

An optimal adaptive finite element method for elastoplasticity

Carsten Carstensen · Andreas Schröder · Sebastian Wiedemann

Received: 12 June 2012 / Revised: 20 June 2014
© Springer-Verlag Berlin Heidelberg 2015

Abstract An adaptive finite element algorithm for problems in elastoplasticity with hardening is proven to be of optimal convergence with respect to the notion of approximation classes. The results rely on the equivalence of the errors of the stresses and energies resulting from Jensen's inequality. Numerical experiments study the influence of the hardening and bulk parameters to the convergence behavior of the AFEM algorithm. This is the first optimal adaptive FEM for a variational inequality.

Mathematics Subject Classification 65N30 · 65N50 · 74C05

1 Introduction

Elastoplasticity with hardening is of great importance in many problems of structural engineering. Well-known models are given by linear kinematic hardening and

This work was supported by the DFG Research Center Matheon, Project C13.

C. Carstensen · S. Wiedemann
Department of Mathematics, Humboldt-Universität zu Berlin, Unter den Linden 6,
10099 Berlin, Germany
e-mail: cc@math.hu-berlin.de

S. Wiedemann
e-mail: wiedeman@math.hu-berlin.de

C. Carstensen
Department of Computational Science and Engineering, Yonsei University,
50 Yonsei-ro, Seodaemun-gu, Seoul 120-749, Korea

A. Schröder (✉)
Department of Mathematics, University of Salzburg, Hellbrunnerstraße 34, 5020 Salzburg, Austria
e-mail: andreas.schroeder@sbg.ac.at

isotropic hardening allowing for the modeling of elastoplastic deformation in an incremental sense [7, 19, 24]. One pseudo time-step of the elastoplastic model is given by a variational inequality which is equivalent to a minimization problem with a non-differentiable dissipation functional arising from the plastic flow law.

Adaptive finite element methods are well-established to address exact solutions of lower regularity. Empirical evidence shows that efficient simulations converge with optimal rates in a sequence of automatic mesh-refinement steps. These methods are usually based on a posteriori error control which has reached a certain maturity for linear problems. We refer to the monographs [1, 3, 28] for an overview. The a posteriori error analysis for variational inequalities in elastoplasticity is established in [2]. In the presence of hardening, the reduced model without all internal variables is uniformly convex and hence the dual norms of the residuals from linear elasticity are applicable and lead to reliable and efficient error control for elastoplasticity with hardening.

Adaptive finite element methods generate a sequence of triangulations and, therewith, a sequence of approximations via the usual four steps of an adaptive finite element algorithm (AFEM)

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}. \quad (1)$$

The convergence analysis of such an AFEM algorithm started with the marking due to Dörfler [15] and was completed with the optimality argument due to Stevenson [25]. We refer to [13] for linear problems, to [8] for nonlinear problems and to [10] for mixed methods. The convergence analysis of AFEM algorithms for variational inequalities is introduced in [6, 9].

In this paper, the convergence of the AFEM algorithm is proven for the discretization of variational inequalities in elastoplasticity. The discretization is based on low-order finite elements on triangles which are refined by newest vertex bisection. A Dörfler marking is used in the marking step of (1). The key in the analysis is to show the equivalence of the errors of the stresses and energies. The convergence analysis was firstly presented in [12] and relies on the pointwise exactness of the material law. Jensen's inequality is applied to the convex dissipation functional and some orthogonality relations given by the L^2 projection are exploited. The convergence follows with some contraction argument.

The main result of this paper is the proof of optimal convergence in terms of degrees of freedom. We emphasize that this result is the first contribution to prove optimality of the AFEM algorithm for a variational inequality. We apply the methodology of [13] where the overlay of the triangulation generated by the AFEM algorithm and an optimal triangulation is considered in terms of an approximation class. Using its convergence property we show that the AFEM algorithm produces triangulations yielding the same convergence rate as those which are optimal with respect to all admissible triangulations given by newest vertex bisection.

The remaining parts of this paper are organized as follows. In the Sect. 2, we introduce the formulation of one quasi-static time step in the primal problem of elastoplasticity with isotropic as well as linear kinematic hardening. In the Sect. 3, the equivalence of the energy difference to the stresses is proven and a residual-based error estimation is introduced along with some basic estimations of the data oscil-

lations. Moreover, the AFEM algorithm with Dörfler marking is introduced in this section. The convergence of the AFEM algorithm is proven in Sect. 4, while its optimal convergence is shown in Sect. 5. Numerical experiments are presented in Sect. 6 to study the influence of the hardening and bulk parameters to the convergence behavior of the AFEM algorithm. Recall that, throughout this paper, the expression $A \lesssim B$ abbreviates $A \leq CB$ with a positive constant C and $A \approx B$ represents $A \lesssim B \lesssim A$. Standard notation on Lebesgue and Sobolev spaces applies throughout the paper and $(\cdot, \cdot)_{L^2}$ denotes the L^2 inner product.

We explicitly acknowledge that this optimality analysis has been discussed years ago with Y. Kondratyuck which led to some unpublished work on uniformly convex minimization.

2 The model of elastoplasticity

Let $\Omega \subset \mathbb{R}^d$ represent the reference configuration of an elastoplastic body in $d = 2, 3$ dimensions with Lipschitz boundary $\partial\Omega$ split into a closed Dirichlet part Γ_D of positive surface measure and a possibly empty Neumann part $\Gamma_N = \partial\Omega \setminus \Gamma_D$ with outer unit normal ν . The surface traction $g \in L^2(\Gamma_N; \mathbb{R}^d)$ acts along Γ_N and the applied volume force reads $f \in L^2(\Omega; \mathbb{R}^d)$. The deformation of the body is described by the displacement field

$$u \in V := \{v \in H^1(\Omega; \mathbb{R}^d) \mid v = 0 \text{ on } \Gamma_D\}.$$

The linearized Green strain tensor is defined by $\varepsilon(u) := (Du + Du^\top)/2$. The plastic strain is given by

$$p \in Q := L^2\left(\Omega; \mathbb{R}_{\text{sym,dev}}^{d \times d}\right)$$

with $\mathbb{R}_{\text{sym,dev}}^{d \times d} := \{q \in \mathbb{R}_{\text{sym}}^{d \times d} \mid \text{tr}(q) = 0\}$ and trace $\text{tr}(q) := \sum_{j=1}^d q_{jj}$. The linearized Green strain is split into an elastic part e and the plastic part p in a small strain framework as

$$\varepsilon(u) = e + p.$$

The stress tensor is defined as

$$\sigma(u, p) := \mathbb{C}(\varepsilon(u) - p)$$

with the isotropic elasticity tensor \mathbb{C} assumed to be constant and uniformly elliptic, i.e.

$$\kappa_{\mathbb{C}} |\tau|^2 \leq \mathbb{C}\tau : \tau \text{ for all } \tau \in L^2\left(\Omega; \mathbb{R}_{\text{sym}}^{d \times d}\right)$$

with some global constant $\kappa_C > 0$. Using this notation, the strong form of the equilibrium conditions reads

$$\operatorname{div} \sigma(u, p) + f = 0 \quad \text{in } \Omega \quad \text{and} \quad \sigma v = g \quad \text{on } \Gamma_N. \tag{2}$$

We consider a linear hardening state law of the form $\chi = \mathbb{H}p$ and $R = H\alpha$ with the isotropic hardening modulus $H > 0$ and the hardening tensor \mathbb{H} assumed to be symmetric with $\kappa_{\mathbb{H}}|\tau|^2 \leq (\mathbb{H}\tau) : \tau$ for all $\tau \in L^2(\Omega; \mathbb{R}_{\text{sym}}^{k \times k})$ and the global constant $\kappa_{\mathbb{H}} > 0$. Here, $\chi \in Q$ denotes the back stress tensor and $\alpha \in M := L^2(\Omega)$ the accumulated plastic strain with $R \in L^2(\Omega)$ as its dual variable. Rate-independent plasticity models is assigned by a dissipation potential $j = j(q, \beta)$ assumed to be lower semicontinuous, convex, as well as positively homogeneous of degree one. In the case of isotropic hardening, the dissipation potential reads

$$j(q, \beta) := \begin{cases} \sigma_y |q| & \text{for } |q| \leq \beta, \\ \infty & \text{else.} \end{cases}$$

Here and throughout this paper, $|q| := (q:q)^{1/2}$ for the product $A:B := \sum_{i,j=1}^d A_{ij}B_{ij}$ of two matrices $A, B \in \mathbb{R}^{d \times d}$ and $\sigma_y > 0$ is a material constant which is referred as the yield stress. Provided that homogeneous initial conditions holds, the subsequent inclusion condition follows from the time discrete form of the elastoplastic evolution law,

$$(p, \alpha) \in \partial j^*(\sigma, \chi, R).$$

Here, j^* is the Legendre–Fenchel transform of j , and ∂f denotes the subdifferential of a convex function f [20]. The equivalent dual form reads

$$(\sigma, \chi, R) \in \partial j(p, \alpha).$$

This is equivalent to the variational inequality

$$\begin{aligned} & (\mathbb{H}p - \sigma(u, p), p - q)_{L^2(\Omega; \mathbb{R}^{d \times d})} + (H\alpha, \alpha - \beta)_{L^2(\Omega)} \\ & + \int_{\Omega} j(p, \alpha) - j(q, \beta) \, dx \leq 0 \quad \text{for all } (p, \beta) \in Q \times M. \end{aligned} \tag{3}$$

The multiplication of (2) with test functions plus an integration by parts yields

$$(\sigma(u, p), \varepsilon(v))_{L^2(\Omega; \mathbb{R}^{d \times d})} = (f, v)_{L^2(\Omega; \mathbb{R}^d)} + (g, v)_{L^2(\Gamma_N; \mathbb{R}^d)} \quad \text{for all } v \in V. \tag{4}$$

The summation of (3) and (4) implies the variational inequality [19,24]

$$b(z - w) \leq a(w, z - w) + \psi(z) - \psi(w) \quad \text{for all } z \in W \tag{5}$$

with the abbreviations $w = (u, p, \alpha)$, $z = (v, q, \beta) \in W := V \times Q \times M$ and

$$\begin{aligned}
 a(w, z) &:= (\sigma(u, p), \varepsilon(v) - q)_{L^2(\Omega; \mathbb{R}^{d \times d})} + (\mathbb{H}p, q)_{L^2(\Omega; \mathbb{R}^{d \times d})} + (H\alpha, \beta)_{L^2(\Omega)}, \\
 \psi(z) &:= \int_{\Omega} j(q, \beta) \, dx, \\
 b(z) &:= (f, v)_{L^2(\Omega; \mathbb{R}^d)} + (g, v)_{L^2(\Gamma_N; \mathbb{R}^d)}.
 \end{aligned}$$

The Hilbert space W is equipped with the inner product

$$\begin{aligned}
 (w, z)_W &:= (u, v)_{H^1(\Omega; \mathbb{R}^d)} + (p, q)_{L^2(\Omega; \mathbb{R}^{d \times d})} + (\alpha, \beta)_{L^2(\Omega)} \\
 &\text{for } w = (u, p, \alpha), z = (v, q, \beta) \in W.
 \end{aligned}$$

The induced norm in W is denoted by $\|z\|_W := (z, z)_W^{1/2}$. The bilinear form a is symmetric and continuous on W . The linear form b is continuous and the functional ψ is convex, lower semi-continuous and positive homogeneous. Korn's inequality implies that a is W -elliptic, i.e.

$$\kappa \|z\|_W^2 \leq a(z, z) \quad \text{for all } z \in W$$

with the ellipticity constant $\kappa > 0$ [19, Eq. (7.52)]. Hence, the weak form (5) has a unique solution equivalently characterized as the unique minimizer of the functional

$$E(z) := \frac{1}{2} a(z, z) - b(z) + \psi(z) \quad \text{for } z \in W.$$

A simpler model of hardening is given by linear kinematic hardening with the dissipation potential

$$j(q, \beta) := \sigma_y |q|.$$

In this case, the accumulated plastic strain does not appear in the modeling given by the same variational inequality (5). In other words, the Hilbert space W reduces to $W := V \times Q$ and the bilinear form a to

$$\begin{aligned}
 a(w, z) &:= (\sigma(u, p), \varepsilon(v) - q)_{L^2(\Omega; \mathbb{R}^{d \times d})} + (\mathbb{H}p, q)_{L^2(\Omega; \mathbb{R}^{d \times d})} \\
 &\text{for } w = (u, p), z = (v, q) \in W.
 \end{aligned}$$

Again, a unique solution is guaranteed [11, 19].

Remark 1 The model given by linear kinematic hardening can be extended to multiyield plasticity to describe nonlinear hardening [11, 12]. This is done through an additive split of the plastic strain into multiple variables each representing the plastic strain associated to a particular yield surface. The yield surfaces differ only in the material dependent parameters $\sigma_{y,j}$ and \mathbb{H}_j describing particular yield stresses and hardening tensors.

Remark 2 The case $\mathbb{H} = 0$ models perfect plasticity characterized by the absence of hardening. In this case, the problem (5) is not well posed and the bilinear form a is not W -elliptic. However, there exist solution of perfect plasticity in a much weaker sense [26, 27]. Furthermore, the discretized problem may lead to difficulties in the numerical solution for H close to zero. In any case, stable approximations of the stress are feasible [12].

In what follows, we will focus on the modeling with isotropic and linear kinematic hardening. But, all arguments can be transferred to the case of linear kinematic hardening by omitting all terms concerning the accumulated plastic strain. The ellipticity constant κ arises in the following estimate.

Lemma 1 *The minimizer $w \in W$ of E and any $z \in W$ satisfies*

$$\kappa \|w - z\|_W^2 \leq E(z) - E(w). \quad (6)$$

Proof From the variational inequality (5), and the W -ellipticity of a , it follows that

$$\begin{aligned} \kappa \|z - w\|_W^2 &\leq \frac{1}{2} a(z - w, z - w) = \frac{1}{2} (a(z, z) - a(w, w)) - a(w, z - w) \\ &\leq \frac{1}{2} (a(z, z) - a(w, w)) - b(z - w) + \psi(z) - \psi(w) = E(z) - E(w). \end{aligned}$$

3 Adaptive finite element discretization

To discretize the variational inequality (5), finite dimensional subspaces V_ℓ of V , Q_ℓ of Q and M_ℓ of M are introduced. The discrete problem seeks a discrete solution $w_\ell = (u_\ell, p_\ell, \alpha_\ell) \in W_\ell := V_\ell \times Q_\ell \times M_\ell$ with

$$b(z_\ell - w_\ell) \leq a(w_\ell, z_\ell - w_\ell) + \psi(z_\ell) - \psi(w_\ell) \quad \text{for all } z_\ell \in W_\ell. \quad (7)$$

The discrete problem admits a unique solution w_ℓ which is also the minimizer of E over W_ℓ .

The subspaces V_ℓ , Q_ℓ and M_ℓ are specified via finite elements on shape regular triangulations $(\mathcal{T}_\ell)_{\ell \in \mathbb{N}_0}$ of Ω in triangles without hanging nodes. Throughout this paper, \mathcal{T}_ℓ is a refinement of an initial triangulation \mathcal{T}_0 . Possible refinements of \mathcal{T}_0 are restricted to those resulting from newest vertex bisection. The set of these admissible refinements is denoted by \mathbb{T} .

Let $\mathcal{P}_1(\mathcal{T}_\ell)$ be the space of piecewise affine linear and $\mathcal{P}_0(\mathcal{T}_\ell)$ be the space of piecewise constant functions with respect to \mathcal{T}_ℓ and set

$$V_\ell := \mathcal{P}_1(\mathcal{T}_\ell; \mathbb{R}^d) \cap V, \quad Q_\ell := \mathcal{P}_0(\mathcal{T}_\ell; \mathbb{R}_{\text{sym,dev}}^{d \times d}), \quad M_\ell := \mathcal{P}_0(\mathcal{T}_\ell).$$

Moreover, let $\sigma := \sigma(u, p)$ and $\sigma_\ell := \sigma(u_\ell, p_\ell)$ with $u_\ell \in V_\ell$ and $p_\ell \in Q_\ell$. The following theorem states that the error of stresses is equivalent to the error of energies.

Theorem 1 *The exact and discrete solutions w and w_ℓ with stress fields σ and σ_ℓ satisfy*

$$\|\sigma - \sigma_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})}^2 \approx E(w_\ell) - E(w) \leq (\sigma_\ell - \sigma, \varepsilon(u_\ell - u))_{L^2(\Omega; \mathbb{R}^{d \times d})}.$$

Proof The variational inequality (5) implies

$$\|\sigma - \sigma_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})}^2 \lesssim a(w_\ell - w, w_\ell - w) \leq E(w_\ell) - E(w).$$

Let $\bar{p}_\ell \in Q_\ell$ with $\bar{p}_\ell|_T := |T|^{-1} \int_T p \, dx$ and $\bar{\alpha}_\ell \in M_\ell$ with $\bar{\alpha}_\ell|_T := |T|^{-1} \int_T \alpha \, dx$ as well as $\bar{w}_\ell := (u_\ell, \bar{p}_\ell, \bar{\alpha}_\ell)$. Since j is convex, Jensen’s inequality yields

$$j(\bar{p}_\ell, \bar{\alpha}_\ell) \leq |T|^{-1} \int_T j(p, \alpha) \, dx.$$

This implies $\psi(\bar{w}_\ell) \leq \psi(w)$. Thus,

$$\begin{aligned} E(w_\ell) - E(w) &\leq a(w_\ell, w_\ell - w) - b(w_\ell - w) + \psi(w_\ell) - \psi(w) \\ &\leq a(w_\ell, \bar{w}_\ell - w) - b(\bar{w}_\ell - w). \end{aligned}$$

From the orthogonality relations

$$(q_\ell, p - \bar{p}_\ell)_{L^2(\Omega; \mathbb{R}^{d \times d})} = 0$$

for all $q_\ell \in \mathcal{P}_0(\mathcal{T}_\ell; \mathbb{R}^{d \times d})$ and

$$(\beta_\ell, \alpha - \bar{\alpha}_\ell) = 0$$

for all $\beta_\ell \in \mathcal{P}_0(\mathcal{T}_\ell)$, we obtain

$$a(w_\ell, \bar{w}_\ell - w) = (\sigma_\ell, \varepsilon(u_\ell - u))_{L^2(\Omega; \mathbb{R}^{d \times d})}.$$

From (4), Young’s inequality, and Lemma 1, we conclude

$$\begin{aligned} E(w_\ell) - E(w) &\leq a(w_\ell, \bar{w}_\ell - w) - (\sigma, \varepsilon(u_\ell - u))_{L^2(\Omega; \mathbb{R}^{d \times d})} \\ &= (\sigma_\ell - \sigma, \varepsilon(u_\ell) - \varepsilon(u))_{L^2(\Omega; \mathbb{R}^{d \times d})} \\ &\leq \frac{1}{2\kappa} \|\sigma - \sigma_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})}^2 + \frac{\kappa}{2} \|w_\ell - w\|_W^2 \\ &\leq \frac{1}{2\kappa} \|\sigma - \sigma_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})}^2 + \frac{1}{2} (E(w_\ell) - E(w)) \end{aligned}$$

with the ellipticity constant κ . This completes the proof.

The assertion of Theorem 1 remains valid for discretization spaces defined on refinements $\mathcal{T}_{\ell+m}$ of \mathcal{T}_ℓ with $m \geq 1$.

Corollary 1 *It holds*

$$\begin{aligned} \|\sigma_\ell - \sigma_{\ell+m}\|_{L^2(\Omega; \mathbb{R}^{d \times d})}^2 &\approx E(w_\ell) - E(w_{\ell+m}) \\ &\leq (\sigma_\ell - \sigma_{\ell+m}, \varepsilon(u_\ell - u_{\ell+m}))_{L^2(\Omega; \mathbb{R}^{d \times d})}. \end{aligned}$$

Proof Since $W_\ell \subset W_{\ell+m}$, the same arguments from the proof of Theorem 1 and Lemma 1 with w replaced by $w_{\ell+m}$ prove the assertion.

The residual-based explicit error estimation for the norm $\|\cdot\|_{L^2(\Omega; \mathbb{R}^{d \times d})}$ of stresses reads [12, 19]

$$\begin{aligned} \eta_\ell^2(T) &:= |T| \|f\|_{L^2(T; \mathbb{R}^d)}^2 + |T|^{1/2} \sum_{E \in \mathcal{E}(T)} R_E^2 \quad \text{for } T \in \mathcal{T}_\ell, \\ R_E &:= \begin{cases} \|[\sigma_\ell] \nu_E\|_{L^2(E; \mathbb{R}^d)} & \text{for } E \in \mathcal{E}_\ell(\Omega), \\ \|g - \sigma_\ell \nu_E\|_{L^2(E; \mathbb{R}^d)} & \text{for } E \in \mathcal{E}_\ell(\Gamma_N), \\ 0 & \text{for } E \in \mathcal{E}_\ell(\Gamma_D). \end{cases} \end{aligned}$$

Here and throughout this paper, $[\cdot]$ denotes the jump along an edge E of \mathcal{T}_ℓ and ν_E is a fixed unit normal vector with respect to E . The set $\mathcal{E}(T)$ is the set of all edges of $T \in \mathcal{T}_\ell$. Furthermore, $\mathcal{E}_\ell(\Omega)$ is the set of all interior edges of \mathcal{T}_ℓ , $\mathcal{E}_\ell(\Gamma_N)$ the set of all edges on Γ_N , and $\mathcal{E}_\ell(\Gamma_D)$ the set of all edges on Γ_D . The residual-based error estimator is

$$\eta_\ell^2 := \eta_\ell^2(\mathcal{T}_\ell), \quad \eta_\ell^2(\mathcal{M}) := \sum_{T \in \mathcal{M}} \eta_\ell^2(T) \quad \text{for } \mathcal{M} \subset \mathcal{T}_\ell.$$

With the integral means $f_T := |T|^{-1} \int_T f \, dx$ and $g_E := |E|^{-1} \int_E g \, ds$ for $E \in \mathcal{E}_\ell(\Gamma_N)$ we define the oscillations by

$$\begin{aligned} \text{osc}^2(f, T) &:= |T| \|f - f_T\|_{L^2(T; \mathbb{R}^d)}^2, \\ \text{osc}(f, \mathcal{M}) &:= \left(\sum_{T \in \mathcal{M}} \text{osc}^2(f, T) \right)^{1/2} \quad \text{for } \mathcal{M} \subset \mathcal{T}_\ell, \\ \text{osc}^2(g, E) &:= |T|^{1/2} \|g - g_E\|_{L^2(E; \mathbb{R}^d)}^2, \\ \text{osc}(g, \mathcal{F}) &:= \left(\sum_{E \in \mathcal{F}} \text{osc}^2(g, E) \right)^{1/2} \quad \text{for } \mathcal{F} \subset \mathcal{E}_\ell(\Gamma_N). \end{aligned}$$

The oscillations on a triangle $K \in \mathcal{T}_\ell$ dominate the oscillations on the subtriangulation $\mathcal{T}_{\ell+m}(K) := \{T \in \mathcal{T}_{\ell+m} \mid T \subset K\}$ and $\mathcal{E}_{\ell+m}(F) := \{E \in \mathcal{E}_{\ell+m}(\Gamma_N) \mid E \subset F\}$ with $F \in \mathcal{E}(K)$. In turn, the oscillations are dominated by the local error estimator $\eta_\ell(K)$. This is stated in the following lemma.

Lemma 2 Any $K \in \mathcal{T}_\ell$ satisfies

$$\text{osc}^2(f, \mathcal{T}_{\ell+m}(K)) \leq \text{osc}^2(f, K) \leq \eta_\ell^2(K),$$

and any $F \in \mathcal{E}(K) \cap \mathcal{E}_\ell(\Gamma_N)$ satisfies

$$\text{osc}^2(g, \mathcal{E}_{\ell+m}(F)) \leq \text{osc}^2(g, F) \leq \eta_\ell^2(K).$$

Proof We find that $\tilde{f} \in \mathcal{P}_0(\mathcal{T}_{\ell+m}(K); \mathbb{R}^d)$ with $\tilde{f}|_T := f_T$ and $T \in \mathcal{T}_{\ell+m}(K)$ is the L^2 projection of f on $\mathcal{P}_0(\mathcal{T}_{\ell+m}(K); \mathbb{R}^d)$ and that $f_K \in \mathcal{P}_0(\mathcal{T}_{\ell+m}(K); \mathbb{R}^d)$. Thus,

$$\begin{aligned} \text{osc}^2(f, \mathcal{T}_{\ell+m}(K)) &= \sum_{T \in \mathcal{T}_{\ell+m}(K)} |T| \|f - f_T\|_{L^2(T; \mathbb{R}^d)}^2 \\ &\leq |K| \sum_{T \in \mathcal{T}_{\ell+m}(K)} \|f - f_K\|_{L^2(T; \mathbb{R}^d)}^2 = \text{osc}^2(f, K). \end{aligned}$$

Since f_K is the L^2 projection of f in $\mathcal{P}_0(K; \mathbb{R}^d)$ and $0 \in \mathcal{P}_0(K; \mathbb{R}^d)$, we have $\text{osc}^2(f, K) \leq |K| \|f\|_{L^2(K; \mathbb{R}^d)}^2 \leq \eta_\ell(K)$. The second assertion follows by the same arguments.

It is well known that the estimator η_ℓ is reliable and efficient with constants that only depend on the material and hardening parameters as well as on the initial triangulation \mathcal{T}_0 .

Theorem 2 It holds

$$\|\sigma - \sigma_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})} \lesssim \eta_\ell \lesssim \|\sigma - \sigma_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})} + \text{osc}(f, \mathcal{T}_\ell) + \text{osc}(g, \mathcal{E}_\ell(\Gamma_N)).$$

Proof See [12] for a proof.

The error estimator remains reliable on the set of triangles refined from level ℓ to level $\ell + m$. This set is denoted by $\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+m} := \{T \in \mathcal{T}_\ell \mid T \notin \mathcal{T}_{\ell+m}\}$. Note the set of unrefined triangles is given by $\mathcal{T}_\ell \cap \mathcal{T}_{\ell+m}$.

Theorem 3 (Discrete reliability) It holds

$$E(w_\ell) - E(w_{\ell+m}) \lesssim \|\sigma_{\ell+m} - \sigma_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})}^2 \lesssim \eta_\ell^2(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+m}).$$

Proof The Scott–Zhang interpolation on $\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+m}$ of $u_\ell - u_{\ell+m}$ defines $v_\ell \in V_\ell$ with $u_\ell - u_{\ell+m} - v_\ell = 0$ on $\mathcal{T}_\ell \cap \mathcal{T}_{\ell+m}$. Corollary 1 and

$$(\sigma_{\ell+m} - \sigma_\ell, \varepsilon(v_\ell))_{L^2(\Omega; \mathbb{R}^{d \times d})} = 0$$

imply

$$\begin{aligned} & \|\sigma_\ell - \sigma_{\ell+m}\|_{L^2(\Omega; \mathbb{R}^{d \times d})}^2 \\ & \lesssim (\sigma_{\ell+m} - \sigma_\ell, \varepsilon(u_{\ell+m} - u_\ell - v_\ell))_{L^2(\Omega; \mathbb{R}^{d \times d})} \\ & = (f, u_{\ell+m} - u_\ell - v_\ell)_{L^2(\Omega; \mathbb{R}^d)} - (\sigma_\ell, \varepsilon(u_\ell - u_{\ell+m} - v_\ell))_{L^2(\Omega; \mathbb{R}^{d \times d})}. \end{aligned}$$

This and other arguments from [13, 25] prove the assertion. Since the remaining details are the same for linear problems, they are omitted here.

Algorithm 1 The AFEM algorithm.

Input: Loads $f \in L^2(\Omega)$, $g \in L^2(\Gamma_N)$, initial triangulation \mathcal{T}_0 , bulk parameter $0 < \theta \leq 1$.

for $\ell = 0, 1, \dots$ **do**

SOLVE: Compute the solution $w_\ell \in W_\ell$ of (7).

ESTIMATE: For all $T \in \mathcal{T}_\ell$, compute the estimated error $\eta_\ell(T)$.

MARK: Determine the set \mathcal{M}_ℓ of all elements marked for refinement by Dörfler marking/bulk chasing, i.e. compute $\mathcal{M}_\ell \subset \mathcal{T}_\ell$ of minimal cardinality $|\mathcal{M}_\ell|$ with

$$\theta \eta_\ell^2 \leq \eta_\ell^2(\mathcal{M}_\ell). \tag{8}$$

REFINE: Refine all elements in \mathcal{M}_ℓ with newest vertex bisection strategy, where green, blue and *bisec3* refinement patterns from Fig. 1 are allowed to be used to compute the regular refinement $\mathcal{T}_{\ell+1}$.

end for

Output: Sequence of shape regular triangulations $(\mathcal{T}_\ell)_{\ell \in \mathbb{N}_0} \subset \mathbb{T}$, sequence of discrete solutions $(w_\ell)_{\ell \in \mathbb{N}_0}$ in the nested spaces $(W_\ell)_{\ell \in \mathbb{N}_0}$.

Algorithm 1 describes the AFEM algorithm used for adaptive refinement. It consists in loops over SOLVE, ESTIMATE, MARK, and REFINE steps as introduced in (1). In the SOLVE step, an appropriate solution scheme for the discretization of elastoplastic problems has to be applied. We refer to modified Newton’s methods [14, 17, 21, 29] and to algorithms of predictor-corrector type [19]. Due to its implementational simplicity, we use Uzawa’s method based on a discretized mixed variational formulation which is equivalent to the discrete variational inequality (7) [16, 18]. We emphasize that efficient Newton-type methods as referred above should be preferred because of their much better convergence properties. Dörfler marking is used in the MARK step [15], and newest vertex bisection is applied in the REFINE step, see Fig. 1.

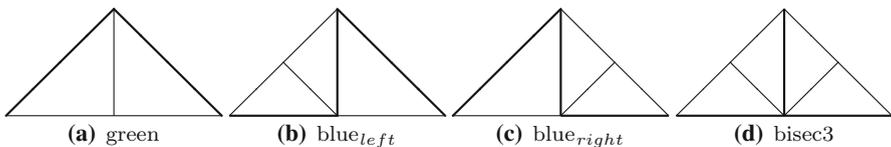


Fig. 1 Possible refinements of a triangle in the AFEM algorithm. *Thick lines* denote refinement edges for subsequent newest vertex bisection

4 Convergence of the AFEM algorithm

The point of departure for the AFEM convergence analysis is the error estimator reduction for increasing refinement levels $\ell \in \mathbb{N}$. For this purpose, define the patches $\omega_E := \text{int}(T \cup T')$ for $T, T' \in \mathcal{T}_\ell$ and $E \in \mathcal{E}(T) \cap \mathcal{E}(T')$ as well as $\omega_T := \bigcup_{E \in \mathcal{E}(T)} \omega_E$. Note that the trace theorem and the shape regularity of \mathcal{T}_ℓ imply

$$\begin{aligned} \|[\sigma_{\ell+m} - \sigma_\ell] \nu_E\|_{L^2(E; \mathbb{R}^d)} &\lesssim |E|^{-1/2} \|\sigma_{\ell+m} - \sigma_\ell\|_{L^2(\omega_E; \mathbb{R}^{d \times d})} \quad \text{for } E \in \mathcal{E}_\ell(\Omega), \\ \|(\sigma_{\ell+m} - \sigma_\ell) \nu_E\|_{L^2(E; \mathbb{R}^d)} &\lesssim |E|^{-1/2} \|\sigma_{\ell+m} - \sigma_\ell\|_{L^2(\omega_E; \mathbb{R}^{d \times d})} \quad \text{for } E \in \mathcal{E}_\ell(\Gamma_N). \end{aligned} \tag{9}$$

In the following, we apply similar techniques as proposed in [13]. In particular, we make use of the identity

$$(a + b)^2 = \min_{0 < t < \infty} \left((1 + t)a^2 + (1 + 1/t)b^2 \right) \quad \text{for } a, b \geq 0. \tag{10}$$

Lemma 3 *There exists a constant $\Lambda_0 > 0$ such that*

$$\begin{aligned} \eta_{\ell+m}^2(T) &\leq (1 + \lambda)\eta^2(T) + \Lambda_0(1 + 1/\lambda)\|\sigma_{\ell+m} - \sigma_\ell\|_{L^2(\omega_T; \mathbb{R}^{d \times d})}^2 \\ &\quad \text{for } \lambda > 0 \text{ and } T \in \mathcal{T}_\ell \cap \mathcal{T}_{\ell+m}. \end{aligned}$$

Proof The triangle inequality leads to

$$\begin{aligned} &|\eta_{\ell+m}(T) - \eta_\ell(T)| \\ &\leq |T|^{1/4} \left(\sum_{E \in \mathcal{E}(T) \cap \mathcal{E}_\ell(\Omega)} (\|[\sigma_{\ell+m}] \nu_E\|_{L^2(E; \mathbb{R}^d)} - \|[\sigma_\ell] \nu_E\|_{L^2(E; \mathbb{R}^d)})^2 \right. \\ &\quad \left. + \sum_{E \in \mathcal{E}(T) \cap \mathcal{E}_\ell(\Gamma_N)} (\|g_E - \sigma_{\ell+m} \nu_E\|_{L^2(E; \mathbb{R}^d)} - \|g_E - \sigma_\ell \nu_E\|_{L^2(E; \mathbb{R}^d)})^2 \right)^{1/2} \\ &\leq |T|^{1/4} \left(\sum_{E \in \mathcal{E}(T) \cap \mathcal{E}_\ell(\Omega)} \|[\sigma_{\ell+m} - \sigma_\ell] \nu_E\|_{L^2(E; \mathbb{R}^d)}^2 \right. \\ &\quad \left. + \sum_{E \in \mathcal{E}(T) \cap \mathcal{E}_\ell(\Gamma_N)} \|(\sigma_\ell - \sigma_{\ell+m}) \nu_E\|_{L^2(E; \mathbb{R}^d)}^2 \right)^{1/2}. \end{aligned}$$

The combination with (9)–(10) concludes the proof.

Lemma 4 *There exists a constant $\Lambda_1 > 0$ such that*

$$\begin{aligned} \eta_{\ell+m}^2(\mathcal{T}_{\ell+m}(K)) &\leq 2^{-1/2}(1 + \lambda)\eta_\ell^2(K) + \Lambda_1(1 + 1/\lambda)\|\sigma_{\ell+m} - \sigma_\ell\|_{L^2(\omega_K; \mathbb{R}^{d \times d})}^2 \\ &\quad \text{for } \lambda > 0 \text{ and } K \in \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+m}. \end{aligned}$$

Proof The refinement patterns of the newest vertex bisection of the AFEM algorithm from Fig. 1 shows $|T| \leq |K|/2$ for all $T \in \mathcal{T}_{\ell+m}(K)$. Thus, (10) implies

$$\begin{aligned} &\eta_{\ell+m}^2(\mathcal{T}_{\ell+m}(K)) \\ &= \sum_{T \in \mathcal{T}_{\ell+m}(K)} |T| \|f\|_{L^2(T; \mathbb{R}^d)}^2 \\ &\quad + \sum_{T \in \mathcal{T}_{\ell+m}(K)} |T|^{1/2} \sum_{E \in \mathcal{E}(T) \cap \mathcal{E}_{\ell+m}(\Omega)} \|[\sigma_{\ell+m}] \nu_E\|_{L^2(E; \mathbb{R}^d)}^2 \\ &\quad + \sum_{E \in \mathcal{E}(T) \cap \mathcal{E}_{\ell+m}(\Gamma_N)} \|(g_E - \sigma_{\ell+m}) \nu_E\|_{L^2(E; \mathbb{R}^d)}^2 \\ &\leq |K|/2 \sum_{T \in \mathcal{T}_{\ell+m}(K)} \|f\|_{L^2(T; \mathbb{R}^d)}^2 \\ &\quad + (1 + 1/\lambda) \sum_{T \in \mathcal{T}_{\ell+m}(K)} |T|^{1/2} \sum_{E \in \mathcal{E}(T) \cap \mathcal{E}_{\ell+m}(\Omega)} \|[\sigma_{\ell+m} - \sigma_\ell] \nu_E\|_{L^2(E; \mathbb{R}^d)}^2 \\ &\quad + \sum_{E \in \mathcal{E}(T) \cap \mathcal{E}_{\ell+m}(\Gamma_N)} \|(\sigma_{\ell+m} - \sigma_\ell) \nu_E\|_{L^2(E; \mathbb{R}^d)}^2 \\ &\quad + (|K|/2)^{1/2} (1 + \lambda) \sum_{F \in \mathcal{E}(K) \cap \mathcal{E}_\ell(\Omega)} \|[\sigma_\ell] \nu_F\|_{L^2(F; \mathbb{R}^d)}^2 \\ &\quad + \sum_{F \in \mathcal{E}(K) \cap \mathcal{E}_\ell(\Gamma_N)} \|(g_F - \sigma_\ell) \nu_F\|_{L^2(F; \mathbb{R}^d)}^2. \end{aligned}$$

The assertion follows from the definition of $\eta^2(K)$ and (9).

Theorem 4 *There exists a constant $\Lambda > 0$ which only depends on the initial triangulation \mathcal{T}_0 such that the estimators η_ℓ and $\eta_{\ell+m}$ satisfy*

$$\eta_{\ell+m} \leq \left(\eta_\ell^2(\mathcal{T}_\ell \cap \mathcal{T}_{\ell+m}) + 2^{-1/2} \eta_\ell^2(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+m}) \right)^{1/2} + \Lambda \|\sigma_{\ell+m} - \sigma_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})}.$$

Proof From Lemmas 3 and 4, we obtain

$$\begin{aligned} \eta_{\ell+m}^2 &= \eta_{\ell+m}^2(\mathcal{T}_\ell \cap \mathcal{T}_{\ell+m}) + \eta_{\ell+m}^2(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+m}) \\ &\leq (1 + \lambda) \left(\eta_\ell^2(\mathcal{T}_\ell \cap \mathcal{T}_{\ell+m}) + 2^{-1/2} \eta_\ell^2(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+m}) \right) \\ &\quad + 4(1 + 1/\lambda) \max\{\Lambda_0, \Lambda_1\} \|\sigma_{\ell+m} - \sigma_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})}^2. \end{aligned}$$

The assertion directly follows from (10).

The following corollary is a direct consequence of the Dörfler marking (8).

Theorem 5 *Given the bulk parameter $0 < \theta \leq 1$ and the constant $\Lambda > 0$ from Theorem 4, then $\rho_1 := (1 - \theta + 2^{-1/2}\theta)^{1/2} < 1$ satisfies*

$$\eta_{\ell+1} \leq \rho_1 \eta_\ell + \Lambda \|\sigma_{\ell+1} - \sigma_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})}.$$

Proof The Dörfler marking (8) implies

$$\theta \eta_\ell^2 \leq \eta_\ell^2(\mathcal{M}_\ell) \leq \eta_\ell^2(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}) = \eta_\ell^2 - \eta_\ell^2(\mathcal{T}_\ell \cap \mathcal{T}_{\ell+1}).$$

This leads to

$$\eta_\ell^2(\mathcal{T}_\ell \cap \mathcal{T}_{\ell+1}) \leq (1 - \theta)\eta_\ell^2.$$

Theorem 4 yields

$$\begin{aligned} \eta_{\ell+1} &\leq \left((1 - \theta)\eta_\ell^2 + 2^{-1/2} \left(\eta_\ell^2 - \eta_\ell(\mathcal{T}_\ell \cap \mathcal{T}_{\ell+1}) \right) \right)^{1/2} + \Lambda \|\sigma_{\ell+1} - \sigma_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})} \\ &\leq \left((1 - \theta)\eta_\ell^2 + 2^{-1/2} \left(\eta_\ell^2 - (1 - \theta)\eta_\ell^2 \right) \right)^{1/2} + \Lambda \|\sigma_{\ell+1} - \sigma_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})} \\ &= \rho_1 \eta_\ell + \Lambda \|\sigma_{\ell+1} - \sigma_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})}. \end{aligned}$$

The convergence of the AFEM algorithm directly results from the convergence of the weighted sum

$$\xi_\ell^2 := \eta_\ell^2 + \beta \delta_\ell$$

with $\delta_\ell := E(w_\ell) - E(w)$ and some $\beta \geq 0$. This is stated in the following theorem.

Theorem 6 *There exist parameters $\beta \geq 0$ and $0 < \rho_2 < 1$ such that*

$$\xi_{\ell+1} \leq \rho_2 \xi_\ell \text{ for all } \ell \in \mathbb{N}_0.$$

Proof For $0 < \lambda < \rho_1^{-2} - 1$ Theorem 5 and (10) imply

$$\eta_{\ell+1}^2 \leq (1 + \lambda)\rho_1^2 \eta_\ell^2 + (1 + 1/\lambda)\Lambda^2 \|\sigma_{\ell+1} - \sigma_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})}^2.$$

For $\rho_\lambda := (1 + \lambda)\rho_1^2 < 1$ and a further constant $\beta_\lambda \geq 0$, which also depends on λ , Corollary 1 shows

$$\eta_{\ell+1}^2 \leq \rho_\lambda \eta_\ell^2 + \beta_\lambda (E(w_\ell) - E(w_{\ell+1})) = \rho_\lambda \eta_\ell^2 + \beta_\lambda \delta_\ell - \beta_\lambda \delta_{\ell+1}.$$

From Theorems 1 and 2, we conclude that there exists some constant $C > 0$ such that $\delta_\ell \leq C \eta_\ell^2$. With

$$\vartheta := \frac{(1 - \rho_\lambda)\beta_\lambda}{\beta_\lambda C + 1} \text{ and } \rho_2 := \rho_\lambda + \vartheta C < 1,$$

we obtain

$$\eta_{\ell+1}^2 + \beta_\lambda \delta_{\ell+1} \leq \rho_\lambda \eta_\ell^2 + \beta_\lambda \delta_\ell \leq \rho_2 \eta_\ell^2 + (\beta_\lambda - \vartheta)\delta_\ell = \rho_2(\eta_\ell + \beta_\lambda \delta_\ell).$$

5 Optimal convergence of the AFEM algorithm

The convergence rate of the AFEM algorithm is described via the notion of an approximation class. For this purpose, set

$$W(\mathcal{T}) := \mathcal{P}_1(\mathcal{T}; \mathbb{R}^d) \times \mathcal{P}_0\left(\mathcal{T}; \mathbb{R}_{\text{sym,dev}}^{d \times d}\right) \times \mathcal{P}_0(\mathcal{T})$$

for a triangulation \mathcal{T} and denote the minimizer of E over $W(\mathcal{T})$ by $w_{\mathcal{T}} \in W(\mathcal{T})$. Furthermore, recall the definition of \mathbb{T} as the set of refinements of the initial triangulation \mathcal{T}_0 resulting from newest vertex bisection and define the subset $\mathbb{T}(N) := \{\mathcal{T} \in \mathbb{T} \mid |\mathcal{T}| - |\mathcal{T}_0| \leq N\}$ for $N \in \mathbb{N}$. Given any $0 < s < \infty$ define the semi-norm

$$\begin{aligned} & |(w, f, g)|_{\mathcal{A}_s} \\ & := \sup_{N \in \mathbb{N}} N^s \min_{\mathcal{T} \in \mathbb{T}(N)} \left(\text{osc}^2(f, \mathcal{T}) + \text{osc}^2(g, \mathcal{E}(\Gamma_N)) + E(w_{\mathcal{T}}) - E(w) \right)^{1/2} \end{aligned}$$

and the approximation class

$$\mathcal{A}_s := \left\{ (w, f, g) \in W \times L^2(\Omega; \mathbb{R}^d) \times L^2(\Gamma_N; \mathbb{R}^d) \mid |(w, f, g)|_{\mathcal{A}_s} < \infty \right\}.$$

The main result of this paper states the optimal convergence of the AFEM algorithm with respect to the energy and in the norm $\|\cdot\|_{L^2(\Omega, \mathbb{R}^{d \times d})}$ of stresses.

Theorem 7 *There exist $0 < \theta_0 \leq 1$ and a constant $C(s) > 0$ such that for all bulk parameters $0 < \theta \leq \theta_0$ of the AFEM algorithm it holds*

$$\begin{aligned} & (|\mathcal{T}_\ell| - |\mathcal{T}_0|)^s (E(w_\ell) - E(w) + \text{osc}^2(f, \mathcal{T}_\ell) + \text{osc}^2(f, \mathcal{E}_\ell(\Gamma_N)))^{1/2} \\ & \leq C(s) |(w, f, g)|_{\mathcal{A}_s}. \end{aligned} \quad (11)$$

Furthermore, there exists a constant $\bar{C}(s) > 0$ such that

$$\begin{aligned} & (|\mathcal{T}_\ell| - |\mathcal{T}_0|)^s (\|\sigma - \sigma_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})}^2 + \text{osc}^2(f, \mathcal{T}_\ell) + \text{osc}^2(g, \mathcal{E}_\ell(\Gamma_N)))^{1/2} \\ & \leq \bar{C}(s) |(w, f, g)|_{\mathcal{A}_s}. \end{aligned} \quad (12)$$

The remaining part of this section is devoted to the proof in four steps.

Step 1. Given $0 < \tau < 1$ and $(w, f, g) \in \mathcal{A}_s$, choose a minimal $N_\ell \in \mathbb{N}$ such that

$$|(w, f, g)|_{\mathcal{A}_s} \leq \tau \xi_\ell N_\ell^s.$$

Such a minimal N_ℓ satisfies

$$N_\ell \leq 2(N_\ell - 1) \leq 2|(w, f, g)|_{\mathcal{A}_s}^{1/s} (\tau \xi_\ell)^{-1/s}. \quad (13)$$

The definition of the approximation class \mathcal{A}_s implies the existence of a triangulation $\tilde{\mathcal{T}}_\ell \in \mathbb{T}(N_\ell)$ and a discrete solution $\tilde{w}_\ell \in W(\tilde{\mathcal{T}}_\ell)$ such that

$$E(\tilde{w}_\ell) - E(w) + \text{osc}^2(f, \tilde{\mathcal{T}}_\ell) + \text{osc}^2(g, \tilde{\mathcal{E}}_\ell(\Gamma_N)) \leq N_\ell^{-2s} |(w, f, g)|_{\mathcal{A}_s}^2 \leq (\tau \xi_\ell)^2 \tag{14}$$

with the set $\tilde{\mathcal{E}}_\ell(\Gamma_N)$ of all edges of $\tilde{\mathcal{T}}_\ell$ on Γ_N .

Step 2 considers the (unique) smallest common refinement $\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell \in \mathbb{T}$ of \mathcal{T}_ℓ and $\tilde{\mathcal{T}}_\ell$ [13], which satisfies

$$|\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell| - |\mathcal{T}_\ell| \leq |\tilde{\mathcal{T}}_\ell| - |\mathcal{T}_0|.$$

Notice carefully that

$$\begin{aligned} |\mathcal{T}_\ell \setminus (\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell)| &\leq \sum_{T \in \mathcal{T}_\ell \setminus (\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell)} (|(\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell)(T)| - 1) \\ &= |(\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell) \setminus \mathcal{T}_\ell| - |\mathcal{T}_\ell \setminus (\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell)| = |\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell| - |\mathcal{T}_\ell| \end{aligned}$$

and conclude

$$|\mathcal{T}_\ell \setminus (\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell)| \leq |\tilde{\mathcal{T}}_\ell| - |\mathcal{T}_0| \leq N_\ell. \tag{15}$$

Step 3 proves

$$\eta_\ell \lesssim \eta_\ell \left(|\mathcal{T}_\ell \setminus (\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell)| \right). \tag{16}$$

In fact, Theorems 2 and 3 imply

$$\begin{aligned} \eta_\ell^2 &\lesssim \|\sigma - \sigma_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})}^2 + \text{osc}^2(f, \mathcal{T}_\ell) + \text{osc}^2(g, \mathcal{E}_\ell(\Gamma_N)) \\ &\lesssim \|\hat{\sigma}_\ell - \sigma_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})}^2 + \|\sigma - \hat{\sigma}_\ell\|_{L^2(\Omega; \mathbb{R}^{d \times d})}^2 + \text{osc}^2(f, \mathcal{T}_\ell) + \text{osc}^2(g, \mathcal{E}_\ell(\Gamma_N)) \\ &\lesssim \eta_\ell^2 (|\mathcal{T}_\ell \setminus (\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell)|) + E(\hat{w}_\ell) - E(w) + \text{osc}^2(f, \mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell) + \text{osc}^2(g, \hat{\mathcal{E}}_\ell(\Gamma_N)). \end{aligned}$$

Here, $\hat{\mathcal{E}}_\ell(\Gamma_N)$ is the set of all edges of $\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell$ on Γ_N and $\hat{\sigma}_\ell := \sigma(\hat{u}_\ell, \hat{p}_\ell)$ with the discrete solution $\hat{w}_\ell := (\hat{u}_\ell, \hat{p}_\ell, \hat{\alpha}_\ell) \in W_{\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell}$. Note

$$\begin{aligned} \text{osc}^2(f, \mathcal{T}_\ell) &= \text{osc}^2(f, \mathcal{T}_\ell \cap (\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell)) + \text{osc}^2(f, \mathcal{T}_\ell \setminus (\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell)) \\ &\leq \text{osc}^2(f, \mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell) + \eta_\ell^2 (|\mathcal{T}_\ell \setminus (\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell)|). \end{aligned}$$

Lemma 2 and (14) yield

$$\begin{aligned} E(\hat{w}_\ell) - E(w) + \text{osc}^2(f, \mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell) + \text{osc}^2(g, \hat{\mathcal{E}}_\ell(\Gamma_N)) \\ \leq E(\tilde{w}_\ell) - E(w) + \text{osc}^2(f, \tilde{\mathcal{T}}_\ell) + \text{osc}^2(g, \tilde{\mathcal{E}}_\ell(\Gamma_N)) \\ \leq (\tau \xi_\ell)^2 \lesssim \tau^2 \eta_\ell^2. \end{aligned}$$

Hence, $\eta_\ell^2 \lesssim \eta_\ell^2(\mathcal{T}_\ell \setminus (\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell)) + \tau^2 \eta_\ell^2$. The estimate (16) follows with a sufficiently small $\tau > 0$.

Step 4 utilizes the BDV-Theorem [5] for newest-vertex bisection, namely

$$|\mathcal{T}_\ell| - |\mathcal{T}_0| \leq |\mathcal{T}_\ell \setminus \mathcal{T}_0| \lesssim |\mathcal{M}_0 \cup \dots \cup \mathcal{M}_{\ell-1}|. \tag{17}$$

The estimate (16) implies the existence of a parameter $0 < \theta_0 \leq 1$ so that $\theta_0 \eta_\ell^2 \leq \eta_\ell^2(\mathcal{T}_\ell \setminus (\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell))$. This means that $\mathcal{T}_\ell \setminus (\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell)$ also satisfies the bulk criterion (8) for all bulk parameters $0 < \theta \leq \theta_0$. Since \mathcal{M}_ℓ is of minimal cardinality, (13) and (15) imply

$$|\mathcal{M}_\ell| \leq |\mathcal{T}_\ell \setminus (\mathcal{T}_\ell \oplus \tilde{\mathcal{T}}_\ell)| \leq N_\ell \leq 2|(w, f, g)|_{\mathcal{A}_s}^{1/s} (\tau \xi_\ell)^{-1/s}.$$

The estimate (17) leads to

$$|\mathcal{T}_\ell| - |\mathcal{T}_0| \lesssim \sum_{k=0}^{\ell-1} |\mathcal{M}_k| \lesssim |(w, f, g)|_{\mathcal{A}_s}^{1/s} \tau^{-1/s} \sum_{k=0}^{\ell-1} \xi_k^{-1/s}.$$

Theorem 6 with $0 < \rho_2 < 1$ yields

$$\xi_\ell \leq \rho_2^{\ell-k} \xi_k \quad \text{for } 0 \leq k \leq \ell.$$

Hence,

$$\sum_{k=0}^{\ell-1} \xi_k^{-1/s} \leq \xi_\ell^{-1/s} \sum_{k=0}^{\ell-1} \rho_2^{(\ell-k)/s} \leq \xi_\ell^{-1/s} \frac{\rho_2^{1/s}}{1 - \rho_2^{1/s}}.$$

Therefore,

$$\xi_\ell (|\mathcal{T}_\ell| - |\mathcal{T}_0|)^s \lesssim \frac{\rho_2}{\tau (1 - \rho_2^{1/s})^s} |(w, f, g)|_{\mathcal{A}_s}.$$

The definition of ξ_ℓ as well as Theorems 1 and 2 eventually yield the assertion (11). The assertion (12) directly follows from Theorem 1 and (11). \square

6 Numerical results

The numerical experiments of this section study the influence of the hardening tensor \mathbb{H} and the choice of the bulk parameter θ to the convergence properties of the AFEM algorithm. In the first experiment, we consider the L-shaped domain $\Omega := [0, 1]^2 \setminus [0, 0.5]^2$ and apply Hooke’s material law with Lamé’s constants $\lambda := 1000$ and $\mu := 1000$. The volume force is set to $f := 0$ on Ω and the surface traction to $g := (0.75, 0)^\top$ on $[0, 1] \times \{1\}$. We prescribe homogeneous Dirichlet boundary

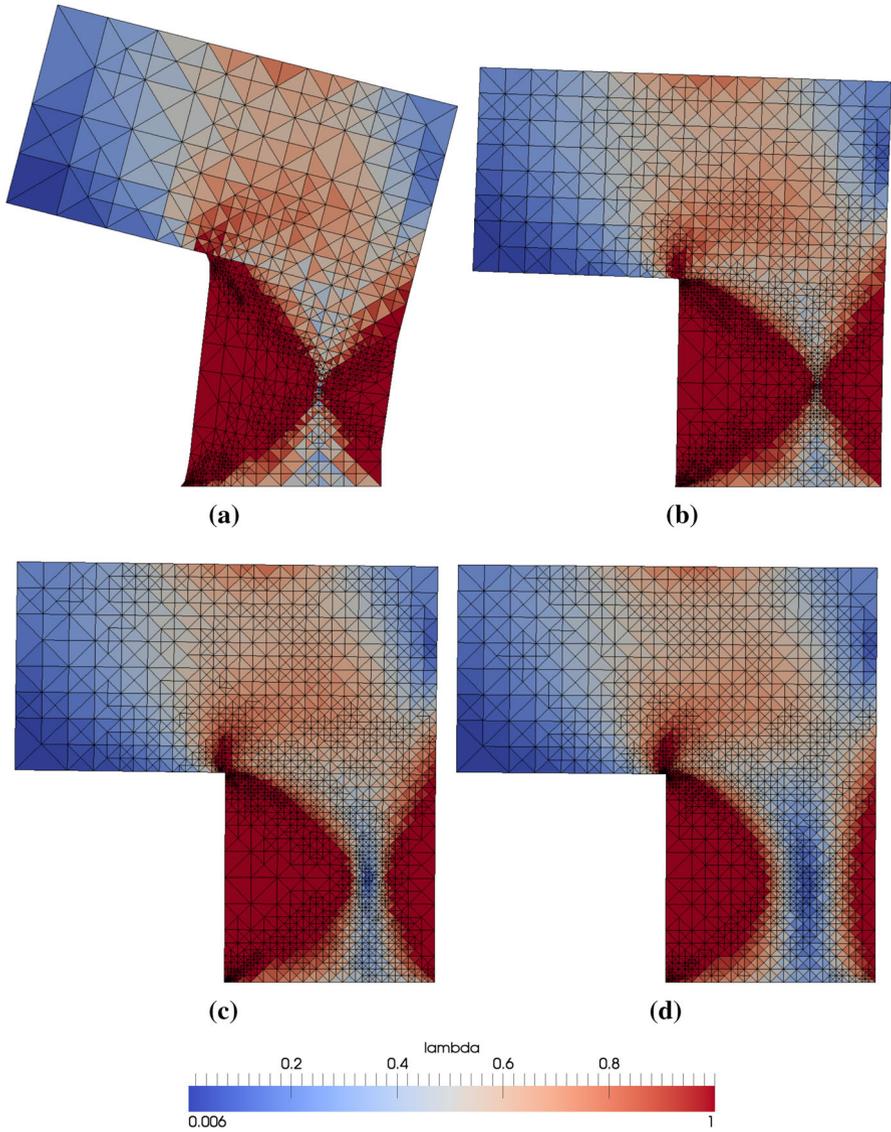
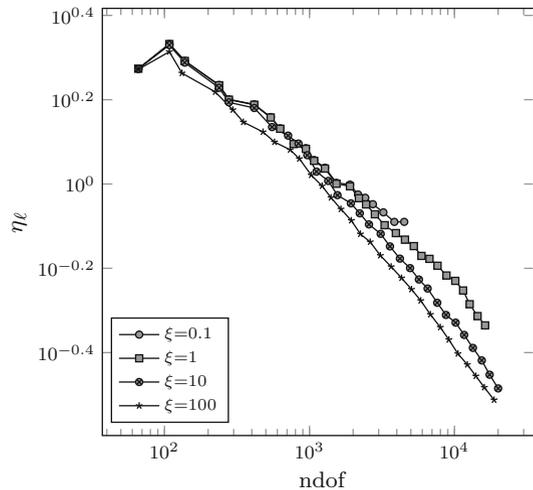


Fig. 2 Displacement and plastic/elastic zone for $\sigma_y = 2.5$ and **a** $\xi = 0.1$, **b** $\xi = 1$, **c** $\xi = 10$, **d** $\xi = 100$

conditions on $\Gamma_D := [0, 0.5] \times \{0\}$. The hardening tensor is defined as $\mathbb{H} := \xi \mathbb{I}$ with $\xi > 0$ and identity tensor \mathbb{I} . In Fig. 2, adaptive meshes after 30 refinements with the bulk parameter $\theta := 0.1$ and the hardening parameter $\xi \in \{0.1, 1, 10, 100\}$ are shown. In this experiment, the yield stress is set to $\sigma_y := 2.5$. The figures show

$$\lambda_\ell := \sigma_y^{-1} |\text{dev}(\sigma(u_\ell, p_\ell) - \mathbb{H}p_\ell)|$$

Fig. 3 Estimated error for different hardening parameters ξ



indicating $p_\ell = 0$ if $\lambda_\ell < 1$ [23]. We observe high yielding of the material for the lowest hardening parameter $\xi = 0.1$ resulting from tension on $[0, 1] \times \{1\}$. The adaptive refinements resolve the singularity at the reentrant corner at $(0, 0)$ as well as the transition regions from elasticity ($\lambda_\ell < 1$) to plasticity ($\lambda_\ell = 1$). In Fig. 3, the estimated error η_ℓ is plotted for the different hardening parameters with respect to the number of degrees of freedom (ndof). For the moderate hardening parameters $\xi = 10$ or $\xi = 100$, we find nearly parallel estimated convergence rates. This is not the case for the small hardening parameters $\xi = 1$ and $\xi = 0.1$. The corresponding convergence rates seem to have slightly smaller slopes than the rates given by the moderate hardening parameters. The reason for this observation could be the essentially smaller ellipticity constant κ which directly influences the efficiency and reliability constants of the estimator described by Theorem 2. This, in turn, may lead to a larger pre-asymptotic region of the convergence. Note that ξ tending to zero leads to the case of perfect plasticity. The problem (5) is no longer well posed in this case since the bilinear form a is not guaranteed to be W -elliptic. For ξ close to zero, numerical algorithms may often fail to solve the discretized problem in practice.

The estimated error for several bulk parameters

$$\theta \in \{0.1, 0.25, 0.5, 0.9, 0.95, 1\} \quad (18)$$

is shown in the Figs. 4 and 5. Here, the hardening parameter is set to $\xi := 100$ and the yield stress to $\sigma_y := 1.25$. Obviously, the convergence rates resulting from the adaptive refinements ($\theta < 1$) have a larger slope than the uniform refinements ($\theta = 1$). In fact, the low regularity properties of the solution resulting from the singularity at the reentrant corner and the elastoplastic material modeling prevents better convergence rates in the case of uniform refinements. Furthermore, we observe that the pre-asymptotic region is more distinct in comparison to the results of Fig. 3.

Fig. 4 Estimated error for adaptive and uniform refinements

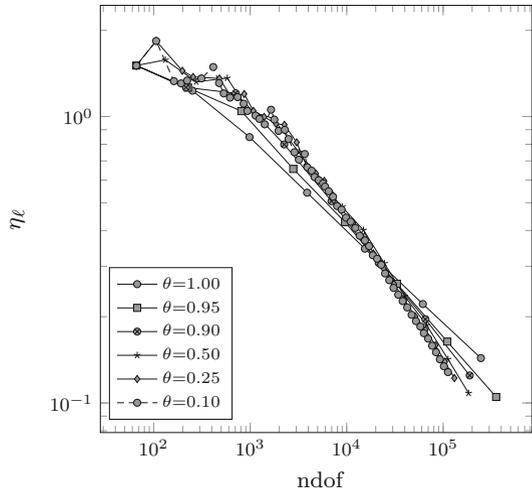
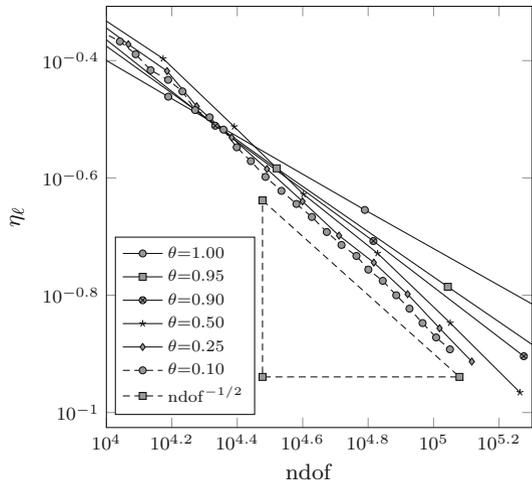


Fig. 5 Zoom of Fig. 4



This may result from the smaller yield stress in this parameter configuration leading to a larger region of plastic deformation, see Fig. 6b.

Theorem 7 states the existence of a certain limit θ_0 so that optimal convergence rates are ensured for all bulk parameters smaller than this limit. Figure 5 shows a zoom of Fig. 4. Therein, we observe that the bulk parameter $\theta = 0.5$ seems to be a candidate of such a limit in the selection (18) of bulk parameters, i.e. the convergence rates ($\approx \mathcal{O}(\text{ndof}^{-1/2})$) are not improved if smaller bulk parameters are chosen. The adaptive mesh for this bulk parameter is depicted in Fig. 6a where, again, the singularity at the reentrant corner and the transition regions are resolved by adaptivity.

To estimate the limit θ_0 more precisely, we may have a closer look at the proof of Theorem 7 (Step 3 and Step 4) where θ_0 is presumed as

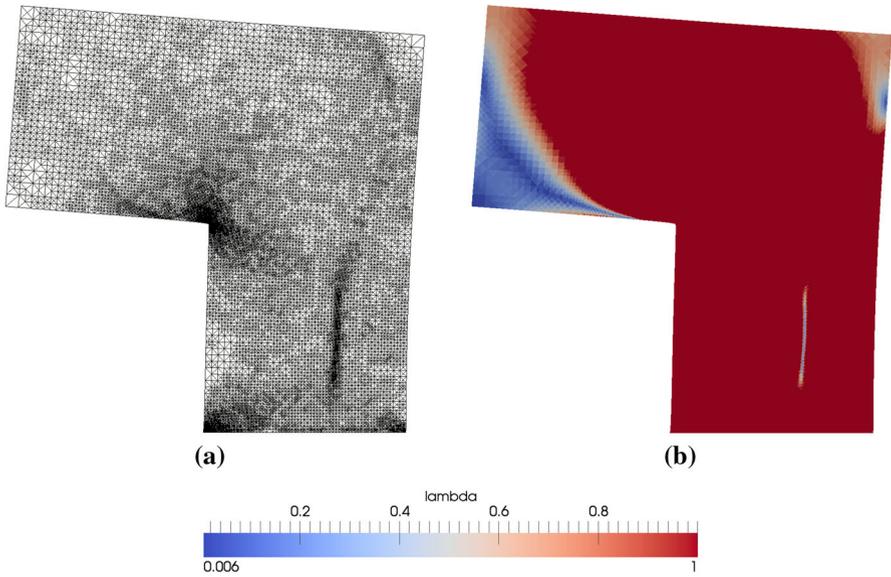
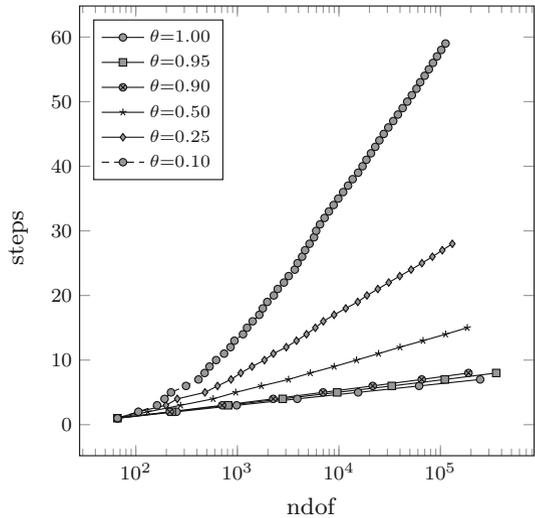


Fig. 6 **a** Adaptive mesh for $\theta = 0.5$. **b** Displacement and plastic/elastic zone for $\sigma_y = 1.25$ and $\xi = 100$

Fig. 7 Number of degrees of freedom (ndof) versus number of refinement steps



$$\theta_0 = \frac{C_{\text{eff}} - \tilde{\tau}}{(1 + \epsilon)C_{\text{rel}}}$$

with the efficiency constant $C_{\text{eff}} > 0$ of Theorem 2 and the discrete reliability constant $C_{\text{rel}} > 0$ of Theorem 3 as well as some parameters $0 < \tilde{\tau} < C_{\text{eff}}$ and $\epsilon > 0$. Thus, $\theta_0 := \min\{C_{\text{eff}} C_{\text{rel}}^{-1}, 1\}$ may be an appropriate choice.

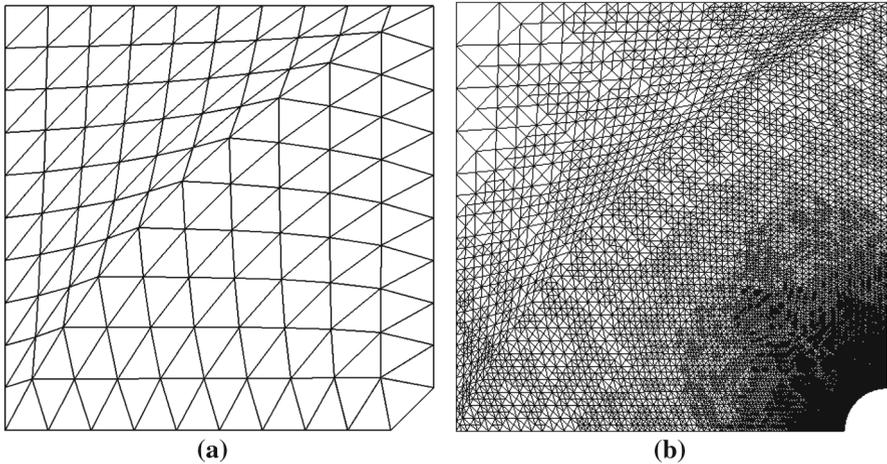


Fig. 8 **a** Coarse mesh of the benchmark problem. **b** Adaptive mesh for $\theta = 0.1$

Recall that in each refinement step a non-linear problem has to be solved which produces the largest computational amount. Hence, it is favorable to choose a bulk parameter which is as large as possible since this reduces the number of refinement steps in the AFEM algorithm. In Fig. 7, the number of degrees of freedom (ndof) versus the number of refinement steps are plotted. We see that these numbers differs strongly, in particular, the smaller the bulk parameter is chosen. Thus, to ensure optimal convergence rates with minimal computational amount, it is reasonable to choose the limit θ_0 as bulk parameter.

In the first experiment, we observe that local refinements result from different types of sources, the reentrant corner and the elastoplastic material modeling. In the second experiment of this section, we consider a problem where local refinements only result from elastoplasticity. The configuration of this experiment is based on the benchmark problem as introduced, for instance, in [4,22]. In this problem, the upper left quarter of a stretched square with a hole is considered. The quarter may be described by

$$\Omega := (0, 10)^2 \setminus \{(x, y) \in \mathbb{R}^2 \mid (x - 10)^2 + y^2 \leq 1\}.$$

Again, we use Hook’s law, here with modulus of elasticity $E := 206900$ and Poisson’s ratio $\nu := 0.29$. The volume force is set to $f := 0$ on Ω and the surface traction is set to $g := (0.75, 0)^T$ on $[0, 10] \times \{10\}$. The yield stress is set to $\sigma_y := 400$. In accordance to the benchmark problem, we use the boundary conditions $u_2 := 0$ on $[0, 9] \times \{0\}$ and $u_1 := 0$ on $\{10\} \times [1, 10]$. The hardening tensor is defined as above via $\xi := 10000$.

In Fig. 8a, the coarse mesh is depicted that is used in the experiment where the hole at the bottom right corner is approximated during the refinement process. Figure 8b shows the adaptive mesh after 39 refinement steps with bulk parameter $\theta := 0.1$. The local refinements coincide with the plastic zone that is indicated by $\lambda_\ell = 1$ in Fig. 9. The convergence of the adaptive scheme is depicted in Fig. 10. We observe

Fig. 9 Plastic and elastic zones indicated by λ_ℓ

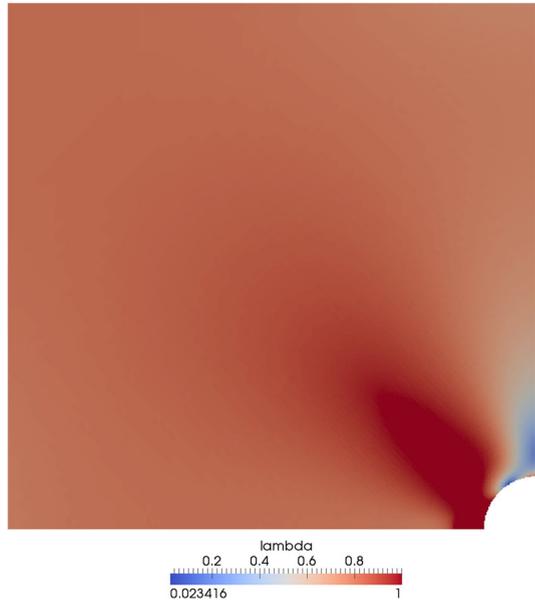
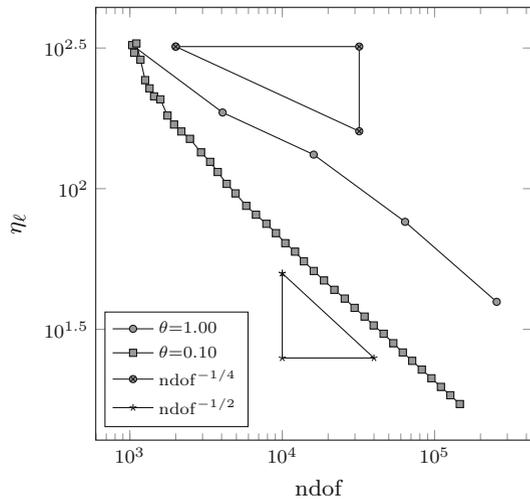


Fig. 10 Estimated error for adaptive and uniform refinements of the benchmark problem



that the convergence rates resulting from the adaptive refinements are clearly of the order $\mathcal{O}(\text{ndof}^{-1/2})$ whereas the convergence rates resulting from uniform refinements ($\theta = 1.0$) are slightly smaller than these rates.

Recall that the mesh size of uniformly refined meshes is $\mathcal{O}(\text{ndof}^{-1/2})$, the convergence rates of the proposed adaptive scheme with low order finite elements seem to be optimal in both experiments as predicted.

References

1. Ainsworth, M., Oden, J.: A posteriori error estimation in finite element analysis. Pure and Applied Mathematics. A Wiley-Interscience Series of Texts, Monographs, and Tracts. Wiley, New York (2000)
2. Albery, J., Carstensen, C., Zarrabi, D.: Adaptive numerical analysis in primal elastoplasticity with hardening. *Comput. Methods Appl. Mech. Eng.* **171**(3–4), 175–204 (1999)
3. Babuska, I., Strouboulis, T.: The Finite Element Method and Its Reliability. Numerical Mathematics and Scientific Computation. Clarendon Press, Oxford (2001)
4. Barthold, F.J., Schmidt, M., Stein, E.: Error indicators and mesh refinements for finite element computations of elastoplastic deformations. *Comput. Mech.* **22**(3), 225–238 (1998)
5. Binev, P., Dahmen, W., DeVore, R.: Adaptive finite element methods with convergence rates. *Numer. Math.* **97**(2), 219–268 (2004)
6. Braess, D., Carstensen, C., Hoppe, R.: Convergence analysis of a conforming adaptive finite element method for an obstacle problem. *J. Numer. Math.* **107**(3), 455–471 (2007)
7. Carstensen, C.: Numerical analysis of the primal problem of elastoplasticity with hardening. *Numer. Math.* **82**(4), 577–597 (1999)
8. Carstensen, C.: Convergence of an adaptive fem for a class of degenerate convex minimisation problems. *IMA J. Numer. Anal.* **28**(3), 423–439 (2007)
9. Carstensen, C.: Convergence of adaptive finite element methods in computational mechanics. *Appl. Numer. Math.* **59**(9), 2119–2130 (2009)
10. Carstensen, C., Rabus, H.: An optimal adaptive mixed finite element method. *Math. Comput.* **80**(274), 649–667 (2011)
11. Carstensen, C., Valdman, J., Brokate, M.: A quasi-static boundary value problem in multi-surface elastoplasticity: part 1—analysis. *Math. Methods Appl. Sci.* **27**, 1697–1710 (2004)
12. Carstensen, C., Valdman, J., Orlando, A.: A convergent adaptive finite element method for the primal problem of elastoplasticity. *Int. J. Numer. Methods Eng.* **67**(13), 1851–1887 (2006)
13. Cascon, J., Kreuzer, C., Nochetto, R.H., Siebert, K.G.: Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.* **46**(5), 2524–2550 (2008)
14. Christensen, P.W.: A nonsmooth newton method for elastoplastic problems. *Comput. Methods Appl. Mech. Eng.* **191**(11–12), 1189–1219 (2002)
15. Dörfler, W.: A convergent adaptive algorithm for poisson’s equation. *SIAM J. Numer. Anal.* **33**(3), 1106–1124 (1996)
16. Glowinski, R., Lions, J.L., Tremolieres, R.: Numerical analysis of variational inequalities. Transl. and rev. ed. Studies in Mathematics and its Applications, vol. 8. North-Holland Publishing Company, Amsterdam (1981)
17. Gruber, P., Valdman, J.: Solution of one-time-step problems in elastoplasticity by a slant newton method. *SIAM J. Sci. Comput.* **31**(2), 1558–1580 (2009)
18. Han, W.: Finite element analysis of a holonomic elastic-plastic problem. *Numer. Math.* **60**(4), 493–508 (1992)
19. Han, W., Reddy, B.: Plasticity. Mathematical theory and numerical analysis. Interdisciplinary Applied Mathematics, vol. 9. Springer, Berlin (1999)
20. Hiriart-Urruty, J.B., Lemaréchal, C.: Fundamentals of Convex Analysis, Text Editions. Springer, Grundlehren (2001)
21. Hofinger, A., Valdman, J.: Numerical solution of the two-yield elastoplastic minimization problem. *Computing* **81**(1), 35–52 (2007)
22. Lang, S., Wieners, C., Wittum, G.: The application of adaptive parallel multigrid methods to problems in nonlinear solid mechanics. In: Stein, E. (ed.) Error-controlled Adaptive Finite Elements in Solid Mechanics, chap. 10, pp. 347–384. Wiley, New York (2002)
23. Schröder, A., Wiedemann, S.: Error estimates in elastoplasticity using a mixed method. *Appl. Numer. Math.* **61**, 1031–1045 (2011)
24. Simo, J., Hughes, T.: Computational inelasticity. Interdisciplinary Applied Mathematics, vol. 7. Springer, Berlin (1998)
25. Stevenson, R.: Optimality of a standard adaptive finite element method. *Found. Comput. Math.* **7**(2), 245–269 (2007)
26. Suquet, P.: Discontinuities and plasticity. Nonsmooth mechanics and applications, CISM Courses Lect. **302**, 279–340 (1988)

27. Temam, R.: Problèmes mathématiques en plasticité. Méthodes Mathématiques de l'Informatique, vol. 12. Publié avec le concours du C.N.R.S. Paris, Gauthier-Villars, Bordas (1983)
28. Verfürth, R.: A review of a posteriori error estimation techniques for elasticity problems. *Comput. Methods Appl. Mech. Eng.* **176**(1–4), 419–440 (1999)
29. Wieners, C.: Nonlinear solution methods for infinitesimal perfect plasticity. *ZAMM Z. Angew. Math. Mech.* **87**(8–9), 643–660 (2007)