RELIABLE AND EFFICIENT APPROXIMATION OF POLYCONVEX ENVELOPES

SÖREN BARTELS

Abstract. An iterative algorithm that approximates the polyconvex envelope $f^{pc}$ of a given function $f : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}$, 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

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

for a bounded Lipschitz domain $\Omega \subseteq \mathbb{R}^n$, $p \geq 1$, 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, 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
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 $F \in \mathbb{R}^{n \times m}$,

\begin{equation}
  f(F) \geq c_f |F|^p - c_f' .
\end{equation}

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

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

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 $\tau$ denotes its length; there holds $|T(A)| \leq c_T |A|_{\min(n,m)}$ if $| \cdot |_{\infty}$ denotes the maximum norm and $|A|_{\infty} \geq 1$. Choosing a set of points $N_{d,r} := d\mathbb{Z}^{n \times m} \cap B_r(0)$ for $r \geq d > 0$ and $B_r(0) := \{ A \in \mathbb{R}^{n \times m} : |A|_{\infty} < r \}$ such that $F \in \text{conv} N_{d,r}$, an approximation of $f^{pc}(F)$ reads

\[ f^{pc}_{d,r}(F) := \inf \left\{ \sum_{A \in N_{d,r}} \theta_A f(A) : \forall A \in N_{d,r}, \theta_A \geq 0, \sum_{A \in N_{d,r}} \theta_A = 1, \sum_{A \in N_{d,r}} \theta_A T(A) = T(F) \right\} . \]
The latter infimum defines a finite-dimensional linear optimization problem and admits a solution and a Lagrange multiplier \( \lambda_{d,r}^F \in \mathbb{R}^\ell \) associated to the constraint \( \sum_{A \in N_{d,r}} \theta_A T(A) = T(F) \). 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 \( f_{pc} \in C^{1,\alpha}_{loc}(\mathbb{R}^{nm}) \) together with explicit bounds on \( |f_{pc}|_{C^{1,\alpha}(B_d(F))} \).

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

\[
ct(n \land m) |\lambda_{d,r}^F| \leq pc_f r^{p-n\land m} \quad \text{and} \quad \ct |\lambda_{d,r}^F| r^{n\land m} - c_f r^p + c' \leq \lambda_{d,r}^F \cdot T(F) - f_{pc_d}(F)
\]

is satisfied, and \( f \in C^{1,\alpha}_{loc}(\mathbb{R}^{nm}) \) for some \( \alpha \in [0, 1] \). Then \( f_{pc_d}(F) = \tilde{f}_{pc_d}(F) \) for a polyconvex function \( \tilde{f}_{pc_d} : \mathbb{R}^{nm} \rightarrow \mathbb{R}, f_{pc_d}(F) = f_{pc_d}(F) \) for all \( s \geq r \), and there exists \( r' \geq r \) such that

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

If, additionally, \( \alpha > 0 \) and \( f_{pc} \in C^{1,\alpha}_{loc}(\mathbb{R}^{nm}) \) then

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

The constants \( c_1, c_2 > 0 \) only depend on \( n \) and \( m \).

It can be shown that \( \lambda_{d,r}^F \) and \( f_{pc_d}(F) \) remain bounded for \( r \to \infty \) so that the a posteriori condition of the theorem is satisfied if \( p > n \land m \) and if \( r \) is large enough. The direct computation of \( f_{pc_d}(F) \) 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 \( f_{pc_d}(F) \).

3. Preliminaries

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

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

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)

\[
\mathcal{N}_{d,r} := d\mathbb{Z}^{nm} \cap B_r(0) \subseteq \mathbb{R}^{nm};
\]

\( \mathbb{Z} \) denotes the set of all integers. We let \( \omega_{d,r} \) be the interior of the union of all closed \((nm)\)-dimensional cubes \( Q \subseteq B_r(0) \) with vertices in \( \mathcal{N}_{d,r} \), and define a uniform triangulation \( T_{d,r} \) of \( \omega_{d,r} \) by setting (cf. the left plot in Figure 1)

\[
T_{d,r} := \{ Q \subseteq \overline{\omega_{d,r}} : Q \text{ is a closed cube with vertices in } \mathcal{N}_{d,r} \text{ and edges of length } d \}.
\]
Note that each \( Q \in \mathcal{T}_{d,r} \) is the convex hull of \( 2^{nm} \) nodes \( M_1, \ldots, M_{2^{nm}} \in \mathcal{N}_{d,r} \), i.e. \( Q = \text{conv} \{ M_1, \ldots, M_{2^{nm}} \} \). To \( \mathcal{T}_{d,r} \) we associate the set of continuous, \( \mathcal{T}_{d,r} \)-elementwise \((nm)\)-linear functions

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

The nodal interpolation operator \( \mathcal{I}_{d,r} \) on \( \mathcal{T}_{d,r} \) is for \( v \in C(\overline{\mathcal{T}_{d,r}}) \) 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 \( c_T > 0 \) such that

\[
\| \mathcal{I}_{d,r} g - g \|_{L^{\infty}(\omega_{d,r})} \leq c_T d^{1+\alpha} |g|_{C^{1,\alpha}(B_r(0))}
\]

for \( \alpha \in (0,1] \) and \( g \in C_{\text{loc}}^{1,\alpha}(\mathbb{R}^{n \times m}) \) 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 \( \alpha \)-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 \( |g|_{C^{1,\alpha}(B_r(0))} = |g|_{\text{Lip}_r} |g|_{\text{Lip}(B_r(0))} \) of \( g \) on \( B_r(0) \) if \( \alpha = 0 \).

\[\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{node_set_diagram}
\caption{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).}
\end{figure}\]

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

\[
T(A) = \left( (A)_{1,1}, \ldots, (A)_{1,m}, (A)_{2,1}, \ldots, (A)_{2,m}, \ldots, (A)_{n,1}, \ldots, (A)_{n,m}, \text{adj}_2 A, \ldots, \text{adj}_{n \times m} A \right),
\]

where for \( 2 \leq \ell \leq n \land m = \min \{ n, m \} \), \( \text{adj}_\ell A \) is a vector containing all \( \ell \times \ell \) minors of \( A \) and

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

There exists \( c_T > 0 \) (which depends on \( n \) and \( m \)) such that \( |T(A)| \leq c_T |A|_{\infty}^{n \land m} \) for all \( A \in \mathbb{R}^{n \times m} \) with \( |A|_{\infty} \geq 1 \); for \( n = m = 2 \) we have \( \text{adj}_2 A = \text{det} A \) 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 $S^1(T_{d,r})$ define a rank-1 decomposition of a matrix $F \in \mathbb{F}_{d,r}$.

**Lemma 3.1.** Let $Q \in T_{d,r}$ and $M_1, ..., M_{2^{nm}} \in N_{d,r}$ be such that $Q = \text{conv} \{M_1, ..., M_{2^{nm}}\}$. Let $F \in Q$ and define $\theta_i := \varphi_{M_i}(F) \geq 0$, $i = 1, ..., 2^{nm}$. There holds

\begin{equation}
\sum_{i=1}^{2^{nm}} \theta_i = 1 \quad \text{and} \quad \sum_{i=1}^{2^{nm}} \theta_i T(M_i) = T(F).
\end{equation}

**Proof.** We construct convex-coefficients that satisfy (3.2) and then show that they equal $\varphi_{M_i}(F)$. 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 $F^0 := F$ and $\varrho^0 := 1$. Then, for $j = 1, ..., n$ and $k = 1, ..., m$ set $\ell := (j-1)m + k$ and define $F^\ell,2\ell-1, F^\ell,2\ell \in \mathbb{R}^{n \times m}$, $\ell = 1, ..., 2^{\ell-1}$, by setting, for $j' = 1, ..., n$ and $k' = 1, ..., m$,

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

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

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

\begin{equation}
\varrho^{\ell,2\ell-1} := \varrho^{\ell-1, \ell}(1 - (F^{\ell-1, \ell})_{j,k}), \quad \varrho^{\ell,2\ell} := \varrho^{\ell-1, \ell}(F^{\ell-1, \ell})_{j,k}.
\end{equation}

(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) $\{F^{nm, \ell} : \ell = 1, ..., 2^{nm}\} = \{0,1\}^{n \times m}$;

(ii) $\theta^\ell,2\ell-1, \theta^\ell,2\ell \geq 0$, $\theta^\ell,2\ell-1 + \theta^\ell,2\ell = 1$, and $F^{\ell-1, \ell} = \theta^\ell,2\ell-1 F^{\ell,2\ell-1} + \theta^\ell,2\ell F^{\ell,2\ell}$ for $\ell = 1, ..., nm$ and $i = 1, ..., 2^{\ell-1}$;

(iii) rank($F^{\ell,2\ell-1} - F^{\ell,2\ell}$) = 1 for $\ell = 1, ..., nm$ and $i = 1, ..., 2^{\ell-1}$;

(iv) $\varphi_{F^{nm, \ell}}(F) = \varrho^{nm, \ell}$ for $\ell = 1, ..., 2^{nm}$.

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

Set $\theta_i := \varrho^{nm, \ell}$ for $\ell = 1, ..., 2^{nm}$. The assertion of the lemma (for $Q = \text{conv} \{0,1\}^{n \times m}$) 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. \hfill $\Box$

### 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 $F \in \mathbb{F}_{n \times m}$ and assume that either $\alpha = 0$ and $f$ is locally Lipschitz continuous or $\alpha \in (0,1]$ and $f \in C^{1,\alpha}_{loc}(\mathbb{R}^{n \times m})$. 5
Proposition 4.1. There exists $r' = r'(F) > 0$ such that

$$|f_{pc}^-(F) - f_{dc,r'}^-(F)| \leq 2c_T d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'(0)})}.$$  

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

$$f_{pc}^-(F) \leq \sum_{\ell=1}^{\tau+1} \varrho_\ell f(A_\ell) \leq f_{pc}^-(F) + c_T d^{1+\alpha} |f|_{C^{1,\alpha}(B_{t}(0))}. \tag{4.1}$$

Choose $r' \geq t$ such that $A_1, \ldots, A_{\tau+1} \in \omega_{d,r'}$. For each $\ell = 1, \ldots, \tau + 1$ Lemma 3.1 guarantees the existence of $M^{(\ell)}(\cdot) \in N_{d,r'}$ and $\varphi_{M^{(\ell)}(A_\ell)} \geq 0$, $\ell = 1, \ldots, 2^{Nm}$, such that $\sum_{i=1}^{2^{Nm}} M_i^{(\ell)} = 1$, $\sum_{i=1}^{2^{Nm}} \varphi_{M_i^{(\ell)}}(M^{(\ell)}(A_\ell)) = T(A_\ell)$, and $A_\ell \in Q_\ell = \text{conv} \{M_1^{(\ell)}, \ldots, M_{2^{Nm}}^{(\ell)}\} \in T_{d,r'}$. For $A \in N_{d,r'}$ 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 $A \in N_{d,r'}$

$$\tilde{\theta}_A = \sum_{\ell=1}^{\tau+1} \sum_{i=1}^{2^{Nm}} \varrho_\ell \varphi_{M_i^{(\ell)}}(A_\ell), \tag{4.2}$$

defines a feasible $(\tilde{\theta}_A : A \in N_{d,r'})$ to compute $f_{dc,r'}^-(F)$ so that

$$f_{pc}^-(F) \leq f_{dc,r'}^-(F) = \sum_{A \in N_{d,r'}} \theta_A f(A) \leq \sum_{A \in N_{d,r'}} \tilde{\theta}_A f(A), \tag{4.3}$$

where $(\theta_A : A \in N_{d,r'})$ is optimal in $f_{dc,r'}^-(F)$. Since $|f|_{C^{1,\alpha}(B_{t}(0))} \leq |f|_{C^{1,\alpha}(B_{r'}(0))}$ and $\theta_{M_i^{(\ell)}} = \varphi_{M_i^{(\ell)}}(A_\ell)$, estimates (3.1) and (4.1)-(4.3) imply

$$0 \leq f_{dc,r'}^-(F) - f_{pc}^-(F)$$

$$\leq \sum_{A \in N_{d,r'}} \tilde{\theta}_A f(A) - \sum_{\ell=1}^{\tau+1} \varrho_\ell f(A_\ell) + c_T d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))}$$

$$= \sum_{\ell=1}^{\tau+1} \varrho_\ell \left( \sum_{i=1}^{2^{Nm}} \varphi_{M_i^{(\ell)}}(M^{(\ell)}(A_\ell)) - f(A_\ell) \right) + c_T d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))}$$

$$= \sum_{\ell=1}^{\tau+1} \varrho_\ell \left( I_{d,r'} f(A_\ell) - f(A_\ell) \right) + c_T d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))}$$

$$\leq 2c_T d^{1+\alpha} |f|_{C^{1,\alpha}(B_{r'}(0))} \tag{4.4}$$

which proves the proposition. \hfill \square

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 $f_{dc,r'}^-(F)$ and which we will refer to through $f_{dc,r'}^-(F)$. In particular, the lemma characterizes the Lagrange multiplier $\lambda_{d,r}^F$ 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 $\lambda_{d,r}^F \in \mathbb{R}^T$ such that

$$
\max_{\mathcal{A} \in \mathcal{N}_{d,r}} (\lambda_{d,r}^F \cdot T(A) - f(A)) \leq \lambda_{d,r}^F \cdot T(F) - f_{d,r}^{pc}(F).
$$

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

Employing the optimality conditions of Lemma 4.1 we can state sufficient conditions that ensure that $p > n \wedge m$ is large enough so that $f_{d,r}^{pc}(F) = f_{d,r}^{pc}(F)$, 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 $\lambda_{d,r}^F \in \mathbb{R}^T$ be as in Lemma 4.1, suppose $p \geq n \wedge m$, $r \geq 1$, and

$$
|\lambda_{d,r}^F| \leq \rho c r^{p-n \wedge m} \text{ and } c_T|\lambda_{d,r}^F|^p + c_f \leq \lambda_{d,r}^F \cdot T(A) - f(A). \tag{4.6}
$$

Then, there holds

$$
\max_{\mathcal{A} \in \mathbb{R}^{n \times m}} (\lambda_{d,r}^F \cdot T(A) - f(A)) \leq \lambda_{d,r}^F \cdot T(F) - f_{d,r}^{pc}(F)
$$

and $f_{d,r}^{pc}(F) = f_{d,r}^{pc}(F)$ for all $r' \geq r$.

**Proof.** Since $c_T(n \wedge m)|\lambda_{d,r}^F| \leq \rho c r^{p-n \wedge m}$ and $p \geq n \wedge m$ the mapping

$$
\{ t \in \mathbb{R} : t \geq r \} \to \mathbb{R}, \quad t \mapsto c_T + c_f t^{p-n \wedge m},
$$

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

$$
\lambda_{d,r}^F \cdot T(A) - f(A) \leq c_T|\lambda_{d,r}^F||A|^n_{\infty} - c_f|A|^p_{\infty} + c_f \leq c_T|\lambda_{d,r}^F|^p r^{n \wedge m} - c_f r^p + c_f
$$

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

$$
f(A) \geq \lambda_{d,r}^F \cdot T(A) - f_{d,r}^{pc}(F)
$$

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

$$
f(A) \geq \lambda_{d,r}^F \cdot T(A) - f_{d,r}^{pc}(F) = f_{d,r}^{pc}(F).
$$

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

$$
f_{d,r'}^{pc}(F) = \sum_{\mathcal{A} \in \mathcal{N}_{d,r'}} \tilde{\theta}_A f(A) \geq \sum_{\mathcal{A} \in \mathcal{N}_{d,r'}} \tilde{\theta}_A (\lambda_{d,r}^F \cdot T(A) - f_{d,r}^{pc}(F)) = f_{d,r}^{pc}(F).
$$

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

We now turn to estimates for $\lambda_{d,r}^F$ for which we need to construct a polyconvex extension of $f_{d,r}^{pc}(F)$ 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 $\tilde{f}_d|_{Q}$ is $(nm)$-linear for each cube $Q \subseteq \mathbb{R}^{n \times m}$ 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 $A \in \mathbb{R}^{n \times m}$,
\[
\tilde{f}_d^{pc}(A) = \inf \left\{ \sum_{\ell=1}^{\tau+1} q_{\ell} \tilde{f}_d(A_{\ell}) : A_{\ell} \in \mathbb{R}^{n \times m}, \, q_{\ell} \geq 0, \, \sum_{\ell=1}^{\tau+1} q_{\ell} = 1, \, \sum_{\ell=1}^{\tau+1} q_{\ell} T(A_{\ell}) = T(A) \right\}.
\]

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 $B_{\kappa} \in d \mathbb{Z}^{n \times m}$, $\gamma_\kappa \geq 0$, $\kappa = 1, \ldots, 2^{nm}(\tau + 1)$, 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
\[
(4.8) \quad \tilde{f}_d^{pc}(A) \leq \sum_{\kappa=1}^{2^{nm}(\tau+1)} \gamma_\kappa f(B_{\kappa}) \leq \tilde{f}_d^{pc}(A) + \varepsilon.
\]

**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 $p > n \land m$. 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 $q_{\ell} \geq 0$, $\ell = 1, \ldots, \tau+1$, such that $\sum_{\ell=1}^{\tau+1} q_{\ell} = 1$, $\sum_{\ell=1}^{\tau+1} q_{\ell} T(A_{\ell}) = T(A)$, and
\[
\sum_{\ell=1}^{\tau+1} q_{\ell} \tilde{f}_d(A_{\ell}) \leq \tilde{f}_d^{pc}(A) + \varepsilon.
\]

For each $\ell = 1, \ldots, \tau+1$ let $\tilde{Q}_\ell = \text{conv} \left\{ M_{1,\ell}^{(\ell)}, \ldots, M_{2^{nm}}^{(\ell)} \right\}$ be a cube in $\mathbb{R}^{n \times m}$ with vertices $M_{1,\ell}^{(\ell)}, \ldots, M_{2^{nm}}^{(\ell)} \in d \mathbb{Z}^{n \times m}$, 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 $\theta_{M_{i}^{(\ell)}} \geq 0$, $\ell = 1, \ldots, 2^{nm}$, such that $\sum_{i=1}^{2^{nm}} \theta_{M_{i}^{(\ell)}} = 1$, $\sum_{i=1}^{2^{nm}} \theta_{M_{i}^{(\ell)}} T(M_{i}^{(\ell)}) = T(A_{\ell})$, and
\[
\tilde{f}_d(A_{\ell}) = \sum_{i=1}^{2^{nm}} \varphi_{M_{i}^{(\ell)}}(A_{\ell}) f(M_{i}^{(\ell)}) = \sum_{i=1}^{2^{nm}} \theta_{M_{i}^{(\ell)}} f(M_{i}^{(\ell)}).
\]

This implies
\[
\sum_{\ell=1}^{\tau+1} \sum_{i=1}^{2^{nm}} q_{\ell} \theta_{M_{i}^{(\ell)}} f(M_{i}^{(\ell)}) \leq \tilde{f}_d^{pc}(A) + \varepsilon
\]

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

The following assertion is due to Ball [B1].

**Lemma 4.3.** There exist convex functions $\hat{f}, \tilde{f}_d : \mathbb{R}^r \to \mathbb{R}$ such that
\[
f^{pc} = \hat{f} \circ T \quad \text{and} \quad \tilde{f}_d^{pc} = \tilde{f}_d \circ T.
\]
For \( g = f \) or \( g = \hat{f}_d \) the function \( \hat{g} = \hat{f} \) or \( \hat{g} = \hat{f}_d \), respectively, can be defined by

\[
\hat{g}(X) = \inf \left\{ \sum_{\ell=1}^{\tau+1} \varrho_\ell g(A_\ell) : A_\ell \in \mathbb{R}^{n \times m}, \varrho_\ell \geq 0, \sum_{\ell=1}^{\tau+1} \varrho_\ell = 1, \sum_{\ell=1}^{\tau+1} \varrho_\ell T(A_\ell) = X \right\}. \quad \Box
\]

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 \( B^r = \{E_1, \ldots, E_r\} \) be the canonical basis in \( \mathbb{R}^r \). There exists \( r' > 0 \) such that, for all \( E \in \pm B^r \), there holds

\[
|\hat{f}_d(T(F) + dE) - \hat{f}(T(F) + dE)| \leq 2c_Td^{1+\alpha}|f|_{C^{1,\alpha}(B_{r'}(0))}.
\]

(ii) Let \( B^{n \times m} = \{E_1, \ldots, E_{nm}\} \) 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

\[
|\hat{f}_d(T(F) + dE) - f^{pc}(F + dE)| \leq 2c_Td^{1+\alpha}|f|_{C^{1,\alpha}(B_{r'}(0))}.
\]

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

\[
\hat{f}(T(F) + dE) \leq \sum_{\ell=1}^{\tau+1} \varrho_\ell f(A_\ell) \leq \hat{f}(T(F) + dE) + c_Td^{1+\alpha}|f|_{C^{1,\alpha}(B_{t}(0))}
\]
as well as

\[
\hat{f}_d(T(F) + dE) \leq \sum_{\ell=1}^{\tau+1} \varrho^{(d)}_\ell f(A^{(d)}_\ell) \leq \hat{f}_d(T(F) + dE) + c_Td^{1+\alpha}|f|_{C^{1,\alpha}(B_{t}(0))}.
\]

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

\[
0 \leq \hat{f}(T(F) + dE) - \hat{f}_d(T(F) + dE)
\]

\[
\leq \sum_{\ell=1}^{\tau+1} \varrho^{(d)}_\ell \left( f(A^{(d)}_\ell) - f_d(A^{(d)}_\ell) \right) + c_Td^{1+\alpha}|f|_{C^{1,\alpha}(B_{r'}(0))}
\]

\[
\leq 2c_Td^{1+\alpha}|f|_{C^{1,\alpha}(B_{r'}(0))}.
\]

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

\[
0 \leq \hat{f}_d(T(F) + dE) - \hat{f}(T(F) + dE)
\]

\[
\leq \sum_{\ell=1}^{\tau+1} \varrho_\ell \left( f_d(A_\ell) - f(A_\ell) \right) + c_Td^{1+\alpha}|f|_{C^{1,\alpha}(B_{r'}(0))}
\]

\[
\leq 2c_Td^{1+\alpha}|f|_{C^{1,\alpha}(B_{r'}(0))}.
\]

Choosing \( r' \) maximal so that for each \( E \in \pm B^r \) 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 \( \hat{f} \circ T = f^{pc} \). \quad \Box
The next lemma is the key observation for the estimates for $\lambda^F_{d,r}$ 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

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

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 $\partial h(v_0) \subseteq [(h(v_0) - h(v_0 - s))/s, (h(v_0 + s) - h(v_0))/s]$ for all $s > 0$.

(iii) If $h(v_0) \leq h(v)$ for all $v \in \mathbb{R}^\ell$ then $0 \in \partial h(v_0)$.

**Lemma 4.5.** Suppose that

$$\max_{A \in \mathbb{R}^{n \times m}} \left( \lambda^F_{d,r} \cdot T(A) - f(A) \right) \leq \lambda^F_{d,r} \cdot T(F) - f^pc_{d,r}(F).$$

Then there holds $f^pc_{d}(F) = f^pc_{d,r}(F)$ and $\lambda^F_{d,r} \in \partial f_d(T(F))$.

**Proof.** We show that $f^pc_{d}(F) = f^pc_{d,r}(F)$ and

$$\max_{A \in \mathbb{R}^{n \times m}} \left( \lambda^F_{d,r} \cdot T(A) - f^pc_{d}(A) \right) = \lambda^F_{d,r} \cdot T(F) - f^pc_{d}(F).$$

Then, the asserted inclusion is deduced from these observations as follows: Let $\varepsilon > 0$. For $X \in \mathbb{R}^7$ there exist $A_\ell \in \mathbb{R}^{n \times m}, \varrho_\ell \geq 0, \ell = 1, \ldots, \tau + 1$, such that $\sum_{\ell=1}^{\tau+1} \varrho_\ell = 1$, $\sum_{\ell=1}^{\tau+1} \varrho_\ell T(A_\ell) = X$, and

$$\hat{f}_d(X) \geq \sum_{\ell=1}^{\tau+1} \varrho_\ell \hat{f}_d(A_\ell) - \varepsilon.$$ 

Using $\tilde{f}^pc_{d} \leq \hat{f}_d$, $\tilde{f}^pc_{d}(F) = \hat{f}_d(T(F))$, and (4.9) we deduce

$$\hat{f}_d(X) - \lambda^F_{d,r} \cdot X \geq \sum_{\ell=1}^{\tau+1} \varrho_\ell (\tilde{f}^pc_{d}(A_\ell) - \tilde{f}^pc_{d,r}(A_\ell)) - \varepsilon$$

$$\geq \sum_{\ell=1}^{\tau+1} \varrho_\ell (\tilde{f}^pc_{d}(A_\ell) - \lambda^F_{d,r} \cdot T(A_\ell)) - \varepsilon$$

$$\geq \tilde{f}^pc_{d}(F) - \lambda^F_{d,r} \cdot T(F) - \varepsilon$$

$$= \hat{f}_d(T(F)) - \lambda^F_{d,r} \cdot T(F) - \varepsilon.$$ 

By arbitrariness of $\varepsilon > 0$, the convex function $X \mapsto \hat{f}_d(X) - \lambda^F_{d,r} \cdot X, X \in \mathbb{R}^7$, has a minimum in $T(F)$ so that (iii) in Remark 4.4 implies the asserted inclusion.

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

$$\sum_{\kappa=1}^{2^m(\tau+1)} \gamma_\kappa f(B_\kappa) \leq \tilde{f}^pc_{d}(F) + \varepsilon.$$
The hypothesis of the lemma implies, for $\kappa = 1, \ldots, 2^{nm}(\tau + 1)$,
\[
f(B_\kappa) \geq \lambda_{d,r}^F \cdot T(B_\kappa) - \lambda_{d,r}^F \cdot T(F) + f_{d,r}^{pc}(F)
\]
and this estimate yields
\[
\tilde{f}_{d}^{pc}(F) \geq \sum_{\kappa=1}^{2^{nm}(\tau+1)} \gamma_\kappa f(B_\kappa) - \varepsilon \geq f_{d,r}^{pc}(F) - \varepsilon,
\]
which, by arbitrariness of $\varepsilon > 0$, shows $\tilde{f}_{d}^{pc}(F) \geq f_{d,r}^{pc}(F)$ and hence yields $\tilde{f}_{d}^{pc}(F) = f_{d,r}^{pc}(F)$.

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 $\delta_\kappa \geq 0$, $\kappa = 1, \ldots, 2^{nm}(\tau + 1)$ such that $\sum_{\kappa=1}^{2^{nm}(\tau+1)} \delta_\kappa = 1$.

Then, the hypothesis of the lemma and $\tilde{f}_{d}^{pc}(F) = f_{d,r}^{pc}(F)$ imply
\[
\max_{A \in \mathbb{R}^{n \times m}} \left( \lambda_{d,r}^F \cdot T(A) - f_{d}^{pc}(A) \right) \leq \lambda_{d,r}^F \cdot T(A^*) - \tilde{f}_{d}^{pc}(A^*)
\]
\[
\leq \sum_{\kappa=1}^{2^{nm}(\tau+1)} \delta_\kappa \left( \lambda_{d,r}^F \cdot T(C_\kappa) - f(C_\kappa) \right) + \varepsilon
\]
\[
\leq \sum_{\kappa=1}^{2^{nm}(\tau+1)} \delta_\kappa \max_{A \in d\mathbb{Z}^{n \times m}} \left( \lambda_{d,r}^F \cdot T(A) - f(A) \right) + \varepsilon
\]
\[
\leq \lambda_{d,r}^F \cdot T(F) - f_{d,r}^{pc}(F) + \varepsilon
\]
\[
= \lambda_{d,r}^F \cdot T(F) - \tilde{f}_{d}^{pc}(F) + \varepsilon,
\]
which, by arbitrariness of $\varepsilon > 0$, is (4.9) and therefore concludes the proof. \qed

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

**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
\[
|\lambda_{d,r}^F - D \hat{f}(T(F))| \leq 4 \sqrt{\tau} d^{\alpha} |f|_{C^{1,\alpha}(B_{r}(0))} + \sqrt{\tau} d^{\alpha} |\hat{f}|_{C^{1,\alpha}(B_{d}(T(F)))}.
\]

**Proof.** Let $B' = \{ E_1, \ldots, E_r \}$ be the canonical basis in $\mathbb{R}^\tau$. Lemma 4.5 proves
\[
|\lambda_{d,r}^F - D \hat{f}(T(F))| \leq \sup_{S \in \partial f_d(T(F))} |S - D \hat{f}(T(F))|
\]
\[
\leq \sqrt{\tau} \sup_{S \in \partial f_d(T(F))} |S - D \hat{f}(T(F))|_{\infty}
\]
\[
= \sqrt{\tau} \sup_{S \in \partial \tilde{f}_d(T(F))} \max_{E \in \pm B'} |(S - D \hat{f}(T(F))) \cdot E|.
\]


Let $S \in \partial \hat{f}_d(T(F))$ and $E \in \pm B^r$. 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

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

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

$$S_+(E) \leq \hat{f}(T(F) + dE) - \hat{f}(T(F)) + 4c_T d^{1+\alpha} |f|_{C^{1,\alpha}(B_{t}(0))} \over d$$

and

$$S_-(E) \geq \hat{f}(T(F)) - \hat{f}(T(F) - dE) - 4c_T d^{1+\alpha} |f|_{C^{1,\alpha}(B_{t}(0))} \over d.$$ 

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

$$\left| (S - D\hat{f}(T(F))) \cdot E \right| \leq 4c_T d^{\alpha} |f|_{C^{1,\alpha}(B_{t}(0))} + d^{\alpha} |\hat{f}|_{C^{1,\alpha}(B_{d}(T(F)))}$$

which concludes the proof. \qed

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

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

or

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

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

$$|\lambda_{d,r}^p \cdot DT(F) - Df^{pc}(F)| \leq 4 \sqrt{nm} c_T d^\alpha |f|_{C^{1,\alpha}(B_{t}(0))} + \sqrt{nm} d^\alpha |f^{pc}|_{C^{1,\alpha}(B_{d}(F))}.$$

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

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

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 $(S \cdot DT(F)) \cdot E \in [S_-(E), S_+(E)]$ (cf. Remark 4.4) for

$$S_\pm(E) = \pm \hat{f}_d(T(F \pm dE) - \hat{f}_d(T(F)) \over d.$$
Assuming $S_-(E) \leq S_+(E)$, Lemma 4.4, $\hat{f}(T(F)) = f_{d,r}^{pc}(F)$, and Proposition 4.1 show

$$S_+(E) \leq \frac{f_{pc}(F + dE) - f_{pc}(F) + 4cTd^{\alpha+1}|f|_{C^{1,\alpha}(B_\epsilon(0))}}{d}$$

and

$$S_-(E) \geq \frac{f_{pc}(F) - f_{pc}(F - dE) - 4cTd^{\alpha+1}|f|_{C^{1,\alpha}(B_\epsilon(0))}}{d}.$$ 

These estimates, the mean value theorem, and Hölder continuity of $Df_{pc}$ imply

$$|(S \cdot DT(F) - Df_{pc}(F)) \cdot E| \leq 4cTd^{\alpha}|f|_{C^{1,\alpha}(B_\epsilon(0))} + d^\alpha|f_{pc}|_{C^{1,\alpha}(B_d(F))}$$

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** $f_{d,r}^{pc}(F)$

As mentioned in the introduction, the direct computation of $f_{d,r}^{pc}(F)$ 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 $N_{d,r}$ to obtain a set of nodes that leads to a good approximation of $f_{d/2,r}^{pc}(F)$ provided that we computed (an approximation of) $f_{d,r}^{pc}(F)$. The first assertion allows to remove nodes $A \in N_{d/2,r}$ that cannot lead to volume fractions $\theta_A$ larger than a given threshold $\delta = c/M$ for a known constant $c$ and a (large) parameter $M$.

**Proposition 5.1.** Let $(\theta_A : A \in N_{d,r})$ be feasible and optimal for $f_{d,r}^{pc}(F)$ with corresponding multiplier $\lambda_{d,r}^F \in \mathbb{R}^r$. For each $A \in N_{d,r}$ let $M(A) \geq M > 0$ and set

$$Z = \{A \in N_{d,r} : \lambda_{d,r}^F \cdot T(A) - f(A) \leq \lambda_{d,r}^F \cdot T(F) - f_{d,r}^{pc}(F) - d M(A)\}.$$

Let

$$Z' = \{A' \in N_{d/2,r} : \exists A \in Z, |A - A'| \leq d\}.$$ 

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

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

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

$$f(A') \geq f(A) - d|f|_{Lip,r}$$

$$\geq f_{d,r}^{pc}(F) + \lambda_{d,r}^F \cdot T(A) - \lambda_{d,r}^F \cdot T(F) + d M(A) - d|f|_{Lip,r}$$

$$\geq f_{d,r}^{pc}(F) + \lambda_{d,r}^F \cdot T(A') - \lambda_{d,r}^F \cdot T(F) + d M(A) - d|f|_{Lip,r} - d|\lambda_{d,r}^F||T|_{Lip,r}.$$ 

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

$$f(B') \geq f_{d,r}^{pc}(F) + \lambda_{d,r}^F \cdot T(B') - \lambda_{d,r}^F \cdot T(F) - d|f|_{Lip,r} - d|\lambda_{d,r}^F||T|_{Lip,r}.$$
Let \((\theta_{A'} : A' \in \mathcal{N}_{d/2,r})\) be feasible and optimal for \(f_{d/2,r}^{pe}(F)\). The previous two estimates imply

\[
f_{d/2,r}^{pe}(F) = \sum_{A' \in \mathcal{N}_{d/2,r}} \theta_{A'} f(A') = \sum_{A' \in Z'} \theta_{A'} f(A') + \sum_{B' \in \mathcal{N}_{d/2,r} \setminus Z'} \theta_{B'} f(B') \\
\geq f_{d/2,r}^{pe}(F) - (|f|_{\text{Lip},r} + |\lambda_{d,r}^F||T|_{\text{Lip},r})d + \sum_{A' \in Z'} \theta_{A'} d_r.
\]

Using \(f_{d/2,r}^{pe}(F) \leq f_{d/2,r}^{pe}(F)\) we deduce \(\theta_{A'} \leq (|f|_{\text{Lip},r} + |\lambda_{d,r}^F||T|_{\text{Lip},r})/M\). □

A more efficient and even more reliable assertion can be formulated if we have explicit estimates for \(|f_{d/2,r}^{pe}(F) - f_{d/2,r}^{pe}(F)|\) and \(|\lambda_{d,r}^F - \lambda_{d/2,r}^F|\). Sufficient for this is the knowledge of \(r'\) in Theorem A and \(\hat{f} \in C^{1,1}_0(\mathbb{R}^s)\) 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 \(|\lambda_{d,r}^F - \lambda_{d/2,r}^F| \leq c_{LM} d\), and let \(Z\) and \(Z'\) be as in Proposition 5.1. Assume that for all \(A \in \mathcal{N}_{d,r}\) there holds

\[
c_{LM}|T|_{\text{Lip},r'} + |\lambda_{d,r}^F||T|_{\text{Lip},r'} + 4c_{T} |f|_{\text{Lip},r'} + c_{LM}|T(A)| + c_{LM}|T(F)| < M(A).
\]

Then, for any \((\theta_A : A \in \mathcal{N}_{d/2,r})\) that is feasible and optimal for \(f_{d/2,r}^{pe}(F)\) 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

\[
\lambda_{d/2,r}^F \cdot T(A') - f(A') \\
= \lambda_{d,r}^F \cdot T(A) - f(A) + (\lambda_{d/2,r}^F - \lambda_{d,r}^F) \cdot T(A') + \lambda_{d,r}^F \cdot ((T(A') - T(A)) + (f(A) - f(A'))) \\
\leq \lambda_{d,r}^F \cdot T(A) - f(A) + c_{LM} d|T(A')| + d|\lambda_{d,r}^F||T|_{\text{Lip},r'} + d|f|_{\text{Lip},r'} \\
\leq \lambda_{d,r}^F \cdot T(F) - f_{d/2,r}^{pe}(F) - dM(A) + c_{LM} d|T(A')| + d|\lambda_{d,r}^F||T|_{\text{Lip},r'} + d|f|_{\text{Lip},r'}.
\]

The definitions of \(Z\) and \(Z'\), Proposition 4.1, and again the assumed estimate for \(|\lambda_{d,r}^F - \lambda_{d/2,r}^F|\) show

\[
\lambda_{d,r}^F \cdot T(F) - f_{d/2,r}^{pe}(F) - dM(A) + c_{LM} d|T(A')| + d|\lambda_{d,r}^F||T|_{\text{Lip},r'} + d|f|_{\text{Lip},r'} \\
= \lambda_{d/2,r}^F \cdot T(F) - f_{d/2,r}^{pe}(F) + (\lambda_{d,r}^F - \lambda_{d/2,r}^F) \cdot T(F) + (f_{d/2,r}^{pe}(F) - f_{d/2,r}^{pe}(F)) \\
- dM(A) + c_{LM} d|T(A')| + d|\lambda_{d,r}^F||T|_{\text{Lip},r'} + 2c_{T} d|f|_{\text{Lip},r'} \\
\leq \lambda_{d/r}^F \cdot T(F) - f_{d/2,r}^{pe}(F) + c_{LM} d|T(F)| + d|\lambda_{d,r}^F||T|_{\text{Lip},r'} - dM(A) \\
+ 2c_{T} d|f|_{\text{Lip},r'} + c_{LM} d|T(F)| + 2c_{T} d|f|_{\text{Lip},r'}.
\]

Employing once more Lipschitz continuity of \(T\) proves

\[
\lambda_{d/2,r}^F \cdot T(F) - f_{d/2,r}^{pe}(F) \\
+ d(c_{LM}|T(A')| + |\lambda_{d,r}^F||T|_{\text{Lip},r'} + 2c_{T} d|f|_{\text{Lip},r'} + c_{LM}|T(F)| - M(A)) \\
\leq \lambda_{d/2,r}^F \cdot T(F) - f_{d/2,r}^{pe}(F) \\
+ d(c_{LM}|T|_{\text{Lip},r'} + c_{LM}|T(A)| + |\lambda_{d,r}^F||T|_{\text{Lip},r'} + 4c_{T} d|f|_{\text{Lip},r'} + c_{LM}|T(F)| - M(A)).
\]
In view of (5.1) the last three estimates imply, for all $A' \in Z'$,

$$\lambda_{d/2,r}^F \cdot T(A') - f(A') < \lambda_{d/2,r}^F \cdot T(F) - f_{d/2,r}^{pc}(F).$$

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

$$\lambda_{d/2,r}^F \cdot T(A) - f(A) \leq \lambda_{d/2,r}^F \cdot T(F) - f_{d/2,r}^{pc}(F).$$

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

$$f_{d/2,r}^{pc}(F) = \sum_{A \in \mathcal{N}_{d/2,r}} \theta_A f(A)$$

$$> \sum_{A \in \mathcal{N}_{d/2,r}} \theta_A (\lambda_{d/2,r}^F \cdot T(A') - \lambda_{d/2,r}^F \cdot T(F) + f_{d/2,r}^{pc}(F)) = f_{d/2,r}^{pc}(F).$$

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

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. $\mathcal{N} \subseteq \mathcal{N}_{d/2,r}$ 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 $\mathcal{N} \subseteq \mathbb{R}^{nxm}$ we define

$$f_{N}^{pc}(F) := \min \left\{ \sum_{A \in \mathcal{N}} \theta_A f(A) : \forall A \in \mathcal{N}, \theta_A \geq 0, \sum_{A \in \mathcal{N}} \theta_A = 1, \sum_{A \in \mathcal{N}} \theta_A T(A) = T(F) \right\}.$$

Optimality conditions for $f_{N}^{pc}(F)$ guarantee the existence of some $\lambda_{N}^F \in \mathbb{R}^r$ such that

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

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

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

$$f_{N,X}^{pc}(F) := \min \left\{ \sum_{A \in X} \theta_A f(A) : \forall A \in X, \theta_A \geq 0, \sum_{A \in X} \theta_A = 1, \sum_{A \in X} \theta_A T(A) = T(F) \right\}.$$

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

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

$$X = \left\{ A \in \mathcal{N} : \lambda^F \cdot T(A) - f(A) \geq \max_{A \in \mathcal{N}} (\lambda^F \cdot T(A') - f(A')) - \varepsilon_{AS} \right\}$$

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

15
Proof. The optimality conditions (5.5) show (cf. (5.4) in the proof of Proposition 5.2), for all \( A \in \mathcal{N} \),

\[
\theta_A > 0 \quad \implies \quad A \in Y := \{ A' \in \mathcal{N} : \lambda^F_N \cdot T(A') - f(A') = \lambda^F_N \cdot T(F) - f^{pc}_{N}(F) \}.
\]

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

\[
\tilde{\lambda}^F \cdot T(A) - f(A) \geq \lambda^F_N \cdot T(A) - f(A) - \varepsilon_{AS}/2
\]

\[
= \lambda^F_N \cdot T(F) - f^{pc}_{N}(F) - \varepsilon_{AS}/2
\]

\[
= \max_{A' \in \mathcal{N}} (\lambda^F_N \cdot T(A') - f(A')) - \varepsilon_{AS}/2
\]

\[
\geq \max_{A' \in \mathcal{N}} (\tilde{\lambda}^F \cdot T(A') - f(A')) - \varepsilon_{AS},
\]

i.e. \( A \in X \). \( \square \)

Given some \( \tilde{\lambda}^F \) we do in general not know \( \varepsilon_{AS} \) to define \( X \) as in the lemma. We may however enlarge \( \varepsilon_{AS} \) 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 for \( f_{d_0/2^J, r_0}^pc(F) \) 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 (\( A^pc_{\text{adap}} \)).**

(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

\[
X := \left\{ A \in \mathcal{N} : \tilde{\lambda}^F \cdot T(A) - f(A) \geq \max_{A' \in \mathcal{N}} (\tilde{\lambda}^F \cdot T(A') - f(A')) - \varepsilon_{AS} \right\}
\]

\[
\cup \{ A \in \mathcal{N}_{d,r} : |F - A| \leq d \}.
\]

(c) Compute \( f^pc_{\mathcal{N},X}(F) \) and obtain a multiplier \( \lambda^F_N \in \mathbb{R}^T \).

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

\[
\lambda^F_N \cdot T(A) - f(A) \leq \lambda^F_N \cdot T(F) - f^{pc}_{N}(F)
\]

go to (f).

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

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

\[
c_T(n \wedge m) |\lambda^F_N| \leq p c_f r^{p-n \wedge m} \quad \text{and} \quad c_T |\lambda^F_N| r^{n \wedge m} - c_T r^p + c'_f \leq \lambda^F_N \cdot T(F) - f^{pc}_{N}(F)
\]

go to (h).

(g) Set \( r := 2r \), \( \tilde{\lambda}^F := \lambda^F_N \), and go to (b).
(h) If $j < J$ define
\[ \mathcal{N} = \{A' \in \mathcal{N}_{d/2,r}: \exists A \in \mathcal{N}, |A - A'| \leq d\}, \]
\[ \lambda_N^F \cdot T(A) - f(A) > \lambda_N^F \cdot T(F) - f_N^p(F) - Md \}, \]
set $\bar{\lambda}^F := \lambda_N^F$, $\varepsilon_{AS} := d$, $d := d/2$, $j := j + 1$, and go to (b).
(j) Stop.

Remarks 5.1. (i) We set $\varepsilon_{AS} = d$ in Step (h) since in the optimal case (if $\hat{f} \in C^{1,1}_{loc}(\mathbb{R}^r)$) Proposition 4.3 guarantees $|\lambda_{d,r}^F - \lambda_{d/2,r}^F| \leq O(d)$ so that the conditions of Lemma 5.1 are satisfied up to some constant of order $O(1)$.
(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 $\{A \in \mathcal{N}_{d,r}: |F - A| \leq d\}$ to $X$ in Step (b) guarantees feasibility of $f_{N,X}^p(F)$.

6. Numerical Experiments I

In this section we report on the practical performance of Algorithm $(A_{pc, \text{adapt}}^{r_0,d_0,J,F,M})$ 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 $f_{pc} = f_{rc} = f^{**}$ where $f^{**}$ is the convex envelope of $f$ and for $F \in \mathbb{R}^{2 \times 2}$ given by
\[ f^{**}(F) = \begin{cases} (|F|^2 - 1)^2 & \text{for } |F| \geq 1, \\ 0 & \text{for } |F| \leq 1. \end{cases} \]

Example 6.2 ([Ko, DW]). For $n = m = 2$,
\[ A_1 := \begin{pmatrix} 5/4 & 0 \\ 0 & 3/4 \end{pmatrix} \quad \text{and} \quad A_2 := \begin{pmatrix} 3\sqrt{2}/8 & 3/8 \\ -5/8 & 5\sqrt{3}/8 \end{pmatrix} \]
and $F \in \mathbb{R}^{2 \times 2}$ let
\[ f(F) := \frac{1}{2} \min\{|F - A_1|^2, |F - A_2|^2\}. \]
Then, (2.1) holds for $p = 2$, $c_f = 1/2^2$, and $c'_f = \max\{|A_1|^2, |A_2|^2\}/2 = 17/16$. Here, $f^{**} \neq f_{pc} = f_{rc}$ and $f_{pc}$ is for $F \in \mathbb{R}^{2 \times 2}$ given by
\[ f_{pc}(F) = \begin{cases} f_1(F) & \text{for } f_1(F) - f_2(F) \leq -\lambda/2, \\ f_2(F) - (f_2(F) - f_1(F) + \lambda/2) / (2\lambda) & \text{for } |f_1(F) - f_2(F)| \leq \lambda/2, \\ f_2(F) & \text{for } f_1(F) - f_2(F) \geq \lambda/2, \end{cases} \]
where $f_j(F) = |F - A_j|^2/2, j = 1, 2$, and $\lambda = |A_1 - A_2|$.

Example 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
\[ f(F) := \begin{cases} 1 + |F|^2 & \text{for } |F| \geq \sqrt{2} - 1, \\ 2\sqrt{2}|F| & \text{for } |F| \leq \sqrt{2} - 1. \end{cases} \]
Then, (2.1) holds for $p = 2$, $c_f = 1$, and $c'_f = 0$. Letting $g(F) := \sqrt{|F|^2 + 2|\det F|}$ for $F \in \mathbb{R}^{2 \times 2}$ there holds

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

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 $e = |f_{d,r}^{pc}(F) - f_{pc}(F)|$ and $e' = |\lambda_{d,r}^F \cdot DT(F) - Df_{pc}^*(F)|$, 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 $f_{d,r}^{pc}(F)$.

<table>
<thead>
<tr>
<th>d</th>
<th>$e$</th>
<th>$e'$</th>
<th>#X</th>
<th>#N</th>
<th>#N_{d,r}</th>
<th>CPU-time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.656637</td>
<td>2.000000</td>
<td>266</td>
<td>6,561</td>
<td>6,561</td>
<td>0.1 s</td>
</tr>
<tr>
<td>1/2</td>
<td>0.037543</td>
<td>0.325927</td>
<td>132</td>
<td>160</td>
<td>83,521</td>
<td>0.2 s</td>
</tr>
<tr>
<td>1/4</td>
<td>0.009873</td>
<td>0.206980</td>
<td>237</td>
<td>768</td>
<td>1,185,921</td>
<td>0.2 s</td>
</tr>
<tr>
<td>1/8</td>
<td>0.000177</td>
<td>0.001392</td>
<td>2,137</td>
<td>3,920</td>
<td>17,850,625</td>
<td>0.7 s</td>
</tr>
<tr>
<td>1/16</td>
<td>0.000010</td>
<td>0.000058</td>
<td>14,360</td>
<td>33,920</td>
<td>276,922,881</td>
<td>3.7 s</td>
</tr>
</tbody>
</table>

Table 1. Discretization parameter $d$, errors $e = |f_{d,r}^{pc}(F) - f_{pc}(F)|$ and $e' = |\lambda_{d,r}^F \cdot DT(F) - Df_{pc}^*(F)|$, number of active, possible, and theoretical nodes, and CPU-time needed to compute $f_{d,r}^{pc}(F)$ 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 $e \leq 10^{-3}$. 

18
To test \( \mathcal{A}^{\text{pc, adapt}}_{r_0,d_0,J,F,M} \) 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 \( \mathcal{N} \).

<table>
<thead>
<tr>
<th>( d )</th>
<th>( e )</th>
<th>( e' )</th>
<th>#X</th>
<th>#N</th>
<th>#( \mathcal{N}_{d,r} )</th>
<th>CPU-time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.165961</td>
<td>0.454711</td>
<td>6,561</td>
<td>6,561</td>
<td>6,561</td>
<td>0.9 s</td>
</tr>
<tr>
<td>1/2</td>
<td>0.072351</td>
<td>0.115344</td>
<td>794</td>
<td>43,568</td>
<td>83,521</td>
<td>1.1 s</td>
</tr>
<tr>
<td>1/4</td>
<td>0.014273</td>
<td>0.042726</td>
<td>1,311</td>
<td>167,136</td>
<td>1,185,921</td>
<td>1.4 s</td>
</tr>
<tr>
<td>1/8</td>
<td>0.004031</td>
<td>0.063871</td>
<td>1,766</td>
<td>715,504</td>
<td>17,850,625</td>
<td>2.4 s</td>
</tr>
<tr>
<td>1/16</td>
<td>0.001139</td>
<td>0.017684</td>
<td>4,889</td>
<td>3,082,192</td>
<td>276,922,881</td>
<td>6.5 s</td>
</tr>
<tr>
<td>1/32</td>
<td>0.000204</td>
<td>0.001032</td>
<td>12,140</td>
<td>13,694,256</td>
<td>4,362,470,401</td>
<td>21.8 s</td>
</tr>
</tbody>
</table>

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

We ran Algorithm \( \mathcal{A}^{\text{pc, adapt}}_{r_0,d_0,J,F,M} \) 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 \( d \leq 1/8 \).

<table>
<thead>
<tr>
<th>( d )</th>
<th>( e )</th>
<th>( e' )</th>
<th>#X</th>
<th>#N</th>
<th>#( \mathcal{N}_{d,r} )</th>
<th>CPU-time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.328922</td>
<td>0.162697</td>
<td>81</td>
<td>81</td>
<td>81</td>
<td>0.1 s</td>
</tr>
<tr>
<td>1/2</td>
<td>0.095387</td>
<td>0.458318</td>
<td>457</td>
<td>625</td>
<td>625</td>
<td>0.2 s</td>
</tr>
<tr>
<td>1/4</td>
<td>0.024647</td>
<td>0.160972</td>
<td>689</td>
<td>6,561</td>
<td>6,561</td>
<td>0.3 s</td>
</tr>
<tr>
<td>1/8</td>
<td>0.000806</td>
<td>0.167216</td>
<td>1,276</td>
<td>83,521</td>
<td>83,521</td>
<td>0.6 s</td>
</tr>
<tr>
<td>1/16</td>
<td>0.000389</td>
<td>0.079923</td>
<td>1,618</td>
<td>513,280</td>
<td>1,185,921</td>
<td>1.2 s</td>
</tr>
<tr>
<td>1/32</td>
<td>0.000180</td>
<td>0.032843</td>
<td>2,213</td>
<td>2,128,048</td>
<td>17,850,625</td>
<td>3.6 s</td>
</tr>
</tbody>
</table>

Table 3. Discretization parameter \( d \), errors \( e = |f_{d,r}^{\text{pc}}(F) - f^{\text{pc}}(F)| \) and \( e' = |\lambda_{d,r}^{F} \cdot DT(F) - Df^{\text{pc}}(F)| \), number of active, possible, and theoretical nodes, and CPU-time needed to compute \( f_{d,r}^{\text{pc}}(F) \) in Example 6.3.
7. Numerical Experiments II

In this section we outline how Algorithm \(A^{pc, adapt}_{d_0, d, J, F, M} \) 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.

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

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 \( \tilde{u} \in C(\overline{\Omega}; \mathbb{R}^m) \). We further suppose that \( \Omega \) is polyhedral and let \( T \) be a regular triangulation of \( \Omega \) such that \( \Gamma_D \) is matched exactly by edges (respectively faces) of elements in \( T \). We let \( S^1(T)^m \) denote the lowest order finite element space on \( T \) which consists of all globally continuous, \( T \)-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 \( T \). 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 \( A^{pc, adapt}_{d_0, d, J, F, M} \) provides an approximation of \( Df^{pc,F} \).

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

1. Set \( j := 0 \).
2. Let \( r_h \in S^1(T)^m \) satisfy \( r_h|_{\Gamma_D} = 0 \) and
   \[
   \int_{\Omega} \nabla r_h \cdot \nabla v_h \, dx = - \int_{\Omega} \left( \lambda \nabla u_h^{(j)} \cdot DT(\nabla u_h^{(j)}) \right) \cdot \nabla v_h \, dx
   \]
   for all \( v_h \in S^1(T)^m \) with \( v_h|_{\Gamma_D} = 0 \).
3. Compute \( t^* \in [0, 1] \) which is a local minimizer in \([0, 1]\) for
   \[
   t \mapsto \int_{\Omega} f^{pc}_{d,r}(\nabla(u_h^{(j)} + tr_h)) \, dx.
   \]
4. Stop and set \( u_h := u_h^{(j)} \) if \( t^* < \delta \).
5. Set \( u_h^{(j+1)} := u_h^{(j)} + t^*r_h \), \( j := j + 1 \), and go to (b).

**Output:** \( u_h \in S^1(T)^m \).

We specify \( f, \Omega, \tilde{u}_D, \) and \( T \) 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 \( T_k \) be the triangulation of \( \Omega \) 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 \((A_{\text{avvp}})\) 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 \( u_h^{(0)} = \bar{u}_{D,h} + \xi_h \) where \( \xi_h \in S^1(\mathcal{T})^2 \) 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 \( \delta \) was set to \( \delta = 0.03 \) and Algorithm \((A_{\text{avvp}})\) terminated after 175, 86, and 8 iterations for \( \ell = 1, 2, \) and 3, respectively. Table 4 displays the numerically relaxed energies for the output \( u_h \) in Example 7.1 and compares them to values that we obtained when \( f_{d,r}(F) \) and \( Df_{d,r} \) are replaced by \( f \) and \( Df \), \( f_{\text{pc}} \) and \( Df_{\text{pc}} \), and a numerical approximation \( f_{d,\text{rc}} \) of \( f_{\text{rc}} \) in steps (b) and (c) of Algorithm \((A_{\text{avvp}})\) (the numbers for \( f_{d,\text{rc}} \) are taken from [DW]). We observe that Algorithm \((A_{\text{avvp}})\) 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 \((A_{\text{avvp}})\) converge to the optimal value \( f_{\text{pc}}(B) = 0.199 \text{761} \) 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 \( f_{d,\text{rc}} \) were employed.

<table>
<thead>
<tr>
<th>( h )</th>
<th>( f )</th>
<th>( f_{\text{pc}} )</th>
<th>( f_{d,r}^{\text{pc}} )</th>
<th>( f_{d,\text{rc}}^{\text{pc}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/8</td>
<td>0.244128</td>
<td>0.199761</td>
<td>0.252177 (0.313327)</td>
<td>0.202507</td>
</tr>
<tr>
<td>1/16</td>
<td>0.234894</td>
<td>0.199761</td>
<td>0.216076 (0.274192)</td>
<td>0.208264</td>
</tr>
<tr>
<td>1/32</td>
<td>0.233769</td>
<td>0.199761</td>
<td>0.202994 (0.261244)</td>
<td>0.245633</td>
</tr>
</tbody>
</table>

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

Figure 2 displays the initial \( u_h^{(0)} \), the numerical solution obtained from direct numerical 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 \((A_{\text{avvp}})\) with stress field \( |\lambda_{d/4,r} \cdot DT(\nabla u_h)| \) 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 \((A_{\text{avvp}})\) 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 $\Gamma_D$.

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

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

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 $(A_{\text{nvvp}})$ (with the same parameters as in the previous experiment) with the approximated polyconvex envelope $f$ approach the value that we obtained with the exact polyconvex envelope $f_{\text{pc}}$. The minimal energies obtained with the nonrelaxed functional and with the discrete approximation of the rank-1 convex envelope of $f$ (numbers for $f_{d_{\text{rc}}}$ 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 $(A_{\text{nvvp}})$ and displayed in the right plot of Figure 3.

<table>
<thead>
<tr>
<th>$h$</th>
<th>$f$</th>
<th>$f_{\text{pc}}$</th>
<th>$f_{d_{\text{rc}}}$</th>
<th>$f_{d_{\text{rc}}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/8</td>
<td>0.186939 0.170023 0.231421 (0.244203) 0.175216</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/16</td>
<td>0.189875 0.168945 0.186094 (0.198793) 0.173716</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/32</td>
<td>0.188556 0.168583 0.173779 (0.186396) 0.183599</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**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 $(A_{\text{nvvp}})$ with stress field $|\Lambda_{d/\delta}\Sigma u_h \cdot D\nabla u_h|$ 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 $(A_{\text{nvvp}})$.

(ii) A post-processing procedure based on the algorithms in [Do, DW, Ba2] may be included.
in Algorithm \((A^{\text{rvvp}})\) which approximates the rank-1 convex envelope applied to \(\nabla u_h\) almost everywhere in \(\Omega\). Then, if (up to numerical tolerances) there holds \(f_{pc}(\nabla u_h) = f_{rc}(\nabla u_h)\) almost everywhere in \(\Omega\) 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 \(\Omega\). In this case, \(u_h\) serves as an approximation of a stationary point of the quasiconvex relaxation of \((M)\).

(iii) The computation of \(f_{d,\tau}^{pc}(\nabla u_h^{(j)})\) and \(\lambda_{d,\tau}^{u_h^{(j)}}\) in steps (b) and (c) of Algorithm \((A^{\text{rvvp}})\) has to be done on each element of the triangulation \(T\) (since \(\nabla u_h\) is \(T\)-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


Department of Mathematics, University of Maryland, College Park, MD 20742, USA

E-mail address: sba@math.umd.edu