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:

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

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 $I : \mathcal{A} \to \mathbb{R}$ is for $v \in \mathcal{A}$ 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 $\alpha \ge 0$. An energy density W that can be derived from a three dimensional model with one-dimensional symmetry [BHJPS] is given by N + 1 wells $s_0, ..., s_N \in \mathbb{R}^n$ and numbers $s_0^0, ..., s_N^0 \in \mathbb{R}$ and reads

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

This function W serves as a model energy density but more generally we will consider mappings $W : \mathbb{R}^n \to \mathbb{R}$ 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 Ω . 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 $(2 \le p < \infty)$.

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 Ω 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 Ω 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 $\mathcal{M}(\mathbb{R}^n)$ 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 $\mu \in \mathcal{M}(\mathbb{R}^n)$ 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

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

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

Infinizing 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)
$$c_1|s|^2 - c_2 \le W(s) \le c_2(1+|s|^2)$$
 for all $s \in \mathbb{R}^n$.

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

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

$$\Omega, \text{ there holds } \nabla u(x) = \int_{\mathbb{R}^n} s \, \nu_x(ds).$$

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

The Young measure ν generated by an infinizing 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 ν we obtain the extended problem (EP).

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

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

$$\overline{I}(v,\mu) := \int_{\Omega} \int_{\mathbb{R}^n} W(s) \,\mu_x(ds) \,dx + \alpha \int_{\Omega} |u_0 - v|^2 \,dx - \int_{\Omega} fv \,dx - \int_{\Gamma_N} gv \,ds_x.$$

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 $\iota : \mathcal{A} \to \mathcal{B}, u \mapsto (u, \delta_{\nabla u})$, 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 $\int_{\mathbb{R}^n} v(s) \,\delta_{\nabla u(x)}(ds) = v(\nabla u(x)), \,\mathcal{A}$ can be embedded continuously into \mathcal{B} . **Theorem 2.1** ([R], Proposition 5.2.1). (i) (EP) admits a solution.

(*ii*) $\inf_{v \in \mathcal{A}} I(v) = \min_{(w,\mu) \in \mathcal{B}} \overline{I}(w,\mu).$

(iii) The embedding $\iota : \mathcal{A} \to \mathcal{B}$ 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 $\iota : \mathcal{A} \to \mathcal{B}$ of an infimizing sequence for (P).

Carathéodory's Theorem implies that there exist solutions $(u, \nu) \in \mathcal{B}$ to (EP) such that for almost all $x \in \Omega$ the probability measure ν_x is a convex combination of at most n + 1Dirac 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 \mathcal{T} be a regular triangulation of Ω 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., $\overline{\Omega} = \bigcup_{T \in \mathcal{T}} T$, and \mathcal{T} satisfies the maximum angle condition. Therefore, $\partial\Omega$ is assumed to be polygonal. The extremal points of $T \in \mathcal{T}$ 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) $E = \operatorname{conv} \{z_1, ..., z_n\} \subseteq \partial T$ for pairwise distinct $z_1, ..., z_n \in \mathcal{N}$ and $T \in \mathcal{T}$ is denoted as \mathcal{E} . A partition $\mathcal{E} = \mathcal{E}_\Omega \cup \mathcal{E}_D \cup \mathcal{E}_N$ is given by $\mathcal{E}_N := \{E \in \mathcal{E} : E \subseteq \overline{\Gamma}_N\}, \mathcal{E}_D := \{E \in \mathcal{E} : E \subseteq \Gamma_D\}, \text{ and } \mathcal{E}_\Omega := \mathcal{E} \setminus (\mathcal{E}_D \cup \mathcal{E}_N)$. The set

$$\mathcal{L}^{k}(\mathcal{T}) := \{ v_{h} \in L^{\infty}(\Omega) : \forall T \in \mathcal{T}, v_{h} |_{T} \in P_{k}(T) \}$$

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

$$\mathcal{S}^{1}(\mathcal{T}) := \mathcal{L}^{1}(\mathcal{T}) \cap C(\overline{\Omega}) \text{ and } \mathcal{S}^{1}_{D}(\mathcal{T}) := \{u_{h} \in \mathcal{S}^{1}(\mathcal{T}) : u_{h}|_{\Gamma_{D}} = 0\} \subseteq W^{1,2}_{D}(\Omega)$$

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

The nodal interpolation operator associated to a triangulation \mathcal{T} is denoted by $P_{\mathcal{T}}$. If τ is a triangulation of a convex domain ω and $v \in C(\overline{\omega})$ we extend $P_{\tau}v$ to \mathbb{R}^n by setting $P_{\tau}v(s) = P_{\tau}v(\mathcal{P}_{\omega}(s))$ where \mathcal{P}_{ω} 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_{Γ_N} on Γ_N is continuous, g is continuous. We set

(3.1)
$$\mathcal{S}_N^1(\mathcal{T},g) := \{ \tau_h \in \mathcal{S}^1(\mathcal{T})^n : \forall E \in \mathcal{E}_N \; \forall z \in E \cap \mathcal{N}, \; \tau_h(z) \cdot n_{\Gamma_N} |_E = g(z) \}$$

and note that $\mathcal{S}_N^1(\mathcal{T},g) \neq \emptyset$ if n=2. We will assume that $\mathcal{S}_N^1(\mathcal{T},g) \neq \emptyset$ 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 $\|\cdot\|_{L^p(\Omega;\mathbb{R}^\ell)}$, and $\|\cdot\|$ abbreviates $\|\cdot\|_{L^2(\Omega)}$. The operator $\partial_{\mathcal{E}} \cdot /\partial s$ 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 $\mathcal{Y}_2(\Omega; \mathbb{R}^n)$ 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 τ of ω with nodes \mathcal{N}_{τ} and \mathcal{T} of Ω we set

$$YM_{d,h}(\Omega; \mathbb{R}^n) := \{ \nu_{d,h} \in \mathcal{Y}_2(\Omega; \mathbb{R}^n) : \forall z \in \mathcal{N}_\tau \, \exists a_z \in \mathcal{L}^0(\mathcal{T}), \, a_z \ge 0 \text{ and} \\ \sum_{z \in \mathcal{N}_\tau} a_z(x) = 1 \text{ a.e. in } \Omega, \, \nu_{d,h,x} = \sum_{z \in \mathcal{N}_\tau} a_z(x)\delta_z \text{ for a.e. } x \in \Omega \},$$

where δ_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 τ and \mathcal{T} , respectively, and refer to τ and \mathcal{T} through these quantities.

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

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

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 ω 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 ω 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 Ω with N^n free nodes and a triangulation τ of ω with N^n 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 $v_h \in \mathcal{S}_D^1(\mathcal{T})$, there holds

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

Remark 3.1. The elementwise constant function $\lambda_{d,h}$ is the Lagrange multiplier for the constraint $\nabla u_{d,h}|_T = \int_{\mathbb{R}^n} s \,\nu_{d,h}|_T(ds), T \in \mathcal{T}$, in $(EP_{d,h})$. For the practical implementation a bounded domain ω and a finite discretization of ω has to be chosen. We formulate appropriate computable conditions that imply Lemma 3.1.

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

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

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 $2r_0c_1 \ge \|\lambda\|_{L^{\infty}(\Omega)}$, and if for almost all $x \in \Omega$

(3.2)
$$r_0 \|\lambda\|_{L^{\infty}(\Omega)} - c_1 r_0^2 + c_2 \le \lambda_{d,h}(x) \cdot s_x^* - W(s_x^*),$$

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

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

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

Since $2r_0c_1 \geq \|\lambda\|_{L^{\infty}(\Omega)}$ the mapping $t \mapsto \|\lambda_{d,h}\|_{L^{\infty}(\Omega)}t - c_1t^2 + c_2$, $t \geq r_0$, is monotonically decreasing and since $\|\lambda_{d,h}\|_{L^{\infty}(\Omega)}r_0 - c_1r_0^2 + c_2 \leq \mathcal{H}_{\lambda_{d,h}}(x, s_x^*)$ we have $\lambda_{d,h}(x) \cdot s - P_{\tilde{\tau}}W(s) \leq \mathcal{H}_{\lambda_{d,h}}(x, s_x^*)$ for all $s \in \mathcal{N}_{\tilde{\tau}} \setminus \overline{B_{r_0}(0)}$ which implies the same estimate for all $s \in \mathbb{R}^n \setminus \overline{B_{r_0}(0)}$. \Box

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^{**}) .

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

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

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

Definition 4.1. For a solution $u \in \mathcal{A}$ 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 $u \in \mathcal{A}$ such that

Theorem 4.1 ([C1], Theorem 2). (1) admits a solution
$$u \in \mathcal{A}$$
 such that

(4.1)
$$\int_{\Omega} DW^{**}(\nabla u) \cdot \nabla v \, dx - 2\alpha \int_{\Omega} (u_0 - u) v \, dx - \int_{\Omega} f v \, dx - \int_{\Gamma_N} g v \, ds_x = 0$$

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

(4.2)
$$|DW^{**}(F) - DW^{**}(G)|^2 \le C \left(DW^{**}(F) - DW^{**}(G) \right) \cdot (F - G)$$

then for two solutions $u, w \in \mathcal{A}$ for (P^{**}) there holds $DW^{**}(\nabla u) = DW^{**}(\nabla w)$, i.e., σ 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 $(u, \nu) \in \mathcal{B}$ 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]

$$\int_{\mathbb{R}^n} DW(s)\,\nu_x(ds) = DW^{**}(\nabla w(x)).$$

(iii) A result in [CaM] shows $\sigma \in W^{1,2}_{loc}(\Omega)$.

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 $P_{\tau}W$. To do this we apply the concept of subdifferentials.

Definition 4.2. For a convex function $V: \mathbb{R}^n \to \mathbb{R}$ and $\varsigma \in \mathbb{R}^n$ the subdifferential of V at ς is defined by

$$\partial V(\varsigma) := \{ \xi \in \mathbb{R}^n : V(\varsigma + \zeta) - V(\varsigma) \ge \zeta \cdot \xi \text{ for all } \zeta \in \mathbb{R}^n \}.$$

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

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 $W_d^{cx} := ((P_\tau W)|_{\overline{\omega}})^{**}$ denote the convexification of the restriction of $P_\tau W$ to $\overline{\omega}$. Assume that $(u_{d,h}, \nu_{d,h}) \in \mathcal{B}_{d,h}$ and $\lambda_{d,h} \in \mathcal{L}^0(\mathcal{T})^n$ satisfy the conditions of Lemma 3.2. Then $(u_{d,h}, \nu_{d,h})$ minimizes the modified energy functional

$$\overline{I}'(v_h,\mu_{d,h}) := \int_{\Omega} \int_{\mathbb{R}^n} W_d^{cx}(s) \,\mu_{d,h,x}(ds) \,dx + \alpha \int_{\Omega} |u_0 - v_h|^2 \,dx - \int_{\Omega} f v_h \,dx - \int_{\Gamma_N} g v_h \,ds_x,$$

among all $(v_h, \mu_{d,h}) \in \mathcal{B}_{d,h}$. Moreover,

$$\lambda_{d,h}(x) \in \partial W_d^{cx}(\nabla u_{d,h}(x)) \qquad for \ a.e. \ x \in \Omega$$

Proof. For $s \in \omega$ we have by Carathéodory's Theorem [R],

$$W_{d}^{cx}(s) = ((P_{\tau}W)|_{\overline{\omega}})^{**}(s) = \inf_{\substack{s_{1},\dots,s_{n+1}\in\overline{\omega},\\\theta_{1},\dots,\theta_{n+1}\in[0,1],\\\sum_{i=1}^{n+1}\theta_{i}=1,\sum_{i=1}^{n+1}\theta_{i}s_{i}=s}} \sum_{i=1}^{n+1} \theta_{i}P_{\tau}W(s_{i})$$

Since $P_{\tau}W|_{\overline{\omega}}$ is τ -elementwise affine, it suffices to use the nodal values of $P_{\tau}W$ in the calculation of W_d^{cx} , i.e.,

(4.3)
$$W_d^{cx}(s) = ((P_\tau W)|_{\overline{\omega}})^{**}(s) = \inf_{\substack{\theta_z \in [0,1], \\ \sum_{z \in \mathcal{N}_\tau} \theta_z = 1, \\ \sum_{z \in \mathcal{N}_\tau} \theta_z z = s}} \sum_{z \in \mathcal{N}_\tau} \theta_z P_\tau W(z).$$

Assume that there exists $s \in \operatorname{conv} \{z_1, ..., z_{n+1}\} = t \in \tau, z_1, ..., z_{n+1} \in \mathcal{N}_{\tau}$ such that $s = \sum_{i=1}^{n+1} \alpha_i z_i$ but $W_d^{cx}(s) \neq \sum_{i=1}^{n+1} \alpha_i W_d^{cx}(z_i)$ with $\alpha_i \in [0, 1], \sum_{i=1}^{n+1} \alpha_i = 1$. If $W_d^{cx}(s) > \sum_{i=1}^{n+1} \alpha_i W_d^{cx}(z_i)$ then $W_d^{cx}(s)$ was not convex. If $W_d^{cx}(s) < \sum_{i=1}^{n+1} \alpha_i W_d^{cx}(z_i)$ then W_d^{cx} was not the largest convex function satisfying $W_d^{cx} \leq P_\tau W|_{\overline{\omega}}$. Therefore, $W_d^{cx}(s) = \sum_{i=1}^{n+1} \alpha_i W_d^{cx}(z_i)$, so that W_d^{cx} is τ -elementwise affine and $P_{\tau}W_d^{cx}|_{\overline{\omega}} = W_d^{cx}$. To prove that $(u_{d,h}, \nu_{d,h})$ minimizes

the functional \overline{I}' it suffices to verify the optimality conditions from Lemma 3.2 with $P_{\tau}W$ replaced by $P_{\tau}W_d^{cx}$. For this it is sufficient to show that, for almost all $x \in \Omega$, there holds

(4.4)
$$\max_{s \in \omega} (\lambda_{d,h}(x) \cdot s - P_{\tau}W(s)) = \max_{s \in \omega} (\lambda_{d,h}(x) \cdot s - W_d^{cx}(s))$$

and

(4.5)
$$\int_{\mathbb{R}^n} (\lambda_{d,h}(x) \cdot s - W_d^{cx}(s)) \nu_{d,h,x}(ds) = \int_{\mathbb{R}^n} (\lambda_{d,h}(x) \cdot s - P_\tau W(s)) \nu_{d,h,x}(ds).$$

Since $W_d^{cx} \leq P_{\tau} W(s)|_{\overline{\omega}}$, we only have to show that

$$\max_{s \in \omega} (\lambda_{d,h}(x) \cdot s - P_{\tau}W(s)) \ge \max_{s \in \omega} (\lambda_{d,h}(x) \cdot s - W_d^{cx}(s))$$

and

$$\int_{\mathbb{R}^n} W_d^{cx}(s) \,\nu_{d,h,x}(ds) \ge \int_{\mathbb{R}^n} P_{\tau}(s) \,\nu_{d,h,x}(ds).$$

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

$$\lambda_{d,h}(x) \cdot \overline{s} - W_d^{cx}(\overline{s}) = \max_{s \in \omega} (\lambda_{d,h}(x) \cdot s - W_d^{cx}(s)).$$

By definition of W_d^{cx} there exist $\theta_1, ..., \theta_{n+1} \in [0, 1]$, $\sum_{i=1}^{n+1} \theta_i = 1$, and $z_1, ..., z_{n+1} \in \mathcal{N}_{\tau}$ such that $\sum_{i=1}^{n+1} \theta_i z_i = \overline{s}$ and $W_d^{cx}(\overline{s}) = \sum_{i=1}^{n+1} \theta_i P_{\tau} W(z_i)$. By linearity of $s \mapsto \lambda_{d,h}(x) \cdot s$ we have

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

which proves (4.4). If

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

the explicit representation of W_d^{cx} contradicts the fact that $(u_{d,h}, \nu_{d,h})$ is minimal for \overline{I} . We have thus shown (4.5) which yields the optimality conditions. The maximum principle of Lemma 3.1, the convexity of the mapping $s \mapsto W_d^{cx}(s) - \lambda_{d,h}(x) \cdot s$ 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 $x \in \Omega$,

$$\max_{s \in \omega} (\lambda_{d,h}(x) \cdot s - W_d^{cx}(s)) = \int_{\mathbb{R}^n} (\lambda_{d,h}(x) \cdot s - W_d^{cx}(s)) \nu_{d,h,x}(ds)$$
$$\leq \lambda_{d,h}(x) \cdot \nabla u_{d,h}(x) - W_d^{cx}(\nabla u_{d,h}(x)).$$

Remark 4.2 shows, for almost all $x \in \Omega$, $0 \in -\lambda_{d,h}(x) + \partial W_d^{cx}(\nabla u_{d,h}(x))$.

Another definition is needed for the *a priori* and *a posteriori* error estimates. It concerns the approximation of DW^{**} by the multi-valued mapping ∂W_d^{cx} .

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 $\lambda_{d,h} \in \mathcal{L}^0(\mathcal{T})^n$ satisfy the conditions of Lemma 3.1. There holds

$$\begin{aligned} \|\sigma - \lambda_{d,h}\| + \alpha \|u - u_{d,h}\| &\leq C \inf_{(v_h, \mu_{d,h}) \in \mathcal{B}_{d,h}} \left(\|\nabla (u - v_h)\| + \alpha \|u - v_h\| \right) \\ &+ C \|\partial W_d^{cx} - DW^{**}\|_{L^{\infty}(\omega; 2^{\mathbb{R}^n})} + |\Omega|\sqrt{C'} \|\partial W_d^{cx} - DW^{**}\|_{L^{\infty}(\omega; 2^{\mathbb{R}^n})}^{1/2}. \end{aligned}$$

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

$$\begin{aligned} \frac{1}{2} \|\sigma - \lambda_{d,h}\|^2 &\leq \|\sigma - DW^{**}(\nabla u_{d,h})\|^2 + \|DW^{**}(\nabla u_{d,h}) - \lambda_{d,h}\|^2 \\ &\leq C \int_{\Omega} (DW^{**}(\nabla u) - DW^{**}(\nabla u_{d,h})) \cdot \nabla(u - u_{d,h}) \, dx + \|DW^{**}(\nabla u_{d,h}) - \lambda_{d,h}\|^2 \\ &= C \int_{\Omega} (DW^{**}(\nabla u) - \lambda_{d,h}) \cdot \nabla(u - u_{d,h}) \, dx + C \int_{\Omega} (\lambda_{d,h} - DW^{**}(\nabla u_{d,h})) \cdot \nabla(u - u_{d,h}) \, dx \\ &+ \|\lambda_{d,h} - DW^{**}(\nabla u_{d,h})\|^2 \\ &\leq C \int_{\Omega} (DW^{**}(\nabla u) - \lambda_{d,h}) \cdot \nabla(u - u_{d,h}) \, dx + C \|\lambda_{d,h} - DW^{**}(\nabla u_{d,h})\| \|\nabla(u - u_{d,h})\| \\ &+ \|\lambda_{d,h} - DW^{**}(\nabla u_{d,h})\|^2. \end{aligned}$$

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

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

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 $\|u - u_{d,h}\|$,

$$\begin{aligned} \|\sigma - \lambda_{d,h}\|^2 + \alpha \|u - u_{d,h}\|^2 &\leq C \big(\|\nabla (u - u_{d,h} - w_h)\|^2 + \alpha \|u - u_{d,h} - w_h\|^2 \\ &+ \|\lambda_{d,h} - DW^{**}(\nabla u_{d,h})\| \|\nabla (u - u_{d,h})\| + \|\lambda_{d,h} - DW^{**}(\nabla u_{d,h})\|^2 \big). \end{aligned}$$

Lemma 4.1 ensures $\lambda_{d,h}(x) \in \partial W_d^{cx}(\nabla u_{d,h}(x))$ 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

$$\begin{aligned} \|\lambda_{d,h} - DW^{**}(\nabla u_{d,h})\|^{2} &\leq \int_{\Omega} \sup_{s \in \partial W_{d}^{cx}(\nabla u_{d,h}(x)) - DW^{**}(\nabla u_{d,h})} |s|^{2} dx \\ &\leq |\Omega| \sup_{t \in \omega} \sup_{s \in \partial W_{d}^{cx}(t) - DW^{**}(t)} |s|^{2} = |\Omega| \|\partial W_{d}^{cx} - DW^{**}\|_{L^{\infty}(\omega;2^{\mathbb{R}^{n}})}^{2}. \end{aligned}$$

Letting $w_h = v_h - u_{d,h}$ for arbitrary $(v_h, \mu_{d,h}) \in \mathcal{B}_{d,h}$ and estimating $\|\nabla(u - u_{d,h})\| \leq C$ (which follows from the growth conditions (2.1)) we verify the assertion of the theorem. \Box

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

Remarks 4.3. (i) Theorem 4.2, Theorem 5.1 below, and density of finite element spaces in \mathcal{A} prove $\lambda_{d,h} \to \sigma$ in $L^2(\Omega)$ for $(d, h_{\mathcal{T}}) \to 0$ and, if $\alpha > 0$, we also have $u_{d,h} \to u$ 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 $u_{d,h} \to u$ 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 ω 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 $\lambda_{d,h} \in \mathcal{L}^0(\mathcal{T})^n$ define

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

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

$$\begin{aligned} \|\sigma - \lambda_{d,h}\|^{2} + \alpha \|u - u_{d,h}\|^{2} &\leq C \Big\{ \Big(\sum_{T \in \mathcal{T}} h_{T}^{2} \| (f + \operatorname{div} \lambda_{d,h} + 2\alpha (u_{0} - u_{d,h})) \|_{L^{2}(T)}^{2} \Big)^{1/2} \\ &+ \Big(\sum_{E \in \mathcal{E}_{\Omega} \cup \mathcal{E}_{N}} h_{E} \| [\lambda_{d,h} \cdot n_{E}] \|_{L^{2}(E)}^{2} \Big)^{1/2} + \|\partial W_{d}^{cx} - DW^{**} \|_{L^{\infty}(\omega; 2^{\mathbb{R}^{n}})} + \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}}^{2} u_{D} / \partial s^{2} \|_{L^{2}(\Gamma_{D})} \Big\} \\ &+ |\Omega| \| \partial W_{d}^{cx} - DW^{**} \|_{L^{\infty}(\omega; 2^{\mathbb{R}^{n}})}^{2}. \end{aligned}$$

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 $v_h \in \mathcal{S}_D^1(\mathcal{T})$, there holds

$$C\|\sigma - \lambda_{d,h}\|^{2} + 2\alpha \|u - u_{d,h}\|^{2} \leq \int_{\Omega} (DW^{**}(\nabla u) - \lambda_{d,h}) \cdot \nabla (u - u_{d,h} - w - v_{h}) dx + 2\alpha \int_{\Omega} (u - u_{d,h}) (u - u_{d,h} - w - v_{h}) dx + |\Omega| \|\partial W^{cx}_{d} - DW^{**}\|^{2}_{L^{\infty}(\omega;2^{\mathbb{R}^{n}})} + C|\Omega| \|\partial W^{cx}_{d} - DW^{**}\|_{L^{\infty}(\omega;2^{\mathbb{R}^{n}})} + \|\sigma - \lambda_{d,h}\| \|\nabla w\| + 2\alpha \|u - u_{d,h}\| \|w\|.$$

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

(4.6)
$$\|\nabla v_h\| + \|h_{\mathcal{T}}^{-1}(v-v_h)\| + \|h_{\mathcal{E}}^{-1/2}(v-v_h)\|_{L^2(\cup\mathcal{E})} \le C \|\nabla v\|_{10}$$

The Euler-Lagrange equations (4.1) for u, an elementwise integration by parts, and (4.6)show for $v := u - u_{d,h} - w \in W_D^{1,2}(\Omega)$

$$\begin{split} \int_{\Omega} (DW^{**}(\nabla u) - \lambda_{d,h}) \cdot \nabla(v - \mathcal{J}v) \, dx + 2\alpha \int_{\Omega} (u - u_{d,h})(v - \mathcal{J}v) \, dx \\ &= \sum_{T \in \mathcal{T}} \int_{T} (f + \operatorname{div} \lambda_{d,h})(v - \mathcal{J}v) \, dx + 2\alpha \int_{\Omega} (u_0 - u_{d,h})(v - \mathcal{J}v) \, dx \\ &\quad + \sum_{E \in \mathcal{E}_{\Omega} \cup \mathcal{E}_N} \int_{E} [\lambda_{d,h} \cdot n_E](v - \mathcal{J}v) \, ds_x \\ &\leq C \Big(\Big(\sum_{T \in \mathcal{T}} h_T^2 \| (f + \operatorname{div} \lambda_{d,h} + 2\alpha(u_0 - u_{d,h})) \|_{L^2(T)}^2 \Big)^{1/2} \\ &\quad + \Big(\sum_{E \in \mathcal{E}_{\Omega} \cup \mathcal{E}_N} h_E \| [\lambda_{d,h} \cdot n_E] \|_{L^2(E)}^2 \Big)^{1/2} \Big) \| \nabla v \|. \end{split}$$

The combination of the last two estimates together with the a priori bound $\|\nabla v\| \leq \|\nabla (u - v)\|$ $|u_{d,h}\rangle \| + \|\nabla w\| \le C$ and $\min_{w|_{\Gamma_D} = u_D - u_{D,h}} \|w\|_{W^{1,2}(\Omega)} \le C \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}}^2 u_D / \partial s^2\|_{L^2(\Gamma_D)}^2$ (cf. [BCD], Lemma 3.1) shows the assertion after absorption of $\|\sigma - \lambda_{d,h}\|$ and $\|u - u_{d,h}\|$.

Remarks 4.4. (i) The term $\|h_{\mathcal{E}}^{3/2}\partial_{\mathcal{E}}^2 u_D/\partial s^2\|_{L^2(\Gamma_D)}$ is of higher order. (ii) The terms $\|\partial W_d^{cx} - DW^{**}\|_{L^{\infty}(\omega;2\mathbb{R}^n)}^2$ and $\|\partial W_d^{cx} - DW^{**}\|_{L^{\infty}(\omega;2\mathbb{R}^n)}$ are of higher order provided $d \ll h_{\mathcal{T}}$ (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 $h_{\tau}^{1/2}$ 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 $u \in \mathcal{A}$ solves (P^{**}) . Let $(u_{d,h}, \nu_{d,h}) \in$ $\mathcal{B}_{d,h}$ and $\lambda_{d,h} \in \mathcal{L}^0(\mathcal{T})^n$ satisfy the conditions of Lemma 3.1. If $\alpha = 0$ and $f \in W^{1,2}(\Omega)$ then

$$\begin{aligned} \|\sigma - \lambda_{d,h}\|^{2} &\leq C \Big\{ \min_{\tau_{h} \in \mathcal{S}_{N}^{1}(\mathcal{T},g)} \|\lambda_{d,h} - \tau_{h}\| + \|h_{\mathcal{T}}^{2} \nabla f\| + \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}}^{2} u_{D} / \partial s^{2}\|_{L^{2}(\Gamma_{D})} \\ &+ \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}} g / \partial s\|_{L^{2}(\Gamma_{N})} + \|\partial W_{d}^{cx} - DW^{**}\|_{L^{\infty}(\omega;2^{\mathbb{R}^{n}})} \Big\} + |\Omega| \|\partial W_{d}^{cx} - DW^{**}\|_{L^{\infty}(\omega;2^{\mathbb{R}^{n}})}^{2}. \end{aligned}$$

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 $v_h \in \mathcal{S}_D^1(\mathcal{T})$

$$C\|\sigma - \lambda_{d,h}\|^{2} \leq \int_{\Omega} (DW^{**}(\nabla u) - \lambda_{d,h}) \cdot \nabla(u - u_{d,h} - w - v_{h}) dx + |\Omega| \|\partial W_{d}^{cx} - DW^{**}\|_{L^{\infty}(\omega;2^{\mathbb{R}^{n}})}^{2} + C|\Omega| \|\partial W_{d}^{cx} - DW^{**}\|_{L^{\infty}(\omega;2^{\mathbb{R}^{n}})}^{2} + \|\sigma - \lambda_{d,h}\| \|\nabla w\|.$$

Letting $\tau_h \in \mathcal{S}_N^1(\mathcal{T}, g)$ and writing $v := u - u_{d,h} - w \in W_D^{1,2}(\Omega)$ and $v_h := \mathcal{J}v \in \mathcal{S}_D^1(\mathcal{T})$ we verify, using div $\lambda_{d,h}|_T = 0$, the Euler-Lagrange equation (4.1), an integration by parts, and Hölder's inequality,

$$\begin{split} \int_{\Omega} (DW^{**}(\nabla u) - \lambda_{d,h}) \cdot \nabla(v - \mathcal{J}v) \, dx &\leq \int_{\Omega} f(v - \mathcal{J}v) \, dx + \sum_{T \in \mathcal{T}} \int_{T} \operatorname{div} \left(\tau_h - \lambda_{d,h}\right) (v - \mathcal{J}v) \, dx \\ &+ \int_{\Gamma_N} (g - \tau_h \cdot n_{\Gamma_N}) (v - \mathcal{J}v) \, ds_x + \|\tau_h - \lambda_{d,h}\| \|\nabla(v - \mathcal{J}v)\|. \end{split}$$

The estimate (cf. [CB], Theorem 2.1) $\int_{\Omega} f(v - \mathcal{J}v) dx \leq C \|\nabla v\| \|h_{\mathcal{T}}^2 \nabla f\|$ and (4.6) yield

$$\int_{\Omega} (DW^{**}(\nabla u) - \lambda_{d,h}) \cdot \nabla(v - \mathcal{J}v) \, dx$$

$$\leq C \Big\{ \|h_T^2 \nabla f\| + \big(\sum_{T \in \mathcal{T}} h_T^2 \|\operatorname{div} \left(\tau_h - \lambda_{d,h}\right)\|_{L^2(T)}^2 \big)^{1/2} + \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}} g / \partial s\|_{L^2(\Gamma_N)} + \|\tau_h - \lambda_{d,h}\| \Big\} \|\nabla v\|.$$

Choosing w as in [BCD] and employing elementary results about nodal interpolation on Γ_N we infer

$$\|w\|_{W^{1,2}(\Omega)} + \|g - \tau_h \cdot n_{\Gamma_N}\|_{L^2(\Gamma_N)} \le C \Big(\|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}}^2 u_D / \partial s^2\|_{L^2(\Gamma_D)} + \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}} g / \partial s\|_{L^2(\Gamma_N)} \Big)$$

and using an elementwise inverse estimate of the form

$$h_T \|\operatorname{div} \left(\tau_h - \lambda_{d,h}\right)\|_{L^2(T)} \le C \|\tau_h - \lambda_{d,h}\|_{L^2(T)} \quad \text{for all } T \in \mathcal{T}$$

we verify the assertion as in the proof of the preceding theorem.

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 $\|\lambda_{d,h} - \lambda_{d,h}^*\|$ for any choice of a smooth approximation $\lambda_{d,h}^* \in \mathcal{S}_N^1(\mathcal{T},g)$ 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 σ is smooth but with different exponents,

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

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 $DW^{**} - \partial W_d^{cx}$ 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 $\omega = (-m,m)^2$, $m \ge 1$, there exists a triangulation τ of ω with maximal mesh-size d = 1/k, k a positive integer, of ω such that

$$\|\partial W_d^{cx} - DW^{**}\|_{L^{\infty}(\omega;2^{\mathbb{R}^n})} \le Cd \, \|D^2 W^{**}\|_{L^{\infty}(\omega)}$$
12

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

 f_{2}

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

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

For d = 1/k, k a positive integer, choose τ 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 τ -elementwise differentiable it therefore suffices to show for each $t \in \tau$

$$\|DW_d^{cx} - DW^{**}\|_{L^{\infty}(t)} \le d\|D^2W^{**}\|_{L^{\infty}(t)}.$$

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

(5.1)
$$\|DW_d^{cx} - DW^{**}\|_{L^{\infty}(t)} \leq \|DW_d^{cx} - DW_d^{**}\|_{L^{\infty}(t)} + \|DW_d^{**} - DW^{**}\|_{L^{\infty}(t)} \\ \leq \|DW_d^{cx} - DW_d^{**}\|_{L^{\infty}(t)} + Cd \|D^2W^{**}\|_{L^{\infty}(t)}.$$

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 (5.2) $W_d^{cx}(x) = \sup_{k \in \tau} a_k(x)$

and $W_d^{cx}|_k = a_k$. If $k \subseteq X_I \cup X_{II} \cup X_{III} \cup X_{IV} \cup X_V \cup X_{VI}$ we define a_k such that $a_k(z) = W^{**}(z)$ for all $z \in k \cap \mathcal{N}_{\tau}$. 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

$$a_k(x) = W(1+jd,jd) + (x_1 - 1 - jd, x_2 - jd) \cdot (1,1)$$

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

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 $k \subset X_I \cup X_{II} \cup ... \cup X_{VI}$ 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 $k \subseteq A_j = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 - x_2 \in [-1, 1], x_1 + x_2 \in 1 + 2d[j, j + 1)\}, j \ge 0$ there holds $W_d^{cx} = W^{**}$ on ∂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

$$\begin{aligned} \|DW_d^{cx} - DW_d^{**}\|_{L^{\infty}(t)} &\leq \|DW_d^{cx} - DW^{**}\|_{L^{\infty}(t)} + \|DW^{**} - DW_d^{**}\|_{L^{\infty}(t)} \\ &\leq Cd \, \|D^2W^{**}\|_{L^{\infty}(t)} \end{aligned}$$

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 dist $(A, B) := \inf_{(a,b) \in A \times B} |a-b|$. We write $\limsup_{\rho \to \rho_0} A_\rho \subseteq A_\rho$ A (i.e., A is the upper Kuratowski limit of A_{ρ} , cf. [Ku]) if

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

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

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

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

(5.3)
$$\operatorname{Lim}\sup_{d\to 0} \{F \in \mathbb{R}^2 : \exists G, S \in \mathbb{R}^2, \{S, T\} \subseteq \partial W_d^{cx}(G), S \in \partial W_d^{cx}(F)\} \\ \subseteq \{F \in \mathbb{R}^2 : T = DW^{**}(F)\},\$$

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

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

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

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

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

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

Since W^{**} is affine on conv supp ν_x , $\int_{\mathbb{R}^2} s \, d\nu_x = \nabla u(x)$, and $\operatorname{supp} \nu_x \subseteq \{E \in \mathbb{R}^2 : W(E) = W^{**}(E)\}$ 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

$$\operatorname{supp} \mu(F) = S(DW^{**}(F)),$$

so that supp $\nu_x = S(\sigma(x))$ for a.e. $x \in \Omega$. Hence

(5.4)
$$\operatorname{conv} S(T) = \{ E \in \mathbb{R}^2 : T = DW^{**}(E) \}.$$

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

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

if $x \in \Omega$ and $\lambda_{d,h}(x) \to \sigma(x)$. Because of (5.3), (5.4) and since $\operatorname{Lim} \sup_{\rho \to \rho_0} B_{\rho} \subseteq A$ if $\operatorname{Lim} \sup_{\rho \to \rho_0} A_{\rho} \subseteq A$ and $B_{\rho} \subseteq A_{\rho}$ for all ρ , we only have to show that for $T = \lambda_{d,h}(x)$

 $\operatorname{conv} \operatorname{supp} \nu_{d,h,x} \subseteq \{F \in \mathbb{R}^2 : \exists G, S \in \mathbb{R}^2, \{S,T\} \subseteq \partial W_d^{cx}(G), S \in \partial W_d^{cx}(F)\}$

in order to prove the second assertion. The set

$$M_1 := \{ G \in \mathbb{R}^2 : \exists S \in \partial W_d^{cx}(\nabla u_{d,h}), \, S \in \partial W_d^{cx}(G) \}$$

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

 W_d^{cx} affine on A and $\nabla u_{d,h}(x) \in A$.

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

 $M_1 \subseteq \{F \in \mathbb{R}^2 : \exists G, S \in \mathbb{R}^2, \{S, T\} \subseteq \partial W_d^{cx}(G), S \in \partial W_d^{cx}(F)\}$

which concludes the proof.

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 $u \in \mathcal{A}$ for the convexified problem (P^{**}) and a solution $(u_{d,h}, \nu_{d,h}) \in \mathcal{B}_{d,h}$ 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 Ω_{ms} is uniquely defined and that an appropriate approximation $\tilde{\Omega}_{m,h}$ of $\Omega_{ms,h}$ converges to Ω_{ms} .

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

 $x \in \Omega_{ms} \quad \Longleftrightarrow \quad \xi(\sigma(x)) = 0.$

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

$$\hat{\Omega}_{m,h} := \{ x \in \Omega : \xi(\lambda_{d,h}(x)) = 0 \}.$$

We then have

(5.5)
$$\|\xi(\sigma) - \xi(\lambda_{d,h})\| \le C \|\sigma - \lambda_{d,h}\|$$

and

$$x \in \tilde{\Omega}_{m,h} \implies \text{dist} \left(\nabla u_{d,h}(x), \overline{M} \right) \le C' \| \partial W_d^{cx} - DW^{**} \|_{L^{\infty}(\omega; 2^{\mathbb{R}^2})}.$$

Conversely, there holds

$$x \in \Omega_{ms,h} \implies |\xi(\lambda_{d,h}(x))| \le \|\partial W_d^{cx} - DW^{**}\|_{L^{\infty}(\omega;2^{\mathbb{R}^2})}.$$

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 $F = \nabla u(x)$

(5.6)
$$x \in \Omega_{ms} \iff F \in \overline{X}_I \cup \overline{X}_V \cup \overline{X}_{VI} \cup \overline{X}_{VII} \\ \iff (s_1 = 0 \land s_2 \le 0) \lor (s_2 = 0 \land s_1 \le 0) \lor (s_1 = s_2 \land s_1 \ge 0).$$

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 Ω_{ms} . The Lipschitz continuity of ξ implies the estimate (5.5). Let $x \in \Omega$ be such that $\xi(\lambda_{d,h}(x)) = 0$. The Lipschitz continuity of ξ and the inclusion $\lambda_{d,h}(x) \in \partial W_d^{cx}(\nabla u_{d,h}(x))$ show

$$\begin{aligned} \xi(DW^{**}(\nabla u_{d,h}(x))) &= |\xi(DW^{**}(\nabla u_{d,h}(x))) - \xi(\lambda_{d,h}(x))| \\ &\leq C|DW^{**}(\nabla u_{d,h}(x)) - \lambda_{d,h}(x)| \leq C \|DW^{**} - \partial W_d^{cx}\|_{L^{\infty}(\omega;2^{\mathbb{R}^2})}. \end{aligned}$$

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

dist
$$(F, \overline{M}) \le c\xi(DW^{**}(F))$$

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

$$\xi(DW^{**}(F)) = \min\{f_1 - 1 + f_2, f_1 - f_2 - 1\} = f_1 - f_2 - 1 \ge (f_1 - f_2 - 1)/\sqrt{2}$$
$$= \min\{f_1 - 1, (f_1 - f_2 - 1)/\sqrt{2}\} = \operatorname{dist}(F, \overline{M}).$$

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 $\lambda_{d,h}(x) \in \partial W_d^{cx}(\nabla u_{d,h}(x)) \text{ and since } \partial W_d^{cx}(\nabla u_{d,h}) = \operatorname{conv} \{ D W_d^{cx}|_t : t \in \tau, \nabla u_{d,h}(x) \in t \},$ there exist $t_1, ..., t_{n+1} \in \tau$ and $\varrho_i \in [0, 1], \sum_{i=1}^{n+1} \varrho_i = 1$ such that $\lambda_{d,h}(x) = \sum_{i=1}^{n+1} \varrho_i DW_d^{cx}|_{t_i}$. The identities

$$\lambda_{d,h}(x) = \sum_{i=1}^{n+1} \varrho_i DW_d^{cx}|_{t_i} = \sum_{i=1}^{n+1} \varrho_i (DW_d^{cx}|_{t_i} - DW^{**}(\nabla u_{d,h})) + DW^{**}(\nabla u_{d,h})$$
16

and $\xi(DW^{**}(\nabla u_{d,h})) = 0$ combined with the Lipschitz continuity of ξ show

$$\begin{aligned} |\xi(\lambda_{d,h}(x))| &= |\xi(\sum_{i=1}^{n+1} \varrho_i(DW_d^{cx}|_{t_i} - DW^{**}(\nabla u_{d,h})) + DW^{**}(\nabla u_{d,h})) - \xi(DW^{**}(\nabla u_{d,h}))| \\ &\leq C|\sum_{i=1}^{n+1} \varrho_i(DW_d^{cx}|_{t_i} - DW^{**}(\nabla u_{d,h}))| \leq C||W^{**} - \partial W_d^{cx}||_{L^{\infty}(\omega;2^{\mathbb{R}^2})}. \quad \Box \end{aligned}$$

6. Combination of A Multilevel Scheme and Adaptive Mesh Refinement

6.1. Active Set Strategy Due to Carstensen & Roubíček. The identity

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

in Lemma 3.1 for a solution $(u_{d,h}, \nu_{d,h}) \in \mathcal{B}_{d,h}$ for $(EP_{d,h})$ with multiplier $\lambda_{d,h} \in \mathcal{L}^0(\mathcal{T})^n$ states that for almost each $x \in \Omega$ the probability measure $\nu_{d,h,x}$ is supported in those atoms $z \in \mathcal{N}_{\tau}$ 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 $\nu_{d,h}$,

$$\operatorname{Supp}(\nu_{d,h}) := \{ (x, z) \in \Omega \times \mathcal{N}_{\tau} : z \in \operatorname{supp}(\nu_{d,h,x}) \}$$

where supp $(\nu_{d,h,x}) \subseteq \mathbb{R}^n$ is the support of the Radon measure $\nu_{d,h,x}$, was known a priori, we could set $A := \text{Supp}(\nu_{d,h})$ and seek $(u_{d,h}, \nu_{d,h})$ as a solution of the following lower-dimensional problem $(EP_{d,h,A})$.

$$(EP_{d,h,A}) \quad \begin{cases} \text{Seek } (u_{d,h}, \nu_{d,h}) \in \mathcal{B}_{d,h} \text{ such that } \operatorname{Supp}(\nu_{d,h}) \subseteq A \\ \text{and } \overline{I}(u_{d,h}, \nu_{d,h}) = \operatorname{inf}_{(\nu_h, \mu_{d,h}) \in \mathcal{B}_{d,h}} \overline{I}(\nu_h, \mu_{d,h}). \end{cases}$$

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 \hat{h} of $\mathcal{H}_{\lambda_{d,h}}$ we define a set of active atoms, called the *active set*, by

(6.1)
$$A = \{ (x, z) \in \Omega \times \mathcal{N}_{\tau} : \tilde{h}(x, z) \ge \max_{s \in \omega} \tilde{h}(x, s) - \varepsilon(x) \},$$

where $\varepsilon \in \mathcal{L}^0(\mathcal{T})$, $\varepsilon > 0$ almost everywhere in Ω , is a given tolerance. If ε 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 $(u_{d,h}, \nu_{d,h})$ be a solution for $(EP_{d,h})$ with corresponding multiplier $\lambda_{d,h}$ and $\mathcal{H}_{\lambda_{d,h}}(x,s) = \lambda_{d,h}(x) \cdot s - P_{\tau}W(s)$. 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 Ω be such that, for each $T \in \mathcal{T}$,

$$\|\mathcal{H}_{\lambda_{d,h}} - h\|_{L^{\infty}(T \times S_T)} \le \varepsilon|_T,$$

with $S_T \subseteq \mathbb{R}^n$ such that, for almost all $x \in T$, we have

$$\{s \in \omega : \mathcal{H}_{\lambda_{d,h}}(x,s) = \max_{\tilde{s} \in \overline{\omega}} \mathcal{H}_{\lambda_{d,h}}(x,\tilde{s})\} \cup \{s \in \omega : \tilde{h}(x,s) = \max_{\tilde{s} \in \overline{\omega}} \tilde{h}(x,\tilde{s})\} \subseteq S_T.$$

If A is defined by (6.1) then any solution for $(EP_{d,h,A})$ is a solution for $(EP_{d,h})$.

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 ($A^{active \, set}$). Let $\tau_1, \tau_2, ..., \tau_J$ be triangulations of $\omega, \, \varepsilon_1, \varepsilon_2, ..., \varepsilon_J > 0$ 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 $\lambda_{d,h} \in \mathcal{L}^0(\mathcal{T})^n$.
- (4) If the conditions of Lemma 3.2 are satisfied then go to (6) otherwise proceed with (5).

(5) Increase m if necessary. Enlarge ε by $\varepsilon|_T := 2\varepsilon|_T$ if for some $x_T \in T$

$$\max_{z \in \mathcal{N}_{\tau}} \mathcal{H}_{\lambda_{d,h}}(x_T, z) > \int_{\mathbb{R}^n} \mathcal{H}_{\lambda_{d,h}}(x_T, s) \,\nu_{h, x_T}(ds),$$

and set $\varepsilon|_T := \varepsilon|_T$ otherwise. Go to (2).

(6) If j < J proceed with (7) otherwise terminate.

(7) Set j := j + 1, $\tilde{h}(x, s) := \lambda_{d,h}(x) \cdot s - P_{\tau}W(s)$, $\varepsilon := \varepsilon_j$ 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 \mathcal{T}' . (ii) Since the tolerance ε 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 $T \in \mathcal{T}$,

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

In regard to Theorem 4.4 we employ the operator $\overline{\mathcal{A}} : L^2(\Omega)^n \to \mathcal{S}^1_N(\mathcal{T};g)$ of [CB], which is for $\Gamma_N = \emptyset$ and $p \in L^2(\Omega)^n$ given by

$$\bar{\mathcal{A}}p = \sum_{z \in \mathcal{N}} p_z \varphi_z, \quad \text{for } p_z = \int_{\varphi_z > 0} p \, dx / \int_{\varphi_z > 0} 1 \, dx,$$

to define, for $T \in \mathcal{T}$,

$$\eta_Z(T) := \|\lambda_{d,h} - \bar{\mathcal{A}}\lambda_{d,h}\|_{L^2(T)}.$$

With these definitions we have

$$\|\sigma - \lambda_{d,h}\|^2 \le C \left(\sum_{\substack{T \in \mathcal{T} \\ 18}} \eta(T)^2\right)^{1/2} + \text{h.o.t.}$$

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 τ which are of higher order provided $d \ll h_T$ and on smoothness of given right-hand sides. We set

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

Remark 4.5 (iii) states

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



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

(2) Compute a discrete solution $(u_{\ell}, \nu_{\ell}, \lambda_{\ell})$ with Algorithm $(A^{active \, set})$ and starting values $\tilde{h}_0(x, s) := \tilde{\lambda}_{\ell}(x) \cdot s - P_{\tau}W(s), J = 2, d_j = 2^{j-1}/k, k = \lfloor 4m \, 2^{-J} \text{card} \, (\mathcal{N}_{\mathcal{T}_{\ell}})^{3/2n} \rfloor \, (\lfloor s \rfloor \text{ is the largest integer } \leq s), \varepsilon_j := 2^{-\ell-j} \, 10^{-4} \text{ for } j = 1, ..., J \, (\varepsilon_1 := \infty \text{ if } \ell = 1 \text{ to activate all atoms}),$ and a triangulation τ_j of ω with maximal mesh-size d_j .

- (3) For each $T \in \mathcal{T}_{\ell}$ compute refinement indicators $\eta_Z(T)$ and $\eta_R(T)$.
- (4) Mark the element T for red-refinement if

$$\eta_R(T) \ge \Theta \max_{T' \in \mathcal{T}_\ell} \eta_R(T').$$

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

Remarks 6.2. (i) We chose k such that $d \propto h^{3/2}$ so that $\|DW^{**} - \partial W_d^{cx}\|_{L^{\infty}(\Omega;2^{\mathbb{R}^n})}$ 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 $\lambda_{\ell} \to DW^{**}(\nabla u)$ in $L^2(\Omega)$ for a solution $u \in \mathcal{A}$ for (P^{**}) , λ_{ℓ} is a Cauchy sequence, and therefore λ_{ℓ} is a good approximation for $\lambda_{\ell+1}$ if ℓ 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 $(EP_{d,h})$.

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

$$f(x) = \begin{cases} 0 & \text{for } x \le x_b, \\ \gamma(x - x_b)/2 & \text{for } x \ge x_b, \end{cases}$$

and

$$u_D(0) = 3x_b^5/128 + x_b^3/3$$
 and $u_D(1) = \gamma(1 - x_b)^3/24 + 1 - x_b$,

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

$$u(x) = \begin{cases} -3(x-x_b)^5/128 - (x-x_b)^3/3 & \text{for } x \le x_b, \\ \gamma(x-x_b)^3/24 + x - x_b & \text{for } x \ge x_b \end{cases}$$

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

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

For Algorithm $(A_{\Theta}^{adaptive})$ we used m = 4 and $\mathcal{T}_1 = \{[0, 1/4], [1/4, 1/2], [1/2, 3/4], [3/4, 1]\}.$

Note that the weighted jumps $h_E \| [\lambda_{d,h} \cdot n_E] \|_{L^2(E)}^2$ of $\lambda_{d,h}$ across edges $E \in \mathcal{E}_{\Omega}$ 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 $z \in \mathcal{K}$, $T_1, T_2 \in \mathcal{T}$ 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 $(A_0^{adaptive})$ and $(A_{1/2}^{adaptive})$ in Example 7.1. The obtained error estimators η_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 \mathcal{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 η_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.

triangulation	1	2	3	4	5	6	7	8	9
# elements	4	8	16	32	64	128	256	512	1,024
# atoms	179	433	1,122	3,034	8,385	23,443	65,921	185,908	$525,\!057$
# active atoms	8.5	12.0	7.9	10.7	9.6	17.4	46.3	64.5	127.1

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 $(A^{active \ set})$. 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.

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

TABLE 2. Possible and active atoms per element on adapted meshes.

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 $t \in [0, 1]$,

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

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

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

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

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

the pair (u, ν) is a solution for (EP). The coarsest triangulation \mathcal{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 $(A^{active set})$ and to calculate error indicators η_R , $\eta_{Z,R}$, and $\eta_{Z,E}$ in order to refine the mesh and to estimate the error in Algorithm $(A^{adaptive})$.

Figure 4 shows the adaptively generated mesh \mathcal{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 τ 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 ν from



FIGURE 4. Adaptively generated mesh and Young measure restricted to three different elements in Example 7.2.

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.

dof	9	35	70	162	255	492
CPU-time [s]	11.7	269.0	830.7	5,792.7	9,797.4	$24,\!317.5$

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 \mathcal{T}_k , k = 1, ..., 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

[B] Ball, J.M. (1989), A version of the fundamental theorem for Young measures. Partial differential equations and continuum models of phase transitions. EDS. M RASCLE, D. SERRE, M. SLEMROD: Lecture Notes in Physics 344, Springer, Berlin, 207-215.

- [BHJPS] Ball, J.M., Holmes, P.J., James, R.D., Pego, R.L., Swart, P.J. (1991), On the dynamics of fine structure. J. Nonlinear Sci. 1, 17-70.
- [BCD] Bartels, S., Carstensen, C., Dolzmann, G. (2001), Inhomogeneous Dirichlet conditions in a priori and a posteriori finite element error analysis. Berichtsreihe des Mathematischen Seminars Kiel, Technical report 02-1, Christian–Albrechts–Universität zu Kiel, Kiel.
- [Ca] Carstensen, C. (1999), Quasi interpolation and a posteriori error analysis in finite element method. M^2AN 33, No. 6, 1187-1202.
- [CB] Carstensen, C., Bartels S. (2002), Each Averaging Technique Yields Reliable A Posteriori Error Control in FEM on Unstructured Grids. Part I: Low Order Conforming, Nonconforming, and Mixed FEM. Math. Comp. 71, No. 239, 945–969.
- [CaM] Carstensen, C., Müller, S. (2002), Local stress regularity in scalar non-convex variational problems. SIAM J. Math. Anal. 34, No. 2, 495–509.
- [CP] Carstensen, C., Plecháč, P. (1997), Numerical solution of the scalar double-well problem allowing microstructure. *Math. Comp.* 66, No. 219, 997-1026.
- [CR] Carstensen, C., Roubiček, T. (2000), Numerical Approximation of Young Measures in Non-convex Variational Problems. Numer. Math. 84, No. 3, 395-414.
- [ChM] Chipot, M., Müller, S. (1997), Sharp energy estimates for finite element approximations of non-convex problems. Variations of domain and free-boundary problems in solid mechanics., Solid Mech. Appl. 66, Kluwer, Dordrecht, 317–325.
- [Ci] Ciarlet, P.G. (1978), The finite element method for elliptic problems. North-Holland, Amsterdam.
- [Cl] Clarke, F.H. (1990), Optimization and Nonsmooth Analysis. SIAM, Philadelphia.
- [CL] Collins, C., Luskin, M. (1991), Optimal order error estimates for the finite element approximation of the solution of a non-convex variational problem. *Math. Comp.* 57, No. 196, 621-637.
- [E] Ericksen, J.L. (1986), Constitutive theory for some constrained elastic crystals. J. Solids and Structures 22, 951-964.
- [F] Friesecke, G. (1994), A necessary and sufficient condition for non-attainment and formation of microstructure almost everywhere in scalar variational problems. *Proc. R. Soc. Edin.* **124A**, 437-471.
- [G] Gondzio, J. (1995), HOPDM (version 2.12) A fast LP solver based on a primal-dual interior point method. European Journal of Operational Research 85, 221-225.
- [KiP] Kinderlehrer, D., Pedregal, P. (1991), Weak convergence of integrands and the Young measure representation. SIAM J. Math. Anal., 23, No. 1, 1-19.
- [Kr] Kružik, M. (1998), Numerical approach to double well problems. SIAM J. Numer. Anal. 35, No. 5, 1833-1849.
- [KrP] Kružík, M., Prohl, A. (2001), Young measure approximation in micromagnetism. Numer. Math. 90, No. 2, 291–307.
- [Ku] Kuratowski, K. (1966, 1968), Topology I, II. Acad. Press, New York, and PWN, Warszawa.
- [L] Luskin, M. (1996), On the computation of crystalline microstructure. Acta Numerica 5, 191-257.
- [MRS] Mataché, A.-M., Roubíček, T., Schwab, C. (2001), Higher-order convex approximations of Young measures for optimal control. Research Report Forschungsinstitut für Mathematik ETH Zürich.
- [M] Müller, S., (1999), Variational models for microstructure and phase transitions. Lect. Notes Math. 1713, Springer, Berlin, 85-210.
- [NW] Nicolaides, R. A., Walkington, N. J. (1993), Computation of microstructure utilizing Young measure representations, Transactions of the Tenth Army Conference on Applied Mathematics and Computing (West Point, NY, 1992), ARO Rep., 93-1, 57-68.
- [P1] Pedregal, P. (1995), Numerical approximation of parametrized measures. Numer. Funct. Anal. Optim. 16, No. 7-8, 1049–1066.
- [P2] Pedregal, P. (1997), Parametrized measures and variational principles. Birkhäuser, Basel.
- [R] Roubíček, T. (1997), Relaxation in optimization theory and variational calculus. De Gruyter, New York.
- [RK] Roubíček, T., Kružik, M. (2001), Adaptive approximation algorithm for relaxed optimization problems. Fast solution of discretized optimization problems, Birkhäuser, Basel, 242–254,

- [V] Verfürth, R. (1996), A review of a posteriori error estimation and adaptive mesh-refinement techniques. Wiley-Teubner, Stuttgart.
- [Y] Young, L.C. (1937), Generalized curves and the existence of an attained absolute minimum in the calculus of variations. Comptes Rendues de la Société des Sciences et des Lettres de Varsovie, classe III 30, 212-234.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MARYLAND, COLLEGE PARK, MD 20742, USA. *E-mail address:* sba@math.umd.edu