RELIABLE AND EFFICIENT APPROXIMATION OF POLYCONVEX ENVELOPES

SÖREN BARTELS

Abstract. An iterative algorithm that approximates the polyconvex envelope f^{pc} of a given function $f : \mathbb{R}^{n\times m} \to \mathbb{R}$, i.e. the largest function below f which is convex in all minors, is established. A rigorous error analysis with a focus on reliability and optimal orders of convergence, an efficient strategy that reduces the large number of unknowns, as well as numerical experiments are presented.

1. Introduction

A non-convex variational problem due to [BJ] modeling phase transitions in crystalline solids and allowing for microstructure reads

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

for a bounded Lipschitz domain $\Omega \subseteq \mathbb{R}^n$, $p \geq 1$, a (non-convex) continuous energy density $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{n\times m} \to \mathbb{R}$ satisfying p-growth conditions, and a space of admissible deformations $\mathcal{A} \subseteq W^{1,p}(\Omega;\mathbb{R}^m)$ containing boundary conditions. Since I may not be weakly lower semicontinuous, minimizing sequences develop oscillations in the gradient variable and their weak limits do in general not minimize I (see e.g. [Da2, M, R2]). Together with a Young measure generated by a minimizing sequence in the sense of [B2], 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. [Da2, M, R2]). 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 a reliable and efficient algorithm that computes a lower bound. Numerical schemes for the approximation of upper bounds can be found in [Do, DW, Ba2].

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. [L, CM, BP]) or f^{qc} is convex (see e.g. [NW2, CP1, CR, Ba1]). 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-dependent oscillations, those approaches cannot be expected to lead to efficient numerical algorithms. In the second case efficient algorithms are available but the

Date: May 5, 2004.

1991 *Mathematics Subject Classification.* 65N12 (65N15 65N30).

Key words and phrases. Nonconvex variational problem, calculus of variations, quasiconvexity, relaxed variational problems, microstructure, adaptive algorithm.

proposed numerical schemes are restricted to scalar problems. An algorithm that checks for different notions of convexity for a class of functions can be found in [DH].

By computing a (polyconvex) Young measure solution for (M) with affine boundary conditions, our iterative algorithm approximates the polyconvex envelope [B1, Da1] f^{pc} of f as a lower bound for f^{qc}. A straightforward discretization linearizes non-linear constraints and results in a large but linear optimization problem. We show that for a large class of functions f the approximation is very accurate. The efficient iterative strategy for the solution of the linear optimization problem is based on results in [R1] that state sharp estimates on the support of a (polyconvex) Young measure solution for (M). Moreover, the strategy employs and generalizes a multilevel scheme of [CR] for the approximation of scalar non-convex variational problems.

The proposed algorithm can be employed for the simultaneous (polyconvex) relaxation and approximation of non-convex variational problems. This approach results in discrete problems with two numerical scales that reflect microscopic and macroscopic effects. We refer to [NW1, HH, ML, Kr, Ba1, Do, DW] for related numerical experiments. Moreover, in combination with the algorithms of [Do, DW, Ba2] for the approximation of an upper bound, the results of this paper allow to numerically check for equality of polyconvex and rank-1 convex envelopes.

The rest of the paper is organized as follows. We present the approximation scheme with an error estimate in Section 2. Some preliminaries in Section 3 lead to the proof of the main result which is given in Section 4. Section 5 is devoted to a reliable and efficient algorithm that realizes the approximation scheme. Numerical experiments that illustrate the high efficiency and accuracy of the proposed algorithm are reported on in Section 6. Section 7 discusses the effective numerical solution of (M) based on the approximation of polyconvex envelopes.

2. Approximation Scheme and Main Results

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

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

The polyconvex envelope f^{pc} of f is for $F \in \mathbb{R}^{n\times m}$ given by [B1, Da1]

$$f^{pc}(F) = \inf\Big\{\sum_{\ell=1}^{\tau+1} \varrho_\ell f(A_\ell) : A_\ell \in \mathbb{R}^{n\times m},\ \varrho_\ell \geq 0,\ \sum_{\ell=1}^{\tau+1} \varrho_\ell = 1,\ \sum_{\ell=1}^{\tau+1} \varrho_\ell T(A_\ell) = T(F)\Big\}.$$

Here, $T(A) \in \mathbb{R}^\tau$ is a vector containing all minors of the matrix $A \in \mathbb{R}^{n\times m}$ in a fixed order and τ denotes its length; there holds $|T(A)| \leq c_T |A|_\infty^{\min\{n,m\}}$ if $|\cdot|_\infty$ denotes the maximum norm and $|A|_\infty \geq 1$. Choosing a set of points $\mathcal{N}_{d,r} := d\mathbb{Z}^{n\times m} \cap \overline{B_r(0)}$ for $r \geq d > 0$ and $B_r(0) := \{A \in \mathbb{R}^{n\times m} : |A|_\infty < r\}$ such that $F \in \operatorname{conv} \mathcal{N}_{d,r}$, an approximation of $f^{pc}(F)$ reads

$$f^{pc}_{d,r}(F) := \inf\Big\{\sum_{A\in\mathcal{N}_{d,r}} \theta_A f(A) : \forall A \in \mathcal{N}_{d,r},\ \theta_A \geq 0,\ \sum_{A\in\mathcal{N}_{d,r}} \theta_A = 1,\ \sum_{A\in\mathcal{N}_{d,r}} \theta_A T(A) = T(F)\Big\}.$$

The latter infimum defines a finite-dimensional *linear* optimization problem and admits a solution and a Lagrange multiplier $\lambda_{d,r}^F \in \mathbb{R}^\tau$ associated to the constraint $\sum_{A\in\mathcal{N}_{d,r}} \theta_A T(A) = T(F)$. Our main results concerning the approximation of polyconvex envelopes are summarized in Theorem A. We refer to Section 4 for more general assertions and to [BKK] for conditions that ensure $f^{pc} \in C^{1,\alpha}_{loc}(\mathbb{R}^{n\times m})$ together with explicit bounds on $|f^{pc}|_{C^{1,\alpha}(B_d(F))}$.

Theorem A. *Suppose that $F \in \operatorname{conv}\mathcal{N}_{d,r}$, $p \geq \min\{n,m\} =: n\wedge m$, $r \geq 1$, the computable a posteriori condition*

$$c_T(n\wedge m)|\lambda_{d,r}^F| \leq pc_f r^{p-n\wedge m} \quad \text{and} \quad c_T|\lambda_{d,r}^F| r^{n\wedge m} - c_f r^p + c_f' \leq \lambda_{d,r}^F \cdot T(F) - f_{d,r}^{pc}(F)$$

is satisfied, and $f \in C^{1,\alpha}_{loc}(\mathbb{R}^{n\times m})$ for some $\alpha \in [0,1]$. Then $f_{d,r}^{pc}(F) = \tilde f_d^{pc}(F)$ for a polyconvex function $\tilde f_d^{pc} : \mathbb{R}^{n\times m} \to \mathbb{R}$, $f_{d,r}^{pc}(F) = f_{d,s}^{pc}(F)$ for all $s \geq r$, and there exists $r' \geq r$ such that

$$|f_{d,r}^{pc}(F) - f^{pc}(F)| \leq c_1\, d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))}.$$

If, additionally, $\alpha > 0$ and $f^{pc} \in C^{1,\alpha}_{loc}(\mathbb{R}^{n\times m})$ then

$$|\lambda_{d,r}^F \cdot DT(F) - Df^{pc}(F)| \leq c_2\, d^\alpha \big(|f|_{C^{1,\alpha}(B_{r'}(0))} + |f^{pc}|_{C^{1,\alpha}(B_d(F))}\big).$$

The constants $c_1, c_2 > 0$ only depend on n and m.

It can be shown that $\lambda_{d,r}^F$ and $f_{d,r}^{pc}(F)$ remain bounded for $r \to \infty$ so that the a posteriori condition of the theorem is satisfied if $p > n\wedge m$ and if r is large enough. The direct computation of $f_{d,r}^{pc}(F)$ requires the solution of a linear optimization problem with $(r/d)^{nm}$ unknowns and would therefore be very expensive. The combination of an active set strategy (due to [CR] for $\min\{n,m\} = 1$) in combination with local grid refinement and coarsening to avoid to check a maximum principle in all nodes of $\mathcal{N}_{d,r}$ leads to a very efficient but still reliable iterative algorithm that computes $f_{d,r}^{pc}(F)$.

3. Preliminaries

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

$$|A|^2 = \sum_{j=1}^n \sum_{k=1}^m (A)_{j,k}^2.$$

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

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 (cf. the left plot in Figure 1)

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

$\mathbb{Z}$ denotes the set of all integers. 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 (cf. the left plot in Figure 1)

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

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

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

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 exists $c_{\mathcal{I}} > 0$ such that

$$\|\mathcal{I}_{d,r}g - g\|_{L^\infty(\omega_{d,r})} \leq c_{\mathcal{I}} d^{1+\alpha} |g|_{C^{1,\alpha}(B_r(0))} \tag{3.1}$$

for $\alpha \in (0,1]$ and $g \in C^{1,\alpha}_{loc}(\mathbb{R}^{n\times m})$ or for $\alpha = 0$ and a locally Lipschitz continuous function $g : \mathbb{R}^{n\times m} \to \mathbb{R}$; $|g|_{C^{1,\alpha}(B_r(0))}$ denotes the α-Hölder constant of Dg on $B_r(0)$ if $\alpha > 0$, i.e.

$$|g|_{C^{1,\alpha}(B_r(0))} := \sup_{G,H \in B_r(0)} \frac{|Dg(G) - Dg(H)|}{|G-H|^\alpha},$$

and the Lipschitz constant $|g|_{C^{1,0}(B_r(0))} := |g|_{Lip,r} = |g|_{Lip(B_r(0))}$ of g on $B_r(0)$ if $\alpha = 0$.

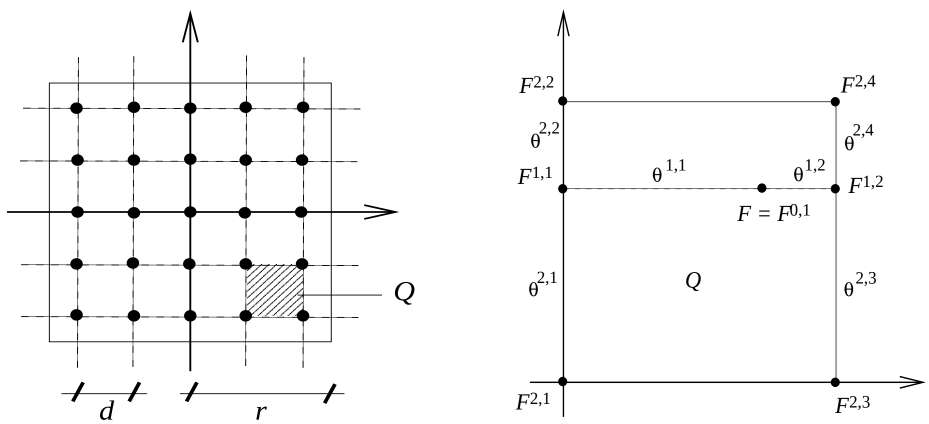

FIGURE 1. Set of nodes $\mathcal{N}_{d,r}$ (filled circles) and $Q \in \mathcal{T}_{d,r}$ (shaded small box) (left). Decomposition of the matrix F in the proof of Lemma 3.1 for $n = 2$ and $m = 1$ (right).

The operator $T : \mathbb{R}^{n\times m} \to \mathbb{R}^\tau$ is for $A \in \mathbb{R}^{n\times m}$ defined by

$$T(A) = \big((A)_{1,1}, ..., (A)_{1,m}, (A)_{2,1}, ..., (A)_{2,m}, ..., (A)_{n,1}, ..., (A)_{n,m}, \operatorname{adj}_2 A, ..., \operatorname{adj}_{n\wedge m} A\big),$$

where for $2 \leq \ell \leq n \wedge m = \min\{n,m\}$, $\operatorname{adj}_\ell A$ is a vector containing all $\ell \times \ell$ minors of A and

$$\tau := \tau(n,m) = \sum_{\ell=1}^{n\wedge m} \sigma(\ell) \quad \text{for} \quad \sigma(\ell) = \frac{m!\,n!}{(\ell!)^2(m-\ell)!(n-\ell)!}.$$

There exists $c_T > 0$ (which depends on n and m) such that $|T(A)| \leq c_T |A|_\infty^{n\wedge m}$ for all $A \in \mathbb{R}^{n\times m}$ with $|A|_\infty \geq 1$; for $n = m = 2$ we have $\operatorname{adj}_2 A = \det A$ and we can choose $c_T = 2\sqrt{2}$.

The following observation is of central importance in our analysis. It shows that the values of the nodal basis functions in $\mathcal{S}^1(\mathcal{T}_{d,r})$ define a rank-1 decomposition of a matrix $F \in \overline{\omega}_{d,r}$.

Lemma 3.1. *Let $Q \in \mathcal{T}_{d,r}$ and $M_1, ..., M_{2^{nm}} \in \mathcal{N}_{d,r}$ be such that $Q = \operatorname{conv}\{M_1, ..., M_{2^{nm}}\}$. Let $F \in Q$ and define $\theta_\iota := \varphi_{M_\iota}(F) \geq 0$, $\iota = 1, ..., 2^{nm}$. There holds*

$$\sum_{\iota=1}^{2^{nm}} \theta_\iota = 1 \quad \textit{and} \quad \sum_{\iota=1}^{2^{nm}} \theta_\iota T(M_\iota) = T(F). \tag{3.2}$$

Proof. We construct convex-coefficients that satisfy (3.2) and then show that they equal $\varphi_{M_\iota}(F)$. Suppose first that $d = 1$ and $Q = [0,1]^{n\times m}$ is the unit cube in $\mathbb{R}^{n\times m}$ so that $\{M_1, ..., M_{2^{nm}}\} = \{0,1\}^{n\times m}$ are the vertices of Q. Set $F^{0,1} := F$ and $\varrho^{0,1} := 1$. Then, for $j = 1, ..., n$ and $k = 1, ..., m$ set $\ell := (j-1)m + k$ and define $F^{\ell,2\iota-1}, F^{\ell,2\iota} \in \mathbb{R}^{n\times m}$, $\iota = 1, ..., 2^{\ell-1}$, by setting, for $j' = 1, ..., n$ and $k' = 1, ..., m$,

$$\left(F^{\ell,2\iota-1}\right)_{j',k'} := \begin{cases} \left(F^{\ell-1,\iota}\right)_{j',k'} & \text{for } (j',k') \neq (j,k), \\ 0 & \text{for } (j',k') = (j,k), \end{cases}$$

$$\left(F^{\ell,2\iota}\right)_{j',k'} := \begin{cases} \left(F^{\ell-1,\iota}\right)_{j',k'} & \text{for } (j',k') \neq (j,k), \\ 1 & \text{for } (j',k') = (j,k). \end{cases}$$

Moreover, set $\theta^{\ell,2\iota-1} := 1 - \left(F^{\ell-1,\iota}\right)_{j,k}$ and $\theta^{\ell,2\iota} := \left(F^{\ell-1,\iota}\right)_{j,k}$ and

$$\varrho^{\ell,2\iota-1} := \varrho^{\ell-1,\iota}\big(1 - \left(F^{\ell-1,\iota}\right)_{j,k}\big), \qquad \varrho^{\ell,2\iota} := \varrho^{\ell-1,\iota}\left(F^{\ell-1,\iota}\right)_{j,k}. \tag{3.3}$$

(The right plot in Figure 1 schematically displays the decomposition for $n = 2$ and $m = 1$.) The decomposition of F has the following properties:
(i) $\{F^{nm,\iota} : \iota = 1, ..., 2^{nm}\} = \{0,1\}^{n\times m}$;
(ii) $\theta^{\ell,2\iota-1}, \theta^{\ell,2\iota} \geq 0$, $\theta^{\ell,2\iota-1} + \theta^{\ell,2\iota} = 1$, and $F^{\ell-1,\iota} = \theta^{\ell,2\iota-1}F^{\ell,2\iota-1} + \theta^{\ell,2\iota}F^{\ell,2\iota}$ for $\ell = 1, ..., nm$ and $\iota = 1, ..., 2^{\ell-1}$;
(iii) $\operatorname{rank}(F^{\ell,2\iota-1} - F^{\ell,2\iota}) = 1$ for $\ell = 1, ..., nm$ and $\iota = 1, ..., 2^{\ell-1}$;
(iv) $\varphi_{F^{nm,\iota}}(F) = \varrho^{nm,\iota}$ for $\iota = 1, ..., 2^{nm}$.

The proofs of (i)-(iii) follow directly from the decomposition. To verify (iv) we note that according to (3.3), each $\varrho^{nm,\iota}$, $\iota = 1, ..., 2^{nm}$, defines a polynomial in F of partial degree ≤ 1. Moreover, if $F \in \{0,1\}^{n\times m}$ then $F = F^{nm,\iota}$ for some $\iota \in \{1, ..., 2^{nm}\}$ and by construction we then have $\varrho^{nm,\iota} = 1$ and $\varrho^{nm,\iota'} = 0$ for $\iota' \in \{1, ..., 2^{nm}\} \setminus \{\iota\}$. This proves (iv).

Set $\theta_\iota := \varrho^{nm,\iota}$ for $\iota = 1, ..., 2^{nm}$. The assertion of the lemma (for $Q = \operatorname{conv}\{0,1\}^{n\times m}$) follows from an induction over $\ell = 1, ..., 2^{nm}$ with (i)-(iv) and the fact that T is affine along rank-1 connections. The case $Q \neq [0,1]^{n\times m}$ follows with a dilation and a translation from the special case. ☐

4. Proof of Theorem A

This section is devoted to the proof of Theorem A which follows from several propositions that state more general results. The first proposition is a partial version of Theorem A but does not state sufficient conditions for an efficient choice of r. Throughout this section we consider a fixed $F \in \mathbb{R}^{n\times m}$ and assume that either $\alpha = 0$ and f is locally Lipschitz continuous or $\alpha \in (0,1]$ and $f \in C^{1,\alpha}_{loc}(\mathbb{R}^{n\times m})$.

Proposition 4.1. *There exists $r' = r'(F) > 0$ such that*
$$|f^{pc}(F) - f^{pc}_{d,r'}(F)| \le 2c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))}.$$

Proof. Let $t > 0$ be such that $|f|_{C^{1,\alpha}(B_t(0))} > 0$. By definition of $f^{pc}(F)$ there exist $\varrho_\ell \ge 0$ and $A_\ell \in \mathbb{R}^{n\times m}$, $\ell = 1, ..., \tau+1$, such that $\sum_{\ell=1}^{\tau+1} \varrho_\ell = 1$, $\sum_{\ell=1}^{\tau+1} \varrho_\ell T(A_\ell) = T(F)$, and

$$f^{pc}(F) \le \sum_{\ell=1}^{\tau+1} \varrho_\ell f(A_\ell) \le f^{pc}(F) + c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_t(0))}. \tag{4.1}$$

Choose $r' \ge t$ such that $A_1, ..., A_{\tau+1} \in \omega_{d,r'}$. For each $\ell = 1, ..., \tau+1$ Lemma 3.1 guarantees the existence of $M_\iota^{(\ell)} \in \mathcal{N}_{d,r'}$ and $\theta_{M_\iota^{(\ell)}} = \varphi_{M_\iota^{(\ell)}}(A_\ell) \ge 0$, $\iota = 1, ..., 2^{nm}$, such that $\sum_{\iota=1}^{2^{nm}} \theta_{M_\iota^{(\ell)}} = 1$, $\sum_{\iota=1}^{2^{nm}} \theta_{M_\iota^{(\ell)}} T(M_\iota^{(\ell)}) = T(A_\ell)$, and $A_\ell \in Q_\ell = \operatorname{conv}\{M_1^{(\ell)}, ..., M_{2^{nm}}^{(\ell)}\} \in \mathcal{T}_{d,r'}$. For $A \in \mathcal{N}_{d,r'}$ and $B \in \mathbb{R}^{n\times m}$ let $\chi_A(B) = 1$ if $A = B$ and $\chi_A(B) = 0$ otherwise. Setting for each $A \in \mathcal{N}_{d,r'}$

$$\tilde{\theta}_A = \sum_{\ell=1}^{\tau+1} \sum_{\iota=1}^{2^{nm}} \varrho_\ell \, \theta_{M_\iota^{(\ell)}} \chi_A(M_\iota^{(\ell)}), \tag{4.2}$$

defines a feasible $(\tilde{\theta}_A : A \in \mathcal{N}_{d,r'})$ to compute $f^{pc}_{d,r'}(F)$ so that

$$f^{pc}(F) \le f^{pc}_{d,r'}(F) = \sum_{A \in \mathcal{N}_{d,r'}} \theta_A f(A) \le \sum_{A \in \mathcal{N}_{d,r'}} \tilde{\theta}_A f(A), \tag{4.3}$$

where $(\theta_A : A \in \mathcal{N}_{d,r'})$ is optimal in $f^{pc}_{d,r'}(F)$. Since $|f|_{C^{1,\alpha}(B_t(0))} \le |f|_{C^{1,\alpha}(B_{r'}(0))}$ and $\theta_{M_\iota^{(\ell)}} = \varphi_{M_\iota^{(\ell)}}(A_\ell)$, estimates (3.1) and (4.1)-(4.3) imply

$$\begin{aligned}
0 &\le f^{pc}_{d,r'}(F) - f^{pc}(F) \\
&\le \sum_{A \in \mathcal{N}_{d,r'}} \tilde{\theta}_A f(A) - \sum_{\ell=1}^{\tau+1} \varrho_\ell f(A_\ell) + c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))} \\
&= \sum_{\ell=1}^{\tau+1} \varrho_\ell \Big(\sum_{\iota=1}^{2^{nm}} \theta_{M_\iota^{(\ell)}} f(M_\iota^{(\ell)}) - f(A_\ell) \Big) + c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))} \\
&= \sum_{\ell=1}^{\tau+1} \varrho_\ell \big(\mathcal{I}_{d,r'} f(A_\ell) - f(A_\ell) \big) + c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))} \\
&\le 2c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))}
\end{aligned} \tag{4.4}$$

which proves the proposition. □

Remark 4.1. The assumption $f \in C^{1,\alpha}_{loc}(\mathbb{R}^{n\times m})$ can be replaced by $f \in C^{1,\alpha}_{loc}(U_d)$ where $U_d = \{G \in \mathbb{R}^{n\times m} : \inf_{H\in U} |G - H|_\infty < d\}$ for $U = \{H \in \mathbb{R}^{n\times m} : f^{pc}(H) = f(H)\}$.

The subsequent lemma states the Kuhn-Tucker optimality conditions for the linear optimization problem that defines $f^{pc}_{d,r}(F)$ and which we will refer to through $f^{pc}_{d,r}(F)$. In particular, the lemma characterizes the Lagrange multiplier $\lambda^F_{d,r}$ mentioned in Section 2.

The equations have first been employed in the context of relaxation in the calculus of variations in [R1, R2] and have further been exploited in the numerical approximation of scalar non-convex variational problems in [CR, Ba1].

Lemma 4.1. *There exists $\lambda^F_{d,r} \in \mathbb{R}^\tau$ such that*

$$\max_{A\in\mathcal{N}_{d,r}} \big(\lambda^F_{d,r}\cdot T(A) - f(A)\big) \le \lambda^F_{d,r}\cdot T(F) - f^{pc}_{d,r}(F). \tag{4.5}$$

Conversely, any $(\theta_A : A \in \mathcal{N}_{d,r})$ that is feasible in $f^{pc}_{d,r}(F)$ is optimal if there exists $\lambda^F_{d,r} \in \mathbb{R}^\tau$ such that (4.5) holds with $f^{pc}_{d,r}(F)$ replaced by $\sum_{A\in\mathcal{N}_{d,r}} \theta_A f(A)$. □

Employing the optimality conditions of Lemma 4.1 we can state sufficient conditions that ensure that $r \le r'$ is large enough so that $f^{pc}_{d,r}(F) = f^{pc}_{d,r'}(F)$, where r' is as in Proposition 4.1. If $p > n\wedge m$ condition (4.6) below can be employed as an a posteriori criterion to iteratively enlarge r.

Proposition 4.2. *Let $\lambda^F_{d,r} \in \mathbb{R}^\tau$ be as in Lemma 4.1, suppose $p \ge n\wedge m$, $r \ge 1$, and*

$$c_T(n\wedge m)|\lambda^F_{d,r}| \le p c_f r^{p-n\wedge m} \quad \text{and} \quad c_T|\lambda^F_{d,r}| r^{n\wedge m} - c_f r^p + c'_f \le \lambda^F_{d,r}\cdot T(F) - f^{pc}_{d,r}(F). \tag{4.6}$$

Then, there holds

$$\max_{A\in d\mathbb{Z}^{n\times m}} \big(\lambda^F_{d,r}\cdot T(A) - f(A)\big) \le \lambda^F_{d,r}\cdot T(F) - f^{pc}_{d,r}(F) \tag{4.7}$$

and $f^{pc}_{d,r'}(F) = f^{pc}_{d,r}(F)$ for all $r' \ge r$.

Proof. Since $c_T(n\wedge m)|\lambda^F_{d,r}| \le p c_f r^{p-n\wedge m}$ and $p \ge n\wedge m$ the mapping

$$\{t\in\mathbb{R} : t \ge r\} \to \mathbb{R}, \quad t \mapsto c_f t^p - c'_f - c_T|\lambda^F_{d,r}| t^{n\wedge m},$$

is monotonically increasing. Since $|T(G)| \le c_T|G|^{n\wedge m}_\infty$ for all $G\in\mathbb{R}^{n\times m}$ with $|G|_\infty \ge 1$, we have, for all $A \in d\mathbb{Z}^{n\times m}\setminus\mathcal{N}_{d,r}$, i.e. for all $A\in d\mathbb{Z}^{n\times m}$ with $|A|_\infty > r \ge 1$,

$$\lambda^F_{d,r}\cdot T(A) - f(A) \le c_T|\lambda^F_{d,r}||A|^{n\wedge m}_\infty - c_f|A|^p_\infty + c'_f \le c_T|\lambda^F_{d,r}|r^{n\wedge m} - c_f r^p + c'_f.$$

Then, the second inequality in (4.6) implies, for all $A \in d\mathbb{Z}^{n\times m}\setminus\mathcal{N}_{d,r}$,

$$f(A) \ge \lambda^F_{d,r}\cdot T(A) - \lambda^F_{d,r}\cdot T(F) + f^{pc}_{d,r}(F)$$

while the optimality conditions (4.5) guarantee, for all $A\in\mathcal{N}_{d,r}$,

$$f(A) \ge \lambda^F_{d,r}\cdot T(A) - \lambda^F_{d,r}\cdot T(F) + f^{pc}_{d,r}(F).$$

The last two estimates prove (4.7). Let $r' \ge r$ and let $(\tilde\theta_A : A \in \mathcal{N}_{d,r'})$ be a solution to $f^{pc}_{d,r'}(F)$. Employing (4.7) and $\sum_{A\in\mathcal{N}_{d,r'}} \tilde\theta_A T(A) = T(F)$ we infer

$$f^{pc}_{d,r'}(F) = \sum_{A\in\mathcal{N}_{d,r'}} \tilde\theta_A f(A) \ge \sum_{A\in\mathcal{N}_{d,r'}} \tilde\theta_A\big(\lambda^F_{d,r}\cdot T(A) - \lambda^F_{d,r}T(F) + f^{pc}_{d,r}(F)\big) = f^{pc}_{d,r}(F).$$

The obvious estimate $f^{pc}_{d,r'}(F) \le f^{pc}_{d,r}(F)$ concludes the proof. □

We now turn to estimates for $\lambda^F_{d,r}$ for which we need to construct a polyconvex extension of $f^{pc}_{d,r}(F)$ to $\mathbb{R}^{n\times m}$. Note that the subsequent definition of this extension does not depend on F, i.e. it only depends on d.

Definition 4.1. Let $\tilde{f}_d$ be the nodal interpolant of f on $\mathbb{R}^{n\times m}$ with respect to nodes in $d\mathbb{Z}^{n\times m}$, i.e. $\tilde{f}_d(A) = f(A)$ for all $A \in d\mathbb{Z}^{n\times m}$, $\tilde{f}_d$ is continuous, and $\tilde{f}_d|_{\tilde{Q}}$ is (nm)-linear for each cube $\tilde{Q} \subseteq \mathbb{R}^{n\times m}$ with vertices in $d\mathbb{Z}^{n\times m}$ and edges of length d. Then, let $\tilde{f}_d^{pc}$ be the polyconvex envelope of $\tilde{f}_d$, i.e. for $A \in \mathbb{R}^{n\times m}$,

$$\tilde{f}_d^{pc}(A) = \inf\Big\{\sum_{\ell=1}^{\tau+1} \varrho_\ell \tilde{f}_d(A_\ell) : A_\ell \in \mathbb{R}^{n\times m},\ \varrho_\ell \geq 0,\ \sum_{\ell=1}^{\tau+1} \varrho_\ell = 1,\ \sum_{\ell=1}^{\tau+1} \varrho_\ell T(A_\ell) = T(A)\Big\}.$$

The following lemma shows that $\tilde{f}_d^{pc}$ can be approximated arbitrarily well by convex combinations of nodal values of f.

Lemma 4.2. *For all $\varepsilon > 0$ and all $A \in \mathbb{R}^{n\times m}$ there exist $B_\kappa \in d\mathbb{Z}^{n\times m}$, $\gamma_\kappa \geq 0$, $\kappa = 1, ..., 2^{nm}(\tau+1)$, such that $\sum_{\kappa=1}^{2^{nm}(\tau+1)} \gamma_\kappa = 1$, $\sum_{\kappa=1}^{2^{nm}(\tau+1)} \gamma_\kappa T(B_\kappa) = T(A)$, and*

$$\tilde{f}_d^{pc}(A) \leq \sum_{\kappa=1}^{2^{nm}(\tau+1)} \gamma_\kappa f(B_\kappa) \leq \tilde{f}_d^{pc}(A) + \varepsilon. \tag{4.8}$$

Remark 4.2. Employing optimality conditions for the minimization problem $\tilde{f}_d^{pc}(A)$ in Definition 4.1 one can show that the infimum is attained if $p > n \wedge m$. In this case one may choose $\varepsilon = 0$ in Lemma 4.2.

Proof of Lemma 4.2. By definition of $\tilde{f}_d^{pc}(A)$ there exist $A_\ell \in \mathbb{R}^{n\times m}$ and $\varrho_\ell \geq 0$, $\ell = 1, .., \tau+1$, such that $\sum_{\ell=1}^{\tau+1} \varrho_\ell = 1$, $\sum_{\ell=1}^{\tau+1} \varrho_\ell T(A_\ell) = T(A)$, and

$$\sum_{\ell=1}^{\tau+1} \varrho_\ell \tilde{f}_d(A_\ell) \leq \tilde{f}_d^{pc}(A) + \varepsilon.$$

For each $\ell = 1, ..., \tau+1$ let $\tilde{Q}_\ell = \operatorname{conv}\big\{M_1^{(\ell)}, ..., M_{2^{nm}}^{(\ell)}\big\}$ be a cube in $\mathbb{R}^{n\times m}$ with vertices $M_1^{(\ell)}, ..., M_{2^{nm}}^{(\ell)} \in d\mathbb{Z}^{n\times m}$, edges of length d, and such that $A_\ell \in \tilde{Q}_\ell$. By Lemma 3.1 (with $r = \tilde{r}$ for some $\tilde{r}$ large enough so that $A_\ell \in \tilde{Q}_\ell \subseteq \omega_{d,\tilde{r}}$) there exist $\theta_{M_\iota^{(\ell)}} \geq 0$, $\iota = 1, ..., 2^{nm}$, such that $\sum_{\iota=1}^{2^{nm}} \theta_{M_\iota^{(\ell)}} = 1$, $\sum_{\iota=1}^{2^{nm}} \theta_{M_\iota^{(\ell)}} T(M_\iota^{(\ell)}) = T(A_\ell)$, and

$$\tilde{f}_d(A_\ell) = \sum_{\iota=1}^{2^{nm}} \varphi_{M_\iota^{(\ell)}}(A_\ell) f(M_\iota^{(\ell)}) = \sum_{\iota=1}^{2^{nm}} \theta_{M_\iota^{(\ell)}} f(M_\iota^{(\ell)}).$$

This implies

$$\sum_{\ell=1}^{\tau+1} \sum_{\iota=1}^{2^{nm}} \varrho_\ell \theta_{M_\iota^{(\ell)}} f(M_\iota^{(\ell)}) \leq \tilde{f}_d^{pc}(A) + \varepsilon$$

which, after appropriate relabeling, is (4.8). $\square$

The following assertion is due to Ball [B1].

Lemma 4.3. *There exist convex functions $\hat{f}, \hat{f}_d : \mathbb{R}^\tau \to \mathbb{R}$ such that*

$$f^{pc} = \hat{f} \circ T \qquad \textit{and} \qquad \tilde{f}_d^{pc} = \hat{f}_d \circ T.$$

For $g = f$ or $g = \tilde{f}_d$ the function $\hat{g} = \hat{f}$ or $\hat{g} = \hat{f}_d$, respectively, can be defined by

$$\hat{g}(X) = \inf\Big\{\sum_{\ell=1}^{\tau+1} \varrho_\ell g(A_\ell) : A_\ell \in \mathbb{R}^{n\times m},\ \varrho_\ell \geq 0,\ \sum_{\ell=1}^{\tau+1} \varrho_\ell = 1,\ \sum_{\ell=1}^{\tau+1} \varrho_\ell T(A_\ell) = X\Big\}. \quad \square$$

Remark 4.3. The function $\hat{g}$ is not unique and the presented formula can be found in [Da2].

An estimate for the difference between $\hat{f}$ and $\hat{f}_d$ follows immediately.

Lemma 4.4. *(i) Let $B^\tau = \{E_1, ..., E_\tau\}$ be the canonical basis in $\mathbb{R}^\tau$. There exists $r' > 0$ such that, for all $E \in \pm B^\tau$, there holds*

$$|\hat{f}_d(T(F) + dE) - \hat{f}(T(F) + dE)| \leq 2c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))}.$$

(ii) Let $B^{n\times m} = \{E_1, ..., E_{nm}\}$ be the canonical basis in $\mathbb{R}^{n\times m}$. There exists $r' > 0$ such that, for all $E \in \pm B^{n\times m}$, there holds

$$|\hat{f}_d(T(F + dE)) - f^{pc}(F + dE)| \leq 2c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))}.$$

Proof. (i) Let $E \in \pm B^\tau$ and let $t > 0$ be such that $|f|_{C^{1,\alpha}(B_t(0))} > 0$. By definition of $\hat{f}$ and $\hat{f}_d$ there exist A_ℓ, $A_\ell^{(d)} \in \mathbb{R}^{n\times m}$ and ϱ_ℓ, $\varrho_\ell^{(d)} \geq 0$, $\ell = 1, ..., \tau + 1$, such that $\sum_{\ell=1}^{\tau+1} \varrho_\ell = \sum_{\ell=1}^{\tau+1} \varrho_\ell^{(d)} = 1$, $\sum_{\ell=1}^{\tau+1} \varrho_\ell A_\ell = \sum_{\ell=1}^{\tau+1} \varrho_\ell^{(d)} A_\ell^{(d)} = T(F) + dE$, and

$$\hat{f}(T(F) + dE) \leq \sum_{\ell=1}^{\tau+1} \varrho_\ell f(A_\ell) \leq \hat{f}(T(F) + dE) + c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_t(0))}$$

as well as

$$\hat{f}_d(T(F) + dE) \leq \sum_{\ell=1}^{\tau+1} \varrho_\ell^{(d)} f(A_\ell^{(d)}) \leq \hat{f}_d(T(F) + dE) + c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_t(0))}.$$

Let $r' \geq t$ be such that $A_\ell, A_\ell^{(d)} \in \omega_{d,r'}$ for all $\ell = 1, ..., \tau+1$. If $\hat{f}(T(F)+dE) \geq \hat{f}_d(T(F)+dE)$ then

$$\begin{aligned} 0 &\leq \hat{f}(T(F) + dE) - \hat{f}_d(T(F) + dE) \\ &\leq \sum_{\ell=1}^{\tau+1} \varrho_\ell^{(d)} \big(f(A_\ell^{(d)}) - f_d(A_\ell^{(d)})\big) + c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))} \\ &\leq 2c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))}. \end{aligned}$$

Otherwise, if $\hat{f}(T(F) + dE) \leq \hat{f}_d(T(F) + dE)$ then

$$\begin{aligned} 0 &\leq \hat{f}_d(T(F) + dE) - \hat{f}(T(F) + dE) \\ &\leq \sum_{\ell=1}^{\tau+1} \varrho_\ell \big(f_d(A_\ell) - f(A_\ell)\big) + c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))} \\ &\leq 2c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))}. \end{aligned}$$

Choosing r' maximal so that for each $E \in \pm B^\tau$ one of the last two estimates holds proves the first part of the lemma.
(ii) The proof of the second assertion follows as (i) together with the fact that $\hat{f} \circ T = f^{pc}$. $\square$

The next lemma is the key observation for the estimates for $\lambda_{d,r}^F$ for which we employ the concept of subgradients. Some elementary facts about the subgradient are cited in the following remark.

Remark 4.4 ([C]). Let $h : \mathbb{R}^\ell \to \mathbb{R}$ be a continuous, convex function. For $v_0 \in \mathbb{R}^\ell$ define

$$\partial h(v_0) := \{s \in \mathbb{R}^\ell : \forall v \in \mathbb{R}^\ell, s \cdot (v - v_0) \le h(v) - h(v_0)\}.$$

There holds: (i) If $g : \mathbb{R} \to \mathbb{R}^\ell$ is affine and $g(0) = v_0$ then $\partial(h \circ g)(0) = \partial h(v_0) \cdot Dg(0)$.
(ii) If $\ell = 1$ then $\partial h(v_0) \subseteq \big[(h(v_0) - h(v_0 - s))/s, (h(v_0 + s) - h(v_0))/s\big]$ for all $s > 0$.
(iii) If $h(v_0) \le h(v)$ for all $v \in \mathbb{R}^\ell$ then $0 \in \partial h(v_0)$.

Lemma 4.5. *Suppose that*

$$\max_{A \in d\mathbb{Z}^{n\times m}} \big(\lambda_{d,r}^F \cdot T(A) - f(A)\big) \le \lambda_{d,r}^F \cdot T(F) - f_{d,r}^{pc}(F).$$

Then there holds $\tilde f_d^{pc}(F) = f_{d,r}^{pc}(F)$ and $\lambda_{d,r}^F \in \partial \hat f_d(T(F))$.

Proof. We show that $\tilde f_d^{pc}(F) = f_{d,r}^{pc}(F)$ and

$$\max_{A \in \mathbb{R}^{n\times m}} \big(\lambda_{d,r}^F \cdot T(A) - \tilde f_d^{pc}(A)\big) \le \lambda_{d,r}^F \cdot T(F) - \tilde f_d^{pc}(F). \tag{4.9}$$

Then, the asserted inclusion is deduced from these observations as follows: Let $\varepsilon > 0$. For $X \in \mathbb{R}^\tau$ there exist $A_\ell \in \mathbb{R}^{n\times m}$, $\varrho_\ell \ge 0$, $\ell = 1, ..., \tau + 1$, such that $\sum_{\ell=1}^{\tau+1} \varrho_\ell = 1$, $\sum_{\ell=1}^{\tau+1} \varrho_\ell T(A_\ell) = X$, and

$$\hat f_d(X) \ge \sum_{\ell=1}^{\tau+1} \varrho_\ell \tilde f_d(A_\ell) - \varepsilon.$$

Using $\tilde f_d^{pc} \le \tilde f_d$, $\tilde f_d^{pc}(F) = \hat f_d(T(F))$, and (4.9) we deduce

$$\begin{aligned}
\hat f_d(X) - \lambda_{d,r}^F \cdot X &\ge \sum_{\ell=1}^{\tau+1} \varrho_\ell\big(\tilde f_d(A_\ell) - \lambda_{d,r}^F \cdot T(A_\ell)\big) - \varepsilon \\
&\ge \sum_{\ell=1}^{\tau+1} \varrho_\ell\big(\tilde f_d^{pc}(A_\ell) - \lambda_{d,r}^F \cdot T(A_\ell)\big) - \varepsilon \\
&\ge \tilde f_d^{pc}(F) - \lambda_{d,r}^F \cdot T(F) - \varepsilon \\
&= \hat f_d(T(F)) - \lambda_{d,r}^F \cdot T(F) - \varepsilon.
\end{aligned}$$

By arbitrariness of $\varepsilon > 0$, the convex function $X \mapsto \hat f_d(X) - \lambda_{d,r}^F \cdot X$, $X \in \mathbb{R}^\tau$, has a minimum in $T(F)$ so that (iii) in Remark 4.4 implies the asserted inclusion.

To verify $\tilde f_d^{pc}(F) = f_{d,r}^{pc}(F)$ we note that $\tilde f_d^{pc}(F) \le f_{d,r}^{pc}(F)$ and let $\varepsilon > 0$. By Lemma 4.2 there exist $B_\kappa \in d\mathbb{Z}^{n\times m}$, $\gamma_\kappa \ge 0$, $\kappa = 1, ..., 2^{nm}(\tau + 1)$, such that $\sum_{\kappa=1}^{2^{nm}(\tau+1)} \gamma_\kappa = 1$, $\sum_{\kappa=1}^{2^{nm}(\tau+1)} \gamma_\kappa T(B_\kappa) = T(F)$, and

$$\sum_{\kappa=1}^{2^{nm}(\tau+1)} \gamma_\kappa f(B_\kappa) \le \tilde f_d^{pc}(F) + \varepsilon.$$

The hypothesis of the lemma implies, for $\kappa = 1, .., 2^{nm}(\tau + 1)$,

$$f(B_\kappa) \geq \lambda^F_{d,r} \cdot T(B_\kappa) - \lambda^F_{d,r} \cdot T(F) + f^{pc}_{d,r}(F)$$

and this estimate yields

$$\tilde{f}^{pc}_d(F) \geq \sum_{\kappa=1}^{2^{nm}(\tau+1)} \gamma_\kappa f(B_\kappa) - \varepsilon \geq f^{pc}_{d,r}(F) - \varepsilon,$$

which, by arbitrariness of $\varepsilon > 0$, shows $\tilde{f}^{pc}_d(F) \geq f^{pc}_{d,r}(F)$ and hence yields $\tilde{f}^{pc}_d(F) = f^{pc}_{d,r}(F)$.

To prove (4.9), let $A^* \in \mathbb{R}^{n\times m}$ be maximal in the left-hand side of (4.9). For $\varepsilon > 0$ Lemma 4.2 guarantees the existence of $C_\kappa \in d\mathbb{Z}^{n\times m}$ and $\delta_\kappa \geq 0$, $\kappa = 1, ..., 2^{nm}(\tau + 1)$ such that $\sum_{\kappa=1}^{2^{nm}(\tau+1)} \delta_\kappa = 1$, $\sum_{\kappa=1}^{2^{nm}(\tau+1)} \delta_\kappa T(C_\kappa) = T(A^*)$, and

$$\sum_{\kappa=1}^{2^{nm}(\tau+1)} \delta_\kappa f(C_\kappa) \leq \tilde{f}^{pc}_d(A^*) + \varepsilon.$$

Then, the hypothesis of the lemma and $\tilde{f}^{pc}_d(F) = f^{pc}_{d,r}(F)$ imply

$$\begin{aligned}
\max_{A\in\mathbb{R}^{n\times m}} \big(\lambda^F_{d,r} \cdot T(A) - \tilde{f}^{pc}_d(A)\big) &\leq \lambda^F_{d,r} \cdot T(A^*) - \tilde{f}^{pc}_d(A^*) \\
&\leq \sum_{\kappa=1}^{2^{nm}(\tau+1)} \delta_\kappa \big(\lambda^F_{d,r} \cdot T(C_\kappa) - f(C_\kappa)\big) + \varepsilon \\
&\leq \sum_{\kappa=1}^{2^{nm}(\tau+1)} \delta_\kappa \max_{A\in d\mathbb{Z}^{n\times m}} \big(\lambda^F_{d,r} \cdot T(A) - f(A)\big) + \varepsilon \\
&\leq \lambda^F_{d,r} \cdot T(F) - f^{pc}_{d,r}(F) + \varepsilon \\
&= \lambda^F_{d,r} \cdot T(F) - \tilde{f}^{pc}_d(F) + \varepsilon,
\end{aligned}$$

which, by arbitrariness of $\varepsilon > 0$, is (4.9) and therefore concludes the proof. $\square$

Provided that $\hat{f}$ is of class $C^{1,\alpha}_{loc}$ we have an estimate for $|\lambda^F_{d,r} - D\hat{f}(T(F))|$.

Proposition 4.3. *Assume that the hypothesis of Lemma 4.5 is satisfied, suppose that $\alpha > 0$, and let $\hat{f} \in C^{1,\alpha}_{loc}(\mathbb{R}^\tau)$. There exists $r' > 0$ such that*

$$|\lambda^F_{d,r} - D\hat{f}(T(F))| \leq 4\sqrt{\tau} c_{\mathcal{I}} d^\alpha |f|_{C^{1,\alpha}(B_{r'}(0))} + \sqrt{\tau} d^\alpha |\hat{f}|_{C^{1,\alpha}(B_d(T(F)))}.$$

Proof. Let $B^\tau = \{E_1, ..., E_\tau\}$ be the canonical basis in $\mathbb{R}^\tau$. Lemma 4.5 proves

$$\begin{aligned}
|\lambda^F_{d,r} - D\hat{f}(T(F))| &\leq \sup_{S\in\partial\hat{f}_d(T(F))} |S - D\hat{f}(T(F))| \\
&\leq \sqrt{\tau} \sup_{S\in\partial\hat{f}_d(T(F))} |S - D\hat{f}(T(F))|_\infty \\
&= \sqrt{\tau} \sup_{S\in\partial\hat{f}_d(T(F))} \max_{E\in\pm B^\tau} |(S - D\hat{f}(T(F))) \cdot E|.
\end{aligned}$$

Let $S \in \partial \hat{f}_d(T(F))$ and $E \in \pm B^\tau$. Since $\hat{f}_d$ is convex (i) and (ii) in Remark 4.4 show $S \cdot E \in \partial \hat{f}_d(T(F)) \cdot E \subseteq [S_-(E), S_+(E)]$ for

$$S_\pm(E) = \pm \frac{\hat{f}_d(T(F) \pm dE) - \hat{f}_d(T(F))}{d}.$$

Assume without loss of generality $S_-(E) \leq S_+(E)$. Lemma 4.4 and Proposition 4.1 (note that $\hat{f}(T(F)) = f^{pc}(F)$ and $\hat{f}_d(T(F)) = f^{pc}_{d,r}(F)$) prove

$$S_+(E) \leq \frac{\hat{f}(T(F) + dE) - \hat{f}(T(F)) + 4c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))}}{d}$$

and

$$S_-(E) \geq \frac{\hat{f}(T(F)) - \hat{f}(T(F) - dE) - 4c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))}}{d}.$$

From these estimates, the mean value theorem, and Hölder continuity of $D\hat{f}$ we infer

$$|(S - D\hat{f}(T(F))) \cdot E| \leq 4c_{\mathcal{I}} d^\alpha |f|_{C^{1,\alpha}(B_{r'}(0))} + d^\alpha |\hat{f}|_{C^{1,\alpha}(B_d(T(F)))}$$

which concludes the proof. □

It is not known under which conditions there holds $\hat{f} \in C^{1,\alpha}_{loc}(\mathbb{R}^\tau)$ or under which conditions there exists a convex function $\tilde{f} \in C^{1,\alpha}_{loc}(\mathbb{R}^\tau)$ such that $\tilde{f} = f^{pc} \circ T$. A result in [BKK] shows that if

$$\liminf_{G \to \infty} \frac{f(G)}{|G|^p} > 0 \quad \text{and} \quad \limsup_{G \to \infty} \frac{f(G)}{|G|^{p+1}} = 0$$

or

$$\frac{f(G)}{|G|^p} \to \infty \text{ as } G \to \infty \quad \text{and} \quad \limsup_{G \to \infty} \frac{f(G)}{|G|^{p+1}} < \infty$$

and if there exists $c > 0$ such that, for all $G \in \mathbb{R}^{n \times m}$, there exists $S \in \partial f(G)$ such that

$$f(G + H) - f(H) - S \cdot H \leq c \max\{f(G), 1\} |H|^{1+\alpha}$$

for all $H \in B_1(0)$ then $f^{pc} \in C^{1,\alpha}_{loc}(\mathbb{R}^{n \times m})$. Observing that $D\hat{f}(T(F)) \cdot DT(F) = Df^{pc}(F)$ a small modification (that does not use *any* regularity of $\hat{f}$) of the proof of Proposition 4.3 yields the following result.

Proposition 4.4. *Assume that the hypothesis of Lemma 4.5 is satisfied, suppose that $\alpha > 0$, and let $f^{pc} \in C^{1,\alpha}_{loc}(\mathbb{R}^{n \times m})$. Then, there exists $r' > 0$ such that*

$$|\lambda^F_{d,r} \cdot DT(F) - Df^{pc}(F)| \leq 4\sqrt{nm} c_{\mathcal{I}} d^\alpha |f|_{C^{1,\alpha}(B_{r'}(0))} + \sqrt{nm} d^\alpha |f^{pc}|_{C^{1,\alpha}(B_d(F))}.$$

Proof. Let $B^{n \times m} = \{E_1, ..., E_{n \times m}\}$ be the canonical basis in $\mathbb{R}^{n \times m}$. Lemma 4.5 proves

$$|\lambda^F_{d,r} \cdot DT(F) - Df^{pc}(F)| \leq \sqrt{nm} \sup_{S \in \partial \hat{f}_d(T(F))} \max_{E \in \pm B^{n \times m}} |(S \cdot DT(F) - Df^{pc}(F)) \cdot E|.$$

Let $S \in \partial \hat{f}_d(T(F))$ and $E \in \pm B^{n \times m}$. Convexity of $t \mapsto \hat{f}_d(T(F + tE))$ shows $\big(S \cdot DT(F)\big) \cdot E \in [S_-(E), S_+(E)]$ (cf. Remark 4.4) for

$$S_\pm(E) = \pm \frac{\hat{f}_d(T(F \pm dE)) - \hat{f}_d(T(F))}{d}.$$

Assuming $S_-(E) \le S_+(E)$, Lemma 4.4, $\hat{f}(T(F)) = f^{pc}_{d,r}(F)$, and Proposition 4.1 show

$$S_+(E) \le \frac{f^{pc}(F+dE) - f^{pc}(F) + 4c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))}}{d}$$

and

$$S_-(E) \ge \frac{f^{pc}(F) - f^{pc}(F-dE) - 4c_{\mathcal{I}} d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))}}{d}.$$

These estimates, the mean value theorem, and Hölder continuity of Df^{pc} imply

$$|(S \cdot DT(F) - Df^{pc}(F)) \cdot E| \le 4c_{\mathcal{I}} d^{\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))} + d^{\alpha} |f^{pc}|_{C^{1,\alpha}(B_d(F))}$$

and thereby prove the proposition. □

Proof of Theorem A. This is a combination of Propositions 4.1, 4.2,4.4, and Lemma 4.5. □

5. Efficient Computation of $f^{pc}_{d,r}(F)$

As mentioned in the introduction, the direct computation of $f^{pc}_{d,r}(F)$ is very expensive. To reduce the number of unknowns we use a multilevel scheme with local mesh refinement and coarsening.

5.1. **Grid coarsening and local refinement.**

The following propositions define criteria that allow to add nodes to and remove nodes from $\mathcal{N}_{d,r}$ to obtain a set of nodes that leads to a good approximation of $f^{pc}_{d/2,r}(F)$ provided that we computed (an approximation of) $f^{pc}_{d,r}(F)$. The first assertion allows to remove nodes $A \in \mathcal{N}_{d/2,r}$ that cannot lead to volume fractions θ_A larger than a given threshold $\delta = c/M$ for a known constant c and a (large) parameter M.

Proposition 5.1. *Let $(\theta_A : A \in \mathcal{N}_{d,r})$ be feasible and optimal for $f^{pc}_{d,r}(F)$ with corresponding multiplier $\lambda^F_{d,r} \in \mathbb{R}^\tau$. For each $A \in \mathcal{N}_{d,r}$ let $M(A) \ge M > 0$ and set*

$$Z = \{A \in \mathcal{N}_{d,r} : \lambda^F_{d,r} \cdot T(A) - f(A) \le \lambda^F_{d,r} \cdot T(F) - f^{pc}_{d,r}(F) - d\, M(A)\}.$$

Let

$$Z' = \{A' \in \mathcal{N}_{d/2,r} : \exists A \in Z,\, |A - A'| \le d\}.$$

Then, for any $(\theta'_{A'} : A' \in \mathcal{N}_{d/2,r})$ that is feasible and optimal for $f^{pc}_{d/2,r}(F)$ there holds

$$\sum_{A' \in Z'} \theta'_{A'} \le (|f|_{Lip,r} + |\lambda^F_{d,r}||T|_{Lip,r})/M.$$

Proof. For $A' \in Z'$ and $A \in Z$ such that $|A - A'| \le d$ there holds

$$\begin{aligned}
f(A') &\ge f(A) - d|f|_{Lip,r} \\
&\ge f^{pc}_{d,r}(F) + \lambda^F_{d,r} \cdot T(A) - \lambda^F_{d,r} \cdot T(F) + dM(A) - d|f|_{Lip,r} \\
&\ge f^{pc}_{d,r}(F) + \lambda^F_{d,r} \cdot T(A') - \lambda^F_{d,r} \cdot T(F) + dM(A) - d|f|_{Lip,r} - d|\lambda^F_{d,r}||T|_{Lip,r}.
\end{aligned}$$

Let $B' \in \mathcal{N}_{d/2,r} \setminus Z'$ and $B \in \mathcal{N}_{d,r}$ be such that $|B' - B| \le d$. Employing (4.5) and arguing similarly as in the previous estimate we infer

$$f(B') \ge f^{pc}_{d,r}(F) + \lambda^F_{d,r} \cdot T(B') - \lambda^F_{d,r} \cdot T(F) - d|f|_{Lip,r} - d|\lambda^F_{d,r}||T|_{Lip,r}.$$

Let $(\theta'_{A'} : A' \in \mathcal{N}_{d/2,r})$ be feasible and optimal for $f^{pc}_{d/2,r}(F)$. The previous two estimates imply

$$\begin{aligned} f^{pc}_{d/2,r}(F) = \sum_{A' \in \mathcal{N}_{d/2,r}} \theta_{A'} f(A') &= \sum_{A' \in Z'} \theta_{A'} f(A') + \sum_{B' \in \mathcal{N}_{d/2,r} \setminus Z'} \theta_{B'} f(B') \\ &\geq f^{pc}_{d,r}(F) - (|f|_{Lip,r} + |\lambda^F_{d,r}||T|_{Lip,r})d + \sum_{A' \in Z'} \theta_{A'} M d. \end{aligned}$$

Using $f^{pc}_{d/2,r}(F) \leq f^{pc}_{d,r}(F)$ we deduce $\sum_{A' \in Z'} \theta_{A'} \leq \big(|f|_{Lip,r} + |\lambda^F_{d,r}||T|_{Lip,r}\big)/M$. $\square$

A more efficient and even more reliable assertion can be formulated if we have explicit estimates for $|f^{pc}_{d,r}(F) - f^{pc}_{d/2,r}(F)|$ and $|\lambda^F_{d,r} - \lambda^F_{d/2,r}|$. Sufficient for this is the knowledge of r' in Theorem A and $\hat{f} \in C^{1,1}_{loc}(\mathbb{R}^\tau)$ with $\hat{f}$ as in Lemma 4.3.

Proposition 5.2. *For each $A \in \mathcal{N}_{d,r}$ let $M(A) > 0$, suppose that $|\lambda^F_{d,r} - \lambda^F_{d/2,r}| \leq c_{LM} d$, and let Z and Z' be as in Proposition 5.1. Assume that for all $A \in \mathcal{N}_{d,r}$ there holds*

$$c_{LM}|T|_{Lip,r'} + |\lambda^F_{d,r}|\,|T|_{Lip,r'} + 4c_{\mathcal{I}}|f|_{Lip,r'} + c_{LM}|T(A)| + c_{LM}|T(F)| < M(A). \tag{5.1}$$

Then, for any $(\theta_A : A \in \mathcal{N}_{d/2,r})$ that is feasible and optimal for $f^{pc}_{d/2,r}(F)$ there holds $\theta_{A'} = 0$ for all $A' \in Z'$.

Proof. Let $A' \in Z'$ and $A \in Z$ such that $|A - A'| \leq d$. By the hypotheses and by Lipschitz continuity of f and T there holds

$$\begin{aligned} &\lambda^F_{d/2,r} \cdot T(A') - f(A') \\ &= \lambda^F_{d,r} \cdot T(A) - f(A) + \big(\lambda^F_{d/2,r} - \lambda^F_{d,r}\big) \cdot T(A') + \lambda^F_{d,r} \cdot \big((T(A') - T(A)\big) + \big(f(A) - f(A')\big) \\ &\leq \lambda^F_{d,r} \cdot T(A) - f(A) + c_{LM} d|T(A')| + d|\lambda^F_{d,r}||T|_{Lip,r'} + d|f|_{Lip,r'} \\ &\leq \lambda^F_{d,r} \cdot T(F) - f^{pc}_{d,r}(F) - dM(A) + c_{LM} d|T(A')| + d|\lambda^F_{d,r}||T|_{Lip,r'} + d|f|_{Lip,r'}. \end{aligned}$$

The definitions of Z and Z', Proposition 4.1, and again the assumed estimate for $|\lambda^F_{d,r} - \lambda^F_{d/2,r}|$ show

$$\begin{aligned} &\lambda^F_{d,r} \cdot T(F) - f^{pc}_{d,r}(F) - dM(A) + c_{LM} d|T(A')| + d|\lambda^F_{d,r}||T|_{Lip,r'} + d|f|_{Lip,r'} \\ &= \lambda^F_{d/2,r} \cdot T(F) - f^{pc}_{d/2,r}(F) + \big(\lambda^F_{d,r} - \lambda^F_{d/2,r}\big) \cdot T(F) + \big(f^{pc}_{d/2,r}(F) - f^{pc}_{d,r}(F)\big) \\ &\qquad - dM(A) + c_{LM} d|T(A')| + d|\lambda^F_{d,r}||T|_{Lip,r'} + 2c_{\mathcal{I}} d|f|_{Lip,r'} \\ &\leq \lambda^F_{d/2,r} \cdot T(F) - f^{pc}_{d/2,r}(F) + c_{LM} d|T(F)| + d|\lambda^F_{d,r}||T|_{Lip,r'} - dM(A) \\ &\qquad + 2c_{\mathcal{I}} d|f|_{Lip,r'} + c_{LM} d|T(F)| + 2c_{\mathcal{I}} d|f|_{Lip,r'}. \end{aligned}$$

Employing once more Lipschitz continuity of T proves

$$\begin{aligned} &\lambda^F_{d/2,r} \cdot T(F) - f^{pc}_{d/2,r}(F) \\ &\qquad + d\big(c_{LM}|T(A')| + |\lambda^F_{d,r}||T|_{Lip,r'} + 2c_{\mathcal{I}} d|f|_{Lip,r'} + c_{LM}|T(F)| - M(A)\big) \\ &\leq \lambda^F_{d/2,r} \cdot T(F) - f^{pc}_{d/2,r}(F) \\ &\qquad + d\big(c_{LM}|T|_{Lip,r'} + c_{LM}|T(A)| + |\lambda^F_{d,r}||T|_{Lip,r'} + 4c_{\mathcal{I}}|f|_{Lip,r'} + c_{LM}|T(F)| - M(A)\big). \end{aligned}$$

In view of (5.1) the last three estimates imply, for all $A' \in Z'$,

$$\lambda^F_{d/2,r} \cdot T(A') - f(A') < \lambda^F_{d/2,r} \cdot T(F) - f^{pc}_{d/2,r}(F). \tag{5.2}$$

The optimality conditions ensure, for all $A \in \mathcal{N}_{d/2,r}$,

$$\lambda^F_{d/2,r} \cdot T(A) - f(A) \le \lambda^F_{d/2,r} \cdot T(F) - f^{pc}_{d/2,r}(F). \tag{5.3}$$

Let $(\theta_A : A \in \mathcal{N}_{d/2,r})$ be feasible and optimal for $f^{pc}_{d/2,r}(F)$ and suppose that there is $A \in Z'$ such that $\theta_A > 0$. Then, (5.2) and (5.3) imply

$$\begin{aligned} f^{pc}_{d/2,r}(F) &= \sum_{A \in \mathcal{N}_{d/2,r}} \theta_A f(A) \\ &> \sum_{A \in \mathcal{N}_{d/2,r}} \theta_A \big(\lambda^F_{d/2,r} \cdot T(A') - \lambda^F_{d/2,r} \cdot T(F) + f^{pc}_{d/2,r}(F)\big) = f^{pc}_{d/2,r}(F). \end{aligned} \tag{5.4}$$

This is a contradiction and proves $\theta_A = 0$ for all $A \in Z'$. □

5.2. **Prediction of the active set.** Following an idea in [CR] for the approximation of scalar nonconvex variational problems we can further remove nodes temporarily from a mesh with nodes $\mathcal{N}$, e.g. $\mathcal{N} \subseteq \mathcal{N}_{d,r}$ is a refinement of $\mathcal{N}_{2d,r}$, using an iterative method that we establish in the following lemma. The method consists in defining an appropriate subset $X \subseteq \mathcal{N}$ and seeking for a solution of a lower dimensional subproblem. For a discrete set $\mathcal{N} \subseteq \mathbb{R}^{n\times m}$ we define

$$f^{pc}_{\mathcal{N}}(F) := \min\Big\{ \sum_{A\in\mathcal{N}} \theta_A f(A) : \forall A \in \mathcal{N},\ \theta_A \ge 0,\ \sum_{A\in\mathcal{N}} \theta_A = 1,\ \sum_{A\in\mathcal{N}} \theta_A T(A) = T(F) \Big\}.$$

Optimality conditions for $f^{pc}_{\mathcal{N}}(F)$ guarantee the existence of some $\lambda^F_{\mathcal{N}} \in \mathbb{R}^\tau$ such that

$$\max_{A\in\mathcal{N}} \big(\lambda^F_{\mathcal{N}} \cdot T(A) - f(A)\big) \le \lambda^F_{\mathcal{N}} \cdot T(F) - f^{pc}_{\mathcal{N}}(F). \tag{5.5}$$

Conversely, any $(\theta_A : A \in \mathcal{N})$ that is feasible in $f^{pc}_{\mathcal{N}}(F)$ is optimal if there exists $\lambda^F_{\mathcal{N}} \in \mathbb{R}^\tau$ such that (5.5) holds with $f^{pc}_{\mathcal{N}}(F)$ replaced by $\sum_{A\in\mathcal{N}} \theta_A f(A)$.

Given $X \subseteq \mathcal{N}$ we consider the following lower dimensional subproblem of $f^{pc}_{\mathcal{N}}(F)$,

$$f^{pc}_{\mathcal{N},X}(F) := \min\Big\{ \sum_{A\in X} \theta_A f(A) : \forall A \in X,\ \theta_A \ge 0,\ \sum_{A\in X} \theta_A = 1,\ \sum_{A\in X} \theta_A T(A) = T(F) \Big\}.$$

The next lemma states sufficient conditions on X such that $f^{pc}_{\mathcal{N},X}(F) = f^{pc}_{\mathcal{N}}(F)$ and directly leads to an iterative algorithm.

Lemma 5.1. *Let $(\theta_A : A \in \mathcal{N})$ be feasible and optimal for $f^{pc}_{\mathcal{N}}(F)$ with multiplier $\lambda^F_{\mathcal{N}} \in \mathbb{R}^\tau$. Assume $\varepsilon_{AS} > 0$ and $\tilde\lambda^F \in \mathbb{R}^\tau$ satisfy $\sup_{A\in\mathcal{N}} |(\tilde\lambda^F - \lambda^F_{\mathcal{N}}) \cdot T(A)| \le \varepsilon_{AS}/2$. If*

$$X = \Big\{ A \in \mathcal{N} : \tilde\lambda^F \cdot T(A) - f(A) \ge \max_{A'\in\mathcal{N}} \big(\tilde\lambda^F \cdot T(A') - f(A')\big) - \varepsilon_{AS} \Big\}$$

and if the optimization problem $f^{pc}_{\mathcal{N},X}(F)$ is feasible then $f^{pc}_{\mathcal{N},X}(F) = f^{pc}_{\mathcal{N}}(F)$.

Proof. The optimality conditions (5.5) show (cf. (5.4) in the proof of Proposition 5.2), for all $A \in \mathcal{N}$,

$$\theta_A > 0 \quad \Longrightarrow \quad A \in Y := \{A' \in \mathcal{N} : \lambda_{\mathcal{N}}^F \cdot T(A') - f(A') = \lambda_{\mathcal{N}}^F \cdot T(F) - f_{\mathcal{N}}^{pc}(F)\}.$$

Hence it suffices to show that $Y \subseteq X$. Let $A \in Y$. By assumption on ε_{AS}, the definitions of X and Y, and (5.5) there holds

$$\begin{aligned}\tilde{\lambda}^F \cdot T(A) - f(A) &\geq \lambda_{\mathcal{N}}^F \cdot T(A) - f(A) - \varepsilon_{AS}/2 \\ &= \lambda_{\mathcal{N}}^F \cdot T(F) - f_{\mathcal{N}}^{pc}(F) - \varepsilon_{AS}/2 \\ &= \max_{A' \in \mathcal{N}} \big(\lambda_{\mathcal{N}}^F \cdot T(A') - f(A')\big) - \varepsilon_{AS}/2 \\ &\geq \max_{A' \in \mathcal{N}} \big(\tilde{\lambda}^F \cdot T(A') - f(A')\big) - \varepsilon_{AS},\end{aligned}$$

i.e. $A \in X$. □

Given some $\tilde{\lambda}^F$ we do in general not know ε_{AS} to define X as in the lemma. We may however enlarge ε_{AS} successively until the optimality conditions (5.5) are satisfied. Having computed a solution for some parameter d we may then use the corresponding multiplier to define X on a finer mesh.

5.3. **An iterative, adaptive algorithm.** The ideas of the preceding propositions lead to the following algorithm that iteratively computes, for prescribed $d_0 > 0$, $r_0 > 0$, $M > 0$, and $J > 0$, a solution to $f^{pc}_{d_0/2^J, 2^\ell r_0}(F)$ where for some $\ell \geq 0$, $r = 2^\ell r_0$ satisfies the conditions of Theorem A for $d = d_0/2^J$ if $p > n \wedge m$. If $p = n \wedge m$ we suppose that r_0 is large enough, i.e. $r_0 \geq r'$ for r' as in Theorem A. We assume that $F \in \omega_{d_0, r_0}$.

Algorithm $(A^{pc,\,adapt}_{r_0, d_0, J, F, M})$**.**

(a) Set $j := 0$, $d := d_0$, $r := r_0$, $\tilde{\lambda}^F := 0$, $\mathcal{N} := \mathcal{N}_{d,r}$, and $\varepsilon_{AS} := \infty$.

(b) Define

$$X := \Big\{A \in \mathcal{N} : \tilde{\lambda}^F \cdot T(A) - f(A) \geq \max_{A' \in \mathcal{N}} \big(\tilde{\lambda}^F \cdot T(A') - f(A')\big) - \varepsilon_{AS}\Big\} \\ \cup \big\{A \in \mathcal{N}_{d,r} : |F - A| \leq d\big\}.$$

(c) Compute $f^{pc}_{\mathcal{N},X}(F)$ and obtain a multiplier $\lambda_{\mathcal{N}}^F \in \mathbb{R}^\tau$.

(d) If for all $A \in \mathcal{N}$ there holds

$$\lambda_{\mathcal{N}}^F \cdot T(A) - f(A) \leq \lambda_{\mathcal{N}}^F \cdot T(F) - f_{\mathcal{N}}^{pc}(F)$$

go to (f).

(e) Set $\tilde{\lambda}^F := \lambda_{\mathcal{N}}^F$, $\varepsilon_{AS} := 2\varepsilon_{AS}$, and go to (b).

(f) If $p = n \wedge m$ or

$$c_T(n \wedge m)|\lambda_{\mathcal{N}}^F| \leq p c_f r^{p - n\wedge m} \quad \text{and} \quad c_T|\lambda_{\mathcal{N}}^F| r^{n \wedge m} - c_f r^p + c_f' \leq \lambda_{\mathcal{N}}^F \cdot T(F) - f_{\mathcal{N}}^{pc}(F)$$

go to (h).

(g) Set $r := 2r$, $\tilde{\lambda}^F := \lambda_{\mathcal{N}}^F$, and go to (b).

(h) If $j < J$ define

$$\mathcal{N} = \big\{A' \in \mathcal{N}_{d/2,r} : \exists A \in \mathcal{N},\ |A - A'| \le d,$$
$$\lambda^F_{\mathcal{N}} \cdot T(A) - f(A) > \lambda^F_{\mathcal{N}} \cdot T(F) - f^{pc}_{\mathcal{N}}(F) - Md\big\},$$

set $\tilde{\lambda}^F := \lambda^F_{\mathcal{N}}$, $\varepsilon_{AS} := d$, $d := d/2$, $j := j+1$, and go to (b).

(j) Stop.

Remarks 5.1. (i) We set $\varepsilon_{AS} = d$ in Step (h) since in the optimal case (if $\hat{f} \in C^{1,1}_{loc}(\mathbb{R}^\tau)$) Proposition 4.3 guarantees $|\lambda^F_{d,r} - \lambda^F_{d/2,r}| \le \mathcal{O}(d)$ so that the conditions of Lemma 5.1 are satisfied up to some constant of order $\mathcal{O}(1)$.
(ii) Note that $\mathcal{N}_{d/2,r}$ need not be computed explicitly since we can add nodes to and remove nodes from $\mathcal{N}$ locally.
(iii) Adding $\big\{A \in \mathcal{N}_{d,r} : |F - A| \le d\big\}$ to X in Step (b) guarantees feasibility of $f^{pc}_{\mathcal{N},X}(F)$.

6. Numerical Experiments I

In this section we report on the practical performance of Algorithm $(A^{pc,\,adapt}_{r_0,d_0,J,F,M})$ when applied to three choices of f for which explicit formulae for f^{pc} and f^{rc} are known.

Example 6.1 ([Da2]). For $n = m = 2$ and $F \in \mathbb{R}^{2\times 2}$ let

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

Then, (2.1) holds for $p = 4$ and $c_f = (c-2)/c$, $c'_f = 2c - 1$ for all $c > 2$ and we choose $c = 3$. In this example $f^{pc} = f^{qc} = f^{**}$ where f^{**} is the convex envelope of f and for $F \in \mathbb{R}^{2\times 2}$ given by

$$f^{**}(F) = \begin{cases} (|F|^2 - 1)^2 & \text{for } |F| \ge 1, \\ 0 & \text{for } |F| \le 1. \end{cases}$$

Example 6.2 ([Ko, DW]). 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

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

Then, (2.1) holds for $p = 2$, $c_f = 1/2^3$, and $c'_f = \max\{|A_1|^2, |A_2|^2\}/2 = 17/16$. Here, $f^{**} \ne f^{pc} = f^{qc}$ and f^{pc} is for $F \in \mathbb{R}^{2\times 2}$ given by

$$f^{pc}(F) = \begin{cases} f_1(F) & \text{for } f_1(F) - f_2(F) \le -\lambda/2, \\ f_2(F) - \big(f_2(F) - f_1(F) + \lambda/2\big)/(2\lambda) & \text{for } |f_1(F) - f_2(F)| \le \lambda/2, \\ f_2(F), & \text{for } f_1(F) - f_2(F) \ge \lambda/2, \end{cases}$$

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

Example 6.3 ([KS, Do]). For $n = m = 2$ and $F \in \mathbb{R}^{2\times 2}$ a modification proposed in [Do] (to ensure continuity of f) of an energy density occurring in an optimal design problem in [KS] reads

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

Then, (2.1) holds for $p = 2$, $c_f = 1$, and $c'_f = 0$. Letting $\varrho(F) := \sqrt{|F|^2 + 2|\det F|}$ for $F \in \mathbb{R}^{2\times 2}$ there holds

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

Note that $f^{**} \neq f^{pc}$ in this example.

We tested Algorithm $(A^{pc,\,adapt}_{r_0,d_0,J,F,M})$ in Examples 6.1-6.3. The implementation of the algorithm was performed in Matlab with a generation of the adaptively refined grids in C. The experiments were performed on a node of a Compaq SC-Cluster with four Alpha-EV68 processors (1 GHz, 8 MB Cache/CPU) and 32 GB RAM.

We set $r_0 = 1$, $d_0 = 1$, $M = 1$, $J = 5$, and

$$F = \frac{1}{5}\begin{pmatrix} \pi & 1 \\ -1 & \pi \end{pmatrix}.$$

to run Algorithm $(A^{pc,\,adapt}_{r_0,d_0,J,F,M})$ in Example 6.1. The a posteriori criterion of Proposition 4.2 enforced the algorithm to enlarge r_0 from 1 to 4 on the first level, i.e. for the largest d. Table 1 presents the errors

$$e = |f^{pc}_{d,r}(F) - f^{pc}(F)| \quad \text{and} \quad e' = |\lambda^F_{d,r} \cdot DT(F) - Df^{pc}(F)|,$$

the number of nodes in the set $\mathcal{N}$, the number of activated nodes in X, the theoretical number of nodes, i.e. the number of nodes in $\mathcal{N}_{d,r}$, and the CPU-time in seconds needed to compute $f^{pc}_{\mathcal{N}}(F)$.

d	e	e'	$\#X$	$\#\mathcal{N}$	$\#\mathcal{N}_{d,r}$	CPU-time
1	0.656 637	2.000 000	266	6,561	6,561	0.1 s
1/2	0.037 543	0.325 927	132	160	83,521	0.2 s
1/4	0.009 873	0.206 980	237	768	1,185,921	0.2 s
1/8	0.000 177	0.001 392	2,137	3,920	17,850,625	0.7 s
1/16	0.000 010	0.000 058	14,360	33,920	276,922,881	3.7 s

TABLE 1. Discretization parameter d, errors $e = |f^{pc}_{d,r}(F) - f^{pc}(F)|$ and $e' = |\lambda^F_{d,r} \cdot DT(F) - Df^{pc}(F)|$, number of active, possible, and theoretical nodes, and CPU-time needed to compute $f^{pc}_{d,r}(F)$ in Example 6.1.

We observe that e converges with experimental rate 4 to 0 and the experimental convergence rate for e' is better than linear. Due to the grid coarsening strategy and the active set strategy the number of activated nodes in X, i.e. the size of each linear optimization problem, is remarkably small when compared to the possible and theoretical numbers of nodes and the CPU-time needed to obtain an absolute error of about 10^{-5} is only 3.7 seconds. We obtained similar numbers e and e' for the more reliable choice $M = 100$ but the number of activated nodes and the CPU-time was significantly larger, e.g. 9.4 seconds were needed to achieve $e \leq 10^{-3}$.

To test $(A^{pc,\,adapt}_{r_0,d_0,J,F,M})$ in Example 6.2 we set $J = 6$, $r_0 = 4$, $d_0 = 1$, $M = 10$, and

$$F = \frac{1}{10}\begin{pmatrix} \sqrt{2}/3 & 1/3 \\ \sqrt{5}/5 & \sqrt{2}/3 \end{pmatrix}.$$

Table 2 presents the errors and the numbers of nodes as in the previous example. The error e converges with an experimental convergence rate 1.8, i.e, $e \approx d^{1.8}$, and there seems to be approximately linear decay in e' but convergence cannot be deduced. The choice $M = 10$ is very optimistic in this example and leads to very small sets $\mathcal{N}$.

d	e	e'	$\#X$	$\#\mathcal{N}$	$\#\mathcal{N}_{d,r}$	CPU-time
1	0.165 961	0.454 711	6,561	6,561	6,561	0.9 s
1/2	0.072 351	0.115 344	794	43,568	83,521	1.1 s
1/4	0.014 273	0.042 726	1,311	167,136	1,185,921	1.4 s
1/8	0.004 031	0.063 871	1,766	715,504	17,850,625	2.4 s
1/16	0.001 139	0.017 684	4,889	3,082,192	276,922,881	6.5 s
1/32	0.000 204	0.011 032	12,140	13,694,256	4,362,470,401	21.8 s

TABLE 2. Discretization parameter d, errors $e = |f^{pc}_{d,r}(F) - f^{pc}(F)|$ and $e' = |\lambda^F_{d,r} \cdot DT(F) - Df^{pc}(F)|$, number of active, possible, and theoretical nodes, and CPU-time needed to compute $f^{pc}_{d,r}(F)$ in Example 6.2.

We ran Algorithm $(A^{pc,\,adapt}_{r_0,d_0,J,F,M})$ in Example 6.3 with $M = 100$, $J = 6$, $r_0 = 4$, $d_0 = 1$, and

$$F = \frac{1}{5}\begin{pmatrix} \pi & 0 \\ 0 & \pi \end{pmatrix}.$$

Table 3 presents the errors e and e', the numbers of activated, possible, and theoretical nodes, and the CPU time as in the previous examples. We observe that e converges at least linearly to 0 while e' seems to converge linearly, at least for $d \leq 1/8$.

d	e	e'	$\#X$	$\#\mathcal{N}$	$\#\mathcal{N}_{d,r}$	CPU-time
1	0.328 922	0.162 697	81	81	81	0.1 s
1/2	0.095 387	0.458 318	457	625	625	0.2 s
1/4	0.024 647	0.160 972	689	6,561	6,561	0.3 s
1/8	0.000 806	0.167 216	1,276	83,521	83,521	0.6 s
1/16	0.000 389	0.079 923	1,618	513,280	1,185,921	1.2 s
1/32	0.000 180	0.032 843	2,213	2,128,048	17,850,625	3.6 s

TABLE 3. Discretization parameter d, errors $e = |f^{pc}_{d,r}(F) - f^{pc}(F)|$ and $e' = |\lambda^F_{d,r} \cdot DT(F) - Df^{pc}(F)|$, number of active, possible, and theoretical nodes, and CPU-time needed to compute $f^{pc}_{d,r}(F)$ in Example 6.3.

7. Numerical Experiments II

In this section we outline how Algorithm ($A^{pc,\,adapt}_{r_0,d_0,J,F,M}$) may be used for the effective numerical simulation of nonconvex vectorial variational problems and we report on two numerical experiments. The proposed algorithm aims to numerically relax and minimize variational problems of the form (M), i.e.

$$(M) \quad \text{Minimize} \quad I(u) := \int_\Omega f(\nabla u)\,dx \quad \text{among } u \in \mathcal{A} := \{v \in W^{1,p}(\Omega;\mathbb{R}^m) : v|_{\Gamma_D} = u_D\},$$

where f is continuous and satisfies p-growth conditions, $\Omega \subseteq \mathbb{R}^n$ is a bounded Lipschitz domain, $\Gamma_D \subseteq \partial\Omega$ is closed and of positive surface measure, and $u_D = \tilde{u}|_{\Gamma_D}$ for some $\tilde{u}_D \in C(\overline{\Omega};\mathbb{R}^m)$. We further suppose that Ω is polyhedral and let $\mathcal{T}$ be a regular triangulation of Ω such that Γ_D is matched exactly by edges (respectively faces) of elements in $\mathcal{T}$. We let $\mathcal{S}^1(\mathcal{T})^m$ denote the lowest order finite element space on $\mathcal{T}$ which consists of all globally continuous, $\mathcal{T}$-elementwise affine functions in $W^{1,p}(\Omega;\mathbb{R}^m)$. Finally, we let $\tilde{u}_{D,h}$ be the nodal interpolant of $\tilde{u}_D$ on $\mathcal{T}$. The following algorithm is capable of finding an approximation of a weak limit of an infimizing sequence for the nonconvex vectorial variational problem (M). The approximation scheme realizes a steepest descent approach and exploits the fact that Algorithm ($A^{pc,\,adapt}_{r_0,d_0,J,F,M}$) provides an approximation of $Df^{pc}(F)$.

Algorithm (A^{nvvp})**.** Input: $u_h^{(0)} \in \mathcal{S}^1(\mathcal{T})^m$ with $u_h^{(0)}|_{\Gamma_D} = \tilde{u}_{D,h}|_{\Gamma_D}$, parameters r_0, d_0, J, M for Algorithm ($A^{pc,\,adapt}_{r_0,d_0,J,\cdot,M}$), and a termination criterion $\delta > 0$.

(a) Set $j := 0$.

(b) Let $r_h \in \mathcal{S}^1(\mathcal{T})^m$ satisfy $r_h|_{\Gamma_D} = 0$ and

$$\int_\Omega \nabla r_h \cdot \nabla v_h\,dx = -\int_\Omega \big(\lambda_{d/4,r}^{\nabla u_h^{(j)}} \cdot DT(\nabla u_h^{(j)})\big) \cdot \nabla v_h\,dx$$

for all $v_h \in \mathcal{S}^1(\mathcal{T})^m$ with $v_h|_{\Gamma_D} = 0$.

(c) Compute $t^* \in [0,1]$ which is a local minimizer in $[0,1]$ for

$$t \mapsto \int_\Omega f^{pc}_{d,r}\big(\nabla(u_h^{(j)} + t r_h)\big)\,dx.$$

(d) Stop and set $u_h := u_h^{(j)}$ if $t^* < \delta$.

(e) Set $u_h^{(j+1)} := u_h^{(j)} + t^* r_h$, $j := j+1$, and go to (b).

Output: $u_h \in \mathcal{S}^1(\mathcal{T})^m$.

We specify f, Ω, $\tilde{u}_D$, and $\mathcal{T}$ in two examples. The first example involves affine boundary data on $\partial\Omega$.

Example 7.1 ([DW]). Set $n = m := 2$, $\Omega := (0,1)^2$, $\tilde{u}_D(x) := Bx + c$ for

$$B = \begin{pmatrix} 1/2 & 1/4 \\ -1/4 & 15/32 \end{pmatrix}$$

and $c := (0, 1/4)$, and let f be as in Example 6.2. Given an integer $k > 0$ let $h_k := 1/k$ and $\mathcal{T}_k$ be the triangulation of Ω that consists of $2k^2$ triangles which are halved squares of sidelength h and with diagonals parallel to $(1,1)$.

For $\ell = 1, 2, 3$ we ran Algorithm (A^{nvvp}) with $r_0 = 2$, $M = 10$, $J_\ell = 2 + \ell$, and $k_\ell = 2^{2+\ell}$ in Example 7.1. The initial $u_h^{(0)}$ was chosen as $u_h^{(0)} = \tilde{u}_{D,h} + \xi_h$ where $\xi_h \in \mathcal{S}^1(\mathcal{T})^2$ with $\xi_h|_{\Gamma_D} = 0$ is obtained from a linear interpolation of random values $\xi_h(z) \in [-h_\ell, h_\ell]^2$ in the free nodes z of $\mathcal{T}$. The termination criterion δ was set to $\delta = 0.03$ and Algorithm (A^{nvvp}) terminated after 175, 86, and 8 iterations for $\ell = 1$, 2, and 3, respectively. Table 4 displays the numerically relaxed energies for the output u_{h_ℓ} in Example 7.1 and compares them to values that we obtained when $f_{d,r}^{pc}(F)$ and $\lambda_{d,r}^F \cdot DT(F)$ are replaced by f and Df, f^{pc} and Df^{pc}, and a numerical approximation f_d^{rc} of f^{rc} in steps (b) and (c) of Algorithm (A^{nvvp}) (the numbers for f_d^{rc} are taken from [DW]). We observe that Algorithm (A^{nvvp}) significantly reduces the initial (numerically relaxed) energies shown in brackets in the fourth column of Table 4. Moreover, we may deduce from the numerical results that the (numerically relaxed) energies of the outputs of Algorithm (A^{nvvp}) converge to the optimal value $f^{pc}(B) = 0.199\,761$ for $(d, h) \to 0$. No experimental convergence can be deduced when the original function f or the numerically obtained approximation of the rank-1 convex envelope f_d^{rc} were employed.

h	f	f^{pc}	$f_{d,r}^{pc}$	f_d^{rc}
1/8	0.244 128	0.199 761	0.252 177 (0.313 327)	0.202 507
1/16	0.234 894	0.199 761	0.216 076 (0.274 192)	0.208 264
1/32	0.233 769	0.199 761	0.202 994 (0.261 244)	0.245 633

TABLE 4. Minimal energies for various approaches to the numerical solution of (M) in Example 7.1.

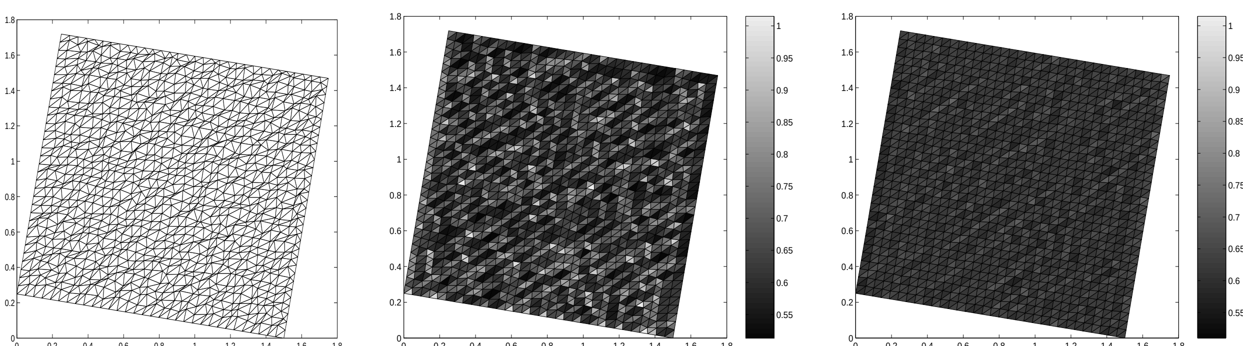

FIGURE 2. Initial deformation $u_h^{(0)}$, numerical solution for direct minimization employing f with modulus of the related stress field $|Df(\nabla u_h)|$, and numerical solution u_h obtained from numerical relaxation with Algorithm (A^{nvvp}) with stress field $|\lambda_{d/4,r}^{\nabla u_h} \cdot DT(\nabla u_h)|$ in Example 7.1 (from left to right).

Figure 2 displays the initial $u_h^{(0)}$, the numerical solution obtained from direct numerical minimization of (M), and the numerical solution obtained from the numerical relaxation realized by Algorithm (A^{nvvp}) in Example 7.1. We observe mesh-dependent oscillations in the stress field defined by the numerical solution when no relaxation is used while we observe a rather smooth stress field when the nonconvex vectorial variational problem is numerically polyconvexified.

The second example incorporates non-affine boundary conditions on Γ_D.

Example 7.2 ([DW]). Let n, m, Ω, Γ_D, f, and $\mathcal{T}$ be as in Example 7.1 and define for $x \in \overline{\Omega}$

$$\tilde{u}_D(x) := \frac{x - (1/2, 1/2)}{\sqrt{\left|x - (1/2, 1/2)\right|^2 + 1/4}}.$$

Table 5 displays the minimal energies for various approaches to the numerical simulation of (M) in Example 7.2. As in the previous example we observe that the results obtained by Algorithm (A^{nvvp}) (with the same parameters as in the previous experiment) with the approximated polyconvex envelope of f approach the value that we obtained with the exact polyconvex envelope f^{pc}. The minimal energies obtained with the nonrelaxed functional and with the discrete approximation of the rank-1 convex envelope of f (numbers for f_d^{rc} are taken from [DW]) do not show a convergent behavior. As in the previous experiment we observe in Figure 3 mesh-dependent oscillations in the numerical solution obtained from a direct minimization scheme. No significant oscillations can be found in the numerical solution computed with Algorithm (A^{nvvp}) and displayed in the right plot of Figure 3.

h	f	f^{pc}	$f_{d,r}^{pc}$	f_d^{rc}
1/8	0.186 939	0.170 023	0.231 421 (0.244 203)	0.175 216
1/16	0.189 875	0.168 945	0.186 094 (0.198 793)	0.173 716
1/32	0.188 556	0.168 583	0.173 779 (0.186 396)	0.183 599

TABLE 5. Minimal energies for various approaches to the numerical solution of (M) in Example 7.2.

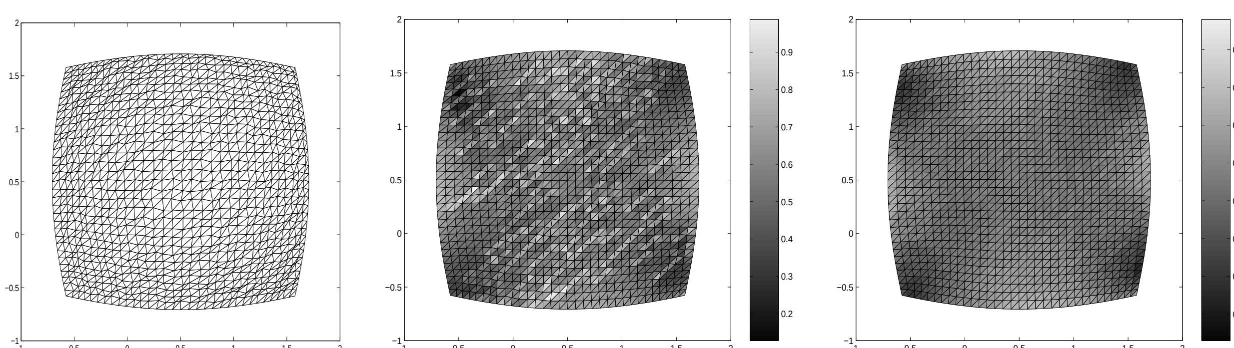

FIGURE 3. Initial deformation $u_h^{(0)}$, numerical solution for direct minimization employing f with modulus of the related stress field $|Df(\nabla u_h)|$, and numerical solution u_h obtained from numerical relaxation with Algorithm (A^{nvvp}) with stress field $|\lambda_{d/4,r}^{\nabla u_h} \cdot DT(\nabla u_h)|$ in Example 7.2 (from left to right).

Remarks 7.1. (i) A good initial $u_h^{(0)}$ may be obtained from solving (M) with f replaced by a convex function, e.g. $F \mapsto |F|^2/2$, as a pre-processing step in Algorithm (A^{nvvp}).
(ii) A post-processing procedure based on the algorithms in [Do, DW, Ba2] may be included

in Algorithm (A^{nvvp}) which approximates the rank-1 convex envelope applied to ∇u_h almost everywhere in Ω. Then, if (up to numerical tolerances) there holds $f^{pc}(\nabla u_h) = f^{rc}(\nabla u_h)$ almost everywhere in Ω one has that $f^{qc}(\nabla u_h) = f^{pc}(\nabla u_h)$ and (provided $f^{pc} \leq f^{qc} \leq f^{rc}$ are smooth enough) that $Df^{qc}(\nabla u_h) = Df^{pc}(\nabla u_h)$ almost everywhere in Ω. In this case, u_h serves as an approximation of a stationary point of the quasiconvex relaxation of (M).

(iii) The computation of $f^{pc}_{d,r}(\nabla u_h^{(j)})$ and $\lambda_{d,r}^{\nabla u_h^{(j)}}$ in steps (b) and (c) of Algorithm (A^{nvvp}) has to be done on each element of the triangulation $\mathcal{T}$ (since ∇u_h is $\mathcal{T}$-elementwise constant). This may be time-consuming but can be parallelized without communication costs.

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

References

[B1] Ball, J.M. (1977), Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Ration. Mech. Anal.* **63**, 337-403.

[B2] Ball, J.M. (1989), A version of the fundamental theorem for Young measures. *Partial differential equations and continuum models of phase transitions.* Eds. M Rascle, D. Serre, M. Slemrod: *Lecture Notes in Physics* **344**, 207—215.

[BJ] Ball, J.M., James, R.D. (1987), Fine phase mixtures as minimizers of energy. *Arch. Rational Mech. Anal.* **100**, 13—52.

[BKK] Ball, J.M., Kirchheim, B., Kristensen, J. (2000), Regularity of quasiconvex envelopes. *Calc. Var. Partial Differ. Equ.* **11**, 333-359.

[Ba1] Bartels, S. (2004), Error estimates for adaptive Young measure approximation. *SIAM J. Numer. Anal.* **42**, 505-529.

[Ba2] Bartels, S. (2002), Linear convergence in the approximation of rank-one convex envelopes. Submitted. Preprints of the DFG Priority Program "Multiscale Problems", No. 77 (www.mathematik.uni-stuttgart.de/˜mehrskalen/).

[BP] Bartels, S., Prohl, A. (2004), Multiscale resolution in the computation of crystalline microstructure. *Numer. Math.* **96**, 641-660.

[CP1] Carstensen, C., Plecháč, P. (1997), Numerical solution of the scalar double-well problem allowing microstructure. *Math. Comp.* **66**, 997-1026.

[CP2] Carstensen, C., Plecháč, P. (2000), Numerical analysis of compatible phase transitions in elastic solids. *SIAM J. Numer. Anal.* **37**, 2061-2081.

[CR] Carstensen, C., Roubiček, T. (2000), Numerical approximation of Young measures in non-convex variational problems. *Numer. Math.* **84**, 395-414.

[CM] Chipot, M., Müller, S. (1997), Sharp energy estimates for finite element approximations of non-convex problems. Max-Planck-Institut Leipzig, Preprint 1997-**8**.

[C] Clarke, F.H. (1990), *Optimization and Nonsmooth Analysis.* Classics in Applied Mathematics. SIAM.

[Da1] Dacorogna, B. (1987), A characterization of polyconvex, quasiconvex and rank one convex envelopes. *Integral functionals in calculus of variations, Proc. Int. Workshop*, Trieste/Italy 1985, Suppl. Rend. Circ. Mat. Palermo, II. Ser. 15, 37-58.

[Da2] Dacorogna, B. (1989), *Direct methods in the calculus of variations.* Applied Math. Sciences **78**, Springer-Verlag, Heidelberg.

[DH] Dacorogna, B., Haeberly, J.-P. (1998), Some numerical methods for the study of the convexity notions arising in the calculus of variations. *M2AN* **32**, 153–175.

[Do] Dolzmann, G. (1999), Numerical computation of rank-one convex envelopes. *SIAM J. Numer. Anal.* **36**, 1621-1635.

[DW] Dolzmann, G., Walkington, N.J. (2000), Estimates for numerical approximations of rank one convex envelopes. *Numer. Math.* **85**, 647-663,

[HH] Hackl, K., Hoppe, U. (2003), On the calculation of microstructures for inelastic materials using relaxed energies. *Solid Mech. Appl.* **108**, 77-86.

[Ko] Kohn, R. V. (1991), The relaxation of a double-well energy. *Contin. Mech. Thermodyn.* **3**, 193-236.

[KS] Kohn, R. V., Strang, G. (1986), Optimal design and relaxation of variational problems. I.-III. *Commun. Pure Appl. Math.* **39**, 353-377.

[Kr] Kružik, M. (1998), Numerical approach to double well problems. *SIAM J. Numer. Anal.* **35**, 1833-1849.

[L] Luskin, M. (1996), On the computation of crystalline microstructure. *Acta Numerica* **5**, 191-257.

[ML] Miehe, C., Lambrecht, M. (2003), Analysis of microstructure development in shearbands by energy relaxation of incremental stress potentials: large-strain theory for standard dissipative solids. *Internat. J. Numer. Methods Engrg.* **58**, 1-41.

[M] Müller, S., (1999), *Variational models for microstructure and phase transitions.* Springer. Lect. Notes Math. 1713, 85-210.

[NW1] Nicolaides, R. A., Walkington, N. J. (1993), Computation of microstructure utilizing Young measure representations, *Transactions of the Tenth Army Conference on Applied Mathematics and Computing* (West Point, NY, 1992), ARO Rep., 93-1, 57-68.

[NW2] Nicolaides, R. A., Walkington, N. J. (1995), Strong convergence of numerical solutions to degenerate variational problems. *Math. Comp.* **64**, 117-127.

[R1] Roubíček, T. (1996), Numerical approximation of relaxed variational problems. *J. Convex Anal.* **3**, 329-347.

[R2] Roubíček, T. (1997), *Relaxation in optimization theory and variational calculus.* De Gruyter Series in Nonlinear Analysis and Applications 4, New York.

Department of Mathematics, University of Maryland, College Park, MD 20742, USA
E-mail address: sba@math.umd.edu