CONVERGENCE OF ADAPTIVE FEM FOR A CLASS OF DEGENERATE CONVEX MINIMIZATION PROBLEMS

CARSTEN CARSTENSEN*

Abstract. A class of degenerate convex minimization problems allows for some adaptive finite element method (AFEM) to compute strongly converging stress approximations. The algorithm AFEM consists of successive loops of the form

$$\texttt{SOLVE} \to \texttt{ESTIMATE} \to \texttt{MARK} \to \texttt{REFINE}$$

and employs the bulk criterion. The convergence in $L^{p'}(\Omega;\mathbb{R}^{m\times n})$ relies on new sharp strict convexity estimates of degenerate convex minimization problems with

$$\mathcal{J}(v) := \int_\Omega W(Dv)\,dx - \int_\Omega f v\,dx \quad \text{for } v \in V := W_0^{1,p}(\Omega;\mathbb{R}^m).$$

The class of minimization problems includes strong convex problems and allows applications in an optimal design task, Hencky elastoplasticity, or relaxation of 2-well problems allowing for microstructures.

1. Class of Convex Minimization Problems

This section specifies a class of C^1 energy densities $W\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ characterized by (H1)-(H2) for some constants $1 < p < \infty$, $1 \le r < \infty$, and $0 \le s < \infty$ with

$$\max\{(1+s/r)/(1-1/r), 2n/(n+2)\} \le p,$$

through the *two-sided growth condition*

$$\text{(H1)} \qquad |F|^p - 1 \lesssim W(F) \lesssim 1 + |F|^p \text{ for all } F \in \mathbb{R}^{m\times n}$$

1991 *Mathematics Subject Classification.* 65 N 12, 65 N 12, 65 N 30.

Key words and phrases. adaptive finite element methods, convex minimization, degenerate convex problems, convergence, error reduction, energy reduction, adaptive mesh-refining, a posteriori error estimates, discrete local efficiency, discrete residual control

* Supported by the DFG Research Center MATHEON “Mathematics for key technologies” in Berlin.

and the *convexity control*

$$\text{(H2)}\quad \begin{aligned}&(1+|A|^s+|B|^s)^{-1}|DW(A)-DW(B)|^r\\ &\lesssim W(B)-W(A)-DW(A):(B-A) \text{ for all } A,B\in\mathbb{R}^{m\times n}.\end{aligned}$$

Here and throughout "$\cdot$" denotes the scalar product in $\mathbb{R}^m$, "$:$" denotes the scalar product in $\mathbb{R}^{m\times n}$, and the expression "$\lesssim$" abbreviates an inequality up to some multiplicative generic constant, i.e., $A\lesssim B$ means $A\le cB$ with some generic constant $c>0$, which is independent of the arguments A,B,F in (H1)-(H2) (but may depend on W and on the aspect ratio of finite element triangulations).

Finally, $t:=1+s/p$ and the Hölder conjugate p' of p satisfy

$$1<p'\le r/t<\infty,\quad \text{and}\quad 1/p+1/p'=1$$

and where r/t and $r/(r-t)$ are conjugate exponents, i.e., $t/r+(r-t)/r=1$.

Section 3 exposes a list of examples with (H1)-(H2). The two-sided growth control (H1) is standard in the form of

$$|F|^p\lesssim W(F)+1\quad\text{and}\quad W(F)\lesssim 1+|F|^p.$$

By adding a constant to $W(F)$, it could be replaced even by

$$|F|^p\lesssim W(F)\lesssim 1+|F|^p.$$

The convexity control (H2) implies the monotonicity condition

$$\text{(H3)}\quad \begin{aligned}&(1+|A|^s+|B|^s)^{-1}\,|DW(A)-DW(B)|\\ &\lesssim \big(DW(A)-DW(B)\big):(A-B)\quad\text{for all } A,B\in\mathbb{R}^{m\times n}\end{aligned}$$

from [10, 11]. Under some conditions, (H2) is in fact equivalent to (H3) [15, 16].

Given such energy density $W:\mathbb{R}^{m\times n}\to\mathbb{R}$ and a bounded Lipschitz domain $\Omega\subset\mathbb{R}^n$, $n=2,3$, and some right-hand side $f\in L^{p'}(\Omega;\mathbb{R}^m)$, define $\mathcal{J}:V\to\mathbb{R}$ by

$$\text{(1.1)}\quad \mathcal{J}(v):=\int_\Omega W(Dv)\,dx-\int_\Omega f\cdot v\,dx\quad\text{for } v\in V:=W_0^{1,p}(\Omega;\mathbb{R}^m).$$

Throughout this paper, $Dv(x)$ denotes the $m\times n$ functional matrix of V at x and we adapt standard notation on Lebesgue and Sobolev spaces, e.g., $W_0^{1,p}(\Omega)$ denotes the subset of functions in $W^{1,p}(\Omega)$ with trace zero on the boundary $\partial\Omega$ of Ω.

The minimization problem reads: *Seek minimizers in $\mathcal{J}$ in V*, written

$$\text{(1.2)}\qquad u\in\arg\min_{v\in V}\mathcal{J}(v).$$

The existence of minimizers u or u_ℓ of (1.1) in V or some closed subspace V_ℓ of V is guaranteed under (H1)-(H2) while, in general, their uniqueness fails. However, the respective exact and discrete stress

$$\sigma := DW(Du) \quad \text{and} \quad \sigma_\ell = DW(Du_\ell) \in L^{r/t}(\Omega;\mathbb{R}^{m\times n})$$

is unique [11], i.e., σ and σ_ℓ do not depend on the choice of u and u_ℓ amongst the set of exact and discrete minimizers. The smoothness of $\sigma \in W^{1,p}_{loc}(\Omega;\mathbb{R}^{m\times n})$ has been analysed in [10, 16], while the smoothness of u is open (recall that u may be non-unique). Therefore the a priori error estimate (valid for any choice of $u \in \operatorname{argmin} J$)

$$\|\sigma - \sigma_\ell\|_{L^q(\Omega;\mathbb{R}^{m\times n})} \lesssim \min_{v_\ell \in V_\ell} \|u - v_\ell\|_V,$$

although it may be regarded as quasi-optimal convergent, has its limitations. The a posteriori error estimates for $\|\sigma - \sigma_\ell\|_{L^q(\Omega;\mathbb{R}^{m\times n})}$ known from the literature even face some reliability-efficiency gap [9], cf. Section 2 and Remark 2.1 below. Surprisingly, this does not prevent the design of convergent adaptive mesh-refining algorithms.

2. AFEM

This section describes the adaptive mesh-refining strategy, proposed in this paper and states the main result.

2.1. **Outline.** Given an initial coarse mesh $\mathcal{T}_0$, an adaptive finite element method (AFEM) successively generates a sequence of meshes $\mathcal{T}_1, \mathcal{T}_2, \dots$ and associated discrete subspaces

$$V_0 \subsetneqq V_1 \subsetneqq \cdots \subsetneqq V_\ell \subsetneqq V_{\ell+1} \subsetneqq \cdots \subsetneqq V \tag{2.1}$$

with discrete problems (P_0), (P_1), (P_2), ... and discrete solutions u_0, u_1, $u_2, \dots$ and discrete stresses $\sigma_0, \sigma_1, \sigma_2, \dots$ steered by refinement rules and indicators. A typical loop from V_ℓ to $V_{\ell+1}$ (at the frozen level ℓ) consists of the steps

$$\texttt{SOLVE} \rightarrow \texttt{ESTIMATE} \rightarrow \texttt{MARK} \rightarrow \texttt{REFINE} \tag{2.2}$$

explained in the following Subsections.

2.2. **Input.** Input a shape-regular triangulation $\mathcal{T}_0$ of $\Omega \subset \mathbb{R}^n$ into closed triangles (if $n = 2$) or closed tetrahedra (if $n = 3$) with associated first-order finite element space V_0; suppose each element domain in $\mathcal{T}_0$ (and furthermore in $\mathcal{T}_1$, $\mathcal{T}_2$, ...) has at least one vertex in the interior of Ω, put level $\ell := 0$.

A triangulation $\mathcal{T}_\ell$ is regular if two distinct closed-element domains are either disjoint or their intersection is one common vertex, one common edge (or, if $n = 3$ possibly one common face). For simplicity, all triangulations in the paper will be regular. Those common faces are called sides $\mathcal{E}_\ell$, if $n = 3$. For $n = 2$, $\mathcal{E}_\ell$ are the interior edges.

2.3. `SOLVE`. Given the triangulation $\mathcal{T}_\ell$ with set of interior sides $\mathcal{E}_\ell$ and interior nodes $\mathcal{K}_\ell$, the piecewise affine space $\mathcal{P}_1(\mathcal{T}_\ell)$ reads

$$\begin{aligned}\mathcal{P}_1(\mathcal{T}_\ell;\mathbb{R}^m) &:= \big\{v \in L^\infty(\Omega;\mathbb{R}^m) : \forall T \in \mathcal{T}_\ell,\, v|_T \in \mathcal{P}_1(T;\mathbb{R}^m)\big\};\\ \mathcal{P}_1(T;\mathbb{R}^m) &:= \big\{v \in C^\infty(T;\mathbb{R}^m) : \exists A \in \mathbb{R}^{m\times n}\, \exists b \in \mathbb{R}^m\\ &\qquad \forall x \in T : v(x) = Ax + b\big\}.\end{aligned}$$

The discrete space $V_\ell := V \cap \mathcal{P}_1(\mathcal{T}_\ell;\mathbb{R}^m)$ is the first-order finite element space and allows for a nodal basis $(\varphi_z : z \in \mathcal{K}_\ell)$. Then the step `SOLVE` reads: Solve the nonlinear discrete problem

$$u_\ell \in \arg\min_{v_\ell \in V_\ell} \mathcal{J}(v_\ell) \quad \text{and set} \quad \sigma_\ell := DW(Du_\ell). \tag{2.3}$$

The $\mathbb{R}^{m\times n}$-valued stress σ_ℓ is piecewise constant with respect to $\mathcal{T}_\ell$.

2.4. `ESTIMATE`. Given any interior side $E \in \mathcal{E}_\ell$ with measure $|E|$, and normal unit vector ν_E, compute the jump

$$J_E := [\sigma_\ell]_E\, \nu_E \in \mathbb{R}^m$$

of the discrete normal stresses $\sigma_\ell \nu_E$ over E, where

$$[\sigma_\ell]_E(x) := \lim_{T_+ \ni a \to x} \sigma_\ell(a) - \lim_{T_- \ni b \to x} \sigma_\ell(b)$$

for all $x \in E = \partial T_+ \cap \partial T_-$, and by convention, ν_E is exterior to T_+. Then define

$$\eta_\ell := \Big(\sum_{E\in\mathcal{E}_\ell} \eta_E^{p'}\Big)^{1/p'} \quad \text{with} \quad \eta_E := h_E^{1/p'} |E|^{1/p'}\, |J_E| \quad \text{for } E \in \mathcal{E}_\ell. \tag{2.4}$$

It is essentially known from [9, 11] that η_ℓ is a reliable a posteriori error estimator in the sense that

$$\|\sigma - \sigma_\ell\|^r_{L^{r/t}(\Omega;\mathbb{R}^{m\times n})} \lesssim \eta_\ell + \mathrm{osc}_\ell, \tag{2.5}$$

cf. Lemma 4.2 below. Here and throughout, osc_ℓ denotes data oscillations. Given any connected open nonvoid $\omega \subset \Omega$, let

$$\mathrm{osc}(f,\omega)^{p'} := \mathrm{diam}(\omega)^{p'} \|f - f_\omega\|^{p'}_{L^{p'}(\omega)} \text{ with } f_\omega := |\omega|^{-1} \int_\omega f\,dx, \tag{2.6}$$

the integral mean of f over ω. For each node z in the triangulation $\mathcal{T}_\ell$ with nodal basis function $\varphi_z \in V_\ell$, let $\omega_z := \{x \in \Omega : \varphi(x) > 0\}$ denote the patch of z. Then, recall $\mathcal{K}_\ell$ denotes the set of all interior nodes,

$$\operatorname{osc}_\ell^{p'} := \sum_{z\in\mathcal{K}_\ell} \operatorname{osc}(f,\omega_z)^{p'}. \tag{2.7}$$

Since osc_ℓ depends on the given data and explicitly on $\mathcal{T}_\ell$, it can easily be made arbitrarily small by additional refinement steps. This data oscillation control allows for $\lim_{\ell\to\infty} \operatorname{osc}_\ell = 0$; cf. [17, 22] for algorithmic details.

Remark 2.1. The upper bound in (2.5) is not sharp, the estimator η_ℓ is not efficient, because of $r > 1$. This is called reliability-efficiency gap [9].

2.5. `MARK`. Select a subset $\mathcal{M}_\ell$ of $\mathcal{E}_\ell$ in the current triangulation $\mathcal{T}_\ell$ with

$$\eta_\ell^{p'} \lesssim \sum_{E\in\mathcal{M}_\ell} \eta_E^{p'}. \tag{2.8}$$

Given a parameter $0 < \Theta < 1$ the selection condition (2.8) results from choosing sufficiently many sides E with bigger η_E in $\mathcal{M}_\ell$ such that the *bulk criterion* [13, 17, 18, 22] holds:

$$\Theta\, \eta_\ell^{p'} \leq \sum_{E\in\mathcal{M}_\ell} \eta_E^{p'}.$$

This is easily arranged with some greedy algorithm.

2.6. `REFINE`. Refine the triangulation $\mathcal{T}_\ell$ and design a refined shape-regular triangulation $\mathcal{T}_{\ell+1}$ such that each interior side $E = \partial T_+ \cap \partial T_- \in \mathcal{M}_\ell$ is refined in $\mathcal{T}_{\ell+1}$, for $T_+, T_- \in \mathcal{T}_\ell$ and $T_+ \cup T_-$ includes at least one new node on E and at least one new node in the interior of either T_+

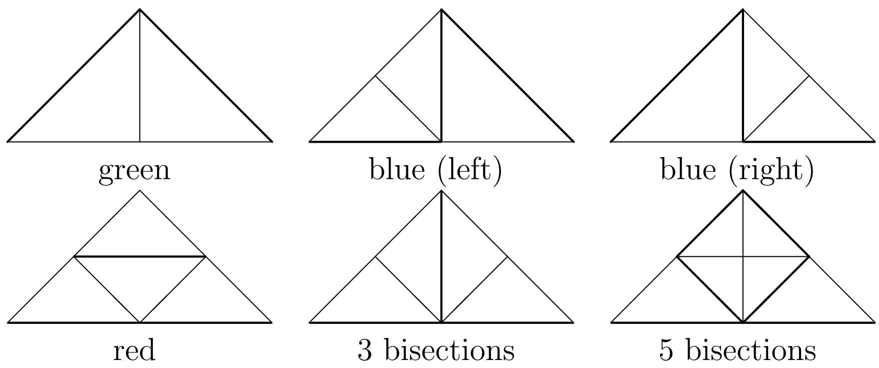

FIGURE 2.1. Possible refinements of a triangle in `REFINE` of AFEM. The 5 bisections allow for an interior node property.

or T_-. For $n = 2$ the inner node property is easily depicted with 5 bisections as in Figure 2.1. More details on the shape-regular refinement strategies can be found in [6].

2.7. **Output.** The AFEM computes a sequence of discrete stresses $\sigma_0, \sigma_1, \sigma_2, \dots$ in $L^{p'}(\Omega; \mathbb{R}^{m\times n})$ as approximations to $\sigma := DW(Du)$. The main result of this paper is the strong convergence of the stresses.

Theorem 2.1 (Convergence Theorem)**.** *Suppose* (H1)-(H2) *and*

$$\lim_{\ell\to\infty} \mathrm{osc}_\ell = 0.$$

Then the sequence of stress fields $\sigma_0, \sigma_1, \sigma_2, \dots$ converges strongly towards the exact stress field σ in $L^{r/t}(\Omega; \mathbb{R}^{m\times n})$.

The technical proof is postponed to Section 4, after the motivating list of examples in Section 3.

3. Examples and Applications

This section briefly summarizes a few applications with explicit proofs of (H1)-(H2) and hence with a convergent AFEM.

3.1. **Uniformly Convex Minimization.** Uniformly convex $\mathcal{C}^1$ function $W : \mathbb{R}^{m\times n} \to \mathbb{R}$ with globally Lipschitz continuous derivative DW, i.e., for all $A, B \in \mathbb{R}^{m\times n}$ there holds

$$\begin{aligned} |A-B|^2 &\lesssim DW(A) : (A-B) - W(A) + W(B) \\ |DW(A) - DW(B)| &\lesssim |A-B|. \end{aligned}$$

This implies (H1)-(H2) with $p = 2 = r$ and $s = 0$ and, thus, the class (i) is included in class (ii). Simple examples are $W(F) = \varphi(|\mathrm{sym}\, F|)|F|^2$ for proper $\mathcal{C}^2$ functions φ (cf., e.g., [23, Sections 62.3, 62.8-9] and [15, Exercise 1.7 on page 21]).

3.2. **Nonlinear Laplacian.** The p-Laplacian satisfies (H1)-(H2) for any $2 \le p < \infty$ and $r = 2$, $s = p - 2$.

Lemma 3.1. *Given $1 \le p < \infty$ define the function $W : \mathbb{R}^{m\times n} \to \mathbb{R}$ by $W(A) := |A|^p/p$. Then there exist a constant $c_1 = c(p)$ such that for all $A, B \in \mathbb{R}^{m\times n}$ there holds*

$$\begin{aligned} |DW(A) - DW(B)|^2 \le c_1(|A|^{p-2} + |B|^{p-2}) \\ \times \left(W(B) - W(A) - DW(A; B - A)\right). \end{aligned}$$

Proof. Given $A, B \in \mathbb{R}^{m\times n}$ with $A \neq B$ set $a := |A|$ and $b := |B|$. A quick check verifies that the assertion holds for either $a = 0$ or $b = 0$ with $c_1 = \max\{p, q\}$. It is therefore assumed that $ab > 0$ in the sequel and $c := A : B/(ab)$. Then $0 < t := b/a < \infty$. The left- and right-hand side of the assertion vanish for $a = b$ and $c = +1$. This situation is therefore excluded in the sequel. Then,

$$\begin{aligned} W(B) - W(A) - DW(A; B - A) &= b^p/p - a^p/p - a^{p-1}(cb - a) \\ &= b^p/p + a^p/q - a^{p-1}bc \end{aligned}$$

is strictly positive (non-negativity immediately follows from Young's inequality and $-1 \leq c \leq 1$). Since

$$|DW(A) - DW(B)|^2 = a^{2(p-1)} + b^{2(p-1)} - 2ca^{p-1}b^{p-1}.$$

The quotient of the left- and the right-hand side of the assertion reads

$$\begin{aligned} \frac{a^{2(p-1)} + b^{2(p-1)} - 2ca^{p-1}b^{p-1}}{(a^{p-2} + b^{p-2})(b^p/p + a^p/q - a^{p-1}bc)} &= \frac{1 + t^{2(p-1)} - 2ct^{p-1}}{(1 + t^{p-2})(t^p/p + 1/q - ct)} \\ &=: f(t, c). \end{aligned}$$

A direct calculation verifies that $\partial f/\partial c$ as a function of c has one sign (which depends on t and p) and hence is monotone increasing or decreasing. Therefore

$$\max_{-1\leq c\leq 1} f(t, c) = \max\{f(t, 1), f(t, -1)\}$$

and the assertion reads $f(t, 1) \leq c_1$ and $f(t, -1) \leq c_1$ for all $0 < t < \infty$. The case $c = +1$ is the crucial one because $t^p/p + 1/q - t$ vanishes for $t = 1$. Hospital's rule yields $f(1, 1) = 0$. Since $f(0, 1) = q$ and $\lim_{t\to\infty} f(t, 1) = p$, one deduces from continuity of $f(t, 1)$ in t that

$$\sup_{0<t<\infty} f(t, 1) =: c_1 < \infty.$$

The analysis for $c = -1$ is simpler and hence omitted. □

3.3. **Optimal Design Problem.** Let $0 < t_1 < t_2$ and $0 < \mu_2 < \mu_1$ be positive real numbers with $t_1\mu_1 = t_2\mu_2$ and consider a convex C^1 function $\psi : [0, \infty) \to \mathbb{R}$ with $\psi(0) = 0$ and

$$\psi'(t) := \begin{cases} \mu_1 t & \text{for } 0 \leq t \leq t_1, \\ t_1\mu_1 = t_2\mu_2 & \text{for } t_1 \leq t \leq t_2, \\ \mu_2 t & \text{for } t_2 \leq t. \end{cases}$$

The energy density $W(A) := \psi(|A|)$, $A \in \mathbb{R}^n$, results from a relaxation process [14]. It satisfies (H1)-(H2) with $p = r = 2$ and $s = 0$. Details can be found in [2].

3.4. **Scalar 2-Well Problem.** The scalar convexified 2-well energy density W results from a relaxation in nonconvex minimization problems allowing for microstructures [11]. It satisfies (H1)-(H2) with $p = 4$ and $r = 2 = s$.

Proposition 3.2. *Given distinct F_1 and F_2 in $\mathbb{R}^n$ set $A := (F_2 - F_1)/2 \neq 0$ and $B := (F_1 + F_2)/2$ where $(\cdot)_+ := \max\{0, \cdot\}$ and $(\cdot)_+^2 := \max\{0, \cdot\}^2$. For any $F \in \mathbb{R}^n$ let*

$$W(F) := (|F - B|^2 - |A|^2)_+^2 + 4\left(|A|^2|F - B|^2 - (A \cdot (F - B))^2\right).$$

Then for any $F, G \in \mathbb{R}^n$ with $\xi := (|F - B|^2 - |A|^2)_+$ and $\eta := (|G - B|^2 - |A|^2)_+$ there holds

$$|DW(G) - DW(F)|^2 \\ \leq 32(|A|^2 + \xi + \eta)(W(G) - W(F) - DW(F) \cdot (G - F)).$$

The proof of Proposition 3.2 is based on two lemmas.

Lemma 3.3. *Given $A, B \in \mathbb{R}^n$ let $W(F) := (|F - B|^2 - |A|^2)_+^2$. For any F and G in $\mathbb{R}^n$ let*

$$\xi := (|F - B|^2 - |A|^2)_+ \quad \textit{and} \quad \eta := (|G - B|^2 - |A|^2)_+.$$

Then there holds

$$|DW(F) - DW(G)|^2 \\ \leq 32(|A|^2 + \xi + \eta)(W(G) - W(F) - DW(F) \cdot (G - F)).$$

Proof. Let $U := F - B$, $V := G - B$, $a := |A|$ and notice that $DW(F) = 4\xi U$ and $DW(G) = 4\eta V$. In the first case suppose that both, $\xi = |U|^2 - a^2$ and $\eta = |V|^2 - a^2$, are positive. Utilizing

$$DW(F) - DW(G) = 4(\xi U - \eta V) = 4\xi(U - V) + 4(\xi - \eta)V$$

one obtains

$$1/32\,|DW(F) - DW(G)|^2 \leq \xi^2|U - V|^2 + (\xi - \eta)^2|V|^2.$$

Since $|V|^2 = \eta + a^2$ this proves

$$1/32\,|DW(F) - DW(G)|^2 \leq (a^2 + \xi + \eta)(\xi|U - V|^2 + (\xi - \eta)^2). \tag{3.1}$$

On the other hand, the preceeding situation allows the direct calculation of

$$\begin{aligned} &W(G) - W(F) - DW(F) \cdot (F - G) \\ &= \eta^2 - \xi^2 + 4\xi U \cdot (U - V) \\ &= \eta^2 - \xi^2 + 2\xi(|U|^2 - |V|^2) + 2\xi|U - V|^2 \\ &= 2\xi|U - V|^2 + (\xi - \eta)^2. \end{aligned}$$

The combination with (3.1) shows the assertion in the present first case of positive ξ and η. For $\xi = 0 < \eta = |V|^2 - a^2$ the assertion reads

$$16\eta^2|V|^2 \le 32(a^2+\eta)\,\eta^2$$

which follows immediately from $|V|^2 \le (a^2+\eta)$. In the remaining case $\eta = a < \xi = |U|^2 - a^2$, whence $|V| \le a < |U|$, the assertion reads

$$16\xi^2|U|^2 \le 32(a^2+\xi)(4\xi U\cdot(U-V)-\xi^2).$$

This is equivalent to

$$\xi^2|U|^2 \le 2(a^2+\xi)(\xi^2+2\xi(a^2-|V|^2)+2\xi|U-V|^2)$$

and hence follows from $|U|^2 = a^2+\xi$ and $0 \le a^2-|V|^2$. $\square$

Lemma 3.4. *Let S be a symmetric and positive semidefinite real $n\times n$ matrix with spectral radius $\varrho(S)$ and pseudo inverse S^+ and induced seminorm $|\cdot|_{S^+}$, i.e.,*

$$|F|_{S^+} := (F\cdot S^+F)^{1/2} \quad \text{for all } F\in\mathbb{R}^n.$$

Then the function $W:\mathbb{R}^n\to\mathbb{R}$ defined by

$$W(F) := {}^1\!/_2\, F\cdot SF \quad \text{for } F\in\mathbb{R}^n$$

satisfies

$$\begin{aligned}\varrho(S)^{-1}|DW(F)-DW(G)|^2 &\le |DW(F)-DW(G)|^2_{S^+}\\ &= (F-G)\cdot S(F-G)\\ &= 2(W(G)-W(F)-(SF)\cdot(G-F)).\end{aligned}$$

Proof. Since S is symmetric, $S = SS^+S$, and so $DW(F) = SF$ satisfies

$$|S(F-G)|^2 \le \varrho(S)|S^{1/2}(F-G)|^2 = \varrho(S)|S(F-G)|^2_{S^+}.$$

The remaining identity results from

$${}^1\!/_2(F-G)\cdot S(F-G) = W(G)-W(F)+F\cdot S(F-G). \quad \square$$

Proof of Proposition 3.2. Notice that $W(F)$ is the sum of the two energy densities of the aforegoing lemmas. Indeed, let $A^0 := A/|A|$ and define the symmetric and positive semidefinite matrix $S := 1-A^0\otimes A^0$. Then

$$4\left(|A|^2|F-B|^2-(A\cdot(F-B))^2\right) = 4|A|^2|F-B|^2_S.$$

Observe the upper bound of S

$$|DW(G)-DW(F)|^2 \le 32|\xi U-\eta V|^2 + 32|A|^4|U-V|^2_S$$

is estimated in Lemma 3.3 and Lemma 3.4, respectively. This concludes the proof. □

3.5. **Vectorial 2-Well Problem.** Given two distinct wells E_1 and E_2 in $\mathbb{R}^{n\times n}_{\rm sym}$ with minimal energies W_1^0 and W_2^0 in $\mathbb{R}$, we consider the quadratic elastic energies

$$W_j(E) := {}^1\!/_2(E-E_j) : \mathbb{C}(E-E_j) + W_j^0 \quad \text{for all } E \in \mathbb{R}^{n\times n}_{\rm sym}.$$

Energy minimization leads to an optimal choice of the configuration of the two phases, and so the strain energy density $\tilde{W}$ is modelled by the minimum

$$\tilde{W}(E) = \min\{W_1(E), W_2(E)\} \quad \text{for all } E \in \mathbb{R}^{n\times n}_{\rm sym}.$$

The two wells (transformation strains) are said to be *compatible* if

$$E_1 = E_2 + {}^1\!/_2(a\otimes b + b\otimes a) \quad \text{for some } a,b \in \mathbb{R}^n. \tag{3.2}$$

Then the constant $\gamma = {}^1\!/_2|E_2 - E_1|^2_{\mathbb{C}}$ and the quasiconvexification W of $\tilde{W} = \{W_1, W_2\}$ [14] is given by

$$W(E) = \begin{cases} W_2(E) & \text{if } W_2(E)+\gamma \le W_1(E),\\ \frac{1}{2}(W_2(E)+W_1(E)) - \frac{1}{4\gamma}(W_2(E)-W_1(E))^2 - \frac{\gamma}{4} & \\ & \text{if } |W_2(E)-W_1(E)| \le \gamma,\\ W_1(E) & \text{if } W_1(E)+\gamma \le W_2(E). \end{cases}$$

The convex W satisfies (H1)-(H2) with $p = 2 = r$ and $s = 0$.

Proposition 3.5. *In the compatible case (3.2) there holds, for all* $A, B \in \mathbb{R}^{n\times n}_{\rm sym}$,

$${}^1\!/_2\,|DW(A) - DW(B)|^2_{\mathbb{C}^{-1}} \le W(B) - W(A) - DW(A) : (B-A).$$

Proof. A translation of the argument in W allows us to assume, without loss of generality, that $E_1 + E_2 = 0$. For $E \in \mathbb{R}^{n\times n}_{\rm sym}$, let

$$\begin{aligned} \varphi(E) &:= \gamma^{-1}(W_2(E) - W_1(E)),\\ \psi(E) &:= \max\{-1, \min\{1, \varphi(E)\}\}. \end{aligned}$$

As in [12] one deduces, for $E \in \mathbb{R}^{n\times n}_{\rm sym}$ and $\gamma\varphi(E) = 2(\mathbb{C}E_1) : E + W_2^0 - W_1^0$,

$$DW(E) = \mathbb{C}E - \psi(E)\mathbb{C}E_1$$

and observes that $\psi(E) = \varphi(E)$ for $E \in \mathbb{R}^{n\times n}_{\rm sym}$ with $-1 \le \varphi(E) \le 1$. The proof of the proposition starts with the discussion of

$$\gamma/2\,(\psi(B) - \psi(A))(\psi(A) - \varphi(A)) \ge 0. \tag{3.3}$$

In fact, $\psi(A) \neq \varphi(A)$ implies either $\psi(A) = 1 < \varphi(A)$ [notice $\psi(B) - 1 \leq 0$] or $\psi(A) = -1 > \varphi(A)$ [notice $\psi(B) + 1 \geq 0$] and in each case (3.3) follows. Algebraic manipulations will show in the sequel that (3.3) is equivalent to the assertion. Abbreviate $\sigma := DW(A)$ and $\tau := DW(B)$ to compute the left-hand side of the assertion, namely

$$^1\!/_2\, |\sigma - \tau|^2_{\mathbb{C}^{-1}} = 1/2\,(\tau - \sigma) : \mathbb{C}^{-1}(\tau + \sigma) + (\sigma - \tau) : \mathbb{C}^{-1}\sigma.$$

With $\mathbb{C}^{-1}(\sigma - \tau) = A - B - \psi(A)E_1 + \psi(B)E_1$, this reads

$$\begin{aligned}\sigma : (A - B) - {}^1\!/_2\, |\sigma - \tau|^2_{\mathbb{C}^{-1}}& \\ &= (\psi(A) - \psi(B))E_1 : \sigma - {}^1\!/_2\, |\tau|^2_{C^{-1}} + {}^1\!/_2\, |\sigma|^2_{\mathbb{C}^{-1}}.\end{aligned}$$

The definition of σ and τ and $\gamma/2 = |E_1|^2_{\mathbb{C}}$ show

$$\begin{aligned}^1\!/_2\, |\sigma|^2_{\mathbb{C}^{-1}} - {}^1\!/_2\, |\tau|^2_{\mathbb{C}^{-1}} = {}^1\!/_2\, |A|^2_{\mathbb{C}} - {}^1\!/_2\, |B|^2_{\mathbb{C}} + \gamma/4\,(\psi(A)^2 - \psi(B)^2)& \\ - \psi(A)A : \mathbb{C}E_1 + \psi(B)B : \mathbb{C}E_1&.\end{aligned}$$

It is a lengthy but direct verification that $W(E)$, $E \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$, can be written as

$$W(E) = {}^1\!/_2\, E : \mathbb{C}E + {}^1\!/_2(W^0_1 + W^0_2) + \gamma/4\,\psi(E)(\psi(E) - 2\varphi(E)).$$

The combination of the preceeding three identities [the last applied to $E = A$ and $E = B$] shows

$$\begin{aligned}&W(B) - W(A) + \sigma : (A - B) - {}^1\!/_2\, |\sigma - \tau|^2_{\mathbb{C}^{-1}}\\ &= (\psi(A) - \psi(B))(E_1 : \mathbb{C}A - \psi(A)\gamma/2)\\ &\quad - \psi(A)A : \mathbb{C}E_1 + \psi(B)B : \mathbb{C}E_1\\ &\quad + \gamma/2\,\varphi(A)\psi(A) - \gamma/2\,\varphi(B)\psi(B)\\ &= -\gamma/2\,\psi(A)^2 + \gamma/2\,\psi(A)\psi(B) - \psi(B)E_1 : \mathbb{C}(A - B)\\ &\quad + \gamma/2\,\varphi(A)\psi(A) - \gamma/2\,\varphi(B)\psi(B).\end{aligned}$$

Since $E_1 : \mathbb{C}(A - B) = \gamma/2(\varphi(A) - \varphi(B))$ shows that the preceeding expression equals the left-hand side of (3.3). □

Remark 3.1. The immediate corollary (H3) of Proposition 3.5 is known from [10, 12] and fundamental for error analysis and regularity.

3.6. **Hencky elastoplasticity with hardening.** One time step within an elastoplastic evolution problem leads to Hencky's model. For various hardening laws and von-Mises yield conditions, an elimination of internal variables [1] leads to the energy function

$$W(E) := \frac{1}{2}\,E : \mathbb{C}E - \frac{1}{4\mu}\,\max\{0, |\operatorname{dev}\mathbb{C}E| - \sigma_y\}^2/(1 + \eta) \tag{3.4}$$

for $E \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$. Here we adopt notation of the previous section and $\mathbb{C}$ is the fourth-order elasticity tensor, $\sigma_y > 0$ is the yield stress, and $\eta > 0$ is the modulus of hardening. The model of perfect plasticity corresponds to $\eta = 0$ [21]. For $\eta > 0$ there holds (H1)-(H2) for $p = 2 = r$ and $s = 0$.

Proposition 3.6. *For all $A, B \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$ there holds*

$$1/2\,|DW(A) - DW(B)|^2_{\mathbb{C}^{-1}} \le W(B) - W(A) - DW(A) : (B - A).$$

Proof. Set $\psi(x) := 1 - \max\{0, 1 - \sigma_y/(2\mu x)\}/(1 + \eta)$ to define the continuous and monotone decreasing function $\psi : [0, \infty) \to (\eta/(1 + \eta), 1]$ which satisfies

$$DW(E) = (\lambda + 2\mu/n)\,\mathrm{tr}(E)\,\mathbf{1} + 2\mu\psi(|\operatorname{dev} E|)\operatorname{dev} E \quad \text{for all } E \in \mathbb{R}^{n\times n}_{\mathrm{sym}}.$$

Given $A, B \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$, the following abbreviations will be used throughout the remaining part of the proof:

$$\begin{aligned} \sigma &:= DW(A), & a &:= |\operatorname{dev} A|, & \alpha &:= \psi(a),\\ \tau &:= DW(B), & b &:= |\operatorname{dev} B|, & \beta &:= \psi(b). \end{aligned}$$

Then the assertion reads

$$\delta := W(B) - W(A) + \sigma : (A - B) - 1/2\,|\sigma - \tau|^2_{\mathbb{C}^{-1}} \ge 0.$$

In the first three steps one computes δ. The aforementioned formulae for $DW(A)$ and $DW(B)$ and elementary calculations with the third formula of Binomi yield in step one that

$$\begin{aligned} &\sigma : \mathbb{C}^{-1}(\sigma - \tau) - 1/2|\sigma - \tau|^2_{\mathbb{C}^{-1}}\\ &= 1/2\,|\sigma|^2_{\mathbb{C}^{-1}} - 1/2\,|\tau|^2_{\mathbb{C}^{-1}}\\ &= (\lambda/2 + \mu/n)(\mathrm{tr}(A)^2 - \mathrm{tr}(B)^2) + \mu(\alpha^2 a^2 - \beta^2 b^2). \end{aligned}$$

Step two employs the definition of ψ to rewrite the energy as

$$W(E) = 1/2\,|E|^2_{\mathbb{C}} - (1 + \eta)\mu\,(1 - \psi(|\operatorname{dev} E|))^2\,|\operatorname{dev} E|^2,$$

for all $E \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$. Step three employs the above formulae for σ and τ to estimate

$$\sigma : (A - B) - \sigma : \mathbb{C}^{-1}(\sigma - \tau) = 2\mu\,\alpha\operatorname{dev} A : ((1 - \alpha)\operatorname{dev} A - (1 - \beta)\operatorname{dev} B).$$

The Cauchy inequality, leads to

$$\sigma : (A - B) - \sigma : \mathbb{C}^{-1}(\sigma - \tau) \ge 2\mu\,\alpha(1 - \alpha)a^2 - 2\mu\,\alpha(1 - \beta)ab.$$

The left-hand sides considered in the first three steps add up to δ and so lead to a lower bound of δ. Elementary manipulations with this lower bound in step four of the proof yield the estimate

$$\begin{aligned}\delta/\mu &\geq \alpha^2a^2 - \beta^2b^2 + b^2 - a^2 + (1+\eta)(1-\alpha)^2a^2 - (1+\eta)(1-\beta)^2b^2\\ &\quad + 2\alpha(1-\alpha)a^2 - 2\alpha(1-\beta)ab\\ &= \eta(1-\alpha)^2a^2 - \eta(1-\beta)^2b^2 + 2(1-\beta)b(\beta b - \alpha a)\\ &= \eta\Big((1-\alpha)a - (1-\beta)b\Big)^2\\ &\quad + 2(1-\beta)b\Big((1+\eta)(\beta b - \alpha a) - \eta(b-a)\Big).\end{aligned}$$

Step five concerns the function $g(x) := x\psi(x)$ which satisfies $g'(x) = 1$ and $g'(x) = \eta/(1+\eta)$ for $2\mu x < \sigma_y$ and $\sigma_y < 2\mu x$, respectively. For $a \leq b$, this and the fundamental theorem of calculus show

$$\eta(b-a) \leq (1+\eta)\int_a^b g'(x)\,dx = (1+\eta)(\beta b - \alpha a). \tag{3.5}$$

This concludes the proof of $\delta \geq 0$ in this case. In the case $b < a$, the above lower bound of δ shows $\delta \geq 0$ if $\beta = 1$. Hence it remains to consider $b < a$ and $\beta < 1$ which implies $\sigma_y < 2\mu b$ and so $g'(x) = \eta/(1+\eta)$ for all $b < x < a$. This yields equality in (3.5) and so proves $\delta \geq 0$. ☐

Remark 3.2. Although (H2) holds for $\eta = 0$ as well, the linear growth condition yields a different functional analytical setting in $BD(\Omega)$ [21].

4. Proof of Convergence

This section provides a proof of Theorem 2.1 on the convergence of the stress fields in $L^{r/t}(\Omega;\mathbb{R}^{m\times n})$. Throughout this section, the focus is on the energy difference

$$\delta_\ell := \mathcal{J}(u_\ell) - \mathcal{J}(u) \geq 0.$$

Due to (2.1), the sequence $(\delta_\ell)_\ell$ is monotone decreasing, and hence convergent to some limit $\delta \geq 0$. It is essential to prove $\delta = 0$, which is not known in the beginning of the proof.

Lemma 4.1. *There holds*

$$\|\sigma_{\ell+1} - \sigma_\ell\|^r_{L^{r/t}(\Omega;\mathbb{R}^{m\times n})} \lesssim \delta_\ell - \delta_{\ell+1}.$$

Proof. The two-sided growth conditions in (H1) lead in [11] to the boundedness of discrete minimizers in $W^{1,p}$ and show

$$\int_\Omega (1 + |Du_\ell|^s + |Du_{\ell+1}|^s)^{p/s}\,dx \lesssim 1. \tag{4.1}$$

Since $\sigma_{\ell+1}$ satisfies the discrete Euler-Lagrange equations, there holds

$$\int_\Omega \sigma_{\ell+1} : D(u_\ell - u_{\ell+1})\, dx = \int_\Omega f \cdot (u_\ell - u_{\ell+1})\, dx.$$

Therefore,

$$\begin{aligned}\delta_\ell - \delta_{\ell+1} &= \int_\Omega \Big(W(Du_\ell) - W(Du_{\ell+1}) - f \cdot (u_\ell - u_{\ell+1}) \Big)\, dx \\ &= \int_\Omega \Big(W(Du_\ell) - W(Du_{\ell+1}) - \sigma_{\ell+1} : D(u_\ell - u_{\ell+1}) \Big)\, dx.\end{aligned}$$

An application of (H2) with $A = Du_{\ell+1}(x)$ and $B = Du_\ell(x)$ leads to an estimate for all x in Ω. The integral of those inequalities reads

$$\begin{aligned}&\int_\Omega (1 + |Du_\ell|^s + |Du_{\ell+1}|^s)^{-1}\, |\sigma_\ell - \sigma_{\ell+1}|^r\, dx \\ &\quad\lesssim \int_\Omega (W(Du_\ell) - W(Du_{\ell+1}) - \sigma_{\ell+1} : D(u_\ell - u_{\ell+1}))\, dx \\ &\quad= \delta_\ell - \delta_{\ell+1}.\end{aligned} \tag{4.2}$$

The Hölder inequality with t and $t' = 1 + p/s$, $1/t + 1/t' = 1$, plus (4.1) with $t'/t = p/s$ lead to

$$\begin{aligned}\|\sigma_{\ell+1} - \sigma_\ell\|^{r/t}_{L^{r/t}(\Omega;\mathbb{R}^{m\times n})} &= \int_\Omega (1 + |Du_\ell|^s + |Du_{\ell+1}|^s)^{-1/t}\, |\sigma_\ell - \sigma_{\ell+1}|^{r/t} \\ &\qquad \times (1 + |Du_\ell|^s + |Du_{\ell+1}|^s)^{1/t}\, dx \\ &\lesssim \Big(\int_\Omega (1 + |Du_\ell|^s + |Du_{\ell+1}|^s)^{-1}\, |\sigma_\ell - \sigma_{\ell+1}|^r\, dx \Big)^{1/t}.\end{aligned}$$

The combination of this estimate with (4.2) proves the lemma. □

Lemma 4.2. *There holds* (2.5)*, namely*

$$\|\sigma - \sigma_\ell\|^r_{L^{r/t}(\Omega;\mathbb{R}^{m\times n})} \lesssim \eta_\ell + \mathrm{osc}_\ell\, .$$

Proof. In slightly different notation, it is proven in [11] that

$$\|\sigma - \sigma_\ell\|^r_{L^{r/t}(\Omega;\mathbb{R}^{m\times n})} \lesssim \eta_\ell + \|h_{\mathcal{T}_\ell} f\|_{L^{p'}(\Omega)}. \tag{4.3}$$

It is known since [19, 20] that the volume contribution $\|h_{\mathcal{T}_\ell} f\|_{L^{p'}(\Omega)}$ can be controlled by $\eta_\ell + \mathrm{osc}_\ell$ and so (4.3) leads to the assertion; cf. [9] for one particular case. The main arguments are recalled here for convenient reading. A triangle inequality yields, for each free node z, that

$$\|f\|_{L^{p'}(\omega_z)} \le \|f - f_{\omega_z}\|_{L^{p'}(\omega_z)} + |f_{\omega_z}|\, |\omega_z|^{1/p'}. \tag{4.4}$$

The integral mean equals

$$f_{\omega_z}|\omega_z| \approx \int_\Omega \varphi_z f_{\omega_z}\,dx = \int_\Omega \varphi_z(f - f_{\omega_z})\,dx + \int_\Omega \varphi_z f\,dx. \tag{4.5}$$

The combination of (4.4)-(4.5) plus a Hölder inequality shows

$$\|f\|_{L^{p'}(\omega_z)} \lesssim \|f - f_{\omega_z}\|_{L^{p'}(\omega_z)} + |\omega_z|^{-1/p}\Big|\int_\Omega \varphi_z f\,dx\Big|. \tag{4.6}$$

On the other hand, the discrete Euler-Lagrange equations show for the j-th component f_j of f and the components $\sigma_{\ell,j} := (\sigma_{\ell,j_1},\dots,\sigma_{\ell,j_n})$ of σ_ℓ, that

$$\int_\Omega \varphi_z f_j\,dx = \int_\Omega \sigma_{\ell,j}\cdot\nabla\varphi_z\,dx = \sum_{E\in\mathcal{E}}\int_E \Big([\sigma_{\ell,j}]\cdot\nu_E\Big)\varphi_z\,ds \tag{4.7}$$

with an elementwise integration by parts. Let $\mathcal{E}(z) := \{E \in \mathcal{E} : z \in E\}$ denote the set of sides which contribute in (4.7). Then for all $j = 1, 2, \dots, m$ components in (4.7) it follows that

$$\Big|\int_{\Omega_z} f\varphi_z\,dx\Big| \le \Big(\sum_{E\in\mathcal{E}(z)} \eta_E^{p'}\Big)^{1/p'}\Big(\sum_{E\in\mathcal{E}(z)} h_E^{-p/p'}\,\|\varphi_z\|^p_{L^p(E)}\Big)^{1/p}. \tag{4.8}$$

Since the last factor in (4.8) is proportional to $h_z^{n/p\;-1}$ for $h_z = \operatorname{diam}(\omega_z)$, (4.7)-(4.8) yield

$$|\omega_z|^{-p'/p}\Big|\int_\Omega f\varphi_z\,dx\Big|^{p'} \lesssim h_z^{-p'}\sum_{E\in\mathcal{E}(z)} \eta_E^{p'}. \tag{4.9}$$

Since $\mathcal{E}(z)$, for free nodes $z \in \mathcal{K}$, have a finite overlap, the combination of (4.6) and (4.9) shows

$$\|h_{\mathcal{T}_\ell} f\|^{p'}_{L^{p'}(\Omega)} \approx \sum_{z\in\mathcal{K}} h_z^{p'}\|f\|^{p'}_{L^{p'}(\omega_z)} \lesssim \operatorname{osc}_\ell(f)^{p'} + \eta_\ell.$$

This and (4.3) proof the assertion. □

Remark 4.1. The condition that each element has at least one vertex, which is a free node, leads to $\Omega = \bigcup_{z\in\mathcal{K}}\omega_z$ in the proof of Lemma 4.2. This can be generalised by enlarging ω_z to Ω_z by some elements near the boundary. We refer to [5, 4, 7, 8] for details.

Lemma 4.3. *For any $E \in \mathcal{M}_\ell$ with $E = \partial T_+ \cup \partial T_-$ for $T_+, T_- \in \mathcal{T}_\ell$ and $\omega_E = \operatorname{int}(T_+ \cup T_-)$ there holds*

$$\eta_E \lesssim \|\sigma_{\ell+1} - \sigma_\ell\|_{L^{p'}(\omega_E;\mathbb{R}^{m\times n})} + \|f - f_{\omega_E}\|_{L^{p'}(\omega_E;\mathbb{R}^m)}.$$

Proof. `REFINE` allows for nodal basis functions φ_E of a new node $\mathrm{mid}(E)$ in E and ψ_E of a new node $\mathrm{mid}(\omega_E)$ in either T_+ or T_-, with respect to the finer triangulation $\mathcal{T}_{\ell+1}$ and E, T_+, T_- from $\mathcal{T}_\ell$. Then, there exists some linear combination

$$V_E := \alpha\varphi_E + \beta\psi_E \in V_{\ell+1} \cap W_0^{1,p}(\omega_E; \mathbb{R}^m)$$

with the following conditions

$$\int_E v_E\, ds = |E|, \; \int_{\omega_E} v_E\, dx = 0, \; \|v_E\|_V \approx h_E^{-1}\, |\omega_E|^{1/p}.$$

The construction of such V_E is the same as in linear problems [3, 13, 17, 18, 22] and hence the remaining details are neglected and the subsequent outline is kept brief. Since J_E is constant along E

$$|E| J_E = \int_E ([\sigma_\ell]\nu_E) \cdot v_E\, ds = \int_{\omega_E} \sigma_\ell : Dv_E\, dx.$$

Since $v_E \in V_{\ell+1}$ and $\sigma_{\ell+1}$ satisfy the discrete Euler-Lagrange equations,

$$\int_{\omega_E} \sigma_\ell : Dv_E\, dx = \int_{\omega_E} (\sigma_\ell - \sigma_{\ell+1}) : Dv_E\, dx + \int_{\omega_E} (f - f_{\omega_E}) \cdot v_E\, dx$$

with the constant integral mean f_{ω_E} of f over ω_E. The combination of the above identity with Friedrichs inequality $\|v_E\|_{L^p(\omega_E;\mathbb{R}^m)} \lesssim h_E \|v_E\|_V$ proves

$$\eta_E = h_E^{1/p'} |E|^{1/p'} |J_E| \lesssim h_E^{1/p'} |E|^{1/p} \big(\|\sigma_\ell - \sigma_{\ell+1}\|_{L^{p'}(\omega_E;\mathbb{R}^{m\times n})} + h_{\omega_E} \|f - f_{\omega_E}\|_{L^{p'}(\omega_E;\mathbb{R}^m)} \big) \|v_E\|_V. \qquad \square$$

Proof of Theorem 2.1. Notice that the patches have a finite overlap and

$$\sum_{E \in \mathcal{E}_\ell} h_E^{p'} \|f - f_{\omega_E}\|_{L^{p'}(\omega_E;\mathbb{R}^m)} \lesssim \mathrm{osc}_\ell^{p'}.$$

Hence Lemma 4.3 leads to

$$\sum_{E \in \mathcal{M}} \eta_E^{p'} \lesssim \|\sigma_{\ell+1} - \sigma_\ell\|_{L^{p'}(\Omega;\mathbb{R}^{m\times n})}^{p'} + \mathrm{osc}_\ell^{p'}.$$

This, (2.8) in `MARK` and Lemma 4.2 show

$$\begin{aligned} \|\sigma - \sigma_\ell\|_{L^{r/t}(\Omega;\mathbb{R}^{m\times n})}^{rp'} &\lesssim \eta_\ell^{p'} + \mathrm{osc}_\ell^{p'} \\ &\lesssim \sum_{E \in \mathcal{M}_\ell} \eta_E^{p'} + \mathrm{osc}_\ell^{p'} \\ &\lesssim \|\sigma_{\ell+1} - \sigma_\ell\|_{L^{p'}(\Omega;\mathbb{R}^{m\times n})}^{p'} + \mathrm{osc}_\ell^{p'}. \end{aligned} \tag{4.10}$$

Since $(\delta_\ell) \to \delta$, the right-hand side in Lemma 4.1 converges to zero, i.e.,

$$\lim_{\ell\to\infty} \|\sigma_{\ell+1} - \sigma_\ell\|_{L^{r/t}(\Omega;\mathbb{R}^{m\times n})} = 0.$$

Since $p' \leq r/t$ and $|\Omega| \lesssim 1$, the right-hand side in (4.10) tends to zero as $\ell \to \infty$. This proves the claimed strong convergence

$$\lim_{\ell\to\infty} \|\sigma - \sigma_\ell\|_{L^{r/t}(\Omega;\mathbb{R}^{m\times n})} = 0. \quad \square$$

Acknowledgements. This work was initiated and announced while the author enjoyed a research visit in 2002 at the Isaac Newton Institute for Mathematical Sciences, Cambridge, UK, continued in 2004 at the Institute for Mathematical Sciences, National University of Singapore, and eventually finalized in 2005 during a visit at the Institute of Mathematics and its Applications, Minneapolis, USA. The support by the EPSRC (N09176/ 01), FWF (P15274 and P16461), and the DFG throught Matheon is thankfully acknowledged. The author thanks Dr. Antonio Orlando for pointing out reference [15].

References

[1] Alberty, J. and Carstensen, C. and Zarrabi, D. (1999). *Adaptive numerical analysis in primal elastoplasticity with hardening.* Comput. Methods Appl. Mech. Engrg., **171**, 3-4, 175–204.

[2] Bartels, S. and Carstensen, C. (2006). *A convergent adaptive finite element method for an optimal design problem.* (in preparation).

[3] Binev, P., Dahmen, W. and DeVore, R. (2004). *Adaptive finite element methods with convergence rates.* Numer. Math., **97**, 2, 219–268.

[4] Carstensen, C. (1999). *Quasi-interpolation and a posteriori error analysis in finite element method.* M2AN Math. Model. Numer. Anal., **33**, 6, 1187–1202.

[5] Carstensen, C. (2004). *Some remarks on the history and future of averaging techniques in a posteriori finite element error analysis.* ZAMM Z. Angew. Math. Mech., **84**, 1, 3–21.

[6] Carstensen, C. (2004). *An adaptive mesh-refining algorithm allowing for an H^1-stable L^2-projection onto Courant finite element spaces.* Constr. Approx. Vol. **20**, No. 4, 549-564.

[7] Carstensen, C. (2006). *Clément interpolation and its role in adaptive finite element error control.* Operator Theory: Advances and Applications, Vol. **168**, 27-43, Birkhäuser Verlag Basel/Switzerland.

[8] Carstensen, C. and Jensen, M. (2006) *Averaging techniques for reliable and efficient a posteriori Finite Element Error Control: Analysis and Applications.* Contemporary Mathematics, 383 (2006), pp. 15-34.

[9] Carstensen, C. and Jochimsen, K. (2003). *Adaptive finite element methods for microstructures? Numerical experiments for a 2-well benchmark.* Computing, **71**, 2, 175–204.

[10] Carstensen, C. and Müller, S. (2002). *Local stress regularity in scalar nonconvex variational problems.* SIAM J. Math. Anal., **34**, 2, 495–509.

[11] Carstensen, C. and Plecháč, P. (1997). *Numerical solution of the scalar double-well problem allowing microstructure.* Math. Comp., **66**, 219, 997–1026.

[12] Carstensen, C. and Plecháč, P. (2000). *Numerical analysis of compatible phase transitions in elastic solids.* SIAM J. Numer. Anal., **37**, 6, 2061–2081.

[13] Dörfler, W. (1996). *A convergent adaptive algorithm for Poisson's equation.* SIAM J. Numer. Anal., **33**, 3, 1106–1124.

[14] Goodman, J. and Kohn, R.V. and Reyna, L. (1986). *Numerical study of a relaxed variational problem from optimal design.* Comput. Methods Appl. Mech. Engrg., **57**, 1, 107–127.

[15] Hiriart-Urruty, J.-B. (1998). *Optimisation et analyse convexe.* Presses Universitaires de France, Paris.

[16] D. Knees (2005). *Global stress regularity for convex and some nonconvex variational problems on Lipschitz domains*, SFB 404, Bericht 2005/04 (submitted).

[17] Morin, P. and Nochetto, R.H. and Siebert, K.G. (2000). *Data oscillation and convergence of adaptive FEM.* SIAM J. Numer. Anal., **38**, 2, 466–488.

[18] Morin, P. and Nochetto, R.H. and Siebert, K.G. (2003). *Local problems on stars: a posteriori error estimators, convergence, and performance.* Math. Comp., **72**, 243, 1067–1097.

[19] R. Rodriguez (1994). *Some remarks on Zienkiewicz-Zhu estimator.* Int. J. Numer. Methods Partial Diff. Equ. **10**, 625-635.

[20] R. Rodriguez (1994). *A posteriori error analysis in the finite element method,* in: *Finite Element Methods. 50 Years of the Courant Element*, Conference held at the Univ. of Jyväskylä, Finland (1993), edited by M. Krizek (Marcel Dekker, New York, NY); Inc. Lect. Notes Pure Appl. Math. **164**, 389-397.

[21] Temam, R. (1983). *Problèmes mathématiques en plasticité.* Méthodes Mathématiques de l'Informatique [Mathematical Methods of Information Science], **12**, Gauthier-Villars, Montrouge, vii+353.

[22] Veeser, A. (2002). *Convergent adaptive finite elements for the nonlinear Laplacian.* Numer. Math., **92**, 4, 743–770.

[23] Zeidler, E. (1988). *Nonlinear functional analysis and its applications. IV Applications to Mathematical Physics.* Springer-Verlag, New York.

Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

E-mail address: cc@math.hu-berlin.de