\mu \geq \varepsilon^{1/3}\ell^{-1/3}$ we choose $\rho = \varepsilon^{1/3}\ell^{2/3}$ and conclude $I_{\tilde{\Omega}_\ell}(u) \geq c\varepsilon^{2/3}\ell^{1/3}$ ($\rho \leq 1$ since $\varepsilon\ell^2 \leq 1$).

If $\mu < \varepsilon^{1/3}\ell^{-1/3}$ we choose $\rho = \mu^{-1/2}\varepsilon^{1/2}\ell^{1/2}$ and conclude $I_{\tilde{\Omega}_\ell}(u) \geq c\mu^{1/2}\varepsilon^{1/2}\ell^{1/2}$ since $\mu^{1/2}\varepsilon^{1/2}\ell^{1/2} \leq \ell^{-1}\rho^2 = \mu^{-1}\varepsilon$ by the assumption on μ (note that $\rho \leq 1$ since $\mu \leq \varepsilon\ell$). This concludes the proof of Lemma 3.4. $\qquad\square$

If θ is small then a more complex procedure is needed, in order to capture the various logarithmic divergences in the energy. Before presenting the main estimate in Lemma 3.7 we need a number of preliminary results, which characterize the behavior of low-energy functions close to the corners in $(0,0)$ and $(0,1)$ and on "good" slices.

We begin by proving estimate for the behavior of u_1 close to the corners $(0,0)$ and $(0,1)$, which captures the logarithmic divergence of the matrix energy, see (3.12).

Lemma 3.5. *For any $c_* \in (0,1]$ there is $C = C(c_*)$ such that whenever $1 \le \ell \le 2L$, $m \in (0, \frac{1}{4}]$ and*

$$0 < \mu \le \theta \le m$$

one of the following holds: either

$$\frac{1}{C}I(u) \ge \min\left\{\mu\theta^2 \ln\left(3 + \frac{\theta}{\mu}\right), \mu\theta^2 \ln\frac{1}{m}, \mu\theta^2 \ln(3+\ell)\right\}$$

or

$$\int_m^1 \frac{|(u_1(x_1,1) - u_1(x_1,0))|}{x_1}\, \mathrm{d}x_1 \le c_*\theta \ln\frac{1}{m}. \tag{3.12}$$

Proof. We show below the following: either (3.12) holds or

$$\frac{1}{c}I(u) \ge \min\left\{\frac{\theta^3}{q}, \mu\theta^2 \ln(3+q), \mu\theta^2 \ln\frac{1}{m}\right\} \quad \text{for any } q \in [1,\ell]. \tag{3.13}$$

We first prove that this implies the assertion. We let $q_* := \frac{2\theta}{\mu \ln(3+\theta/\mu)}$, observe that $\mu \le \theta$ implies $q_* \ge 1$, and distinguish two cases. If $q_* \le \ell$ we set $q = q_*$. Then $\theta^3/q_* = \frac{1}{2}\mu\theta^2 \ln(3 + \frac{\theta}{\mu})$. Further, observe that for any $t > 0$ one has $3 + \frac{2t}{\ln(3+t)} \ge (3+t)^{1/2}$, which implies $\ln(3 + \frac{2t}{\ln(3+t)}) \ge \frac{1}{2}\ln(3+t)$, and obtain $\mu\theta^2 \ln(3 + q_*) \ge \frac{1}{2}\mu\theta^2 \ln(3 + \frac{\theta}{\mu})$, which concludes the proof. If instead $q_* \ge \ell$ we set $q = \ell$ and observe that in this case $\frac{\theta^3}{q} \ge \frac{\theta^3}{q_*} = \frac{1}{2}\mu\theta^2 \ln(3 + \frac{\theta}{\mu})$, which also concludes the proof. Therefore it suffices to show that one of (3.12) and (3.13) holds.

Fix $q \in [1,\ell]$. We can assume that $I(u) \le 2^{-9}c_*^3 \frac{\theta^3}{q}$ (if not, (3.13) holds and the proof is finished). Since $\|\partial_1 u_1\|_{L^2(\Omega_\ell)}^2 \le I(u)$, we can choose $\delta \in (0, \frac{1}{16}c_*\theta)$ such that $u_1(\cdot, \delta), u_1(\cdot, 1-\delta) \in W^{1,2}((0,\ell))$ with

$$\int_0^\ell |\partial_1 u_1|^2(x_1, \delta) + |\partial_1 u_1|^2(x_1, 1-\delta)\, \mathrm{d}x_1 \le \frac{1}{32}c_*^2 \frac{\theta^2}{q}.$$

This implies, setting $\bar{u}_1^- := u_1(0,\delta)$ and $\bar{u}_1^+ := u_1(0, 1-\delta)$,

$$\left|u_1(\tilde{x}_1, \delta) - \bar{u}_1^-\right| + \left|u_1(\tilde{x}_1, 1-\delta) - \bar{u}_1^+\right| \le \frac{1}{4}c_*\theta \text{ for all } \tilde{x}_1 \in [0,q]. \tag{3.14}$$

To shorten notation we define

$$f(x) := \begin{cases} \min\{|e(u)(x) - \theta e_1 \odot e_2|, |e(u)(x) - (\theta-1)e_1 \odot e_2|\}, & \text{if } x \in (0,\ell) \times (0,1), \\ |\nabla u|(x), & \text{otherwise.} \end{cases}$$

For $x_1 \in (\delta, q)$ we define $z : [0, 1] \to \mathbb{R}$ by $z(s) := (u_1 - u_2)(x_1 - s, s)$. For almost every x_1 we can estimate, similar to (1.10)-(1.11),

$$
\begin{aligned}
|z'(s)| &= |\partial_1 u_2 + \partial_2 u_1 - \partial_1 u_1 - \partial_2 u_2|(x_1 - s, s) \\
&\leq \min_{\sigma \in \{\theta, \theta-1\}} |\sigma| + |(e_{12}(u) + e_{21}(u) - \sigma) - (e_{11}(u) + e_{22}(u))|(x_1 - s, s) \\
&\leq 1 + 2 \min \{|e(u)(x_1 - s, s) - \theta e_1 \odot e_2|, |e(u)(x_1 - s, s) - (\theta - 1)e_1 \odot e_2|\} \leq 1 + 2f(x_1 - s, s).
\end{aligned}
\tag{3.15}
$$

With multiple triangular inequalities,

$$
\begin{aligned}
|u_1(x_1, 0) - u_1^-| &\leq |u_1(x_1, 0) - u_1(x_1 - \delta, \delta)| + |u_1(x_1 - \delta, \delta) - u_1^-| \\
&\leq |z(0) - z(\delta)| + |u_2(x_1, 0) - u_2(x_1 - \delta, \delta)| + |u_1(x_1 - \delta, \delta) - u_1^-| \\
&\leq |z(0) - z(\delta)| + |u_2(x_1, 0) - u_2(x_1 - \delta, -\delta)| + |u_2(x_1 - \delta, -\delta) - u_2(x_1 - \delta, \delta)| \\
&\quad + |u_1(x_1 - \delta, \delta) - u_1^-|.
\end{aligned}
\tag{3.16}
$$

By the fundamental theorem of calculus, using (3.15), $|\partial_2 u_2|(x) \leq f(x)$ everywhere and $|\nabla u_2|(x) \leq f(x)$ for $x_2 \leq 0$,

$$
|u_1(x_1, 0) - u_1^-| \leq |u_1(x_1 - \delta, \delta) - u_1^-| + \delta + 2 \int_{S_{x_1}^-} f \, d\mathcal{H}^1,
\tag{3.17}
$$

where $S_{x_1}^-$ is the polygonal joining $(x_1, 0), (x_1 - \delta, \delta), (x_1 - \delta, -\delta), (x_1, 0)$ (see Fig. 10). Repeating the computation on the other side with $\tilde{z}(s) := (u_1 + u_2)(x_1 - s, 1 - s)$ leads to

$$
|u_1(x_1, 1) - u_1^+| \leq |u_1(x_1 - \delta, 1 - \delta) - u_1^+| + \delta + 2 \int_{S_{x_1}^+} f \, d\mathcal{H}^1,
\tag{3.18}
$$

where $S_{x_1}^+$ is the polygonal joining $(x_1, 1), (x_1 - \delta, 1 - \delta), (x_1 - \delta, 1 + \delta), (x_1, 1)$. Adding the two, and recalling (3.14) for $\tilde{x}_1 = x_1 - \delta$ yields, since $\mathcal{H}^1(S_{x_1}^+) = (2 + 2\sqrt{2})\delta \leq 5\delta$, and $2\delta \leq \frac{1}{8} c_* \theta$,

$$
|u_1(x_1, 0) - u_1^-| + |u_1(x_1, 1) - u_1^+| \leq \frac{1}{4} c_* \theta + 2\delta + 2 \int_{S_{x_1}^- \cup S_{x_1}^+} f \, d\mathcal{H}^1 \leq \frac{3}{8} c_* \theta + 8\delta^{1/2} \left(\int_{S_{x_1}^- \cup S_{x_1}^+} f^2 \, d\mathcal{H}^1 \right)^{1/2}
$$

for almost every $x_1 \in (\delta, q)$. We divide by x_1 and integrate over $x_1 \in (a, b)$, for some a, b with $\delta \leq a < b \leq q$,

$$
\begin{aligned}
\int_a^b \frac{|u_1(x_1, 0) - u_1^-| + |u_1(x_1, 1) - u_1^+|}{x_1} \, dx_1 &\leq \frac{3}{8} c_* \theta \ln \frac{b}{a} + 8\delta^{1/2} \int_a^b \frac{1}{x_1} \left(\int_{S_{x_1}^- \cup S_{x_1}^+} f^2 \, d\mathcal{H}^1 \right)^{1/2} dx_1 \\
&\leq \frac{3}{8} c_* \theta \ln \frac{b}{a} + 8\delta^{1/2} \left(\int_a^b \frac{1}{x_1^2} \, dx_1 \right)^{1/2} \left(\int_a^b \int_{S_{x_1}^- \cup S_{x_1}^+} f^2 \, d\mathcal{H}^1 \, dx_1 \right)^{1/2}.
\end{aligned}
$$

The first integral is controlled by $1/a$. For the second one we use Fubini's theorem,

$$
\int_a^b \int_{S_{x_1}^- \cup S_{x_1}^+} f^2 \, d\mathcal{H}^1 \, dx_1 \leq 2 \int_{(0,b) \times (-\delta, 1+\delta)} f^2 \, dx \leq \max\{1, \mu^{-1}\} 2I(u) = \frac{2I(u)}{\mu},
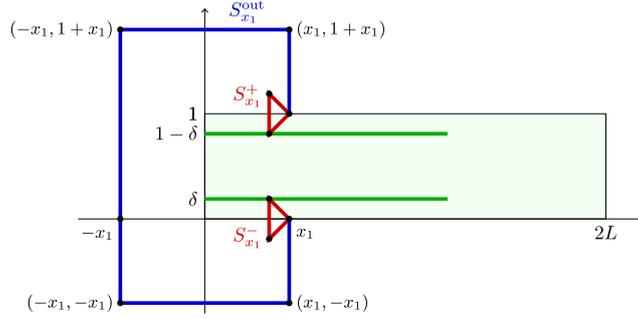$$

FIGURE 10. Sketch of the geometry in the proof of Lemma 3.5. The two horizontal lines entering
(3.14), at $x_2 = \delta$ and $x_2 = 1 - \delta$, are shown in green. The two polygonal lines $S_{x_1}^+$, $S_{x_1}^-$, used
in (3.16) and (3.17) are shown in red. The exterior polygonal line $S_{x_1}^{\text{out}}$ used to obtain (3.21) is
shown in blue.

since by assumption $\mu \leq \theta \leq 1$. Therefore

$$\int_a^b \frac{|u_1(x_1, 0) - u_1^-| + |u_1(x_1, 1) - u_1^+|}{x_1}\, \mathrm{d}x_1 \leq \frac{3}{8} c_* \theta \ln \frac{b}{a} + \frac{12 \delta^{1/2} I(u)^{1/2}}{a^{1/2} \mu^{1/2}} \quad \text{whenever } \delta \leq a < b \leq q. \qquad (3.19)$$

We first use (3.19) with $a = m$, $b = 1$. This gives, recalling $\delta \leq \theta \leq m$,

$$\int_m^1 \frac{|u_1(x_1, 0) - u_1^-| + |u_1(x_1, 1) - u_1^+|}{x_1}\, \mathrm{d}x_1 \leq \frac{3}{8} c_* \theta \ln \frac{1}{m} + \frac{12}{\mu^{1/2}} I(u)^{1/2}.$$

If the second term is larger than $\frac{1}{8} c_* \theta \ln \frac{1}{m}$ then $I(u) \geq c \mu \theta^2 \ln \frac{1}{m}$, (3.13) holds and we are done. Otherwise the
right-hand side is not larger than $\frac{1}{2} c_* \theta \ln \frac{1}{m}$, so that

$$\int_m^1 \frac{|u_1(x_1, 0) - u_1(x_1, 1)|}{x_1}\, \mathrm{d}x_1 \leq |u_1^- - u_1^+| \ln \frac{1}{m} + \frac{1}{2} c_* \theta \ln \frac{1}{m}.$$

If

$$|u_1^- - u_1^+| \leq \frac{1}{2} c_* \theta \qquad (3.20)$$

then (3.12) holds and we are done.

It remains to consider the case that (3.20) does not hold. For $x_1 \in (0, \ell)$ we let $S_{x_1}^{\text{out}} := (\partial((-x_1, x_1) \times (-x_1, 1 + x_1))) \setminus (0, L) \times (0, 1)$ (see Fig. 10). Then

$$|u_1(x_1, 0) - u_1(x_1, 1)| \leq \int_{S_{x_1}^{\text{out}}} |\nabla u_1|\, \mathrm{d}\mathcal{H}^1 \leq \int_{S_{x_1}^{\text{out}}} f\, \mathrm{d}\mathcal{H}^1 \leq \left(\mathcal{H}^1(S_{x_1}^{\text{out}})\right)^{1/2} \left(\int_{S_{x_1}^{\text{out}}} f^2\, \mathrm{d}\mathcal{H}^1\right)^{1/2}.$$

Therefore

$$
\begin{aligned}
\int_{1/8}^{q} \frac{|u_1(x_1,0) - u_1(x_1,1)|}{x_1} \, \mathrm{d}x_1 &\leq \int_{1/8}^{q} \frac{\left(\mathcal{H}^1(S_{x_1}^{\mathrm{out}})\right)^{1/2}}{x_1} \left(\int_{S_{x_1}^{\mathrm{out}}} f^2 \, \mathrm{d}\mathcal{H}^1\right)^{1/2} \mathrm{d}x_1 \\
&\leq \left(\int_{1/8}^{q} \frac{\mathcal{H}^1(S_{x_1}^{\mathrm{out}})}{x_1^2} \, \mathrm{d}x_1\right)^{1/2} \left(\int_{0}^{q} \int_{S_{x_1}^{\mathrm{out}}} f^2 \, \mathrm{d}\mathcal{H}^1 \, \mathrm{d}x_1\right)^{1/2} \\
&\leq c\mu^{-1/2} \ln^{1/2}(8q) I^{1/2}(u),
\end{aligned}
\tag{3.21}
$$

where in the last step we used that $\mathcal{H}^1(S_{x_1}^{\mathrm{out}}) = 8x_1 + 1 \leq 16x_1$ for $x_1 \geq \frac{1}{8}$. We use (3.19) with $a = \frac{1}{8}$, $b = q$, and obtain

$$
\int_{1/8}^{q} \frac{|u_1(x_1,0) - u_1^-| + |u_1(x_1,1) - u_1^+|}{x_1} \, \mathrm{d}x_1 \leq \frac{3}{8} c_* \theta \ln(8q) + \frac{c\delta^{1/2}}{\mu^{1/2}} I^{1/2}(u) \leq \frac{3}{8} c_* \theta \ln(8q) + \frac{c \ln^{1/2}(8q) I^{1/2}(u)}{\mu^{1/2}}
$$

where we used $\delta \leq 1/8$ and $q \geq 1$. Combining with (3.21) and $\delta \leq \frac{1}{8}$ yields

$$
|u_1^- - u_1^+| \ln(8q) = \int_{1/8}^{q} \frac{|u_1^- - u_1^+|}{x_1} \, \mathrm{d}x_1 \leq \frac{3}{8} c_* \theta \ln(8q) + \frac{c \ln^{1/2}(8q) I^{1/2}(u)}{\mu^{1/2}}.
$$

Since (3.20) does not hold, we have $I(u) \geq c\mu\theta^2 \ln(8q) \geq c\mu\theta^2 \ln(3+q)$ since $q \geq 1$. Therefore (3.13) holds and the proof is concluded also in this case. □

The next Lemma proves that u_2 is, up to a small exceptional set, very well controlled by the energy. In particular, it is significantly smaller than θ, in different measures. In this section (Lem. 3.7) we shall use (i). Since the proofs are naturally connected, to avoid repetition we present here also the proof of two estimates that will be used in the next section, specifically, (iii) in Lemma 3.10 and (ii) in Lemma 3.11.

Lemma 3.6 (Local estimates for u_2). *For all $\bar{C} > 0$ there exist $C = C(\bar{C}) > 0$ such that for any $\delta \in (0,1]$ and any $u \in W_{\mathrm{loc}}^{1,2}(\mathbb{R}^2, \mathbb{R}^2)$ there is a set $\mathcal{G} \subseteq [0, L - \xi_1]$ such that*

$$
I(u) \geq C \min\{\mu, 1\} \theta^2 \mathcal{L}^1(\mathcal{G})
$$

and

i. for any $m \in (0, 1/4]$ and any $x_1 \in [0, L - \xi_1] \setminus \mathcal{G}$, one has

$$
\int_{m}^{1} \frac{|u_2(x_1 - s\xi_1, 1) - u_2(x_1 + (1+s)\xi_1, 1)|}{s} \, \mathrm{d}s \leq 3\bar{C}\theta \ln \frac{1}{m},
$$

ii. for any $m \in (0, 1/4]$ and any $x_1 \in [0, L - \xi_1] \setminus \mathcal{G}$, one has

$$
\int_{m}^{1} \frac{|u_2(x_1 + s\xi_1, s) - u_2(x_1 + (1-s)\xi_1, 1-s)|}{s} \, \mathrm{d}s \leq 6\bar{C}\theta \ln \frac{1}{m},
$$

iii. for any $x_1 \in [0, L - \xi_1] \setminus \mathcal{G}$ and any $\delta \in (0,1]$ one has

$$
\frac{1}{\delta} \int_{(0,\delta)} |u_2(x_1 + (1-s)\xi_1, 1-s) - u_2(x_1 + s\xi_1, s)| \, \mathrm{d}s \leq 5\bar{C}\theta.
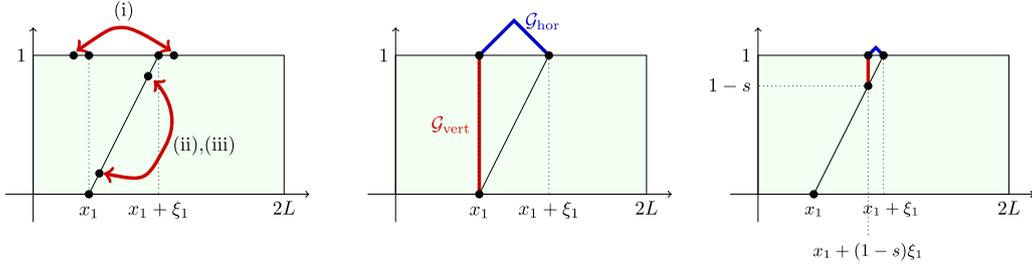$$

FIGURE 11. Sketch of the geometry in Lemma 3.6. Left panel: the boundary estimate (i) compares values across the top boundary, at distance $\xi_1(1 + 2s)$. The other two estimates compare points along the diagonal $(x_1, 0) + \mathbb{R}\xi$. Middle panel: the sets $\mathcal{G}_{\text{vert}}$ and $\mathcal{G}_{\text{hor}}$ compare points which are separated vertically or horizontally, and are estimated integrating along the sketched curves, see (3.22) and (3.23). Right panel: the sets $\mathcal{G}_{\text{hor}}^{\text{av},+}$ and $\mathcal{G}_{\text{vert}}^{\text{av}}$ are used for similar estimates, but integrated over s.

Proof.

Step 1. Construction of $\mathcal{G}$.

We construct $\mathcal{G}$ as the union of different pieces, which are all defined and estimated similarly.

The first one contains points with large vertical differences (see Fig. 11 (middle)),

$$\mathcal{G}_{\text{vert}} := \left\{ x_1 \in [0, L - \xi_1] : \bar{C}\theta \leq |u_2(x_1, 0) - u_2(x_1, 1)| \right\}. \tag{3.22}$$

By the fundamental theorem of calculus, for almost every x_1 we have

$$|u_2(x_1, 0) - u_2(x_1, 1)| \leq \int_{(0,1)} |\partial_2 u_2|(x_1, s) \, \mathrm{d}s$$

which gives, using first Hölder's inequality and then Fubini's theorem,

$$\bar{C}^2 \theta^2 \mathcal{L}^1(\mathcal{G}_{\text{vert}}) \leq \int_{(0, L - \xi_1)} |u_2(x_1, 0) - u_2(x_1, 1)|^2 \, \mathrm{d}x_1 \leq \int_{\Omega_L} |\partial_2 u_2|^2 \, \mathrm{d}x \leq I(u).$$

The second one contains points with large horizontal differences along the top boundary,

$$\mathcal{G}_{\text{hor}} := \left\{ x_1 \in [0, L - \xi_1] : \bar{C}\theta \leq |u_2(x_1, 1) - u_2(x_1 + \xi_1, 1)| \right\}. \tag{3.23}$$

Let S_{x_1} be the polygonal joining the points (see Fig. 11 (middle))

$$(x_1, 1), \left(x_1 + \frac{1}{2}\xi_1, 1 + \frac{1}{2}\xi_1 \right), (x_1 + \xi_1, 1).$$

By the fundamental theorem of calculus, for almost every x_1 we have

$$|u_2(x_1, 1) - u_2(x_1 + \xi_1, 1)| \leq \int_{S_{x_1}} |\nabla u_2| \, \mathrm{d}\mathcal{H}^1 \leq \left(\mathcal{H}^1(S_{x_1}) \int_{S_{x_1}} |\nabla u_2|^2 \, \mathrm{d}\mathcal{H}^1 \right)^{1/2}$$

so that, squaring, integrating over $\mathcal{G}_{\text{hor}}$, and using $\mathcal{H}^1(S_{x_1}) = \sqrt{2}\xi_1 = \frac{1}{2\sqrt{2}}$,

$$
\begin{aligned}
\bar{C}^2\theta^2\mathcal{L}^1(\mathcal{G}_{\text{hor}}) &\leq \int_{(0,L-\xi_1)} |u_2(x_1,1) - u_2(x_1+\xi_1,1)|^2\,\mathrm{d}x_1 \leq \frac{1}{2\sqrt{2}}\int_{\mathbb{R}}\int_{S_{x_1}} |\nabla u_2|^2\,\mathrm{d}\mathcal{H}^1\,\mathrm{d}x_1 \\
&\leq \frac{1}{2}\int_{\mathbb{R}}\int_{(0,\frac{1}{2}\xi_1)} \left[|\nabla u_2|^2(x_1+t,1+t) + |\nabla u_2|^2(x_1+\xi_1-t,1+t)\right]\mathrm{d}t\,\mathrm{d}x_1 \\
&\leq \int_{\mathbb{R}\times(1,2)} |\nabla u_2|^2\,\mathrm{d}x \leq \mu^{-1}I(u).
\end{aligned}
$$

The next one controls vertical fluctuations, it will be used to estimate $u_2(x_1+s\xi_1,0) - u_2(x_1+s\xi_1,s)$, and the same term on the other side. Precisely, we set

$$
\mathcal{G}_{\text{vert}}^{\text{av}} := \left\{ x_1 \in [0, L-\xi_1] : \bar{C}^2\theta^2 \leq \int_{(0,1)}\int_{(0,1)} \left(|\partial_2 u_2|^2(x_1+s\xi_1,t) + |\partial_2 u_2|^2(x_1+(1-s)\xi_1,t)\right)\mathrm{d}s\,\mathrm{d}t \right\}. \tag{3.24}
$$

Integrating over all $x_1 \in \mathcal{G}_{\text{vert}}^{\text{av}}$ and swapping the order of integration gives

$$
\bar{C}^2\theta^2\mathcal{L}^1(\mathcal{G}_{\text{vert}}^{\text{av}}) \leq 2\int_{(0,1)} \mathrm{d}s \int_{(0,L)}\int_{(0,1)} |\partial_2 u_2|^2(x_1,t)\,\mathrm{d}t\,\mathrm{d}x_1 \leq 2\int_{\Omega_L} |\partial_2 u_2|^2\,\mathrm{d}x \leq 2I(u).
$$

And finally we consider an averaged version of $\mathcal{G}_{\text{hor}}$,

$$
\mathcal{G}_{\text{hor}}^{\text{av},+} := \left\{ x_1 \in \mathbb{R} : \bar{C}^2\theta^2 \leq \int_{(-1,1)} \frac{|u_2(x_1+s\xi_1,1) - u_2(x_1,1)|^2}{|s|}\,\mathrm{d}s \right\}. \tag{3.25}
$$

Let $S_{x_1,s}^+$ be the polygonal line which joins the points

$$
(x_1,1), \ \left(x_1 + \frac{1}{2}s\xi_1, 1 + \frac{1}{2}|s|\xi_1\right), \ (x_1 + s\xi_1, 1)
$$

(see Fig. 11). By the fundamental theorem of calculus for Sobolev functions, Hölder's inequality and $\mathcal{H}^1(S_{x_1,s}^+) = |s|\xi_1\sqrt{2} \leq |s|$ we get

$$
|u_2(x_1+s\xi_1,1) - u_2(x_1,1)| \leq \int_{S_{x_1,s}^+} |\nabla u_2|\,\mathrm{d}\mathcal{H}^1 \leq |s|^{1/2}\left(\int_{S_{x_1,s}^+} |\nabla u_2|^2\,\mathrm{d}\mathcal{H}^1\right)^{1/2}.
$$

Squaring, dividing by $|s|$ and integrating over $s \in (-1,1)$ gives

$$
\int_{(-1,1)} \frac{|u_2(x_1+s\xi_1,1) - u_2(x_1,1)|^2}{|s|}\,\mathrm{d}s \leq \int_{(-1,1)}\int_{S_{x_1,s}^+} |\nabla u_2|^2\,\mathrm{d}\mathcal{H}^1\,\mathrm{d}s.
$$

We integrate over all $x_1 \in \mathcal{G}_{\text{hor}}^{\text{av},+}$ and obtain, using Fubini's theorem as above,

$$
\begin{aligned}
\bar{C}^2 \theta^2 \mathcal{L}^1(\mathcal{G}_{\text{hor}}^{\text{av},+}) &\leq \int_{\mathbb{R}} \int_{(-1,1)} \frac{|u_2(x_1 + s\xi_1, 1) - u_2(x_1, 1)|^2}{|s|} \, \mathrm{d}s \, \mathrm{d}x_1 \\
&\leq \int_{\mathbb{R}} \int_{(-1,1)} \int_{S_{x_1,s}^+} |\nabla u_2|^2 \, \mathrm{d}\mathcal{H}^1 \, \mathrm{d}s \, \mathrm{d}x_1 \\
&\leq 2\sqrt{2} \int_{(-1,1)} \int_{(0,\frac{1}{2}|s|\xi_1)} \int_{\mathbb{R}} |\nabla u_2|^2 (x_1, 1+t) \, \mathrm{d}x_1 \, \mathrm{d}t \, \mathrm{d}s \\
&\leq 4\sqrt{2} \int_{\mathbb{R} \times (1,2)} |\nabla u_2|^2 \, \mathrm{d}x \leq 6\mu^{-1} I(u).
\end{aligned}
$$

The analogue estimate holds for

$$
\mathcal{G}_{\text{hor}}^{\text{av},-} := \left\{ x_1 \in \mathbb{R} : \bar{C}^2 \theta^2 \leq \int_{(-1,1)} \frac{|u_2(x_1 + s\xi_1, 0) - u_2(x_1, 0)|^2}{|s|} \, \mathrm{d}s \right\}. \tag{3.26}
$$

We finally define

$$
\mathcal{G} := \mathcal{G}_{\text{vert}} \cup \mathcal{G}_{\text{hor}} \cup \mathcal{G}_{\text{vert}}^{\text{av}} \cup \mathcal{G}_{\text{hor}}^{\text{av},+} \cup \mathcal{G}_{\text{hor}}^{\text{av},-} \cup (\xi_1 + \mathcal{G}_{\text{hor}}^{\text{av},+}).
$$

The previous estimates imply $I(u) \geq C_4 \min\{\theta^2, \mu\theta^2\} \mathcal{L}^1(\mathcal{G})$, with $C_4 := \frac{1}{36}\bar{C}^2$.

Step 2. Proof of (i) (estimate on the boundary). For any $f \in L^2((0,1))$ one has for $m \in (0,1)$

$$
\int_{(m,1)} \frac{|f|}{|s|} \, \mathrm{d}s \leq \left(\int_{(m,1)} \frac{1}{|s|} \, \mathrm{d}s \right)^{1/2} \left(\int_{(m,1)} \frac{|f|^2}{|s|} \, \mathrm{d}s \right)^{1/2} \leq \ln^{1/2} \frac{1}{m} \left(\int_{(0,1)} \frac{|f|^2}{|s|} \, \mathrm{d}s \right)^{1/2}, \tag{3.27}
$$

and similarly on $(-1, -m)$. Therefore, for $x_1 \in \mathbb{R} \setminus \mathcal{G}_{\text{hor}}^{\text{av},+}$ (recall (3.25)) we have

$$
\int_{(m,1)} \frac{|u_2(x_1 + s\xi_1, 1) - u_2(x_1, 1)|}{|s|} \, \mathrm{d}s \leq \bar{C}\theta \ln^{1/2} \frac{1}{m}, \tag{3.28}
$$

and analogously

$$
\int_{(m,1)} \frac{|u_2(x_1 - s\xi_1, 1) - u_2(x_1, 1)|}{|s|} \, \mathrm{d}s \leq \bar{C}\theta \ln^{1/2} \frac{1}{m}. \tag{3.29}
$$

For later reference we notice that a similar computation shows that for $x_1 \in \mathbb{R} \setminus \mathcal{G}_{\text{hor}}^{\text{av},-}$ (recall (3.26)) we have

$$
\int_{(m,1)} \frac{|u_2(x_1 + s\xi_1, 0) - u_2(x_1, 0)|}{|s|} \, \mathrm{d}s \leq \bar{C}\theta \ln^{1/2} \frac{1}{m}. \tag{3.30}
$$

For $x_1 \in [0, L - \xi_1] \setminus \mathcal{G}$ we have

$$\int_m^1 \frac{|u_2(x_1 - s\xi_1, 1) - u_2(x_1 + (1+s)\xi_1, 1)|}{s} \, \mathrm{d}s \leq \int_m^1 \frac{|u_2(x_1 - s\xi_1, 1) - u_2(x_1, 1)|}{s} \, \mathrm{d}s$$

$$+ |u_2(x_1, 1) - u_2(x_1 + \xi_1, 1)| \int_m^1 \frac{1}{s} \, \mathrm{d}s$$

$$+ \int_m^1 \frac{|u_2(x_1 + \xi_1, 1) - u_2(x_1 + (1+s)\xi_1, 1)|}{s} \, \mathrm{d}s$$

$$\leq 3\bar{C}\theta \ln \frac{1}{m},$$

where we used $x_1 \notin \mathcal{G}_{\mathrm{hor}}^{\mathrm{av},+}$ and (3.29) to estimate the first term, $x_1 \notin \mathcal{G}_{\mathrm{hor}}$ to estimate the second one, and $x_1 + \xi_1 \notin \mathcal{G}_{\mathrm{hor}}^{\mathrm{av},+}$ and (3.28) to estimate the third one, and then $\ln^{1/2} \frac{1}{m} \leq \ln \frac{1}{m}$ to simplify the estimate.

This concludes the proof of (i).

Step 3. Proof of (ii) (estimate on the diagonal).

Let $x_1 \in [0, L - \xi_1] \setminus \mathcal{G}$. For any $s \in (0, 1)$ we have, by the fundamental theorem of calculus and Hölder's inequality,

$$\frac{|u_2(x_1 + s\xi_1, s) - u_2(x_1 + s\xi_1, 0)|^2}{s} \leq \frac{1}{s} \left(\int_{(0,s)} |\partial_2 u_2|(x_1 + s\xi_1, t) \, \mathrm{d}t \right)^2 \leq \int_{(0,1)} |\partial_2 u_2|^2(x_1 + s\xi_1, t) \, \mathrm{d}t.$$

Integrating over $s \in (m, 1)$ and using $x_1 \notin \mathcal{G}_{\mathrm{vert}}^{\mathrm{av}}$ (recall (3.24)),

$$\int_0^1 \frac{|u_2(x_1 + s\xi_1, s) - u_2(x_1 + s\xi_1, 0)|^2}{s} \, \mathrm{d}s \leq \bar{C}^2\theta^2,$$

and combining with $x_1 \notin \mathcal{G}_{\mathrm{hor}}^{\mathrm{av},-}$ (recall (3.26),

$$\int_0^1 \frac{|u_2(x_1 + s\xi_1, s) - u_2(x_1, 0)|^2}{s} \, \mathrm{d}s \leq 4\bar{C}^2\theta^2. \tag{3.31}$$

The same estimate on the other side gives, using $x_1 + \xi_1 \notin \mathcal{G}_{\mathrm{hor}}^{\mathrm{av},+}$ (recall (3.25)),

$$\int_0^1 \frac{|u_2(x_1 + (1-s)\xi_1, 1-s) - u_2(x_1 + \xi_1, 1)|^2}{s} \, \mathrm{d}s \leq 4\bar{C}^2\theta^2. \tag{3.32}$$

By the triangular inequality,

$$\int_m^1 \frac{|u_2(x_1 + s\xi_1, s) - u_2(x_1 + (1-s)\xi_1, 1-s)|}{s} \, \mathrm{d}s \leq \int_m^1 \frac{|u_2(x_1 + s\xi_1, s) - u_2(x_1, 0)|}{s} \, \mathrm{d}s$$

$$+ |u_2(x_1, 0) - u_2(x_1 + \xi_1, 1)| \int_m^1 \frac{1}{s} \, \mathrm{d}s$$

$$+ \int_m^1 \frac{|u_2(x_1 + \xi_1, 1) - u_2(x_1 + (1-s)\xi_1, 1-s)|}{s} \, \mathrm{d}s.$$

By (3.31) and (3.27), the first term is estimated by $2\bar{C}\theta \ln^{1/2}\frac{1}{m} \leq 2\bar{C}\theta \ln\frac{1}{m}$. The same holds for the last one, by (3.32) and (3.27). For the middle one we use $x_1 \notin \mathcal{G}_{\text{vert}}$ and $x_1 \notin \mathcal{G}_{\text{hor}}$, which give $|u_2(x_1, 0) - u_2(x_1 + \xi_1, 1)| \leq 2\bar{C}\theta$. Adding these three estimates leads to

$$\int_m^1 \frac{|u_2(x_1 + s\xi_1, s) - u_2(x_1 + (1-s)\xi_1, 1-s)|}{s}\,\mathrm{d}s \leq 6\bar{C}\theta \ln\frac{1}{m},$$

which concludes the proof.

Step 4. Proof of (iii) (estimate close to the boundary).

For any $f \in L^2((0,1))$ and any $\delta \in (0,1)$ one has

$$\frac{1}{\delta}\int_{(0,\delta)}|f|\,\mathrm{d}s \leq \frac{1}{\delta}\Big(\int_{(0,\delta)}\frac{|f|^2}{|s|}\,\mathrm{d}s\Big)^{1/2}\Big(\int_{(0,\delta)}|s|\,\mathrm{d}s\Big)^{1/2} \leq \frac{1}{\sqrt{2}}\Big(\int_{(0,1)}\frac{|f|^2}{|s|}\,\mathrm{d}s\Big)^{1/2}, \qquad (3.33)$$

and analogously on $(-\delta, 0)$. Let $x_1 \in [0, L-\xi_1] \setminus \mathcal{G}$. Using (3.33) and (3.32) we obtain

$$\frac{1}{\delta}\int_{(0,\delta)}|u_2(x_1 + (1-s)\xi_1, 1-s) - u_2(x_1 + \xi_1, 1)|\,\mathrm{d}s \leq \sqrt{2}\bar{C}\theta.$$

Analogously, with (3.33) and (3.31) we obtain

$$\frac{1}{\delta}\int_{(0,\delta)}|u_2(x_1 + s\xi_1, s) - u_2(x_1, 0)|\,\mathrm{d}s \leq \sqrt{2}\bar{C}\theta.$$

As above, $x_1 \notin \mathcal{G}_{\text{vert}}$ and $x_1 \notin \mathcal{G}_{\text{hor}}$ give $|u_2(x_1, 0) - u_2(x_1 + \xi_1, 1)| \leq 2\bar{C}\theta$, so that by triangle inequality

$$\frac{1}{\delta}\int_{(0,\delta)}|u_2(x_1 + (1-s)\xi_1, 1-s) - u_2(x_1 + s\xi_1, s)|\,\mathrm{d}s \leq 5\bar{C}\theta$$

which concludes the proof of (iii).

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

At this point we are ready to present the main result of this section, which basically gives the proof of the lower bound in the cases with fine microstructure and small θ. Following [27] we introduce two new parameters, $\lambda, m > 0$, which will be chosen below (see the proof of Prop. 3.3) in different ways depending on the regime. This permits to unify different parts of the proof of the lower bound. Roughly speaking, the parameters λ and m correspond to the length scales of the martensitic laminate deep inside the nucleus and on the vertical austenite/martensite interface.

Lemma 3.7 (The case of small θ). *There exists $m_0 \in (0, 1/4]$ and $c > 0$ such that for all $u \in \mathcal{X}$, $\theta > 0$, $\mu > 0$, $\varepsilon > 0$, $\lambda > 0$, $\ell > 0$, and $m > 0$ which obey*

$$\theta \leq m \leq m_0\,, \qquad 1 \leq \ell \leq 2L\,, \qquad \varepsilon \leq \theta^2\lambda\,, \qquad and \qquad \lambda \leq 1$$

one has

$$I_{\bar{\Omega}_\ell}(u) \geq c\min\Big\{\frac{\varepsilon\ell}{\lambda}, \mu m^2, \mu\theta^2\lambda\ln\frac{1}{m}, \theta^2\ell^{-1}\lambda^2 m\Big(\ln\frac{1}{m}\Big)^2, \mu\theta^2\ln\Big(3 + \frac{\theta}{\mu}\Big), \mu\theta^2\ln(3+\ell)\Big\}.$$

This proof follows the strategy of Section 5.2 in [27], with important modifications to treat both the vectorial nature of this problem and the additional logarithmic terms which appear here, due to the different boundary conditions and the fact that we do not have a hard constraint on the order parameter.

Proof of Lemma 3.7.
Step 1: Preparation.
We start by choosing a 'good' slice, parametrized as usual by x_1. The slice is chosen so as to have boundary values at the upper and lower boundary of Ω_{2L} which are close together, in a sense similar to the one used in the definition of the set $\mathcal{P}$ defined in (3.2). However, in order to capture the different logarithmic factor, we need to use a larger set, in which the difference between the boundary values is controlled in the scale of $m \geq \theta$. Specifically, we consider the set

$$\mathcal{P}_* := \left\{ x_1 \in (0, \ell - \xi_1) \, : \, |u((x_1, 0) + \xi) - u(x_1, 0)| \leq \frac{1}{10} m \right\}.$$

For almost every $x_1 \in (0, \ell - \xi_1) \setminus \mathcal{P}_*$ we obtain by Lemma 3.2(i) that

$$\mu \int_{S_{x_1}} |\nabla u|^2 \; d\mathcal{H}^1 \geq c \frac{\mu m^2}{1 + x_1}$$

so that, using Fubini and monotonicity of $1/(1 + x_1)$ as in the proof of Lemma 3.2(iv),

$$I_{\tilde{\Omega}_\ell}^{\text{ext}}(u) \geq c\mu m^2 \int_{(0, \ell - \xi_1) \setminus \mathcal{P}_*} \frac{1}{x_1 + 1} \; dx_1 \geq c\mu m^2 \int_p^{\ell - \xi_1} \frac{1}{x_1 + 1} \; dx_1 = c\mu m^2 \ln \frac{\ell + 1 - \xi_1}{p + 1},$$

where $p := \mathcal{L}^1(\mathcal{P}_*)$. If $p \leq \frac{1}{2}\ell$, then, recalling $1 \leq \ell$, we see that $p + 1 \leq \frac{1}{2}\ell + 1 \leq \frac{1}{2}\ell + \frac{5}{14}\ell + \frac{9}{14} = \frac{6}{7}(\ell + \frac{3}{4})$. Therefore in this case $\ln \frac{\ell + 1 - \xi_1}{p+1} \geq \ln \frac{7}{6} > 0$, so that $I(u) \geq c\mu m^2$ and we are done. Therefore we can assume $\mathcal{L}^1(\mathcal{P}_*) > \frac{1}{2}\ell$ in the following.

We set $\hat{C} := 2^{-10}$ and consider the set where the surface energy or the elastic energy are large along slices in the ξ direction. For $x_1 \in (0, \ell - \xi_1)$ we recall that in (1.6) we defined $u_{x_1}^\xi : [0, 1] \to \mathbb{R}^2$ by $u_{x_1}^\xi := (4u_1 + u_2)((x_1, 0) + s\xi)$. We set

$$\mathcal{R} := \left\{ x_1 \in (0, \ell - \xi_1) \, : \, |\partial_s \partial_s u_{x_1}^\xi|((0,1)) \geq \frac{1}{4}\hat{C}\lambda^{-1} \text{ or } \left\| \min\left\{ |\partial_s u_{x_1}^\xi - \theta|, |\partial_s u_{x_1}^\xi + (1 - \theta)| \right\} \right\|_{L^2((0,1))}^2 \geq \hat{C}\varepsilon\lambda^{-1} \right\}.$$

By (1.11) and Fubini's theorem,

$$I_{\tilde{\Omega}_\ell}(u) \geq c\mathcal{L}^1(\mathcal{R})\frac{\varepsilon}{\lambda}.$$

If $\mathcal{L}^1(\mathcal{R}) \geq \frac{1}{8}\ell$, then we have $I_{\tilde{\Omega}_\ell}(u) \geq c\varepsilon\ell\lambda^{-1}$ and the proof is concluded.

Let $\mathcal{G}$ be as in Lemma 3.6, with $\bar{C} = 2^{-7}$. If $\mathcal{L}^1(\mathcal{G}) \geq \frac{1}{8}\ell$ then $I_{\tilde{\Omega}_\ell}(u) \geq c \min\{\theta^2\ell, \mu\theta^2\ell\} \geq c \min\{\theta^2\ell^{-1}\lambda^2 m \ln^2 \frac{1}{m}, \mu\theta^2 \ln(3 + \ell)\}$ and we are done (recall that $\ell^{-1}\lambda^2 m \ln^2 \frac{1}{m} \leq \ell$ and $\ln(3 + \ell) \leq c\ell$ by our assumptions). Therefore we can assume $\mathcal{L}^1(\mathcal{R} \cup \mathcal{G}) < \frac{1}{4}\ell$.

We choose $x_1^* \in \mathcal{P}_* \setminus \mathcal{R} \setminus \mathcal{G}$ such that $v := u_{x_1^*}^\xi \in W^{1,2}((0,1))$ is the trace of u, which necessarily satisfies

$$|v(1) - v(0)| \leq \frac{1}{2}m, \qquad |\partial_s \partial_s v|((0,1)) < \frac{1}{4}\hat{C}\lambda^{-1}, \qquad \left\| \min\left\{ |v' - \theta|, |v' + (1 - \theta)| \right\} \right\|_{L^2((0,1))}^2 < \hat{C}\varepsilon\lambda^{-1} \quad (3.34)$$

and, since $x_1^* \notin \mathcal{G}$, by Lemma 3.6(i)

$$\int_{(m,1)} \frac{1}{s} |u_2(x_1^* - s\xi_1, 1) - u_2(x_1^* + (s+1)\xi_1, 1)| \, \mathrm{d}s \leq \frac{1}{32} \theta \ln \frac{1}{m}. \tag{3.35}$$

For $t \in \mathbb{R}$, we define $\omega_t := \{s \in (0,1) \, : \, v'(s) \leq \theta - t\}$ and $\chi_t := \chi_{\omega_t}$. We observe that

$$|v'(s) - \theta + \chi_{1/2}(s)| = \min\{|v'(s) - \theta|, |v'(s) + (1 - \theta)|\}. \tag{3.36}$$

By the coarea formula we have

$$\int_{\frac{1}{2}}^{\frac{3}{4}} \mathcal{H}^0\left(\partial \omega_t \cap (0,1)\right) \, \mathrm{d}t \leq \int_{\mathbb{R}} \mathcal{H}^0\left(\partial \omega_t \cap (0,1)\right) \, \mathrm{d}t = |\partial_s \partial_s v| \left((0,1)\right),$$

and hence, by (3.34), there is $t^* \in (1/2, 3/4)$ such that $\omega := \omega_{t^*}$ consists of at most $\hat{C}\lambda^{-1}$ many intervals.

We compute

$$v(1) - v(0) = \int_{(0,1)} v' \, \mathrm{d}s = \theta - \mathcal{L}^1(\omega_{1/2}) + \int_{(0,1)} (v' - \theta + \chi_{1/2}) \, \mathrm{d}s$$

which, by Hölder's inequality, (3.34), $\varepsilon \leq \theta^2 \lambda$, (3.36), and the choice of $\hat{C}$ gives

$$|v(1) - v(0) - \theta + \mathcal{L}^1(\omega_{1/2})| \leq \|v' - \theta + \chi_{1/2}\|_{L^1((0,1))} \leq \hat{C}^{1/2} \varepsilon^{1/2} \lambda^{-1/2} \leq 2^{-5}\theta.$$

Using (3.34), $\omega \subset \omega_{1/2}$ and $\varepsilon \leq \theta^2 \lambda$ as above, (note that $v'(s) - \theta \in [-3/4, -1/2]$ implies $v'(s) + 1 - \theta \in [1/4, 1/2]$)

$$0 \leq \mathcal{L}^1(\omega_{1/2}) - \mathcal{L}^1(\omega) \leq \mathcal{L}^1(\{s : v'(s) - \theta \in [-3/4, -1/2]\}) \leq 16 \int_0^1 \min\{|v' - \theta|^2, |v' + (1-\theta)|^2\} \, \mathrm{d}s \tag{3.37}$$
$$\leq 16\hat{C}\varepsilon\lambda^{-1} \leq 16\hat{C}\theta^2 \leq 2^{-6}\theta,$$

so that

$$|v(1) - v(0) - \theta + \mathcal{L}^1(\omega)| \leq |v(1) - v(0) - \theta + \mathcal{L}^1(\omega_{1/2})| + |\mathcal{L}^1(\omega_{1/2}) - \mathcal{L}^1(\omega_1)| \leq 2^{-4}\theta. \tag{3.38}$$

We conclude that ω consists of at most $\hat{C}\lambda^{-1}$ many intervals and obeys (recall (3.34) and $\theta \leq m$)

$$\mathcal{L}^1(\omega) \leq |v(1) - v(0)| + \theta + 2^{-4}\theta \leq 2m. \tag{3.39}$$

Step 2: A test function for the logarithmic scaling in the interior.
We denote the connected components of ω by $(y_i - r_i, y_i + r_i)$ and define $g_i := \min\{r_i + \lambda m, m\}$. Notice that $2r_i \leq \mathcal{L}^1(\omega) \leq 2m$ implies $r_i \leq g_i$ for all i. Recall that the number n of these components is at most $\hat{C}\lambda^{-1}$. We then have

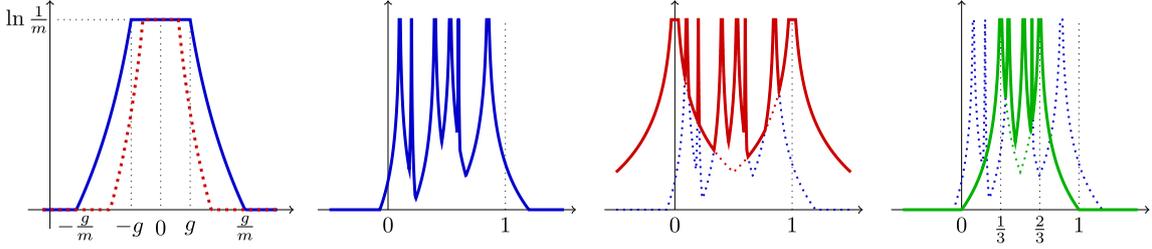$$\lambda m \leq g_i \leq m \qquad \text{for all } i, \tag{3.40}$$

FIGURE 12. Sketch of the construction of the test function ψ in the proof of Lemma 3.7. From left to right: the first panel shows two of the functions ψ_i with different g_i, the second one the function $\hat{\psi}$ defined in Step 2, the third one the function ψ defined in Step 3, which equals $\ln \frac{1}{m}$ at the boundary points, and the fourth one the function ψ defined in Step 4, which equals $\ln \frac{1}{m}$ at $\frac{1}{3}$ and $\frac{2}{3}$ and vanishes at the boundary points. In the last two panels $\hat{\psi}$, $\max\{\psi_B(t), \psi_B(1-t)\}$ and $\max\{\psi_T(t - \frac{1}{3}), \psi_T(t - \frac{2}{3})\}$ are shown dotted for comparison. The vertical axis is compressed for clarity.

and

$$\sum_{i=1}^{n} 2g_i \leq \mathcal{L}^1(\omega) + 2n\lambda m \leq 2m + 2\hat{C}m \leq 3m.$$

We consider the test function $\hat{\psi} : \mathbb{R} \to \mathbb{R}$ defined by (see Fig. 12)

$$\hat{\psi}(t) := \max_{1,\ldots n} \psi_i(t - y_i), \qquad \psi_i(t) := \left[\ln \frac{1}{m} - \left(\ln \frac{|t|}{g_i}\right)_+\right]_+ = \begin{cases} \ln \frac{1}{m}, & \text{if } |t| \leq g_i, \\ \ln \frac{g_i}{m|t|}, & \text{if } g_i < |t| \leq \frac{g_i}{m}, \\ 0, & \text{if } |t| > \frac{g_i}{m}, \end{cases}$$

where $a_+ := \max\{a, 0\}$. Since $m_0 \leq 1/4$, we have $\ln 1/m \geq \ln 4 > 0$ for any $m \leq m_0$. Recalling $y_i \in \omega \subset (0, 1)$ and $g_i \leq m$ we see that $\operatorname{supp} \hat{\psi} \subset (-1, 2)$.

We compute

$$\|\psi_i\|_{L^1(\mathbb{R})} \leq 2\int_0^{g_i/m} \ln \frac{g_i}{mt}\, dt = \frac{2g_i}{m} \int_0^1 \ln \frac{1}{t}\, dt = \frac{2g_i}{m},$$

and analogously $\|\psi_i\|_{L^2(\mathbb{R})}^2 \leq 4g_i/m$, which imply, recalling that $\sum_i 2g_i \leq 3m$,

$$\|\hat{\psi}\|_{L^1(\mathbb{R})} \leq \sum_{i=1}^{n} \|\psi_i\|_{L^1(\mathbb{R})} \leq \frac{1}{m}\sum_{i=1}^{n} 2g_i \leq 3 \qquad \text{and} \qquad \|\hat{\psi}\|_{L^2(\mathbb{R})}^2 \leq \int_{\mathbb{R}} \max_i |\psi_i|^2(t - y_i)\, dt \leq \sum_{i=1}^{n} \|\psi_i\|_{L^2(\mathbb{R})}^2 \leq 6.$$

To estimate the L^2 norm of $\hat{\psi}'$ we first compute

$$|\psi_i'|(t) = \frac{1}{|t|}\chi_{\{g_i \leq |t| \leq g_i/m\}}$$

and observe that $|\hat{\psi}'|^2(t) \le \max_i |\psi_i'|^2(t-y_i) \le \sum_i |\psi_i'|^2(t-y_i)$. This implies

$$\|\hat{\psi}'\|_{L^2(\mathbb{R})}^2 \le \sum_{i=1}^n \int_{\mathbb{R}} |\psi_i'(t)|^2 \, \mathrm{d}t = 2 \sum_{i=1}^n \int_{g_i}^{g_i/m} \frac{1}{t^2} \, \mathrm{d}t \le 2 \sum_{i=1}^n \frac{1}{g_i} \le \frac{2n}{\lambda m} \le \frac{1}{\lambda^2 m}.$$

We then estimate the $H^{1/2}$ norm of $\hat{\psi}$. Specifically, we define an extension and estimate its homogeneous H^1 norm. Following Lemma 5.2 in [27] we let $\Psi_i(x) := \psi_i(|x|)$ be the radially symmetric extension of ψ_i to $\mathbb{R}^2$ and compute

$$\int_{\mathbb{R}^2} |\nabla \Psi_i|^2 \, \mathrm{d}x = 2\pi \int_{g_i}^{g_i/m} r \frac{1}{r^2} \, \mathrm{d}r = 2\pi \ln \frac{1}{m}.$$

We define the function $\hat{\Psi}(x_1, x_2) := \max_i \Psi_i(x_1, x_2 - y_i)$, which obeys $\hat{\Psi}(0, t) = \hat{\psi}(t)$ for $t \in \mathbb{R}$ and $|\nabla \hat{\Psi}|(x) \le \max_i |\nabla \Psi_i|(x_1, x_2 - y_i)$ for almost every $x \in \mathbb{R}^2$. This implies (recall (3.40) and $n \le \hat{C}\lambda^{-1} \le (2\pi)^{-1}\lambda^{-1}$)

$$\|\nabla \hat{\Psi}\|_{L^2(\mathbb{R}^2)}^2 = \int_{\mathbb{R}^2} |\nabla \hat{\Psi}|^2 \, \mathrm{d}x \le \sum_{i=1}^n \int_{\mathbb{R}^2} |\nabla \Psi_i(x_1, x_2 - y_i)|^2 \, \mathrm{d}x = 2\pi n \ln \frac{1}{m} \le \frac{1}{\lambda} \ln \frac{1}{m}.$$

Step 3: Boundary correction for $\mu \le \theta$.
In this situation we take the largest value, $\ln \frac{1}{m}$. Specifically, we set $\psi(t) := \max\{\hat{\psi}(t), \psi_B(t), \psi_B(1-t)\}$, where

$$\psi_B(t) := \left[\ln \frac{1}{m} - \left(\ln \frac{|t|}{m} \right)_+ \right]_+ = \begin{cases} \ln \frac{1}{m}, & \text{if } |t| \le m, \\ \ln \frac{1}{|t|}, & \text{if } m < |t| \le 1, \\ 0, & \text{if } |t| > 1. \end{cases}$$

We remark that ψ_B has the same form as the functions ψ_i, with the only difference that the width of the central region is not $g_i \in [\lambda m, m]$ but exactly m. This is important to ensure symmetry of the boundary conditions.

One computes $\|\psi_B\|_{L^1((0,1))} \le 1$, $\|\psi_B\|_{L^2((0,1))}^2 \le 2$ and $\|\psi_B'\|_{L^2((0,1))}^2 \le \frac{1}{m}$. The previous estimates for $\hat{\psi}$ lead then to

$$\|\psi\|_{L^1((0,1))} \le 5, \quad \|\psi\|_{L^2((0,1))}^2 \le 10, \quad \|\psi'\|_{L^2(\mathbb{R})} \le \frac{3}{\lambda m^{1/2}}, \quad \|\nabla \Psi\|_{L^2(\mathbb{R}^2)} \le c \frac{1}{\lambda^{1/2}} \ln^{1/2} \frac{1}{m} \qquad (3.41)$$

where $\Psi(x) := \max\{\hat{\Psi}(x), \psi_B(|x|), \psi_B(|x - e_2|)\}$ obeys $\operatorname{supp} \Psi \subseteq [-1,1] \times [-1,2]$, $\Psi(0,t) = \psi(t)$, $\Psi(t,0) = \Psi(t,1) = \psi_B(t)$. Here it is important that $g_i \le m$, so that, in taking the maximum, ψ_B is always the larger one on the top and bottom boundaries, $x_2 \in \{0,1\}$.

Since we are working in the case $\mu \le \theta$, from Lemma 3.5 with $c_* := \frac{1}{8}$ we obtain the following: either

$$\frac{1}{c} I(u) \ge \min\{\mu\theta^2 \ln\left(3 + \frac{\theta}{\mu}\right), \mu\theta^2 \ln \frac{1}{m}, \mu\theta^2 \ln(3 + \ell)\}$$

and (since $\lambda \le 1$) we are done, or

$$\int_m^1 \frac{|(u_1(s,1) - u_1(s,0))|}{s} \, \mathrm{d}s \le \frac{1}{8} \theta \ln \frac{1}{m}.$$

Since $\psi_B' = 0$ on $(0, m)$ and $\psi_B'(s) = -1/s$ on $(m, 1)$, this implies

$$\int_{(0,1)} \psi_B'(s) \big[u_1(s, 1) - u_1(s, 0) \big] \, ds \leq \frac{1}{8} \theta \ln \frac{1}{m}. \tag{3.42}$$

We now turn to u_2 and recall that (3.35) implies

$$\int_{(0,1)} \psi_B'(s) \big[u_2(x_1^* - s\xi_1, 1) - u_2(x_1^* + (s+1)\xi_1, 1) \big] \, ds \leq \frac{1}{32} \theta \ln \frac{1}{m}. \tag{3.43}$$

Step 4: Energy estimate for $\mu \leq \theta$.

We compute, recalling that $\omega \subset \omega_{1/2}$ and that $\psi = \ln \frac{1}{m}$ on $\omega \cup \{0, 1\}$,

$$
\begin{aligned}
\mathcal{L}^1(\omega) \ln \frac{1}{m} &= \int_{(0,1)} \chi_\omega \psi \, ds \leq \int_{(0,1)} \chi_{\omega_{1/2}} \psi \, ds \\
&= \int_{(0,1)} (\chi_{\omega_{1/2}} - \theta + v') \psi \, ds + \int_{(0,1)} (\theta - v') \psi \, ds \\
&\leq \| v' + \chi_{\omega_{1/2}} - \theta \|_{L^2((0,1))} \| \psi \|_{L^2((0,1))} + \theta \| \psi \|_{L^1((0,1))} + (v(0) - v(1)) \ln \frac{1}{m} + \int_{(0,1)} v\psi' \, ds,
\end{aligned}
$$

where we integrated by parts in the last term. We recall that (3.38) gives

$$(v(0) - v(1)) \ln \frac{1}{m} \leq (\mathcal{L}^1(\omega) - \theta + 2^{-4}\theta) \ln \frac{1}{m}$$

and that $\| v' + \chi_{\omega_{1/2}} - \theta \|_{L^2((0,1))} \leq 2^{-5}\theta$ by (3.34) using $\varepsilon\lambda^{-1} \leq \theta^2$.
Inserting in the previous expression gives

$$\theta \ln \frac{1}{m} \leq 2^{-5}\theta \| \psi \|_{L^2((0,1))} + \theta \| \psi \|_{L^1((0,1))} + 2^{-4}\theta \ln \frac{1}{m} + \int_{(0,1)} v\psi' \, ds.$$

Since the estimates in (3.41) give $\| \psi \|_{L^2((0,1))} \leq 4$ and $\| \psi \|_{L^1((0,1))} \leq 5$, choosing m_0 such that $5 \leq 2^{-4} \ln \frac{1}{m}$ for $m \in (0, m_0]$ we get

$$\frac{1}{2} \theta \ln \frac{1}{m} \leq \int_{(0,1)} v\psi' \, ds.$$

Using the definition of v, this gives

$$\frac{1}{2} \theta \ln \frac{1}{m} \leq \int_{(0,1)} u_1((x_1^*, 0) + s\xi) \psi'(s) \, ds + 4 \int_{(0,1)} u_2((x_1^*, 0) + s\xi) \psi'(s) \, ds. \tag{3.44}$$

At this point we distinguish two cases, depending on which of the two terms in (3.44) is larger. In the first case, by the fundamental theorem and the trace theorem we have

$$\frac{1}{4} \theta \ln \frac{1}{m} \leq \int_{(0,1)} (u_1(x_1^*, 0) + s\xi) \psi'(s) \, ds = \int_{(0,1)} \int_0^{x_1^* + s\xi_1} \partial_1 u_1(x_1, s\xi_2) \, dx_1 \, \psi'(s) \, ds + \int_{(0,1)} u_1(0, s) \psi'(s) \, ds.$$

The first integral can be estimated by $\|\psi'\|_{L^2((0,1))}\ell^{1/2}I^{1/2}(u)$. Recalling (3.42),

$$\frac{1}{8}\theta\ln\frac{1}{m} \le \|\psi'\|_{L^2((0,1))}\ell^{1/2}I^{1/2}(u) + \int_{(0,1)}u_1(0,s)\psi'(s)\,\mathrm{d}s + \int_{(0,1)}(u_1(s,1)-u_1(s,0))\psi'_B(s)\,\mathrm{d}s.$$

For brevity in we write here $I(u)$ for $I_{\tilde{\Omega}_\ell}(u)$. We define $F_1 := (-1,1)\times(-1,2)\setminus[0,1]^2$ and observe that the last two integrals can be written as a boundary integral of u_1 times the tangential derivative $\partial_\tau\Psi$, and that Ψ vanishes on the rest of the boundary of F_1. Therefore

$$\frac{1}{8}\theta\ln\frac{1}{m} \le \int_{\partial F_1}\partial_\tau\psi u_1\,\mathrm{d}\mathcal{H}^1 + \|\psi'\|_{L^2((0,1))}\ell^{1/2}I^{1/2}(u).$$

With Lemma 1.3 and the estimates for ψ in (3.41) this gives

$$\frac{1}{8}\theta\ln\frac{1}{m} \le \|\nabla\Psi\|_{L^2(F_1)}\|\nabla u_1\|_{L^2(F_1)} + \|\psi'\|_{L^2((0,1))}\ell^{1/2}I^{1/2}(u) \le c\frac{1}{\lambda^{1/2}}\ln^{1/2}\frac{1}{m}\mu^{-1/2}I^{1/2}(u)$$
$$+c\frac{1}{\lambda m^{1/2}}\ell^{1/2}I^{1/2}(u),$$

which gives $I(u) \ge c\min\{\mu\theta^2\lambda\ln\frac{1}{m}, \frac{\theta^2\lambda^2 m}{\ell}\ln^2\frac{1}{m}\}$ and concludes the proof in this case.

We now turn to the second case, in which the second term in (3.44) is the largest, and write correspondingly using the fundamental lemma of calculus

$$\frac{1}{16}\theta\ln\frac{1}{m} \le \int_{(0,1)}u_2(x_1^*+s\xi_1,s\xi_2)\psi'(s)\,\mathrm{d}s$$
$$= -\int_{(0,1)}\int_{(s\xi_2,1)}\partial_2 u_2(x_1^*+s\xi_1,x_2)\,\psi'(s)\,\mathrm{d}x_2\,\mathrm{d}s + \int_{(0,1)}u_2(x_1^*+s\xi_1,1)\psi'(s)\,\mathrm{d}s,$$

with the first integral being estimated by $2\|\psi'\|_{L^2((0,1))}I^{1/2}(u)$.

We recall that (3.43) states, after changing variables separately in the two terms,

$$-\int_{(-1,0)}u_2(x_1^*+s\xi_1,1)\psi'_B(s)\,\mathrm{d}s - \int_{(1,2)}u_2(x_1^*+s\xi_1,1)\psi'_B(s-1)\,\mathrm{d}s \le \frac{1}{32}\theta\ln\frac{1}{m},$$

sum and obtain

$$\frac{1}{32}\theta\ln\frac{1}{m} \le 2\|\psi'\|_{L^2((0,1))}I^{1/2}(u) + \int_{\mathbb{R}}u_2(x_1^*+s\xi_1,1)\psi'(s)\,\mathrm{d}s.$$

As above, using the estimates for ψ in (3.41) and Lemma 1.3 with the extension to $F_2 := (-2,\ell+2)\times(1,2)$, and using $\tilde{\Psi}(x_1^*+s\xi_1,x_2) := \Psi(x_2-1,s)$, leads to

$$\frac{1}{32}\theta\ln\frac{1}{m} \le 2\|\psi'\|_{L^2((0,1))}I^{1/2}(u)+2\|\nabla\tilde{\Psi}\|_{L^2(F_2)}\|\nabla u_2\|_{L^2(F_2)} \le c\frac{1}{\lambda m^{1/2}}I^{1/2}(u) + c\lambda^{-1/2}\ln^{1/2}\frac{1}{m}\mu^{-1/2}I^{1/2}(u),$$

which gives $I(u) \ge c\min\{\mu\theta^2\lambda\ln\frac{1}{m}, \theta^2\lambda^2 m\ln^2\frac{1}{m}\}$ and, since $1 \le \ell$, concludes the proof also in this case.

Step 5: Boundary correction for $\theta < \mu$.

In this case we truncate, so that the new function vanishes at $s = 0$ and $s = 1$. We set

$$\psi_T(t) := \left[\ln\frac{1}{m} - \left(\ln\frac{3|t|}{m}\right)_+\right]_+ = \begin{cases} \ln\frac{1}{m}, & \text{if } |t| \leq \frac{1}{3}m, \\ \ln\frac{1}{3|t|}, & \text{if } \frac{1}{3}m < |t| \leq \frac{1}{3}, \\ 0, & \text{if } |t| > \frac{1}{3}. \end{cases}$$

We remark that ψ_T has the same form as the functions ψ_i, with the only difference that the width of the central region is not $g_i \in [\lambda m, m]$ but exactly $\frac{1}{3}m$, so that $\operatorname{supp}\psi_T = [-\frac{1}{3}, \frac{1}{3}]$. This is important to ensure symmetry of the boundary conditions. We then define

$$\psi(t) := \begin{cases} \max\{\hat{\psi}(t), \psi_T(t - \frac{1}{3}), \psi_T(t - \frac{2}{3})\}, & \text{if } t \in (\frac{1}{3}, \frac{2}{3}), \\ \psi_T(t - \frac{1}{3}), & \text{if } t \leq \frac{1}{3}, \\ \psi_T(t - \frac{2}{3}), & \text{if } t \geq \frac{2}{3}, \end{cases}$$

and observe that $\psi = 0$ on $\mathbb{R} \setminus (0, 1)$. Correspondingly,

$$\Psi(x) := \begin{cases} \max\{\hat{\Psi}(x), \psi_T(|(\frac{1}{3}x_1, x_2 - \frac{1}{3})|), \psi_T(|(\frac{1}{3}x_1, x_2 - \frac{2}{3})|)\}, & \text{if } t \in (\frac{1}{3}, \frac{2}{3}), \\ \psi_T(|(\frac{1}{3}x_1, x_2 - \frac{1}{3})|), & \text{if } t \leq \frac{1}{3}, \\ \psi_T(|(\frac{1}{3}x_1, x_2 - \frac{2}{3})|), & \text{if } t \geq \frac{2}{3}. \end{cases}$$

It is apparent that $\Psi(0, t) = \psi(t)$, and that $\Psi = 0$ outside $(-1, 1) \times (0, 1)$. For $x_2 = \frac{1}{3}$ we observe that $\hat{\Psi}((x_1, \frac{1}{3})) \leq \max_i \psi_i(x_1) \leq \min\{\ln\frac{1}{m}, (\ln\frac{\max g_i}{m|x_1|})_+\} \leq \min\{\ln\frac{1}{m}, (\ln\frac{1}{|x_1|})_+\} = \psi_T(\frac{1}{3}|x_1|)$. Therefore Ψ is continuous across the boundaries $x_2 \in \{\frac{1}{3}, \frac{2}{3}\}$. Repeating the same estimates as above we obtain

$$\|\psi\|_{L^1((0,1))} \leq 5, \quad \|\psi\|_{L^2((0,1))}^2 \leq 10, \quad \|\psi'\|_{L^2(\mathbb{R})} \leq \frac{c}{\lambda m^{1/2}}, \quad \|\nabla\Psi\|_{L^2(\mathbb{R}^2)} \leq c\frac{1}{\lambda^{1/2}}\ln^{1/2}\frac{1}{m}. \tag{3.45}$$

At this point we need to check that restricting to the central one-third of $(0, 1)$ we did not loose most of the minority phase. Specifically, we claim that we may assume that

$$\mathcal{L}^1\left(\omega \cap \left(\frac{1}{3}, \frac{2}{3}\right)\right) \geq \frac{1}{8}\theta. \tag{3.46}$$

To prove (3.46), we first show that

$$\min_{\alpha \in \mathbb{R}} \|v - \alpha\|_{L^1((\frac{1}{3}, \frac{2}{3}))} \leq c(\mu^{-1/2} + \ell^{1/2})I^{1/2}(u). \tag{3.47}$$

Let $\alpha_1 := \int_{(-1,0) \times (0,1)} u_1(0, s) \, \mathrm{d}s$. By Poincaré's inequality and the trace theorem in $W^{1,2}$ we have

$$\|u_1(0, \cdot) - \alpha_1\|_{L^1((\frac{1}{3}, \frac{2}{3}))} \leq \|u_1(0, \cdot) - \alpha_1\|_{L^1((0,1))} \leq c\|\nabla u_1\|_{L^2((-1,0) \times (0,1))} \leq c\mu^{-1/2}I^{1/2}(u)$$

and

$$\int_{1/3}^{2/3} |u_1((x_1^*, 0) + \xi s) - \alpha_1| \, \mathrm{d}s \leq \int_{1/3}^{2/3} |u_1(0, s) - \alpha_1| \, \mathrm{d}s + \int_{1/3}^{2/3} \int_{(0, x_1^* + \xi_1)} |\partial_1 u_1|(t, s) \, \mathrm{d}t \, \mathrm{d}s \leq c\mu^{-1/2}I^{1/2}(u)$$
$$+ c\ell^{1/2}I^{1/2}(u).$$

Analogously, with $\alpha_2 := \int_{(x_1^*, x_1^*+1) \times (-1,0)} u_2(0,s) \, \mathrm{d}s$,

$$\int_{1/3}^{2/3} |u_2((x_1^*, 0) + \xi s) - \alpha_2| \, \mathrm{d}s \leq c\mu^{-1/2} I^{1/2}(u) + cI^{1/2}(u).$$

Recalling that $v(s) = (u_1 + 4u_2)((x_1^*, 0) + \xi s)$ concludes the proof of (3.47) since $\ell \geq 1$.

We now prove (3.46). If it does not hold, then

$$\int_{1/3}^{2/3} |v' - \theta| \, \mathrm{d}s \leq \int_{1/3}^{2/3} |v' - \theta + \chi_{1/2}| \, \mathrm{d}s + \mathcal{L}^1 \left(\omega_{1/2} \cap \left(\frac{1}{3}, \frac{2}{3} \right) \right)$$
$$\leq \int_{(0,1)} |v' - \theta + \chi_{1/2}| \, \mathrm{d}s + \mathcal{L}^1 \left(\omega_{1/2} \setminus \omega \right) + \mathcal{L}^1(\omega \cap (\frac{1}{3}, \frac{2}{3}))$$
$$\leq \hat{C}^{1/2}\theta + 2^{-6}\theta + \frac{1}{8}\theta \leq \frac{1}{4}\theta$$

where we used (3.36), (3.34), $\varepsilon\lambda^{-1} \leq \theta^2$ and (3.37). This implies $v(s) - v(s') \geq \theta(s - s') - \frac{1}{4}\theta$ for all $s, s' \in (\frac{1}{3}, \frac{2}{3})$, and therefore

$$\min_{\alpha \in \mathbb{R}} \|v - \alpha\|_{L^1((\frac{1}{3}, \frac{2}{3}))} \geq \frac{1}{2} \int_0^{1/6} |v\left(\frac{1}{2} + s\right) - v\left(\frac{1}{2} - s\right)| \, \mathrm{d}s \geq \frac{1}{2}\theta \int_{1/8}^{1/6} \left(2s - \frac{1}{4}\right) \mathrm{d}s = c\theta.$$

Recalling (3.47), this implies $I(u) \geq c\min\{\mu\theta^2, \ell^{-1}\theta^2\}$ which, since $\lambda^2 m \ln^2 \frac{1}{m} \leq 1$ and $\theta \leq \mu$, concludes the proof. Therefore we can assume that (3.46) holds.

Step 6: Energy estimate for $\theta < \mu$.

The computation is similar to Step 4, but with significant differences in the treatment of the boundary terms. Recalling (3.46), that $\psi = \ln \frac{1}{m}$ and $-v'\psi \geq t^* \ln \frac{1}{m}$ on $\omega \cap (\frac{1}{3}, \frac{2}{3})$, with $t^* \geq \frac{1}{2}$, and $\psi \geq 0$, we have

$$\frac{1}{16}\theta \ln \frac{1}{m} \leq t^*\mathcal{L}^1 \left(\omega \cap \left(\frac{1}{3}, \frac{2}{3} \right) \right) \ln \frac{1}{m} \leq - \int_{(0,1)} v'(s)\psi(s) \, \mathrm{d}s + \int_{\{v' \geq 0\} \cap (0,1)} v'(s)\psi(s) \, \mathrm{d}s.$$

First we observe that

$$\int_{\{v' \geq 0\} \cap (0,1)} v'(s)\psi(s) \, \mathrm{d}s \leq \theta\|\psi\|_{L^1((0,1))} + \|\min\{|v' - \theta|, |v' + (1-\theta)|\}\|_{L^2((0,1))} \|\psi\|_{L^2((0,1))}$$
$$\leq 5\theta + \hat{C}^{1/2}\theta 10^{1/2} \leq \frac{1}{32}\theta \ln \frac{1}{m},$$

where in the first step we used (3.45) and in the second we assumed that m_0 is chosen such that $5 + \hat{C}^{1/2} 10^{1/2} \leq \frac{1}{32} \ln \frac{1}{m}$. Inserting in the previous expression and integrating by parts we get

$$\frac{1}{32}\theta \ln \frac{1}{m} \leq \int_{(0,1)} v\psi' \, \mathrm{d}s.$$

The rest of the proof is very close to the one of Step 4, with some simplifications in the treatment of the exterior field. For the convenience of the reader we repeat the computation here. Recalling the definition of v,

$$\frac{1}{32}\theta\ln\frac{1}{m} \le \int_{(0,1)} u_1((x_1^*,0)+s\xi)\psi'(s)\,\mathrm{d}s + 4\int_{(0,1)} u_2((x_1^*,0)+s\xi)\psi'(s)\,\mathrm{d}s. \tag{3.48}$$

If the first term in (3.48) is larger than the second, by the fundamental theorem and the trace theorem

$$\frac{1}{64}\theta\ln\frac{1}{m} \le \int_{(0,1)} u_1(x_1^*,0)+s\xi)\psi'(s)\,\mathrm{d}s = \int_{(0,1)}\int_0^{x_1^*+s\xi_1} \partial_1 u_1(x_1,s\xi_2)\,\mathrm{d}x_1\,\psi'(s)\,\mathrm{d}s + \int_{(0,1)} u_1(0,s)\psi'(s)\,\mathrm{d}s,$$

with the first integral being estimated by $\|\psi'\|_{L^2((0,1))}\ell^{1/2}I^{1/2}(u)$.

Letting $F_3 := (-1,0)\times(0,1)$ and recalling that $\Psi(0,t)=\psi(t)$ and $\Psi(-1,t)=\Psi(-t,0)=\Psi(-t,1)=0$ for $t\in(0,1)$, we see that the last two integrals can be written as a boundary integral of u_1 times the tangential derivative $\partial_\tau\Psi$, and that Ψ vanishes on the rest of the boundary of F_3. Therefore

$$\frac{1}{64}\theta\ln\frac{1}{m} \le \int_{\partial F_3} u_1\,\partial_\tau\psi\,\mathrm{d}\mathcal{H}^1 + \|\psi'\|_{L^2((0,1))}\ell^{1/2}I^{1/2}(u).$$

With Lemma 1.3 and the estimates for ψ in (3.45) this gives

$$\frac{1}{64}\theta\ln\frac{1}{m} \le \|\nabla\Psi\|_{L^2(F_3)}\|\nabla u_1\|_{L^2(F_3)} + \|\psi'\|_{L^2((0,1))}\ell^{1/2}I^{1/2}(u) \le c\frac{1}{\lambda^{1/2}}\ln^{1/2}\frac{1}{m}\mu^{-1/2}I^{1/2}(u)$$
$$+ c\frac{1}{\lambda m^{1/2}}\ell^{1/2}I^{1/2}(u),$$

which gives $I(u) \ge c\min\{\mu\theta^2\lambda\ln\frac{1}{m}, \frac{\theta^2\lambda^2 m}{\ell}\ln^2\frac{1}{m}\}$ and concludes the proof in this case.

If instead the second term in (3.48) is the larger one, we write

$$\frac{1}{256}\theta\ln\frac{1}{m} \le \int_{(0,1)} u_2(x_1^*+s\xi_1,s\xi_2)\psi'(s)\,\mathrm{d}s$$
$$= -\int_{(0,1)}\int_{(s\xi_2,1)} \partial_2 u_2(x_1^*+s\xi_1,x_2)\,\psi'(s)\,\mathrm{d}x_2\,\mathrm{d}s + \int_{(0,1)} u_2(x_1^*+s\xi_1,1)\psi'(s)\,\mathrm{d}s,$$

with the first integral being estimated by $2\|\psi'\|_{L^2((0,1))}I^{1/2}(u)$. As above, using the estimates for ψ in (3.45) and Lemma 1.3 with the extension to $F_4 := (-2,\ell+2)\times(-1,0)$ leads to

$$\frac{1}{256}\theta\ln\frac{1}{m} \le 2\|\nabla\Psi\|_{L^2(F_4)}\|\nabla u_2\|_{L^2(F_4)} + 2\|\psi'\|_{L^2((0,1))}I^{1/2}(u) \le c\ln^{1/2}\frac{1}{m}\mu^{-1/2}\frac{1}{\lambda^{1/2}}I^{1/2}(u) + c\frac{1}{\lambda m^{1/2}}I^{1/2}(u),$$

which gives $I(u) \ge c\min\{\mu\theta^2\lambda\ln\frac{1}{m}, \theta^2\lambda^2 m\ln^2\frac{1}{m}\}$ and, since $1\le\ell$, concludes the proof also in this case. □

We finally turn to the proof of Proposition 3.3, in which the different ingredients proven in this Section are put together.

Proof of Proposition 3.3. Let $m_0 \in (0,\frac{1}{4}]$ be given as in Lemma 3.7, and define $m_1 \in (0,m_0)$ as the unique solution to $m_1\ln\frac{1}{m_1} = m_0$. If $\theta \ge m_1$ the statement follows directly from Lemma 3.4. Otherwise we use Lemma 3.7 with the following choices of the parameters λ and m:

i) Consider first the case $\mu \geq \varepsilon^{1/3}\theta^{-2/3}\ell^{-1/3}m_0^{2/3}$. Choose $m := m_0$ and $\lambda := \varepsilon^{1/3}\theta^{-2/3}\ell^{2/3}$. Then $\varepsilon\ell^2 \leq \theta^2$ implies on the one hand $\lambda \leq 1$, and on the other hand $\varepsilon \leq \theta^2\ell^{-2} \leq \theta^2\ell$, which gives $\varepsilon \leq \theta^2\lambda$. Thus, λ is admissible. In this case, $\mu\theta^2\lambda \ln\frac{1}{m} \leq c\mu\theta^2$, hence the second and the last two terms in the minimum can be ignored. Therefore, Lemma 3.7 yields that

$$I_{\tilde{\Omega}_\ell}(u) \geq c\min\{\varepsilon^{2/3}\theta^{2/3}\ell^{1/3},\ \mu\varepsilon^{1/3}\theta^{4/3}\ell^{2/3}\} \geq c\varepsilon^{2/3}\theta^{2/3}\ell^{1/3}$$

where we used $\mu\theta^2\lambda \geq \varepsilon^{2/3}\theta^{2/3}\ell^{1/3}m_0^{2/3}$ by the assumption on μ and the definition of λ.

ii) Suppose now that $\mu < \varepsilon^{1/3}\theta^{-2/3}\ell^{-1/3}m_0^{2/3}$. We set $m := \max\{\mu^{3/2}\varepsilon^{-1/2}\theta\ell^{1/2}, \theta\ln\frac{1}{\theta}\}$ and $\lambda := \max\{\varepsilon\theta^{-2}, \frac{1}{2}\mu^{-1/2}\varepsilon^{1/2}\theta^{-1}\ell^{1/2}\ln^{-1/2}\frac{1}{m}\}$. By the assumption on μ and $0 < \theta < m_1$ we obtain $m \in [\theta, m_0]$. Further, $\lambda \geq \varepsilon\theta^{-2}$ by definition. It remains to show $\lambda \leq 1$. Since $\varepsilon\ell^2 \leq \theta^2$ implies $\varepsilon\theta^{-2} \leq 1$, it suffices to prove that

$$\min\{\ln\left(3 + \frac{\varepsilon}{\mu^3\theta^2\ell}\right), \ln\left(3 + \frac{1}{\theta^2}\right)\} \leq 4\ln\frac{1}{m}. \tag{3.49}$$

Indeed, (3.49) and the assumption on $\varepsilon\ell$ give $\varepsilon\ell \leq 4\mu\theta^2 \ln\frac{1}{m}$ which immediately implies $\lambda \leq 1$.

It remains to prove the algebraic inequality (3.49). If $m = \mu^{3/2}\varepsilon^{-1/2}\theta\ell^{1/2}$, then (3.49) follows from the fact that for any $x \in (0, \frac{1}{4})$ we have $\ln(3 + \frac{1}{x^2}) \leq 3\ln\frac{1}{x}$. If instead $m = \theta\ln\frac{1}{\theta}$, we use analogously

$$\ln\left(3 + \frac{1}{x^2}\right) \leq 4\ln\frac{1}{x\ln\frac{1}{x}} \text{ for all } x \in \left(0, \frac{1}{4}\right).$$

The last inequality is equivalent to $3 + \frac{1}{x^2} \leq \frac{1}{x^4\ln^4\frac{1}{x}}$, which is true, since $x\ln^2\frac{1}{x} \leq 4e^{-2} \leq \frac{2}{3}$ for all $x \in (0, 1)$. Therefore (3.49) holds.

We use Lemma 3.7 and estimate the terms separately below. First, using that $\lambda^{-1} = \min\{\varepsilon^{-1}\theta^2, 2\mu^{1/2}\varepsilon^{-1/2}\theta\ell^{-1/2}\ln^{1/2}\frac{1}{m}\}$ and then that $\varepsilon \leq \theta^2\ell^{-2} \leq \theta^2\ell$, we find

$$\frac{\varepsilon\ell}{\lambda} = \min\left\{\theta^2\ell,\ 2\mu^{1/2}\varepsilon^{1/2}\theta\ell^{1/2}\ln^{1/2}\frac{1}{m}\right\} \geq \min\left\{\varepsilon^{2/3}\theta^{2/3}\ell^{1/3},\ 2\mu^{1/2}\varepsilon^{1/2}\theta\ell^{1/2}\ln^{1/2}\frac{1}{m}\right\}.$$

From $m \geq \theta\ln\frac{1}{\theta}$ we get $\mu m^2 \geq \frac{1}{3}\mu\theta^2\ln(3 + \frac{1}{\theta^2})$, and recalling the assumption on $\varepsilon\ell$ we get

$$\mu m^2 \geq \frac{1}{3}\mu\theta^2\ln\left(3 + \frac{1}{\theta^2}\right) \geq \frac{1}{4}(\varepsilon\ell\mu\theta^2)^{1/2}\ln^{1/2}\left(3 + \frac{1}{\theta^2}\right).$$

Next, by definition of λ, we have $\lambda \geq \frac{1}{2}\mu^{-1/2}\varepsilon^{1/2}\theta^{-1}\ell^{1/2}\ln^{-1/2}\frac{1}{m}$ and hence

$$\mu\theta^2\lambda\ln\frac{1}{m} \geq \frac{1}{2}\mu^{1/2}\varepsilon^{1/2}\theta\ell^{1/2}\ln^{1/2}\frac{1}{m}.$$

Finally, using that $\lambda \geq \frac{1}{2}\mu^{-1/2}\varepsilon^{1/2}\theta^{-1}\ell^{1/2}\ln^{-1/2}\frac{1}{m}$, $m \geq \mu^{3/2}\varepsilon^{-1/2}\theta\ell^{1/2}$ and then $\ln^{1/2}\frac{1}{m} \leq \ln\frac{1}{m}$, we find

$$\theta^2\ell^{-1}\lambda^2 m\ln^2\frac{1}{m} \geq \frac{1}{4}\theta^2\ell^{-1}\left(\mu^{-1}\varepsilon\theta^{-2}\ell\ln^{-1}\frac{1}{m}\right)\left(\mu^{3/2}\varepsilon^{-1/2}\theta\ell^{1/2}\right)\ln^2\frac{1}{m} \geq \frac{1}{4}\mu^{1/2}\varepsilon^{1/2}\theta\ell^{1/2}\ln^{1/2}\frac{1}{m}.$$

Putting things together, and recalling (3.49) we obtain

$$I_{\tilde{\Omega}_\ell}(u) \geq c \min \Big\{ \varepsilon^{2/3}\theta^{2/3}\ell^{1/3},$$
$$\mu^{1/2}\varepsilon^{1/2}\theta\ell^{1/2}\ln^{1/2}\left(3 + \frac{1}{\theta^2}\right), \mu^{1/2}\varepsilon^{1/2}\theta\ell^{1/2}\ln^{1/2}\left(3 + \frac{\varepsilon}{\mu^3\theta^2\ell}\right), \mu\theta^2\ln\left(3 + \frac{\theta}{\mu}\right), \mu\theta^2\ln(3 + \ell)\Big\}$$

which concludes the proof also in this case.

$\square$

3.3. A lower bound for small ε, but εL not small

We will now consider the remaining cases. We focus here on the situation in which ε is small, but L is so large that a straight interface along the entire martensitic sample is not optimal. Although this condition does not appear explicitly in the assumptions of Proposition 3.8, the result will only be useful in this situation, as the term εL appears as one of the options in the estimate.

The proof of the lower bound in this case has a structure similar to the one of Section 3.1. We shall use, as above, the subdivision of the set of diagonal slices into various subsets. In particular, we shall show in Lemma 3.10 that the interface between a $\mathcal{P}$ and a $\mathcal{C}$ slice is, energetically speaking, expensive. At variance with Section 3.1, the interpolation between an affine region and a region with periodic boundary values will no longer be penalized with a D^2u term but with a $\partial_1 u_1$ term. This requires estimates on u, and not only on its derivatives.

Proposition 3.8 (A lower bound in the case $\varepsilon \leq \mu\theta^2$ and $\varepsilon \leq \theta^2$)**.** *There exists $c > 0$ such that for all $L \in [1/2, \infty)$, $\theta \in (0, 1/2]$, $\varepsilon > 0$, $\mu > 0$, and $u \in \mathcal{X}$ with*

$$\varepsilon \leq \mu\theta^2 \qquad and \qquad \varepsilon \leq \theta^2$$

we have

$$I(u) \geq c \min \Big\{ \varepsilon L, \mu\theta^2\ln(3 + L), \mu\theta^2\ln\left(3 + \frac{\varepsilon L}{\mu\theta^2}\right) + \varepsilon^{1/2}\theta^{3/2} + \mu\theta^2\ln\left(3 + \frac{\varepsilon}{\mu^2\theta^2}\right), \mu\theta^2\ln\left(3 + \frac{\varepsilon L}{\mu\theta^2}\right)$$
$$+ \mu\theta^2\ln\left(3 + \frac{\theta}{\mu}\right) \Big\}.$$

We first prove some lemmata used in the proof of Proposition 3.8. The first one concerns a local variant of the set $\mathcal{C}$, for which a sharper estimate on the volume is possible.

Lemma 3.9 (Estimates near the boundary)**.** *Assume that $\delta \in (0, \frac{1}{64}\theta]$, $u \in W^{1,2}(\Omega_L, \mathbb{R}^2)$, and let*

$$\mathcal{R} := \Big\{ x_1 \in (0, L - \xi_1) \,:\, \frac{1}{32}\theta \leq \max\{\|u_{x_1}^\xi(s) - u_{x_1}^\xi(0)\|_{L^\infty((0,\delta))}, \|u_{x_1}^\xi(1-s) - u_{x_1}^\xi(1)\|_{L^\infty((0,\delta))}\}.$$

Then

$$I(u) \geq c\theta\mathcal{L}^1(\mathcal{R}).$$

Proof. For almost every $x_1 \in \mathcal{R}$ we have $v := u_{x_1}^\xi \in W^{1,2}((0,1))$. For $s \in (0, \delta)$ we estimate, using (1.11),

$$|v(1-s) - v(1)| \leq s^{1/2}\|v'\|_{L^2((1-s,1))}$$
$$\leq s^{1/2}(\|1\|_{L^2((1-s,1))} + \|\min\{|v' - \theta|, |v' + (1-\theta)|\}\|_{L^2((0,1))})$$

$$\leq \delta + \delta^{1/2} 5 \| \min\{|e(u) - \theta e_1 \odot e_2|, |e(u) + (1-\theta)e_1 \odot e_2|\} \|_{L^2(\Delta_{x_1}^\xi)}$$

and correspondingly for $|v(s) - v(0)|$. Therefore for almost any $x_1 \in \mathcal{R}$ we have

$$\frac{1}{32}\theta \leq \delta + \delta^{1/2} 5 \| \min\{|e(u) - \theta e_1 \odot e_2|, |e(u) + (1-\theta)e_1 \odot e_2|\} \|_{L^2(\Delta_{x_1}^\xi)}.$$

For $\delta \leq \frac{1}{64}\theta$ we deduce

$$c\theta \leq \| \min\{|e(u) - \theta e_1 \odot e_2|, |e(u) + (1-\theta)e_1 \odot e_2|\} \|_{L^2(\Delta_{x_1}^\xi)}^2$$

and integrating over $x_1 \in \mathcal{R}$ we obtain the assertion. $\qquad\square$

Lemma 3.10 (Interpolation estimate). *Let $\mathcal{P}$ be defined as in (3.2), $p := \mathcal{L}^1(\mathcal{P})$. Assume*

$$\varepsilon \leq \mu\theta^2 \qquad and \qquad \varepsilon \leq \theta^2.$$

Then

$$I(u) \geq c \min\{\varepsilon^{1/2}\theta^{3/2}, \mu\theta^2 p, \theta^2 p, \varepsilon L\}.$$

Proof. The proof is based on selecting a good slice in which u is approximately affine, and another one in which the boundary values are close to each other, and estimating the energy in between. We shall work on a thin slice around the boundary, of width $\delta := \frac{1}{64}\theta$.

Step 1. Estimate on good $\mathcal{C}$-slices.
We recall that $\mathcal{C}$ was defined in (3.1) as the set of slices such that the deformation is close to one of the two martensite variants in the interior, and that by Lemma 3.2(iv) it obeys (recall that $\varepsilon \leq \theta^2$)

$$I(u) \geq c\varepsilon \mathcal{L}^1([0, L-\xi_1] \setminus \mathcal{C}). \tag{3.50}$$

We let $\mathcal{G}$ be the set from Lemma 3.6 with $\bar{C} := 2^{-7}$, which obeys

$$I(u) \geq c \min\{\theta^2, \mu\theta^2\} \mathcal{L}^1(\mathcal{G}). \tag{3.51}$$

By Lemma 3.6(iii),

$$\frac{1}{\delta} \int_{(0,\delta)} |u_2(x_1 + (1-s)\xi_1, 1-s) - u_2(x_1 + s\xi_1, s)| \, \mathrm{d}s < \frac{1}{16}\theta \text{ for any } x_1 \in [0, L-\xi_1] \setminus \mathcal{G}.$$

We can assume $\mathcal{L}^1(\mathcal{C} \setminus \mathcal{G}) > 0$. Indeed, if this were not the case, then (up to null sets) $\mathcal{C} \subset \mathcal{G}$ and $([0, L-\xi_1] \setminus \mathcal{C}) \cup \mathcal{G} = [0, L-\xi_1]$, which implies $I(u) \geq cL \min\{\theta^2, \mu\theta^2, \varepsilon\} = c\varepsilon L$ and concludes the proof.

We claim that

$$\frac{1}{2}\theta \leq \frac{1}{\delta} \int_{(0,\delta)} |u_1((x_c, 0) + s\xi) - u_1((x_c, 0) + (1-s)\xi)| \, \mathrm{d}s \text{ for any } x_c \in \mathcal{C} \setminus \mathcal{G}. \tag{3.52}$$

To see this, assume $x_1 \in \mathcal{C}$ and for $\sigma \in \{0,1\}$ let $f_\sigma(s) := u_{x_1}^\xi(0) + s(\theta - \sigma)$. Then $|f_\sigma(1-s) - f_\sigma(s)| = |1 - 2s|\,|\theta - \sigma| \geq |1 - 2s|\theta$, therefore for any $s \in (0, \delta)$ we have

$$\frac{7}{8}\theta \leq (1-2s)\theta \leq \min_{\sigma \in \{0,1\}} |f_\sigma(s) - f_\sigma(1-s)|$$

$$\leq |u_{x_1}^\xi(s) - u_{x_1}^\xi(1-s)| + \min_{\sigma \in \{0,1\}} \left(|u_{x_1}^\xi(s) - f_\sigma(s)| + |u_{x_1}^\xi(1-s) - f_\sigma(1-s)| \right)$$

$$\leq |u_{x_1}^\xi(s) - u_{x_1}^\xi(1-s)| + \frac{2}{16}\theta.$$

Therefore, recalling that $u_{x_1}^\xi(s) = u_1((x_1, 0) + s\xi) + 4u_2((x_1, 0) + s\xi)$,

$$\frac{3}{4}\theta \leq |u_{x_1}^\xi(s) - u_{x_1}^\xi(1-s)|$$

$$\leq |u_1((x_1,0) + s\xi) - u_1((x_1,0) + (1-s)\xi)| + 4|u_2((x_1,0) + s\xi) - u_2((x_1,0) + (1-s)\xi)|$$

for any $x_1 \in \mathcal{C}$. Averaging over $s \in (0, \delta)$, and using that $x_1 \notin \mathcal{G}$,

$$\frac{3}{4}\theta \leq \frac{1}{\delta} \int_{(0,\delta)} \left| u_{x_1}^\xi(s) - u_{x_1}^\xi(1-s) \right| \, ds$$

$$\leq \frac{1}{\delta} \int_{(0,\delta)} |u_1((x_1,0) + s\xi) - u_1((x_1,0) + (1-s)\xi)| \, ds + \frac{1}{4}\theta$$

which proves (3.52).

Step 2. Estimate on good $\mathcal{P}$-slices.
We consider the set $\mathcal{R}$ defined in Lemma 3.9, which obeys

$$I(u) \geq c\theta \mathcal{L}^1(\mathcal{R}). \tag{3.53}$$

We can assume $\mathcal{L}^1(\mathcal{P} \setminus \mathcal{R} \setminus \mathcal{G}) > 0$. Indeed, if this were not the case, then $\mathcal{L}^1(\mathcal{R} \cup \mathcal{G}) \geq p$, $I(u) \geq cp\min\{\theta^2, \mu\theta^2\}$, and the proof is concluded.

We claim that

$$\frac{1}{\delta} \int_{(0,\delta)} |u_1((x_p,0) + s\xi) - u_1((x_p,0) + (1-s)\xi)| \, ds \leq \frac{3}{8}\theta \text{ for any } x_p \in \mathcal{P} \setminus \mathcal{R} \setminus \mathcal{G}. \tag{3.54}$$

Indeed, let $x_1 \in \mathcal{P}$. Then $|u(x_1, 0) - u(x_1 + \xi_1, 1)| \leq 2^{-7}\theta$, therefore $|u_{x_1}^\xi(0) - u_{x_1}^\xi(1)| \leq 2^{-4}\theta$. If $x_1 \notin \mathcal{R}$, then a triangular inequality shows that for any $s \in (0, \delta)$

$$|u_{x_1}^\xi(s) - u_{x_1}^\xi(1-s)| \leq |u_{x_1}^\xi(0) - u_{x_1}^\xi(1)| + |u_{x_1}^\xi(s) - u_{x_1}^\xi(0)| + |u_{x_1}^\xi(1-s) - u_{x_1}^\xi(1)| \leq \frac{1}{16}\theta + \frac{2}{32}\theta = \frac{1}{8}\theta.$$

A similar computation as above, using

$$|u_1((x_1,0) + s\xi) - u_1((x_1,0) + (1-s)\xi)| \leq |u_{x_1}^\xi(s) - u_{x_1}^\xi(1-s)| + 4|u_2((x_1,0) + s\xi) - u_2((x_1,0) + (1-s)\xi)|$$

and $x_1 \notin \mathcal{G}$, leads to

$$\frac{1}{\delta} \int_{(0,\delta)} |u_1((x_1,0) + s\xi) - u_1((x_1,0) + (1-s)\xi)| \, ds \leq \frac{1}{\delta} \int_{(0,\delta)} |u_{x_1}^\xi(s) - u_{x_1}^\xi(1-s)| \, ds + \frac{1}{4}\theta \leq \frac{3}{8}\theta.$$

This concludes the proof of (3.54).

Step 3. Interpolation.
We choose $x_p \in \mathcal{P} \setminus \mathcal{R} \setminus \mathcal{G}$ and $x_c \in \mathcal{C} \setminus \mathcal{G}$ and compute

$$
\begin{aligned}
|u_1((x_c,0) + s\xi) - u_1((x_c,0) + (1-s)\xi)| &\leq |u_1((x_c,0) + s\xi) - u_1((x_p,0) + s\xi)| \\
&\quad + |u_1((x_p,0) + s\xi) - u_1((x_p,0) + (1-s)\xi)| \\
&\quad + |u_1((x_p,0) + (1-s)\xi) - u_1((x_c,0) + (1-s)\xi|.
\end{aligned}
$$

Averaging over $s \in (0,\delta)$ and using (3.52), (3.54) and Hölder gives

$$
\begin{aligned}
\frac{1}{2}\theta &\leq \frac{1}{\delta} \int_{(0,\delta)} |u_1((x_c,0) + s\xi) - u_1((x_c,0) + (1-s)\xi)| \, ds \\
&\leq \frac{1}{\delta} \int_{(0,\delta)} |u_1((x_p,0) + s\xi) - u_1((x_p,0) + (1-s)\xi)| \, ds + \frac{1}{\delta} \int_{(0,\delta)\cup(1-\delta,1)} |u_1((x_p,0) + s\xi) - u_1((x_c,0) + s\xi)| \, ds \\
&\leq \frac{3}{8}\theta + \frac{\sqrt{2}}{\delta^{1/2}} \Big(\int_{(0,\delta)\cup(1-\delta,1)} |u_1((x_p,0) + s\xi) - u_1((x_c,0) + s\xi)|^2 \, ds \Big)^{1/2}.
\end{aligned}
$$

Therefore

$$\frac{1}{128}\theta^2 \delta \leq \int_{(0,1)} |u_1((x_p,0) + s\xi) - u_1((x_c,0) + s\xi)|^2 \, ds.$$

By the fundamental theorem of calculus and Hölder's inequality, for any $s \in (0,1)$ we have

$$|u_1((x_p,0) + s\xi) - u_1((x_c,0) + s\xi)|^2 \leq |x_p - x_c| \int_{(0,L)} |\partial_1 u_1|^2(t,s) \, dt.$$

Integrating over s gives

$$\frac{\theta^3}{c} \leq \int_0^1 |u_1((x_p,0) + s\xi) - u_1((x_c,0) + s\xi)|^2 \, ds \leq |x_c - x_p| \, \|\partial_1 u_1\|_{L^2((x_c,x_p)\times(0,1))}^2 \leq |x_c - x_p| \, I(u).$$

Step 4. Conclusion of the proof.
Let $\beta := \mathcal{L}^1(\mathcal{R} \cup \mathcal{G} \cup ([0, L - \xi_1] \setminus \mathcal{C} \setminus \mathcal{P}))$. On the one hand, (3.50), (3.51), (3.53) give $I(u) \geq c \min\{\varepsilon, \theta^2, \mu\theta^2\}\beta = c\varepsilon\beta$. On the other hand, $\inf\{|x_c - x_p| : \text{(3.52) and (3.54) hold}\} \leq \beta$. Therefore

$$I(u) \geq c \min_{\beta' \in (0,L]} \Big[\frac{\theta^3}{\beta'} + \varepsilon\beta' \Big].$$

The minimum is attained at $\beta' = \varepsilon^{-1/2}\theta^{3/2}$ or at $\beta' = L$, and gives

$$I(u) \geq c \min\{\theta^{3/2}\varepsilon^{1/2}, \varepsilon L\}$$

which concludes the proof. $\qquad\square$

Lemma 3.11 (The boundary logarithm). *There are $c > 0$ and $m_2 \in (0, \frac{1}{4}]$ such that the following holds. If*

$$\frac{1}{2} \leq L, \qquad 0 < \theta \leq m_2, \qquad \varepsilon \leq \mu\theta^2 \qquad \mu \leq \theta, \quad and \qquad \mu^2\theta^2 \leq \varepsilon,$$

then for any $u \in \mathcal{X}$ one has

$$\frac{1}{c}I(u) \geq \min\left\{\varepsilon L, \mu\theta^2 \ln\left(3 + \frac{1}{\theta^2}\right), \mu\theta^2 \ln(3 + L), \mu\theta^2 \ln\left(3 + \frac{\theta}{\mu}\right), \mu\theta^2 \ln\left(3 + \frac{\varepsilon}{\mu^2\theta^2}\right)\right\}.$$

Proof.
Step 1. Energy estimate.
We show that there are $c > 0$, $m_2 \in (0, \frac{1}{4}]$ such that for any $m \in [\theta, m_2]$ there is $q \in [0, L]$ such that

$$\frac{1}{c}I(u) \geq \min\left\{\varepsilon L, \mu\theta^2 \ln\left(3 + \frac{\theta}{\mu}\right), \varepsilon q + \frac{\theta^2 m}{q+1}\ln^2\frac{1}{m}, \mu\theta^2 \ln\frac{1}{m}, \mu\theta^2 \ln(3 + L)\right\}. \tag{3.55}$$

We define, similar to Lemma 3.7, $\psi_C(t) := \max\{\psi_B(t), \psi_B(1-t)\}$, where

$$\psi_B(t) := \left[\ln\frac{1}{m} - \left(\ln\frac{|t|}{m}\right)_+\right]_+ = \begin{cases} \ln\frac{1}{m}, & \text{if } |t| \leq m, \\ \ln\frac{1}{|t|}, & \text{if } m < |t| \leq 1, \\ 0, & \text{if } |t| > 1, \end{cases}$$

and compute $\|\psi_B\|_{L^1((0,1))} \leq 1$, $\|\psi_B'\|_{L^1((0,1))} \leq \ln\frac{1}{m}$, and $\|\psi_B'\|_{L^2((0,1))} \leq m^{-1/2}$, which imply

$$\|\psi_C\|_{L^1((0,1))} \leq 2, \quad \|\psi_C'\|_{L^1((0,1))} \leq 2\ln\frac{1}{m}, \quad \text{and} \quad \|\psi_C'\|_{L^2((0,1))} \leq \frac{2}{m^{1/2}}. \tag{3.56}$$

We first claim that

$$\frac{7}{8}\theta \ln\frac{1}{m} \leq \left|\int_0^1 u_{x_1}^\xi(s)\psi_C'(s)\,ds\right| \quad \text{for any } x_1 \in \mathcal{C}, \tag{3.57}$$

where $\mathcal{C}$ was defined in (3.1). To see this we compute, for any $x_1 \in \mathcal{C}$ and $\sigma \in \mathbb{R}$,

$$\sigma \ln\frac{1}{m} = \int_0^1 \frac{d}{ds}(s\sigma\psi_C(s))\,ds = \sigma\int_0^1 \psi_C\,ds + \int_0^1 (s\sigma)\psi_C'(s)\,ds$$

$$= \sigma\int_0^1 \psi_C\,ds + \int_0^1 (s\sigma + u_{x_1}^\xi(0) - u_{x_1}^\xi(s))\psi_C'(s)\,ds + \int_0^1 (u_{x_1}^\xi(s) - u_{x_1}^\xi(0))\psi_C'(s)\,ds.$$

Since $\psi_C(0) = \psi_C(1)$, the last term disappears, and

$$|\sigma| \ln \frac{1}{m} \leq |\sigma| \, \|\psi_C\|_{L^1((0,1))} + \|s\sigma + u_{x_1}^\xi(0) - u_{x_1}^\xi(s)\|_{L^\infty((0,1))} \|\psi_C'\|_{L^1((0,1))} + \left| \int_0^1 u_{x_1}^\xi(s)\psi_C'(s)\,\mathrm{d}s \right|.$$

At this point we recall (3.56) and that $x_1 \in \mathcal{C}$. Therefore there is a choice of $\sigma \in \{\theta, \theta - 1\}$ such that

$$|\sigma| \ln \frac{1}{m} \leq 2|\sigma| + \frac{1}{16} \theta \ln \frac{1}{m} + \left| \int_0^1 u_{x_1}^\xi(s)\psi_C'(s)\,\mathrm{d}s \right|.$$

If m_2 is sufficiently small, $2 \leq 2^{-4} \ln \frac{1}{m}$. For both choices of σ we have $\theta \leq |\sigma|$. Therefore

$$\frac{7}{8} \theta \ln \frac{1}{m} \leq \left| \int_0^1 u_{x_1}^\xi(s)\psi_C'(s)\,\mathrm{d}s \right|,$$

which concludes the proof of (3.57).

Let $\mathcal{G}$ be as in Lemma 3.6 with $\bar{C} := 2^{-8}$. Since $\psi_C'(s) = -\psi_C'(1-s) = -1/s$ on $(m, 1/2)$ and $\psi_C'(s) = \psi_C'(1-s) = 0$ on $(0, m)$, using Lemma 3.6(ii) for $x_1 \notin \mathcal{G}$ we have

$$\left| \int_0^1 4u_2((x_1, 0) + s\xi)\psi_C'(s)\,\mathrm{d}s \right| \leq \int_m^{1/2} \frac{4}{s} |u_2((x_1, 0) + s\xi) - u_2((x_1, 0) + (1-s)\xi)|\,\mathrm{d}s \leq \frac{1}{8}\theta \ln \frac{1}{m}.$$

Recalling that $u_{x_1}^\xi(s) = (u_1 + 4u_2)((x_1, 0) + s\xi)$, we see that

$$\frac{3}{4}\theta \ln \frac{1}{m} \leq \left| \int_0^1 u_1((x_1, 0) + s\xi)\psi_C'(s)\,\mathrm{d}s \right| \qquad \text{for any } x_1 \in \mathcal{C} \setminus \mathcal{G}. \tag{3.58}$$

By Lemma 3.2(iv) and $\varepsilon \leq \theta^2$ we have $I(u) \geq c\varepsilon \mathcal{L}^1([0, L - \xi_1] \setminus \mathcal{C})$. By Lemma 3.6 and $\varepsilon \leq \min\{\theta^2, \mu\theta^2\}$ we have $I(u) \geq c\varepsilon \mathcal{L}^1(\mathcal{G})$. Then

$$I(u) \geq c\varepsilon q, \quad \text{with} \quad q := \mathcal{L}^1(\mathcal{G} \cup ([0, L - \xi_1] \setminus \mathcal{C})).$$

If $q \geq \frac{1}{2}L$ we are done. Otherwise we pick $x_1 \in [0, q + \frac{1}{4}] \cap \mathcal{C} \setminus \mathcal{G}$. We compute

$$\left| \int_0^1 (u_1((x_1, 0) + s\xi) - u_1(0, s))\psi_C'(s)\,\mathrm{d}s \right| \leq \int_0^1 \int_0^{x_1 + \xi_1} |\partial_1 u_1|(t, s)|\psi_C'|(s)\,\mathrm{d}t\,\mathrm{d}s$$

$$\leq \|\psi_C'\|_{L^2((0,1))} \left(x_1 + \frac{1}{4} \right)^{1/2} I(u)^{1/2} \leq cm^{-1/2}(q+1)^{1/2}I(u)^{1/2}$$

where we used (3.56). If the right-hand side is larger than $\frac{1}{4}\theta \ln \frac{1}{m}$ then $I(u) \geq c(q+1)^{-1}\theta^2 m \ln^2 \frac{1}{m}$ and (3.55) is proven. Otherwise, with (3.58) we obtain

$$\frac{1}{2}\theta \ln \frac{1}{m} \leq \left| \int_0^1 u_1(0, s)\psi_C'(s)\,\mathrm{d}s \right|.$$

We define $\Psi_C : \mathbb{R}^2 \to \mathbb{R}$ by

$$\Psi_C(x) := \max\{\psi_B(|x - (0,1)|), \psi_B(|x|)\}.$$

One easily checks that $\Psi_C(0,t) = \psi_C(t)$ and $\Psi_C(t,0) = \Psi_C(t,1) = \psi_B(t)$ for $t \in (0,1)$, $\Psi_C = 0$ on the rest of the boundary of $F_1 := (-1,1) \times (-1,2) \setminus (0,1)^2$. Further, $\|\nabla\Psi_C\|_{L^2(\mathbb{R}^2)}^2 \leq c \ln \frac{1}{m}$.

By Lemma 3.5 with $c_* = 2^{-4}$ and $\ell = 2L$, either $I(u) \geq c \min\{\mu\theta^2 \ln \frac{1}{m}, \mu\theta^2 \ln(3 + \frac{\theta}{\mu}), \mu\theta^2 \ln(3 + L)\}$ and (3.55) holds, so that we are done, or

$$\left| \int_0^1 \psi_B'(s)(u_1(s,1) - u_1(s,0)) \, ds \right| \leq \frac{1}{16} \theta \ln \frac{1}{m}.$$

We conclude that (recalling Lem. 1.3)

$$\frac{1}{8} \theta \ln \frac{1}{m} \leq \int_{\partial F_1} u_1 \partial_\tau \Psi_C \, d\mathcal{H}^1 \leq c \|\nabla\Psi_C\|_{L^2(F_1)} \|\nabla u_1\|_{L^2(F_1)} \leq c \ln^{1/2} \frac{1}{m} \mu^{-1/2} I(u)^{-1/2}$$

which implies $I(u) \geq c\mu\theta^2 \ln \frac{1}{m}$ and concludes the proof of (3.55).
Step 2. Choice of the parameters.
We first remark that

$$\min_{x \geq 0} \left[ax + \frac{b}{x+1} \right] = \min_{x \geq 0} \left[a(x+1) + \frac{b}{x+1} \right] - a \geq 2a^{1/2}b^{1/2} - a \geq a^{1/2}b^{1/2} \quad \text{whenever } 0 < a \leq b. \quad (3.59)$$

We use (3.55) with $m := \max\{\theta, m_2 \frac{\mu^2\theta^2}{\varepsilon}\} \in [\theta, m_2]$. If the first, the second or the last term are the smallest, the proof is concluded. Assume that the smallest is the third or the fourth one. We remark that $\varepsilon \leq \mu\theta^2$ and $\mu \leq \theta$ imply $\varepsilon \leq \theta^3$, hence $\varepsilon \leq \theta^2 m \ln^2 \frac{1}{m}$. Optimizing the third term in (3.55) in q with (3.59) leads to

$$\frac{1}{c} I(u) \geq \min\left\{ \varepsilon^{1/2} \theta m^{1/2} \ln \frac{1}{m}, \mu\theta^2 \ln \frac{1}{m} \right\}.$$

At this point we distinguish two cases. If $\varepsilon \leq m_2 \mu^2 \theta$ then $m = m_2 \frac{\mu^2\theta^2}{\varepsilon}$, we insert and obtain $I(u) \geq c\mu\theta^2 \ln \frac{\varepsilon}{m_2\mu^2\theta^2} \geq c\mu\theta^2 \ln(3 + \frac{\varepsilon}{\mu^2\theta^2})$, which concludes the proof. If instead $m_2\mu^2\theta < \varepsilon$ then $m = \theta$, and the above estimate gives $I(u) \geq c\mu\theta^2 \ln \frac{1}{\theta} \geq c\mu\theta^2 \ln(3 + \frac{1}{\theta^2})$, which also concludes the proof. $\qquad\square$

Proof of Proposition 3.8. We first recall that by Lemma 3.2(iv) we have, since $\varepsilon \leq \theta^2$,

$$\frac{1}{c} I(u) \geq \mu\theta^2 \ln \frac{L - \xi_1 + 1}{p + 1} + \varepsilon p. \quad (3.60)$$

The interpolation estimate from Lemma 3.10 gives

$$I(u) \geq c \min\{\varepsilon^{1/2}\theta^{3/2}, \mu\theta^2 p, \theta^2 p, \varepsilon L\}. \quad (3.61)$$

We treat the four cases separately.
If the minimum in (3.61) is εL, we are done.
If the minimum in (3.61) is $\mu\theta^2 p$, then, as in (i) in the proof of Proposition 3.1,

$$\frac{1}{c} I(u) \geq \min_{p' \in [0, L-\xi_1]} \left(\mu\theta^2 p' + \mu\theta^2 \ln \frac{L - \xi_1 + 1}{p' + 1} \right) = \mu\theta^2 \ln(L - \xi_1 + 1) \geq c\mu\theta^2 \ln(3 + L),$$

and we are done.

If the minimum in (3.61) is $\theta^2 p$, then we can assume $\mu \geq 1$ and, as in (ii) in the proof of Proposition 3.1,

$$\frac{1}{c}I(u) \geq \min_{p' \in [0, L-\xi_1]} \left(\theta^2 p' + \mu\theta^2 \ln \frac{L - \xi_1 + 1}{p' + 1} \right) \geq c \min \left\{ \mu\theta^2 \ln(3 + L), \mu\theta^2 \ln(3 + \frac{L}{\mu}), \theta^2 L \right\}.$$

Also in this case we are done. Indeed, recalling $\varepsilon \leq \theta^2$, we have $\theta^2 L \geq \varepsilon L$ and $\varepsilon^{1/2}\theta^{3/2} \leq \theta^2 \leq \mu\theta^2$, $\frac{\varepsilon L}{\mu\theta^2} \leq \frac{L}{\mu}$ and $\frac{\varepsilon}{\mu^2\theta^2} \leq \frac{1}{\mu^2} \leq 1$, so that $\mu\theta^2 \ln(3 + \frac{\varepsilon L}{\mu\theta^2}) + \mu\theta^2 \ln(3 + \frac{\varepsilon}{\mu^2\theta^2}) + \varepsilon^{1/2}\theta^{3/2} \leq 4\mu\theta^2 \ln(3 + \frac{L}{\mu}) \leq 4cI(u)$.

We are left with the case that (3.61) states $I(u) \geq c\varepsilon^{1/2}\theta^{3/2}$.

We now show that (3.60) implies $I(u) \geq c \min\{\varepsilon L, \mu\theta^2 \ln(3 + \frac{\varepsilon L}{\mu\theta^2}), \mu\theta^2(3 + L)\}$. Indeed, the minimum of the expression in the right-hand side of (3.60) is attained at $p = 0$, or at $p = L - \xi_1$, or at $p + 1 = \mu\theta^2/\varepsilon$. If it is at $p = 0$ then $I(u) \geq c\mu\theta^2 \ln(L - \xi_1 + 1) \geq c\mu\theta^2 \ln(3 + L)$ and the proof is concluded. If it is at some $p \geq \frac{1}{5}L$ then $I(u) \geq c\varepsilon p \geq c\varepsilon L$ and the proof is concluded. We are left with the case that the first term is at least $\mu\theta^2$ and $p + 1 = \mu\theta^2/\varepsilon$. Then, recalling the previous result from the interpolation estimate, we have

$$\frac{1}{c}I(u) \geq \varepsilon^{1/2}\theta^{3/2} + \mu\theta^2 \ln \left(3 + \frac{\varepsilon L}{\mu\theta^2} \right). \tag{3.62}$$

Let m_2 be as in Lemma 3.11. We next show that we can assume (with a constant c depending on m_2) that

$$\frac{1}{c}I(u) \geq \min \left\{ \mu\theta^2 \ln \left(3 + \frac{1}{\theta^2} \right), \mu\theta^2 \ln \left(3 + \frac{\theta}{\mu} \right), \mu\theta^2 \ln \left(3 + \frac{\varepsilon}{\mu^2\theta^2} \right) \right\}. \tag{3.63}$$

If at least one of $m_2 \leq \theta$, $\theta \leq \mu$, $\varepsilon \leq \mu^2\theta^2$ holds then the minimum in (3.63) is below $c\mu\theta^2$ and (3.63) follows from (3.62). If instead $\theta < m_2$, $\mu < \theta$, $\mu^2\theta^2 < \varepsilon$ then Lemma 3.11 shows that either $\frac{1}{c}I(u) \geq \min\{\varepsilon L, \mu\theta^2 \ln(3 + L)\}$, and we are done, or (3.63) holds.

It remains to show that (3.62) and (3.63) imply the assertion. This is clear if $\varepsilon \leq \mu^2$, since in this case the first term in (3.63) is not relevant and $\varepsilon^{1/2}\theta^{3/2} \geq 0$. If instead $\mu^2 < \varepsilon$, then $\mu^2\theta^2 < \varepsilon$ and therefore $(\frac{\mu^2\theta^2}{\varepsilon})^{1/4} \ln(3 + \frac{\varepsilon}{\mu^2\theta^2}) \leq C$. This implies $\mu\theta^2 \ln(3 + \frac{\varepsilon}{\mu^2\theta^2}) \leq \mu^{1/2}\theta^{3/2}\varepsilon^{1/4}C \leq C\varepsilon^{1/2}\theta^{3/2}$, so that (3.62) concludes the proof. □

Remark 3.12. In the last step of the proof, we just removed the scaling $\varepsilon^{1/2}\theta^{3/2}$ in the regime $\mu\theta^2 \ln(3 + \frac{\varepsilon L}{\mu\theta^2}) + \mu\theta^2 \ln(3 + \frac{\theta}{\mu}) + \varepsilon^{1/2}\theta^{3/2}$. We note that this does not change the scaling behaviour of our lower bound: We distinguish two possibilities. If $\varepsilon \leq \mu^2\theta$, then $\varepsilon^{1/2}\theta^{3/2} \leq \mu\theta^2$, and we have

$$\mu\theta^2 \ln \left(3 + \frac{\varepsilon L}{\mu\theta^2} \right) + \mu\theta^2 \ln \left(3 + \frac{\theta}{\mu} \right) \leq \mu\theta^2 \ln \left(3 + \frac{\varepsilon L}{\mu\theta^2} \right) + \mu\theta^2 \ln \left(3 + \frac{\theta}{\mu} \right) + \varepsilon^{1/2}\theta^{3/2}$$

$$\leq 2 \left(\mu\theta^2 \ln(3 + \frac{\varepsilon L}{\mu\theta^2}) + \mu\theta^2 \ln \left(3 + \frac{\theta}{\mu} \right) \right).$$

Otherwise, if $\varepsilon > \mu^2\theta$, we have as in the proof of the upper bound (Thm. 2.7 (d))

$$\mu\theta^2 \ln \left(3 + \frac{\varepsilon L}{\mu\theta^2} \right) + \mu\theta^2 \ln \left(3 + \frac{\theta}{\mu} \right) + \varepsilon^{1/2}\theta^{3/2} \geq \mu\theta^2 \ln \left(3 + \frac{\varepsilon L}{\mu\theta^2} \right) + \mu\theta^2 \ln \left(3 + \frac{\theta}{\mu} \right) \geq c\mu\theta^2 \ln(3 + L).$$

3.4. Conclusion of the lower bound

We finally bring together the bounds proven in the previous Sections to obtain the desired lower bound.

Theorem 3.13 (Lower bound). *For any $\varepsilon > 0$, $\mu > 0$, $L \geq \frac{1}{2}$, $\theta \in (0, \frac{1}{2}]$ and any $u \in \mathcal{X}$ we have*

$$I(u) \geq c\mathcal{I}(\mu, \varepsilon, \theta, L),$$

where $\mathcal{I}$ was defined in Theorem 1.1.

Proof. We distinguish several cases.

i. Assume that at least one of $\theta^2 \leq \varepsilon$ and $\mu\theta^2 \leq \varepsilon$ holds. Then Proposition 3.1 gives

$$I(u) \geq c\min\left\{\mu\theta^2 \ln(3 + L), \theta^2 L, \mu\theta^2 \ln\left(3 + \frac{L}{\mu}\right) + \varepsilon\theta\right\},$$

   and the proof is concluded.

ii. From now on we have $\varepsilon \leq \theta^2$ and $\varepsilon \leq \mu\theta^2$.
   We start by applying Proposition 3.8, see also the remark afterwards. This gives that

$$I(u) \geq c\min\left\{\varepsilon L, \mu\theta^2 \ln(3 + L), \mu\theta^2 \ln\left(3 + \frac{\varepsilon L}{\mu\theta^2}\right) + \varepsilon^{1/2}\theta^{3/2} + \mu\theta^2 \ln\left(3 + \frac{\varepsilon}{\mu^2\theta^2}\right), \mu\theta^2 \ln\left(3 + \frac{\varepsilon L}{\mu\theta^2}\right)\right.$$
$$\left. +\mu\theta^2 \ln\left(3 + \frac{\theta}{\mu}\right)\right\}.$$

   Note that the proof is concluded unless the first term is the smallest. Assume now that

$$I(u) \geq c\varepsilon L. \tag{3.64}$$

   If $\theta^2 \leq \varepsilon(2L)^2$, then

$$\frac{1}{c}I(u) \geq 2\varepsilon L = \varepsilon L + \varepsilon^{2/3}L^{1/3}(\varepsilon L^2)^{1/3} \geq \varepsilon L + \frac{1}{2}\varepsilon^{2/3}\theta^{2/3}L^{1/3}$$

   concludes the proof.
   In the following we can assume that $\varepsilon(2L)^2 < \theta^2$ and $\varepsilon \leq \mu\theta^2$ and (3.64) hold. We distinguish more subcases, depending on the competition between the interfacial energy and the austenite elasticity. Specifically, the critical condition is whether

$$2\varepsilon L \leq \mu\theta^2 \min\left\{\ln\left(3 + \frac{\varepsilon}{2\mu^3\theta^2 L}\right), \ln\left(3 + \frac{1}{\theta^2}\right)\right\} \tag{3.65}$$

   holds or not.
   Assume that (3.65) does not hold. Then

$$2\varepsilon L \geq \min\left\{\mu^{1/2}\varepsilon^{1/2}\theta L^{1/2}\left(\ln\left(3 + \frac{1}{\theta^2}\right)\right)^{1/2} + \varepsilon L, \mu^{1/2}\varepsilon^{1/2}\theta L^{1/2}\left(\ln\left(3 + \frac{\varepsilon}{\mu^3\theta^2 L}\right)\right)^{1/2} + \varepsilon L\right\}$$

   and with (3.64) the proof is concluded.

   Finally, assume that (3.65) holds. In this situation we can use Proposition 3.3 with $\ell = 2L$. Recalling (3.64),

this gives

$$\frac{1}{c}I(u) \geq \varepsilon L + \min\left\{\mu\theta^2\ln(3+L), \mu\theta^2\ln\left(3+\frac{\theta}{\mu}\right), \varepsilon^{2/3}\theta^{2/3}L^{1/3},\right.$$
$$\left.\mu^{1/2}\varepsilon^{1/2}\theta L^{1/2}\left(\ln\left(3+\frac{1}{\theta^2}\right)\right)^{1/2}, \mu^{1/2}\varepsilon^{1/2}\theta L^{1/2}\left(\ln\left(3+\frac{\varepsilon}{\mu^3\theta^2 L}\right)\right)^{1/2}\right\}.$$

We remark that the term $\mu\theta^2\ln(3+\frac{\theta}{\mu})$ can be dropped. Indeed, if $\theta \geq \mu L$ then it is larger than $\ln(3+L)$. If instead $\theta < \mu L$, then it is equivalent to the term $\mu\theta^2\ln(3+\frac{\varepsilon L}{\mu\theta^2}) + \mu\theta^2\ln(3+\frac{\theta}{\mu}) + \varepsilon^{1/2}\theta^{3/2}$, that we already included before. Indeed, in this case $\varepsilon^{1/2}\theta^{3/2} \leq \varepsilon^{1/2}L^{1/2}\mu^{1/2}\theta \leq \varepsilon L + \mu\theta^2$ and, using $\ln(3+x) \leq 2+x$ in the first term,

$$\mu\theta^2\ln(3+\frac{\varepsilon L}{\mu\theta^2}) + \mu\theta^2\ln\left(3+\frac{\theta}{\mu}\right) + \varepsilon^{1/2}\theta^{3/2} \leq 2\mu\theta^2 + \varepsilon L + \mu\theta^2\ln\left(3+\frac{\theta}{\mu}\right) + \varepsilon L + \mu\theta^2$$
$$\leq 4\left(\mu\theta^2\ln\left(3+\frac{\theta}{\mu}\right) + \varepsilon L\right).$$

This concludes the proof.

$\square$

REFERENCES

[1] L. Ambrosio, A. Coscia and G. Dal Maso, Fine properties of functions with bounded deformation. *Arch. Rat. Mech. Anal.* **139** (1997) 201–238.

[2] L. Ambrosio, N. Fusco and D. Pallara, Functions of bounded variation and free discontinuity problems. Oxford University Press, Oxford, (2000).

[3] J.M. Ball, Mathematical models of martensitic microstructure. *Mater. Sci. Eng. A* **378** (2004) 61–69.

[4] J.M. Ball and R.D. James, Fine phase mixtures as minimizers of energy. *Arch. Rat. Mech. Anal.* **100** (1987) 13–52.

[5] J.M. Ball and R.D. James, Proposed experimental tests of a theory of fine microstructure, and the two-well problem. *Phil. Trans. R. Soc. London A* **338** (1992) 389–450.

[6] P. Bella and M. Goldman, Nucleation barriers at corners for cubic-to-tetragonal phase transformation. *Proc. Roy. Soc. Edinburgh A* **145** (2015) 715–724.

[7] P. Bella and R.V. Kohn, Wrinkles as the Result of Compressive Stresses in an Annular Thin Film. *Comm. Pure Appl. Math.* **67** (2014) 693–747.

[8] H. Ben Belgacem, S. Conti, A. De Simone and S. Müller, Energy scaling of compressed elastic films. *Arch. Rat. Mech. Anal.* **164** (2002) 1–37.

[9] K. Bhattacharya, Self-accomodation in martensite. *Arch. Rat. Mech. Anal.* **120** (1992) 201–244.

[10] A. Capella and F. Otto, A rigidity result for a perturbation of the geometrically linear three-well problem. *Comm. Pure Appl. Math.* **62** (2009) 1632–1669.

[11] A. Capella and F. Otto, A quantitative rigidity result for the cubic-to-tetragonal phase transition in the geometrically linear theory with interfacial energy. *Proc. Roy. Soc. Edinburgh Sect. A* **142** (2012) 273–327.

[12] P. Cesana, F. Della Porta, A. Rüland, C. Zillinger and B. Zwicknagl, Exact constructions in the (non-linear) planar theory of elasticity: from elastic crystals to nematic elastomers. *Arch. Rat. Mech. Anal.* **237** (2020) 383–445.

[13] A. Chan, *Energieskalierung, Gebietsverzweigung und SO(2)-Invarianz in einem fest-fest Phasenübergangsproblem*. Ph.D. thesis, Bonn University (2013). Available from: http://hss.ulb.uni-bonn.de/2013/3388/3388.htm.

[14] A. Chan and S. Conti, Energy Scaling and Domain Branching in Solid-Solid Phase Transitions, in Singular Phenomena and Scaling in Mathematical Models, edited by M. Griebel. Springer International Publishing, Cham (2014) 243–260.

[15] A. Chan and S. Conti, Energy scaling and branched microstructures in a model for shape-memory alloys with SO(2) invariance. *Math. Models Methods App. Sci.* **25** (2015) 1091–1124.

[16] R. Choksi, Scaling laws in microphase separation of diblock copolymers. *J. Nonlinear Sci.* **11** (2001) 223–236.

[17] R. Choksi and R.V. Kohn, Bounds on the micromagnetic energy of a uniaxial ferromagnet. *Comm. Pure Appl. Math.* **51** (1998) 259–289.

[18] R. Choksi, R.V. Kohn and F. Otto, Energy minimization and flux domain structure in the intermediate state of a type-I superconductor. *J. Nonlinear Sci.* **14** (2004) 119-171.

[19] R. Choksi, R.V. Kohn and F. Otto, Domain branching in uniaxial ferromagnets: a scaling law for the minimum energy. *Comm. Math. Phys.* **201** (1998) 61–79.

[20] S. Conti, Branched microstructures: scaling and asymptotic self-similarity. *Comm. Pure Appl. Math.* **53** (2000) 1448–1474.

[21] S. Conti, A lower bound for a variational model for pattern formation in shape-memory alloys. *Cont. Mech. Thermodyn.* **17** (2006) 469–476.

[22] S. Conti, J. Diermeier and B. Zwicknagl, Deformation concentration for martensitic microstructures in the limit of low volume fraction. *Calc. Var. Partial Differ. Equ.* **56** (2017) 16.

[23] S. Conti, M. Goldman, F. Otto and S. Serfaty, A branched transport limit of the Ginzburg-Landau functional. *J. l'École Polytech. Math.* **5** (2018) 317-375.

[24] S. Conti, M. Klar and B. Zwicknagl, Piecewise affine stress-free martensitic inclusions in planar nonlinear elasticity. *Proc. R. Soc. A* **473** (2017) 20170235.

[25] S. Conti, F. Otto and S. Serfaty, Branched microstructures in the Ginzburg-Landau model of type-I superconductors. *SIAM J. Math. Anal.* **48** (2016) 2994–3034.

[26] S. Conti and M. Ortiz, Optimal scaling in solids undergoing ductile fracture by crazing. *Arch. Ration. Mech. Anal.* **219** (2016) 607-636.

[27] S. Conti and B. Zwicknagl, Low volume-fraction microstructures in martensites and crystal plasticity. *Math. Models Methods App. Sci.* **26** (2016) 1319–1355.

[28] J. Cui, Y.S. Chu, O.O. Famodu, Y. Furuya, J. Hattrick-Simpers, R.D. James, A. Ludwig, S. Thienhaus, M. Wuttig, Z. Zhang and I. Takeuchi, Combinatorial search of thermoelastic shape-memory alloys with extremely small hysteresis width. *Nat. Mater.* **5** (2006) 286–290.

[29] B. Dacorogna, Direct methods in the calculus of variations, Vol. 78. Springer, Berlin (2007).

[30] J. Diermeier, Nichtkonvexe Variationsprobleme und Mikrostrukturen. Bachelor's thesis, Universität Bonn (2010).

[31] J. Diermeier, *Domain branching in linear elasticity.* Master's thesis, Universität Bonn (2013).

[32] J. Diermeier, *Analysis of martensitic microstructures in shape-memory-alloys: A low volume-fraction limit.* Ph.D. thesis, Bonn University (2016). http://hss.ulb.uni-bonn.de/2016/4499/4499.htm.

[33] G. Dolzmann and S. Müller, Microstructures with finite surface energy: the two-well problem. *Arch. Ration. Mech. Anal.* **132** (1995) 101–141.

[34] R.D. James, Materials from mathematics. *Bull. Am. Math. Soc.* **56** (2019) 1–28.

[35] R.D. James and Z. Zhang, A way to search for multiferroic materials with ¨unlikely¨ combinations of physical properties, in The Interplay of Magnetism and Structure in Functional Materials, edited by L. Manosa, A. Planes and A.B. Saxena, Vol. 79. Springer, Berlin (2005).

[36] W. Jin and P. Sternberg, Energy estimates of the von Kármán model of thin-film blistering. *J. Math. Phys.* **42** (2001) 192–199.

[37] B. Kirchheim, Rigidity and Geometry of Microstructures. *MPI-MIS lecture notes* (2003).

[38] H. Knüpfer and R.V. Kohn, Minimal energy for elastic inclusions. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **467** (2011) 695–717.

[39] H. Knüpfer, R.V. Kohn and F. Otto, Nucleation Barriers for the Cubic-to-Tetragonal Phase Transformation. *Comm. Pure Appl. Math.* **66** (2013) 867–904.

[40] H. Knüpfer and C. Muratov, Domain Structure of Bulk Ferromagnetic Crystals in Applied Fields Near Saturation. *J. Nonlinear Sci.* **21** (2011) 1–42.

[41] H. Knüpfer and F. Otto, Nucleation barriers for the cubic-to-tetragonal phase transformation in the absence of self-accommodation. *Z. Angew. Math. Mech.* **99** (2019) e201800179.

[42] R.V. Kohn, Energy-driven pattern formation, in *International Congress of Mathematicians, ICM 2006*, Vol. 1 (2006) 359–383.

[43] R.V. Kohn and S. Müller, Branching of twins near an austenite-twinned martensite interface. *Phil. Mag. A* **66** (1992) 697–715.

[44] R.V. Kohn and S. Müller, Surface energy and microstructure in coherent phase transitions. *Comm. Pure Appl. Math.* **XLVII** (1994) 405–435.

[45] R.V. Kohn and B. Wirth, Optimal fine-scale structures in compliance minimization for a shear load. *Comm. Pure Appl. Math.* **69** (2016) 1572–1610.

[46] V.A. Kondrat'ev and O.A. Oleinik, Boundary-value problems for the system of elasticity theory in unbounded domains. Korn's inequalities. *Russ. Math. Surv.* **43** (1988) 65–119.

[47] D. Melching, *Microstructures in shape memory alloys.* Master's thesis, Universität Bonn (2015).

[48] S. Müller, Variational models for microstructure and phase transitions, in Calculus of variations and geometric evolution problems, edited by F. Bethuel *et al. Springer Lecture Notes in Math. 1713.* Springer, Berlin (1999) 85–210.

[49] A. Rüland, A Rigidity Result for a Reduced Model of a Cubic-to-Orthorhombic Phase Transition in the Geometrically Linear Theory of Elasticity. *J. Elast.* **123** (2016) 137–177.

[50] A. Rüland, The Cubic-to-Orthorhombic Phase Transition: Rigidity and Non-Rigidity Properties in the Linear Theory of Elasticity. *Arch. Ration. Mech. Anal.* **221** (2016) 23–106.

[51] T. Simon, Rigidity of branching microstructures in shape memory alloys. Preprint arXiv:1705.03664 (2017).

[52] Z. Zhang, R.D. James and S. Müller, Energy barriers and hysteresis in martensitic phase transformations. *Acta Mater.* **57** (2009) 4332–4352.

[53] B. Zwicknagl, Microstructures in Low-Hysteresis Shape Memory Alloys: Scaling Regimes and Optimal Needle Shapes. *Arch. Ration. Mech. Anal.* **213** (2014) 355–421.