Error estimates in elastoplasticity using a mixed method

Andreas Schröder [a,1,*], Sebastian Wiedemann[a,1]

[a] *Department of Mathematics, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin*

Abstract

In this paper, a mixed formulation and its discretization are introduced for elastoplasticity with linear kinematic hardening. The mixed formulation relies on the introduction of a Lagrange multiplier to resolve the non-differentiability of the plastic work function. The main focus is on the derivation of a priori and a posteriori error estimates based on general discretization spaces. The estimates are applied to several low-order finite elements. In particular, a posteriori estimates are expressed in terms of standard residual estimates. Numerical experiments are presented, confirming the applicability of the a posteriori estimates within an adaptive procedure.

Keywords:
error estimation, elastoplasticity, mixed method, adaptivity

---

1. Introduction

Elastoplasticity with hardening is of great importance in many problems of mechanical engineering. A well-established case of elastoplasticity with linear kinematic hardening is the holonomic constitutive law which allows for the modeling of elastoplastic deformation in an incremental sense, cf. [15, 16]. One pseudo time-step of the holonomic model is given by a variational inequality of second kind which is equivalent to a minimization problem with a non-differentiable term arising from the plastic flow law.

This non-differentiable term can be resolved by the regularization of the plastic work function using an appropriate regularization so that Newton's method can be applied to obtain a suitable discretization scheme. However, the influence of the regularization to the discretization can cause numerical difficulties which must be handled very carefully. This is possible, for instance, through the adaptation of a regularization parameter during the Newton iteration and through an appropriate damping strategy of Newton's method. We refer to [8, 9, 12, 17] for more details on solution schemes in elastoplasticity.

An alternative approach not requiring regularization is given by the mixed formulation which captures the plastic work function as a supremum over a convex and closed set of bounded Lagrange multipliers. The corresponding minimization problem is then equivalent to a saddle point problem of which the stationary conditions yield a variational equation in terms of the deformation and plastic variable as well as a variational inequality in terms of the introduced Lagrange multiplier. The mixed formulation can be discretized by finite element approaches and solved, for instance, by the simple Uzawa's method with projection, cf. [15].

At the first glance, this mixed method seems to be unfavorable in comparison to discretization schemes based on a pure primal formulation, since the Lagrange multiplier is an additional variable which also has to be discretized involving the same order of unknowns as the primal variables (in contrast to mixed methods, e.g., for contact problems where the Lagrange multiplier is only defined on the boundary.) However, the discrete Lagrange multiplier can be used to detect regions of pure elastic deformations which is often a difficult task.

In the literature, mixed methods that resolve the non-differentiable part of the plastic work function by a Lagrange multiplier are rarely studied. We refer to [15, 16] for the derivation of such mixed methods. To the best of the authors' knowledge there are no a priori results including the discretization of Lagrange multipliers. In [21], some

[*] Corresponding author, email: schroder@math.hu-berlin.de
[1] Supported by the German Research Foundation (DFG) within the Priority Program 1180.

recent results on a posteriori estimates are published. For some goal-oriented error estimates for a model problem in elastoplasticity, we refer to [23].

In this paper, we discuss the mixed formulation and its discretization for elastoplasticity with linear kinematic hardening. Since we assume the set of discrete Lagrange multipliers to be a subset of the discretization space of the plastic variable, the stability of the discretization is straightforward. Moreover, under this assumption, the mixed method is equivalent to the discretization of the variational inequality of second kind given by the pure primal approach.

The main focus is on the derivation of a priori and a posteriori error estimates for the mixed method with general discretization spaces. We apply the estimates to several low-order finite elements on triangle, quadrilateral, tetrahedron and hexahedron meshes. Constricting to the primal variable and using piecewise linear elements, we find the same results as for the pure primal approach which is, of course, expected due to the equivalence of the discretization approaches. Moreover, we obtain the same a posteriori estimates expressed in terms of standard residuals as already known in the literature for this case, cf. [3, 7].

The piecewise linear approach on triangles or tetrahedron provides the deviatoric part of the discrete stress tensor to be contained in the discretization space of the plastic variable which is usually assumed to be piecewise constant. For piecewise bilinear or trilinear approaches on quadrilaterals or hexahedrons this property is not given if the plastic variables are approximated by piecewise constant functions as well. We find that in this case, an additional term measuring the error of the L^2-projection of some deviatoric part is unavoidable in the a posteriori estimates.

The paper is organized as follows: In the Section 2, we introduce the formulation of one quasi-static time step in the primal problem of elastoplasticity with linear kinematic hardening. In the Sections 3 and 4, we consider the mixed formulation and introduce a discretization for general discretization spaces. A priori and a posteriori estimates are derived in Section 5 and Section 6, respectively. They are based on general discretization spaces as well. Low-order finite elements are introduced in Section 7 and the application of the estimates to them is contained in the Sections 8 and 9. Finally, numerical results are presented in Section 10, confirming the applicability of the a posteriori estimates within an adaptive procedure. Furthermore, we introduce Uzawa's method with projection and discuss the detection of regions with pure elastic deformations.

2. Elastoplasticity with linear kinematic hardening

We consider the deformation of a body which is represented by a domain $\Omega \subset \mathbb{R}^k$, $k \in \{2, 3\}$, with a sufficiently smooth boundary $\Gamma := \partial\Omega$. The body is clamped at a boundary part given by the closed set $\Gamma_D \subset \Gamma$ with positive measure. Volume and surface forces act on the body. They are described by functions $f \in L^2(\Omega; \mathbb{R}^k)$ and $g \in L^2(\Gamma_N; \mathbb{R}^k)$ with $\Gamma_N := \Gamma\backslash\Gamma_D$, respectively. The resulting deformation is described by displacement fields $v \in H^1(\Omega; \mathbb{R}^k)$ with the linearized strain tensor $\varepsilon(u) := \frac{1}{2}(\nabla u + (\nabla u)^\top)$. The elasticity tensor $\mathbb{C}$ with $\mathbb{C}_{ijkl} \in L^\infty(\Omega)$ satisfies the standard symmetry condition $\mathbb{C}_{ijkl} = \mathbb{C}_{jilk} = \mathbb{C}_{klij}$ and is uniformly elliptic, i.e. with a constant $\kappa > 0$ there holds $(\mathbb{C}\tau) : \tau \geq \kappa\tau : \tau$ for all $\tau \in L^2(\Omega; \mathbb{R}^{k\times k}_{\text{sym}})$ where $\mathbb{R}^{k\times k}_{\text{sym}} := \{M \in \mathbb{R}^{k\times k} \mid M_{ij} = M_{ji}\}$ and $\tau : \tau := \sum_{i,j=1}^k \tau_{ij}\tau_{ij}$. The stress tensor is defined as $\sigma(u, p) := \mathbb{C}(\varepsilon(u) - p)$ where $p \in L^2(\Omega; \mathbb{R}^{k\times k}_{\text{sym}})$ with trace $\operatorname{tr}(p) := \sum_{j=1}^k p_{jj} = 0$ is the plastic strain. We set $H^1_D(\Omega) := \{v \in H^1(\Omega; \mathbb{R}^k) \mid \gamma_{|\Gamma_D}(v_i) = 0,\ i = 1, \ldots, k\}$ with the trace operator $\gamma \in L(H^1(\Omega), L^2(\Gamma))$ and define $\sigma_n(u, p) := \sigma(u, p)n$ with outer normal n of Γ. The primal problem of elastoplasticity with linear kinematic hardening is to find a displacement field u and a plastic strain p such that

$$
\begin{aligned}
-\operatorname{div}\sigma(u, p) &= f \text{ in } \Omega,\\
u &= 0 \text{ on } \Gamma_D,\\
\sigma_n(u, p) &= g \text{ on } \Gamma_N,\\
\sigma(u, p) - \mathbb{H}p &\in \partial j(p).
\end{aligned}
\tag{2.1}
$$

Here, j is the non-differentiable part of the minimum plastic work function with $j(q) := \sigma_y(q : q)^{1/2}$ and the yield stress $\sigma_y > 0$ in uniaxial tension. The hardening tensor $\mathbb{H}$ is assumed to be symmetric and positive definite, i.e., $(\mathbb{H}\tau) : \tau > 0$ for all $\tau \in L^2(\Omega; \mathbb{R}^{k\times k}_{\text{sym}})$, $\tau > 0$.

Remark 2.1. Replacing $\partial j(p)$ by $\partial j(\dot{p})$ with $\dot{p} = \partial p/\partial t$ in the inclusion condition we observe that the formulation (2.1) describes one time step of quasi-static elastoplasticity with hardening and initial conditions $p_0 = 0$, cf. [7, 15, 17].

3. Mixed variational formulation of elastoplasticity with linear kinematic hardening

To derive a variational formulation, let $(\cdot,\cdot)_{0,\omega}$, $(\cdot,\cdot)_{0,\Gamma'}$ be the usual L^2-scalar products for vector-valued or matrix-valued functions on $\omega \subset \mathbb{R}^k$ and $\Gamma' \subset \Gamma\backslash\Gamma_D$, respectively. For $v \in H^1_D(\Omega)$, $v' \in L^2(\Gamma';\mathbb{R}^k)$ and $q \in L^2(\Omega;\mathbb{R}^{k\times k})$, we define $\|v\|^2_{0,\omega} := (v,v)_{0,\omega}$, $\|v'\|^2_{0,\Gamma'} := (v',v')_{0,\Gamma'}$ and $\|q\|^2_{0,\omega} := (q,q)_{0,\omega}$. We write $(\cdot,\cdot)_0$ instead of $(\cdot,\cdot)_{0,\Omega}$ if not stated otherwise. Moreover, we state $|v|_1^2 := (\epsilon(v),\epsilon(v))_0$ and $\|v\|_1^2 := \|v\|_0^2 + |v|_1^2$ which is equivalent to the usual H^1-norm on $H^1_D(\Omega)$ due to Korn's inequality. We set

$$Q := \{q \in L^2(\Omega;\mathbb{R}^{k\times k}_{\mathrm{sym}}) \mid \mathrm{tr}(q) = 0 \text{ a.e. in } \Omega\}$$

and define $W := H^1_D(\Omega)\times Q$, which is a Hilbert space with the norm $\|(v,q)\|^2 := \|v\|_1^2 + \|q\|_0^2$ for $(v,q)\in W$.
It is well-known, that the solution $w = (u,p) \in W$ of (2.1) fulfills the variational equation

$$(\sigma(w),\varepsilon(v))_0 = (f,v)_0 + (g,\gamma_{|\Gamma_N}(v))_{0,\Gamma_N} \tag{3.1}$$

for all $v \in H^1_D(\Omega)$ and, additionally, the variational inequality

$$(\mathbb{H}p - \sigma(w), q-p)_0 + \int_\Omega j(q) - j(p)\,dx \geq 0 \tag{3.2}$$

for all $q \in Q$, cf. [5, 15, 17]. Defining

$$\begin{aligned} a((u,p),(v,q)) &:= (\sigma(u,p),\varepsilon(v)-q)_0 + (\mathbb{H}p,q)_0,\\ \psi(v,q) &:= \int_\Omega j(q)\,dx,\\ \ell(v,q) &:= (f,v)_0 + (g,\gamma_{|\Gamma_N}(v))_{0,\Gamma_N}, \end{aligned}$$

we observe that $w\in W$ fulfills (3.1) and (3.2) if and only if the variational inequality of second kind

$$a(w,z-w) + \psi(z) - \psi(w) \geq \ell(z-w) \tag{3.3}$$

holds for all $z \in W$. Note, the bilinear form a is continuous and W-elliptic, i.e., there exist constants $\nu_0,\nu_1 > 0$ such that

$$a(z,z') \leq \nu_0\|z\|\,\|z'\|, \quad \nu_1\|z\|^2 \leq a(z,z) \tag{3.4}$$

for all $z, z' \in W$, see [17].

Remark 3.1. For $\mathbb{H}$ tending to zero, we would arrive at the case of perfect plasticity. Thus, problem (3.3) is no longer well posed in this case. In fact, if $\mathbb{H}$ tends to a non-positive definite tensor the bilinear form a is no longer guaranteed to be W-elliptic. In this article we do not focus on the problem of perfect plasticity and therefore neglect the influence of $\mathbb{H}$ on the error estimates. We emphasize that some of the following estimates no longer hold in the case of $\mathbb{H} = 0$. Furthermore, numerical algorithms to solve the discretized problem may face difficulties in practice for $\mathbb{H}$ close to zero. In any case, stable approximations of the stress remain possible, cf. [10].

The inequality (3.3) is fulfilled if and only if w is a minimizer of the functional $E := H + \psi$ with $H(z) := \frac{1}{2}a(z,z) - \ell(z)$ in W. The functional H is strictly convex, continuous and coercive, cf. [17]. This implies the existence of a unique minimizer w due to $\psi \geq 0$. It is easy to see, that there holds

$$\psi(z) = \sup_{\mu\in\Lambda}(\mu,\chi(z))_0 \tag{3.5}$$

with $\chi(v,q) = \sigma_y q$ and $\Lambda := \{\mu \in Q \mid \mu : \mu \leq 1\}$. Indeed, from Cauchy's inequality we have

$$\mu : q \leq (q:q)^{1/2} \tag{3.6}$$

for all $\mu\in\Lambda$ and all $q \in Q$. Therefore,

$$\sup_{\mu\in\Lambda}(\mu,\chi(z))_0 = \int_\Omega \sigma_y\mu : q\,dx \leq \psi(z).$$

Then again, we obtain

$$\sup_{\mu\in\Lambda}(\mu,\chi(z))_0 \geq (\Xi(q),\chi(z))_0 = \psi(z)$$

with $\Xi(q) := q/(q:q)^{1/2}$, where $q \neq 0$, and zero elsewhere.

Given the Lagrange functional $\mathcal{L}(z,\mu) := H(z) + (\mu,\chi(z))_0$ on $W \times \Lambda$, from (3.5) we have

$$E(w) = \inf_{z\in W}\sup_{\mu\in\Lambda}\mathcal{L}(z,\mu). \tag{3.7}$$

The identity (3.7) implies that w is a minimizer of E, whenever $(w,\lambda) = ((u,p),\lambda) \in W \times \Lambda$ is a saddle point of $\mathcal{L}$. The existence of a saddle point is guaranteed, since Λ is convex, closed and bounded, cf. [13, Rem. IV.2.1 and Prop. IV.2.3]. Due to the stationary condition, $(w,\lambda) \in W \times \Lambda$ is a saddle point of $\mathcal{L}$, if and only if it fulfills the mixed variational formulation

$$\begin{aligned} a(w,z) &= \ell(z) - (\lambda,\chi(z))_0, \\ (\mu-\lambda,\chi(w))_0 &\leq 0 \end{aligned} \tag{3.8}$$

for all $z \in W$ and $\mu \in \Lambda$. Since

$$\sigma_y\|\lambda_0-\lambda_1\|_0 = \sigma_y \sup_{q\in Q,\|q\|_0=1}(\lambda_0-\lambda_1,q)_{0,\Omega} \leq \sup_{z\in W,\|z\|=1}(\lambda_0-\lambda_1,\chi(z))_{0,\Omega} = 0$$

for some saddle points $(w,\lambda_0),(w,\lambda_1) \in W \times \Lambda$, also the uniqueness of the saddle point is guaranteed.

With the identity tensor $\mathbb{I}$, we define the deviatoric part of $\tau \in L^2(\Omega,\mathbb{R}^{k\times k})$ by $\operatorname{dev}(\tau) := \tau - \frac{1}{k}\operatorname{tr}(\tau)\mathbb{I}$ and observe for $q \in Q$

$$(\operatorname{tr}(\tau)\mathbb{I},q)_0 = (\operatorname{tr}(\tau),\operatorname{tr}(q))_0 = 0. \tag{3.9}$$

Proposition 3.2. *There holds*

$$\operatorname{dev}(\sigma(w) - \mathbb{H}p) = \sigma_y\lambda.$$

Proof. Let $z = (v,q) \in W$. From (3.9) we obtain

$$\begin{aligned}(\operatorname{dev}(\sigma(w)-\mathbb{H}p)-\sigma_y\lambda,q)_0 &= (\sigma(w)-\mathbb{H}p-\frac{1}{k}\operatorname{tr}(\sigma(w)-\mathbb{H}p)\mathbb{I},q)_0-(\sigma_y\lambda,q)_0 = (\sigma(w)-\mathbb{H}p,q)_0-(\sigma_y\lambda,q)_0\\ &= (\sigma(w),\varepsilon(v))_0 - a(w,z) - (\sigma_y\lambda,q)_0 = \ell(v) - a(w,z) - (\sigma_y\lambda,q)_0 = 0,\end{aligned}$$

where we use (3.1) in the last line. Since $\operatorname{dev}(\sigma(w)-\mathbb{H}p) - \sigma_y\lambda \in Q$, we obtain the assertion. □

Remark 3.3. The inequality in (3.8) is equivalent to

$$\lambda : p = (p:p)^{1/2}. \tag{3.10}$$

In fact, it follows from Cauchy's inequality (3.6) and $\mu := \Xi(p)$ in (3.8),

$$0 \leq (\lambda - \Xi(p),\chi(w))_0 = \sigma_y\int_\Omega \lambda : p - (p:p)^{1/2}\,dx \leq 0$$

and, hence, $\int_\Omega \lambda : p - (p:p)^{1/2}\,dx = 0$. Because of (3.6) we have $\lambda : p - (p:p)^{1/2} \leq 0$ and, therefore, (3.10) holds. Furthermore, (3.6) and (3.10) yield

$$(\mu-\lambda,\chi(w))_0 = \sigma_y\int_\Omega \mu : p - (p:p)^{1/2}\,dx \leq 0$$

which is the inequality in (3.8). A simple consequence of (3.10) together with Cauchy's inequality (3.6) is that if $p \neq 0$, there holds $\lambda : \lambda = 1$ and, therefore,

$$(\operatorname{dev}(\sigma(w)-\mathbb{H}p) : \operatorname{dev}(\sigma(w)-\mathbb{H}p))^{1/2} = \sigma_y. \tag{3.11}$$

Remark 3.4. In this work we consider elastoplasticity with linear kinematic hardening for a single yield surface. In the case of multiple yield surfaces we have to add further plastic variables $\mathbb{H}_i$, $\sigma_{y,i}$, and p_i with $i = 0, \ldots, r$, where $r+1$ is the number of surfaces. Moreover, we have to replace Q by $\tilde{Q} := (Q)^r$ and W by $\tilde{W} := H^1_D(\Omega) \times \tilde{Q}$. The bilinear form then becomes

$$\tilde{a}(\tilde{w}, \tilde{z}) := (\mathbb{C}(\varepsilon(u) - \sum_{i=0}^{r} p_i), \varepsilon(v) - \sum_{i=0}^{r} q_i)_0 + \sum_{i=0}^{r} (\mathbb{H}_i p_i, q_i)_0$$

for $\tilde{w} = (u, p_0, \ldots, p_r), \tilde{z} = (v, q_0, \ldots, q_r) \in W$. The nonlinear functional $\tilde{\psi}$ instead of ψ reads

$$\tilde{\psi}(\tilde{z}) := \sum_{i=0}^{r} \int_\Omega j_i(q_i)\, dx$$

with $j_i(q_i) := \sigma_{y,i}|q_i|$. The saddle point formulation of multi-yield plasticity requires Lagrange multipliers for each additional plastic strain variable. Its mixed formulation is then given by

$$\tilde{a}(\tilde{w}, \tilde{z}) = \ell(\tilde{z}) - \sum_{i=0}^{r} (\lambda_i, \chi_i(\tilde{z}))_0,$$

$$\sum_{i=0}^{r} (\mu_i - \lambda_i, \chi_i(\tilde{z}))_0 \leq 0,$$

where $\chi_i(\tilde{z}) := \sigma_{y,i} q_i$. For more details on the formulation of multi-yield plasticity, we refer to [8].

As we can see, modeling with multi-yield surfaces does not cause additional mathematical complexity. Therefore, the following results only deal with single yield surfaces and can easily be transferred to the case of multi-yield surfaces.

4. Discretization of the mixed variational formulations

Let V_h be a finite dimensional subspace of $H^1_D(\Omega)$ and Q_h be a finite dimensional subspace of Q. The discrete saddle point problem of the primal problem of elastoplasticity with linear kinematic hardening is to find a discrete saddle point $(w_h, \lambda_h) \in W_h \times \Lambda_h$ with $W_h := V_h \times Q_h$ and $\Lambda_h := \Lambda \cap Q_h$ such that

$$\mathcal{L}(w_h, \lambda_h) = \inf_{z_h \in W_h} \sup_{\mu_h \in \Lambda_h} \mathcal{L}(z_h, \mu_h). \tag{4.1}$$

By the stationary condition, we conclude that the discrete saddle point is equivalently characterized by

$$\begin{aligned} &\forall z_h \in W_h : a(w_h, z_h) = \ell(z_h) - (\lambda_h, \chi(z_h))_0, \\ &\forall \mu_h \in \Lambda_h : (\mu_h - \lambda_h, \chi(w_h))_0 \leq 0. \end{aligned} \tag{4.2}$$

Again, the existence of a discrete saddle point is guaranteed, since Λ_h is convex, closed and bounded. The first component is the unique minimizer of the minimization problem

$$E(w_h) = \min_{z_h \in W_h} H(z_h) + \psi(z_h).$$

As in Section 3, we conclude from $\Lambda_h \subset Q_h$ that also the second component is unique. Note that

$$(\lambda_h, \chi(w_h))_0 = \sup_{\mu_h \in \Lambda_h} (\mu_h, \chi(w_h))_0 = \psi(w_h) = \sup_{\mu \in \Lambda} (\mu, \chi(w_h))_0 \geq (\lambda, \chi(w_h))_0 \tag{4.3}$$

for all $w_h \in W_h$.

Proposition 4.1. *There holds*

$$(\mathrm{dev}(\sigma(w_h) - \mathbb{H} p_h) - \sigma_y \lambda_h, q_h)_0 = 0 \tag{4.4}$$

for all $q_h \in Q_h$. *If* $\mathrm{dev}(\sigma(W_h)) \subset Q_h$, *there holds*

$$\mathrm{dev}(\sigma(w_h) - \mathbb{H} p_h) = \sigma_y \lambda_h. \tag{4.5}$$

Proof. The assertions hold by the same arguments as in Proposition 3.2. □

Remark 4.2. Condition (4.4) implies that $\sigma_y \lambda_h$ is the orthogonal L^2-projection of $\operatorname{dev}(\sigma(w_h) - \mathbb{H} p_h)$ in Q_h.

Remark 4.3. By the same arguments as in Remark 3.3, the inequality in (4.2) is equivalent to $\lambda_h : p_h = (p_h : p_h)^{1/2}$ and there holds $\lambda_h : \lambda_h = 1$, if $p_h \neq 0$. The relation (3.11) holds for the discrete variables, if $p_h \neq 0$ and $\operatorname{dev}(\sigma(W_h)) \subset Q_h$.

Remark 4.4. As mentioned, the uniqueness of the discrete Lagrange multiplier is guaranteed through the canonical assumption $\Lambda_h \subset Q_h$. Choosing the set Λ_h as a subset of a further space $\tilde{Q}_h \neq Q_h$ such that $\Lambda_h \not\subset Q_h$, one has to ensure, that W_h and $\tilde{Q}_h$ satisfy a discrete inf-sup condition,

$$\alpha \|\tilde{\mu}_h\|_0 \leq \sup_{w_h \in W_h} (\tilde{\mu}_h, \chi(w_h))_0$$

for all $\tilde{\mu}_h \in \tilde{Q}_h$ and a constant $\alpha > 0$.

The convergence of the mixed method can be stated without any regularity assumptions using some standard techniques of convex analysis. Only, the coercivity of a and the approximation properties of W_h and Q_h are used. Here, we present a modification of Theorem 1.1.5.3 in [18].

Theorem 4.5. *Assume that for all $(z, \mu) \in W \times \Lambda$ there exists a sequence $\{(z_h, \mu_h)\}$ with $(z_h, \mu_h) \in W_h \times \Lambda_h$ and $(z_h, \mu_h) \to (z, \mu)$ as $h \to 0$. Then, the sequence $\{w_h\}$ strongly converges to w and the sequence of Lagrange multipliers $\{\lambda_h\}$ weakly converges to λ as $h \to 0$.*

Proof. From (4.2), we obtain $\nu_1 \|w_h\| \leq \|\ell\| + \|\lambda_h\|_0$, so that the boundedness of $\{\lambda_h\}$ implies that $\{w_h\}$ is also bounded. Due to the reflexivity of $W \times Q$ there exists a subsequence $\{(w_{\tilde{h}}, \lambda_{\tilde{h}})\} \subset \{(w_h, \lambda_h)\}$ which weakly converges to $(w^*, \lambda^*) \in W \times Q$. Since Λ is convex and closed and, therefore, weakly closed, we have $(w^*, \lambda^*) \in W \times \Lambda$. It is easy to see that $\lim_{\tilde{h}\to 0} a(w_{\tilde{h}}, z_{\tilde{h}}) = a(w^*, z)$ and $\lim_{\tilde{h}\to 0} (\mu_{\tilde{h}}, \chi(w_{\tilde{h}}))_0 = (\mu, \chi(w^*))$. Passing to the limit in (4.2) yields

$$a(w^*, z) = \ell(z) - (\lambda^*, \chi(z))_0, \tag{4.6}$$

$$(\mu, \chi(w^*))_0 \leq \liminf_{\tilde{h}\to 0} (\lambda_{\tilde{h}}, \chi(w_{\tilde{h}}))_0. \tag{4.7}$$

Since $z \mapsto a(z, z)$ is convex and continuous and, therefore, weakly lower semicontinuous, we obtain

$$a(w^*, w^*) + \liminf_{\tilde{h}\to 0} (\lambda_{\tilde{h}}, \chi(w_{\tilde{h}}))_0 \leq \liminf_{\tilde{h}\to 0} \left(a(w_{\tilde{h}}, w_{\tilde{h}}) + (\lambda_{\tilde{h}}, \chi(w_{\tilde{h}}))_0 \right) = \liminf_{\tilde{h}\to 0} \ell(w_{\tilde{h}}) = \ell(w^*)$$

from (4.2). Hence, using (4.6) with $z := w^*$ and (4.7), we find

$$(\mu, \chi(w^*))_0 \leq \liminf_{\tilde{h}\to 0} (\lambda_{\tilde{h}}, \chi(w_{\tilde{h}}))_0 \leq \ell(w^*) - a(w^*, w^*) = (\lambda^*, \chi(w^*))_0. \tag{4.8}$$

Since (z, μ) is arbitrarily chosen, (4.6) and (4.8) imply that (w^*, λ^*) is a saddle point. Due to the uniqueness, we conclude $(w^*, \lambda^*) = (w, \lambda)$ and, additionally, that the entire sequence $\{(w_h, \lambda_h)\}$ converges to (w, λ) weakly. To show that $\{w_h\}$ converges to w strongly, we conclude from (4.3)

$$\begin{aligned} a(w - w_h, w - w_h) &= a(w, w) - 2a(w, w_h) + \ell(w_h) - (\lambda_h, \chi(w_h))_0 \\ &\leq a(w, w) - 2a(w, w_h) + \ell(w_h) - (\lambda, \chi(w_h))_0 \to 0 \end{aligned}$$

as $h \to 0$. □

5. A priori error estimates

In the following, $A \lesssim B$ abbreviates $A \leq CB$ with a positive constant C which is independent of V_h and Q_h. Furthermore, $A \approx B$ represents $A \lesssim B \lesssim A$.

Lemma 5.1. *There holds*

$$\|\lambda - \lambda_h\|_0 \lesssim \|w - w_h\| + \|\lambda - \mu_h\|_0$$

for all $\mu_h \in \Lambda_h$.

Proof. From (3.4) we obtain

$$\begin{aligned}
\sigma_y\|\mu_h-\lambda_h\|_0 &= \sup_{q_h\in Q_h,\ \|q_h\|_0=1}\sigma_y(\mu_h-\lambda_h,q_h)_0 \le \sup_{z_h\in W_h,\ \|z_h\|=1}(\mu_h-\lambda_h,\chi(z_h))_0\\
&= \sup_{z_h\in W_h,\ \|z_h\|=1}((\mu_h,\chi(z_h))_0 + a(w_h,z_h)-\ell(z_h))\\
&= \sup_{z_h\in W_h,\ \|z_h\|=1}((\mu_h-\lambda,\chi(z_h))_0 + a(w_h-w,z_h)) \le \sigma_y\|\lambda-\mu_h\|_0+\nu_0\|w-w_h\|.
\end{aligned}$$

The assertion follows from $\|\lambda-\lambda_h\|_0\le\|\lambda-\mu_h\|_0+\|\mu_h-\lambda_h\|_0$. □

Theorem 5.2. *There holds*

$$\|w-w_h\|+\|\lambda-\lambda_h\|_0 \lesssim \|w-z_h\|+\|\lambda-\mu_h\|_0+(\lambda-\mu_h,\chi(w))_0^{1/2} \tag{5.1}$$

for all $z_h\in W_h$ *and all* $\mu_h\in\Lambda_h$.

Proof. Since $\Lambda_h\subset\Lambda$, we obtain

$$(\lambda_h-\lambda,\chi(w-w_h))_0\le(\lambda-\lambda_h,\chi(w_h))_0\le(\lambda-\mu_h,\chi(w_h))_0\le(\lambda-\mu_h,\chi(w))_0+(\lambda-\mu_h,\chi(w_h-w))_0.$$

Consequently, we obtain from (3.4) and Lemma 5.1

$$\begin{aligned}
\|w-w_h\|^2 &\lesssim a(w-w_h,w-z_h)+a(w-w_h,z_h-w_h)\\
&= a(w-w_h,w-z_h)+(\lambda_h-\lambda,\chi(z_h-w))_0+(\lambda_h-\lambda,\chi(w-w_h))_0\\
&\lesssim \|w-w_h\|\|w-z_h\|+\|\lambda-\lambda_h\|_0\|w-z_h\|+\|\lambda-\mu_h\|_0\|w-w_h\|+(\lambda-\mu_h,\chi(w))_0\\
&\lesssim \|w-w_h\|(\|w-z_h\|+\|\lambda-\mu_h\|_0)+(\lambda-\mu_h,\chi(u))_0.
\end{aligned}$$

Since $x^2\le ax+b$ implies $x\le a+b^{1/2}$ for $x,a,b\in\mathbb{R}_+$, we have

$$\|w-w_h\|\lesssim\|w-z_h\|+\|\lambda-\mu_h\|_0+(\lambda-\mu_h,\chi(w))_0^{1/2}.$$

Again with Lemma 5.1, we conclude $\|w-w_h\|+\|\lambda-\lambda_h\|_0\lesssim\|w-w_h\|+\|\lambda-\mu_h\|_0$, which completes the proof. □

Remark 5.3. Given the assumptions in Theorem 4.5, Theorem 5.2 yields strong convergence of both w_h and λ_h, whereas Theorem 4.5 only yields weak convergence of λ_h.

6. A posteriori error estimates

Let the residual Res : $W\to V^*$ be defined as $\langle\mathrm{Res}(w),v\rangle := (f,v)_0+(g,\gamma_{\Gamma_N}(v))_{0,\Gamma_N}-(\sigma(w),\varepsilon(v))_0$ for $w\in W$ and $v\in V$, where V^* is the topological dual space of V.

Lemma 6.1. *There holds*

$$\|w-w_h\|\lesssim\|\mathrm{Res}(w_h)\|+\|\operatorname{dev}(\sigma(w_h)-\mathbb{H}p_h)-\sigma_y\lambda_h\|_0.$$

Proof. With $w-w_h=(u-u_h,p-p_h)$, we obtain from (3.4)

$$\begin{aligned}
\|w-w_h\|^2 &\lesssim a(w-w_h,w-w_h)=\ell(w-w_h)-(\lambda,\chi(w-w_h))_0-a(w_h,w-w_h)\\
&=\ell(w-w_h)+(\lambda_h-\lambda,\chi(w))_0+(\lambda-\lambda_h,\chi(w_h))_0-(\lambda_h,\chi(w-w_h))_0-a(w_h,w-w_h)\\
&\le \ell(w-w_h)-(\lambda_h,\chi(w-w_h))_0-a(w_h,w-w_h)\\
&=\ell(w-w_h)-(\sigma(w_h),\varepsilon(u-u_h))_0+(\sigma(w_h)-\mathbb{H}p_h-\sigma_y\lambda_h,p-p_h)_0\\
&=\ell(w-w_h)-(\sigma(w_h),\varepsilon(u-u_h))_0+(\operatorname{dev}(\sigma(w_h)-\mathbb{H}p_h)-\sigma_y\lambda_h,p-p_h)_0\\
&\le(\|\mathrm{Res}(w_h)\|+\|\operatorname{dev}(\sigma(w_h)-\mathbb{H}p_h)-\sigma_y\lambda_h\|_0)\|w-w_h\|,
\end{aligned}$$

where we use $(\lambda-\lambda_h,\chi(w_h))\le 0$ as shown in (4.3). □

To include the error $\|\lambda - \lambda_h\|_0$ in the a posteriori estimation, we consider the auxiliary problem: Find $w_\star \in W$ so that

$$a(w_\star, z) = \ell(z) - (\lambda_h, \chi(z))_0 \tag{6.1}$$

for all $z \in W$. Obviously, the solution $w_\star$ of (6.1) exists and is unique.

Lemma 6.2. *There holds*

$$\|\lambda - \lambda_h\|_0 \lesssim \|w - w_\star\|.$$

Proof. The assertion follows from

$$\sigma_y \|\lambda - \lambda_h\|_{0,\Omega} = \sigma_y \sup_{q\in Q, \|q\|_0=1} (\lambda - \lambda_h, q)_{0,\Omega} \leq \sup_{z\in W, \|z\|_0=1} (\lambda - \lambda_h, \chi(z))_{0,\Omega} = \sup_{z\in W, \|z\|=1} a(w_\star - w, z) \leq \nu_0 \|w - w_\star\|.$$

□

Lemma 6.3. *There holds*

$$\|\operatorname{dev}(\sigma(w_h) - \mathbb{H}p_h) - \sigma_y\lambda_h\|_0 \lesssim \|w - w_h\| + \|\lambda - \lambda_h\|_0.$$

Proof. From Proposition 3.2, we have

$$\begin{aligned}\|\operatorname{dev}(\sigma(w_h) - \mathbb{H}p_h) - \sigma_y\lambda_h\|_0 &\lesssim \|\operatorname{dev}(\sigma(w_h) - \mathbb{H}p_h) - \sigma_y\lambda\|_0 + \|\lambda - \lambda_h\|_0 \\ &= \|\operatorname{dev}(\sigma(w_h - w) - \mathbb{H}(p_h - p))\|_0 + \|\lambda - \lambda_h\|_0 \lesssim \|w - w_h\| + \|\lambda - \lambda_h\|_0.\end{aligned}$$

□

Theorem 6.4. *There holds*

$$\|w - w_h\| + \|\lambda - \lambda_h\|_0 \approx \|\operatorname{Res}(w_h)\| + \|\operatorname{dev}(\sigma(w_h) - \mathbb{H}p_h) - \sigma_y\lambda_h\|_0. \tag{6.2}$$

Proof. With $w_\star - w_h = (u_\star - u_h, p_\star - p_h)$, we obtain from (3.4)

$$\begin{aligned}\|w_\star - w_h\|^2 &\lesssim a(w_\star - w_h, w_\star - w_h) = \ell(w_\star - w_h) - (\lambda_h, \chi(w_\star - w_h))_0 - a(w_h, w_\star - w_h) \\ &= \ell(w_\star - w_h) - (\sigma(w_h), \varepsilon(u_\star - u_h))_0 + (\sigma(w_h) - \mathbb{H}p_h - \sigma_y\lambda_h, p_\star - p_h)_0 \\ &= \ell(w_\star - w_h) - (\sigma(w_h), \varepsilon(u_\star - u_h))_0 + (\operatorname{dev}(\sigma(w_h) - \mathbb{H}p_h) - \sigma_y\lambda_h, p_\star - p_h)_0 \\ &\leq (\|\operatorname{Res}(w_h)\| + \|\operatorname{dev}(\sigma(w_h) - \mathbb{H}p_h) - \sigma_y\lambda_h\|_0)\|w_\star - w_h\|\end{aligned}$$

Together with the triangle inequality, Lemma 6.1 and Lemma 6.2 yield one of the estimates in (6.2). The other estimate follows from the definition of the residual, relations (3.1) and (3.4) as well as Lemma 6.3. □

Remark 6.5. If $\operatorname{dev}(\sigma(W_h))$ is a subset of Q_h, then $\|\lambda - \lambda_h\|_0 \approx \|w - w_h\|$ and the term $\|\operatorname{dev}(\sigma(w_h) - \mathbb{H}p_h) - \sigma_y\lambda_h\|_0$ in (6.2) vanishes, i.e. $\|w - w_h\| \approx \|\operatorname{Res}(w_h)\|$.

7. Low-order finite elements

In this section, we propose three low-order finite element discretizations based on triangles and tetrahedrons as well as quadrilaterals and hexahedrons. We use continuous piecewise linear, bilinear or trilinear functions, respectively, to define the discretization space V_h. The main differences of the three approaches are in the definition of the space Q_h. The simplest discretization space consists of piecewise constant functions. Since $\operatorname{dev}(\sigma(W_h)) \subset Q_h$ does not hold on quadrilaterals and hexahedrons in this case, the additional term $\|\operatorname{dev}(\sigma(w_h) - \mathbb{H}p_h) - \sigma_y\lambda_h\|_0$ in (6.2) has to be taken into account in the a posteriori analysis. To avoid this term, we also introduce a second discretization approach for Q_h using piecewise, discontinuous bilinear and trilinear functions on quadrilaterals and hexahedrons, respectively, and still piecewise constant functions on triangles and tetrahedrons. The third approach uses piecewise linear, bilinear or trilinear, but continuous functions to define Q_h, which may lead to a reduction of the degrees of freedom. In the latter case the additional term in (6.2) has to be considered.

Let $\{\mathcal{T}_h\}$ be a shape regular family of finite element meshes of Ω with mesh size h. We define

$$\mathcal{P}_h := \{p : \Omega \to \mathbb{R} \mid p_{|T} \in \mathcal{P}(T),\ T \in \mathcal{T}_h\},$$

where

$$\mathcal{P}(T) := \begin{cases} \mathcal{P}_1(T), & T \text{ is a triangle or a tetrahedron,} \\ Q_1(T), & T \text{ is a quadrilateral or a hexahedron.} \end{cases}$$

Here, $\mathcal{P}_1(T)$ denotes the space of linear polynomials on $T \in \mathcal{T}_h$ and $Q_1(T)$ is the space of bilinear or trilinear polynomials on T, respectively. Furthermore, C_h denotes the space of piecewise constant functions on $\mathcal{T}_h$. We define finite element discretizations of low-order by

$$\begin{aligned}
&\hat{V}_h := (\mathcal{P}_h)^k \cap H^1_D(\Omega), \\
&\overline{Q}_h := \{q_h \in Q \mid\ q_{h,ij} \in C_h\}, \quad \hat{Q}_h := \{q_h \in Q \mid\ q_{h,ij} \in \mathcal{P}_h\}, \quad \hat{Q}_h^\circ := \hat{Q}_h \cap H^1(\Omega; \mathbb{R}^{3\times 3}) \\
&\overline{W}_h := \hat{V}_h \times \overline{Q}_h, \quad \hat{W}_h := \hat{V}_h \times \hat{Q}_h, \quad \hat{W}_h^\circ := \hat{V}_h \times \hat{Q}_h^\circ \\
&\overline{\Lambda}_h := \Lambda \cap \overline{Q}_h, \quad \hat{\Lambda}_h := \Lambda \cap \hat{Q}_h, \quad \hat{\Lambda}_h^\circ := \Lambda \cap \hat{Q}_h^\circ.
\end{aligned}$$

The following Proposition summarizes the main approximation and interpolation results on the spaces $\hat{V}_h$, $\overline{Q}_h$, $\hat{Q}_h$ and $\hat{Q}_h^\circ$. For this purpose, let $v \in H^1(\Omega; \mathbb{R}^k)$ with $v_i \in H^{1+\theta}(\Omega)$, $q \in Q$ with $q_{ij} \in H^{\theta'}(\Omega)$, and $\theta, \theta' > 0$. Moreover, we assume quadrilaterals to be parallelograms and hexahedron to be parallelepipeds. Using the L^2-projection $\tilde{\Pi}_h$ onto the space C_h, we define $(\Pi_h(q))_{ij} := \tilde{\Pi}_h(q_{ij})$. Furthermore, let $\{\varphi_{T,m}\}_{m=1}^{n_T} \subset \mathcal{P}(T)$ denote the usual nodal basis on T, i.e., with Kronecker's delta δ there holds $\varphi_{T,m}(x_{T,l}) = \delta_{ml}$ for the vertices $x_{T,l}$ of T, $l = 1, \ldots, n_T$. We have $\sum_{m=1}^{n_T} \varphi_{T,m} = 1$ and $0 \le \varphi_{T,m} \le 1$. In particular, there holds

$$1 = \left(\sum_{m=1}^{n_T} \varphi_{T,m} \right)^2 = \sum_{l,m=1}^{n_T} \varphi_{T,l} \varphi_{T,m}. \tag{7.1}$$

Clement's interpolant $\tilde{J}_h$ is defined as

$$\tilde{J}_h(\tilde{v}_{|T})_{|T} := \sum_{l=1}^{n_T} |\omega_{x_{T,l}}|^{-1} \int_{\omega_{x_{T,l}}} \tilde{v}\, dx\ \varphi_{T,l}$$

for $\tilde{v} \in H^{\tilde{\theta}}(\Omega)$ with $\tilde{\theta} \ge 0$ and $\omega_x := \{T \in \mathcal{T}_h \mid x \in \overline{T}\}$, cf. [2]. Thus, we set $J_h(q))_{ij} := \tilde{J}_h(q_{ij})$.

Proposition 7.1. *There holds*

(i) $\Pi_h(\Lambda) \subset \overline{\Lambda}_h \subset \hat{\Lambda}_h$ *and* $J_h(\Lambda) \subset \hat{\Lambda}_h^\circ$,
(ii) $\|q - \Pi_h(q)\|_0 \le C_\Pi(q) h^{\min\{\theta',1\}}$ *with a constant* $C_\Pi(q) > 0$,
(iii) $\|q - J_h(q)\|_0 \le C_J(q) h^{\min\{\theta',2\}}$ *with a constant* $C_J(q) > 0$,

Proof. Proofs of the assertions (ii) and (iii) can be found in, e.g., [2, 22]. To show (i), let $\mu \in \Lambda$ and consider $\Pi_h(\mu)$ and $J_h(\mu)$ on a $T \in \mathcal{T}_h$. Using Cauchy's Inequality, we observe

$$\sum_{i,j=1}^{k} |\omega|^{-2} \left(\int_\omega \mu_{ij}\, dx \right)^2 \le |\omega|^{-1} \int_\omega \mu : \mu\, dx \le 1 \tag{7.2}$$

for $\omega \subset \Omega$. Thus, we obtain

$$\Pi_h(\mu) : \Pi_h(\mu) = \sum_{i,j=1}^{k} |T|^{-2} \left(\int_T \mu_{ij}\, dx \right)^2 \le 1$$

which gives us $\Pi_h(\Lambda) \subset \overline{\Lambda}_h$. To prove $J_h(\Lambda) \subset \hat{\Lambda}_h^\circ$, we derive from Young's inequality and (7.2)

$$\sum_{i,j=1}^{k} |\omega_{z_l}|^{-1} \left(\int_{\omega_{z_l}} \mu_{ij}\, dx \right) |\omega_{z_m}|^{-1} \left(\int_{\omega_{z_m}} \mu_{ij}\, dx \right) \le \frac{1}{2} \left(\sum_{i,j=1}^{k} |\omega_{z_l}|^{-2} \left(\int_{\omega_{z_l}} \mu_{ij}\, dx \right)^2 + \sum_{i,j=1}^{k} |\omega_{z_m}|^{-2} \left(\int_{\omega_{z_m}} \mu_{ij}\, dx \right)^2 \right) \le 1.$$

From (7.1), we obtain

$$J_h(\mu) : J_h(\mu) = \sum_{i,j=1}^{k} \sum_{l,m=1}^{n_T} |\omega_{z_l}|^{-1} \left(\int_{\omega_{z_l}} \mu_{ij}\, dx \right) \varphi_{T,l}\, |\omega_{z_m}|^{-1} \left(\int_{\omega_{z_m}} \mu_{ij}\, dx \right) \varphi_{T,m} \leq \sum_{l,m=1}^{n_T} \varphi_{T,l} \varphi_{T,m} = 1.$$

□

Remark 7.2. The definition of $\mathcal{P}_h$ allows for the use of hybrid meshes consisting of triangles and quadrilaterals. In three dimensions, the use of hybrid meshes is more complicated due to the continuity assumptions on $\hat{V}_h$ and involves further shapes as pyramids and prism. Therefore, we assume non-hybrid meshes in three dimensions.

Remark 7.3. Instead of piecewise bilinear functions, we can also apply piecewise linear functions on parallelograms or parallelepipeds, respectively, to define a space Q_h with $\operatorname{dev}(\sigma(W_h)) \subset Q_h$. In this case, we obtain the same a priori results as using $\hat{Q}_h$, cf. Section 8.

8. A priori estimates for low-order finite elements

In the following, let $u_i \in H^{1+\theta}(\Omega)$ and $p_{ij} \in H^{\theta'}(\Omega)$ with $\theta, \theta' > 0$. From Proposition 3.2, we observe that $\lambda_{ij} \in H^{\bar{\theta}}(\Omega)$ with $\bar{\theta} := \min\{\theta, \theta'\}$. Choosing $Q_h := \bar{Q}_h$ or $Q_h := \hat{Q}_h$, we can make use of the orthogonality relation of the L^2-projection Π_H.

Corollary 8.1. *Let $W_h := \bar{W}_h$ or $W_h := \hat{W}_h$. Hence, there holds*

$$\|w - w_h\| + \|\lambda - \lambda_h\|_0 \lesssim h^{\min\{1,\bar{\theta}\}}.$$

Proof. Form the orthogonality relation of the L^2-projection and Proposition 7.1, we obtain

$$\begin{aligned}(\lambda - \Pi_h(\lambda), \chi(w))_0 &= \sigma_y(\lambda - \Pi_h(\lambda), p - \Pi_h(p))_0 \\ &\leq \sigma_y \|\lambda - \Pi_h(\lambda)\|_0 \|p - \Pi_h(p)\|_0 \leq \sigma_y C_\Pi(\lambda) C_\Pi(p) h^{\min\{1,\bar{\theta}\}+\min\{1,\theta'\}}.\end{aligned}$$

This gives us the assertion using Theorem 5.2. □

For $Q_h := \hat{Q}_h^\circ$, we obtain the following result by utilizing Cauchy's inequality:

Corollary 8.2. *Let $W_h := \hat{W}_h^\circ$. Hence, there holds*

$$\|w - w_h\| + \|\lambda - \lambda_h\|_0 \lesssim h^{\min\{1,\bar{\theta}/2\}}.$$

Proof. Cauchy's inequality and Proposition 7.1 yield

$$(\lambda - J_h(\lambda), \chi(w))_0 \leq \sigma_y \|q\|_0 \|\lambda - J_h(\lambda)\|_0 \leq \sigma_y \|q\|_0 C_J(\lambda) h^{\min\{2,\bar{\theta}\}}.$$

The assertion follows by Theorem 5.2. □

Remark 8.3. For triangle meshes Corollary 8.1 yields the same optimal convergence result $O(h)$ as shown in [1]. The Corollary 8.2 implies that the use of globally continuous elements for Q_h leads to an optimal convergence rate $O(h)$ if $\bar{\theta} \geq 2$ holds, which, however, requires $u_i \in H^3(\Omega)$ and $p_{ij} \in H^2(\Omega)$.

Remark 8.4. Corollaries 8.1 and 8.2 are dissatisfying in the sense that the regularity given by θ and θ' is unclear in general. We refer to [19, 20] for some results on the regularity in elastoplasticity.

9. A posteriori estimates for low-order finite elements

As stated in Theorem 6.4, we have to estimate $\|\operatorname{Res}(w_h)\|$ which can be done by standard techniques known from the context of linear elasticity. Here, we present the standard residual approach. For this purpose, we define the residuals

$$R_T := \|f_T + \operatorname{div}\sigma(w_h)\|_{0,T}, \quad R_E := \begin{cases} \|[\sigma_n(w_h)]\|_{0,E}, & E \in \mathcal{E}_h^\circ, \\ \|g_E - \sigma_n(w_h)\|_{0,E}, & E \in \mathcal{E}_h^N, \end{cases}$$

where

$$f_T := |T|^{-1} \int_T f\,dx, \quad g_E := |E|^{-1} \int_E g\,ds,$$

$\mathcal{E}_h^\circ$ is the set of all interior edges of $\mathcal{T}$, and $\mathcal{E}_h^N$ the set of edges E with $E \subset \Gamma_N$. Here, we denote the maximum diameter of a mesh element $T \in \mathcal{T}_h$ by h_T and the length of an edge $E \in \mathcal{E}_h^\circ \cup \mathcal{E}_h^N$ by h_E. Moreover, we define the residual-based error estimator by

$$\eta_h := \left(\sum_{T \in \mathcal{T}} h_T^2 R_T^2 + \sum_{E \in \mathcal{E}^0 \cup \mathcal{E}} h_E R_E^2 \right)^{1/2}$$

and the oscillations by

$$\operatorname{osc}(f, \mathcal{T}_h) := \left(\sum_{T \in \mathcal{T}_h} h_T^2 \|f - f_T\|_{0,T}^2 \right)^{1/2}, \quad \operatorname{osc}(g, \mathcal{E}_h^N) := \left(\sum_{T \in \mathcal{T}_h} h_E \|g - g_E\|_{0,T}^2 \right)^{1/2}.$$

Theorem 9.1. *There holds*

$$\|\operatorname{Res}(w_h)\| \lesssim \eta_h + \operatorname{osc}(f, \mathcal{T}_h) + \operatorname{osc}(g, \mathcal{E}_h^N).$$

Proof. Using standard arguments, cf. [4], and Clement's interpolation defining $J_h(v)_i := \tilde{J}_h(v_i)$ for $v \in H_D^1(\Omega)$, we obtain

$$\langle \operatorname{Res}(w_h), v - J_h(v) \rangle \lesssim \|v\|_1 (\eta_h + \operatorname{osc}(f, \mathcal{T}_h) + \operatorname{osc}(g, \mathcal{E}_h^N)).$$

From (4.2) with $z_h := (v_h, 0)$, we conclude

$$\langle \operatorname{Res}(w_h), v_h \rangle = (f, v_h)_0 + (g, \gamma_{|\Gamma_N}(v_h))_{0,\Gamma_N} - (\sigma(w_h), \varepsilon(v_h))_0 = 0$$

for all $v_h \in V_h$. Thus, we have $\langle \operatorname{Res}(w_h), v \rangle = \langle \operatorname{Res}(w_h), v - J_h(v) \rangle$. The definition of $\|\operatorname{Res}(w_h)\|$ yields the assertion. □

We set $\overline{\eta}_h := \eta_h + \|\operatorname{dev}(\sigma(w_h) - \mathbb{H}p_h) - \sigma_y \lambda_h\|_0$ and use Theorem 6.4 and Theorem 9.1 to conclude

$$\|w - w_h\| + \|\lambda - \lambda_h\|_0 \lesssim \overline{\eta}_h + \operatorname{osc}(f, \mathcal{T}_h) + \operatorname{osc}(g, \mathcal{E}_h^N)$$

which implies that $\overline{\eta}_h$ is a reliable error estimator (except for oscillations). Using $W_h = \overline{W}_h$ or the discretization approach described in Remark 7.3, we are able to follow the arguments proposed, for instance, in [24], to show the efficiency of $\overline{\eta}_h$ where we obtain

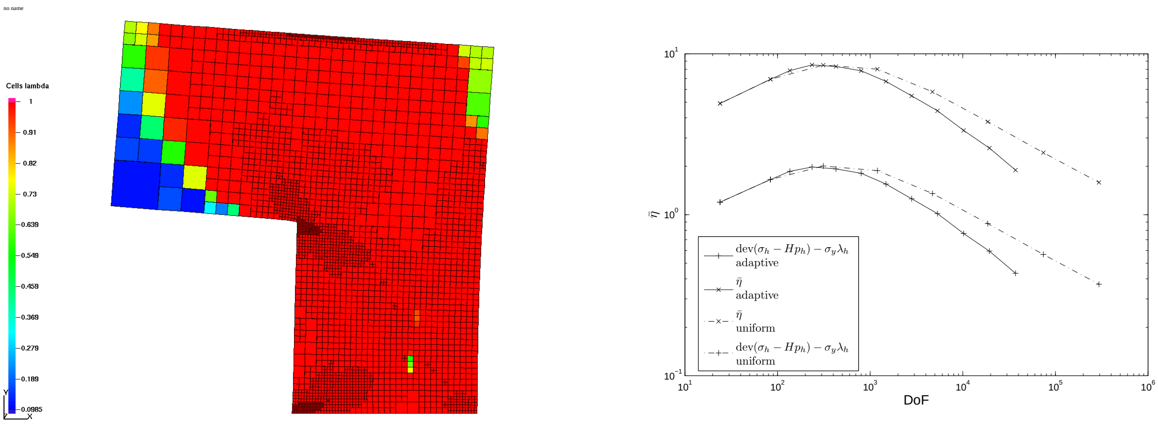
$$\eta_h \lesssim \|w - w_h\| + \operatorname{osc}(f, \mathcal{T}_h) + \operatorname{osc}(g, \mathcal{E}_h^N),$$

in particular, exploiting equation (3.1) and the fact that $f_T + \operatorname{div}\sigma(w_h)$ is constant. From Lemma 6.3, we obtain

$$\overline{\eta}_h \lesssim \|w - w_h\| + \|\lambda - \lambda_h\|_0 + \operatorname{osc}(f, \mathcal{T}_h) + \operatorname{osc}(g, \mathcal{E}_h^N)$$

which is the efficiency of $\overline{\eta}_h$. Unfortunately, we can not argue in the same way if $W_h = \hat{W}_h$ or $W_h = \hat{W}_h^\circ$ since $f_T + \operatorname{div}\sigma(w_h)$ is not constant in these cases.

Remark 9.2. In the same way, techniques for a posteriori error control for meshes with hanging nodes can be transferred from linear elasticity, cf. [6].



(a) (b)

Figure 1: **(a)** Adaptive refinements, the colors indicate $(\lambda_h : \lambda_h)^{1/2}$. **(b)** Estimated convergence rates for adaptive and uniform refinements.

Remark 9.3. Applying Theorem 6.4 with $W_h = \overline{W}_h$ on triangle meshes, we obtain the same well-known error estimation as in [10]. The estimations seem to be new for meshes with quadrilaterals or hexahedrons where we have to determine

$$\|\operatorname{dev}(\sigma(w_h) - \mathbb{H}p_h) - \sigma_y\lambda_h\|^2 = \sum_{T\in Q_h} \|\operatorname{dev}(\sigma(w_h) - \mathbb{H}p_h) - \sigma_y\lambda_h\|_{0,T}^2$$

with the set of all quadrilaterals or hexahedrons $Q_h \subset \mathcal{T}_h$. This term measures the (local) error between $\operatorname{dev}(\sigma(w_h) - \mathbb{H}p_h)$ and its L^2-projection. Under the regularity assumption of Section 8 we obtain from Corollary 8.1 and Lemma 6.3 that it behaves like $O(h^{\min\{1,\bar\theta\}})$.

10. Numerical results

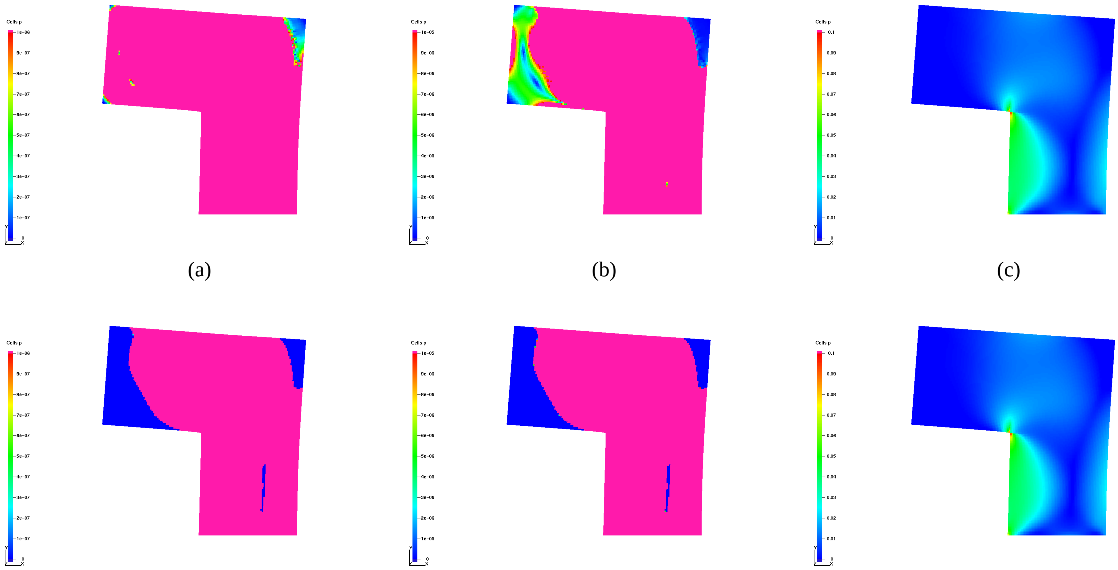
In this section, we study some numerical experiments to show the applicability of the derived estimates within adaptive schemes and discuss some numerical properties of the mixed method. We consider problems of elasto-plasticity with linear kinematic hardening and either one or two yield surfaces. The elasticity tensor is defined as $\mathbb{C}\tau := \lambda\,\mathrm{tr}(\tau)\mathbb{I} + 2\mu\tau$ with the first and second Lamé parameters λ and μ, respectively. Moreover, we set the hardening tensor $\mathbb{H} := \xi\mathbb{I}$ with a positive real number ξ. We restrict ourselves to the two-dimensional case where we discretize with piecewise bilinear and piecewise constant functions on a quadrilateral mesh, i.e. we use the finite element spaces $V_h := \hat{V}_h$ and $Q_h := \overline{Q}_h$. This discretization approach is easy to implement and seems to be new in the context of elastoplasticity, see also Remark 9.3.

We solve the discrete mixed variational formulation by Uzawa's method, cf. [14, 15]. For this purpose, we introduce the standard nodal basis $\{\varphi_i\}_{0\le i<n}$ of $\hat{V}_h$ and the basis $\{\psi_j\}_{0\le j<m}$ of $\overline{Q}_h$ which is given by

$$\psi_{r(T,1)} := \chi_T \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \psi_{r(T,2)} := \chi_T \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

with $n := \dim \hat{V}_h$ and $m := \dim \overline{Q}_h$. Here, $\chi_T(x)$ is 1 if $x \in T$, and zero otherwise. Furthermore, $r : \mathcal{T}_h \times \{1,2\} \to \{1,\dots,m\}$ denotes an appropriate bijective numbering. Consequently, (4.2) is equivalent to find $(x, y, z) \in \mathbb{R}^n \times \mathbb{R}^m \times \tilde{\Lambda}$, such that,

$$\begin{aligned} &Ax + By = L, \\ &B^\top x + Cy + Dz = 0, \\ &(z - \tilde{z})^\top Dy \le 0 \end{aligned}$$



(a) (b) (c)

(d) (e) (f)

Figure 2: $(\tilde{p}_h : \tilde{p}_h)^{1/2}$ for different ranges and tolerances, **(a)** $[0, 10^{-6}]$, **(b)** $[0, 10^{-5}]$, **(c)** $[0, 10^{-1}]$ with tol $= 10^{-5}$, and **(d)** $[0, 10^{-6}]$, **(e)** $[0, 10^{-5}]$, **(f)** $[0, 10^{-1}]$ with tol $= 10^{-10}$.

for all $\tilde{z} \in \tilde{\Lambda} := \{z \in \mathbb{R}^m \mid \sum_{j=1}^m z_j \psi_j \in \overline{\Lambda}_h\}$. Here, the matrices $A \in \mathbb{R}^{n\times n}$, $L \in \mathbb{R}^n$, $B \in \mathbb{R}^{m\times n}$ and $C, D \in \mathbb{R}^{m\times m}$ are defined as

$$A_{ij} := (\mathbb{C}\varepsilon(\varphi_j), \varepsilon(\varphi_i))_0, \quad L_i := (f, \varphi_i)_0 + (g, \gamma_{|\Gamma_N}(\varphi_i))_{0,\Gamma_N},$$
$$B_{ij} := (-\mathbb{C}\psi_j, \varepsilon(\varphi_i))_0, \quad C_{ij} := ((\mathbb{C} + \mathbb{H})\psi_j, \psi_i)_0, \quad D_{ij} := \sigma_y(\psi_j, \psi_i)_0.$$

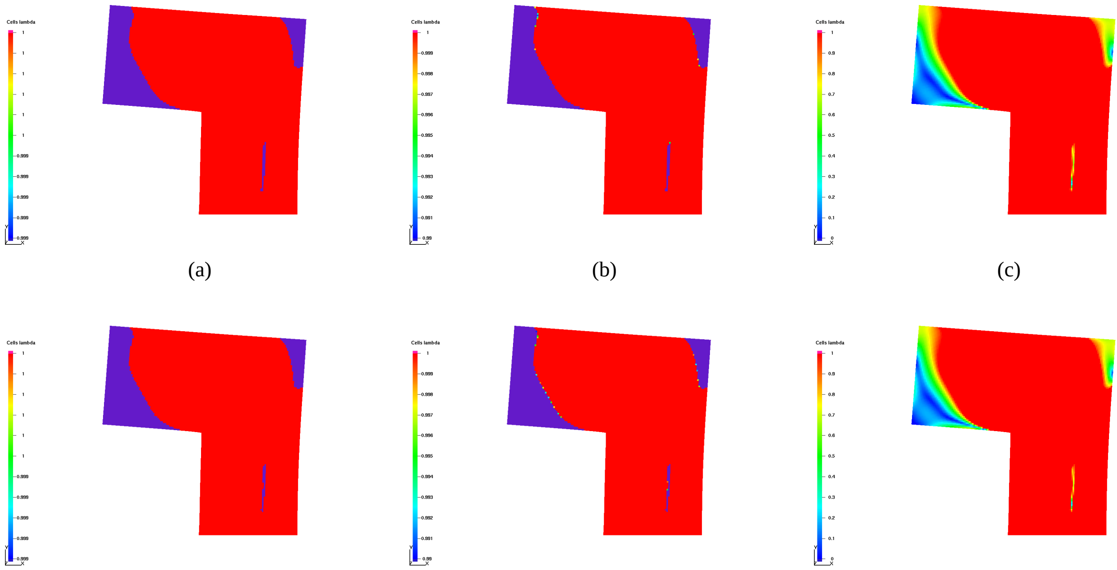
Obviously, the matrices A, C and D are symmetric and positive definite. Moreover, the matrices C and D are diagonal matrices with the diagonal entries $C_{jj} = 2|T|(2\mu + \xi)$ and $D_{jj} = 2|T|\sigma_y$ for $T \in \mathcal{T}_h$, $j = r(T, s)$ and $s = 1, 2$. Using a projection $P : \mathbb{R}^m \to \tilde{\Lambda}$ and an invertible matrix $S \in \mathbb{R}^{(n+m)\times(n+m)}$, we obtain an iterative scheme by

$$\begin{aligned} \begin{pmatrix} x^{k+1} \\ y^{k+1} \end{pmatrix} &= \begin{pmatrix} x^k \\ y^k \end{pmatrix} - \rho_1 S \begin{pmatrix} Ax^k + By^k - L \\ B^\top x^k + Cy^k + Dz^k \end{pmatrix}, \\ z^{k+1} &= P(z^k + \rho_2 D y^{k+1}). \end{aligned} \tag{10.1}$$

The convergence of this method for some parameters $\rho_1, \rho_2 > 0$ is proven in [14]. In Uzawa's method with projection the matrix S is chosen as

$$S := \begin{pmatrix} A & B \\ B^\top & C \end{pmatrix}^{-1}$$

or, in the case of the inexact Uzawa's method, as an appropriate approximation of it, cf. [11]. To define a suitable projection P, let $i, j \in \{1, \dots, m\}$, $z \in \mathbb{R}^m$ and $d(i, j) := 2(z_i^2 + z_j^2)$. We set $P_{i,j}(z) := d(i, j)^{-1/2} z_i$ if $d(i, j) > 1$ and to z_i otherwise. Therewith, the projection P is given by $P_{r(T,1)} := P_{r(T,1),r(T,2)}$ and $P_{r(T,2)} := P_{r(T,2),r(T,1)}$ for $T \in \mathcal{T}_h$. We emphasize that we do not focus on efficient solution algorithms in this paper. We primarily introduce Uzawa's method because of its implementational simplicity to solve the mixed discretization (4.3). Since the sparsity structure of the matrix $B^\top C^{-1} B$ is included in the structure of the matrix A, it is easy to determine the matrix $E := A - BC^{-1}B^\top$ by extending the usual assembling process. It is straightforward to see that E is symmetric and positive definite, too. The



(a) (b) (c)

(d) (e) (f)

Figure 3: $(\tilde{\lambda}_h : \tilde{\lambda}_h)^{1/2} \approx (\lambda_h : \lambda_h)^{1/2}$ for different ranges and tolerances, **(a)** $[0.999, 1]$, **(b)** $[0.99, 1]$, **(c)** $[0, 1]$ with tol $= 10^{-5}$, and **(d)** $[0.999, 1]$, **(e)** $[0.99, 1]$, **(e)** $[0, 1]$ with tol $= 10^{-10}$.

scheme (10.1) is then simplified to

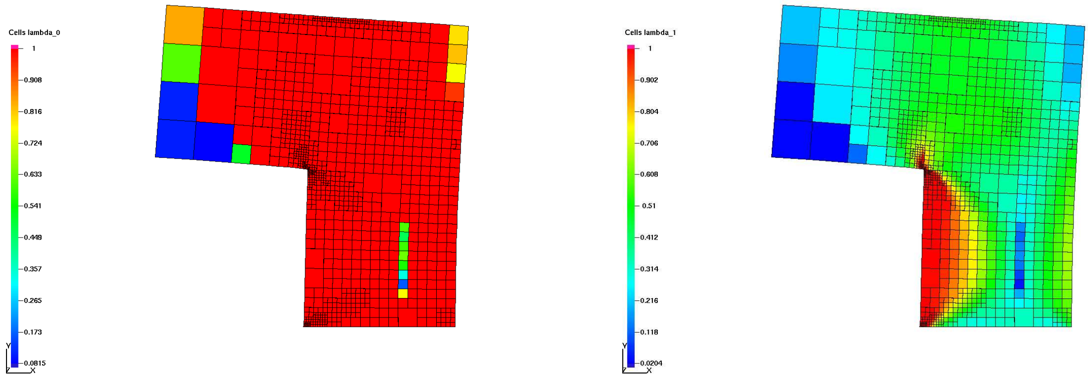
$$
\begin{aligned}
\zeta^k &= E^{-1}(L + \kappa B z^k),\\
x^{k+1} &= (1-\rho_1)x^k - \rho_1 \zeta^k,\\
y^{k+1} &= (1-\rho_1)y^k - \rho_1(\kappa z^k - C^{-1}B^\top \zeta^k),\\
z^{k+1} &= P(z^k + \rho_2 D y^{k+1})
\end{aligned}
$$

with $\kappa := \sigma_y(2\mu + \xi)^{-1}$. Further aspects on solution schemes in elastoplasticity based on the proposed mixed method will be considered in future work.

To adapt the finite element meshes, we use the derived error estimates within a simple fixed fraction strategy, where a fixed fraction of all mesh elements with the largest error contributions to the total error is refined. The quadrilateral elements are isotropically refined into four new elements. We allow for (multilevel) hanging nodes possibly resulting from the adaptive refinement process. The first test example is given by the standard L-shape domain, where Ω is set to $(0, 1)^2 \backslash (0, 0.5)^2$. Furthermore, we assume homogeneous Dirichlet boundary conditions on $\Gamma_D := [0.5, 1] \times \{0\}$. In the first instance, we consider single yield plasticity and assume the surface traction g to be non-zero, namely $g := 1.25$, only on $[0, 1] \times \{1\}$. The material parameters are chosen as $\lambda := 1000$, $\mu = 1000$, $\xi := 100$ and $\sigma_y = 1.25$. The volume force f is set to zero. Even though, the exact solution for this problem is not known, we expect singular behavior at the reentrant corner and at the points where the boundary conditions change. Indeed, we observe adaptive refinements towards those points as we can see in Figure 1a.

To check the performance of the adaptive refinements, we compare the estimated convergence rates obtained by the adaptive algorithm with the rates obtained by uniform refinements. Obviously, we gain better convergence rates using adaptive schemes in comparison to uniform mesh refinement. Moreover, we observe that the additional term $\|\operatorname{dev}(\sigma_h - \mathbb{H} p_h) - \sigma_y \lambda_h\|$ is of the same order as $\overline{\eta}$.

The use of the mixed methods as proposed in this paper may be motivated by the lack of regularization parameters as required in Newton's method. Another motivation to apply the mixed method in conjunction with Uzawa's method is to detect regions of pure elastic deformation. Clearly, such regions are characterized by $p = 0$. However, using



(a) (b)

Figure 4: Lagrange multipliers describing the first and second yield surface: **(a)** $(\lambda_{0,h} : \lambda_{0,h})^{1/2}$, **(b)** $(\lambda_{1,h} : \lambda_{1,h})^{1/2}$.

discretization as well as solution schemes it is not clear whether the calculated discrete plastic variable $\tilde{p}_h \approx p_h$ approximates zero or just a small value. In Figure 2, $(\tilde{p}_h : \tilde{p}_h)^{1/2}$ is depicted in several ranges and for different tolerances

$$(|x^{k+1} - x^k| + |y^{k+1} - y^k| + |z^{k+1} - z^k|)/(|x^k| + |y^k| + |z^k|) < \mathrm{tol}$$

with tol $= 10^{-5}$ (Figure 2a-c) and tol $= 10^{-10}$ (Figure 2d-f) using Uzawa's method. We observe that the accuracy of Uzawa's method has a significant influence on the plastic variable close to zero. Without additional information, the regions of pure elastic deformations can not be detected for the larger tolerance $tol = 10^{-5}$.

As discussed in Remark 3.3 and Remark 4.3, p_h is equal to zero if $\lambda_h : \lambda_h < 1$. Involving the calculated Lagrange multiplier $\tilde{\lambda}_h \approx \lambda_h$, we obtain a very sharp criterion as we can see in Figure 3a-f where the same tolerances are used, tol $= 10^{-5}$ (Figure 3a-c) and tol $= 10^{-10}$ (Figure 3d-f). We already observe sharp distributions for tol $= 10^{-5}$.

Finally, we consider multi-yield plasticity with two yield surfaces as introduced in Remark 3.4 with the parameters $\sigma_{y,0} := 1.25$, $\sigma_{y,1} := 5$, $\xi_0 := 100$, and $\xi_1 := 50$. Again, Ω is the L-shape domain as in the example of single yield plasticity. The boundary conditions and the exterior forces remain the same as well. Figure 4 shows the Lagrange multipliers $\lambda_{0,h}$ and $\lambda_{1,h}$ describing the first and second yield surface on an adaptively refined mesh.

References

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

[2] C. Bernardi, V. Girault, A local regularization operator for triangular and quadrilateral finite elements., SIAM J. Numer. Anal. 35 (5) (1998) 1893–1916.

[3] V. Bostan, W. Han, B. Reddy, A posteriori error estimation and adaptive solution of elliptic variational inequalities of the second kind., Appl. Numer. Math. 52 (1) (2005) 13–38.

[4] D. Braess, Finite elements. Theory, fast solvers and applications in solid mechanics. Translated from German by Larry L. Schumaker., Cambridge: Cambridge University Press. xvii, 365 p., 2007.

[5] C. Carstensen, Numerical analysis of the primal problem of elastoplasticity with hardening., Numer. Math. 82 (4) (1999) 577–597.

[6] C. Carstensen, J. Hu, Hanging nodes in the unifying theory of a posteriori finite element error control., J. Comput. Math. 27 (2-3) (2009) 215–236.

[7] C. Carstensen, R. Klose, A. Orlando, Reliable and efficient equilibrated a posteriori error finite element error control in elastoplasticity and elastoviscoplasticity with hardening, Comput. Methods Appl. Mech. Engrg. 195 (19-22) (2006) 2574–2598.

[8] C. Carstensen, J. Valdman, M. Brokate, A quasi-static boundary value problem in multi-surface elastoplasticity: part 1 - analysis, Math. Methods Appl. Sci. 27 (2004) 1697–1710.

[9] C. Carstensen, J. Valdman, M. Brokate, A quasi-static boundary value problem in multi-surface elastoplasticity: part 2 - numerical solution, Math. Methods Appl. Sci. 28 (2005) 881–901.

[10] C. Carstensen, J. Valdman, A. Orlando, A convergent adaptive finite element method for the primal problem of elastoplasticity, Internat. J. Numer. Methods Engrg. 67 (13) (2006) 1851–1887.

[11] X.-l. Cheng, W. Han, Inexact uzawa algorithms for variational inequalities of the second kind., Comput. Methods Appl. Mech. Eng. 192 (11-12) (2003) 1451–1462.

[12] P. W. Christensen, A nonsmooth newton method for elastoplastic problems., Comput. Methods Appl. Mech. Eng. 191 (11-12) (2002) 1189–1219.

[13] I. Ekeland, R. Temam, Convex analysis and variational problems., Studies in Mathematics and its Applications., North-Holland Publishing Company, Amsterdam, 1976.

[14] R. Glowinski, J. L. Lions, R. Trémolierès, Numerical analysis of variational inequalities., Studies in Mathematics and its Applications, North-Holland Publishing Company, Amsterdam, 1981.

[15] W. Han, Finite element analysis of a holonomic elastic-plastic problem., Numer. Math. 60 (4) (1992) 493–508.

[16] W. Han, B. Reddy, On the finite element method for mixed variational inequalities arising in elastoplasticity., SIAM J. Numer. Anal. 32 (6) (1995) 1778–1807.

[17] W. Han, B. Reddy, Plasticity. Mathematical theory and numerical analysis., Interdisciplinary Applied Mathematics. 9. New York, NY: Springer. xiii, 371 p., 1999.

[18] I. Hlávaček, J. Haslinger, J. Nečas, J. Lovíšek, Solution of variational inequalities in mechanics., Applied Mathematical Sciences, Springer-Verlag, New York, 1988.

[19] D. Knees, On global spatial regularity in elasto-plasticity with linear hardening., Calc. Var. Partial Differ. Equ. 36 (4) (2009) 611–625.

[20] P. Neff, D. Knees, Regularity up to the boundary for nonlinear elliptic systems arising in time-incremental infinitesimal elasto-plasticity., SIAM J. Math. Anal. 40 (1) (2008) 21–43.

[21] S. I. Repin, J. Valdman, Functional a posteriori error estimates for incremental models in elasto-plasticity., Cent. Eur. J. Math. 7 (3) (2009) 506–519.

[22] A. Sanchez, R. Arcangeli, Estimations des erreurs de meilleure approximation polynomiale et d'interpolation de lagrange dans les espaces de sobolev d'ordre non entier, Numer. Math. 45 (1984) 301–321.

[23] F. Suttmeier, Localised FE-analysis of Strang's problem based on Lagrange techniques., J. Numer. Math. 18 (2) (2010) 135–141.

[24] R. Verfürth, A review of a posteriori error estimation techniques for elasticity problems., Comput. Methods Appl. Mech. Eng. 176 (1-4) (1999) 419–440.