Numer. Math. (1999) 82: 577–597

Numerische Mathematik

Electronic Edition

Numerical analysis of the primal problem of elastoplasticity with hardening

Carsten Carstensen

Mathematisches Seminar, Christian-Albrechts-Universität zu Kiel,
Ludewig-Meyn-Strasse 4, D-24098 Kiel, Germany; e-mail: cc@numerik.uni-kiel.de

Received May 7, 1997 / Revised version received August 31, 1998

Summary. The finite element method is a reasonable and frequently utilised tool for the spatial discretization within one time-step in an elastoplastic evolution problem. In this paper, we analyse the finite element discretization and prove a priori and a posteriori error estimates for variational inequalities corresponding to the primal formulation of (Hencky) plasticity. The finite element method of lowest order consists in minimising a convex function on a subspace of continuous piecewise linear resp. piecewise constant trial functions. An a priori error estimate is established for the fully-discrete method which shows *linear* convergence as the mesh-size tends to zero, provided the exact displacement field u is smooth. Near the boundary of the plastic domain, which is unknown a priori, it is most likely that u is non-smooth. In this situation, automatic mesh-refinement strategies are believed to improve the quality of the finite element approximation. We suggest such an adaptive algorithm on the basis of a computable a posteriori error estimate. This estimate is reliable and efficient in the sense that the quotient of the error by the estimate and its inverse are bounded from above. The constants depend on the hardening involved and become larger for decreasing hardening.

Mathematics Subject Classification (1991): 65N30, 65R20, 73C50

1. Introduction

The time-independent elastoplastic material behaviour can be modeled in what Han and Reddy [11] call the primal formulation. This is to minimize a function

(1.1) $$\phi + \psi : X \times Y \to \mathbb{R} \cup \{+\infty\}$$

where $\phi : X \times Y \to \mathbb{R}$ is uniformly convex and has a Lipschitz continuous Fréchet derivative $D\phi$ while $\psi : Y \to \mathbb{R} \cup \{+\infty\}$ is convex and lower semicontinuous (and possibly non-smooth). We emphasize that this is a nonlinear variational inequality (see, e.g., [9] for details and further references on the numerical analysis of variational inequalities), where we seek $z = (x, y) \in Z := X \times Y$ satisfying

$$D\phi(z; z - \zeta) \leq \psi(\eta) - \psi(y) \qquad (\zeta := (\xi, \eta) \in Z). \tag{1.2}$$

The minimization problem (1.1) as the variational inequality (1.2) are dual to and so equivalent to the classical model in plasticity. The numerical analysis of the latter model was analyzed by Johnson (see, e.g., [13–16]) who proved a priori linear convergence and also established a posteriori error control as adaptive mesh-refining algorithms.

The aim of this paper is to establish the analog for (1.1) and (1.2), the primal form for plasticity with hardening in the spatially discrete situation. First, we improve the a priori error estimates from [10] (and a list of earlier references quoted in [11]). Secondly, we prove an a posteriori error estimate which justifies an adaptive algorithm for automatic mesh-refinement. The estimate is reliable and efficient. But, the constants involved rely on the hardening and so the estimates become worse for vanishing hardening parameters. (Then one should follow corresponding arguments in [16].)

The proof argues with Jensen's inequality for constant coefficients in the material laws. Hence, it is not too obvious that the improved convergence order is not destroyed by varying coefficients. Our proof relies on a closer study of the hardening laws (our analysis covers perfect plasticity as well, but there are only much weaker implications). To keep the representation short and precise, we focus on one (quite general master) example, which models combined isotropic and kinematic hardening, instead of stating general conditions in an abstract fashion. However, the technique applies to other situations as well.

An outline of the paper is as follows. In Sect. 2 we state the continuous problem in its strong, weak, primal, and dual form. In addition, preliminary consequences of the action of the hardening law are established. The corresponding spatially discrete problem is introduced in Sect. 3. For simplicity, we treat the lowest order method, but it should be stressed that the advantage of the primal formulation (over the more elaborated dual formulation) is that conform ansatz functions of arbitrary order can be employed. Corresponding a priori resp. a posteriori error estimates are stated in Sect. 4 resp. Sect. 5. The analysis in the proofs provided in Sect. 6 covers effects of numerical integration as well as discrete evaluation of the material law. Thereby, since ψ may be infinite on piecewise constant ansatz functions, a discrete counterpart $\psi_{\mathcal{S}}$ is required and we face a variational crime in the sense that the exact resp. the discrete solution (x, y) resp. $(x_{\mathcal{S}}, y_{\mathcal{S}})$ may yield $\psi_{\mathcal{S}}(y) = \infty = \psi(y_{\mathcal{S}})$.

Numerical evidence of the linear convergence of the lowest order scheme in the time-independent case will be provided in [2] and in the time-depending in [1].

2. The strong, weak, primal, and dual form of the continuous problem

In the small-strain models of solid mechanics, a bounded Lipschitz domain Ω in $\mathbb{R}^d$, $d = 1, 2, 3$, serves as reference and current configuration of a body. The strong form of equilibrium conditions states that the (Cauchy) stress field $\sigma \in L^2(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}})$, $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ being the set of all real symmetric $d \times d$ matrices, satisfies

$$\operatorname{div}\sigma + f = 0, \quad \sigma = \sigma^{\mathrm{T}} \qquad \text{in } \Omega, \tag{2.1}$$

$$\sigma \cdot n = g \qquad \text{on } \Gamma_{\mathrm{N}}, \tag{2.2}$$

where $f \in L^2(\Omega; \mathbb{R}^d)$ is a given applied volume force and $g \in L^2(\Gamma_{\mathrm{N}}; \mathbb{R}^d)$ is a given applied surface force. (The Lebesgue and Sobolev spaces in the definition are defined in a standard way [12, 17, 23].) The boundary $\Gamma = \partial\Omega$ is split into the Dirichlet boundary Γ_{D}, a compact set of positive surface measure, and the (possibly empty) Neumann boundary $\Gamma_{\mathrm{N}} = \Gamma \setminus \Gamma_{\mathrm{D}}$. The exterior unit vector on Γ (which exists almost everywhere on Γ) is denoted as n (see (2.2)). The displacement field

$$u \in H^1_{\mathrm{D}}(\Omega) := \{w \in H^1(\Omega)^d : w|_{\Gamma_{\mathrm{D}}} = 0\} \tag{2.3}$$

is linked to the (linear Green) strain field $\epsilon(u) \in L^2(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}})$,

$$(\epsilon(u))_{jk} := \frac{1}{2}\left(\frac{\partial u_j}{\partial x_k} + \frac{\partial u_k}{\partial x_j}\right) \qquad (j, k = 1, \ldots, d). \tag{2.4}$$

The constitutive relations in small strain elastoplasticity are based on an additive split of the total strain $\epsilon(u)$ into an elastic part e and a plastic part p,

$$\epsilon(u) = e + p. \tag{2.5}$$

A free energy is assumed in an uncoupled form as

$$F(e, \xi) = \frac{1}{2} e : \boldsymbol{C} e + \xi \cdot \boldsymbol{H}\xi, \tag{2.6}$$

where $\boldsymbol{C} \in L^\infty(\Omega; \mathbb{R}^{d\times d\times d\times d})$ is the fourth order elasticity tensor, with the Lamé constants λ and μ,

$$\boldsymbol{C}q := 2\mu q + (\lambda \cdot \operatorname{tr} q) 1_{d\times d} \qquad (q \in \mathbb{R}^{d\times d}_{\mathrm{sym}}), \tag{2.7}$$

and $\boldsymbol{H} \in L^\infty(\Omega; \mathbb{R}^{m\times m}_{\mathrm{sym}})$ is the positive definite modulus of hardening. In (2.7), $1_{d\times d}$ denotes the $(d\times d)$-unit matrix and we define $\mathrm{tr}\, q := q : 1_{d\times d}$ as the trace of q. Here, in (2.6), and below, the scalar product of two matrices in $p, q \in \mathbb{R}^{d\times d}$ is written with a colon, e.g., $p : q := \sum_{i,j=1}^d p_{ij} q_{ij}$.

The elastic part e and the internal variable ξ are linked to the stress σ and an internal stress χ through the free energy (2.6),

$$\sigma = \frac{\partial F}{\partial e} \quad \text{and} \quad \chi = -\frac{\partial F}{\partial \xi}. \tag{2.8}$$

Finally, the material law of plastic evolution is the principle of maximal dissipation,

$$(p, \xi) \in \partial\varphi(\sigma, \chi), \tag{2.9}$$

where $\varphi : \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ is the dissipation functional and $\partial\varphi$ denotes its sub-gradient.

Remark 2.1. According to the definition of the sub-gradient (or sub-differential) in convex analysis, the inclusion (2.9) (with respect to some scalar product $*$ in $\mathbb{R}^m$ which is specified below) equivalently reads

$$\begin{aligned} p : (\tilde\sigma - \sigma) + \xi * (\tilde\chi - \chi) &\le \varphi(\tilde\sigma, \tilde\chi) - \varphi(\sigma, \chi) \\ &\qquad ((\tilde\sigma, \tilde\chi) \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^m). \end{aligned} \tag{2.10}$$

The weak form of (2.1)–(2.9) is obtained straightforwardly as discussed, e.g., in [11], and is (formally) equivalent to the problem which Han and Reddy call dual formulation of elastoplasticity (we refer to [11] and omit details).

Definition 2.1 (Dual form). For $u \in X := H^1_{\mathrm{D}}(\Omega)^d$ and $(\sigma, \chi) \in Y := L^2(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^m)$, set

$$\begin{aligned} \phi(u, \sigma, \chi) := {} & \frac{1}{2}\int_\Omega \sigma : \boldsymbol{C}^{-1}\sigma\, dx + \frac{1}{2}\int_\Omega \chi \cdot \boldsymbol{H}^{-1}\chi\, dx \\ & - \int_\Omega \sigma : \epsilon(u)\, dx + \int_\Omega f u\, dx + \int_{\Gamma_{\mathrm{N}}} g u\, ds, \end{aligned} \tag{2.11}$$

$$\psi(\sigma, \chi) := \int_\Omega \varphi(\sigma, \chi)\, dx. \tag{2.12}$$

Then, the dual problem consists in finding a minimizer (u, σ, χ) of $\phi + \psi$ in $Z := X \times Y$.

The following important class of dissipation functionals is usually applied to metals or other ductile materials.

Example 2.1. Von-Mises' yield function with combined kinematic and isotropic hardening states that a (generalized) stress (σ, χ) is *admissible* if $\chi = (a, b) \in \mathbb{R} \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \equiv \mathbb{R}^m$, $m = 1 + d(d+1)/2$, with $a \geq 0$ and

$$\Phi(\sigma, a, b) := |\operatorname{dev}\sigma - \operatorname{dev} b| - \sigma_Y(1 + Ha) \leq 0. \tag{2.13}$$

Here, $\sigma_Y > 0$ is the yield stress and $H \geq 0$ is the hardening modulus. Furthermore,

$$\operatorname{dev}\sigma := \sigma - \frac{\operatorname{tr}\sigma}{d} \cdot 1_{d\times d}. \tag{2.14}$$

Then, the dissipation functional in (2.9) is given as the characteristic functional of the admissible stresses (2.7), i.e.,

$$\varphi(\sigma, a, b) := \begin{cases} 0 & \text{if } a \geq 0 \wedge \Phi(\sigma, a, b) \leq 0 \\ \infty & \text{if } a < 0 \vee \Phi(\sigma, a, b) > 0 \end{cases}$$
$$((\sigma, a, b) \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}). \tag{2.15}$$

By definition of the sub-gradient, (2.9) reads $\Phi(\sigma, a, b) \leq 0$ and for all $(\tilde\sigma, \tilde a, \tilde b) \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ with $\Phi(\tilde\sigma, \tilde a, \tilde b) \leq 0$ there holds

$$p : (\tilde\sigma - \sigma) - \alpha \cdot \mathsf{A}^{-1}(\tilde a - a) - \beta : \boldsymbol{B}^{-1}(\tilde b - b) \leq 0. \tag{2.16}$$

Here, $\xi = (\alpha, \beta)$ and $\chi = (a, b) = -\boldsymbol{H}\xi$ (recall (2.8)) and we introduced a scalar product $*$ in $\mathbb{R}^m$ represented by $(\mathsf{A}^{-1}, \boldsymbol{B}^{-1})$. In other words, if $\Phi(\sigma, a, b) < 0$ then there is no plastic evolution, $(p, -\alpha, -\beta) = 0$, and if $\Phi(\sigma, a, b) = 0$ then the vector of plastic evolution $(p, -\alpha, -\beta)$ is perpendicular to the surface of admissible (generalized) stresses (with respect to the scalar product given in (2.16)). Therefore, the maximal dissipation principle is also called the normal rule.

Remarks 2.2. 1. Example 2.1 models combined isotropic and kinematic hardening for the von-Mises yield condition; in particular, kinematic hardening for $H = 0$ and isotropic hardening for $\boldsymbol{B} = 0$, and perfect plasticity for $H = 0$ and $\boldsymbol{B} = 0$. In the sequel, we will say that a constant is hardening-independent if it does not depend on A or $\boldsymbol{B}$. In particular, $\boldsymbol{H}$ is expected to be bounded and independent of A and $\boldsymbol{B}$.
2. In the presence of hardening, the dual problem has a unique solution [15]. In case of perfect plasticity we have of no hardening, and solutions do, in general, *not* exist in $H^1_{\mathrm{D}}(\Omega) \times L^2(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}})$, cf. [13,20,21] for details in this case.
3. A dual formulation is obtained by using the dual φ^* of φ, i.e.,

$$\varphi^*(b) := \sup_a \{a \cdot b - \varphi(a)\}, \tag{2.17}$$

and is based on the equivalence of $a \in \partial\varphi(b)$ and $b \in \partial\varphi^*(a)$. Therefore, the dual form to (2.9) reads

$$(\sigma, \chi) \in \partial\varphi^*(p, \xi). \tag{2.18}$$

Definition 2.2 (Primal form). For $u \in X := H^1_{\mathrm{D}}(\Omega)^d$ and $(p,\xi) \in Y := L^2(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^m)$, set

$$\begin{aligned}\phi(u,p,\xi) := \frac{1}{2}\int_\Omega (p-\epsilon(u)) : \boldsymbol{C}(p-\epsilon(u))\,dx \\ +\frac{1}{2}\int_\Omega \xi\cdot \boldsymbol{H}\xi\,dx - \int_\Omega fu\,dx - \int_{\Gamma_{\mathrm{N}}} gu\,ds,\end{aligned} \tag{2.19}$$

$$\psi(p,\xi) := \int_\Omega \varphi^*(p,\xi)\,dx. \tag{2.20}$$

Then, the primal problem consists in finding a minimiser (u,p,ξ) of $\phi+\psi$ in $Z := X\times Y$.

Remarks 2.3. 1. The primal and dual problem are equivalent to each other and of the form (1.1) which is equivalent to (1.2). (For an elementary proof of "(1.1) implies (1.2)", we infer from convexity of ψ that $\phi(x,y)+\psi(y) \le \phi(x,\lambda\eta+(1-\lambda)y)+\lambda\psi(\eta)+(1-\lambda)\psi(y)$. Rearranging this and letting $\lambda\to 0$ we obtain (1.2).)
2. The quadratic form ϕ is known to be uniformly convex and ψ is convex, lower semi-continuous, and non-negative [10,11,5]. Hence, the primal problem has exactly one solution.

The numerical analysis of the primal problem is under consideration in the next sections. This section is concluded by computing the functional φ^* and illustrating the action of the hardening law related to Example 2.1.

Proposition 2.1. *If* $\varphi : \mathbb{R}^{d\times d}_{\mathrm{sym}}\times\mathbb{R}\times\mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}\cup\{\infty\}$ *is defined by* (2.13)–(2.16) *then its dual functional* $\varphi^* : \mathbb{R}^{d\times d}_{\mathrm{sym}}\times\mathbb{R}\times\mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}\cup\{\infty\}$ *is, for* $(p,\alpha,\beta)\in\mathbb{R}^{d\times d}_{\mathrm{sym}}\times\mathbb{R}\times\mathbb{R}^{d\times d}_{\mathrm{sym}}$, *given by*

$$\varphi^*(p,\alpha,\beta) = \begin{cases} \sigma_Y|p| & \text{if } \operatorname{tr} p = 0 \wedge p = -\boldsymbol{B}^{-1}\beta \\ & \wedge\ \mathsf{A}^{-1}\alpha + \sigma_Y H|p| \le 0, \\ \infty & \text{if not.}\end{cases} \tag{2.21}$$

Proof. According to (2.17), (2.15) and (2.16),

$$\varphi^*(p,\alpha,\beta) := \sup_{\Phi(\sigma,a,b)\le 0} \{p:\sigma + \alpha\mathsf{A}^{-1}a + \beta : \boldsymbol{B}^{-1}b\}, \tag{2.22}$$

where the supremum is taken over all $(\sigma,a,b)\in\mathbb{R}^{d\times d}_{\mathrm{sym}}\times\mathbb{R}\times\mathbb{R}^{d\times d}_{\mathrm{sym}}$ satisfying $\Phi(\sigma,a,b)\le 0$. First, we consider $(\sigma,a,b) = (r\cdot 1_{d\times d},0,0)$ and get $\varphi^*(p,\alpha,\beta)\ge r\cdot\operatorname{tr} p$ for all $r\in\mathbb{R}$. Thus, either $\operatorname{tr} p = 0$ or $\varphi^*(p,\alpha,\beta)=\infty$. Secondly, we consider $(\sigma,a,b) = (r\cdot E_{ij},0,r\cdot E_{ij})$ where $E_{ij} := \operatorname{sym} e_i\otimes e_j$ is the symmetric part of a matrix that has one non-vanishing entry 1 at the position (i,j), $i,j=1,\dots,d$. Then, $\varphi^*(p,\alpha,\beta)\ge r\cdot(p_{ij}+\boldsymbol{B}^{-1}b_{ij})$ and we infer that $\varphi^*(p,\alpha,\beta)=\infty$ or $p=-\boldsymbol{B}^{-1}\beta$. Thirdly, we may assume $\operatorname{tr} p=0$ and consider $(\sigma,a,b)=(\sigma_Y(1+rH)\operatorname{sign} p,r,0)$ where $r\geq 0$ and $\operatorname{sign} p=p/|p|$ if $p\neq 0$ and $\operatorname{sign} 0:=0$. A minor calculation shows that (σ,a,b) is admissible and that $\varphi^*(p,\alpha,\beta)\geq \sigma_Y|p|+r(\mathsf{A}^{-1}\alpha+\sigma_Y H|p|)$. Thus, $\varphi^*(p,\alpha,\beta)=\infty$ or $\sigma_Y H|p|+\mathsf{A}^{-1}\alpha\leq 0$ because $r\geq 0$ may be arbitrary large. Moreover, letting $r=0$ we see $\varphi^*(p,\alpha,\beta)\geq\sigma_Y|p|$.

Finally, we assume $\sigma_Y H|p|+\mathsf{A}^{-1}\alpha\leq 0$, $\beta=-\boldsymbol{B}p$, $\operatorname{tr} p=0$ and $\Phi(\sigma,a,b)\leq 0$, $a\geq 0$. According to Cauchy's inequality and orthogonality of deviatoric and unit matrices yield

$$
\begin{aligned}
p:\sigma+\alpha\mathsf{A}^{-1}\alpha+\beta:\boldsymbol{B}^{-1}b &= p(\operatorname{dev}\sigma-\operatorname{dev} b)+a\mathsf{A}^{-1}a\\
&\leq \sigma_Y(1+Ha)|p|+\alpha\mathsf{A}^{-1}a\\
&\leq \sigma_Y|p|+a(\sigma_Y H|p|+\mathsf{A}^{-1}\alpha)\leq\sigma_Y|p|, \qquad (2.23)
\end{aligned}
$$

whence $\varphi^*(p,\alpha,\beta)\leq\sigma_Y|p|$. Thus, under the present assumptions on (p,α,β), $\sigma_Y|p|=\varphi^*(p,\alpha,\beta)$ (and otherwise $\varphi^*(p,\alpha,\beta)=\infty$). □

Proposition 2.2. *If* $(\sigma,\chi)\in\partial\varphi^*(p,\xi)$ *and* $\chi=(a,b)$, $\xi=(\alpha,\beta)\in\mathbb{R}\times\mathbb{R}^{d\times d}_{\mathrm{sym}}$ *such that* $p\neq 0$, *then, with* $k:=\mathsf{A}\sigma_Y H$, *we have*

$$\frac{\operatorname{dev}(\sigma-b)}{\sigma_Y(1+Ha)}=p/|p| \quad \textit{and} \quad \alpha=-k|p|. \qquad (2.24)$$

Proof. Note that $tr\,p=0$, $p=-B^{-1}\beta$, $\alpha+k\cdot|p|\leq 0$. For any $q\in\mathbb{R}^{d\times d}_{\mathrm{sym}}$ with $tr\,q=0$ we consider $\tilde p:=p+q$, $\tilde\alpha:=-k|p+q|$, $\tilde\beta:=-B(p+q)$, and $\tilde\xi:=(\tilde\alpha,\tilde\beta)$, such that $\varphi^*(\tilde p,\tilde\xi)=\sigma_Y|p+q|$. Thus, according to the definition of the subgradient, we have

$$\sigma:q+a\cdot\mathsf{A}^{-1}(\tilde\alpha-\alpha)+b:\boldsymbol{B}^{-1}(\tilde\beta-\beta)\leq\sigma_Y(|p+q|-|p|). \qquad (2.25)$$

The estimate on α and the equalities for $\tilde\alpha$, $\tilde\beta$, β lead in (2.25) to

$$\frac{\operatorname{dev}(\sigma-b)}{\sigma_Y(1+Ha)}:q\leq|p+q|-|p|. \qquad (2.26)$$

This is the definition for

$$\tau:=\frac{\operatorname{dev}(\sigma-b)}{\sigma_Y(1+Ha)}\in\partial|\operatorname{dev}\cdot|(p)=\operatorname{sign}\operatorname{dev} p, \qquad (2.27)$$

where $\operatorname{sign} p:=\{p/|p|\}$ if $p\neq 0$, and $\operatorname{sign} 0=\{q\in\mathbb{R}^{d\times d}_{\mathrm{sym}}|\,|q|\leq 1\}$. We remark that the last identity is known in convex analysis and so we give a proof only for convenient reading. Letting $q=\tau$ in (2.26) (notice $tr\,\tau=0$), we infer with the triangle inequality

$$|\tau|^2\leq|\tau+p|-|p|\leq|\tau|, \qquad (2.28)$$

whence $|\tau| \leq 1$. Letting $q = -p$ in (2.26) (notice $tr\, p = 0$), we infer after a multiplication with -1 that

$$|p| \leq \tau : p \leq |\tau| \cdot |p| \leq |p| \tag{2.29}$$

because of Cauchy's inequality and $|\tau| \leq 1$. So we have equality in Cauchy's inequality, and this shows $\operatorname{sign} p \subseteq \operatorname{sign} \tau$ and concludes the proof of (2.27). We add that $p \neq 0$ implies $\alpha = -k|p|$, because equality in (2.29) is possible only if we have equality in (2.25), which is a strict inequality if $\alpha < -k|p|$. □

3. Spatially discrete problem

In this section, we analyse the discrete finite element method based on piecewise linear resp. constant ansatz functions on a regular triangulation $\mathcal{T}$. Also, we specify notation of the model in Example 2.1.

Definition 3.1 (Triangulation). The triangulation $\mathcal{T}$ of the polyhedral domain Ω is assumed to be regular in the sense of [6] and satisfies the minimum angle condition such that there is a constant $c_1 > 0$ with

$$c_1^{-1} \cdot h_T^2 \leq |T| \leq c_1 \cdot h_T^2 \qquad (T \in \mathcal{T}), \tag{3.1}$$

where $|T|$ is the area and h_T is the diameter of T. (The triangulation $\mathcal{T}$ is assumed to match Ω exactly and a change of boundary conditions is only allowed in nodes.)

We define $S^0(\mathcal{T}) \subset L^2(\Omega)$ as the piecewise constant and $S^1(\mathcal{T}) \subset H^1(\Omega)$ or $S^1_{\mathrm{D}}(\mathcal{T}) \subset H^1_{\mathrm{D}}(\Omega)$ as continuous and piecewise affine functions; piecewise is understood with respect to $\mathcal{T}$. Define

$$X := H^1_{\mathrm{D}}(\Omega) \quad \text{and}$$

$$Y := \{(p,\xi) \in L^2(\Omega; \mathbb{R}^{d\times d} \times \mathbb{R}^m) : \operatorname{tr} p = 0 \wedge p = p^{\mathrm{T}}\}, \tag{3.2}$$

$$S_X := S^1_{\mathrm{D}}(\mathcal{T}) \quad \text{and}$$

$$S_Y := \{(p,\xi) \in S^0(\mathcal{T})^{d\times d+m} : \operatorname{tr} p = 0 \wedge p = p^{\mathrm{T}}\}, \tag{3.3}$$

$$\mathcal{S} := S_X \times S_Y \subset X \times Y = Z. \tag{3.4}$$

Remark 3.1. The restrictions $p = p^{\mathrm{T}}$ and $\operatorname{tr} p = 0$ are easily implemented by identifying the symmetric 2×2-matrix $p(x)$ with the 4 entries $r(x)$, $s(x)$, $s(x)$, and $-r(x)$ with scalar functions r, s for $d = 2$ (and corresponding formulae for $d = 3$). Since there are no other restrictions involved, the implementation of conform higher order methods is quite simple.

The material parameters A, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{H}$, and σ_Y may vary in the domain Ω and so the integrals have to be approximated by quadrature or interpolation by $\mathcal{T}$-piecewise constant functions $\mathsf{A}_{\mathcal{T}}$, $\boldsymbol{B}_{\mathcal{T}}$, $\boldsymbol{C}_{\mathcal{T}}$, $\boldsymbol{H}_{\mathcal{T}}$, $\sigma_{Y\mathcal{T}}$.

Definition 3.2 (Data Approximation). Given bounded and measurable functions A, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{H}$, σ_Y in Ω, we define their piecewise constant approximants $\mathsf{A}_{\mathcal{T}}$, $\boldsymbol{B}_{\mathcal{T}}$, $\boldsymbol{C}_{\mathcal{T}}$, $\boldsymbol{H}_{\mathcal{T}}$, $\sigma_{Y\mathcal{T}}$ in Ω, with respect to the triangulation $\mathcal{T}$, such that, e.g.,

(3.5) $$(\mathsf{A}_{\mathcal{T}})|_T \in \mathbb{R} \qquad (T \in \mathcal{T})$$

(and corresponding formulae for the remaining quantities). The above mentioned properties are preserved such as $\mathsf{A}_{\mathcal{T}}$, $\boldsymbol{B}_{\mathcal{T}}$, $\boldsymbol{C}_{\mathcal{T}}$, $\boldsymbol{H}_{\mathcal{T}}$ are symmetric and positive definite as A, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{H}$ are and $\sigma_{Y\mathcal{T}} > 0$ as $\sigma_Y > 0$.

The mean operator $\mathcal{M}$ is defined with respect to $\mathcal{T}$ by

(3.6) $$(\mathcal{M}f)|_T := \int_T f\,dx/\text{meas}(T) \qquad (f \in L^2(\Omega), T \in \mathcal{T})$$

such that $\mathcal{M}f$ is constant on each element T with volume $\text{meas}(T)$ and equals its integral mean there. (We apply $\mathcal{M}$ to each component if the argument is a vector or matrix.) Finally, let $(f_{\mathcal{T}}, g_{\mathcal{T}}) \in S^0(\mathcal{T})^d \times \{w|_{\Gamma_\text{N}} : w \in S^0(\mathcal{T})^d\}$ be constant on each $T \in \mathcal{T}$.

The spatially discrete problem is simply the original problem when we replace all the material parameters and functions of the right-hand side and in initial values by their discrete counterparts.

Definition 3.3 (Discrete FEM). For $u \in X := H^1_\text{D}(\Omega)$, $(p,\xi) \in Y := L^2(\Omega; \mathbb{R}^{d\times d}_\text{sym} \times \mathbb{R}^m)$,

(3.7) $$\begin{aligned}\phi_{\mathcal{T}}(u,p,\xi) := &\frac{1}{2}\int_\Omega (p-\epsilon(u)) : \boldsymbol{C}_{\mathcal{T}}(p-\epsilon(u))\,dx \\ &+\frac{1}{2}\int_\Omega \xi : \boldsymbol{H}_{\mathcal{T}}\xi\,dx \\ &-\int_\Omega f_{\mathcal{T}} u\,dx - \int_{\Gamma_\text{N}} g_{\mathcal{T}} u\,ds,\end{aligned}$$

(3.8) $$\psi_{\mathcal{T}}(p,\xi) := \int_\Omega \varphi^*_{\mathcal{T}}(p,\xi)\,dx,$$

where $\varphi^*_{\mathcal{T}}(p,\alpha,\beta) := \sigma_{Y\mathcal{T}}|p|$ if simultaneously $\operatorname{tr} p = 0$, $p = -\boldsymbol{B}^{-1}_{\mathcal{T}}\beta$, and, $\mathsf{A}^{-1}_{\mathcal{T}}\alpha + \sigma_{Y\mathcal{T}}H_{\mathcal{T}}|p| \le 0$, while $\varphi^*(p,\alpha,\beta) := \infty$ if not. Then, the fully-discrete finite element scheme consists of minimising $\phi_{\mathcal{T}} + \psi_{\mathcal{T}}$ on $\mathcal{S}$.

Remarks 3.2. 1. The approximations $\phi_{\mathcal{T}}$ and $\psi_{\mathcal{T}}$ inherit the convexity properties of ϕ and ψ and so there exists exactly one solution to the discrete problem.
2. Since possibly $\varphi^*(p_{\mathcal{S}}, \alpha_{\mathcal{S}}, \beta_{\mathcal{S}}) = \infty$ or $\varphi^*_{\mathcal{T}}(p,\alpha,\beta) = \infty$ for the exact and discrete solution (p,α,β) and $(p_{\mathcal{S}}, \alpha_{\mathcal{S}}, \beta_{\mathcal{S}})$, respectively, we encounter what sometimes is called a variational crime.
3. We establish the convention that all known discrete quantities get an index $\mathcal{T}$ while the discrete unknowns are labelled by $\mathcal{S}$.

4. A priori error analysis

Adopting notation for the primal problem, let $z = (x, y) \in Z$ and $z_{\mathcal{S}} = (x_{\mathcal{S}}, y_{\mathcal{S}}) \in \mathcal{S} \subset Z$ be the minimisers of $\phi + \psi$ in Z and $\phi_{\mathcal{T}} + \psi_{\mathcal{T}}$ in $\mathcal{S}$, respectively.

Owing to the data approximation, further notation is provided to formulate a priori error estimates.

Definition 4.1 (Data-Errors). Let $\mathcal{M}$ be the mean operator, let id denote identity, and set $k := \mathsf{A} H \sigma_Y$ and $k_{\mathcal{T}} := \mathsf{A}_{\mathcal{T}} H_{\mathcal{T}} \sigma_{Y\mathcal{T}}$. Then,

$$\begin{aligned}
\delta_1 := \max\{ \| ((\mathsf{id} - \mathcal{M})(k, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{H}), \boldsymbol{C}_{\mathcal{T}} - \mathcal{M}\boldsymbol{C}, \\
\boldsymbol{H}_{\mathcal{T}} - \mathcal{M}\boldsymbol{H}) \|_{L^\infty(\Omega)}, \| (\mathsf{id} - \mathcal{M})\sigma_Y \|_{L^2(\Omega)} \}, \qquad (4.1)
\end{aligned}$$

$$\begin{aligned}
\delta_2 := \max\{ \| (k_{\mathcal{T}} - \mathcal{M}k, \boldsymbol{B}_{\mathcal{T}} - \mathcal{M}\boldsymbol{B}) \|_{L^\infty(\Omega)}, \\
\| \sigma_{Y\mathcal{T}} - \mathcal{M}\sigma_Y \|_{L^2(\Omega)} \}. \qquad (4.2)
\end{aligned}$$

The proof of the a priori error estimate will be provided in Sect. 6.

Theorem 4.1. *There exists a hardening-independent constant $C_1 > 0$ such that*

$$\begin{aligned}
& C_1^{-1} \| (\sigma - \sigma_{\mathcal{S}}, \xi - \xi_{\mathcal{S}}) \|^2_{L^2(\Omega)} \\
& \leq \delta_1^2 + \delta_2 (1 + \| p_{\mathcal{S}} \|_{L^2(\Omega)}) \\
& \quad + \| (\mathsf{id} - \mathcal{M})(|p|, p, \xi) \|^2_{L^2(\Omega)} + \inf_{V \in S_X} \Big\{ \| \epsilon(u - V) \|^2_{L^2(\Omega)} \\
& \quad + \int_\Omega (f - f_{\mathcal{T}})(u_{\mathcal{S}} - V)\, dx + \int_{\Gamma_{\mathrm{N}}} (g - g_{\mathcal{T}})(u_{\mathcal{S}} - V)\, ds \Big\}, \qquad (4.3)
\end{aligned}$$

where the stresses are $\sigma := \boldsymbol{C}(\epsilon(u) - p)$ and $\sigma_{\mathcal{S}} := \boldsymbol{C}(\epsilon(u_{\mathcal{S}}) - p_{\mathcal{S}})$.

Proofs of the following discussion of the right-hand side in the theorem will be given implicitly in Sect. 6.

Remarks 4.1. 1. In general, the constant C_1 depends on Ω, Γ_{D}, $\| (\boldsymbol{C}, \boldsymbol{C}^{-1}, \boldsymbol{H}, \boldsymbol{H}^{-1}) \|_{L^\infty(\Omega)}$ and $\|(p,\xi,\xi_{\mathcal{S}},\sigma_{\mathcal{S}})\|_{L^2(\Omega)}$.
2. The convergence estimate (4.3) implies that $\| (\xi_{\mathcal{S}}, \sigma_{\mathcal{S}}) \|_{L^2(\Omega)}$ is bounded. The boundedness of $\|(p, p_{\mathcal{S}})\|_{L^2(\Omega)}$ is implied by hardening (cf. Theorem 4.2 below).
3. If $\delta_2 = 0$, the right-hand side is independent of $\| p_{\mathcal{S}} \|_{L^2(\Omega)}$.
4. The dependence of the right-hand side on $\| (p, p_{\mathcal{S}}) \|_{L^2(\Omega)}$ can be relaxed. If we replaced $\| \sigma_{Y\mathcal{T}} - \mathcal{M}\sigma_Y \|_{L^2(\Omega)}$ in the definition of δ_2 by $\| \sigma_{Y\mathcal{T}} - \mathcal{M}\sigma_Y \|_{L^\infty(\Omega)}$, then the right-hand side would depend on $\| p_{\mathcal{S}} \|_{L^1(\Omega)}$ only. The latter norm is bounded because the minimisation involves $\psi_{\mathcal{T}}$.
5. For piecewise constant data (with respect to $\mathcal{T}$), we can arrange that $\delta_1 = 0 = \delta_2$. Then, the approximation error $\| (\mathsf{id} - \mathcal{M})(|p|, p, \xi) \|_{L^2(\Omega)}$

can be omitted on the right-hand side in (4.3).
6. To ensure $\delta_1^2 + \delta_2 = O(h^2)$ for $h := \max_{T\in\mathcal{T}} h_T$, h_T being the diameter of $T \in \mathcal{T}$, an elementwise one-point evaluation on the points of inertia is sufficient for piecewise smooth data.
7. Exact integration of f resp. g leads to $\int_\Omega (f - f_\mathcal{T})(u_\mathcal{S} - V)\,dx = 0$ resp. $\int_{\Gamma_\mathrm{N}} (g - g_\mathcal{T})(u_\mathcal{S} - V)\,ds = 0$. We refer to standard literature for effects of inexact integration [3,6] and mention that elementwise one-point Gauss-quadrature is sufficient in case that $\| u_\mathcal{S} \|_{H^1(\Omega)}$ is bounded (which is guaranteed by hardening, cf. Theorem 4.2 below).
8. If $\delta_1 = 0 = \delta_2$ and f, g are exactly integrated, the theorem specifies to

$$C_1^{-1} \| \sigma - \sigma_\mathcal{S} \|_{L^2(\Omega)}^2 \le \inf_{V\in S_X} \| \epsilon(u - V) \|_{L^2(\Omega)}^2 \tag{4.4}$$

and we stress that there is *no* contribution of the best-approximation error of the other variable p. This clearly indicates that a finer mesh for the approximation of the plastic strain variable p is pointless.
9. The estimate (4.3) holds also in case of perfect plasticity (where $m = 0$ and ξ, $\boldsymbol{H}$, etc. are omitted).
10. At first glance, if $u \in H^2(\Omega)$ and if the data are piecewise smooth, the theorem leads to the optimal linear convergence for approximants of the stress field and the internal variables.
11. At second glance, the estimate (4.4) and so (4.3) is poor in perfect plasticity, because we cannot expect that u is very smooth. There is evidence that σ is much smoother than u [18,19] and so (4.3) cannot be regarded as a quasi-optimal convergence estimate.

The hardening law allows further estimates. In case that the modulus of kinematic hardening $\boldsymbol{B}$ is absent or too small (i.e. $\| \boldsymbol{B}^{-1} \|_{L^\infty}(\Omega)$ is too large) we need further restrictions on the energetic coupling of kinematic and isotropic hardening. Usually, this coupling is omitted. Here, we allow a sufficiently small interaction. With $(1,0) \in \mathbb{R} \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \equiv \mathbb{R}^m$, define

$$\gamma := \operatorname*{ess\,inf}_{x\in\Omega} \min_{|q|\le 1} (1,0) \cdot \boldsymbol{H}(1, k^{-1}(x)\boldsymbol{B}(x)q)^{\mathrm{T}}, \tag{4.5}$$

$$\gamma_\mathcal{T} := \operatorname*{ess\,inf}_{x\in\Omega} \min_{|q|\le 1} (1,0) \cdot \boldsymbol{H}_\mathcal{T}(1, k_\mathcal{T}^{-1}(x)\boldsymbol{B}_\mathcal{T}(x)q)^{\mathrm{T}}, \tag{4.6}$$

where the argument in the minimum is amongst all $q \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ in the closed unit ball.

Remark 4.2. Notice that $\gamma, \gamma_T > 0$ if there is *no* energetic coupling of the hardening parameters (i.e., $\boldsymbol{H}, \boldsymbol{H}_\mathcal{T}$ are block-diagonal). The physical justification and interpretation of off-diagonal entries in $\boldsymbol{H}$ is unclear to the author. In the literature, it is usually assumed to be very small or even negligible.

Theorem 4.2. *Suppose either* $\|(\boldsymbol{B}_{\mathcal{T}}^{-1}, \boldsymbol{B}^{-1})\|_{L^\infty(\Omega)} < \infty$ *or* $k, k_{\mathcal{T}}, \gamma, \gamma_{\mathcal{T}} > 0$*. Then, there exists a hardening-dependent constant* $C_2 > 0$ *such that*

$$\begin{aligned} &C_2^{-1}\|(p - p_{\mathcal{S}}, \epsilon(u - u_{\mathcal{S}}))\|_{L^2(\Omega)}^2 \\ &\leq \|(\sigma - \sigma_{\mathcal{S}}, \xi - \xi_{\mathcal{S}})\|_{L^2(\Omega)}^2 + \delta_1^2 + \delta_2^2. \end{aligned} \tag{4.7}$$

Proof. Since $\sigma - \sigma_{\mathcal{S}} = \boldsymbol{C}(\epsilon(u - u_{\mathcal{S}}) - (p - p_{\mathcal{S}}))$, it is sufficient to estimate one of the summands, i.e., we will prove that $\|p - p_{\mathcal{S}}\|_2^2$ is bounded by the right-hand side in (4.7).

In the first case, suppose that kinematic hardening is present, i.e. $\|\boldsymbol{B}^{-1}\|_\infty$ and $\|\boldsymbol{B}_{\mathcal{T}}^{-1}\|_\infty$ are bounded from above, so that $\|\boldsymbol{B}^{-1} - \overline{\boldsymbol{B}}^{-1}\|_\infty$ and $\|\overline{\boldsymbol{B}}^{-1} - \boldsymbol{B}_{\mathcal{T}}^{-1}\|_\infty$, where $\overline{\boldsymbol{B}} := \mathcal{M}\boldsymbol{B}$, are bounded by $c \cdot (\delta_1 + \delta_2)$. Then, we infer from $p = -\boldsymbol{B}^{-1}\beta$ and $p_{\mathcal{S}} = -\boldsymbol{B}_{\mathcal{T}}^{-1}\beta_{\mathcal{S}}$ that

$$\begin{aligned} \|p - p_{\mathcal{S}}\|_2 &\leq \|(\boldsymbol{B}^{-1} - \overline{\boldsymbol{B}}^{-1})\beta\|_2 + \|(\overline{\boldsymbol{B}}^{-1} - \boldsymbol{B}_{\mathcal{T}}^{-1})\beta\|_2 \\ &\quad + \|\boldsymbol{B}_{\mathcal{T}}^{-1}(\beta - \beta_{\mathcal{S}})\|_2 \\ &\leq c\Big((\delta_1 + \delta_2)\|\beta\|_2 + \|\xi - \xi_{\mathcal{S}}\|_2\Big). \end{aligned} \tag{4.8}$$

In the second case, suppose that isotropic hardening is present, i.e., $\|1/k\|_\infty$ and $\|1/k_{\mathcal{T}}\|_\infty$ are bounded from above and $\gamma, \gamma_{\mathcal{T}}$ are positive. As in (4.8), we obtain from Proposition 2.2 that

$$\begin{aligned} \||p| - |p_{\mathcal{S}}|\|_2 &= \|\alpha/k - \alpha_{\mathcal{S}}/k_{\mathcal{T}}\|_2 \\ &\leq c\Big((\delta_1 + \delta_2)\|\alpha\|_2 + \|\xi - \xi_{\mathcal{S}}\|_2\Big). \end{aligned} \tag{4.9}$$

Almost everywhere in Ω, we have

$$|p - p_{\mathcal{S}}|^2 = (|p| - |p_{\mathcal{S}}|)^2 + 2(|p||p_{\mathcal{S}}| - p : p_{\mathcal{S}}). \tag{4.10}$$

The first term on the right-hand side can be bounded by (4.9). Moreover, writing $\tilde{\sigma}_{\mathcal{S}} := \boldsymbol{C}_{\mathcal{T}}\boldsymbol{C}^{-1}\sigma_{\mathcal{S}}$, we infer from Proposition 2.2 that $|\operatorname{dev}\sigma| = \sigma_Y(1 + Ha)$ and $|\operatorname{dev}\tilde{\sigma}_{\mathcal{S}}| = \sigma_{Y\mathcal{T}}(1 + H_{\mathcal{T}}a_{\mathcal{S}})$, where $(a, b) := -\boldsymbol{H}\xi$ and $(a_{\mathcal{S}}, b_{\mathcal{S}}) := -\boldsymbol{H}_{\mathcal{T}}\xi_{\mathcal{S}}$. If the second term $|p(x)||p_{\mathcal{S}}(x)| - p(x) : p_{\mathcal{S}}(x)$ on the right-hand side in (4.10) is nonzero it equals (by Proposition 2.2)

$$\begin{aligned} &\frac{\alpha\alpha_{\mathcal{S}}k^{-1}k_{\mathcal{T}}^{-1}\sigma_Y^{-1}\sigma_{Y\mathcal{T}}^{-1}}{(1 + Ha)(1 + H_{\mathcal{T}}a_{\mathcal{S}})} \\ &\Big(|\operatorname{dev}(\sigma - b)||\operatorname{dev}(\tilde{\sigma}_{\mathcal{S}} - b_{\mathcal{S}})| - \operatorname{dev}(\sigma - b) : \operatorname{dev}(\tilde{\sigma}_{\mathcal{S}} - b_{\mathcal{S}})\Big), \end{aligned} \tag{4.11}$$

where we neglected the argument $x \in \Omega$. We have to analyse the term $\alpha_{\mathcal{S}}/(1 + H_{\mathcal{T}}a_{\mathcal{S}})$ and may utilise that $\gamma > 0$. By $p_{\mathcal{S}} \neq 0$, we have $\alpha_{\mathcal{S}} =$

$-k_{\mathcal{T}}|p_{\mathcal{S}}|$ and $\beta_{\mathcal{S}} = -\boldsymbol{B}_{\mathcal{T}} p_{\mathcal{S}}$ and thus, writing $q := p_{\mathcal{S}}/|p_{\mathcal{S}}|$ and $\boldsymbol{H}_{\mathcal{T}} = (\boldsymbol{H}_{ij} : i, j = 1, 2)$ for $\boldsymbol{H}_{11} > 0$ and $\boldsymbol{H}_{12} \in \mathbb{R}^{1\times(m-1)}$,

$$\begin{aligned} |a_{\mathcal{S}}| &= |\boldsymbol{H}_{11}\alpha_{\mathcal{S}} + \boldsymbol{H}_{12}\beta_{\mathcal{S}}| = (\boldsymbol{H}_{11} + k_{\mathcal{T}}^{-1}\boldsymbol{H}_{12}\boldsymbol{B}_{\mathcal{T}} q)|\alpha_{\mathcal{S}}| \\ &\geq \gamma_{\mathcal{T}}|\alpha_{\mathcal{S}}|. \end{aligned} \tag{4.12}$$

Using this and its analog $|a| \geq \gamma|\alpha|$, the factor in (4.11) is seen to be bounded from above by a universal constant. Therefore, we conclude almost everywhere in Ω that

$$\begin{aligned} &|p - p_{\mathcal{S}}|^2 \leq (|p| - |p_{\mathcal{S}}|)^2 \\ &+2\frac{|\operatorname{dev}(\sigma - b)||\operatorname{dev}(\tilde{\sigma}_{\mathcal{S}} - b_{\mathcal{S}})| - \operatorname{dev}(\sigma - b) : \operatorname{dev}(\tilde{\sigma}_{\mathcal{S}} - b_{\mathcal{S}})}{k k_{\mathcal{T}} \sigma_Y \sigma_{Y\mathcal{T}} H H_{\mathcal{T}} \gamma\gamma_{\mathcal{T}}}. \end{aligned} \tag{4.13}$$

Arguing as in (4.10), we achieve

$$\begin{aligned} &|\operatorname{dev}(\sigma - b)||\operatorname{dev}(\tilde{\sigma}_{\mathcal{S}} - b_{\mathcal{S}})| - \operatorname{dev}(\sigma - b) : \operatorname{dev}(\tilde{\sigma}_{\mathcal{S}} - b_{\mathcal{S}}) \\ &= \frac{1}{2}|\operatorname{dev}(\sigma - \tilde{\sigma}_{\mathcal{S}} + b_{\mathcal{S}} - b)|^2 \\ &\quad -\frac{1}{2}(|\operatorname{dev}(\sigma - b)| - |\operatorname{dev}(\tilde{\sigma}_{\mathcal{S}} - b_{\mathcal{S}})|)^2 \\ &\leq \frac{1}{2}|\operatorname{dev}(\sigma - \tilde{\sigma}_{\mathcal{S}} + b_{\mathcal{S}} - b)|^2 \leq |\sigma - \tilde{\sigma}_{\mathcal{S}}|^2 + |b - b_{\mathcal{S}}|^2, \end{aligned} \tag{4.14}$$

which, together with (4.9) and (4.13) leads to

$$\begin{aligned} c^{-1}\| p - p_{\mathcal{S}} \|_2^2 \leq \delta_1^2 + \delta_2^2 + \| \xi - \xi_{\mathcal{S}} \|_2^2 \\ +\| \sigma - \tilde{\sigma}_{\mathcal{S}} \|_2^2 + \| b - b_{\mathcal{S}} \|_2^2. \end{aligned} \tag{4.15}$$

Here, the constant c depends on the constant in (4.9) and on $\| (k k_{\mathcal{T}} \sigma_Y \times \sigma_{Y\mathcal{T}} H H_{\mathcal{T}} \gamma\gamma_{\mathcal{T}})^{-1} \|_\infty$. It remains to bound $\| \sigma - \tilde{\sigma}_{\mathcal{S}} \|_2$ and $\| b - b_{\mathcal{S}} \|_2$. According to

$$\begin{aligned} \| (\sigma - \tilde{\sigma}_{\mathcal{S}}, b - b_{\mathcal{S}}) \|_2 \leq \| (\sigma - \sigma_{\mathcal{S}}, \boldsymbol{H}(\xi - \xi_{\mathcal{S}})) \|_2 \\ +\| ((1 - \boldsymbol{C}_{\mathcal{T}}\boldsymbol{C}^{-1})\sigma_{\mathcal{S}}, (\boldsymbol{H} - \boldsymbol{H}_{\mathcal{T}})\xi_{\mathcal{S}}) \|_2 \end{aligned} \tag{4.16}$$

we consider terms like

$$\begin{aligned} \| (\boldsymbol{C}^{-1} - \boldsymbol{C}_{\mathcal{T}}^{-1})\sigma_{\mathcal{S}} \|_2^2 &\leq \| \boldsymbol{C}^{-1}(\boldsymbol{C} - \boldsymbol{C}_{\mathcal{T}})\boldsymbol{C}_{\mathcal{T}}^{-1} \|_\infty^2 \| \sigma_{\mathcal{S}} \|_2^2 \\ &\leq \delta_1^2 \| \boldsymbol{C}^{-1} \|_\infty^2 \| \boldsymbol{C}_{\mathcal{T}}^{-1} \|_\infty^2 \| \sigma_{\mathcal{S}} \|_2^2 \end{aligned} \tag{4.17}$$

to establish a bound of (4.16) and conclude the proof. □

5. A posteriori error estimates

For each $T \in \mathcal{T}$, let h_T denote its diameter and let

$$\eta_T^2 = h_T^2 \int_T |f + \operatorname{div} \sigma_{\mathcal{S}}|^2 \, dx + \int_{\partial T} h_E |J(\sigma_{\mathcal{S}} \cdot n_E)|^2 \, ds \tag{5.1}$$

where $J(\sigma_{\mathcal{S}} \cdot n_E)$ is the jump of the discrete stress field along an edge E with normal n_E and size h_E with the usual modification $J(\sigma_{\mathcal{S}} \cdot n_E) := \sigma_{\mathcal{S}} \cdot n_E - g$ if $E \subset \overline{\Gamma_{\mathrm{N}}}$.

The proof of the a posteriori error estimate will be provided in Sect. 6.

Theorem 5.1. *Under the assumptions of Theorem 4.2, there exists a hardening-dependent constant $C_3 > 0$ such that*

$$\begin{aligned} &C_3^{-1} \| (p - p_{\mathcal{S}}, \epsilon(u - u_{\mathcal{S}}), \sigma - \sigma_{\mathcal{S}}, \xi - \xi_{\mathcal{S}}) \|_{L^2(\Omega)}^2 \\ &\leq \sum_{T \in \mathcal{T}} \eta_T^2 + \delta_1^2 + \delta_2 + \| (\mathsf{id} - \mathcal{M})(|p|, p, \xi) \|_{L^2(\Omega)}^2 \\ &+ \max_{V \in S_X \setminus \{0\}} \Big(\int_\Omega (f - f_{\mathcal{T}}) V \, dx + \int_{\Gamma_{\mathrm{N}}} (g - g_{\mathcal{T}}) V \, ds \Big) / \| \nabla V \|_{L^2(\Omega)} . \end{aligned} \tag{5.2}$$

Remarks 5.1. 1. The estimator η_T is the same as in pure elasticity (utilising the stress field from a discrete elasto-plastic problem).
2. In perfect plasticity, Theorem 5.1 is expected to be false. A closer inspection then shows that it is required to follow the arguments in [16] and to derive weaker estimates.
3. The discussion of the right-hand side is analogous to the comments in Remark 4.1. In particular, if $\delta_1 = 0 = \delta_2$, the term $\| (\mathsf{id} - \mathcal{M})(|p|, p, \xi) \|_{L^2(\Omega)}$ can be neglected.
4. In case of exact integration of f and g, the maximum in the upper bound is zero. Otherwise, the maximum can be computed. Standard arguments (involving Poincaré's inequality) show that elementwise one-point Gauss-quadrature is sufficient for this term to be of order $O(h^2)$ (provided f and g are piecewise smooth).

As in the case of pure elasticity, the a posteriori error estimate is efficient [22] in a local sense.

Theorem 5.2. *There exists a hardening-independent constant $C_4 > 0$ such that, for all $T \in \mathcal{T}$,*

$$\begin{aligned} C_4^{-1} \eta_T^2 \leq \| \sigma - \sigma_{\mathcal{S}} \|_{L^2(\omega_T)}^2 \\ + \operatorname{diam}(\omega_T)^2 (\| f - \mathcal{M} f \|_{L^2(\omega_T)}^2 + \| g - \overline{g} \|_{L^2(\omega_T \cap \Gamma_{\mathrm{N}})}^2) , \end{aligned} \tag{5.3}$$

where ω_T is the set of all neighbouring triangles which share one edge with T and $\overline{g}$ is the piecewise best-approximation of g in $L^2(\Gamma_{\mathrm{N}})$.

Proof. Since the estimator η_T involves stress terms only, the proof can follow arguments in [22] which are independent of the material laws. To illustrate this we prove the first halve of (5.3). Given $T \in \mathcal{T}$, let $b := \lambda_1\lambda_2\lambda_3$ with the barycentric coordinates λ_j on T. Direct calculations verify, $b \leq 1/27$,

$$h_T^2 \leq c_2 \int_T b\,dx, \quad \text{and} \quad \|\nabla b\|_{L^2(T)} \leq c_3, \tag{5.4}$$

where the positive constants c_2 and c_3 depend on c_1 (from (3.1)) but do not depend on h_T. Then, for $f|_T = \operatorname{div}(\sigma_{\mathcal{S}} - \sigma)|_T$ and $f_T := (\mathcal{M}f)|_T \in \mathbb{R}$, we have

$$\begin{aligned} c_2^{-1}\|\mathcal{M}f\|^2_{L^2(T)} &\leq f_T \int_T b f_T\,dx \\ &= f_T \int_T bf\,dx + f_T \int_T b(f_T - f)\,dx. \end{aligned} \tag{5.5}$$

Since the bubble function b vanishes on ∂T, integration by parts leads to

$$\int_T bf\,dx = \int_T (\sigma - \sigma_{\mathcal{S}})\epsilon(b)\,dx \leq c_3\|\sigma - \sigma_{\mathcal{S}}\|_{L^2(T)}. \tag{5.6}$$

Incorporating (5.6) in (5.5), we finally obtain

$$\begin{aligned} c_2^{-1}\|f_T\|_{L^2(T)} \leq c_3 c_1^{-1/2} h_T^{-1}\|\sigma - \sigma_{\mathcal{S}}\|_{L^2(T)} \\ +\|f_T - f\|_{L^2(T)}/27. \end{aligned} \tag{5.7}$$

From this we can estimate the volume contribution to η_T, that is (for some $c_4 > 0$)

$$\begin{aligned} h_T^2/2 \int_T |f|^2\,dx &\leq h_T^2 \int_T |f_T|^2\,dx + h_T^2 \int_T |f - f_T|^2\,dx \\ &\leq c_4(\|\sigma - \sigma_{\mathcal{S}}\|^2_{L^2(T)} + h_T^2\|f_T - f\|^2_{L^2(T)}). \end{aligned} \tag{5.8}$$

The estimation of the edge contributions is similar and replaces b by a product of two barycentric coordinates on two neighbouring elements. We refer to [22] for details. □

6. Proof of Theorem 4.1 and 5.1

First, we argue with (1.2) for both the continuous and the discrete problem. For all $\tilde{z} = (\tilde{x}, \tilde{y}) \in Z$ and all $\tilde{z}_{\mathcal{S}} = (\tilde{x}_{\mathcal{S}}, \tilde{y}_{\mathcal{S}}) \in S$, this yields

$$D\phi(z; z - z_{\mathcal{S}}) - D\phi(z_{\mathcal{S}}; z - z_{\mathcal{S}}) \tag{6.1}$$
$$\leq D\phi(z_{\mathcal{S}}; z_{\mathcal{S}} - \tilde{z}_{\mathcal{S}}) - D\phi_{\mathcal{T}}(z_{\mathcal{S}}; z_{\mathcal{S}} - \tilde{z}_{\mathcal{S}}) \tag{6.2}$$
$$+ D\phi(z_{\mathcal{S}}; \tilde{z}_{\mathcal{S}} - z) + \psi_{\mathcal{T}}(\tilde{y}_{\mathcal{S}}) - \psi(y) \tag{6.3}$$
$$+ D\phi(z; \tilde{z} - z_{\mathcal{S}}) + \psi(\tilde{y}) - \psi_{\mathcal{T}}(y_{\mathcal{S}}). \tag{6.4}$$

The remaining part of the proof consists of a careful estimation of the terms in (6.1)–(6.4) where we let $\tilde{x} = \tilde{u} \in X$ and $\tilde{x}_{\mathcal{S}} = \tilde{u}_{\mathcal{S}} \in S_X$ and fix $\tilde{y} = (\tilde{p}, \tilde{\xi})$ and $\tilde{\xi} = (\tilde{\alpha}, \tilde{\beta})$ etc. as

$$\tilde{y}_{\mathcal{S}} := (\tilde{p}_{\mathcal{S}}, \tilde{\alpha}_{\mathcal{S}}, \tilde{\beta}_{\mathcal{S}}) := (\mathcal{M}p, \min\{\mathcal{M}\alpha, -k_{\mathcal{T}}|\mathcal{M}p|\}, -\boldsymbol{B}_{\mathcal{T}}\mathcal{M}p), \tag{6.5}$$

$$\tilde{y} := (\tilde{p}, \tilde{\alpha}, \tilde{\beta}) := (p_{\mathcal{S}}, \alpha_{\mathcal{S}} + (k_{\mathcal{T}} - k)|p_{\mathcal{S}}|, -\boldsymbol{B}p_{\mathcal{S}}). \tag{6.6}$$

Secondly, one shows $\psi(\tilde{y}) = \int_\Omega \sigma_Y |p_{\mathcal{S}}|\, dx$ and $\psi_{\mathcal{T}}(\tilde{y}_{\mathcal{S}}) = \int_\Omega \sigma_{Y\mathcal{T}} \cdot |\mathcal{M}p|\, dx$. As in the following, we abbreviate $k := \mathsf{A}\sigma_Y H$ and $k_{\mathcal{T}} := \mathsf{A}_{\mathcal{T}}\sigma_{Y\mathcal{T}} H_{\mathcal{T}} \geq 0$ and, $e := \epsilon(u) - p$ and $e_{\mathcal{S}} := \epsilon(u_{\mathcal{S}}) - p_{\mathcal{S}}$, $\mathcal{Y} := (e, \xi)$ as $\mathcal{Y}_{\mathcal{S}} = (e_{\mathcal{S}}, \xi_{\mathcal{S}})$ etc. and $\mathcal{G}$ is the block diagonal tensor with diagonal entries $\boldsymbol{C}$ and $\boldsymbol{H}$ with corresponding modifications for $\overline{\mathcal{G}} = \mathcal{M}\mathcal{G}$ and $\mathcal{G}_{\mathcal{T}}$.

In step three, we consider (6.2) and calculate

$$\begin{aligned} &D\phi(z_{\mathcal{S}}; z_{\mathcal{S}} - \tilde{z}_{\mathcal{S}}) - D\phi_{\mathcal{T}}(z_{\mathcal{S}}; z_{\mathcal{S}} - \tilde{z}_{\mathcal{S}}) \\ &= \int_\Omega (\mathcal{Y}_{\mathcal{S}} - \tilde{\mathcal{Y}}_{\mathcal{S}}) : (\overline{\mathcal{G}} - \mathcal{G}_{\mathcal{T}})\mathcal{Y}_{\mathcal{S}}\, dx - \ell_{\mathcal{T}}(u_{\mathcal{S}} - \tilde{u}_{\mathcal{S}}) \end{aligned} \tag{6.7}$$

where $\ell_{\mathcal{T}}(w) := \int_\Omega (f - f_{\mathcal{T}}) w\, dx + \int_{\Gamma_{\mathrm{N}}} (g - g_{\mathcal{T}}) w\, ds$. Writing $\|\cdot\|_p$ for the norm in (any product of) $L^p(\Omega)$, we infer from Cauchy's inequality and with the definition of δ_1 that

$$\begin{aligned} &D\phi(z_{\mathcal{S}}; z_{\mathcal{S}} - \tilde{z}_{\mathcal{S}}) - D\phi_{\mathcal{T}}(z_{\mathcal{S}}; z_{\mathcal{S}} - \tilde{z}_{\mathcal{S}}) \\ &\leq \delta_1 \|\mathcal{Y}_{\mathcal{S}} - \tilde{\mathcal{Y}}_{\mathcal{S}}\|_2 \|\mathcal{Y}_{\mathcal{S}}\|_2 - \ell_{\mathcal{T}}(u_{\mathcal{S}} - \tilde{u}_{\mathcal{S}}). \end{aligned} \tag{6.8}$$

In step four, Jensen's inequality, $|\mathcal{M}p| \leq \mathcal{M}|p|$, $\mathcal{M}(\mathsf{id} - \mathcal{M}) = 0$, and $\psi_{\mathcal{T}}(\tilde{y}_{\mathcal{S}}) < \infty$ yield

$$\begin{aligned} \psi_{\mathcal{T}}(\tilde{y}_{\mathcal{S}}) - \psi(y) &= \int_\Omega (\sigma_{Y\mathcal{T}}|\mathcal{M}p| - \sigma_Y|p|)\, dx \\ &\leq \int_\Omega (\sigma_{Y\mathcal{T}} - \mathcal{M}\sigma_Y)|\mathcal{M}p|\, dx \\ &\quad + \int_\Omega (\sigma_Y - \mathcal{M}\sigma_Y)(\mathcal{M} - \mathsf{id})|p|\, dx \\ &\leq \delta_1 \|(\mathsf{id} - \mathcal{M})|p|\|_2 + \delta_2 \|p\|_2. \end{aligned} \tag{6.9}$$

In step five, we provide estimates for $\tilde{\alpha}_{\mathcal{S}} - \mathcal{M}\alpha$. On each element $T \in \mathcal{T}$, $\tilde{\alpha}_{\mathcal{S}} - \mathcal{M}\alpha$ is constant, non-positive, and either vanishes or satisfies

$$\begin{aligned} 0 > \tilde{\alpha}_{\mathcal{S}} - \mathcal{M}\alpha &= -\mathcal{M}\alpha - k_{\mathcal{T}}|\mathcal{M}p| \\ &\geq -\mathcal{M}\alpha - k_{\mathcal{T}}|\mathcal{M}p| + \alpha + k|p|, \end{aligned} \tag{6.10}$$

because $\psi(p, \xi) < \infty$. Taking averages in (6.10), we deduce

$$\tilde{\alpha}_{\mathcal{S}} - \mathcal{M}\alpha \geq -k_{\mathcal{T}}|\mathcal{M}p| + \mathcal{M}(k|p|). \tag{6.11}$$

According to Jensen's inequality again, this yields

$$0 \geq \tilde{\alpha}_{\mathcal{S}} - \mathcal{M}\alpha \geq (\overline{k} - k_{\mathcal{T}})\mathcal{M}|p| + \mathcal{M}((k - \overline{k})(\mathrm{id} - \mathcal{M})|p|), \tag{6.12}$$

where we over-line averages, such as $\overline{k} := \mathcal{M}k$ and below $\overline{\boldsymbol{B}} := \mathcal{M}\boldsymbol{B}$. Thus, in all cases we conclude

$$\| \tilde{\alpha}_{\mathcal{S}} - \mathcal{M}\alpha \|_2 \leq \delta_1 \| (\mathrm{id} - \mathcal{M})|p| \|_2 + \delta_2 \| p \|_2. \tag{6.13}$$

In step six, we provide estimates for $\tilde{\beta}_{\mathcal{S}} - \mathcal{M}\beta$. The definition of and the side restriction on ξ and the fact that $\mathcal{M}((\boldsymbol{B} - \overline{\boldsymbol{B}})\mathcal{M}p) = 0$ yield

$$\tilde{\beta}_{\mathcal{S}} - \mathcal{M}\beta = \mathcal{M}((\boldsymbol{B} - \overline{\boldsymbol{B}})(\mathrm{id} - \mathcal{M})p) + (\bar{\boldsymbol{B}} - \boldsymbol{B}_{\mathcal{T}})\mathcal{M}p, \quad \text{whence}$$

$$\| \tilde{\beta}_{\mathcal{S}} - \mathcal{M}\beta \|_2 \leq \delta_1 \| (\mathrm{id} - \mathcal{M})p \|_2 + \delta_2 \| p \|_2. \tag{6.14}$$

In step seven, we consider the stress field $\sigma := \boldsymbol{C}(\epsilon(u) - p)$ (resp. $\sigma_{\mathcal{S}} := \boldsymbol{C}(\epsilon(u_{\mathcal{S}}) - p_{\mathcal{S}})$) which is in equilibrium with applied forces and so

$$\begin{aligned} D\phi(z_{\mathcal{S}}; \tilde{z}_{\mathcal{S}} - z) = \int_\Omega (\sigma_{\mathcal{S}} - \sigma) : \epsilon(\tilde{u}_{\mathcal{S}} - u)\, dx + \int_\Omega \sigma_{\mathcal{S}} : (p - \tilde{p}_{\mathcal{S}})\, dx \\ + \int_\Omega \xi_{\mathcal{S}} \cdot \boldsymbol{H}(\tilde{\xi}_{\mathcal{S}} - \xi)\, dx. \end{aligned} \tag{6.15}$$

The first integral on the right-hand side of (6.15) will be estimated by Cauchy's inequality. According to the definition of $\tilde{p}_{\mathcal{S}}$, the second integral on the right-hand side of (6.15) equals

$$\int_\Omega \sigma_{\mathcal{S}} : (p - \tilde{p}_{\mathcal{S}})\, dx = \int_\Omega \sigma_{\mathcal{S}} : (\mathrm{id} - \mathcal{M})p\, dx = 0. \tag{6.16}$$

The third integral on the right-hand side of (6.15) can be estimated according to (6.13) and (6.14) and we obtain

$$\begin{aligned} &\int_\Omega \xi_{\mathcal{S}} \cdot \boldsymbol{H}(\tilde{\xi}_{\mathcal{S}} - \xi)\, dx \\ &= \int_\Omega \xi_{\mathcal{S}} \cdot \boldsymbol{H}(\tilde{\xi}_{\mathcal{S}} - \mathcal{M}\xi)\, dx \\ &\quad - \int_\Omega ((\mathrm{id} - \mathcal{M})\boldsymbol{H}\xi_{\mathcal{S}}) \cdot (\mathrm{id} - \mathcal{M})\xi\, dx \\ &\leq \| \boldsymbol{H}\xi_{\mathcal{S}} \|_2 \Big(\delta_1 \| (\mathrm{id} - \mathcal{M})(p, |p|) \|_2 + \delta_2 \| (p, |p|) \|_2 \Big) \\ &\quad + \delta_1 \| \xi_{\mathcal{S}} \|_2 \| (\mathrm{id} - \mathcal{M})\xi \|_2. \end{aligned} \tag{6.17}$$

In step eight, we find with (6.6) that

$$\psi(\tilde{y}) - \psi_{\mathcal{T}}(y_{\mathcal{S}}) = \int_\Omega (\mathcal{M}\sigma_Y - \sigma_{Y\mathcal{T}})|p_{\mathcal{S}}|\, dx \leq \delta_2 \| p_{\mathcal{S}} \|_2. \tag{6.18}$$

In step nine, we note that the stress field σ is in equilibrium with the applied forces and that $\tilde p = p_{\mathcal S}$. Hence, arguing as in step seven,

$$\begin{aligned} D\phi(z;\tilde z - z_{\mathcal S}) &= \int_\Omega \xi\cdot \boldsymbol{H}(\tilde\xi - \xi_{\mathcal S})\,dx \\ &= \int_\Omega (\boldsymbol{H}\xi - \mathcal{M}(\boldsymbol{H}\xi))\cdot \begin{pmatrix} (\overline{k}-k)|p_{\mathcal S}| \\ (\overline{\boldsymbol{B}}-\boldsymbol{B})p_{\mathcal S}\end{pmatrix} dx \\ &\quad + \int_\Omega \mathcal{M}(\boldsymbol{H}\xi)\cdot \begin{pmatrix} (k_{\mathcal T}-\overline{k})|p_{\mathcal S}| \\ (\boldsymbol{B}_{\mathcal T}-\overline{\boldsymbol{B}})p_{\mathcal S}\end{pmatrix} dx \\ &\le \Big\{\delta_1 \|\,(\mathsf{id}-\mathcal{M})\boldsymbol{H}\xi\,\|_2 + \delta_2\|\,\boldsymbol{H}\xi\,\|_2\Big\}\|\,(p_{\mathcal S},|p_{\mathcal S}|)\,\|_2. \end{aligned} \tag{6.19}$$

Furthermore, we have $\|\,(\mathsf{id}-\mathcal{M})\boldsymbol{H}\xi\,\|_2 \le \delta_1\|\,\xi\,\|_2 + \|\,\boldsymbol{H}\,\|_\infty \|\,(\mathsf{id}-\mathcal{M}) \times \xi\,\|_2$.

So far, we estimated all the terms in (6.2)–(6.4). In the final step ten, we put all those estimates together and, since

$$\begin{aligned} &D\phi(z;z-z_{\mathcal S}) - D\phi(z_{\mathcal S};z-z_{\mathcal S}) \\ &= \|\,(\boldsymbol{C}^{-1/2}(\sigma-\sigma_{\mathcal S}),\boldsymbol{H}^{1/2}(\xi-\xi_{\mathcal S}))\,\|_2^2, \end{aligned} \tag{6.20}$$

we obtain a constant $c_5 > 0$ such that

$$\begin{aligned} &c_5\|\,(\boldsymbol{C}^{-1/2}(\sigma-\sigma_{\mathcal S}),\boldsymbol{H}^{1/2}(\xi-\xi_{\mathcal S}))\,\|_2^2 \\ &\le \delta_1(\|\,\mathcal{Y}_{\mathcal S}-\tilde{\mathcal{Y}}_{\mathcal S}\,\|_2 + \|\,(\mathsf{id}-\mathcal{M})(p,|p|,\xi)\,\|_2)(1+\|\,\mathcal{Y}_{\mathcal S}\,\|_2) \\ &\quad +\delta_2\|\,(1,p,p_{\mathcal S},\xi,\xi_{\mathcal S})\,\|_2^2 + \ell(u_{\mathcal S}-\tilde u_{\mathcal S}) \\ &\quad +\|\,\boldsymbol{C}^{-1/2}(\sigma-\sigma_{\mathcal S})\,\|_2\|\,\boldsymbol{C}^{1/2}\epsilon(u-\tilde u_{\mathcal S})\,\|_2. \end{aligned} \tag{6.21}$$

According to the above estimates for $\tilde{\mathcal{Y}}_{\mathcal S} - \mathcal{M}\mathcal{Y}$ and $\|\,\mathcal{Y}_{\mathcal S}-\mathcal{M}\mathcal{Y}\,\|_2 \le \|\,\mathcal{Y}_{\mathcal S}-\mathcal{Y}\,\|_2 = \|\,(\boldsymbol{C}^{-1}(\sigma-\sigma_{\mathcal S}),\xi-\xi_{\mathcal S})\,\|_2$,

$$\begin{aligned} \|\,\mathcal{Y}_{\mathcal S}-\tilde{\mathcal{Y}}_{\mathcal S}\,\|_2 &\le \|\,\mathcal{M}\mathcal{Y}-\tilde{\mathcal{Y}}_{\mathcal S}\,\|_2 + \|\,\mathcal{Y}_{\mathcal S}-\mathcal{M}\mathcal{Y}\,\|_2 \\ &\le \|\,\epsilon(u-\tilde u_{\mathcal S})\,\|_2 + \delta_1\|\,(\mathsf{id}-\mathcal{M})(p,|p|)\,\|_2 \\ &\quad +\delta_2\|\,(p,|p|)\,\|_2 + \|\,(\boldsymbol{C}^{-1}(\sigma-\sigma_{\mathcal S}),\xi-\xi_{\mathcal S})\,\|_2. \end{aligned} \tag{6.22}$$

By incorporating (6.22) in (6.21) and by absorbing the factors $\|\,(\sigma-\sigma_{\mathcal S},\xi-\xi_{\mathcal S})\,\|_2$ on the right-hand side, we verify (4.3).

The proof of Theorem 4.1 is finished and it remains to prove Theorem 5.1.

In step eleven, we focus on (6.15) and, now, do not use equilibrium of σ with the applied forces. Hence,

$$\begin{aligned} D\phi(z_{\mathcal S};\tilde z_{\mathcal S}-z) &= \int_\Omega \sigma_{\mathcal S} : \epsilon(\tilde u_{\mathcal S}-u)\,dx - \int_\Omega f(\tilde u_{\mathcal S}-u)\,dx \\ &\quad - \int_{\Gamma_{\mathrm N}} g(\tilde u_{\mathcal S}-u)\,ds \end{aligned} \tag{6.23}$$

$$+ \int_\Omega \sigma_{\mathcal S} : (p-\tilde p_{\mathcal S})\,dx + \int_\Omega \xi_{\mathcal S}\cdot\boldsymbol{H}(\tilde\xi_{\mathcal S}-\xi)\,dx. \tag{6.24}$$

The terms in (6.24) are treated as in step seven, so we concentrate on the first term on the right-hand side in (6.23). The treatment of this term follows established techniques for pure elasticity, so we give only a sketch. Let $I_{\mathcal{T}}(u - u_{\mathcal{S}})$ be the Clement interpolant to $u - u_{\mathcal{S}}$ [6, 7, 3, 22] and let $\tilde{u}_{\mathcal{S}} := I_{\mathcal{T}}(u - u_{\mathcal{S}}) + u_{\mathcal{S}}$ such that we obtain a constant c_6, which depends on c_1 only,

$$\sum_{T \in \mathcal{T}} h_T^{-2} \| u - \tilde{u}_{\mathcal{S}} \|_{L^2(T)}^2 \leq c_6 \cdot \| \nabla(u - u_{\mathcal{S}}) \|_2, \tag{6.25}$$

$$\sum_{E \in \mathcal{E}} h_E^{-1} \| u - \tilde{u}_{\mathcal{S}} \|_{L^2(E)}^2 \leq c_6 \cdot \| \nabla(u - u_{\mathcal{S}}) \|_2, \tag{6.26}$$

$$\| \nabla I_{\mathcal{T}}(u - u_{\mathcal{S}}) \|_2 \leq c_6 \cdot \| \nabla(u - u_{\mathcal{S}}) \|_2. \tag{6.27}$$

Here, $\mathcal{E}$ denotes the set of all edges in $\mathcal{T}$. By elementwise integration by parts, we obtain from the strong form (2.1)–(2.2) that

$$\begin{aligned}
&\int_\Omega \sigma_{\mathcal{S}} : \epsilon(\tilde{u}_{\mathcal{S}} - u)\, dx - \int_\Omega f(\tilde{u}_{\mathcal{S}} - u)\, dx - \int_{\Gamma_{\mathrm{N}}} g(\tilde{u}_{\mathcal{S}} - u)\, ds \\
&= \int_{\cup \mathcal{E}} J(\sigma_{\mathcal{S}} \cdot n_E)(\tilde{u}_{\mathcal{S}} - u)\, ds - \int_{\cup \mathcal{T}} (f + \operatorname{div} \sigma_{\mathcal{S}})(\tilde{u}_{\mathcal{S}} - u)\, dx \\
&\leq \sum_{E \in \mathcal{E}} \| h_E^{1/2} J(\sigma_{\mathcal{S}} \cdot n_E) \|_{L^2(E)} \| h_E^{-1/2} (\tilde{u}_{\mathcal{S}} - u) \|_{L^2(E)} \\
&\quad + \sum_{T \in \mathcal{T}} \| h_T (f + \operatorname{div} \sigma_{\mathcal{S}}) \|_{L^2(T)} \| h_T^{-1} (\tilde{u}_{\mathcal{S}} - u) \|_{L^2(T)} \\
&\leq c_6 \| \nabla(u - u_{\mathcal{S}}) \|_2 \Big(\sum_{T \in \mathcal{T}} \| h_T (f + \operatorname{div} \sigma_{\mathcal{S}}) \|_{L^2(T)}^2 \\
&\quad + \sum_{E \in \mathcal{E}} \| h_E^{1/2} J(\sigma_{\mathcal{S}} \cdot n_E) \|_{L^2(E)}^2 \Big)^{1/2} \\
&\leq c_6 \| \nabla(u - u_{\mathcal{S}}) \|_2 (\sum_{T \in \mathcal{T}} \eta_T^2)^{1/2}. \qquad (6.28)
\end{aligned}$$

In step twelve, we incorporate the results of Theorem 4.1 and 4.2 and (as in (6.21)) obtain, with some $c_7 > 0$,

$$\begin{aligned}
&c_7 \| (p - p_{\mathcal{S}}, \epsilon(u - u_{\mathcal{S}}), \sigma - \sigma_{\mathcal{S}}, \xi - \xi_{\mathcal{S}}) \|_2^2 \\
&\leq \delta_1 (\| \mathcal{Y}_{\mathcal{S}} - \tilde{\mathcal{Y}}_{\mathcal{S}} \|_2 + \| (\mathsf{id} - \mathcal{M})(p, |p|, \xi) \|_2)(1 + \| \mathcal{Y}_{\mathcal{S}} \|_2) \\
&\quad + \delta_2 \| (1, p, p_{\mathcal{S}}, \xi, \xi_{\mathcal{S}}) \|_2^2 + \ell(I_{\mathcal{T}}(u - u_{\mathcal{S}})) \\
&\quad + \| \nabla(u - u_{\mathcal{S}}) \|_2 \cdot (\sum_{T \in \mathcal{T}} \eta_T^2)^{1/2}. \qquad (6.29)
\end{aligned}$$

Notice that $\ell(I_{\mathcal{T}}(u - u_{\mathcal{S}}))$ is bounded by the maximum in (5.2) times $\| \nabla I_{\mathcal{T}}(u - u_{\mathcal{S}}) \|_2$. Owing to Korn's inequality and (6.27), the latter term is bounded by $c_8 \| \epsilon(u - u_{\mathcal{S}}) \|_2$ and can be absorbed in (6.29). Then, the proof is concluded with the above arguments.

Acknowledgement. This work was initiated while the author was guest of the Forschungsinstitut für Mathematik, ETH Zürich, whose hospitality is gratefully acknowledged. The research was partly supported by the German Research Foundation (DFG) under Grant No. Ca 151/4.

References

1. Alberty, J., Carstensen, C. (1998): Numerical analysis of time-depending primal elastoplasticy with hardening. SIAM J. Numer. Anal. (accepted for publication). Berichtsreihe des Mathematischen Seminars Kiel, 98-18
2. Alberty, J., Carstensen, C., Zarrabi, D. (1998): Adaptive numerical analysis in primal elastoplasticity with hardening. Computer Methods in Applied Mechanics and Engineering
3. Brenner, S.C., Scott, L.R. (1994): The mathematical theory of finite element methods. Texts in Applied Mathematics **15**, Springer, New York
4. Carstensen, C. (1996): Coupling of FEM and BEM for interface problems in viscoplasticity and plasticity with hardening. SIAM J. Numer. Anal. **33**, 171–207
5. Carstensen, C. (1993): Interface problems in holonomic elastoplasticity. Math. Meth. Appl. Sci. **16**, 819–835
6. Ciarlet, P.G. (1978): The finite element method for elliptic problems. North-Holland, Amsterdam
7. Clement, P. (1975): Approximation by finite element functions using local regularization. RAIRO Sér. Rouge Anal. Numér. **R-2**, 77–84 (1975)
8. Ekeland, I., Temam, R. (1976): Convex analysis and variational problems. North-Holland, Amsterdam
9. Glowinski, R., Lions, J.L., Trémolières, R. (1981): Numerical analysis of variational inequalities. North-Holland, Amsterdam
10. Han, W. (1992): Finite element Analysis of a holonomic elastic-plastic problem. Numer. Math. **60**, 493–508
11. Han, W., Reddy, B.D. (1995): Computational plasticity: the variational basis and numerical analysis. Computational Mechanics Advances **2**, 285–400
12. Hörmander, L. (1963): Linear partial differential operators. Springer, Berlin Heidelberg New York
13. Johnson, C. (1976): Existence Theorems for Plasticity Problems. J. Math. pures et appl. **55**, 431–444
14. Johnson, C. (1977): A mixed finite element method for plasticity problems with hardening. SIAM J. Numer. Anal. **14**, 575–583
15. Johnson, C. (1978): On Plasticity with Hardening. Journ. of Math. Anal. and Appl. **62**, 325–336
16. Johnson, C., Hansbo, P.: Adaptive finite element methods in computational mechanics. Comput. Meth. Appl. Mech. Engin. **101**, 143–181
17. Lions, J.L., Magenes, E. (19): Non-homogeneous boundary value problems and applications. Vol. I. Springer, Berlin Heidelberg New York
18. Seregin, G.A. (1987): Differentiability of extremals of variational problems in the mechanics of perfect viscoplastic media. Diff. Equations **23**, 1349–1358

19. Seregin, G.A. (1993): On the regularity of the minimizers of some variational problems of plasticity theory. St. Petersburg Math. J. **4**, 989–1020
20. Suquet, P.-M. (1988): Dicontinuities and Plasticity. In Nonsmooth mechanics and applications, CISM Courses **302**, Moreau, J.J., Panagiotopoulos, P.D. (eds.) Springer, New York, 279–341
21. Temam, R. (1985): Mathematical problems in plasticity. Gauthier-Villars, Paris
22. Verfürth, R. (1996): A review of a posteriori error estimation and adaptive mesh-refinement techniques. Teubner-Wiley
23. Zeidler, E. (1985,1988): Nonlinear functional analysis and its applications III and IV. Springer, New York