

## LINEAR CONVERGENCE IN THE APPROXIMATION OF RANK-ONE CONVEX ENVELOPES

SÖREN BARTELS<sup>1</sup>

**Abstract.** A linearly convergent iterative algorithm that approximates the rank-1 convex envelope  $f^{rc}$  of a given function  $f : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}$ , *i.e.* the largest function below  $f$  which is convex along all rank-1 lines, is established. The proposed algorithm is a modified version of an approximation scheme due to Dolzmann and Walkington.

**Mathematics Subject Classification.** 65K10, 74G15, 74G65, 74N99.

Received: May 27, 2003.

### 1. INTRODUCTION

A well known mathematical model of phase transitions in crystalline solids due to [2] allowing for microstructure reads

$$(M) \quad \text{Minimize } I(u) := \int_{\Omega} f(x, u, \nabla u) \, dx \quad \text{among } u \in \mathcal{A}$$

for a bounded Lipschitz domain  $\Omega \subseteq \mathbb{R}^n$ , a (non-convex) continuous energy density  $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{n \times m} \rightarrow \mathbb{R}$  satisfying  $p$ -growth conditions for some parameter  $p \geq 1$ , and a space of admissible deformations  $\mathcal{A} \subseteq W^{1,p}(\Omega; \mathbb{R}^m)$  containing boundary conditions. It is well established that minimizing sequences for  $I$  develop oscillations in the gradient and that their weak limits do not minimize  $I$  (see *e.g.* [9, 20, 22]). Together with a Young measure generated by a minimizing sequence in the sense of [1], weak limits contain the most relevant information about microscopic and macroscopic effects. Moreover, each weak limit of a minimizing sequence is a solution of a relaxed problem in which  $f$  is replaced by its quasiconvex envelope  $f^{qc}$  (see *e.g.* [9, 20, 22]). In general, it is not possible to compute  $f^{qc}$  explicitly or even approximately in order to define the relaxed problem. Therefore, it is desirable to know upper and lower bounds for  $f^{qc}$  and it is the aim of this paper to establish an algorithm that computes an upper bound for  $f^{qc}$  with optimal orders of convergence. Approximated upper bounds have been employed in [14, 17, 19] for effective numerical simulations of inelastic and plastic materials. The approach of simultaneous relaxation and approximation of non-convex variational problems results in discrete problems with two scales that reflect microscopic and macroscopic effects.

Error estimates for the approximation of  $(M)$  are available for the case that either  $\mathcal{A}$  contains affine boundary conditions on  $\partial\Omega$  defined through certain  $F \in \mathbb{R}^{n \times m}$  (see *e.g.* [5, 8, 18]) or  $f^{qc}$  is convex (see *e.g.* [4, 6, 7]).

---

*Keywords and phrases.* Nonconvex variational problem, calculus of variations, relaxed variational problems, rank-1 convex envelope, microstructure, iterative algorithm.

<sup>1</sup> Department of Mathematics, University of Maryland, College Park, MD 20742-4015, USA. e-mail: [sba@math.umd.edu](mailto:sba@math.umd.edu)

In the first case theoretical convergence rates for the approximation of  $(M)$  and thereby of  $f^{qc}(\cdot, \cdot, F)$  are stated but, owing to mesh-size dependent oscillations, those approaches cannot be expected to lead to efficient numerical algorithms. In the second case efficient algorithms are available but the proposed numerical schemes are restricted to scalar problems. A numerical scheme that checks for different notions of convexity for a class of functions can be found in [10].

An algorithm that computes an upper bound, *i.e.* the rank-1 convex envelope [9], for  $f^{qc}$  is given in [11, 12]. By choosing a larger set of discrete rank-1 matrices and allowing interpolated values we state a version of that algorithm which leads to improved convergence rates under the cost of a slightly increased numerical effort. Thereby, we efficiently solve  $(M)$  for a large class of functions  $f$  and affine boundary data on  $\partial\Omega$ . The combination of the presented result with an approximation scheme for a lower bound for  $f^{qc}$  recently given in [3] further allows to numerically check for equality of rank-1 convex and polyconvex envelopes and thereby to detect  $f^{qc}$ .

The rest of the paper is organized as follows. We present the iterative algorithm and the approximation result in Section 2. Some notation and preliminaries are stated in Section 3 and lead to the proof of the main result which is given in Section 4. Numerical experiments are reported on in Section 5. In Section 6 we numerically check for equality of the rank-1 convex and the polyconvex envelope of a two-dimensional version of the Ericksen-James energy density.

## 2. ITERATIVE ALGORITHM AND MAIN RESULTS

Throughout this article we suppose that  $f$  in  $(M)$  is independent of  $x$  and  $u$ , *i.e.*  $f : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}$ , is continuous, and satisfies, for certain  $c_f > 0$ ,  $c'_f \geq 0$ ,  $p > 0$ , and all  $F \in \mathbb{R}^{n \times m}$ ,

$$f(F) \geq c_f |F|^p - c'_f. \tag{2.1}$$

Due to [16] the rank-1 convex envelope  $f^{rc}$  can be defined iteratively by setting  $f^{(0)} := f$  and, for  $k \geq 1$  and  $F \in \mathbb{R}^{n \times m}$ ,

$$f^{(k)}(F) := \inf \{ \theta f^{(k-1)}(A) + (1 - \theta) f^{(k-1)}(B) : \theta \in [0, 1], \\ A, B \in \mathbb{R}^{n \times m}, \theta A + (1 - \theta) B = F, \text{rank}(A - B) = 1 \}.$$

Then,  $f^{rc}$  is the pointwise limit of  $f^{(k)}$  for  $k \rightarrow \infty$ . The idea for the approximation of  $f^{rc}$  is to choose a discrete set of rank-1 matrices  $a_d \otimes b_d$  and then to approximate  $f^{(k)}$  in certain nodes employing the finite set of rank-1 matrices. We choose the set of nodes  $\mathcal{N}_{d,r} := d\mathbb{Z}^{n \times m} \cap B_r(0)$  for  $0 < d \leq r$ , let  $\omega_{d,r} := \text{conv } \mathcal{N}_{d,r}$ , and define

$$\mathcal{R}_{d,r}^1 := \{ (a_d, b_d) \in d\mathbb{Z}^n \times d\mathbb{Z}^m : (1 - 2nd)^{1/2} \leq |a_d| \leq 1 + nd, \\ |b_d| \leq 2\sqrt{nm}r + md \}.$$

We then let  $f_{d,r}^{(0)} := \mathcal{I}_{d,r} f$  be the nodal interpolant of  $f$  in  $\omega_{d,r}$  and solve, for  $k \geq 1$  and  $F \in \mathcal{N}_{d,r}$ , (with  $f_{d,r}^{(k-1)} := \infty$  in  $\mathbb{R}^{n \times m} \setminus \overline{\omega}_{d,r}$  and nodal interpolation of  $f_{d,r}^{(k-1)}$  in  $\overline{\omega}_{d,r}$ ) the linear optimization problem

$$f_{d,r}^{(k)}(F) := \inf_{(a_d, b_d) \in \mathcal{R}_{d,r}^1} \inf \left\{ \sum_{\ell \in \mathbb{Z}} \theta_\ell f_{d,r}^{(k-1)}(F + d\ell a_d \otimes b_d) : \right. \\ \left. \theta_\ell \geq 0, \sum_{\ell \in \mathbb{Z}} \theta_\ell = 1, \sum_{\ell \in \mathbb{Z}} \theta_\ell \ell = 0 \right\}. \tag{2.2}$$

Thereby, we obtain the following approximation result.

**Theorem A.** *There holds  $f_{d,r}^{(k)} \geq f^{rc}$  in  $\overline{\omega}_{d,r}$ . For  $c_1 := 1 + (2\sqrt{nm}r + (1+d)\sqrt{m}(2nr+1+nd))$  set  $\tilde{r} := r - c_1d$  and suppose that  $f^{rc} = f$  in  $\mathbb{R}^{n \times m} \setminus B_{\tilde{r}}(0)$ . If  $f^{(L)} = f^{rc}$  for some  $L \geq 0$  there holds*

$$\|f_{d,r}^{(L)} - f^{rc}\|_{L^\infty(\omega_{d,r})} \leq (c_2L + \sqrt{nm})d \|f\|_{Lip(B_{3r}(0))}$$

with  $c_2 := 2(6\sqrt{nm}r + (1+d)\sqrt{m}(2nr+1+nd)) + \sqrt{nm}$ .

The computational effort to calculate  $f_{d,r}^{(k)}$  is of order  $k(r/d)^{nm+n+m+1}$  while the cheaper algorithm of Dolzmann and Walkington requires  $\mathcal{O}(k(r/d)^{nm+n/3+m/3+1})$  operations but leads to an approximation error  $\mathcal{O}(d^{1/3})$ . There is no *a priori* bound for  $L$  available but there holds  $f^{rc} = f^{(L)}$  if and only if  $f^{(L)} = f^{(L+1)}$  which motivates to stop the discrete iterative process (2.2) if  $|f_{d,r}^{(k+1)} - f_{d,r}^{(k)}|_{L^\infty(\omega_{d,r})} \leq d$ . The condition  $f^{rc} = f$  in  $\mathbb{R}^{n \times m} \setminus B_{\tilde{r}}(0)$  is not always true and hard to verify in practice. We stress however that even if we do not know  $L$  and  $\tilde{r}$  in Theorem A we have  $f_{d,r}^{(k)}(F) \geq f^{rc}(F)$  for all  $k \geq 0$ ,  $r \geq d > 0$ , and  $F \in \omega_{d,r}$  which is important if one is interested in a reliable upper bound for  $f^{qc}$  in order to check for equality with a lower bound. A similar approximation result to Theorem A can be obtained if  $\mathcal{N}_{d,r}$  is replaced by a non-uniform set of nodes but then it is not clear whether  $f_{d,r}^{(k)}$  is a reliable upper bound for  $f^{rc}$  in  $\omega_{d,r}$ .

### 3. PRELIMINARIES

Throughout this article,  $|\cdot|$  denotes the Frobenius norm of a vector or a matrix in  $\mathbb{R}^n$ ,  $\mathbb{R}^m$ , or  $\mathbb{R}^{n \times m}$ , e.g. for  $A \in \mathbb{R}^{n \times m}$  with entries  $A_{jk} \in \mathbb{R}$  for  $j = 1, \dots, n$  and  $k = 1, \dots, m$  we have

$$|A|^2 = \sum_{j=1}^n \sum_{k=1}^m A_{jk}^2.$$

The maximum norm of a vector or a matrix is denoted by  $|\cdot|_\infty$ , e.g.  $|A|_\infty = \max_{j,k} |A_{jk}|$ ; there holds  $|v|_\infty \leq |v| \leq \sqrt{\ell}|v|_\infty$  for all  $v \in \mathbb{R}^\ell$ .

For  $s \in \mathbb{R}$  the integer  $\lfloor s \rfloor \in \mathbb{Z}$  is defined by  $\lfloor s \rfloor := \max\{\ell \in \mathbb{Z} : \ell \leq s\}$ . For vectors the operator  $\lfloor \cdot \rfloor$  is defined by applying  $\lfloor \cdot \rfloor$  to each component.

Given  $r > 0$  and  $G \in \mathbb{R}^\ell$  we set  $B_r(G) := \{A \in \mathbb{R}^\ell : |A - G|_\infty < r\}$  and, for a positive parameter  $d > 0$  with  $d \leq r$ , define

$$\mathcal{N}_{d,r} := d\mathbb{Z}^{n \times m} \cap \overline{B_r(0)} \subseteq \mathbb{R}^{n \times m}.$$

We let  $\omega_{d,r}$  be the interior of the union of all closed  $(nm)$ -dimensional cubes  $Q \subseteq \overline{B_r(0)}$  with vertices in  $\mathcal{N}_{d,r}$ , and define a uniform triangulation  $\mathcal{T}_{d,r}$  of  $\omega_{d,r}$  by setting

$$\mathcal{T}_{d,r} := \{Q \subseteq \overline{\omega}_{d,r} : Q \text{ is a closed cube with vertices in } \mathcal{N}_{d,r} \text{ and edges of length } d\}.$$

Note that each  $Q \in \mathcal{T}_{d,r}$  is the convex hull of  $2^{nm}$  nodes  $M_1, \dots, M_{2^{nm}} \in \mathcal{N}_{d,r}$ , i.e.  $Q = \text{conv}\{M_1, \dots, M_{2^{nm}}\}$ . To  $\mathcal{T}_{d,r}$  we associate the set of continuous,  $\mathcal{T}_{d,r}$ -elementwise  $(nm)$ -linear functions

$$\mathcal{S}^1(\mathcal{T}_{d,r}) := \{v_h \in C(\overline{\omega}_{d,r}) : \forall Q \in \mathcal{T}_{d,r}, v_h|_Q \text{ is a polynomial of partial degree } \leq 1\}.$$

The nodal interpolation operator  $\mathcal{I}_{d,r}$  on  $\mathcal{T}_{d,r}$  is for  $v \in C(\overline{\omega}_{d,r})$  defined by

$$\mathcal{I}_{d,r}v := \sum_{A \in \mathcal{N}_{d,r}} v(A)\varphi_A.$$

Here, for each  $A \in \mathcal{N}_{d,r}$  the function  $\varphi_A \in \mathcal{S}^1(\mathcal{T}_{d,r})$  satisfies  $\varphi_A(A) = 1$  and  $\varphi_A(B) = 0$  for all  $B \in \mathcal{N}_{d,r} \setminus \{A\}$ . There holds

$$\|\mathcal{T}_{d,r}g - g\|_{L^\infty(\omega_{d,r})} \leq d\sqrt{nm} |g|_{Lip(B_r(0))} \tag{3.1}$$

for locally Lipschitz continuous functions  $g : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}$ ;  $|g|_{Lip(B_r(0))}$  denotes the Lipschitz constant of  $g$  on  $B_r(0)$  and will be abbreviated by  $|g|_{Lip,r}$ .

#### 4. PROOF OF THEOREM A

This section is devoted to the proof of Theorem A. The first lemma gives an estimate for the local Lipschitz constant of  $f^{(k)}$ .

**Lemma 4.1.** *Suppose that  $f^{rc} = f$  in  $\mathbb{R}^{n \times m} \setminus B_r(0)$  and  $f$  is locally Lipschitz continuous. Then, for all  $k \geq 0$  the function  $f^{(k)}$  is Lipschitz continuous on  $B_r(0)$  with Lipschitz constant  $|f^{(k)}|_{Lip,r} \leq |f|_{Lip,3r}$ .*

*Proof.* Let  $k \geq 0$  and  $F, G \in B_r(0)$ . Without loss of generality assume  $f^{(k)}(F) \geq f^{(k)}(G)$  and let  $\varepsilon > 0$ . By assumption on  $f$  and definition of  $f^{(k)}(G)$  there exist  $A_\iota \in \overline{B_r(0)}$  and  $\theta_\iota \geq 0$ ,  $\iota = 1, \dots, 2^k$ , (which satisfy the condition  $(\mathcal{H}_{2^k})$  of [9] with  $\sum_{\iota=1}^{2^k} \theta_\iota A_\iota = G$ ) such that  $\sum_{\iota=1}^{2^k} \theta_\iota = 1$ ,  $\sum_{\iota=1}^{2^k} \theta_\iota T(A_\iota) = T(G)$ , and

$$\sum_{\iota=1}^{2^k} \theta_\iota f(A_\iota) \leq f^{(k)}(G) + \varepsilon.$$

For each  $\iota = 1, \dots, 2^k$  define  $\tilde{A}_\iota := (F - G) + A_\iota$ . Then, (since  $\tilde{A}_\iota$ ,  $\iota = 1, \dots, 2^k$ , satisfy the condition  $(\mathcal{H}_{2^k})$  of [9] with  $\sum_{\iota=1}^{2^k} \theta_\iota \tilde{A}_\iota = F$ ) there holds  $f^{(k)}(F) \leq \sum_{\iota=1}^{2^k} \theta_\iota f(\tilde{A}_\iota)$ . Moreover,  $\tilde{A}_\iota \in B_{3r}(0)$ ,  $\tilde{A}_\iota - A_\iota = F - G$ , for  $\iota = 1, \dots, 2^k$  and hence

$$\begin{aligned} |f^{(k)}(F) - f^{(k)}(G)| &= f^{(k)}(F) - f^{(k)}(G) \leq \sum_{\iota=1}^{2^k} \theta_\iota f(\tilde{A}_\iota) - f^{(k)}(G) \\ &\leq \sum_{\iota=1}^{2^k} \theta_\iota (f(\tilde{A}_\iota) - f(A_\iota)) + \varepsilon \leq |f|_{Lip,3r} |F - G| + \varepsilon. \end{aligned} \quad \square$$

**Remark 4.1.** (i) As a consequence of Carathéodory’s theorem (see *e.g.* [22]), the second minimum in (2.2) can be attained by a convex combination of two points, *i.e.* there exist  $\ell_1, \ell_2 \in \mathbb{Z}$  such that  $F + d\ell_j a_d \otimes b_d \in \overline{\omega}_{d,r}$ ,  $j = 1, 2$ , and  $\theta_{\ell_1}, \theta_{\ell_2} \geq 0$ ,  $\theta_{\ell_1} + \theta_{\ell_2} = 1$ ,  $\theta_{\ell_1} \ell_1 + \theta_{\ell_2} \ell_2 = 0$ , such that, for the minimizing  $(\theta_\ell : \ell \in \mathbb{Z})$ , there holds

$$\begin{aligned} f_{d,r}^{(k+1)}(F) &= \sum_{\ell \in \mathbb{Z}} \theta_\ell f_{d,r}^{(k)}(F + d\ell a_d \otimes b_d) \\ &= \theta_{\ell_1} f_{d,r}^{(k)}(F + d\ell_1 a_d \otimes b_d) + \theta_{\ell_2} f_{d,r}^{(k)}(F + d\ell_2 a_d \otimes b_d). \end{aligned}$$

(ii) If  $A, B \in \mathbb{R}^{n \times m}$  with  $\text{rank}(A - B) = 1$  then there exist  $a \in \mathbb{R}^n$  and  $b \in \mathbb{R}^m$  such that  $|a| = 1$  and  $A - B = a \otimes b$ .

**Proposition 4.1.** *Let  $k \geq 0$  and suppose that for all  $\varepsilon > 0$  and all  $F \in \omega_{d,r} \cap B_{r-c_1d}(0)$  there exist  $A, B \in \omega_{d,r} \cap B_{r-c_1d}(0)$ ,  $\varrho \in [0, 1]$  such that  $\text{rank}(A - B) = 1$ ,  $F = \varrho A + (1 - \varrho)B$ , and*

$$\varrho f^{(k)}(A) + (1 - \varrho) f^{(k)}(B) \leq f^{(k+1)}(F) + \varepsilon.$$

Assume that for all  $F \in \omega_{d,r} \setminus B_{r-c_1d}(0)$  there holds  $f^{(k+1)}(F) = f^{(k)}(F)$ . Then

$$\|f_{d,r}^{(k+1)} - f^{(k+1)}\|_{L^\infty(\omega_{d,r})} \leq dc_2|f|_{Lip,3r} + \|f_{d,r}^{(k)} - f^{(k)}\|_{L^\infty(\omega_{d,r})}.$$

*Proof.* Let  $F \in \mathcal{N}_{d,r}$  and set

$$\tilde{f}_{d,r}^{(k+1)}(F) := \min_{(a_d, b_d) \in \mathcal{R}_{d,r}^1} \min \left\{ \sum_{\ell \in \mathbb{Z}: F + d\ell a_d \otimes b_d \in \bar{\omega}_{d,r}} \theta_\ell f^{(k)}(F + d\ell a_d \otimes b_d) : \theta_\ell \geq 0, \sum \theta_\ell = 1, \sum \theta_\ell \ell = 0 \right\}.$$

Let  $(\tilde{a}_d, \tilde{b}_d) \in \mathcal{R}_{d,r}^1$ ,  $\tilde{\ell}_1, \tilde{\ell}_2, \tilde{\theta}_{\tilde{\ell}_1}, \tilde{\theta}_{\tilde{\ell}_2}$  be minimizing in the definition of  $\tilde{f}_{d,r}^{(k+1)}(F)$ , i.e.

$$\tilde{f}_{d,r}^{(k+1)}(F) = \tilde{\theta}_{\tilde{\ell}_1} f^{(k)}(F + d\tilde{\ell}_1 \tilde{a}_d \otimes \tilde{b}_d) + \tilde{\theta}_{\tilde{\ell}_2} f^{(k)}(F + d\tilde{\ell}_2 \tilde{a}_d \otimes \tilde{b}_d).$$

Similarly, let  $(a'_d, b'_d) \in \mathcal{R}_{d,r}^1$ ,  $\ell'_1, \ell'_2 \in \mathbb{Z}$ ,  $\theta'_{\ell'_1}, \theta'_{\ell'_2}$ , be minimizing in the definition of  $f_{d,r}^{(k+1)}(F)$ , i.e.

$$f_{d,r}^{(k+1)}(F) = \theta'_{\ell'_1} f_{d,r}^{(k)}(F + d\ell'_1 a'_d \otimes b'_d) + \theta'_{\ell'_2} f_{d,r}^{(k)}(F + d\ell'_2 a'_d \otimes b'_d).$$

Assume that  $\tilde{f}_{d,r}^{(k+1)}(F) \leq f_{d,r}^{(k+1)}(F)$ . Then there holds, since  $\tilde{a}_d, \tilde{b}_d, \tilde{\ell}_1, \tilde{\ell}_2, \tilde{\theta}_{\tilde{\ell}_1}$ , and  $\tilde{\theta}_{\tilde{\ell}_2}$  are feasible to define  $f_{d,r}^{(k+1)}(F)$ ,

$$\begin{aligned} |\tilde{f}_{d,r}^{(k+1)}(F) - f_{d,r}^{(k+1)}(F)| &= f_{d,r}^{(k+1)}(F) - \tilde{f}_{d,r}^{(k+1)}(F) \\ &\leq f_{d,r}^{(k)}(F + d\tilde{\ell}_1 \tilde{a}_d \otimes \tilde{b}_d) + \tilde{\theta}_{\tilde{\ell}_2} f_{d,r}^{(k)}(F + d\tilde{\ell}_2 \tilde{a}_d \otimes \tilde{b}_d) \\ &\quad - \tilde{\theta}_{\tilde{\ell}_1} f^{(k)}(F + d\tilde{\ell}_1 \tilde{a}_d \otimes \tilde{b}_d) - \tilde{\theta}_{\tilde{\ell}_2} f^{(k)}(F + d\tilde{\ell}_2 \tilde{a}_d \otimes \tilde{b}_d) \\ &\leq \|f_{d,r}^{(k)} - f^{(k)}\|_{L^\infty(\omega_{d,r})}. \end{aligned}$$

Conversely, if  $f_{d,r}^{(k+1)}(F) < \tilde{f}_{d,r}^{(k+1)}(F)$  there holds

$$\begin{aligned} |\tilde{f}_{d,r}^{(k+1)}(F) - f_{d,r}^{(k+1)}(F)| &= \tilde{f}_{d,r}^{(k+1)}(F) - f_{d,r}^{(k+1)}(F) \\ &\leq \theta'_{\ell'_1} f^{(k)}(F + d\ell'_1 a'_d \otimes b'_d) + \theta'_{\ell'_2} f^{(k)}(F + d\ell'_2 a'_d \otimes b'_d) \\ &\quad - \theta'_{\ell'_1} f_{d,r}^{(k)}(F + d\ell'_1 a'_d \otimes b'_d) - \theta'_{\ell'_2} f_{d,r}^{(k)}(F + d\ell'_2 a'_d \otimes b'_d) \\ &\leq \|f_{d,r}^{(k)} - f^{(k)}\|_{L^\infty(\omega_{d,r})}. \end{aligned}$$

In either case we have

$$|\tilde{f}_{d,r}^{(k+1)}(F) - f_{d,r}^{(k+1)}(F)| \leq \|f_{d,r}^{(k)} - f^{(k)}\|_{L^\infty(\omega_{d,r})}. \quad (4.1)$$

We now estimate  $|f^{(k+1)}(F) - \tilde{f}_{d,r}^{(k+1)}(F)|$ . Assume first that  $F \in \omega_{d,r} \setminus B_{r-c_1d}(0)$  so that  $f^{(k+1)}(F) = f^{(k)}(F)$ . Then, since  $f^{(k+1)}(F) \leq \tilde{f}_{d,r}^{(k+1)}(F) \leq f^{(k)}(F)$  we deduce  $|f^{(k+1)}(F) - \tilde{f}_{d,r}^{(k+1)}(F)| = 0$ . Suppose now that

$F \in \omega_{d,r} \cap B_{r-c_1d}(0)$  and set  $\varepsilon := dc_2|f^{(k)}|_{Lip,r}/2$ . Let  $a \in \mathbb{R}^n$ ,  $|a| = 1$ ,  $b \in \mathbb{R}^m$ ,  $A, B \in \omega_{d,r} \cap B_{r-c_1d}(0)$ ,  $B - A = a \otimes b$ , and  $\varrho \in [0, 1]$  be such that  $F = \varrho A + (1 - \varrho)B$  and

$$\varrho f^{(k)}(A) + (1 - \varrho)f^{(k)}(B) \leq f^{(k+1)}(F) + \varepsilon.$$

We set

$$a_d := d\lfloor a/d \rfloor, \quad b_d := d\lfloor b/d \rfloor,$$

and

$$\ell_1 := -\lfloor (1 - \varrho)/d \rfloor, \quad \ell_2 := \lfloor \varrho/d \rfloor.$$

Noting that  $A = F - (1 - \varrho)a \otimes b$  and inserting  $\ell_1 da \otimes b$  we infer

$$|A - (F + \ell_1 da_d \otimes b_d)| = |(1 - \varrho)a \otimes b + \ell_1 da_d \otimes b_d| \leq |((1 - \varrho) + \ell_1 d)a \otimes b| + |\ell_1 d(a \otimes b - a_d \otimes b_d)|.$$

The definition of  $\ell_1$  and  $|a| = 1$  prove

$$|((1 - \varrho) + \ell_1 d)a \otimes b| \leq d|b|.$$

We use  $|d\ell_1| \leq 1 + d$ , insert  $a_d \otimes b$ , and employ  $|a - a_d| \leq \sqrt{nd}$ ,  $|b - b_d| \leq \sqrt{md}$ , to verify

$$\begin{aligned} |\ell_1 d(a \otimes b - a_d \otimes b_d)| &\leq (1 + d)|(a - a_d) \otimes b| + (1 + d)|a_d \otimes (b - b_d)| \\ &\leq d(1 + d)(\sqrt{n}|b| + \sqrt{m}|a_d|). \end{aligned}$$

Since for the components  $a_{d,j}$ ,  $j = 1, \dots, n$ , of  $a_d$  and  $a_j$ ,  $j = 1, \dots, n$ , of  $a$  we have  $a_j - d \leq a_{d,j} \leq a_j$  and since  $|a_j| \leq 1$  there holds

$$1 - 2nd \leq \sum_{j=1}^n a_{d,j}^2 \leq (1 + nd)^2.$$

Since  $|b| = |a \otimes b| = |A - B| \leq 2\sqrt{nm}r$  we have  $|b_d| \leq 2\sqrt{nm}r + md$ . The combination of the previous estimates verifies

$$|A - (F + \ell_1 da_d \otimes b_d)| \leq d(2\sqrt{nm}r + (1 + d)\sqrt{m}(2nr + 1 + nd)) =: dc_3$$

and the same arguments prove

$$|B - (F + \ell_2 da_d \otimes b_d)| \leq dc_3.$$

By assumption on  $A$  and definition of  $c_1 = 1 + c_3$  we have

$$\begin{aligned} |F + \ell_1 da_d \otimes b_d|_\infty &\leq |A|_\infty + |A - (F + \ell_1 da_d \otimes b_d)| \\ &\leq r - c_1d + dc_3 = r - d \end{aligned}$$

and this implies  $F + \ell_1 da_d \otimes b_d \in \omega_{d,r}$ . Similarly, we have  $F + \ell_2 da_d \otimes b_d \in \omega_{d,r}$ .

If  $\ell_1 = \ell_2 = 0$  we set  $\theta_{\ell_1} := \varrho$  and  $\theta_{\ell_2} := (1 - \varrho)$ . Otherwise, we define  $\theta_{\ell_1} := \ell_2/(|\ell_1| + \ell_2)$  and  $\theta_{\ell_2} := 1 - \theta_{\ell_1}$ . Note that  $(1 - \varrho) \leq d|\ell_1| \leq (1 - \varrho) + d$  and  $\varrho - d \leq d\ell_2 \leq \varrho$ . If  $\varrho \leq \theta_{\ell_1}$  and if  $d \leq 1/2$  then

$$|\varrho - \theta_{\ell_1}| = \theta_{\ell_1} - \varrho = \frac{\ell_2}{|\ell_1| + \ell_2} - \varrho \leq \frac{\varrho}{1 - d} - \varrho = \frac{d}{1 - d} \leq 2d.$$

If  $d > 1/2$  we simply estimate  $|\varrho - \theta_{\ell_1}| \leq 1 \leq 2d$ . Similarly, if  $\theta_{\ell_1} \leq \varrho$  there holds

$$|\varrho - \theta_{\ell_1}| = \varrho - \theta_{\ell_1} \leq \varrho - \frac{\varrho - d}{\varrho + (1 - \varrho) + d} \leq 2d.$$

Finally, observe that  $\theta_{\ell_2} = |\ell_1|/(|\ell_1| + \ell_2)$  implies

$$\theta_{\ell_1} \ell_1 + \theta_{\ell_2} \ell_2 = 0.$$

Noting that  $f^{(k+1)}(F) \leq \tilde{f}_{d,r}^{(k+1)}(F)$  and that  $a_d, b_d, \ell_1, \ell_2, \theta_{\ell_1}, \theta_{\ell_2}$  are feasible in the definition of  $\tilde{f}_{d,r}^{(k+1)}(F)$  we observe

$$\begin{aligned} |f^{(k+1)}(F) - \tilde{f}_{d,r}^{(k+1)}(F)| &= \tilde{f}_{d,r}^{(k+1)}(F) - f^{(k+1)}(F) \\ &\leq \theta_{\ell_1} f^{(k)}(F + \ell_1 da_d \otimes b_d) + \theta_{\ell_2} f^{(k)}(F + \ell_2 da_d \otimes b_d) \\ &\quad - \varrho f^{(k)}(A) - (1 - \varrho) f^{(k)}(B) + \varepsilon. \end{aligned}$$

Inserting  $\theta_{\ell_1} f^{(k)}(A)$  and  $\theta_{\ell_2} f^{(k)}(B)$ , using  $\theta_{\ell_2} - (1 - \varrho) = -(\theta_{\ell_1} - \varrho)$ , and employing  $|B - A| \leq 2\sqrt{nm}r$  we verify

$$\begin{aligned} |f^{(k+1)}(F) - \tilde{f}_{d,r}^{(k+1)}(F)| &\leq \theta_{\ell_1} (f^{(k)}(F + \ell_1 da_d \otimes b_d) - f^{(k)}(A)) \\ &\quad + \theta_{\ell_2} (f^{(k)}(F + \ell_2 da_d \otimes b_d) - f^{(k)}(B)) \\ &\quad + (\theta_{\ell_1} - \varrho) (f^{(k)}(A) - f^{(k)}(B)) + \varepsilon \\ &\leq dc_3 |f^{(k)}|_{Lip,r} + 2d 2\sqrt{nm}r \frac{|f^{(k)}(A) - f^{(k)}(B)|}{|A - B|} + \varepsilon \\ &\leq d(c_3 + 4\sqrt{nm}r) |f^{(k)}|_{Lip,r} + \varepsilon. \end{aligned}$$

The triangle inequality, the choice of  $\varepsilon$ , (4.1), and Lemma 4.1 conclude the proof.  $\square$

The following lemma is due to [11].

**Lemma 4.2.** *Suppose that  $f^{rc} = f$  in  $\mathbb{R}^{n \times m} \setminus B_{r-c_1 d}(0)$ . Then the conditions of Proposition 4.1 are satisfied.  $\square$*

The next proposition shows that  $f_{d,r}^{(k)}$  is a reliable upper bound for  $f^{rc}$ .

**Proposition 4.2.** *For all  $k \geq 0$ ,  $r \geq d > 0$ , and all  $G \in \bar{\omega}_{d,r}$  there holds  $f_{d,r}^{(k)}(G) \geq f^{rc}(G)$ .*

*Proof.* The proof follows from the definition of  $f_{d,r}^{(k)}(G)$ , the fact that nodal basis functions (on symmetric grids) evaluated at some  $F \in Q \in \mathcal{T}_{d,r}$  define a rank-1 decomposition of  $F$  (see [3] Lem. 3.1), and the characterization of  $f^{rc}$  through the condition  $(\mathcal{H}_N)$  of [9].  $\square$

*Proof of Theorem A.* The first assertion is Proposition 4.2. The error estimate follows from Lemma 4.2 and an induction over  $\ell = 0, \dots, L$  with Proposition 4.1.  $\square$

## 5. NUMERICAL EXPERIMENTS

In this section we report on the practical performance of the following Algorithm ( $A_{r,d}^{rc}$ ) which realizes (2.2). Input are the parameters  $r \geq d > 0$  and the function  $f$ .

**Algorithm** ( $A_{r,d}^{rc}$ ).

- (a) Set  $f_{d,r}^{(0)} := \mathcal{I}_{d,r}f$  and  $k := 1$ .
- (b) Set  $f_{d,r}^{(k)} := f_{d,r}^{(k-1)}$ ,  $\mathcal{R} := \mathcal{R}_{d,r}^1$ , and  $\mathcal{N} := \mathcal{N}_{d,r}$ .
- (c) Choose  $F \in \mathcal{N}$  and set  $\mathcal{N} := \mathcal{N} \setminus \{F\}$ .
- (d) Choose  $(a_d, b_d) \in \mathcal{R}$  and set  $\mathcal{R} := \mathcal{R} \setminus \{(a_d, b_d)\}$ .
- (e) Solve the linear optimization problem

$$m := \min \left\{ \sum_{\ell \in \mathbb{Z}: F + d\ell a_d \otimes b_d \in \overline{\omega}_{d,r}} \theta_\ell f_{d,r}^{(k-1)}(F + d\ell a_d \otimes b_d) : \theta_\ell \geq 0, \sum_\ell \theta_\ell = 1, \sum_\ell \theta_\ell \ell = 0 \right\}$$

and set  $f_{d,r}^{(k)}(F) := \min\{m, f_{d,r}^{(k)}(F)\}$ .

- (f) If  $\mathcal{R} \neq \emptyset$  go to (d).
- (g) If  $\mathcal{N} \neq \emptyset$  go to (c).
- (h) If  $\|f_{d,r}^{(k)} - f_{d,r}^{(k-1)}\|_{L^\infty(B_r(0))} > d$  set  $k := k + 1$  and go to (b).
- (j) Stop.

**Remark 5.1.** If we are only interested in an upper bound for  $f^{rc}$  and hence for  $f^{qc}$  in  $\omega_{d,r}$  we may stop the iteration in Step (h) of the algorithm if  $k = M$  for some given  $M \geq 0$ .

**Example 5.1** [9]. For  $n = m = 2$  and  $F \in \mathbb{R}^{2 \times 2}$  let

$$f(F) := (|F|^2 - 1)^2.$$

For  $F \in \mathbb{R}^{2 \times 2}$  there holds

$$f^{rc}(F) = \begin{cases} (|F|^2 - 1)^2 & \text{for } |F| \geq 1, \\ 0 & \text{for } |F| \leq 1. \end{cases}$$

**Example 5.2** [12, 15]. For  $n = m = 2$ ,

$$A_1 := \begin{pmatrix} 5/4 & 0 \\ 0 & 3/4 \end{pmatrix} \quad \text{and} \quad A_2 := \begin{pmatrix} 3\sqrt{8}/8 & 3/8 \\ -5/8 & 5\sqrt{3}/8 \end{pmatrix}$$

and  $F \in \mathbb{R}^{2 \times 2}$  let

$$f(F) := \frac{1}{2} \min\{|F - A_1|^2, |F - A_2|^2\}.$$

Then,  $f^{rc}$  is for  $F \in \mathbb{R}^{2 \times 2}$  given by

$$f^{rc}(F) = \begin{cases} f_1(F) & \text{for } f_1(F) - f_2(F) \leq -\lambda/2, \\ f_2(F) - (f_2(F) - f_1(F) + \lambda/2)/(2\lambda) & \text{for } |f_1(F) - f_2(F)| \leq \lambda/2, \\ f_2(F), & \text{for } f_1(F) - f_2(F) \geq \lambda/2, \end{cases}$$

where  $f_j(F) = |F - A_j|^2/2$ ,  $j = 1, 2$ , and  $\lambda = |A_1 - A_2|$ .

**Example 5.3** [11, 16]. For  $n = m = 2$  and  $F \in \mathbb{R}^{2 \times 2}$  let

$$f(F) := \begin{cases} 1 + |F|^2 & \text{for } |F| \geq \sqrt{2} - 1, \\ 2\sqrt{2}|F| & \text{for } |F| \leq \sqrt{2} - 1. \end{cases}$$



Letting  $\varrho(F) := \sqrt{|F|^2 + 2|\det F|}$  for  $F \in \mathbb{R}^{2 \times 2}$  there holds

$$f^{rc}(F) = \begin{cases} 1 + |F|^2 & \text{for } \varrho(F) \geq 1, \\ 2(\varrho(F) - |\det F|) & \text{for } \varrho(F) \leq 1. \end{cases}$$

Linear convergence has been observed in [12] for Examples 5.2 and 5.3 for a related algorithm with a smaller set of discrete rank-one connections, *i.e.* with a small set  $\mathcal{R}'$  (which may be regarded as a subset of  $\mathcal{R}_{d,r}^1$ ) instead of  $\mathcal{R}_{d,r}^1$  in Algorithm  $(A_{r,d}^{rc})$ . For the parameters  $r = 2$  and  $d = 1, 1/2, 1/4$  we obtained a maximal error in the nodes  $\mathcal{N}_{d,r} \cap B_{r'}(0)$  for  $r' = 2, 1, 2$  in Examples 5.1, 5.2, 5.3, respectively, as displayed in Table 1. The table also displays the final iteration level  $k$ , *i.e.* the smallest positive integer  $k$  for which  $\|f_{d,r}^{(k)} - f_{d,r}^{(k+1)}\|_{L^\infty(B_r(0))} \leq d$ . Note that  $f = f^{rc}$  in  $\mathbb{R}^{2 \times 2} \setminus B_2(0)$  in Examples 5.1 and 5.3 so that we used  $r' = 2$  in those examples. The condition  $f = f^{rc}$  in the complement of some compact set is not satisfied in Example 5.2. The argumentation in [12] shows however that we can expect convergence in  $B_1(0)$  so that we chose  $r' = 1$  to compute an error in that example. The displayed error  $e$  is part of an upper bound for  $\|f_{d,r}^{(k)} - f^{rc}\|_{L^\infty(B_{r'}(0))}$  since

$$\begin{aligned} \|f_{d,r}^{(k)} - f^{rc}\|_{L^\infty(B_{r'}(0))} &\leq e + \|f^{rc} - \mathcal{I}_{d,r} f^{rc}\|_{L^\infty(B_{r'}(0))} \\ &\leq e + d\sqrt{nm} \|f^{rc}\|_{Lip,r} \end{aligned}$$

and the sum on the right-hand side converges linearly in all examples. We did not run the algorithm for smaller values of  $d$  since the expected CPU time (for our implementation in C) was about 1000 hours for  $d = 1/8$  in all examples.

TABLE 1. Error  $e = \max_{F \in \mathcal{N}_{d,r'}} |f_{d,r}^{(k)}(F) - f^{rc}(F)|$  in Examples 5.1, 5.2, and 5.3 for  $r' = 2, 1, 2$ , respectively, and termination level  $k$  for  $d = 1, 1/2, 1/4$  and  $r = 2$ .

$d$	$e$ and $k$ in Ex. 5.1	$e$ and $k$ in Ex. 5.2	$e$ and $k$ in Ex. 5.3
1	0.000 001 ( $k = 2$ )	0.125 150 ( $k = 1$ )	0.000 001 ( $k = 1$ )
1/2	0.000 001 ( $k = 2$ )	0.075 128 ( $k = 1$ )	0.085 786 ( $k = 1$ )
1/4	0.003 906 ( $k = 2$ )	0.022 229 ( $k = 1$ )	0.043 861 ( $k = 3$ )

## 6. NUMERICAL STUDY OF A 2D ERICKSON-JAMES ENERGY DENSITY

A reliable and efficient algorithm that approximates a lower bound for  $f^{qc}$ , *i.e.* the polyconvex envelope  $f^{pc}$  (see *e.g.* [9]), has recently been designed and analyzed in [3]. In combination with the results of this article one can therefore numerically check for equality of  $f^{rc}$  and  $f^{pc}$  in order to characterize  $f^{qc}$ . We numerically study a two-dimensional version of the Erickson-James energy density [13].

**Example 6.1** [21]. Given parameters  $\delta, \gamma > 0$ ,  $k_1, k_2, k_3 > 0$ , and letting

$$C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} := F^T F$$

for  $F \in \mathbb{R}^{2 \times 2}$ , define

$$f(F) := k_1(C_{11} + C_{22} - \delta - \gamma)^2 + k_2 C_{12}^2 + k_3(C_{11} - \delta)^2(C_{22} - \delta)^2.$$

For the numerical experiment we set  $k_1 := 1$ ,  $k_2 := 0.3$ ,  $k_3 := 1$ , and  $\delta := 1.1$ ,  $\gamma := 0.9$ .

We employed the algorithm of [3] to compute  $f_{1/8,1/2}^{pc}$  as an approximation of  $f^{pc}$  in the nodes of  $\mathcal{N}_{1/4,1/2}$  with an accuracy  $\mathcal{O}(1/8^2)$ . Moreover, we used algorithm  $(A_{r,d}^{rc})$  with  $r = 2$  and  $d = 1/4$  to obtain an approximation of  $f^{rc}$  with an error of order  $\mathcal{O}(1/4)$  on  $B_{1/2}(0)$  (presuming that the conditions of Theorem A are satisfied). Thereby, we found

$$\max_{F \in \mathcal{N}_{1/4,1/2}} |f_{1/8,1/2}^{pc}(F) - f_{1/4,2}^{(4)}(F)| \leq 0.000897.$$

From this very small distance we may conjecture that  $f^{pc} = f^{rc}$  in  $B_{1/2}(0)$  for  $f$  as in Example 6.1. For a final conclusion one would however have to consider smaller discretization parameters.

*Acknowledgments.* The author gratefully acknowledges support by the DFG through the priority program 1095 “Analysis, Modeling and Simulation of Multiscale Problems”.

## REFERENCES

- [1] J.M. Ball, A version of the fundamental theorem for Young measures. Partial differential equations and continuum models of phase transitions. M. Rascle, D. Serre, M. Slemrod Eds. *Lect. Notes Phys.* **344** (1989) 207–215.
- [2] J.M. Ball and R.D. James, Fine phase mixtures as minimizers of energy. *Arch. Rational Mech. Anal.* **100** (1987) 13–52.
- [3] S. Bartels, Reliable and efficient approximation of polyconvex envelopes. *SIAM J. Numer. Anal.* (accepted) [Preprints of the DFG Priority Program “Multiscale Problems”, No. 76 (2002) ([www.mathematik.uni-stuttgart.de/~mehrskalen/](http://www.mathematik.uni-stuttgart.de/~mehrskalen/))].
- [4] S. Bartels, Error estimates for adaptive Young measure approximation in scalar nonconvex variational problems. *SIAM J. Numer. Anal.* **42** (2004) 505–529.
- [5] S. Bartels and A. Prohl, Multiscale resolution in the computation of crystalline microstructure. *Numer. Math.* **96** (2004) 641–660.
- [6] C. Carstensen and P. Plecháč, Numerical solution of the scalar double-well problem allowing microstructure. *Math. Comp.* **66** (1997) 997–1026.
- [7] C. Carstensen and T. Roubíček, Numerical approximation of Young measures in non-convex variational problems. *Numer. Math.* **84** (2000) 395–414.
- [8] M. Chipot and S. Müller, Sharp energy estimates for finite element approximations of non-convex problems, in Variations of domain and free boundary problems, in solid mechanics, *Solid Mech. Appl.* **66** (1997) 317–327.
- [9] B. Dacorogna, Direct methods in the calculus of variations. *Appl. Math. Sci.* **78** (1989).
- [10] B. Dacorogna and J.-P. Haeberly, Some numerical methods for the study of the convexity notions arising in the calculus of variations. *RAIRO Modél. Math. Anal. Numér.* **32** (1998) 153–175.
- [11] G. Dolzmann, Numerical computation of rank-one convex envelopes. *SIAM J. Numer. Anal.* **36** (1999) 1621–1635.
- [12] G. Dolzmann and N.J. Walkington, Estimates for numerical approximations of rank one convex envelopes. *Numer. Math.* **85** (2000) 647–663.
- [13] J.L. Ericksen, Constitutive theory for some constrained elastic crystals. *Int. J. Solids Struct.* **22** (1986) 951–964.
- [14] K. Hackl and U. Hoppe, On the calculation of microstructures for inelastic materials using relaxed energies. IUTAM symposium on computational mechanics of solid materials at large strains, C. Miehe Ed., *Solid Mech. Appl.* **108** (2003) 77–86.
- [15] R.V. Kohn, The relaxation of a double-well energy. *Contin. Mech. Thermodyn.* **3** (1991) 193–236.
- [16] R.V. Kohn and G. Strang, Optimal design and relaxation of variational problems. I-III. *Commun. Pure Appl. Math.* **39** (1986) 353–377.
- [17] M. Kružík, Numerical approach to double well problems. *SIAM J. Numer. Anal.* **35** (1998) 1833–1849.
- [18] M. Luskin, On the computation of crystalline microstructure. *Acta Numerica* **5** (1996) 191–257.
- [19] C. Miehe and M. Lambrecht, Analysis of micro-structure development in shearbands by energy relaxation of incremental stress potentials: large-strain theory for standard dissipative materials. *Internat. J. Numer. Methods Engrg.* **58** (2003) 1–41.
- [20] S. Müller, Variational models for microstructure and phase transitions. *Lect. Notes Math.* **1713** (1999) 85–210.
- [21] R.A. Nicolaides, N. Walkington and H. Wang, Numerical methods for a nonconvex optimization problem modeling martensitic microstructure. *SIAM J. Sci. Comput.* **18** (1997) 1122–1141.
- [22] T. Roubíček, *Relaxation in optimization theory and variational calculus*. De Gruyter Series in Nonlinear Analysis Appl. **4** New York (1997).