

FINITE ELEMENT ANALYSIS OF A CLASS OF STRESS-FREE MARTENSITIC MICROSTRUCTURES

BO LI

ABSTRACT. This work is concerned with the finite element approximation of a class of stress-free martensitic microstructures modeled by multi-well energy minimization. Finite element energy-minimizing sequences are first constructed to obtain bounds on the minimum energy over all admissible finite element deformations. A series of error estimates are then derived for finite element energy minimizers.

1. INTRODUCTION

A martensitic microstructure is a fine-scale mixture of coherent phases or phase variants of a martensitic crystal. Such a microstructure can often be modeled by multi-well energy minimization. The total free energy does not in general attain its infimum. Energy-minimizing sequences can, however, develop fine-scale oscillations and define stress-free microstructures by the notion of Young measures, cf. [1, 2] and the references therein.

There are several approaches to the numerical analysis of nonconvex variational problems modeling martensitic microstructures. One of them is the direct finite element approximation, in which sequences of finite element energy minimizers indexed by the finite element mesh size are studied. Such an approach has been used in the numerical analysis of a simply laminated microstructure that is uniquely determined by the multi-well energy minimization with a boundary condition that is consistent with the underlying microstructure, see [12] for a survey and [5, 9, 11] for details.

In this work, we consider the direct finite element approximation for a more general and physically important situation in which the underlying microstructure can be nonunique but its macroscopic deformation is unique. Moreover, such a microstructure is essentially a simple or high-order laminate. These properties of microstructure are determined by our assumptions on the Dirichlet boundary data, cf. $F1$ – $F3$ in Section 2.

We shall first construct admissible finite element deformations for a laminate of arbitrary order $q \geq 1$, leading to a bound $O(h^{1/(q+1)})$ on the minimum energy over all admissible finite element deformations, where h is the finite element mesh size,

Received by the editor July 28, 2000 and, in revised form, March 15, 2002.

2000 *Mathematics Subject Classification*. Primary 65N30, 74N15.

Key words and phrases. Martensitic microstructure, energy minimization, finite element deformations, error estimates.

This work was partially supported by the NSF through grant DMS-0072958 and by the Graduate School of the University of Maryland through a GRB Summer Research Award.

cf. Theorem 3.1 and Corollary 3.1. We shall then derive a series of error estimates for finite element energy minimizers on the possible reduction of martensitic variants, the closeness of the deformation gradient to a fixed subset of the energy wells, the strong convergence of deformations, and the weak convergence of deformation gradients, cf. Corollary 4.1.

2. THE MULTI-WELL ENERGY MINIMIZATION PROBLEM AND ITS FINITE ELEMENT SOLUTIONS

Let $\Omega \subset \mathbb{R}^3$ be the reference configuration of a martensitic crystal in discussion. We assume that Ω is a bounded domain with a Lipschitz continuous boundary $\partial\Omega$. For a deformation $y : \Omega \rightarrow \mathbb{R}^3$, we denote by $\nabla y : \Omega \rightarrow \mathbb{R}^{3 \times 3}$ its gradient, where $\mathbb{R}^{3 \times 3}$ denotes the set of all 3×3 real matrices. We also denote by $\phi : \mathbb{R}^{3 \times 3} \rightarrow \mathbb{R}$ the free energy density per unit volume of the reference configuration of the crystal. We consider the variational problem of infimizing the total free energy functional

$$(2.1) \quad \mathcal{E}(y) := \int_{\Omega} \phi(\nabla y(x)) \, dx$$

over a set of admissible deformations \mathcal{A} .

We assume that the free energy density $\phi : \mathbb{R}^{3 \times 3} \rightarrow \mathbb{R}$ is continuous and satisfies the following properties.

$\phi 1$. *Absolute minimizers:*

$$(2.2) \quad \begin{aligned} \phi(F) &\geq 0 \quad \forall F \in \mathbb{R}^{3 \times 3}, \\ \phi(F) &= 0 \quad \text{if and only if} \quad F \in \mathcal{U} := \mathcal{U}_1 \cup \cdots \cup \mathcal{U}_N, \end{aligned}$$

where

$$\mathcal{U}_i := \text{SO}(3)U_i := \{RU_i : R \in \text{SO}(3)\}, \quad 1 \leq i \leq N,$$

$\text{SO}(3)$ is the set of all real 3×3 rotation matrices, and $U_1, \dots, U_N \in \mathbb{R}^{3 \times 3}$ are distinct symmetric positive definite matrices.

$\phi 2$. *Growth condition:*

$$(2.3) \quad \phi(F) \geq \kappa [\text{dist}(F, \mathcal{U})]^2 \quad \forall F \in \mathbb{R}^{3 \times 3},$$

where $\kappa > 0$ is a constant and

$$\text{dist}(F, \mathcal{U}) := \inf_{G \in \mathcal{U}} \|F - G\|,$$

where $\|F\| := \sqrt{\sum_{i,j=1}^3 F_{ij}^2}$ for $F = (F_{ij}) \in \mathbb{R}^{3 \times 3}$.

We define the set of admissible deformations to be

$$(2.4) \quad \mathcal{A} = \{y \in W^{1,\infty}(\Omega; \mathbb{R}^3) : y(x) = y_0(x), x \in \partial\Omega\},$$

where $y_0 : \Omega \rightarrow \mathbb{R}^3$ is a homogeneous deformation defined for a given $F_0 \in \mathbb{R}^{3 \times 3}$ by

$$(2.5) \quad y_0(x) = F_0 x \quad \forall x \in \Omega.$$

We assume that the boundary data $F_0 \in \mathbb{R}^{3 \times 3}$ satisfies the following conditions.

$F 1$. *Uniqueness of macroscopic deformation:* There exist a permutation $(i_1 \cdots i_N)$ of $(1 \cdots N)$, an integer s with $1 \leq s \leq N$, and a unit vector $e_0 \in \mathbb{R}^3$ such that

$$(2.6) \quad |F_0 e_0| = |U_{i_1} e_0| = \cdots = |U_{i_s} e_0|.$$

F2. Variant reduction: If $s < N$, then for each $j \in \{s + 1, \dots, N\}$, either there exists a unit vector $a_j \in \mathbb{R}^3$ such that

$$(2.7) \quad \begin{aligned} |F_0 a_j| &= |U_{i_1} a_j| = \dots = |U_{i_s} a_j| \geq \max_{s+1 \leq k \leq N} |U_{i_k} a_j|, \\ |F_0 a_j| &\neq |U_{i_j} a_j|, \end{aligned}$$

or there exists a unit vector $b_j \in \mathbb{R}^3$ such that

$$(2.8) \quad \begin{aligned} |(\text{Cof } F_0) b_j| &= |(\text{Cof } U_{i_1}) b_j| = \dots = |(\text{Cof } U_{i_s}) b_j| \geq \max_{s+1 \leq k \leq N} |(\text{Cof } U_{i_k}) b_j|, \\ |(\text{Cof } F_0) b_j| &\neq |(\text{Cof } U_{i_j}) b_j|, \end{aligned}$$

where $\text{Cof } F \in \mathbb{R}^{3 \times 3}$ is the cofactor matrix of $F \in \mathbb{R}^{3 \times 3}$.

F3. Laminates of arbitrary order: $F_0 \in \mathcal{S}^{lc} := \bigcup_{i=0}^{\infty} \mathcal{S}^{(i)}$, the lamination convex hull of the set $\mathcal{S} := \mathcal{U}_{i_1} \cup \dots \cup \mathcal{U}_{i_s}$, where $\mathcal{S}^{(0)} := \mathcal{S}$, and for each integer $i \geq 1$

$$\mathcal{S}^{(i)} := \left\{ \lambda A + (1 - \lambda) B : A, B \in \mathcal{S}^{(i-1)}, \text{rank}(A - B) \leq 1, 0 \leq \lambda \leq 1 \right\}.$$

We shall denote by q the smallest nonnegative integer such that $F_0 \in \mathcal{S}^{(q)}$.

Our idea of identifying unified conditions on the boundary data stems from [6]. See similar conditions in [3], [4], [7]. Independently, we formulate such conditions, slightly more general, based on our work [5] on the simply laminated microstructure modeled by a six-well problem.

Examples of martensitic transformations and boundary data that satisfy our assumptions $\phi 1$, $\phi 2$, and $F1$ – $F3$ can be found in [2, 3, 4, 5].

We now define finite element spaces and admissible finite element deformations. For simplicity of exposition, we assume that the reference configuration of the crystal $\Omega \subset \mathbb{R}^3$ is a polygonal domain. (For a treatment of a more general Lipschitz domain, we refer to [10].) Let $\{\tau_h : 0 < h \leq h_0\}$ be a family of finite element meshes of Ω , where h_0 is a constant such that $0 < h_0 < 1$. We assume for each $h \in (0, h_0]$ that τ_h is composed of polyhedra with the maximum diameter h , and that $\bar{\Omega} = \bigcup_{K \in \tau_h} K$.

For each $h \in (0, h_0]$, let V_h be a conforming finite element space defined by

$$V_h = \{v_h \in W^{1,\infty}(\Omega) : v_h|_K \in P(K), \forall K \in \tau_h\},$$

where $P(K)$ is the restriction to K of a linear space of polynomials P fixed for all $K \in \tau_h$ and all $h \in (0, h_0]$. We assume that

H1. $P_1 \subseteq P$, where P_1 is the space of all polynomials of degree ≤ 1 .

We also assume that there exists for each $h \in (0, h_0]$ an interpolation operator $I_h : W^{1,\infty}(\Omega) \rightarrow V_h$ with the following properties:

H2. If $v \in W^{1,\infty}(\Omega)$ and $K \in \tau_h$ satisfy that $v|_K \in P_1(K)$, then $(I_h v)|_K = v|_K$;

H3. There exists a constant $\sigma > 0$ such that

$$(2.9) \quad \|\nabla I_h v\|_{L^\infty(\Omega)} \leq \sigma \|\nabla v\|_{L^\infty(\Omega)} \quad \forall v \in W^{1,\infty}(\Omega) \quad \forall h \in (0, h_0].$$

We define for each $h \in (0, h_0]$ the set of admissible finite element deformations $\mathcal{A}_h := \mathcal{A} \cap \mathcal{V}_h$, where \mathcal{A} is the set of admissible deformations defined in (2.4) and $\mathcal{V}_h = V_h \times V_h \times V_h$. Notice by the assumption *H1* that $y_0 \in \mathcal{A}_h$, where y_0 is the homogeneous deformation defined in (2.5). Define for each $h \in (0, h_0]$ the interpolation operator $\mathcal{I}_h : W^{1,\infty}(\Omega; \mathbb{R}^3) \rightarrow \mathcal{A}_h$ by

$$\mathcal{I}_h y = (I_h y_1, I_h y_2, I_h y_3) \quad \forall y = (y_1, y_2, y_3) \in W^{1,\infty}(\Omega; \mathbb{R}^3).$$

The operator \mathcal{I}_h has properties similar to those of I_h , cf. *H1*–*H3*.

Since \mathcal{A}_h is finite dimensional for each $h \in (0, h_0]$, it follows from a usual argument of compactness and the growth condition (2.3) that there exists a $y_h \in \mathcal{A}_h$ such that

$$(2.10) \quad \mathcal{E}(y_h) = \min_{z_h \in \mathcal{A}_h} \mathcal{E}(z_h).$$

3. CONSTRUCTION OF ENERGY-MINIMIZING SEQUENCES OF ADMISSIBLE FINITE ELEMENT DEFORMATIONS

Recall from the condition F3 that the boundary data $F_0 \in \mathcal{S}^{(q)}$, where \mathcal{S} is the subset of energy wells defined in F3.

Theorem 3.1. *For each $h \in (0, h_0]$, there exists a $y_h \in \mathcal{A}_h$ such that*

$$(3.1) \quad \sup_{0 < h \leq h_0} \|\nabla y_h\|_{L^\infty(\Omega; \mathbb{R}^{3 \times 3})} \leq C$$

and

$$(3.2) \quad \text{meas} \{x \in \Omega : \nabla y_h(x) \notin \mathcal{L}\} \leq Ch^{1/(q+1)}$$

for a fixed finite subset $\mathcal{L} \subset \mathcal{S}$.

The following result is a direct consequence of (2.10), Theorem 3.1, and the assumptions $\phi 1$ and $\phi 2$. It provides a bound on the minimum energy over all the admissible finite element deformations.

Corollary 3.1. *For each $h \in (0, h_0]$,*

$$\min_{z_h \in \mathcal{A}_h} \mathcal{E}(z_h) \leq Ch^{1/(q+1)}.$$

Proof of Theorem 3.1. The case that $q = 0$ is trivial. So, assume $q \geq 1$. Since $F_0 \in \mathcal{S}^{(q)}$, there exist matrices $F_{ij} \in \mathcal{S}$ ($i = 0, \dots, q, j = 1, \dots, 2^i$) with $F_{01} := F_0$ such that

$$(3.3) \quad F_{i-1,j} = \lambda_{ij} F_{i,2j-1} + (1 - \lambda_{ij}) F_{i,2j}$$

for some $\lambda_{ij} \in [0, 1]$ and

$$(3.4) \quad F_{i,2j-1} - F_{i,2j} = a_{ij} \otimes n_{ij}$$

for some $a_{ij}, n_{ij} \in \mathbb{R}^3$ with $|n_{ij}| = 1$ ($i = 1, \dots, q, j = 1, \dots, 2^{i-1}$). See Figure 3.1 for a $(q + 1)$ -level binary tree of these matrices.

We construct the desired $y_h \in \mathcal{A}_h$ ($0 < h \leq h_0$) in five steps and refer to Figure 3.2 for the geometry:

- (1) For each $i \in \{1, \dots, q\}$, decompose Ω into subdomains that represent a laminate of order i ;
- (2) Define piecewise affine mappings $\tilde{y}^{(i)}$ ($1 \leq i \leq q$) on these domains for such laminates;
- (3) Define the transition regions for all laminates, and estimate their volumes;
- (4) Define admissible deformations $y^{(i)} \in \mathcal{A}$ ($1 \leq i \leq q$) by interpolation on transition regions;
- (5) Optimize the thickness of layers and size of transition regions, and define $y_h \in \mathcal{A}_h$ ($0 < h \leq h_0$) to be the finite element interpolation of $y^{(q)} : \Omega \rightarrow \mathbb{R}^3$.

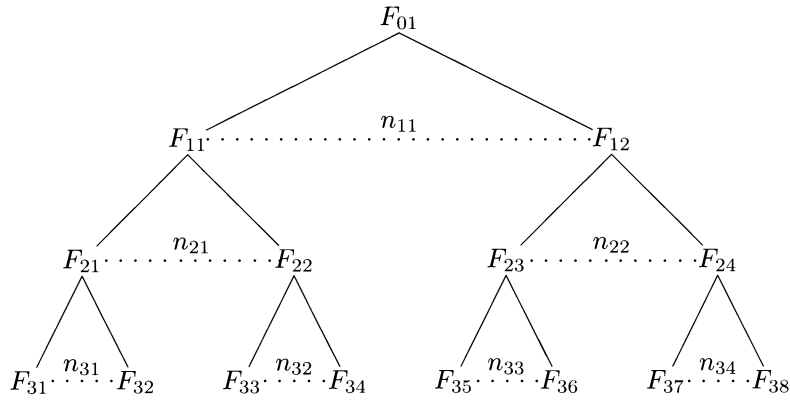


FIGURE 3.1. A $(q + 1)$ -level binary tree of matrices F_{ij} ($i = 0, \dots, q, j = 1, \dots, 2^i$) with $q = 3$. Each parent matrix $F_{i-1,j}$ is an average with volume fractions λ_{ij} and $1 - \lambda_{ij}$ of its two child matrices $F_{i,2j-1}$ and $F_{i,2j}$ that are rank-one connected with normal n_{ij} .

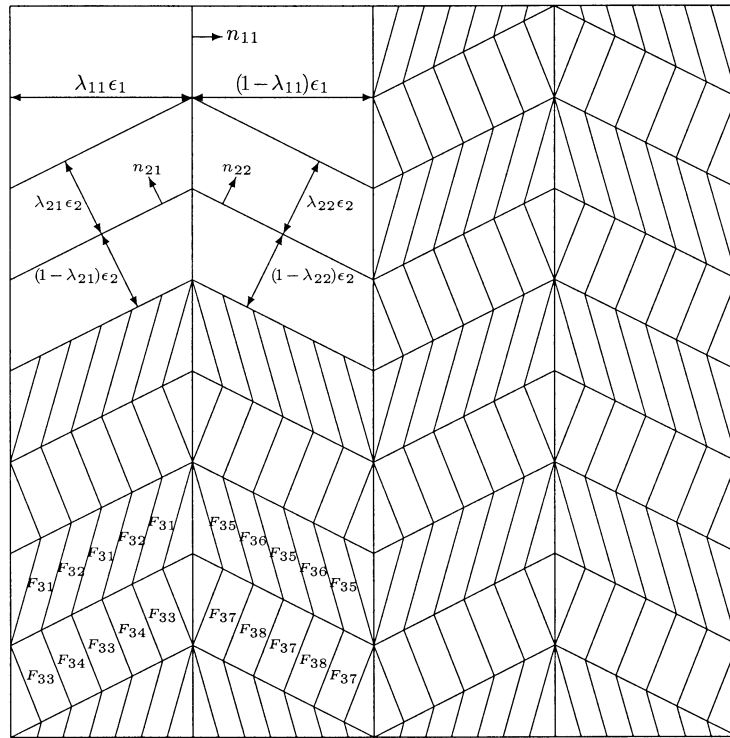


FIGURE 3.2. The geometry of a laminate of order $q = 3$.

Step 1. Let $(\epsilon_0, \dots, \epsilon_q) \in \mathbb{R}^{q+1}$ be such that

$$(3.5) \quad 0 < \epsilon_q < \dots < \epsilon_1 < \epsilon_0 = 1.$$

Each ϵ_i ($1 \leq i \leq q$) will denote the thickness of layers in a laminate of order i under construction. Thus, all ϵ_i ($i = 1, \dots, q$) will be small. Their values are to be specified later. Set

$$\begin{aligned} D_{i,2j-1}^{(k)} &:= \{x \in \mathbb{R}^3 : k\epsilon_i < x \cdot n_{ij} < (k + \lambda_{ij})\epsilon_i\}, \\ D_{i,2j}^{(k)} &:= \{x \in \mathbb{R}^3 : (k + \lambda_{ij})\epsilon_i < x \cdot n_{ij} < (k + 1)\epsilon_i\}, \\ i &= 1, \dots, q, \quad j = 1, \dots, 2^{i-1}, \quad k = 0, \pm 1, \dots \end{aligned}$$

Notice that

$$\bigcup_{k=-\infty}^{\infty} \overline{D_{i,2j-1}^{(k)} \cup D_{i,2j}^{(k)}} = \mathbb{R}^3, \quad i = 1, \dots, q, \quad j = 1, \dots, 2^{i-1}.$$

Set also $\Omega_{01} := \Omega$, and define recursively

$$\begin{aligned} \Omega_{i,2j-1} &:= \Omega_{i-1,j} \cap \left(\bigcup_{k=-\infty}^{\infty} D_{i,2j-1}^{(k)} \right), \\ \Omega_{i,2j} &:= \Omega_{i-1,j} \cap \left(\bigcup_{k=-\infty}^{\infty} D_{i,2j}^{(k)} \right), \end{aligned} \quad i = 1, \dots, q, \quad j = 1, \dots, 2^{i-1}.$$

Obviously,

$$\begin{aligned} \Omega_{i,2j-1} \cap \Omega_{i,2j} &= \emptyset \quad \text{and} \quad \overline{\Omega_{i-1,j}} = \overline{\Omega_{i,2j-1} \cup \Omega_{i,2j}}, \\ i &= 1, \dots, q, \quad j = 1, \dots, 2^{i-1}. \end{aligned}$$

We assume all ϵ_i ($i = 1, \dots, q$) are small enough so that

$$\emptyset \neq \Omega_{i,2j-1} \cup \Omega_{i,2j} \subsetneq \Omega_{i-1,j}, \quad i = 1, \dots, q, \quad j = 1, \dots, 2^{i-1}.$$

Set finally

$$(3.6) \quad \Omega_i := \bigcup_{j=1}^{2^i} \Omega_{ij}, \quad i = 0, \dots, q.$$

Here and below, when no confusion arises, we use ij to denote the double index i, j . It is easy to see that

$$(3.7) \quad \Omega = \Omega_0 = \Omega_1 \supsetneq \dots \supsetneq \Omega_q \quad \text{and} \quad \overline{\Omega} = \overline{\Omega_0} = \overline{\Omega_1} = \dots = \overline{\Omega_q}.$$

The difference $\Omega_i \setminus \Omega_{i-1}$ for $1 \leq i \leq q$ consists of planar boundaries of layers with normals n_{ij} ($j = 1, \dots, 2^{i-1}$) and layer thickness ϵ_i . See Figure 3.3 for a binary tree of these subdomains Ω_{ij} ($i = 0, \dots, q, j = 1, \dots, 2^i$). The structure of this domain tree is identical to that of the matrix tree in Figure 3.1.

Step 2. For any $\lambda \in (0, 1)$, let $\chi_\lambda : \mathbb{R} \rightarrow \mathbb{R}$ be the 1-periodic function defined by

$$\chi_\lambda(t) = \begin{cases} 1 & \text{if } t \in [0, \lambda), \\ 0 & \text{if } t \in [\lambda, 1). \end{cases}$$

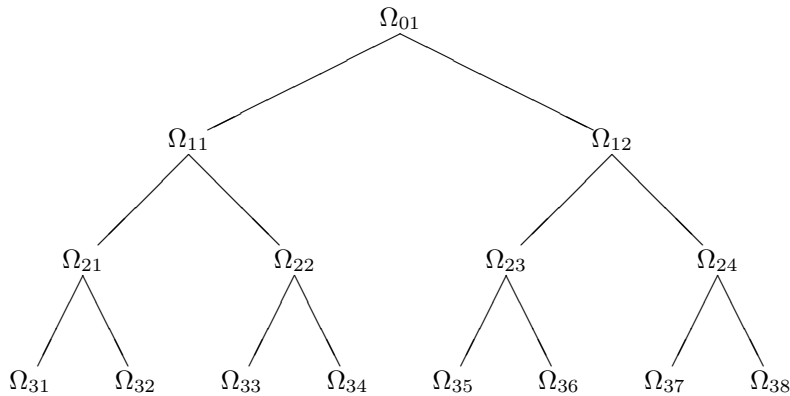


FIGURE 3.3. A $(q + 1)$ -level binary tree of subdomains Ω_{ij} ($i = 0, \dots, q$, $j = 1, \dots, 2^i$) with $q = 3$. Any pair of child matrices $\Omega_{i,2j-1}$ and $\Omega_{i,2j}$ constitute a decomposition of their parent domain $\Omega_{i-1,j}$. All subdomains at the same level constitute a decomposition of Ω .

For convenience, let also $\chi_0 : \mathbb{R} \rightarrow \mathbb{R}$ and $\chi_1 : \mathbb{R} \rightarrow \mathbb{R}$ be defined by $\chi_0(x) = 0$ and $\chi_1(x) = 1$ for all $x \in \mathbb{R}$. Set $\Omega_{-1} := \Omega$ and define $\tilde{y}^{(0)} : \Omega_{-1} \rightarrow \mathbb{R}^3$ by $\tilde{y}^{(0)}(x) = F_0 x$ for all $x \in \Omega_{-1}$. For $i = 1, \dots, q$, we recursively define $\tilde{y}^{(i)} : \Omega_{i-1} \rightarrow \mathbb{R}^3$ by

$$\tilde{y}^{(i)}(x) = F_{i,2j}x + \left[\int_0^{x \cdot n_{ij}} \chi_{\lambda_{ij}} \left(\frac{t}{\epsilon_i} \right) dt \right] a_{ij} + \tilde{y}^{(i-1)}(x) - F_{i-1,j}x$$

$$\forall x \in \Omega_{i-1,j}, \quad j = 1, \dots, 2^{i-1}, \quad i = 1, \dots, q.$$

We claim the following.

1. For each $i \in \{1, \dots, q\}$, $\tilde{y}^{(i)}(x)$ is well defined for any $x \in \Omega_{i-1}$. Moreover,

$$(3.8) \quad \tilde{y}^{(i)} \in W^{1,\infty}(\Omega_{i-1}; \mathbb{R}^3), \quad i = 1, \dots, q.$$

2. Each $\tilde{y}^{(i)} : \Omega_{i-1} \rightarrow \mathbb{R}^3$ is piecewise affine,

$$(3.9) \quad \nabla \tilde{y}^{(i)}(x) = F_{ij} \quad \forall x \in \Omega_{ij}, \quad i = 1, \dots, q, \quad j = 1, \dots, 2^i.$$

3. If $\Omega_{i,j} \cap D_{i,l}^{(k)} \neq \emptyset$ ($1 \leq i \leq q$, $1 \leq j \leq 2^{i-1}$, $l = 2j - 1$ or $2j$, $k \in \mathbb{Z}$), then

$$(3.10) \quad \tilde{y}^{(i)}(x) - F_{i,j}x = \text{constant on } \Omega_{i,j} \cap D_{i,l}^{(k)}.$$

4. The difference of $\tilde{y}^{(i)}$ and $\tilde{y}^{(i-1)}$ is small on Ω_{i-1} ,

$$(3.11) \quad \left| \tilde{y}^{(i)}(x) - \tilde{y}^{(i-1)}(x) \right| \leq \frac{1}{4} \epsilon_i |a_{ij}| \quad \forall x \in \Omega_{i-1}, \quad i = 1, \dots, q.$$

The fact that $\tilde{y}^{(i)}(x)$ is well defined for all $x \in \Omega_{i-1}$ ($1 \leq i \leq q$) and the relation (3.8) follow by induction using (3.6) and (3.7). For $1 \leq i \leq q$ and $1 \leq j \leq 2^{i-1}$, we have by a simple calculation using (3.4) that

$$\begin{aligned} & \nabla \left\{ F_{i,2j}x + \left[\int_0^{x \cdot n_{ij}} \chi_{\lambda_{ij}} \left(\frac{t}{\epsilon_i} \right) dt \right] a_{ij} \right\} \\ &= F_{i,2j} + \chi_{\lambda_{ij}} \left(\frac{x \cdot n_{ij}}{\epsilon_i} \right) a_{ij} \otimes n_{ij} \\ &= \begin{cases} F_{i,2j-1} & \text{if } x \in \Omega_{i,2j-1}, \\ F_{i,2j} & \text{if } x \in \Omega_{i,2j}. \end{cases} \end{aligned}$$

This, together with (3.6), (3.7), and the definition of $\tilde{y}^{(i)} : \Omega_{i-1} \rightarrow \mathbb{R}^3$ ($1 \leq i \leq q$), implies (3.9) by induction. If $\Omega_{i,j} \cap D_{i,l}^{(k)} \neq \emptyset$ ($l = 2j - 1$ or $2j$), then (3.10) follows from (3.9). Notice that $\tilde{y}^{(i)}(x) - F_{i-1,j}x$ is not necessary a constant on $\Omega_{i-1,j}$, since $\Omega_{i-1,j}$ is in general not connected. Finally, in view of the definition of $\tilde{y}^{(i)} : \Omega_{i-1} \rightarrow \mathbb{R}^3$, (3.3), (3.4), and the definition of $\chi_\lambda : \mathbb{R} \rightarrow \mathbb{R}$, we have for $1 \leq i \leq q$ that

$$\begin{aligned} & \left| \tilde{y}^{(i)}(x) - \tilde{y}^{(i-1)}(x) \right| \\ &= \left| F_{i,2j}x + \left[\int_0^{x \cdot n_{ij}} \chi_{ij} \left(\frac{t}{\epsilon_i} \right) dt \right] a_{ij} - [\lambda_{ij}F_{i,2j-1} + (1 - \lambda_{i,2j})F_{i,2j}x] \right| \\ &= \left| \left\{ \int_0^{x \cdot n_{ij}} \left[\chi_{\lambda_{ij}} \left(\frac{t}{\epsilon_i} \right) - \lambda_{ij} \right] dt \right\} a_{ij} \right| \\ &= \epsilon_i \left| \int_0^{\frac{x \cdot n_{ij}}{\epsilon_i}} [\chi_{\lambda_{ij}}(\tilde{t}) - \lambda_{ij}] d\tilde{t} \right| |a_{ij}| \\ &\leq \epsilon_i \lambda_{ij} (1 - \lambda_{ij}) |a_{ij}| \\ &\leq \frac{1}{4} \epsilon_i |a_{ij}| \quad \forall x \in \Omega_{i,2j-1} \cup \Omega_{i,2j}, \quad j = 1, \dots, 2^i. \end{aligned}$$

Now, (3.11) follows from the definition of Ω_{i-1} ($1 \leq i \leq q$), cf. (3.6).

With what has been proved we see on each subdomain $\Omega_{i-1,j}$ ($1 \leq i \leq q, 1 \leq j \leq 2^{i-1}$) that $\tilde{y}^{(i)} : \Omega_{i-1} \rightarrow \mathbb{R}^3$ is a continuous, piecewise affine mapping whose gradient takes alternatively the values $F_{i,2j-1}$ and $F_{i,2j}$ with volume fractions λ_{ij} and $1 - \lambda_{ij}$ on parallel layers that have normal n_{ij} and layer thickness ϵ_i .

Step 3. Let $(\eta_0, \dots, \eta_q) \in \mathbb{R}^{q+1}$ be such that

$$(3.12) \quad 0 < \eta_q < \dots < \eta_0 < 1 \quad \text{and} \quad \eta_i < \epsilon_i, \quad i = 0, \dots, q.$$

Each η_i ($1 \leq i \leq q$) shall denote the size of a transition region in a laminate of order i . All η_i ($i = 0, \dots, q$) shall be small, and their values are to be specified later. Denote

$$\omega(\eta) := \{x \in \omega : \text{dist}(x, \partial\omega) > \eta\}$$

for any $\omega \subseteq \Omega$ and $\eta > 0$. Set $\tilde{\Omega}_{-1} := \Omega$, and define recursively

$$\tilde{\Omega}_i := \left(\Omega_i \cap \tilde{\Omega}_{i-1} \right) (\eta_i), \quad i = 0, \dots, q.$$

We assume all $\eta_i > 0$ are small enough so that $\tilde{\Omega}_i \neq \emptyset$ ($i = 0, \dots, q$). Obviously,

$$(3.13) \quad \tilde{\Omega}_q \subsetneq \dots \subsetneq \tilde{\Omega}_0 \subsetneq \tilde{\Omega}_{-1} = \Omega.$$

Denoting by $C_T > 0$ a generic constant which can only depend on Ω , q , and all the unit normals n_{ij} ($i = 1, \dots, q$, $j = 1, \dots, 2^{i-1}$), we claim that

$$(3.14) \quad \text{meas}(\Omega \setminus \tilde{\Omega}_i) \leq C_T \left(\frac{\eta_0}{\epsilon_0} + \frac{\eta_1}{\epsilon_1} + \dots + \frac{\eta_i}{\epsilon_i} \right), \quad i = 0, \dots, q.$$

Since by (3.13)

$$\text{meas}(\Omega \setminus \tilde{\Omega}_i) = \sum_{l=0}^i \text{meas}(\tilde{\Omega}_{l-1} \setminus \tilde{\Omega}_l), \quad i = 0, \dots, q,$$

and by the definition of $\tilde{\Omega}_i$ ($1 \leq i \leq q$)

$$\text{meas}(\tilde{\Omega}_{i-1} \setminus \tilde{\Omega}_i) \leq C_T \sum_{j=1}^{2^{i-1}} \text{meas}(\Omega_{ij} \setminus \Omega_{ij}(\eta_i)), \quad i = 0, \dots, q,$$

we need only to prove that

$$(3.15) \quad \text{meas}(\Omega_{ij} \setminus \Omega_{ij}(\eta_i)) \leq C_T \frac{\eta_i}{\epsilon_i}, \quad i = 0, \dots, q, \quad j = 1, \dots, 2^{i-1}.$$

The inequality in (3.15) is trivially true for $i = 0$. Consider now $1 \leq i \leq q$. We only show that the inequality in (3.15) holds true for $j = 1$, since the same argument can be used for all j ($1 \leq j \leq 2^{i-1}$). Notice that each connected component of Ω_{i1} is a small band or thin plate. If its closure is in the interior of Ω , then it is a parallelepiped. Otherwise, it is part of a parallelepiped. In fact, all these parallelepipeds at the level i have the same face normals and face areas: they are translations of a single parallelepiped, say P_{i1} . Denoting the number of these parallelepipeds by N_{i1} and setting $V_{i1} := \text{meas}(P_{i1} \setminus P_{i1}(\eta_i))$, we easily see that $\text{meas}(\Omega_{i1} \setminus \Omega_{i1}(\eta_i))$ is bounded by $C_T N_{i1} V_{i1}$. To estimate N_{i1} and V_{i1} , we let $\epsilon_{-2} = \epsilon_{-1} = \epsilon_0 = 1$ and $n_l := n_{l1}$ for $l = 1, \dots, i$. We also let $n_{-2} = n_1$ and $n_{-1}, n_0 \in \mathbb{R}^3$ be unit vectors such that n_{-2}, n_{-1} , and n_0 form an orthonormal basis for \mathbb{R}^3 . We claim that there exist a permutation $(t_{-2}t_{-1} \dots t_{i-1})$ of $(-2, -1, \dots, i-1)$ such that the face normals of P_{i1} are $n_{t_{i-2}}, n_{t_{i-1}}, n_i$, and

$$(3.16) \quad N_{i1} \leq C_T \frac{\epsilon_{t_{-2}} \epsilon_{t_{-1}} \dots \epsilon_{t_{i-3}}}{\epsilon_1 \epsilon_2 \dots \epsilon_i} \quad \text{and} \quad V_{i1} \leq C_T \eta_i \epsilon_{t_{i-2}} \epsilon_{t_{i-1}}.$$

This is obviously true for the case $i = 1$. Suppose it is true for a general i with $1 \leq i \leq q-1$. Denoting $\epsilon_{t_i} := \epsilon_i$ and $n_{t_i} := n_i$, we see that there exists $m \in \{i-2, i-1, i\}$ such that the set of face normals of $P_{i+1,1}$ is $n_{i+1} \cup \{n_{t_{i-2}}, n_{t_{i-1}}, n_{t_i}\} \setminus \{n_{t_m}\}$, the number of bands at the level $i+1$ is

$$N_{i+1,1} \leq C_T N_{i1} \frac{\epsilon_{t_m}}{\epsilon_{i+1}},$$

and the volume is

$$V_{i+1,1} \leq C_T \eta_{i+1} \frac{\epsilon_{t_{i-2}} \epsilon_{t_{i-1}} \epsilon_{t_i}}{\epsilon_{t_m}},$$

since $0 < \epsilon_{i+1} < \min\{\epsilon_{t_{i-2}}, \epsilon_{t_{i-1}}, \epsilon_{t_i}\}$. Therefore, setting $s_l := t_l$ for $-2 \leq l \leq i-1$, $s_{i-2} := t_m$, and $\{s_{i-1}, s_i\} := \{t_{i-2}, t_{i-1}, t_i\} \setminus \{t_m\}$, we see that $(s_{-2} s_{-1} \dots s_i)$ is a permutation of $(-2, -1, \dots, i)$ such that the face normals of $P_{i+1,1}$ are $n_{s_{i-1}}, n_{s_i}$, and n_{i+1} , and such that

$$N_{i+1,1} \leq C_T \frac{\epsilon_{s_{-2}} \epsilon_{s_{-1}} \dots \epsilon_{s_{i-2}}}{\epsilon_1 \epsilon_2 \dots \epsilon_{i+1}} \quad \text{and} \quad V_{i+1,1} \leq C_T \eta_{i+1} \epsilon_{s_{i-1}} \epsilon_{s_i}.$$

This proves (3.16) for any i ($1 \leq i \leq q$). By (3.16), we have that

$$\text{meas}(\Omega_{i1} \setminus \Omega_{i1}(\eta_i)) \leq C_T N_{i1} V_{i1} \leq C_T \frac{\eta_i}{\epsilon_i},$$

proving (3.15).

Step 4. For each $i \in \{0, \dots, q-1\}$, let $\rho_i \in C_0^\infty(\mathbb{R}^3)$ be such that

$$\begin{aligned} 0 \leq \rho_i(x) \leq 1 & \quad \forall x \in \mathbb{R}^3, \\ \rho_i(x) = 1 & \quad \forall x \in \tilde{\Omega}_i, \\ \rho_i(x) = 0 & \quad \forall x \in \mathbb{R}^3 \setminus (\Omega_i \cap \tilde{\Omega}_{i-1}), \\ |\nabla \rho_i(x)| \leq \frac{2}{\eta_i} & \quad \forall x \in \mathbb{R}^3. \end{aligned}$$

Let $y^{(0)} : \Omega \rightarrow \mathbb{R}^3$ be defined by $y^{(0)}(x) := F_0 x$ for all $x \in \Omega$. Define $y^{(i)} : \Omega \rightarrow \mathbb{R}^3$ ($i = 1, \dots, q$) recursively by

$$(3.17) \quad y^{(i)}(x) = \rho_{i-1}(x) \tilde{y}^{(i)}(x) + [1 - \rho_{i-1}(x)] y^{(i-1)}(x) \quad \forall x \in \Omega.$$

We claim the following.

1. For each i ($1 \leq i \leq q$), $y^{(i)} \in W^{1,\infty}(\Omega; \mathbb{R}^3)$ and

$$(3.18) \quad y^{(i)}(x) = \begin{cases} \tilde{y}^{(i)}(x) & \forall x \in \tilde{\Omega}_{i-1}, \\ y^{(i-1)}(x) & \forall x \in \Omega \setminus (\Omega_i \cap \tilde{\Omega}_{i-1}). \end{cases}$$

2. We have

$$(3.19) \quad y^{(k)}(x) = F_{i-1,j} x \quad \forall x \in \partial \tilde{\Omega}_{i-1}, \quad k = i, \dots, q.$$

3. We have

$$(3.20) \quad \begin{aligned} \nabla y^{(i)}(x) &\in \{F_{ij} : j = 1, \dots, 2^{i-1}\} \\ &\text{a.e. } x \in \tilde{\Omega}_{i-1}, \quad i = 1, \dots, q, \end{aligned}$$

$$(3.21) \quad \begin{aligned} \left| \nabla y^{(i)}(x) \right| &\leq \frac{C_a \epsilon_i}{2\eta_{i-1}} + C_F \\ &\text{a.e. } x \in \Omega \setminus \tilde{\Omega}_{i-1}, \quad i = 1, \dots, q, \end{aligned}$$

where

$$C_a := \max_{1 \leq i \leq q, 1 \leq j \leq 2^{i-2}} |a_{ij}| > 0 \quad \text{and} \quad C_F := \max_{1 \leq i \leq q, 1 \leq j \leq 2^i} \|F_{ij}\| > 0.$$

The fact that $y^{(i)} \in W^{1,\infty}(\Omega; \mathbb{R}^3)$ and (3.18) follow from the definition of $y^{(i)}$ ($1 \leq i \leq q$), and (3.19) follows from (3.18). A simple calculation leads to

$$(3.22) \quad \begin{aligned} \nabla y^{(i)}(x) &= \left[\tilde{y}^{(i)}(x) - y^{(i-1)}(x) \right] \otimes \nabla \rho_{i-1}(x) + \rho_{i-1}(x) \nabla \tilde{y}^{(i)}(x) \\ &\quad + [1 - \rho_{i-1}(x)] \nabla y^{(i-1)}(x) \quad \text{a.e. } x \in \Omega, \quad i = 1, \dots, q. \end{aligned}$$

This, together with (3.9) and (3.11), leads to (3.20) and (3.21).

Step 5. Set $\epsilon_i := h^{\alpha_i}$ and $\eta_i := h^{\beta_i}$ for $i = 0, \dots, q$, where all α_i and β_i are real numbers such that

$$(3.23) \quad 0 = \alpha_0 < \beta_0 \leq \alpha_1 < \beta_1 \leq \alpha_2 < \dots < \beta_{q-1} \leq \alpha_q < \beta_q = 1.$$

It is easy to see that the assumptions (3.5) and (3.12) are satisfied with this choice of α_i and β_i ($i = 0, \dots, q$). Define $y_h \in \mathcal{A}_h$ by $y_h := \mathcal{I}_h y^{(q)} \in \mathcal{A}_h$ for $h \in (0, h_0]$.

The uniform boundedness (3.1) follows from (3.20), (3.21), (3.23), and (2.9). By the property $H2$, (3.20), the definition of $\tilde{\Omega}_q$, and the fact that $\eta_q = h$, we have that $\nabla y(x) \in \mathcal{L}$ for all $x \in \tilde{\Omega}_q$, where

$$\mathcal{L} := \{F_{qj} : j = 1, \dots, 2^{q-1}\}.$$

Thus,

$$\{x \in \Omega : \nabla y(x) \notin \mathcal{L}\} \subseteq \Omega \setminus \tilde{\Omega}_q.$$

It remains now to choose all the parameters α_i and β_i that satisfy (3.23) and that minimize $\text{meas}(\Omega \setminus \tilde{\Omega}_q)$.

It follows from (3.14) and (3.23) that

$$\text{meas}(\Omega \setminus \tilde{\Omega}_q) \leq C (h^{\beta_0} + h^{\beta_1 - \alpha_1} + \dots + h^{\beta_{q-1} - \alpha_{q-1}} + h^{1 - \alpha_q}).$$

This is minimized if we choose, according to (3.23), all $\alpha_i = \beta_{i-1}$ for $i = 1, \dots, q$. With such a choice, we have

$$(3.24) \quad \text{meas}(\Omega \setminus \tilde{\Omega}_q) \leq C (h^{\beta_0} + h^{\beta_1 - \beta_0} + \dots + h^{\beta_{q-1} - \beta_{q-2}} + h^{1 - \beta_{q-1}}),$$

where

$$0 < \beta_0 < \dots < \beta_{q-1} < \beta_q = 1.$$

The sum in the inequality (3.24) attains its minimum value $h^{1/(q+1)}$, leading to (3.2), if

$$\beta_0 = \beta_1 - \beta_0 = \dots = \beta_{q-1} - \beta_{q-2} = 1 - \beta_{q-1},$$

i.e.,

$$\beta_i = \frac{i+1}{q+1}, \quad i = 0, \dots, q-1.$$

The proof is complete. \square

4. ERROR ESTIMATES FOR FINITE ELEMENT ENERGY MINIMIZERS

We define a projection $\pi : \mathbb{R}^{3 \times 3} \rightarrow \mathcal{U}$ by

$$(4.1) \quad \|F - \pi(F)\| = \text{dist}(F, \mathcal{U}), \quad F \in \mathbb{R}^{3 \times 3}.$$

It is shown in [8] that, with a possible modification of its definition on a subset of $\mathbb{R}^{3 \times 3}$ of Lebesgue measure zero, this projection is well defined and Borel measurable. We denote by C a generic, positive constant that is always assumed to be independent of the finite element mesh size h .

Theorem 4.1. *The following estimates hold true for all $y_h \in \mathcal{A}_h$.*

(1) *Estimate on the variant reduction in measure:*

$$(4.2) \quad \text{meas}\{x \in \Omega : \pi(\nabla y_h(x)) \notin \mathcal{S}\} \leq C [\mathcal{E}(y_h)^{1/2} + \mathcal{E}(y_h)],$$

where \mathcal{S} is the subset of energy wells defined in the condition F3.

(2) *Estimate on directional derivatives in the L^2 norm:*

$$(4.3) \quad \int_{\Omega} |[\nabla y_h(x) - F_0] e_0|^2 dx \leq C [\mathcal{E}(y_h)^{1/2} + \mathcal{E}(y_h)],$$

where e_0 is the unit vector defined in F1.

(3) Estimate on deformations in the L^2 norm:

$$(4.4) \quad \int_{\Omega} |y_h(x) - y_0(x)|^2 dx \leq C \left[\mathcal{E}(y_h)^{1/2} + \mathcal{E}(y_h) \right],$$

where y_0 is the deformation in the boundary condition, cf. (2.4) and (2.5).

(4) Estimate on deformation gradients in a weak topology:

$$(4.5) \quad \left\| \int_{\omega} [\nabla y_h(x) - F_0] dx \right\| \leq C(\omega) \left[\mathcal{E}(y_h)^{1/8} + \mathcal{E}(y_h)^{1/2} \right],$$

where $\omega \subseteq \Omega$ is a Lipschitz domain and $F_0 = \nabla y_0$ is the gradient of the homogeneous deformation y_0 in the boundary condition, cf. (2.4) and (2.5).

Replacing a general y_h in the theorem by any finite element energy minimizer and using the energy bound established in Corollary 3.1, we immediately obtain the following error estimates for all the finite element energy minimizers defined by (2.10).

Corollary 4.1. For any finite element minimizers $y_h \in \mathcal{A}_h$,

$$\begin{aligned} \text{meas} \{x \in \Omega : \pi(\nabla y_h(x)) \notin \mathcal{S}\} &\leq Ch^{1/2(q+1)}, \\ \int_{\Omega} |[\nabla y_h(x) - F_0] e_0|^2 dx &\leq Ch^{1/2(q+1)}, \\ \int_{\Omega} |y_h(x) - y_0(x)|^2 dx &\leq Ch^{1/2(q+1)}, \\ \left\| \int_{\omega} [\nabla y_h(x) - F_0] dx \right\| &\leq C(\omega) h^{1/(8q+8)}, \end{aligned}$$

where $\omega \subseteq \Omega$ is a Lipschitz domain.

Proof of Theorem 4.1. Let $y_h \in \mathcal{A}_h$.

(1) Let $w \in \mathbb{R}^3$ with $|w| = 1$. By the minors relations and the growth condition (2.3), using arguments similar to those in [5, 8], we obtain

$$(4.6) \quad \int_{\Omega} |[\pi(\nabla y_h(x)) - F_0] w|^2 dx - \int_{\Omega} [|\pi(\nabla y_h(x)) w|^2 - |F_0 w|^2] dx \leq C \mathcal{E}(y_h)^{1/2}$$

and

$$(4.7) \quad \begin{aligned} &\int_{\Omega} |[\text{Cof } \pi(\nabla y_h(x)) - \text{Cof } F_0] w|^2 dx \\ &\quad - \int_{\Omega} \{ |[\text{Cof } \pi(\nabla y_h(x))] w|^2 - |(\text{Cof } F_0) w|^2 \} dx \\ &\leq C \left[\mathcal{E}(y_h)^{1/2} + \mathcal{E}(y_h) \right]. \end{aligned}$$

Denote

$$\Omega_i(y_h) := \{x \in \Omega : \pi(\nabla y_h(x)) \in \mathcal{U}_i\}, \quad i = 1, \dots, N.$$

Notice that all $\Omega_i(y_h)$ ($i = 1, \dots, N$) are pairwise disjoint. Fix $j \in \{s+1, \dots, N\}$, cf. the conditions F1–F3. If (2.7) holds true for some $a_j \in \mathbb{R}^3$, then by (4.6) with

$w = a_j$ we have

$$\begin{aligned} C\mathcal{E}(y_h)^{1/2} &\geq \int_{\Omega} [|F_0 a_j|^2 - |\pi(\nabla y_h(x)) a_j|^2] dx \\ &= \sum_{i=1}^N \int_{\Omega_i(y_h)} [|F_0 a_j|^2 - |\pi(\nabla y_h(x)) a_j|^2] dx \\ &= \sum_{i=1}^N \text{meas } \Omega_i(y_h) [|F_0 a_j|^2 - |U_i a_j|^2] \\ &\geq \text{meas } \Omega_{i_j}(y_h) [|F_0 a_j|^2 - |U_{i_j} a_j|^2], \end{aligned}$$

which, together with (2.7), leads to

$$(4.8) \quad \text{meas } \Omega_{i_j}(y_h) \leq C\mathcal{E}(y_h)^{1/2}.$$

If instead (2.8) holds true for some $b_j \in \mathbb{R}^3$, then a similar argument using (4.7) with $w = b_j$ leads to

$$(4.9) \quad \text{meas } \Omega_{i_j}(y_h) \leq C [\mathcal{E}(y_h)^{1/2} + \mathcal{E}(y_h)].$$

Now, the estimates (4.8) and (4.9), together with the fact that

$$(4.10) \quad \{x \in \Omega : \pi(\nabla y_h(x)) \notin \mathcal{S}\} = \bigcup_{j=s+1}^N \Omega_{i_j}(y_h),$$

lead to (4.2).

(2) Since $|e_0| = 1$, we have, by the growth condition (2.3)

$$(4.11) \quad \int_{\Omega} |[\nabla y_h(x) - \pi(\nabla y_h(x))] e_0|^2 dx \leq \int_{\Omega} \|\nabla y_h(x) - \pi(\nabla y_h(x))\|^2 dx \leq \kappa^{-1} \mathcal{E}(y_h).$$

In view of (4.6) with $w = e_0$, (2.6), (4.10), and (4.2), we have

$$\begin{aligned} &\int_{\Omega} |[\pi(\nabla y_h(x)) - F_0] e_0|^2 dx \\ &\leq C\mathcal{E}(y_h)^{1/2} + \int_{\Omega} [|\pi(\nabla y_h(x)) e_0|^2 - |F_0 e_0|^2] dx \\ &= C\mathcal{E}(y_h)^{1/2} + \int_{\{x \in \Omega : \pi(\nabla y_h(x)) \in \mathcal{S}\}} [|\pi(\nabla y_h(x)) e_0|^2 - |F_0 e_0|^2] dx \\ (4.12) \quad &+ \int_{\{x \in \Omega : \pi(\nabla y_h(x)) \notin \mathcal{S}\}} [|\pi(\nabla y_h(x)) e_0|^2 - |F_0 e_0|^2] dx \\ &\leq C\mathcal{E}(y_h)^{1/2} + \sum_{j=1}^s \text{meas } \Omega_{i_j} [|U_{i_j} e_0|^2 - |F_0 e_0|^2] \\ &\quad + C \text{meas } \{x \in \Omega : \pi(\nabla y_h(x)) \notin \mathcal{S}\} \\ &\leq C [\mathcal{E}(y_h)^{1/2} + \mathcal{E}(y_h)]. \end{aligned}$$

Now, (4.11), (4.12), and an application of the triangle inequality imply (4.3).

(3) This follows from the Poincaré inequality and (4.3), cf. [5, 11].

(4) We obtain (4.5) from (4.4) by using the same argument as in [5, 11]. \square

ACKNOWLEDGMENT

The author thanks the referee for suggestions that helped improve the presentation of this work.

REFERENCES

1. J. M. Ball and R. D. James, *Fine phase mixtures as minimizers of energy*, Arch. Rational Mech. Anal. **100** (1987), 13–52. MR **89c**:80005
2. ———, *Proposed experimental tests of a theory of fine microstructure and the two-well problem*, Phil. Trans. R. Soc. Lond. A **338** (1992), 389–450.
3. K. Bhattacharya and G. Dolzmann, *Relaxed constitutive relations for phase transforming materials*, J. Mech. Phys. Solids **48** (2000), no. 6–7, 1493–1517. MR **2001g**:74053
4. ———, *Relaxation of some multi-well problems*, Proc. R. Soc. Edinburgh: Sect. A **131** (2001), 279–320. MR **2002c**:49011
5. K. Bhattacharya, B. Li, and M. Luskin, *The simply laminated microstructure in martensitic crystals that undergo a cubic to orthorhombic phase transformation*, Arch. Rational Mech. Anal. **149** (1999), no. 2, 123–154. MR **2001i**:74059
6. G. Dolzmann, *private communication*, 1998.
7. ———, *Variational methods for crystalline microstructure – analysis and computation*, Habilitationsschrift, Universität Leipzig (2001).
8. B. Li, *Approximation of martensitic microstructure with general homogeneous boundary data*, J. Math. Anal. Appl. **266** (2002), 451–467.
9. B. Li and M. Luskin, *Finite element analysis of microstructure for the cubic to tetragonal transformation*, SIAM J. Numer. Anal. **35** (1998), no. 1, 376–392. MR **99b**:73015
10. ———, *Approximation of a martensitic laminate with varying volume fractions*, RAIRO Math. Model. Numer. Anal. **33** (1999), no. 1, 67–87. MR **2000c**:74022
11. M. Luskin, *Approximation of a laminated microstructure for a rotationally invariant, double well energy density*, Numer. Math. **75** (1996), 205–221. MR **97k**:73026
12. ———, *On the computation of crystalline microstructure*, Acta Numerica (1996), 191–257. MR **99f**:73030

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MARYLAND, COLLEGE PARK, MARYLAND 20742
E-mail address: bli@math.umd.edu