ADAPTIVE APPROXIMATION OF YOUNG MEASURE SOLUTIONS IN SCALAR NON-CONVEX VARIATIONAL PROBLEMS

Sören Bartels

Abstract. This paper addresses the numerical approximation of Young measures appearing as generalized solutions to scalar non-convex variational problems. We prove a priori and a posteriori error estimates for a macroscopic quantity, the stress. For a scalar three well problem we show convergence of other quantities such as Young measure support and microstructure region. Numerical experiments indicate that the computational effort in the solution of the large optimization problem is significantly reduced by using an adaptive mesh refinement strategy based on a posteriori error estimates in combination with an active set strategy due to Carstensen and Roubíček (2000).

1. Introduction

A scalar model example in the context of phase transitions in crystalline solids reads:

\[(P) \left\{ \begin{array}{l} \text{Seek } u \in A := \{ v \in W^{1,2}(\Omega) : v|_{\Gamma_D} = u_D \} \\ \text{such that } I(u) = \inf_{v \in A} I(v). \end{array} \right. \]

Here, \( \Omega \subseteq \mathbb{R}^n \) is a bounded Lipschitz domain, \( \Gamma_D \subseteq \partial \Omega \) a closed subset of \( \partial \Omega \) with positive surface measure, and \( u_D \in W^{1/2,2}(\Gamma_D) \) is the trace of some function \( \tilde{u}_D \in W^{1,2}(\Omega) \). The energy functional \( I : A \rightarrow \mathbb{R} \) is for \( v \in A \) defined by

\[ I(v) := \int_{\Omega} W(\nabla v(x)) \, dx + \alpha \int_{\Omega} |u_0(x) - v(x)|^2 \, dx - \int_{\Omega} f(x)v(x) \, dx - \int_{\Gamma_N} g(x)v(x) \, ds_x, \]

where \( u_0, f \in L^2(\Omega), g \in L^2(\Gamma_N) \) for \( \Gamma_N := \partial \Omega \setminus \Gamma_D \), and \( \alpha \geq 0 \). An energy density \( W \) that can be derived from a three dimensional model with one-dimensional symmetry [BHJPS] is given by \( N + 1 \) wells \( s_0, \ldots, s_N \in \mathbb{R}^n \) and numbers \( s_0^0, \ldots, s_N^0 \in \mathbb{R} \) and reads

\[(1.1) \quad W(s) = \min_{j=0, \ldots, N} (|s - s_j|^2 + s_j^0) \quad \text{for all } s \in \mathbb{R}^n.\]

This function \( W \) serves as a model energy density but more generally we will consider mappings \( W : \mathbb{R}^n \rightarrow \mathbb{R} \) which are continuous and satisfy quadratic growth conditions.

The contributions in \( I \) which involve \( f \) and \( g \) represent outer body forces while the integral of \( W(\nabla v) \) measures the stored energy in \( \Omega \). A mechanical interpretation of the term \( \alpha \|u_0 - v\|_{L^2(\Omega)}^2 \) may be obtained from a model of a thin crystal plate glued to a rigid substrate [CL]. Similar scalar minimization problems arise in optimal control theory [R]. For ease of

Date: May 5, 2004.

1991 Mathematics Subject Classification. 65K10, 65N15, 49M20.

Key words and phrases. Non-convex variational problems, microstructure, finite elements, error estimation, adaptivity, multiple scales.
presentation, we restrict the analysis to quadratic growth conditions \( (p = 2) \) but stress that the estimates can easily be generalized to other growth conditions \( (2 \leq p < \infty) \).

It is well known that existence of solutions for \((P)\) depends on convexity properties of \(W\): If \(W\) is convex then there exists a solution which is unique provided \(W\) is strictly convex or \(\alpha > 0\). In case that \(W\) fails to be convex then \(I\) is not weakly lower semicontinuous and solutions may not exist. In the latter case, infimizing sequences are generically enforced to develop oscillations and do therefore not converge to a global minimizer of \(I\). To be able to deal with this phenomenon, we will consider appropriate (weak*) limits of infimizing sequences for \((P)\) which contain the most important information and which show where infimizing sequences develop oscillations. Those limits are measure valued functions called Young measures and arise as solutions for an extended problem \((EP)\). The numerical approximation of the extended problem has been proposed in [NW, R, Kr, CR, P1, RK] and in [KrP] for a non-convex variational problem in the theory of micromagnetics. It is our aim to establish error estimates for the numerical treatment of the extended problem. We note however that our analysis is restricted to scalar problems. The practically more relevant case of non-convex vectorial variational problems requires an efficient characterization of gradient Young measures and is excluded from our considerations.

The idea for the derivation of a priori and a posteriori error estimates is that the discretized extended problem may be regarded as a perturbation of a discretization of a relaxed (convexified) problem which has been analyzed in [CP]. This perturbation consists in the difference between the convex hull of the energy density itself and the convex hull of a discrete approximation of the energy density. Employing the concept of subdifferentials in the theory of non-smooth optimization we show that a dual variable, occurring in the discretized extended problem, converges to a macroscopic quantity of the relaxed problem and prove related error estimates. Moreover, we prove computable error estimates that allow for adaptive mesh refinement and which characterize a reliable relation between the two scales involved.

The “Active Set Strategy” of [CR] to solve a discretization of \((EP)\) efficiently for a fixed triangulation of \(\Omega\) is a multilevel scheme and depends on a good guess of a solution. Based on our error estimates we propose the embedding of that scheme into an adaptive mesh refining algorithm. We report the performance of the resulting algorithm for two examples. Our overall observation is that the algorithm performs very efficient but depends on a good solver for large optimization problems. For a two dimensional problem a numerical experiment indicates linear complexity of our solving strategy.

The outline of the rest of this paper is as follows. We state the extended problem in Section 2 and proceed in Section 3 with some notation, a construction of discrete Young measures, and the formulation of the discrete problem. Section 4 gives the announced error analysis as the main contribution of this work. Section 5 is devoted to the analysis of convergence of various quantities in a scalar three well problem. The “Active Set Strategy” of [CR] and its embedding into an adaptive mesh refinement algorithm are given in Section 6. Finally, in Section 7, we report on numerical results for two specifications of \((P)\) which illustrate the theoretical results of this article.
2. Young Measures and Extended Problem

In this section we recall the notion of Young measures which are mappings from \( \Omega \) into the space of probability measures on \( \mathbb{R}^n \) and allow for the computation of certain limits of weakly\(^*\) convergent sequences in Lebesgue spaces.

**Definition 2.1.** Let \( M(\mathbb{R}^n) \) be the set of all signed Radon measures on \( \mathbb{R}^n \) and let \( PM(\mathbb{R}^n) \) be the subset of probability measures on \( \mathbb{R}^n \), i.e., the set of all non-negative Radon measures \( \mu \in M(\mathbb{R}^n) \) satisfying \( \int_{\mathbb{R}^n} \mu(ds) = 1 \). The set of \( L^2 \)-Young measures \( \mathcal{Y}_2(\Omega; \mathbb{R}^n) \) is defined as

\[
\mathcal{Y}_2(\Omega; \mathbb{R}^n) := \{ \nu \in L^\infty(\Omega; M(\mathbb{R}^n)) : \nu_x \in PM(\mathbb{R}^n) \text{ for a.a. } x \in \Omega, \int_\Omega \int_{\mathbb{R}^n} |s|^2 \nu_x(ds) dx < \infty \}. 
\]

Here \( \nu_x := \nu(x) \) for \( x \in \Omega \) and \( L^\infty(\Omega; M(\mathbb{R}^n)) \) consists of those mappings \( \nu \in L^\infty(\Omega; M(\mathbb{R}^n)) \) for which the mapping \( x \mapsto \int_{\mathbb{R}^n} \nu(v) \nu_x(ds) \) is measurable whenever \( v \in C(\mathbb{R}^n) \) satisfies \( \lim_{|s| \to \infty} v(s) = 0 \).

Infimizing sequences for (\( P \)) generate Young measures in the sense of the following statement which is a consequence of the fundamental theorem on Young measures [Y, B, KiP, R]. Throughout this paper we assume that there exist constants \( c_1, c_2 > 0 \) such that

\[
(2.1) \quad c_1 |s|^2 - c_2 \leq W(s) \leq c_2 (1 + |s|^2) \quad \text{for all } s \in \mathbb{R}^n. 
\]

**Lemma 2.1** ([P2], Lemma 4.3). Let \( (u_j) \subseteq A \) be an infimizing sequence for (\( P \)), i.e. \( I(u_j) \to \inf_{v \in A} I(v) \). Then, there exist \( u \in A, \nu \in \mathcal{Y}_2(\Omega; \mathbb{R}^n) \), and a subsequence \((u_k)\) such that \( u_k \rightharpoonup u \) (weakly) in \( W^{1,2}(\Omega) \),

\[
\int_\Omega W(\nabla u_k(x)) dx \to \int_\Omega \int_{\mathbb{R}^n} W(s) \nu_x(ds) dx, 
\]

and, for almost all \( x \in \Omega \), there holds \( \nabla u(x) = \int_{\mathbb{R}^n} s \nu_x(ds) \).

The Young measure \( \nu \) generated by an infimizing sequence \( (u_j) \) for (\( P \)) describes oscillations in that sequence in a statistical way [B]. Together with the weak limit \( u \), we obtain the most relevant information about (\( P \)). If we express the limit of \( I(u_j) \) in terms of \( u \) and \( \nu \) we obtain the extended problem (\( EP \)).

\[
(EP) \quad \begin{cases} 
\text{Seek } (u, \nu) \in \mathcal{B} := \{ (v, \mu) \in W^{1,2}(\Omega) \times \mathcal{Y}_2(\Omega; \mathbb{R}^n) : \\
 v|_{\Gamma_D} = u_D, \nabla v(x) = \int_{\mathbb{R}^n} s \mu_x(ds) \text{ for a.a. } x \in \Omega \}, \\
\text{such that } \mathcal{T}(u, \nu) = \inf_{(v, \mu) \in \mathcal{B}} \mathcal{T}(v, \mu). 
\end{cases} 
\]

The extended energy functional \( \mathcal{T} \) is for \( (v, \mu) \in \mathcal{B} \) defined by

\[
\mathcal{T}(v, \mu) := \int_\Omega \int_{\mathbb{R}^n} W(s) \mu_x(ds) dx + \alpha \int_\Omega |u_0 - v|^2 dx - \int_\Omega f v dx - \int_{\Gamma_N} g v ds_x. 
\]

The following theorem shows that (\( EP \)) is a correct extension of (\( P \)). Limits in \( \mathcal{B} \) refer to the (weak, weak*)-topology in \( W^{1,2}(\Omega) \times \mathcal{Y}_2(\Omega; \mathbb{R}^n) \) (cf. [R] for details). Via the mapping \( \iota : A \to \mathcal{B}, u \mapsto (u, \delta_{\nabla u}) \), where for almost all \( x \in \Omega \) and all \( v \in C(\mathbb{R}^n) \) with \( \lim_{|s| \to \infty} v(s) = 0 \) the Dirac measure \( \delta_{\nabla u(x)} \in PM(\mathbb{R}^n) \) is defined by \( \int_{\mathbb{R}^n} v(s) \delta_{\nabla u(x)}(ds) = v(\nabla u(x)) \), \( A \) can be embedded continuously into \( \mathcal{B} \).
Theorem 2.1 ([R], Proposition 5.2.1). (i) \((EP)\) admits a solution.  
(ii) \(\inf_{w \in A} I(w) = \min_{(w, \mu) \in B} T(w, \mu)\).  
(iii) The embedding \(i : A \rightarrow B\) of each infimizing sequence for \((P)\) has a convergent subsequence whose limit is a solution to \((EP)\).  
(iv) Each solution to \((EP)\) is the limit of the embedding \(i : A \rightarrow B\) of an infimizing sequence for \((P)\). \(\square\)

Carathéodory’s Theorem implies that there exist solutions \((u, \nu) \in B\) to \((EP)\) such that for almost all \(x \in \Omega\) the probability measure \(\nu_x\) is a convex combination of at most \(n + 1\) Dirac measures (cf. [R], Corollary 5.3.3). This fact motivates the discretization of \((EP)\) introduced in Section 3 and the algorithm of [CR] to efficiently approximate \((EP)\).

3. Discretization of \((EP)\)

3.1. Finite Element Spaces and Notation. Let \(T\) be a regular triangulation of \(\Omega\) into triangles \((n = 2)\) or tetrahedra \((n = 3)\) in the sense of [Ci], i.e., there are no hanging nodes, the domain is matched exactly, i.e., \(\overline{\Omega} = \bigcup_{T \in T} T\), and \(T\) satisfies the maximum angle condition. Therefore, \(\partial \Omega\) is assumed to be polygonal. The extremal points of \(T \in T\) are called nodes and \(\mathcal{N}\) denotes the set of all such nodes. Let \(\mathcal{K} := \mathcal{N} \setminus \Gamma_D\) be the subset of free nodes. The set of edges (respectively faces if \(n = 3\)) \(\mathcal{E} = \text{conv}\{z_1, \ldots, z_n\} \subseteq \partial T\) for pairwise distinct \(z_1, \ldots, z_n \in \mathcal{N}\) and \(T \in T\) is denoted as \(\mathcal{E}\). A partition \(\mathcal{E} = \mathcal{E}_1 \cup \mathcal{E}_D \cup \mathcal{E}_N\) is given by \(\mathcal{E}_N := \{E \in \mathcal{E} : E \subseteq \overline{\Gamma}_N\}\), \(\mathcal{E}_D := \{E \in \mathcal{E} : \emptyset \subseteq \gamma_E \}\), and \(\mathcal{E}_\Omega := \mathcal{E} \setminus (\mathcal{E}_D \cup \mathcal{E}_N)\). The set

\[
\mathcal{L}^k(T) := \{v_h \in L^\infty(\Omega) : \forall T \in T, v_h|_T \in P_k(T)\}
\]

consists of all (possibly discontinuous) \(T\)-elementwise polynomials of degree at most \(k\). Define

\[
\mathcal{S}^1(T) := \mathcal{L}^1(T) \cap C(\overline{\Omega}) \quad \text{and} \quad \mathcal{S}^1_D(T) := \{u_h \in \mathcal{S}^1(T) : u_h|_{\Gamma_D} = 0\} \subseteq W^{1,2}_D(\Omega)
\]

where \(W^{1,2}_D(\Omega) := \{v \in W^{1,2}(\Omega) : v|_{\Gamma_D} = 0\}\). Let \((\varphi_z : z \in \mathcal{N})\) be the nodal basis of \(\mathcal{S}^1(T)\), i.e., \(\varphi_z \in \mathcal{S}^1(T)\) satisfies \(\varphi_z(x) = 0\) if \(x \in \mathcal{N} \setminus \{z\}\) and \(\varphi_z(z) = 1\). A function \(h_T \in \mathcal{L}^0(T)\) is defined by \(h_T|_T = h_T := \text{diam} (T)\) for all \(T \in T\). Moreover, let \(h_E \in L^\infty(\cup \mathcal{E})\) be defined by \(h_E|_E = h_E := \text{diam} (E)\) for all \(E \in \mathcal{E}\).

The nodal interpolation operator associated to a triangulation \(T\) is denoted by \(P_T\). If \(\tau\) is a triangulation of a convex domain \(\omega\) and \(v \in C(\overline{\omega})\) we extend \(P_T v\) to \(\mathbb{R}^n\) by setting \(P_T v(s) = P_T v(\mathcal{P}_\omega (s))\) where \(\mathcal{P}_\omega\) denotes the orthogonal projection onto \(\overline{\omega}\).

Suppose \(g \in L^2(\Gamma_N)\) is such that \(g|_E \in W^{1,2}(E)\) for all \(E \in \mathcal{E}_N\) and, for each node \(z \in \mathcal{N} \cap \overline{\Gamma}_N\) where the outer unit normal \(n_{\Gamma_N}\) on \(\Gamma_N\) is continuous, \(g\) is continuous. We set

\[
\mathcal{S}^1_N(T, g) := \{\tau_h \in \mathcal{S}^1(T)^n : \forall E \in \mathcal{E}_N \forall z \in E \cap \mathcal{N}, \tau_h(z) \cdot n_{\Gamma_N}|_E = g(z)\}
\]

and note that \(\mathcal{S}^1_N(T, g) \neq \emptyset\) if \(n = 2\). We will assume that \(\mathcal{S}^1_N(T, g) \neq \emptyset\) if \(n = 3\).

Throughout this article \(c, C > 0\) denote mesh-size independent, generic constants. For \(1 \leq p \leq \infty\) and an integer \(\ell > 0\), \(\|| \cdot ||_{L^p(\Omega)}\) stands for \(|| \cdot ||_{L^p(\Omega; \mathbb{R}^n)}\), and \(|| \cdot ||\) abbreviates \(|| \cdot ||_{L^2(\Omega)}\). The operator \(\partial_E \cdot /\partial s\) denotes the edgewise derivative along (subsets of) \(\partial \Omega\).
3.2. Discrete Young Measures. We define a convex, discrete (i.e., finite-dimensional) subset of the set of $L^2$-Young measures $\mathcal{Y}_d(\Omega; \mathbb{R}^n)$ following ideas of [R, CR, MRS].

**Definition 3.1** ([R], Example 3.5.4). Given a convex polygonal set $\omega \subseteq \mathbb{R}^n$ and regular triangulations $\tau$ of $\omega$ with nodes $\mathcal{N}_\tau$ and $\mathcal{T}$ of $\Omega$ we set

$$YM_{d,h}(\Omega; \mathbb{R}^n) := \{ \nu_{d,h} \in \mathcal{Y}_d(\Omega; \mathbb{R}^n) : \forall z \in \mathcal{N}_\tau \exists a_z \in \mathcal{L}^0(\mathcal{T}), a_z \geq 0 \text{ and } \sum_{z \in \mathcal{N}_\tau} a_z(x) = 1 \text{ a.e. in } \Omega, \nu_{d,h,x} = \sum_{z \in \mathcal{N}_\tau} a_z(x) \delta_z \text{ for a.e. } x \in \Omega \},$$

where $\delta_z$ denotes the Dirac measure supported in the atom $z \in \mathbb{R}^n \cap \mathcal{N}_\tau$. By $d$ and $h$ we denote the maximal mesh-size in $\tau$ and $\mathcal{T}$, respectively, and refer to $\tau$ and $\mathcal{T}$ through these quantities.

3.3. Discretized Extended Problem. For regular triangulations $\mathcal{T}$ of $\Omega$ and $\tau$ of a convex Lipschitz domain $\omega \subseteq \mathbb{R}^n$ and an approximation $u_{D,h} \in S^1(\mathcal{T})|_{\Gamma_D}$ of $u_D$ we consider the following discrete problem $(EP_{d,h})$.

$$(EP_{d,h}) \quad \left\{ \begin{array}{l}
\text{Seek } (u_{d,h}, \nu_{d,h}) \in \mathcal{B}_{d,h} := \{ (v_h, \mu_{d,h}) \in S^1(\mathcal{T}) \times YM_{d,h}(\Omega; \mathbb{R}^n) : v_h|_{\Gamma_D} = u_{D,h}, \\
\nabla v_h(x) = \int_{\mathbb{R}^n} s \mu_{d,h,x}(ds) \text{ for a.e. } x \in \Omega \},
\end{array} \right.$$

such that $\mathcal{T}(u_{d,h}, \nu_{d,h}) = \inf_{(v_h, \mu_{d,h}) \in \mathcal{B}_{d,h}} \mathcal{T}(v_h, \mu_{d,h})$.

An existence result for $(EP_{d,h})$ follows as for $(EP)$.  

**Proposition 3.1** ([R], Proposition 5.5.1). If $\mathcal{B}_{d,h} \neq \emptyset$ then $(EP_{d,h})$ admits a solution.  

**Remarks 3.1.** (i) There holds $\mathcal{B}_{d,h} \neq \emptyset$ if the diameter of $\omega$ is large enough.

(ii) For efficient approximations one has to assume a uniform bound on the gradient of a solution for $(EP)$. Based on optimality conditions stated below one may however enlarge $\omega$ successively to obtain a correct discrete solution. Therefore, no a priori bound on the gradient of an exact solution for $(EP)$ will be assumed.

(iii) For a triangulation $\mathcal{T}$ of $\Omega$ with $N^d$ free nodes and a triangulation $\tau$ of $\omega$ with $N^\tau$ atoms the number of degrees of freedom in $(EP_{d,h})$ is $N^{2n}$ if $h \approx d \approx 1/N$.

3.4. Optimality Conditions. The following lemma describes optimality conditions for $(EP_{d,h})$ which are key ingredients for the subsequent analysis.

**Lemma 3.1** ([CR], Proposition 4.3). Assume $\omega = \mathbb{R}^n$. The pair $(u_{d,h}, \nu_{d,h}) \in \mathcal{B}_{d,h}$ is a solution for $(EP_{d,h})$ if and only if there exists $\lambda_{d,h} \in \mathcal{L}^0(\mathcal{T})^n$ such that, for almost all $x \in \Omega$, we have

$$\max_{s \in \omega} \mathcal{H}_{\lambda_{d,h}}(x, s) = \int_{\mathbb{R}^n} \mathcal{H}_{\lambda_{d,h}}(x, s) \nu_{d,h,x}(ds),$$

where $\mathcal{H}_{\lambda_{d,h}}(x, s) := \lambda_{d,h}(x) \cdot s - P_\tau W(s)$, and, for all $v_h \in S^1_D(\mathcal{T})$, there holds

$$\int_{\Omega} \lambda_{d,h} \cdot \nabla v_h \, dx = 2\alpha \int_{\Omega} (u_0 - u_{d,h}) v_h \, dx + \int_{\Omega} f v_h \, dx + \int_{\Gamma_N} gv_h \, ds_x. \quad \square$$

**Remark 3.1.** The elementwise constant function $\lambda_{d,h}$ is the Lagrange multiplier for the constraint $\nabla u_{d,h}|_{\Gamma} = \int_{\mathbb{R}^n} s \nu_{d,h}|_{\Gamma}(ds)$, $T \in \mathcal{T}$, in $(EP_{d,h})$.  


For the practical implementation a bounded domain $\omega$ and a finite discretization of $\omega$ has to be chosen. We formulate appropriate computable conditions that imply Lemma 3.1.

**Lemma 3.2.** Assume that $\omega$ is bounded and $B_{r_0}(0) := \{s \in \mathbb{R}^n : |s| < r_0\} \subseteq \omega$ for some $r_0 > 0$. Let $(u_{d,h}, v_{d,h}) \in B_{d,h}$, $\lambda_{d,h} \in C^0(T)^n$, and assume

$$\int_\Omega \lambda_{d,h} \cdot \nabla v_h \, dx = 2\alpha \int_\Omega (u_{d,h} - u_{d,h}) v_h \, dx + \int_{\Gamma_N} f v_h \, dx + \int_{\Gamma_N} g v_h \, ds_x.$$  

If for almost all $x \in \Omega$ the mapping $s \mapsto \lambda_{d,h}(x) \cdot s - P_\tau W(s)$, $s \in \overline{\omega}$, attains its maximum in some $s_x^* \in \overline{B_{r_0}(0)}$, if $2r_0c_1 \geq \|\lambda\|_{L^\infty(\Omega)}$, and if for almost all $x \in \Omega$

$$r_0\|\lambda\|_{L^\infty(\Omega)} - c_1 r_0^2 + c_2 \leq \lambda_{d,h}(x) \cdot s_x^* = W(s_x^*),$$  

then the conditions of Lemma 3.1 are satisfied, i.e., $(u_{d,h}, v_{d,h})$ is a solution for $(EP_{d,h})$.

**Proof.** It suffices to show that for almost all $x \in \Omega$, any extension $\bar{\tau}$ of $\tau$ to $\mathbb{R}^n$, and all $s \in \mathcal{N}_\tau \setminus \overline{B_{r_0}(0)}$ there holds $\lambda_{d,h}(x) \cdot s - P_\tau W(s) \leq \mathcal{H}_{\lambda_{d,h}}(x, s_x^*)$, since then the optimality conditions of Lemma 3.1 are satisfied (with $\bar{\omega} = \mathbb{R}^n$ and $\bar{\tau}$). In view of (2.1) there holds

$$\lambda_{d,h}(x) \cdot s - P_\tau W(s) \leq \|\lambda_{d,h}\|_{L^\infty(\Omega)} |s| - c_1 |s|^2 + c_2.$$  

Since $2r_0c_1 \geq \|\lambda\|_{L^\infty(\Omega)}$ the mapping $t \mapsto \|\lambda_{d,h}\|_{L^\infty(\Omega)} t - c_1 t^2 + c_2$, $t \geq r_0$, is monotonically decreasing and since $\|\lambda_{d,h}\|_{L^\infty(\Omega)} r_0 - c_1 r_0^2 + c_2 \leq \mathcal{H}_{\lambda_{d,h}}(x, s_x^*)$ we have $\lambda_{d,h}(x) \cdot s - P_\tau W(s) \leq \mathcal{H}_{\lambda_{d,h}}(x, s_x^*)$ for all $s \in \mathcal{N}_\tau \setminus \overline{B_{r_0}(0)}$ which implies the same estimate for all $s \in \mathbb{R}^n \setminus \overline{B_{r_0}(0)}$. □

4. Error Estimates for $(EP_{d,h})$

We now turn to the formulation of error estimates for solutions for $(EP_{d,h})$. We prove that the Lagrange multiplier $\lambda_{d,h}$ converges to a macroscopic quantity, the stress, that appears naturally in $(P)$ and also in the convexified problem $(P^*)$. To estimate the distance between $\lambda_{d,h}$ and the exact stress we will regard $(EP_{d,h})$ as a perturbation of a discretization of $(P^*)$.

$$(P^*) \text{ Seek } u \in A \text{ such that } I^*(u) = \inf_{v \in A} I^*(v).$$

Here, the energy functional $I^*$ is defined for $v \in A$ and the convex envelope $W^*$ of $W$ by

$$I^*(v) := \int_\Omega W^*(\nabla v(x)) \, dx + \alpha \int_\Omega |u_0 - v|^2 \, dx - \int_\Omega f v \, dx - \int_{\Gamma_N} g v \, ds_x.$$  

**Definition 4.1.** For a solution $u \in A$ for $(P^*)$ we define the stress $\sigma := DW^*(\nabla u) \in L^2(\Omega)^n$.

**Theorem 4.1** ([CP], Theorem 2). $(P^*)$ admits a solution $u \in A$ such that

$$\int_\Omega DW^*(\nabla u) \cdot \nabla v \, dx - 2\alpha \int_\Omega (u_0 - u) v \, dx - \int_\Omega f v \, dx - \int_{\Gamma_N} g v \, ds_x = 0$$

for all $v \in W^{1,2}_D(\Omega)$. If $DW^*$ satisfies, for all $F, G \in \mathbb{R}^n$,

$$|DW^*(F) - DW^*(G)|^2 \leq C (DW^*(F) - DW^*(G)) \cdot (F - G)$$

then for two solutions $u, w \in A$ for $(P^*)$ there holds $DW^*(\nabla u) = DW^*(\nabla w)$, i.e., $\sigma$ is unique. If in addition to (4.2) $\alpha > 0$ or $W^*$ is strictly convex then $u = w$. □
Remarks 4.1. (i) If $W$ is as in (1.1) then $DW^{**}$ satisfies (4.2).
(ii) For a solution $(u, ν) \in B$ for $(EP)$ and a solution $w$ for $(P^{**})$ we have, provided $W, W^{**} \in C^1(\mathbb{R}^n)$, for almost all $x \in \Omega$, [F, KiP]
\[
\int_{\mathbb{R}^n} DW(s) ν_x(ds) = DW^{**}(∇w(x)).
\]
(iii) A result in [CaM] shows $σ \in W^{1,2}_{\text{loc}}(Ω)$.

In order to obtain a version of (4.1) in the discrete setting $(EP_{d,h})$ we need to differentiate the non-smooth convexification of $P_τ W$. To do this we apply the concept of subdifferentials.

**Definition 4.2.** For a convex function $V : \mathbb{R}^n \to \mathbb{R}$ and $ς \in \mathbb{R}^n$ the subdifferential of $V$ at $ς$ is defined by
\[
\partial V(ς) := \{ξ \in \mathbb{R}^n : V(ς + ξ) − V(ς) ≥ ξ \cdot ζ \text{ for all } ζ \in \mathbb{R}^n\}.
\]

**Remarks 4.2 ([C]).** (i) If $V$ is Gâteaux differentiable in $ς \in \mathbb{R}^n$ then $\partial V(ς) = \{∇V(ς)\}$. (ii) $V$ has a minimum in $ς \in \mathbb{R}^n$ if and only if $0 \in \partial V(ς)$.

The following lemma shows that the finite-dimensional minimization problem $(EP_{d,h})$ may be seen as a perturbation of a discretization of $(P^{**})$.

**Lemma 4.1.** Let $W^{cx}_d := ((P_τ W)|_{[\underline{ω}]}^{**})$ denote the convexification of the restriction of $P_τ W$ to $[\underline{ω}]$. Assume that $(u_{d,h}, ν_{d,h}) \in B_{d,h}$ and $λ_{d,h} \in L^0(T)^n$ satisfy the conditions of Lemma 3.2. Then $(u_{d,h}, ν_{d,h})$ minimizes the modified energy functional
\[
T(v_h, μ_{d,h}) := \int_Ω \int_{\mathbb{R}^n} W^{cx}_d(s) μ_{d,h,x}(ds) dx + α \int_Ω |u_0 − v_h|^2 dx − \int_Ω f v_h dx − \int_{Γ_0^N} gv_h ds_x,
\]
among all $(v_h, μ_{d,h}) \in B_{d,h}$. Moreover,
\[
λ_{d,h}(x) \in ∂W^{cx}_d(∇u_{d,h}(x)) \quad \text{for a.e. } x \in Ω.
\]

**Proof.** For $ς \in ω$ we have by Carathéodory’s Theorem [R],
\[
W^{cx}_d(ς) = ((P_τ W)|_{[ς]}^{**})(ς) = \inf_{s_1, ..., s_{n+1} \in ω} \sum_{i=1}^{n+1} \theta_i P_τ W(s_i).
\]

Since $P_τ W|_{[ς]}$ is $τ$-elementwise affine, it suffices to use the nodal values of $P_τ W$ in the calculation of $W^{cx}_d$, i.e.,
\[
W^{cx}_d(ς) = ((P_τ W)|_{[ς]}^{**})(ς) = \inf_{\sum_{z \in N_τ, ς = z} \sum_{\theta_1, z_i, i = 1, ..., n+1} \theta_1 P_τ W(z)} \sum_{\theta_1 \in [0,1], z_i \in N_τ, i = 1, ..., n+1} \sum_{\theta_1, z_i, i = 1, ..., n+1} \theta_1 P_τ W(z).
\]

Assume that there exists $ς \in \text{conv}\{z_1, ..., z_{n+1}\} = t ∈ τ, z_1, ..., z_{n+1} \in N_τ$ such that $ς = \sum_{i=1}^{n+1} \alpha_i z_i$ but $W^{cx}_d(ς) \neq \sum_{i=1}^{n+1} \alpha_i W^{cx}_d(z_i)$ with $α_i \in [0,1], \sum_{i=1}^{n+1} \alpha_i = 1$. If $W^{cx}_d(ς) > \sum_{i=1}^{n+1} \alpha_i W^{cx}_d(z_i)$ then $W^{cx}_d(ς)$ was not convex. If $W^{cx}_d(ς) < \sum_{i=1}^{n+1} \alpha_i W^{cx}_d(z_i)$ then $W^{cx}_d$ was not the largest convex function satisfying $W^{cx}_d ≤ P_τ W|_{[ς]}$. Therefore, $W^{cx}_d(ς) = \sum_{i=1}^{n+1} \alpha_i W^{cx}_d(z_i)$, so that $W^{cx}_d$ is $τ$-elementwise affine and $P_τ W^{cx}_d|_{[ς]} = W^{cx}_d$. To prove that $(u_{d,h}, ν_{d,h})$ minimizes
the functional $\mathcal{T}$ it suffices to verify the optimality conditions from Lemma 3.2 with $P_\tau W_d$ replaced by $P_\tau W^{c\!x}_d$. For this it is sufficient to show that, for almost all $x \in \Omega$, there holds

$$\max_{s \in \omega} (\lambda_{d,h}(x) \cdot s - P_\tau W(s)) = \max_{s \in \omega} (\lambda_{d,h}(x) \cdot s - W^{c\!x}_d(s))$$

and

$$\int_{\mathbb{R}^n} (\lambda_{d,h}(x) \cdot s - W^{c\!x}_d(s)) \nu_{d,h,x}(ds) = \int_{\mathbb{R}^n} (\lambda_{d,h}(x) \cdot s - P_\tau W(s)) \nu_{d,h,x}(ds).$$

Since $W^{c\!x}_d \leq P_\tau W(s)|_\omega$, we only have to show that

$$\max_{s \in \omega} (\lambda_{d,h}(x) \cdot s - P_\tau W(s)) \geq \max_{s \in \omega} (\lambda_{d,h}(x) \cdot s - W^{c\!x}_d(s))$$

and

$$\int_{\mathbb{R}^n} W^{c\!x}_d(s) \nu_{d,h,x}(ds) \geq \int_{\mathbb{R}^n} P_\tau(s) \nu_{d,h,x}(ds).$$

Let $\bar{s} \in \omega$ be maximizing in the right-hand side of (4.4), i.e.,

$$\lambda_{d,h}(x) \cdot \bar{s} - W^{c\!x}_d(\bar{s}) = \max_{s \in \omega} (\lambda_{d,h}(x) \cdot s - W^{c\!x}_d(s)).$$

By definition of $W^{c\!x}_d$ there exist $\theta_1, \ldots, \theta_{n+1} \in [0,1]$, $\sum_{i=1}^{n+1} \theta_i = 1$, and $z_1, \ldots, z_{n+1} \in N_\tau$ such that $\sum_{i=1}^{n+1} \theta_i z_i = \bar{s}$ and $W^{c\!x}_d(\bar{s}) = \sum_{i=1}^{n+1} \theta_i P_\tau W(z_i)$. By linearity of $s \mapsto \lambda_{d,h}(x) \cdot s$ we have

$$\lambda_{d,h}(x) \cdot \bar{s} - W^{c\!x}_d(\bar{s}) = \sum_{i=1}^{n+1} \theta_i (\lambda_{d,h}(x) \cdot z_i - P_\tau W(z_i))$$

$$\leq \sum_{i=1}^{n+1} \theta_i \max_{s \in \omega} (\lambda_{d,h}(x) \cdot s - P_\tau W(s)) = \max_{s \in \omega} (\lambda_{d,h}(x) \cdot s - P_\tau W(s)),$$

which proves (4.4). If

$$\int_{\mathbb{R}^n} W^{c\!x}_d(s) \nu_{d,h,x}(ds) < \int_{\mathbb{R}^n} P_\tau W(s) \nu_{d,h,x}(ds),$$

the explicit representation of $W^{c\!x}_d$ contradicts the fact that $(u_{d,h}, \nu_{d,h})$ is minimal for $\mathcal{T}$. We have thus shown (4.5) which yields the optimality conditions. The maximum principle of Lemma 3.1, the convexity of the mapping $s \mapsto W^{c\!x}_d(s) - \lambda_{d,h}(x) \cdot s$ together with Jensen’s inequality, and the identity $\nabla u_{d,h}(x) = \int_{\mathbb{R}^n} s d\nu_{d,h,x}(s)$ yield, for almost all $x \in \Omega$,

$$\max_{s \in \omega} (\lambda_{d,h}(x) \cdot s - W^{c\!x}_d(s)) = \int_{\mathbb{R}^n} (\lambda_{d,h}(x) \cdot s - W^{c\!x}_d(s)) \nu_{d,h,x}(ds)$$

$$\leq \lambda_{d,h}(x) \cdot \nabla u_{d,h}(x) - W^{c\!x}_d(\nabla u_{d,h}(x)).$$

Remark 4.2 shows, for almost all $x \in \Omega$, $0 \in -\lambda_{d,h}(x) + \partial W^{c\!x}_d(\nabla u_{d,h}(x))$. \hfill $\Box$

Another definition is needed for the a priori and a posteriori error estimates. It concerns the approximation of $DW^{\ast\ast}$ by the multi-valued mapping $\partial W^{c\!x}_d$.

**Definition 4.3.** For $A \subseteq \mathbb{R}^n$ and a multi-valued mapping $S : A \to 2^{\mathbb{R}^n}$, where $2^{\mathbb{R}^n}$ denotes the power set of $\mathbb{R}^n$, let

$$\|S\|_{L^\infty(A, 2^{\mathbb{R}^n})} := \sup_{t \in A} \sup_{s \in S(t)} |s|.$$
4.1. A Priori Error Estimates. The following theorem shows that the multiplier $\lambda_{d,h}$ for a solution $(u_{d,h}, \nu_{d,h}) \in \mathcal{B}_{d,h}$ for $(EP_{d,h})$ approximates the unique quantity $\sigma = DW^{**}(\nabla u)$ for a solution $u \in \mathcal{A}$ for $(P^{**})$.

**Theorem 4.2.** Assume that $DW^{**}$ satisfies (4.2) and $u \in \mathcal{A}$ solves $(P^{**})$. Assume that $(u_{d,h}, \nu_{d,h}) \in \mathcal{B}_{d,h}$ and $\lambda_{d,h} \in \mathcal{L}^0(\mathbb{T})^n$ satisfy the conditions of Lemma 3.1. There holds

$$\|\sigma - \lambda_{d,h}\| + \alpha\|u - u_{d,h}\| \leq C \inf_{(v_h, \nu_{d,h}) \in \mathcal{B}_{d,h}} \left( \|\nabla (u - v_h)\| + \alpha\|u - v_h\| \right) + C\|D W^{cx}_d - DW^{**}\|_{L^\infty(\omega; 2^n)} + \|\sqrt{\mathcal{C}}\|\|\partial W^{cx}_d - DW^{**}\|^{1/2}_{L^\infty(\omega; 2^n)}.$$

**Proof.** The triangle inequality, estimate (4.2), and Hölder’s inequality show

$$\frac{1}{2}\|\sigma - \lambda_{d,h}\|^2 \leq \|\sigma - DW^{**}(\nabla u_{d,h})\|^2 + \|DW^{**}(\nabla u_{d,h}) - \lambda_{d,h}\|^2$$

$$\leq C \int_\Omega (DW^{**}(\nabla u) - DW^{**}(\nabla u_{d,h})) \cdot (\nabla (u - u_{d,h}) + \alpha\|u - u_{d,h}\|) dx + \|\lambda_{d,h} - DW^{**}(\nabla u_{d,h})\|^2$$

$$= C \int_\Omega (DW^{**}(\nabla u) - \lambda_{d,h}) \cdot \nabla (u - u_{d,h}) dx + C \int_\Omega (\lambda_{d,h} - DW^{**}(\nabla u_{d,h})) \cdot \nabla (u - u_{d,h}) dx + \|\lambda_{d,h} - DW^{**}(\nabla u_{d,h})\|^2$$

$$\leq C \int_\Omega (DW^{**}(\nabla u) - \lambda_{d,h}) \cdot \nabla (u - u_{d,h}) dx + C\|\lambda_{d,h} - DW^{**}(\nabla u_{d,h})\|\|\nabla (u - u_{d,h})\|$$

$$+ \|\lambda_{d,h} - DW^{**}(\nabla u_{d,h})\|^2.$$

The Euler-Lagrange equations (4.1) for $u$ and Lemma 3.1 yield, for all $w_h \in S^1_D(\mathbb{T})$,

$$\int_\Omega (\sigma - \lambda_{d,h}) \cdot \nabla w_h dx + 2\alpha \int_\Omega (u - u_{d,h})w_h dx = 0.$$

We thus have

$$\int_\Omega (\sigma - \lambda_{d,h}) \cdot \nabla (u - u_{d,h}) dx + 2\alpha \int_\Omega (u - u_{d,h})^2 dx$$

$$= \int_\Omega (\sigma - \lambda_{d,h}) \cdot \nabla (u - u_{d,h} - w_h) dx + 2\alpha \int_\Omega (u - u_{d,h})(u - u_{d,h} - w_h) dx$$

$$\leq \|\sigma - \lambda_{d,h}\|\|\nabla (u - u_{d,h} - w_h)\| + 2\alpha\|u - u_{d,h}\|\|u - u_{d,h} - w_h\|.$$

The combination of the last two estimates shows after absorption of $\|\sigma - \lambda_{d,h}\|$ and $\|u - u_{d,h}\|$

$$\|\sigma - \lambda_{d,h}\|^2 + \alpha\|u - u_{d,h}\|^2 \leq C (\|\nabla (u - u_{d,h} - w_h)\|^2 + \alpha\|u - u_{d,h} - w_h\|^2$$

$$+ \|\lambda_{d,h} - DW^{**}(\nabla u_{d,h})\|\|\nabla (u - u_{d,h})\| + \|\lambda_{d,h} - DW^{**}(\nabla u_{d,h})\|^2).$$

Lemma 4.1 ensures $\lambda_{d,h}(x) \in \partial W^{cx}_d(\nabla u_{d,h}(x))$ and by construction of $\mathcal{B}_{d,h}$ we have $\nabla u_{d,h}(x) \in \omega$ for almost all $x \in \Omega$. This implies

$$\|\lambda_{d,h} - DW^{**}(\nabla u_{d,h})\|^2 \leq \int_{\Omega \in \partial W^{cx}_d(\nabla u_{d,h}(x)) - DW^{**}(\nabla u_{d,h})} |s|^2 dx$$

$$\leq |\Omega| \sup_{t \in \omega} \sup_{s \in \partial W^{cx}_d(t) - DW^{**}(t)} |s|^2 = |\Omega|\|\partial W^{cx}_d - DW^{**}\|_{L^\infty(\omega; 2^n)}^2.$$
Letting \( w_h = v_h - u_{d,h} \) for arbitrary \((v_h, \mu_{d,h}) \in B_{d,h}\) and estimating \( \|\nabla(u - u_{d,h})\| \leq C \) (which follows from the growth conditions (2.1)) we verify the assertion of the theorem. □

For a given energy density \( W \) and an appropriate triangulation \( \tau \) of \( \omega \) the term \( \| \partial W_d^{\text{cr}} - DW^{**} \|_{L^{\infty}(\omega;2^{\mathbb{R}_n})} \) can be estimated by the mesh-size of the discretization \( \tau \) of \( \omega \). We refer to Theorem 5.1 below for an estimate for a three well energy density.

**Remarks**

(i) Theorem 4.2, Theorem 5.1 below, and density of finite element spaces in \( A \) prove \( \lambda_{d,h} \rightarrow \sigma \) in \( L^2(\Omega) \) for \((d, h_T) \rightarrow 0\) and, if \( \alpha > 0 \), we also have \( u_{d,h} \rightarrow u \) in \( L^2(\Omega) \). If \( u \in C(\overline{\Omega}) \) we may choose \( v_h \) in Theorem 4.2 as the nodal interpolant of \( u \) and then we can estimate the error in powers of the mesh-size depending on smoothness properties of \( u \). Since in general \( u \) has no higher regularity properties, computable error bounds are needed.

(ii) Owing to non-uniqueness of \( u \) and degeneracy of (EP) we cannot expect strong convergence \( u_{d,h} \rightarrow u \) in \( W^{1,2} \).

### 4.2. A Posteriori Error Estimates

In this section two a posteriori error estimates, which are computable bounds for the error \( \| \sigma - \lambda_{d,h} \| \), are given. The first error estimate is similar to classical residual based a posteriori error estimates for elliptic partial differential equations [V] and employs jumps of normal components of \( \lambda_{d,h} \) across edges. Recall from the definition of (EP_{d,h}) that \( \omega \) is a fixed convex subset of \( \mathbb{R}^n \).

**Definition 4.4.** For \( E \in \mathcal{E}_\Omega \) and \( T_1, T_2 \in \mathcal{T} \) with \( E = T_1 \cap T_2 \) let \( n_E \) be the unit vector normal to \( E \), pointing from \( T_1 \) into \( T_2 \). For \( \lambda_{d,h} \in L^0(\mathcal{T})^n \) define

\[
[\lambda_{d,h} \cdot n_E] := \begin{cases} 
(\lambda_{d,h}|_{T_2} - \lambda_{d,h}|_{T_1}) \cdot n_E & \text{for } E \in \mathcal{E}_\Omega, T_1, T_2 \in \mathcal{T}, E = T_1 \cap T_2, \\
g - \lambda_{d,h}|_T \cdot n_{\Gamma}\|_E & \text{for } E \in \mathcal{E}_N, T \in \mathcal{T}, E \subseteq \partial T.
\end{cases}
\]

**Theorem 4.3.** Assume that \( DW^{**} \) satisfies (4.2) and \( u \in A \) solves \( (P^{**}) \). Let \((u_{d,h}, \nu_{d,h}) \in B_{d,h} \) and \( \lambda_{d,h} \in L^0(\mathcal{T})^n \) satisfy the conditions of Lemma 3.1. Then,

\[
\begin{align*}
\| \sigma - \lambda_{d,h} \|^2 + \alpha \| u - u_{d,h} \|^2 & \leq C \left\{ \left( \sum_{T \in \mathcal{T}} h_T^2 \| (f + \text{div} \lambda_{d,h} + 2\alpha (u_0 - u_{d,h})) \|_{L^2(T)} \right)^{1/2} \\
+ \left( \sum_{E \in \mathcal{E}_\Omega \cup \mathcal{E}_N} h_E \| [\lambda_{d,h} \cdot n_E]\|_{L^2(E)} \right)^{1/2} + \| \partial W_d^{\text{cr}} - DW^{**} \|_{L^{\infty}(\omega;2^{\mathbb{R}_n})} + \| h_\varepsilon^{3/2} \partial_s^2 u_D / \partial s^2 \|_{L^2(\Gamma_D)} \right\} \\
+ |\Omega| \| \partial W_d^{\text{cr}} - DW^{**} \|_{L^{\infty}(\omega;2^{\mathbb{R}_n})}.
\end{align*}
\]

**Proof.** Recall from the proof of Theorem 4.2 that, for \( w \in W^{1,2}(\Omega) \) satisfying \( w|_{\Gamma_D} = u_D - u_{D,h} \) and \( v_h \in S^1_D(\mathcal{T}) \), there holds

\[
C \| \sigma - \lambda_{d,h} \|^2 + 2\alpha \| u - u_{d,h} \|^2 \leq \int_{\Omega} (DW^{**}(\nabla u) - \lambda_{d,h}) \cdot \nabla(u - u_{d,h} - w - v_h) \, dx \\
+ 2\alpha \int_{\Omega} (u - u_{d,h})(u - u_{d,h} - w - v_h) \, dx + |\Omega| \| \partial W_d^{\text{cr}} - DW^{**} \|_{L^{\infty}(\omega;2^{\mathbb{R}_n})}^2 \\
+ C|\Omega| \| \partial W_d^{\text{cr}} - DW^{**} \|_{L^{\infty}(\omega;2^{\mathbb{R}_n})} + \| \sigma - \lambda_{d,h} \| \| \nabla w \| + 2\alpha \| u - u_{d,h} \| \| w \|.
\]

We employ the weak approximation operator \( \mathcal{J} : W_D^{1,2}(\Omega) \rightarrow S^1_D(\mathcal{T}) \) of [Ca, CB] and set \( v_h := \mathcal{J} v \). We then have (cf. [CB], Theorem 2.1)

\[
(4.6) \quad \| \nabla v_h \| + \| h_\varepsilon^{-1}(v - v_h) \| + \| h_\varepsilon^{-1/2}(v - v_h) \|_{L^2(\mathcal{E})} \leq C \| \nabla v \|
\]
The Euler-Lagrange equations (4.1) for \(u\), an elementwise integration by parts, and (4.6) show for \(v := u - u_{d,h} - w \in W_{D}^{1,2}(\Omega)\)

\[
\int_{\Omega} (DW^{**}(\nabla u) - \lambda_{d,h}) \cdot \nabla (v - J v) \, dx + 2\alpha \int_{\Omega} (u - u_{d,h}) (v - J v) \, dx
\]

\[
= \sum_{T \in T} \int_{T} (f + \text{div} \lambda_{d,h}) (v - J v) \, dx + 2\alpha \int_{\Omega} (u_{0} - u_{d,h}) (v - J v) \, dx
\]

\[
+ \sum_{E \in \varepsilon_{G} \cup \varepsilon_{N}} \int_{E} [\lambda_{d,h} \cdot n_{E}] (v - J v) \, ds_{x}
\]

\[
\leq C \left( \left( \sum_{T \in T} h_{T}^{2} \| (f + \text{div} \lambda_{d,h} + 2\alpha (u_{0} - u_{d,h})) \|^2_{L^{2}(T)} \right)^{1/2}
\]

\[
+ \left( \sum_{E \in \varepsilon_{G} \cup \varepsilon_{N}} h_{E} \| \lambda_{d,h} \cdot n_{E} \|^2_{L^{2}(E)} \right)^{1/2} \right) \| \nabla v \|.
\]

The combination of the last two estimates together with the a priori bound \(\| \nabla v \| \leq C\) and \(\min_{w|_{\Gamma_{D} = u_{D} - u_{d,h}}} \| w \|^2_{W^{1,2}(\Omega)} \leq C \| h_{T}^{2} \partial_{E} u_{D}/\partial s_{2} \|^2_{L^{2}(\Gamma_{D})}\) (cf. [BCD], Lemma 3.1) shows the assertion after absorption of \(\| \sigma - \lambda_{d,h} \|\) and \(\| u - u_{d,h} \|.\)

\( \square \)

**Remarks 4.4.**

(i) The term \(\| h_{T}^{3/2} \partial_{E} g_{D}/\partial s_{2} \|^2_{L^{2}(\Gamma_{D})}\) is of higher order.

(ii) The terms \(\| \partial W_{d}^{cx} - DW^{**} \|^2_{L^{\infty}(\omega;2^{n})}\) and \(\| \partial W_{d}^{cx} - DW^{**} \|^2_{L^{\infty}(\omega;2^{n})}\) are of higher order provided \(d \ll h_{T}\) (cf. Theorem 5.1). It will be shown later in Section 6 that the assumption \(d \ll h_{T}\) does not lead to inefficiency of our numerical schemes.

(iii) The a priori error estimate of Theorem 4.2 and the a posteriori error estimate of Theorem 4.3 yield a gap between reliability and efficiency of the error estimates with respect to the discretization parameter \(h\). While the a priori estimate gives optimal convergence results (for smooth solutions) we face a loss of a factor \(h_{T}^{1/2}\) in the a posteriori estimate due to degeneracy of the problem.

Our second error estimate is related to Zienkiewicz-Zhu (ZZ) error estimators (see, e.g., [CB]) for elliptic partial differential equations.

**Theorem 4.4.** Assume that \(DW^{**}\) satisfies (4.2) and \(u \in A\) solves \((P^{**})\). Let \((u_{d,h}, v_{d,h}) \in B_{d,h}\) and \(\lambda_{d,h} \in L^{0}(T)^{n}\) satisfy the conditions of Lemma 3.1. If \(\alpha = 0\) and \(f \in W^{1,2}(\Omega)\) then

\[
\| \sigma - \lambda_{d,h} \|^2 \leq C \left( \min_{\tau_{h} \in S_{d}^{1}(T,g)} \| \lambda_{d,h} - \tau_{h} \| + \| h_{T}^{3/2} \partial_{E} g_{D}/\partial s_{2} \|^2_{L^{2}(\Gamma_{D})}
\]

\[
+ \| h_{T}^{3/2} \partial_{E} g_{D}/\partial s_{2} \|^2_{L^{2}(\Gamma_{D})} + |\Omega| \| \partial W_{d}^{cx} - DW^{**} \|^2_{L^{\infty}(\omega;2^{n})} \right).
\]

**Proof.** As in the proof of Theorem 4.2 we have for \(w \in W^{1,2}(\Omega)\) with \(w|_{\Gamma_{D}} = u_{D} - u_{d,h}\) and \(v_{h} \in S_{d}^{1}(T)\)

\[
C \| \sigma - \lambda_{d,h} \|^2 \leq \int_{\Omega} (DW^{**}(\nabla u) - \lambda_{d,h}) \cdot \nabla (u - u_{d,h} - w - v_{h}) \, dx
\]

\[
+ |\Omega| \| \partial W_{d}^{cx} - DW^{**} \|^2_{L^{\infty}(\omega;2^{n})} + C |\Omega| \| \partial W_{d}^{cx} - DW^{**} \|^2_{L^{\infty}(\omega;2^{n})} + \| \sigma - \lambda_{d,h} \| \| \nabla w \|.
\]
Letting \( \tau_h \in S^1_N(\mathcal{T},g) \) and writing \( v := u - u_{d,h} - w \in W^1_0(\Omega) \) and \( v_h := Jv \in S^1_D(\mathcal{T}) \) we verify, using div \( \lambda_{d,h} \big|_T = 0 \), the Euler-Lagrange equation (4.1), an integration by parts, and Hölder’s inequality,

\[
\int_\Omega (DW^{**}(\nabla u) - \lambda_{d,h}) \cdot \nabla (v - Jv) \, dx \leq \int_\Omega f(v - Jv) \, dx + \sum_{T \in \mathcal{T}} \int_{\partial D} \text{div} (\tau_h - \lambda_{d,h})(v - Jv) \, ds_x + \| \tau_h - \lambda_{d,h} \| \| \nabla (v - Jv) \|.
\]

Choosing \( w \) as in [BCD] and employing elementary results about nodal interpolation on \( \Gamma_N \) we infer

\[
\int_\Omega (DW^{**}(\nabla u) - \lambda_{d,h}) \cdot \nabla (v - Jv) \, dx \leq C \left\{ \| h_T^2 \nabla f \| + \sum_{T \in \mathcal{T}} \| h_T^2 \| \text{div} (\tau_h - \lambda_{d,h}) \|_{L^2(T)} \right\} \| \nabla v \|.
\]

Choosing \( w \) as in [BCD] and employing elementary results about nodal interpolation on \( \Gamma_N \) we infer

\[
\| w \|_{W^{1,2}(\Omega)} + \| g - \tau_h \cdot n_{\Gamma_N} \|_{L^2(\Gamma_N)} \leq C \left( \| h_T^3/2 \partial^2_{s^2} u_{D}/\partial s^2 \|_{L^2(\Gamma_D)} + \| h_T^3/2 \partial \varepsilon g / \partial s \|_{L^2(\Gamma_N)} \right)
\]

and using an elementwise inverse estimate of the form

\[
h_T^\| \text{div} (\tau_h - \lambda_{d,h}) \| \|_{L^2(T)} \leq C \| \tau_h - \lambda_{d,h} \| \|_{L^2(T)} \quad \text{for all } T \in \mathcal{T}
\]

we verify the assertion as in the proof of the preceding theorem. \( \square \)

Remarks 4.5. (i) Terms including derivatives of \( u_D, g, \) or \( f \) are of higher order. Moreover, Remarks 4.4 (ii) and (iii) are valid here as well.

(ii) Theorem 4.4 shows, up to higher order terms, reliability of the error estimate \( \| \lambda_{d,h} - \lambda^*_h \| \) for any choice of a smooth approximation \( \lambda^*_h \in S^1_N(\mathcal{T},g) \) to \( \lambda_{d,h} \).

(iii) A triangle inequality proves an inverse, efficiency, estimate of Theorem 4.4 which holds up to higher order terms, provided \( \sigma \) is smooth but with different exponents,

\[
\min_{\tau_h \in S^1_N(\mathcal{T},g)} \| \lambda_{d,h} - \tau_h \| \leq \| \sigma - \lambda_{d,h} \| + \min_{\tau_h \in S^1_N(\mathcal{T},g)} \| \sigma - \tau_h \|.
\]

This efficiency estimate can be made rigorous but then without explicit constants.

5. CONVERGENCE OF OTHER QUANTITIES

In this section we present an estimate for \( DW^{**} - \partial W^{ce}_d \) and results concerning the convergence behavior of other quantities such as Young measure support and microstructure region in a three well problem. Ideas behind the proofs are adapted from [CP, F].

5.1. Approximation of \( DW^{**} \).

Theorem 5.1. For \( W : \mathbb{R}^2 \to \mathbb{R} \), \( s \mapsto \min_{j=0,1,2} |s - s_j|^2 \) with \( s_0 = (0,0), s_1 = (1,0), \) and \( s_2 = (0,1) \) and \( \omega = (-m,m)^2, m \geq 1, \) there exists a triangulation \( \tau \) of \( \omega \) with maximal mesh-size \( d = 1/k \), \( k \) a positive integer, of \( \omega \) such that

\[
\| \partial W^{ce}_d - DW^{**} \|_{L^\infty(\omega;2^{sk})} \leq C d \| D^2 W^{**} \|_{L^\infty(\omega)}.
\]
Moreover, the mapping \( DW^{**} \) satisfies (4.2).

Proof. A careful analysis shows that \( W^{**} \in C^1(\mathbb{R}^n) \) satisfies (4.2) and is for \( F = (f_1, f_2) \in \mathbb{R}^2 \) given by

\[
W^{**}(F) = \begin{cases} 
0, & F \in X_I, \\
W(F), & F \in X_{II} \cup X_{III} \cup X_{IV}, \\
f_2^2, & F \in X_V, \\
f_1^2, & F \in X_{VII}, \\
\frac{1}{2}(f_1 + f_2 - 1)^2, & F \in X_{VIII}.
\end{cases}
\]

Figure 1. \( W^{**} \) and triangulation of \( \omega \subseteq \mathbb{R}^2 \) to resolve the discontinuities of \( D^{2}W^{**} \).

For \( d = 1/k \), \( k \) a positive integer, choose \( \tau \) as in Figure 1. Since \( W_{d}^{cx} \) is affine on each \( t \in \tau \) we have \( \partial W_{d}^{cx}(s) = \operatorname{conv} \{ DW_{d}^{cx}|_t : t \in \tau, s \in t \} \). Since \( DW^{**} \) is continuous and \( \tau \)-elementwise differentiable it therefore suffices to show for each \( t \in \tau \)

\[
\| DW_{d}^{cx} - DW^{**} \|_{L^\infty(t)} \leq d \| D^{2}W^{**} \|_{L^\infty(t)}.
\]

Letting \( W_d^{**} = P_\tau W^{**} \) denote the nodal interpolant of \( W^{**} \) we have by standard interpolation results

\[
\| DW_{d}^{cx} - DW^{**} \|_{L^\infty(t)} \leq \| DW_{d}^{cx} - DW_{d}^{**} \|_{L^\infty(t)} + \| DW_{d}^{**} - DW^{**} \|_{L^\infty(t)}
\]

\[
\leq \| DW_{d}^{cx} - DW_{d}^{**} \|_{L^\infty(t)} + C d \| D^{2}W^{**} \|_{L^\infty(t)}.
\]

For each \( k \in \tau \) we define an affine function \( a_k : \mathbb{R}^2 \to \mathbb{R} \) such that, for all \( x \in \mathbb{R}^2 \), there holds

\[
W_{d}^{cx}(x) = \sup_{k \in \tau} a_k(x)
\]

and \( W_{d}^{cx}|_k = a_k \). If \( k \subseteq X_I \cup X_{II} \cup X_{III} \cup X_{IV} \cup X_V \cup X_{VII} \) we define \( a_k \) such that \( a_k(z) = W^{**}(z) \) for all \( z \in k \cap N_{\tau} \). If \( k \subseteq X_{VII} \) and there exists \( y = (y_1, y_2) \in k \) with \( y_1 + y_2 \in 1 + 2d[j, j + 1], j \geq 0 \) then we define

\[
a_k(x) = W(1 + j d, j d) + (x_1 - 1 - j d, x_2 - j d) \cdot (1, 1)
\]

\[
\times \frac{W(1 + (j + 1)d, (j + 1)d) - W(1 + jd, jd)}{2d}.
\]

Then, \( \sup_{k \in \tau} a_k \) is convex as it is the supreme of countably many affine functions. A proof for (5.2) then follows as above for the convexification of \( W \). Note that \( W_{d}^{cx} \) is mesh dependent. We now prove the remaining estimates. For \( k \subseteq X_I \cup X_{II} \cup ... \cup X_{VII} \) we have \( DW_{d}^{cx}|_k = DW_{d}^{**}|_k \) so that the asserted estimate follows from (5.1). For \( k \subseteq X_{VII} \) such that \( k \subseteq A_j = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 - x_2 \in [-1, 1], x_1 + x_2 \in 1 + 2d[j, j + 1] \}; j \geq 0 \) there holds \( W_{d}^{cx} = W^{**} \) on \( \partial A_j \) and \( W_{d}^{cx} \) is affine on \( A_j \). Therefore, \( W_{d}^{cx} \) interpolates \( W^{**} \) along each line segment in \( A_j \) parallel to \( (1, 1) \). The estimate

\[
\| DW_{d}^{cx} - DW_{d}^{**} \|_{L^\infty(t)} \leq \| DW_{d}^{cx} - DW_{d}^{**} \|_{L^\infty(t)} + \| DW^{**} - DW_{d}^{**} \|_{L^\infty(t)}
\]

\[
\leq C d \| D^{2}W^{**} \|_{L^\infty(t)}
\]
follows from the fact that the line segments have a length $d$.

5.2. Convergence of Young Measure Support.

Definition 5.1. For $A, B \subseteq \mathbb{R}^n$ let $\text{dist} (A, B) := \inf_{(a,b) \in A \times B} |a - b|$. We write $\text{Lim sup}_{\rho \to \rho_0} A_\rho \subseteq A$ (i.e., $A$ is the upper Kuratowski limit of $A_\rho$, cf. [Ku]) if

$$\forall \varepsilon > 0 \exists \delta > 0 \forall \rho, \ |\rho - \rho_0| \leq \delta \forall x \in A_\rho, \ \text{dist} (x, A) \leq \varepsilon.$$

Theorem 5.2. Let $W$ be as in Theorem 5.1, $u \in \mathcal{A}$ a solution for $(P^*)$, and $(u_j)_{j \geq 0}$ an infimizing sequence for $(P)$. Let $(u_{d,h}, \nu_{d,h}) \in \mathcal{B}_{d,h}$ and $\lambda_{d,h} \in \mathcal{L}^0(T)^n$ satisfy the conditions of Lemma 3.1. Assume that a subsequence of $(u_j)_{j > 0}$ converges weakly to $u$ and $\nu_{d,h} \in \mathcal{B}_{d,h}$ generates the Young measure $\nu$. Then, there exists a mapping $S : \mathbb{R}^2 \to 2^{\mathbb{R}^2}$ such that

$$\text{dist} (S(\lambda_{d,h}(x)), \text{supp } \nu_x) \to 0$$

if $x \in \Omega$ and $\lambda_{d,h}(x) \to \sigma(x)$. If for all $T \in \mathbb{R}^2$ there holds

$$\text{Lim sup}_{\delta \to 0} \{F \in \mathbb{R}^2 : \exists G, S \in \mathbb{R}^2, \{S, T\} \subseteq \partial W_d^{ex}(G), S \in \partial W_d^{ex}(F)\}$$

(5.3)

$$\subseteq \{F \in \mathbb{R}^2 : T = DW^{**}(F)\},$$

then we also have, if $x \in \Omega$ and $\lambda_{d,h}(x) \to \sigma(x)$,

$$\text{Lim sup}_{\lambda_{d,h}(x) \to \sigma(x)} \text{conv supp } \nu_{d,h,x} \subseteq \text{conv } S(\sigma(x)).$$

Remark 5.1. If $W_d^{ex}$ is continuously differentiable then

$$\{F \in \mathbb{R}^2 : \exists G, S \in \mathbb{R}^2, \{S, T\} \subseteq \partial W_d^{ex}(G), S \in \partial W_d^{ex}(F)\} = \{F \in \mathbb{R}^2 : T = DW_d^{ex}(F)\}.$$

Proof. Define $\mu : \mathbb{R}^2 \to PM(\mathbb{R}^2)$ by

$$F \mapsto \begin{cases} (1 - f_1 - f_2)\delta_{(0,0)} + f_1\delta_{(1,0)} + f_2\delta_{(0,1)} & \text{for } F \in X_I \\ \delta_F & \text{for } F \in X_{II} \cup X_{III} \cup X_{IV}, \\ (1 - f_1)\delta_{(0,f_2)} + f_1\delta_{(1,f_2)} & \text{for } F \in X_V, \\ (1 - f_2)\delta_{(f_1,0)} + f_2\delta_{(f_1,1)} & \text{for } F \in X_{VI}, \\ \frac{1}{2}(f_1 - f_2 + 1)\delta_{2(f_1 + f_2, f_1 + f_2 - 1)} + \frac{1}{2}(1 - f_1 + f_2)\delta_{2(f_1 + f_2 - 1, f_1 + f_2 + 1)} & \text{for } F \in X_{VII}, \end{cases}$$

Since $W^{**}$ is affine on $\text{conv supp } \nu_x$, $\int_{\mathbb{R}^2} s \, d\nu_x = \nabla u(x)$, and $\text{supp } \nu_x \subseteq \{E \in \mathbb{R}^2 : W(E) = W^{**}(E)\}$ for almost all $x \in \Omega$ [CP, F], one can show $\nu_x = \mu(\nabla u(x))$ for almost all $x \in \Omega$. For $S : \mathbb{R}^2 \to 2^{\mathbb{R}^2}$ defined by

$$(t_1, t_2) \mapsto \begin{cases} ((0,0), (1,0), (0,1)) & \text{for } (t_1, t_2) = (0,0), \\ (t_1 + 2, t_2)/2 & \text{for } t_1 > 0 \text{ and } t_2 < t_1, \\ (t_1, t_2)/2 & \text{for } t_1 < 0 \text{ and } t_2 < 0, \\ (t_1, t_2 + 2)/2 & \text{for } t_2 > 0 \text{ and } t_1 < t_2, \\ (0, t_2)/2, (2, t_2)/2 & \text{for } t_1 = 0, \\ (t_1, 0)/2, (t_1, 2)/2 & \text{for } t_2 = 0, \\ (t_1, t_2 + 2)/2, (t_1 + 2, t_2)/2 & \text{for } t_1 = t_2 \text{ and } t_1 > 0, \end{cases}$$

the explicit representation of $W^{**}$ shows

$$\text{supp } \mu(F) = S(DW^{**}(F)).$$
so that \( \text{supp } \nu_x = S(\sigma(x)) \) for a.e. \( x \in \Omega \). Hence

\[
(5.4) \quad \text{conv } S(T) = \{ E \in \mathbb{R}^2 : T = DW^{**}(E) \}.
\]

Moreover, for each \( \Sigma \in \mathbb{R}^2 \) the mapping \( \text{dist } (S(\cdot), \Sigma) : \mathbb{R}^2 \rightarrow \mathbb{R} \) is continuous and therefore

\[
\text{dist } (S(\lambda_{d,h}(x)), \text{supp } \nu_x) = \text{dist } (S(\lambda_{d,h}(x)), S(\sigma(x))) \rightarrow 0
\]

if \( x \in \Omega \) and \( \lambda_{d,h}(x) \rightarrow \sigma(x) \). Because of (5.3), (5.4) and since \( \text{Lim sup}_{t \rightarrow 0} B_\rho \subseteq A \) if \( \text{Lim sup}_{t \rightarrow 0} A_\rho \subseteq A \) and \( B_\rho \subseteq A_\rho \) for all \( \rho \), we only have to show that for \( T = \lambda_{d,h}(x) \)

\[
\text{conv sup}_d \nu_{d,h,x} \subseteq \{ F \in \mathbb{R}^2 : \exists G, \, S \in \mathbb{R}^2, \{ S, T \} \subseteq \partial W^{cx}_d(G), S \in \partial W^{cx}_d(F) \}
\]

in order to prove the second assertion. The set

\[
M_1 := \{ G \in \mathbb{R}^2 : \exists S \in \partial W^{cx}_d(\nabla u_{d,h}), S \in \partial W^{cx}_d(G) \}
\]

contains each subset \( A \subseteq \mathbb{R}^2 \) with

\[
W^{cx}_d \text{ affine on } A \text{ and } \nabla u_{d,h}(x) \in A.
\]

Since \( W^{cx}_d \) is affine \( \text{conv sup}_d \nu_{d,h,x} \) and \( \nabla u_{d,h}(x) \in \text{conv sup}_d \nu_{d,h,x} \) we have \( \text{conv sup}_d \nu_{d,h,x} \subseteq M_1 \). The inclusion \( \lambda_{d,h}(x) \in \partial W^{cx}_d(\nabla u_{d,h}(x)) \) and the choice \( G = \nabla u_{d,h}(x) \) yield

\[
M_1 \subseteq \{ F \in \mathbb{R}^2 : \exists G, \, S \in \mathbb{R}^2, \{ S, T \} \subseteq \partial W^{cx}_d(G), S \in \partial W^{cx}_d(F) \}
\]

which concludes the proof. \( \square \)

5.3. Convergence of the Microstructure Region.

**Definition 5.2.** Let \( \overline{M} \) denote the closure of \( M := \{ F \in \mathbb{R}^n : W(F) \neq W^{**}(F) \} \). For a solution \( u \in A \) for the convexified problem \( (P^{**}) \) and a solution \( (u_{d,h}, \nu_{d,h}) \in B_{d,h} \) for \( (EP_{d,h}) \) the microstructure region \( \Omega_{ms} \subseteq \Omega \) and the discrete microstructure region \( \Omega_{ms,h} \subseteq \Omega \) are defined by

\[
\Omega_{ms} := \{ x \in \Omega : \nabla u(x) \in \overline{M} \} \quad \text{and} \quad \Omega_{ms,h} := \{ x \in \Omega : \nabla u_{d,h}(x) \in \overline{M} \},
\]

respectively.

The following theorem shows that \( \Omega_{ms} \) is uniquely defined and that an appropriate approximation \( \tilde{\Omega}_{m,h} \) of \( \Omega_{ms,h} \) converges to \( \Omega_{ms} \).

**Theorem 5.3.** Let \( W \) be as in Theorem 5.1 and let \( u \) solve \( (P^{**}) \). There exists a Lipschitz-continuous mapping \( \xi : \mathbb{R}^2 \rightarrow \mathbb{R} \) such that, for almost all \( x \in \Omega \), we have

\[
x \in \Omega_{ms} \iff \xi(\sigma(x)) = 0.
\]

If \( v \in A \) is another solution for \( (P^{**}) \) then \( \xi(DW^{**}(\nabla u)) = \xi(DW^{**}(\nabla v)) \) a.e. in \( \Omega \). For a solution \( (u_{d,h}, \nu_{d,h}) \in B_{d,h} \) for \( (EP_{d,h}) \) with multiplier \( \lambda_{d,h} \in \mathcal{L}^0(T)^2 \) let

\[
\tilde{\Omega}_{m,h} := \{ x \in \Omega : \xi(\lambda_{d,h}(x)) = 0 \}.
\]

We then have

\[
(5.5) \quad \| \xi(\sigma) - \xi(\lambda_{d,h}) \| \leq C\| \sigma - \lambda_{d,h} \|
\]

and

\[
x \in \tilde{\Omega}_{m,h} \implies \text{dist } (\nabla u_{d,h}(x), \overline{M}) \leq C\| \partial W^{cx}_d - DW^{**} \|_{L^\infty(\omega;2\mathbb{R}^2)}.
\]
Conversely, there holds

\[ x \in \Omega_{ms,h} \quad \implies \quad |\xi(\lambda_{d,h}(x))| \leq \|\partial W_d^{cx} - DW^{**}\|_{L^\infty(\Omega;\mathbb{R}^2)}. \]

**Proof.** The explicit representation of $W^{**}$ in the proof of Theorem 5.1 shows, for almost all $x \in \Omega$, with $(s_1, s_2) = \sigma(x)$ and $F = \nabla u(x)$

\[ (s_1 = 0 \land s_2 \leq 0) \lor (s_2 = 0 \land s_1 \leq 0) \lor (s_1 = s_2 \land s_1 \geq 0). \]

The explicit representation of $W^{**}$ in the proof of Theorem 5.1 shows, for almost all $x \in \Omega$, with $(s_1, s_2) = \sigma(x)$ and $F = \nabla u(x)$

\[ x \in \Omega_{ms} \iff F \in \overline{X}_I \cup \overline{X}_V \cup \overline{X}_{VI} \cup \overline{X}_{VII} \]

\[ \iff (s_1 = 0 \land s_2 \leq 0) \lor (s_2 = 0 \land s_1 \leq 0) \lor (s_1 = s_2 \land s_1 \geq 0). \]

The mapping $\xi : \mathbb{R}^2 \to \mathbb{R}_{\geq 0}$ defined by

\[ (s_1, s_2) \mapsto \min\{|s_1| - \min\{-s_2, 0\}, |s_2| - \min\{-s_1, 0\}, |s_1 - s_2| - \min\{s_1, 0\}\} \]

is Lipschitz continuous with bounded Lipschitz-constant $C > 0$ and satisfies because of (5.6) the equivalence

\[ \xi(\sigma(x)) = 0 \iff x \in \Omega_{ms} \]

for almost all $x \in \Omega$. Since the quantity $\sigma := DW^{**}(\nabla u)$ is independent of the choice of a solution (cf. Remark 4.1) we have uniqueness of $\Omega_{ms}$. The Lipschitz continuity of $\xi$ implies the estimate (5.5). Let $x \in \Omega$ be such that $\xi(\lambda_{d,h}(x)) = 0$. The Lipschitz continuity of $\xi$ and the inclusion $\lambda_{d,h}(x) \in \partial W_d^{cx}(\nabla u_{d,h}(x))$ show

\[ \xi(DW^{**}(\nabla u_{d,h}(x))) = \xi(DW^{**}(\nabla u_{d,h}(x))) - \xi(\lambda_{d,h}(x)) \]

\[ \leq C|DW^{**}(\nabla u_{d,h}(x)) - \lambda_{d,h}(x)| \leq C\|DW^{**} - \partial W_d^{cx}\|_{L^\infty(\Omega;\mathbb{R}^2)}. \]

To prove the asserted estimate for $\text{dist}(\nabla u_{d,h}(x), \overline{M})$ it now suffices to prove

\[ \text{dist}(F, \overline{M}) \leq c\xi(DW^{**}(F)) \]

for a constant $c > 0$ and all $F \in \mathbb{R}^2$. The assertion is obvious if $F \in X_I \cup X_V \cup X_{VI} \cup X_{VII}$. We prove the case $F \in X_{II}$, the remaining cases $F \in X_{III}, X_{IV}$ follow analogously. Let $F = (f_1, f_2) \in X_{II}$. Then $f_1 - 1 \geq 0$ and $f_1 - 1 \geq f_2$. A short calculation shows

\[ \text{dist}(F, \overline{M}) = \min\{f_1 - 1, (f_1 - f_2 - 1)/\sqrt{2}\}. \]

Since $DW^{**}(F) = 2(f_1 - 1, f_2)$ we have

\[ \xi(DW^{**}(F)) = 2\min\{f_1 - 1 - \min\{-f_2, 0\}, f_2 + f_1 - 1, f_1 - f_2 - 1\}. \]

If $f_2 \leq 0$ then this term can be simplified to $\xi(DW^{**}(F)) = \min\{f_1 - 1, f_1 - f_2 - 1\}$ and the assertion follows. If $f_2 \geq 0$ we have

\[ \xi(DW^{**}(F)) = \min\{f_1 - 1 + f_2, f_1 - f_2 - 1\} = f_1 - f_2 - 1 \geq (f_1 - f_2 - 1)/\sqrt{2} \]

\[ = \min\{f_1 - 1, (f_1 - f_2 - 1)/\sqrt{2}\} = \text{dist}(F, \overline{M}). \]

To prove the inverse implication let $x \in \Omega_{ms,h}$, i.e., $\nabla u_{d,h}(x) \in X_I \cup X_V \cup X_{VI} \cup X_{VII}$. Since $\lambda_{d,h}(x) \in \partial W_d^{cx}(\nabla u_{d,h}(x))$ and since $\partial W_d^{cx}(\nabla u_{d,h}(x)) = \text{conv}\{DW_d^{cx}|_t : t \in \tau, \nabla u_{d,h}(x) \in t\}$, there exist $t_1, \ldots, t_{n+1} \in \tau$ and $\varrho_i \in [0, 1], n+1 \sum_{i=1}^{n+1} \varrho_i = 1$ such that $\lambda_{d,h}(x) = n+1 \sum_{i=1}^{n+1} \varrho_i DW_d^{cx}|_{t_i}$. The identities

\[ \lambda_{d,h}(x) = \sum_{i=1}^{n+1} \varrho_i DW_d^{cx}|_{t_i} = \sum_{i=1}^{n+1} \varrho_i (DW_d^{cx}|_{t_i} - DW^{**}(\nabla u_{d,h})) + DW^{**}(\nabla u_{d,h}) \]
and \( \xi(DW^*(\nabla u_{d,h})) = 0 \) combined with the Lipschitz continuity of \( \xi \) show
\[
|\xi(\lambda_{d,h}(x))| = |\xi\left(\sum_{i=1}^{n+1} \varrho_i(D_W^{cz}|_{t_i} - DW^*(\nabla u_{d,h}) + DW^*(\nabla u_{d,h})) - (DW^*(\nabla u_{d,h}))\right)|
\leq C \sum_{i=1}^{n+1} \varrho_i(D_W^{cz}|_{t_i} - DW^*(\nabla u_{d,h})) \leq C\|W^* - \partial W_{d,h}^{cz}\|_{L^\infty(\omega;2^{\mathbb{R}^2})}. \quad \square
\]

6. Combination of A Multilevel Scheme and Adaptive Mesh Refinement

6.1. Active Set Strategy Due to Carstensen & Roubíček. The identity
\[
\max_{s \in \omega} \mathcal{H}_{\lambda_{d,h}}(x, s) = \int_{\mathbb{R}^n} \mathcal{H}_{\lambda_{d,h}}(x, s) v_{d,h,x}(ds)
\]
in Lemma 3.1 for a solution \((u_{d,h}, v_{d,h}) \in \mathcal{B}_{d,h}\) for \((EP_{d,h})\) with multiplier \(\lambda_{d,h} \in \mathcal{L}^0(\mathcal{T})\) states that for almost each \(x \in \Omega\) the probability measure \(v_{d,h,x}\) is supported in those atoms \(z \in \mathcal{N}_x\) for which \(\mathcal{H}_{\lambda_{d,h}}(x, \cdot)\) attains its maximum. Typically, these are only a few atoms.

If the support of the Young measure \(v_{d,h}\),
\[
\text{Supp}(v_{d,h}) := \{(x, z) \in \Omega \times \mathcal{N}_x : z \in \text{supp}(v_{d,h,x})\},
\]
where \(\text{supp}(v_{d,h,x}) \subseteq \mathbb{R}^n\) is the support of the Radon measure \(v_{d,h,x}\), was known a priori, we could set \(A := \text{Supp}(v_{d,h})\) and seek \((u_{d,h}, v_{d,h})\) as a solution of the following lower-dimensional problem \((EP_{d,h,A})\).

\[(EP_{d,h,A}) \quad \left\{ \begin{array}{l}
\text{Seek } (u_{d,h}, v_{d,h}) \in \mathcal{B}_{d,h} \text{ such that } \text{Supp}(v_{d,h}) \subseteq A \\
\text{and } \overline{T}(v_{d,h}) = \inf_{(v_{h}, \mu_{d,h}) \in \mathcal{B}_{d,h}} \overline{T}(v_{h}, \mu_{d,h}).
\end{array} \right.\]

Proposition 5.4 in [CR] gives a necessary condition on \(A\) which ensures that \((EP_{d,h,A})\) is a correct reduction of \((EP_{d,h})\). Conversely, Lemma 3.2 states a sufficient criterion for a solution of \((EP_{d,h,A})\) to solve \((EP_{d,h})\).

Given an approximation \(\tilde{h}\) of \(\mathcal{H}_{\lambda_{d,h}}\) we define a set of active atoms, called the active set, by
\[
(6.1) \quad A = \{(x, z) \in \Omega \times \mathcal{N}_x : \tilde{h}(x, z) \geq \max_{s \in \omega} \tilde{h}(x, s) - \varepsilon(x)\},
\]
where \(\varepsilon \in \mathcal{L}^0(\mathcal{T})\), \(\varepsilon > 0\) almost everywhere in \(\Omega\), is a given tolerance. If \(\varepsilon\) is large enough then any solution for \((EP_{d,h,A})\) with \(A\) as in (6.1) is a solution for \((EP_{d,h})\).

**Lemma 6.1.** Let \((u_{d,h}, v_{d,h})\) be a solution for \((EP_{d,h})\) with corresponding multiplier \(\lambda_{d,h}\) and \(\mathcal{H}_{\lambda_{d,h}}(x, s) = \lambda_{d,h}(x) \cdot s - P_sW(s)\). Moreover, let \(\tilde{h} : \Omega \times \mathbb{R}^n \to \mathbb{R}\) and \(\varepsilon \in \mathcal{L}^0(\mathcal{T})\), \(\varepsilon > 0\) almost everywhere in \(\Omega\) be such that, for each \(T \in \mathcal{T}\),
\[
\|\mathcal{H}_{\lambda_{d,h}} - \tilde{h}\|_{L^\infty(\omega;2^{\mathbb{R}^2})} \leq \varepsilon|T|,
\]
with \(S_T \subseteq \mathbb{R}^n\) such that, for almost all \(x \in T\), we have
\[
\{s \in \omega : \mathcal{H}_{\lambda_{d,h}}(x, s) = \max_{\tilde{s} \in \omega} \mathcal{H}_{\lambda_{d,h}}(x, \tilde{s})\} \cup \{s \in \omega : \tilde{h}(x, s) = \max_{\tilde{s} \in \omega} \tilde{h}(x, \tilde{s})\} \subseteq S_T.
\]
If \(A\) is defined by (6.1) then any solution for \((EP_{d,h,A})\) is a solution for \((EP_{d,h})\). \quad \square

**Proof.** The proof follows the arguments of [CR].
The idea to guess the support of a Young measure solution in a multilevel scheme together with Lemma 6.1 motivates the following algorithm in which a sequence of refining triangulations, elementwise constant tolerances, and an initial guess \( \tilde{h}_0 \) for \( \mathcal{H}_{\lambda_{d,h}} \), e.g., \( \tilde{h}_0 = 0 \), are given. Figure 2 includes a schematic flow chart of the algorithm.

Algorithm \((A^{\text{active set}})\). Let \( \tau_1, \tau_2, \ldots, \tau_J \) be triangulations of \( \omega, \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_J > 0 \) be elementwise constant, and \( \tilde{h}_0 \in L^1(\Omega; C(\mathbb{R}^n)) \).

1. Set \( \varepsilon := \varepsilon_1, \tilde{h} := \tilde{h}_0, \tau := \tau_1, \) and \( j := 1 \).
2. Compute \( A \) from (6.1).
3. Compute a solution \( (u_{d,h}, \nu_{d,h}) \in \mathcal{B}_{d,h,A} \) for \((EP_{d,h,A})\) and the multiplier \( \lambda_{d,h} \in \mathcal{L}^0(T)^n \).
4. If the conditions of Lemma 3.2 are satisfied then go to (6) otherwise proceed with (5).
5. Increase \( m \) if necessary. Enlarge \( \varepsilon \) by \( \varepsilon|_T := 2\varepsilon|_T \) if for some \( x_T \in \tau \) max \( z \in \mathbb{N} \) \( \mathcal{H}_{\lambda_{d,h}}(x_T, z) > \int_{\mathbb{R}^n} \mathcal{H}_{\lambda_{d,h}}(x_T, s) \nu_{h,x_T}(ds) \), and set \( \varepsilon|_T := \varepsilon|_T \) otherwise. Go to (2).
6. If \( j < J \) proceed with (7) otherwise terminate.
7. Set \( j := j + 1, \tilde{h}(x, s) := \lambda_{d,h}(x) \cdot s - P_T W(s), \varepsilon := \varepsilon_j \) and go to (2).

Remarks 6.1. (i) The approximation \( \tilde{h}_0 \) may initially be chosen as \( \tilde{h}_0 = 0 \) and then all atoms are activated in (6.1) or \( \tilde{h}_0 \) is defined through the solution on a coarser triangulation \( T' \).

(ii) Since the tolerance \( \varepsilon \) is increased successively the optimality conditions of Lemma 6.1 are satisfied after a finite number of iterations.

6.2. Adaptive Mesh Refinement. Theorems 4.3 and 4.4 allow the introduction of local refinement indicators which may be used for automatic mesh refinement. Let \( (u_{d,h}, \nu_{d,h}) \) be a solution for \((EP_{d,h})\) with corresponding multiplier \( \lambda_{d,h} \).

Theorem 4.3 motivates the elementwise contributions, for \( T \in \mathcal{T}, \)

\[
\eta_R(T)^2 := h_T^2 f + \text{div} \lambda_{d,h} + 2\alpha(u_0 - u_{d,h})^2_{L^2(T)} + \sum_{E \in \mathcal{E}_0 \cup \mathcal{E}_N} \sum_{E \subseteq \partial T} h_E \|[\lambda_{d,h} \cdot n_E]\|_{L^2(E)}^2.
\]

In regard to Theorem 4.4 we employ the operator \( \tilde{A} : L^2(\Omega)^n \rightarrow S_{N}(T; g) \) of [CB], which is for \( \Gamma_N = \emptyset \) and \( p \in L^2(\Omega)^n \) given by

\[
\tilde{A} p = \sum_{z \in \mathcal{N}} p_z \varphi_z, \text{ for } p_z := \int_{\varphi_z > 0} p dx / \int_{\varphi_z > 0} 1 dx,
\]

to define, for \( T \in \mathcal{T}, \)

\[
\eta_z(T) := \|\lambda_{d,h} - \tilde{A} \lambda_{d,h}\|_{L^2(T)}.
\]

With these definitions we have

\[
\|\sigma - \lambda_{d,h}\|^2 \leq C \left( \sum_{T \in \mathcal{T}} \eta(T)^2 \right)^{1/2} + \text{h.o.t.}
\]
where $\eta(T) = \eta_Z(T)$ or $\eta(T) = \eta_R(T)$ and the terms h.o.t. depend on the mesh-size of the triangulation $\tau$ which are of higher order provided $d \ll h_T$ and on smoothness of given right-hand sides. We set

$$\eta_R := \left( \sum_{T \in \mathcal{T}} \eta_R(T)^2 \right)^{1/4} \quad \text{and} \quad \eta_{Z,R} := \left( \sum_{T \in \mathcal{T}} \eta_Z(T)^2 \right)^{1/4}.$$

Remark 4.5 (iii) states

$$\eta_{Z,E} := \left( \sum_{T \in \mathcal{T}} \eta_Z(T)^2 \right)^{1/2} \leq \|\sigma - \lambda_{d,h}\| + \text{h.o.t.}$$

**Figure 2.** Schematic flow chart for the combination of the Active Set Strategy (as in [CR], inside the dashed box) with adaptive mesh refinement.

The following algorithm generates the triangulations in the numerical examples of the subsequent section. The parameter $\Theta$ allows to use the algorithm for uniform mesh refinement which corresponds to $\Theta = 0$ and adaptive mesh refinement where $\Theta = 1/2$. For details on adaptive mesh refinement we refer to [V]. A schematical flow chart for the combination of the Active Set Strategy with the Adaptive Mesh Refinement Algorithm is shown in Figure 2.

**Algorithm ($A^\text{adaptive}_\Theta$).** (1) Start with a coarse triangulation $\mathcal{T}_1$ of $\Omega$ and set $\omega := (-m, m)^n$, $\ell = 1$, and $\tilde{\lambda}_\ell = 0$. 

(2) Compute a discrete solution \((u_\ell, v_\ell, \lambda_\ell)\) with Algorithm \((A^{\text{active set}})\) and starting values \(\hat{h}_0(x, s) := \hat{\lambda}(x) \cdot s - P_x W(s)\). \(J = 2\), \(d_j = 2^{j-1}/k\), \(k = \lfloor 4m \cdot 2^{-j} \cdot \text{card} (\mathcal{N}_T)^{3/2n} \rfloor\) ([\(s\] is the largest integer \(\leq s\), \(\varepsilon_j := 2^{-\ell-j} \cdot 10^{-4}\) for \(j = 1, \ldots, J\) \((\varepsilon_1 := \infty\) if \(\ell = 1\) to activate all atoms), and a triangulation \(T_\ell\) of \(\omega\) with maximal mesh-size \(d_j\).

(3) For each \(T \in T_\ell\) compute refinement indicators \(\eta_Z(T)\) and \(\eta_R(T)\).

(4) Mark the element \(T\) for red-refinement if

\[
\eta_R(T) \geq \Theta \max_{T' \in T_\ell} \eta_R(T').
\]

(5) Mark further elements \((\text{red-blue-green-refinement})\) to avoid hanging nodes. Terminate if the stopping criterion is satisfied, generate a new triangulation \(T_{\ell+1}\), define \(\hat{\lambda}_{\ell+1} := \lambda_\ell\), increment \(\ell\), and go to (b) otherwise.

Remarks 6.2. (i) We chose \(k\) such that \(d \propto h^{3/2}\) so that \(\|DW^{**} - \partial W^c_d\|_{L^\infty(\Omega;\mathbb{R}^n)}\) is of the same order as the presumed higher order terms involving \(g\) and \(u_D\) in Theorems 4.3 and 4.4.

(ii) Since \(\lambda_\ell \rightarrow DW^{**}(\nabla u)\) in \(L^2(\Omega)\) for a solution \(u \in A\) for \((P^{**})\), \(\lambda_\ell\) is a Cauchy sequence, and therefore \(\lambda_\ell\) is a good approximation for \(\lambda_{\ell+1}\) if \(\ell\) is large enough.

7. Numerical Experiments

In this section we present numerical results for two specifications of \((P)\). The first example has been investigated in [CR] and is modified here to obtain quadratic growth conditions. The second example is a two-dimensional problem that reveals limitations of our approach to solve \((P)\) but thereby underlines the necessity of the design of efficient algorithms for the solution for \((EP_{d,b})\).

The implementation of the algorithms was performed in Matlab as described in [CR] for the part concerning the Active Set Strategy. We solved the linear optimization problems with the interior point linear program solver HOPDM [G].

Example 7.1 (One-dimensional two-well problem.). Let \(n = 1\), \(\Omega = (0,1)\), \(\Gamma_D = \{0,1\}\), \(\alpha = 0\), \(\Gamma_N = \emptyset\), and \(W(s) = \min\{(s-1)^2, (s+1)^2\}\). The right-hand sides are defined by

\[
f(x) = \begin{cases} 
0 & \text{for } x \leq x_b, \\
\gamma (x-x_b)/2 & \text{for } x \geq x_b,
\end{cases}
\]

and

\[
u_D(0) = 3x_b^5/128 + x_b^3/3 \quad \text{and} \quad u_D(1) = \gamma(1-x_b)^3/24 + 1 - x_b,
\]

where \(\gamma = 100\) and \(x_b = \pi/6\). A solution for \((P^{**})\) is then given by

\[
u(x) = \begin{cases} 
-3(x-x_b)^5/128 - (x-x_b)^3/3 & \text{for } x \leq x_b, \\
\gamma (x-x_b)^3/24 + x - x_b & \text{for } x \geq x_b
\end{cases}
\]

and allows to compute the unique quantity \(\sigma := DW^{**}(u')\). The microstructure region is \((0,x_b)\) in which \(\sigma = 0\) and \(u'\) lies between the wells \(-1\) and \(1\), i.e., \(u'(x) \in (-1,1)\) for \(x \in (0,x_b)\). A Young measure corresponding to \(u\) is given by

\[
\nu_x = \begin{cases} 
\frac{1-u(x)}{2} \delta_{-1} + \frac{1+u(x)}{2} \delta_{+1} & \text{for } x \leq x_b \\
\delta_{u'(x)} & \text{for } x > x_b.
\end{cases}
\]

For Algorithm \((A_\Theta^{\text{adaptive}})\) we used \(m = 4\) and \(\mathcal{T}_\ell = \{[0, 1/4], [1/4, 1/2], [1/2, 3/4], [3/4, 1]\}\. 
Note that the weighted jumps $h_E \| [\lambda_{d,h} \cdot n_E] \|_{L^2(E)}^2$ of $\lambda_{d,h}$ across edges $E \in E_\Omega$ are in the one-dimensional situation given by
\[
\max\{h_{T_1}, h_{T_2}\}(\lambda_{d,h}|_{T_1} - \lambda_{d,h}|_{T_2})^2
\]
for $z \in K$, $T_1, T_2 \in T$ such that $z = T_1 \cap T_2$.

**Figure 3.** Error and error estimators in Example 7.1 for uniform and adaptive mesh refinement.

We ran Algorithm $(A_0^{\text{adaptive}})$ and $(A_{1/2}^{\text{adaptive}})$ in Example 7.1. The obtained error estimators $\eta_R$, $\eta_{Z,R}$, and $\eta_{Z,E}$ and the exact error $\|\sigma - \lambda_{d,h}\|$ for each triangulation are plotted against the degrees of freedom in $T$ in Figure 3 with a logarithmic scaling used for both axes. Both, uniform and adaptive, refinement strategies yield the same experimental convergence rates but the adaptive scheme yields a comparable error reduction at similar numbers of degrees of freedom. The error estimators $\eta_R$ and $\eta_{Z,R}$ converge much slower than the error itself while the efficient error estimator $\eta_{Z,E}$ approximates the error very well and converges with the same order.

<table>
<thead>
<tr>
<th>triangulation</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td># elements</td>
<td>4</td>
<td>8</td>
<td>16</td>
<td>32</td>
<td>64</td>
<td>128</td>
<td>256</td>
<td>512</td>
<td>1,024</td>
</tr>
<tr>
<td># atoms</td>
<td>179</td>
<td>433</td>
<td>1,122</td>
<td>3,034</td>
<td>8,385</td>
<td>23,443</td>
<td>65,921</td>
<td>185,908</td>
<td>525,057</td>
</tr>
<tr>
<td># active atoms</td>
<td>8.5</td>
<td>12.0</td>
<td>7.9</td>
<td>10.7</td>
<td>9.6</td>
<td>17.4</td>
<td>46.3</td>
<td>64.5</td>
<td>127.1</td>
</tr>
</tbody>
</table>

**Table 1.** Possible and active atoms per element on uniform meshes.

In Tables 1 and 2 we displayed for uniform and adapted meshes, respectively, the number of possible atoms per element and the average number of active atoms per element selected by $(A^{\text{active set}})$. We observe that the numbers of atoms is significantly reduced by the active set strategy. Moreover, the average number of active atoms seems to be bounded or maybe grows very slowly on the adapted meshes while on the uniform meshes the number of active atoms grows linearly.
Table 2. Possible and active atoms per element on adapted meshes.

<table>
<thead>
<tr>
<th>triangulation</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td># elements</td>
<td>45</td>
<td>81</td>
<td>120</td>
<td>187</td>
<td>321</td>
<td>492</td>
<td>741</td>
<td>1,280</td>
</tr>
<tr>
<td># atoms</td>
<td>4,992</td>
<td>11,881</td>
<td>21,297</td>
<td>41,244</td>
<td>92,450</td>
<td>175,143</td>
<td>323,390</td>
<td>733,574</td>
</tr>
<tr>
<td># active atoms</td>
<td>6.6</td>
<td>7.4</td>
<td>11.6</td>
<td>38.9</td>
<td>31.2</td>
<td>29.6</td>
<td>27.4</td>
<td>30.7</td>
</tr>
</tbody>
</table>

Example 7.2 (Two-dimensional, scalar three-well problem). Let $n = 2$, $\Omega = (0,1)^2$, $W$ as in Theorem 5.1, $\alpha = 0$, $\Gamma_D = \partial \Omega$, and, for $(x,y) \in \overline{\Omega}$, $u_D(x,y) = v(x) + v(y)$, where, for $t \in [0,1],$

$$v(t) = \begin{cases} (t - 1/4)^3/6 + (t - 1/4)/8 & \text{for } t \leq 1/4, \\ -(t - 1/4)^5/40 - (t - 1/4)^3/8 & \text{for } t \geq 1/4. \end{cases}$$

Setting $f := -\text{div} DW^*(\nabla u_D)$, i.e., for $(x,y) \in (0,1)^2,$

$$f(x,y) = \begin{cases} 0 & \text{for } x \leq 1/4 \text{ and } y \leq 1/4, \\ -2v''(y) & \text{for } x \leq 1/4 \text{ and } 1/4 \leq y, \\ -2v''(x) & \text{for } 1/4 \leq x \text{ and } y \leq 1/4, \\ -2(v''(x) + v''(y)) & \text{for } 1/4 \leq x \text{ and } 1/4 \leq y, \end{cases}$$

we have that $u = u_D$ is the weak limit of an infimizing sequence for $(P)$. If $u_x$ and $u_y$ abbreviate $\partial u/\partial x$ and $\partial u/\partial y$, respectively, then for

$$\nu(x,y) := \begin{cases} (1 - u_x(x,y) - u_y(x,y))\delta_{(0,0)} & \text{for } x \leq 1/4 \text{ and } y \leq 1/4, \\ +u_x(x,y)\delta_{(1,0)} + u_y(x,y)\delta_{(0,1)} & \text{for } x \leq 1/4 \text{ and } 1/4 \leq y, \\ (1 - u_x(x,y))\delta_{(u_x(x,y),0)} + u_x(x,y)\delta_{(1,u_y(x,y))} & \text{for } x \leq 1/4 \text{ and } 1/4 \leq y, \\ (1 - u_y(x,y))\delta_{(u_x(x,y),0)} + u_y(x,y)\delta_{(u_x(x,y),0)} & \text{for } 1/4 \leq x \text{ and } y \leq 1/4, \\ \delta_{\nabla u(x,y)} & \text{for } 1/4 \leq x \text{ and } 1/4 \leq y, \end{cases}$$

the pair $(u, \nu)$ is a solution for $(EP)$. The coarsest triangulation $T_1$ consists of 32 triangles which are halved squares and we set $m = 1.5$.

Our numerical results in Example 7.2 are not as satisfying as those for Example 7.1. The Lagrange multiplier provided by the linear program solver did not satisfy the optimality conditions even when $m$ was large and all atoms were activated. We suspect that this is caused by the huge complexity of the problem. Other solvers for the linear programming problem did not find a solution when the problem became large. This indicates that efficient methods for the solution of $(EP_{d,h})$ are very important. We found however, that the quantity $DW^*(\nabla u_{d,h})$ satisfied the maximum principle and the equilibrium equation up to an absolute error of about 0.05 in Example 7.2 so that we used this quantity to activate atoms in Algorithm $(\text{A}^{\text{active set}}_{1/2})$ and to calculate error indicators $\eta_R$, $\eta_{Z,E}$, and $\eta_{Z,E}$ in order to refine the mesh and to estimate the error in Algorithm $(\text{A}^{\text{adaptive}}_{1/2})$.

Figure 4 shows the adaptively generated mesh $T_6$ and the support of the discrete Young measure solution and the corresponding volume fractions restricted to three different elements. The three meshes show every tenth atom in $\tau$ and circles indicate that an atom is active. Numbers next to circles are volume fractions provided they are larger than 0.01. We observe that the discrete Young measure approximates the Young measure solution $\nu$ from
Example 7.2 very well. Moreover, the adaptive algorithm refines the mesh in those regions where the stress is large. Since the error estimators and the active set strategy show the same behavior as in the previous example we omit the corresponding plots and tables here.

<table>
<thead>
<tr>
<th>dof</th>
<th>9</th>
<th>35</th>
<th>70</th>
<th>162</th>
<th>255</th>
<th>492</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU-time [s]</td>
<td>11.7</td>
<td>269.0</td>
<td>830.7</td>
<td>5,792.7</td>
<td>9,797.4</td>
<td>24,317.5</td>
</tr>
</tbody>
</table>

**Table 3.** CPU-times for $(EP_{d,h})$ on adaptively refined meshes in Example 7.2.

Table 3 displays the CPU-time needed to solve $(EP_{d,h})$ in Example 7.2 on a sequence of adaptively refined triangulations against the number of degrees of freedom in $T_k$, $k = 1, \ldots, 6$. The numerical solutions were obtained on a SUN Enterprise with 14 processors and 14 GB RAM and the numbers suggest that the CPU-time depends linearly on the number of degrees of freedom.

**Acknowledgments.** The author wishes to thank Professor C. Carstensen for stimulating discussions. The author gratefully acknowledges support within the Priority Program “Analysis, Modeling and Simulation of Multiscale Problems” and within the Graduate School “Efficient Algorithms and Multiscale Methods” of the German Research Foundation (DFG).

**References**


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