ELSEVIER

Compput. Methods Appl. Mech. Engrg. 171 (1999) 327–340

Computer methods in applied mechanics and engineering

Convergence analysis of discrete approximations of problems in hardening plasticity

Weimin Han[a,*], B. Daya Reddy[b,1]

[a]Department of Mathematics, University of Iowa, Iowa City, IA 52242, USA

[b]Department of Mathematics and Applied Mathematics, University of Cape Town, 7700 Rondenbosch, South Africa

Received 1 April 1998; revised 18 May 1998

Abstract

The initial boundary value problem of quasistatic elastoplasticity is considered here as a variational inequality and equation in the displacement and stress. A variational inequality for the stress only may be obtained by eliminating the displacement. Semidiscrete approximations of the stress problem and fully discrete finite element approximations of the full problem are considered under assumptions of minimum regularity of the solution. It is shown that the resulting families of approximations converge to the solution of the original problem.

1. Introduction

Computational elastoplasticity is, at the present time, at a mature stage of development, and there is general agreement that most relevant computational issues have been settled, particularly in the regime of the infinitesimal theory. Thus, there exist at the present time, well-established procedures for obtaining solutions by means of finite element approximations, and a variety of algorithms exist for integration of the initial value problem. A comprehensive overview of the status quo, for both the infinitesimal and finite strain situations, may be found in the monograph by Simo and Hughes [7].

The mathematical theory of elastoplasticity has also seen some significant developments during the past two decades. The existence theory for the small strain perfectly plastic problem has been settled, with a number of authors contributing to the achievement of this goal; see, for example, the works of Anzellotti and Luckhaus [1] and Matthies [6].

Well-posedness of the initial boundary value problem of hardening elastoplasticity may be established within the framework of Sobolev spaces. The work by Han and Reddy [3] gives a comprehensive account of existence and uniqueness, convergence of semi- and fully discrete finite element approximations, as well as the convergence of some algorithms, for the case of hardening materials, and for infinitesimal strains. In that work the authors show that the initial boundary value problem of elastoplasticity may be formulated in two alternative variational forms, each being an evolutionary variational inequality. The two formulations arise from the use of the flow law in two forms: either in the form that uses the dissipation function, or in the more conventional form, which makes use of the yield function. These are referred to in that work as the primal and dual forms of the problem, respectively.

It is the dual problem that is of interest here; this is essentially a weak form of the equilibrium equation and

* Corresponding author. E-mail: whan@math.uiowa.edu

[1] E-mail: bdr@maths.uct.ac.za. The work of this author was supported by the Foundation for Research Development.

0045-7825/99/$ – see front matter
PII: S0045-7825(98)00214-X

normality law (see the next section). The principal unknown variables are the displacement $\boldsymbol{u}$ and generalized stress $\boldsymbol{\Sigma}$. Let $\Omega \subset \mathbb{R}^d$ be the initial configuration of the elastoplastic body. It has been shown in [3] that, under assumptions that are given in detail in Section 2, the problem has a unique solution $(\boldsymbol{u}, \boldsymbol{\Sigma})$ with $\boldsymbol{u} \in H^1(0,T;[H^1(\Omega)]^d)$ and $\boldsymbol{\Sigma} \in H^1(0,T;[L^2(\Omega)]^{d\times d})$. However, convergence of semi- and fully discrete finite element approximations has been established in [3] under the additional assumption that $\boldsymbol{\Sigma} \in W^{3,1}(0,T;[L^2(\Omega)]^{d\times d})$ and $\boldsymbol{u} \in W^{1,\infty}(0,T;[H^2(\Omega)]^d)$. The aim of this contribution is to establish the convergence of various discrete approximations, under the conditions of minimal regularity. We notice that in [5], convergence is proved for a fully discrete approximation for the dual problem with a linear isotropic hardening material. Our convergence analysis is given for more general hardening plastic materials.

The rest of this work is organized as follows. In Section 2 we formulate the variational problem and also summarize a number of results that will be subsequently required. Section 3 is devoted to a convergence analysis of time-discrete schemes of the stress problem, that is, the problem obtained when the displacement is eliminated as a variable. These approximations are based on a generalized midpoint rule. Finally, in Section 4 we consider fully discrete approximations of the full problem, and verify convergence under conditions of minimal regularity.

2. Formulation of the problem

2.1. Function spaces

Let $\Omega \subset \mathbb{R}^d$ be a nonempty open bounded set with a Lipschitz boundary. We will use the standard Lebesgue space $L^2(\Omega)$ and Sobolev spaces $H^k(\Omega)$, $k \geqslant 1$, and $H_0^1(\Omega)$. For a normed space V, we denote its topological dual by V'. The Cartesian product $V \times W$ of two normed spaces V and W is the space of all the ordered pairs (v, w) for $v \in V$ and $w \in W$, and

$$\|(v,w)\|_{V\times W} = \|v\|_V + \|w\|_W .$$

We will also need some vector-valued function spaces. For a normed space V and a positive number T, the space $C^m([0,T;V)$, $m \geqslant 0$, consists of all continuous functions u from $[0,T]$ to V that have continuous derivatives up to and including those of order m. Let

$$C^\infty([0,T];V) = \bigcap_{m=0}^{\infty} C^m([0,T];V) .$$

For $1 \leqslant p < \infty$ the space $L^p(0,T;V)$ consists of all measurable functions u from $[0,T]$ to V for which

$$\|u\|_{L^p(0,T;V)} \equiv \left(\int_0^T \|u(t)\|_V^p \,\mathrm{d}t\right)^{1/p} < \infty .$$

This is a Banach space with the norm $\|u\|_{L^p(0,T;X)}$, provided that the members are understood to represent equivalence classes of functions which are equal a.e. on $(0,T)$. Vector-valued Sobolev spaces $H^k(0,T;V)$, $k \geqslant 0$, are defined similarly.

We will need the following inequality

$$\|v(t) - v(s)\|_V \leqslant \int_s^t \|\dot{v}\|_V \quad \forall\, v \in H^1(0,T;V), \quad 0 \leqslant s \leqslant t \leqslant T .$$

Here, $\dot{v}$ denotes the time derivative of v. The continuous embedding property

$$H^1(0,T;V) \hookrightarrow C([0,T];V)$$

assures that for any $v \in H^1(0,T;V)$, the value $v|_{t=0}$ in V is well-defined.

Finally, we record a density result, a proof of which is found in [4].

THEOREM 2.1. Assume Ω has a Lipschitz continuous boundary $\partial\Omega$, $l,k \geqslant 0$. Then, the space $C^\infty([0,T];C^\infty(\bar{\Omega}))$ is dense in $H^l(0,T;H^k(\Omega))$.

2.2. Description of the problem

Consider the initial boundary value problem for quasistatic behavior of an elastoplastic body which occupies a bounded domain $\Omega \subset \mathbb{R}^d$ ($d \leq 3$ for practical applications) with Lipschitz boundary Γ. We assume that deformations are sufficiently small to warrant adoption of the small strain assumption. The plastic behavior of the material is assumed to be describable within the classical framework of a convex, elastic domain coupled with the normality law. The yield surface, which is the boundary of the elastic domain, need not be smooth, however. The material is assumed to undergo kinematic or isotropic hardening, or a combination of both, and the features of hardening behavior are captured through the introduction of a set of scalar and/or tensorial internal variables, denoted collectively here by an m-dimensional vector $\boldsymbol{\xi}$.

Suppose that the system is initially at rest, and that it is initially undeformed and unstressed. A time-dependent field of body force $\boldsymbol{f}(\boldsymbol{x}, t)$ is given, with $\boldsymbol{f}(\boldsymbol{x}, 0) = \boldsymbol{0}$. Then, the problem is governed by the following set of equations in Ω:

the equilibrium equation

$$\operatorname{div} \boldsymbol{\sigma} + \boldsymbol{f} = \boldsymbol{0}, \tag{2.1}$$

the additive decomposition of strain

$$\boldsymbol{\epsilon} = \boldsymbol{e} + \boldsymbol{p}, \tag{2.2}$$

and the strain–displacement relation

$$\boldsymbol{\epsilon}(\boldsymbol{u}) = \frac{1}{2}(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^{\mathrm{T}}). \tag{2.3}$$

Here, $\boldsymbol{\sigma}$ is the stress tensor, $\boldsymbol{\epsilon}$ is the strain tensor, $\boldsymbol{u}$ the displacement vector, $\boldsymbol{p} = \boldsymbol{p}(\boldsymbol{\xi})$ the plastic strain tensor and $\boldsymbol{e}$ the elastic strain. All the tensors encountered here are symmetric. The plastic deformation is assumed to be incompressible so that

$$\operatorname{tr} \boldsymbol{p} = \boldsymbol{0} \quad \text{or} \quad \sum_{i=1}^{d} p_{ii} = 0. \tag{2.4}$$

For simplicity, and with little loss in generality, we take the boundary condition to be the homogeneous Dirichlet condition

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{on } \Gamma, \tag{2.5}$$

while the initial conditions are assumed to be

$$\boldsymbol{u}(\boldsymbol{x}, 0) = \boldsymbol{0} \quad \text{and} \quad \boldsymbol{\sigma}(\boldsymbol{x}, 0) = \boldsymbol{0}. \tag{2.6}$$

A complete description of the problem requires that a set of constitutive equations be added to (2.1)–(2.6). We have a linear relation between the stress and the elastic strain $\boldsymbol{e} = \boldsymbol{\epsilon}(\boldsymbol{u}) - \boldsymbol{p}$:

$$\boldsymbol{\sigma} = \boldsymbol{C}(\boldsymbol{\epsilon}(\boldsymbol{u}) - \boldsymbol{p}) \tag{2.7}$$

where $\boldsymbol{C}$ is the elasticity modulus. We also introduce a stress-like variable $\boldsymbol{\chi}$ conjugate to the internal variable $\boldsymbol{\xi}$, which is defined by

$$\boldsymbol{\chi} = -\boldsymbol{H}\boldsymbol{\xi} \tag{2.8}$$

for some hardening modulus $\boldsymbol{H}$. We call the ordered pairs $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi})$ and $\boldsymbol{P} = (\boldsymbol{p}, \boldsymbol{\xi})$ the *generalized stress* and *generalized plastic strain*. The relationship between $\boldsymbol{\xi}$ and $\boldsymbol{\chi}$ in a thermodynamic context is discussed in [3].

The generalized stress takes values only in a closed convex set K; the interior of K contains the origin and is called the elastic region, while its boundary is known as the yield surface. The yield surface may be represented by the level set of a continuous convex function ϕ called the yield function, with K then being defined by

$$K = \{\boldsymbol{\Sigma} \in M^{d \times d} \times \mathbb{R}^m : \phi(\boldsymbol{\Sigma}) \leq 0\}. \tag{2.9}$$

Here and elsewhere, $M^{d \times d}$ denotes the space of $d \times d$ symmetric second-order tensors or matrices.

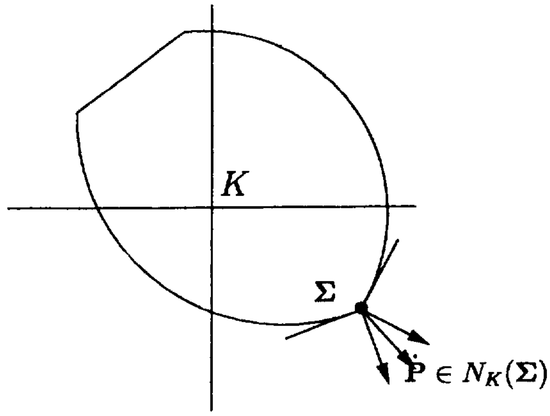

Fig. 1. The yield surface and normality law, in generalized stress space.

The evolution of the generalized plastic strains is governed by the normality law, which takes the form

$$\dot{\boldsymbol{P}} \equiv (\dot{\boldsymbol{p}}, \dot{\boldsymbol{\xi}}) \in N_K(\boldsymbol{\Sigma}) . \tag{2.10}$$

Here, $N_K(\boldsymbol{\Sigma}) = \{\boldsymbol{M} : \boldsymbol{M} \cdot (\boldsymbol{T} - \boldsymbol{\Sigma}) \leq 0 \ \forall \boldsymbol{T} \in K\}$ denotes the normal cone to K at $\boldsymbol{\Sigma}$ (see Fig. 1).

An alternative way of describing the flow law is by using the support function of K, defined by

$$D(\dot{\boldsymbol{P}}) = \sup\{\boldsymbol{T} \cdot \dot{\boldsymbol{P}} : \boldsymbol{T} \in K\} . \tag{2.11}$$

The function D is non-negative and may take on the value $+\infty$. Then from the theory of convex analysis, the flow law (2.10) is equivalent to the relation (cf. [3])

$$\boldsymbol{\Sigma} \in \partial D(\dot{\boldsymbol{P}}) , \tag{2.12}$$

where $\partial D(\dot{\boldsymbol{P}}) = \{(\boldsymbol{\sigma}, \boldsymbol{\chi})\}$ denotes the subdifferential of D at $\dot{\boldsymbol{P}}$, defined by

$$D(\boldsymbol{q}, \boldsymbol{\eta}) \geq D(\dot{\boldsymbol{p}}, \dot{\boldsymbol{\xi}}) + \boldsymbol{\sigma} \cdot (\boldsymbol{q} - \dot{\boldsymbol{p}}) + \boldsymbol{\chi} \cdot (\boldsymbol{\eta} - \dot{\boldsymbol{\xi}}) \quad \forall (\boldsymbol{q}, \boldsymbol{\eta}) . \tag{2.13}$$

In the context of plasticity, the function D is a measure of the rate of irreversible or plastic work, and is known as the dissipation function.

Two possible variational formulations of the initial boundary value problem for elastoplasticity may be obtained from the set of equations summarized here: one, which we refer to as the *primal* formulation, is based on the flow law in the form (2.12), while the other is based on the dual form (2.10) of this law, and is therefore referred to as the *dual* formulation. We focus on the dual formulation in this work. The unknown variables are the generalized stress $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi})$ and displacement $\boldsymbol{u}$. The space V of displacements is defined by

$$V = [H_0^1(\Omega)]^d , \tag{2.14}$$

the space of stresses by

$$S = \{\boldsymbol{\tau} = (\tau_{ij})_{d \times d} : \tau_{ji} = \tau_{ij}, \ \tau_{ij} \in L^2(\Omega)\} , \tag{2.15}$$

and the space M of conjugate forces by

$$M = \{\boldsymbol{\mu} = (\mu_j) : \mu_j \in L^2(\Omega), \ j = 1, \ldots, m\} . \tag{2.16}$$

Further, we set

$$\mathcal{T} = S \times M .$$

This space is endowed with the inner products induced by the natural inner products on S and M.

Admissible generalized stresses are those that belong to the set K pointwise. We accordingly define the convex subset

$$\mathcal{P} = \{\boldsymbol{T} = (\boldsymbol{\tau}, \boldsymbol{\mu}) \in \mathcal{T} : (\boldsymbol{\tau}(\boldsymbol{x}), \boldsymbol{\mu}(\boldsymbol{x})) \in K \text{ a.e. in } \Omega\} .$$

2.3. Properties of the material parameters

The elasticity tensor $\boldsymbol{C}$ has the symmetry properties

$$C_{ijkl} = C_{jikl} = C_{klij} , \tag{2.17}$$

and we assume that

$$C_{ijkl} \in L^{\infty}(\Omega) \tag{2.18}$$

and that $\boldsymbol{C}$ is pointwise stable: there exists a constant $C_0 > 0$ such that

$$C_{ijkl}(\boldsymbol{x})\zeta_{ij}\zeta_{kl} \geq C_0|\boldsymbol{\zeta}|^2 \quad \forall\, \boldsymbol{\zeta} \in M^{d\times d} , \quad \text{a.e. in } \Omega . \tag{2.19}$$

The compliance tensor $\boldsymbol{C}^{-1}$ has the same symmetry properties as $\boldsymbol{C}$, and is also pointwise stable in the sense that a constant $C_0' > 0$ exists such that

$$C^{-1}_{ijkl}(\boldsymbol{x})\zeta_{ij}\zeta_{kl} \geq C_0'|\boldsymbol{\zeta}|^2 \quad \forall\, \boldsymbol{\zeta} \in M^{d\times d} , \quad \text{a.e. in } \Omega . \tag{2.20}$$

The hardening modulus $\boldsymbol{H}$, viewed as a linear operator from $\mathbb{R}^m$ into itself, is assumed to possess the symmetry property

$$\boldsymbol{\xi} \cdot \boldsymbol{H}\boldsymbol{\lambda} = \boldsymbol{\lambda} \cdot \boldsymbol{H}\boldsymbol{\xi} , \tag{2.21}$$

and it is further assumed that

$$H_{ij} \in L^{\infty}(\Omega) \tag{2.22}$$

and that a constant $H > 0$ exists such that

$$\boldsymbol{\xi} \cdot \boldsymbol{H}\boldsymbol{\xi} \geq H|\boldsymbol{\xi}|^2 \quad \text{for all } \boldsymbol{\xi} \in \mathbb{R}^m , \quad \text{a.e. in } \Omega . \tag{2.23}$$

The inverse $\boldsymbol{H}^{-1}$ of the hardening modulus possesses the same properties as $\boldsymbol{H}$: it is a symmetric operator whose matrix representation has uniformly bounded components. Furthermore, there exists a constant $H' > 0$ such that

$$\boldsymbol{\chi} \cdot \boldsymbol{H}^{-1}\boldsymbol{\chi} \geq H'|\boldsymbol{\chi}|^2 \quad \text{for all } \boldsymbol{\chi} \in \mathbb{R}^m , \quad \text{a.e. in } \Omega . \tag{2.24}$$

2.4. The variational formulations

We now introduce the bilinear forms

$$a : S \times S \to \mathbb{R} , \quad a(\boldsymbol{\sigma}, \boldsymbol{\tau}) = \int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{C}^{-1}\boldsymbol{\tau}\, \mathrm{d}x , \tag{2.25}$$

$$b : V \times S \to \mathbb{R} , \quad b(\boldsymbol{v}, \boldsymbol{\tau}) = -\int_{\Omega} \boldsymbol{\epsilon}(\boldsymbol{v}) : \boldsymbol{\tau}\, \mathrm{d}x , \tag{2.26}$$

$$c : M \times M \to \mathbb{R} , \quad c(\boldsymbol{\chi}, \boldsymbol{\mu}) = \int_{\Omega} \boldsymbol{\chi} \cdot \boldsymbol{H}^{-1}\boldsymbol{\mu}\, \mathrm{d}x , \tag{2.27}$$

and

$$A : \mathcal{T} \times \mathcal{T} \to \mathbb{R} , \quad A(\boldsymbol{\Sigma}, \boldsymbol{T}) = a(\boldsymbol{\sigma}, \boldsymbol{\tau}) + c(\boldsymbol{\chi}, \boldsymbol{\mu}) \tag{2.28}$$

for $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi})$ and $\boldsymbol{T} = (\boldsymbol{\tau}, \boldsymbol{\mu})$. Here, $\boldsymbol{C}^{-1}$ is the *compliance tensor*, which is inverse to the elasticity tensor $\boldsymbol{C}$ in the sense that

$$\boldsymbol{C}^{-1}[\boldsymbol{C}\boldsymbol{\epsilon}] = \boldsymbol{\epsilon} \quad \text{and} \quad \boldsymbol{C}[\boldsymbol{C}^{-1}\boldsymbol{\sigma}] = \boldsymbol{\sigma}$$

for all symmetric matrices or second-order tensors $\boldsymbol{\epsilon}$ and $\boldsymbol{\sigma}$.

We will also need the linear functional

$$l(t) : V \to \mathbb{R}, \quad \langle l(t), \boldsymbol{v} \rangle = -\int_\Omega \boldsymbol{f}(t) \cdot \boldsymbol{v} \, dx . \tag{2.29}$$

We are now in a position to define the variational problem

PROBLEM DUAL. Given $l \in H^1(0, T; V')$, $l(0) = 0$, find $(\boldsymbol{u}, \boldsymbol{\Sigma}) = (\boldsymbol{u}, \boldsymbol{\sigma}, \boldsymbol{\chi}) : [0, T] \to V \times \mathcal{P}$ with $(\boldsymbol{u}(0), \boldsymbol{\Sigma}(0)) = (\mathbf{0}, \mathbf{0})$, such that for almost all $t \in (0, T)$,

$$b(\boldsymbol{v}, \boldsymbol{\sigma}(t)) = \langle l(t), \boldsymbol{v} \rangle \quad \forall \, \boldsymbol{v} \in V , \tag{2.30}$$

$$A(\dot{\boldsymbol{\Sigma}}(t), \boldsymbol{T} - \boldsymbol{\Sigma}(t)) + b(\dot{\boldsymbol{u}}(t), \boldsymbol{\tau} - \boldsymbol{\sigma}(t)) \geq 0 \quad \forall \, \boldsymbol{T} = (\boldsymbol{\tau}, \boldsymbol{\mu}) \in \mathcal{P} . \tag{2.31}$$

Problem DUAL is easily obtained from the governing equations (2.1)–(2.10) (see [3]).

Introducing the space

$$\mathcal{P}(t) = \{ \boldsymbol{T} = (\boldsymbol{\tau}, \boldsymbol{\mu}) \in \mathcal{P} : b(\boldsymbol{v}, \boldsymbol{\tau}) = \langle l(t), \boldsymbol{v} \rangle \ \forall \, \boldsymbol{v} \in V \} ,$$

we can eliminate the variable $\dot{\boldsymbol{u}}(t)$ from Problem DUAL to obtain the following stress problem.

PROBLEM STRESS. Given $l \in H^1(0, T; V')$, $l(0) = 0$, find $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi}) : [0, T] \to \mathcal{P}$ with $\boldsymbol{\Sigma}(0) = \mathbf{0}$, such that for almost all $t \in (0, T)$, $\boldsymbol{\Sigma}(t) \in \mathcal{P}(t)$ and

$$A(\dot{\boldsymbol{\Sigma}}(t), \boldsymbol{T} - \boldsymbol{\Sigma}(t)) \geq 0 \quad \forall \, \boldsymbol{T} = (\boldsymbol{\tau}, \boldsymbol{\mu}) \in \mathcal{P}(t) . \tag{2.32}$$

It has been shown in [3] that $A(\cdot, \cdot)$ is continuous and $\mathcal{T}$-elliptic, that is, there exist positive constants α_A and β_A such that

$$\begin{aligned} |A(\boldsymbol{\Sigma}, \boldsymbol{T})| &\leq \alpha_A \|\boldsymbol{\Sigma}\|_{\mathcal{T}} \|\boldsymbol{T}\|_{\mathcal{T}} \quad \text{for all } \boldsymbol{\Sigma}, \boldsymbol{T} \in \mathcal{T} , \\ A(\boldsymbol{T}, \boldsymbol{T}) &\geq \beta_A \|\boldsymbol{T}\|^2_{\mathcal{T}} \quad \text{for all } \boldsymbol{T} \in \mathcal{T} , \end{aligned} \tag{2.33}$$

and that the problem STRESS has a unique solution $\boldsymbol{\Sigma} \in H^1(0, T; \mathcal{T})$ under the following assumption, commonly known as the safe load condition.

ASSUMPTION 2.2. For any $\boldsymbol{\Sigma}_1 = (\boldsymbol{\sigma}_1, \boldsymbol{\chi}_1) \in \mathcal{P}$, and for any $\boldsymbol{\sigma}_2 \in S$, there exists $\boldsymbol{\chi}_2 \in M$ such that $|\boldsymbol{\chi}_2| \leq c|\boldsymbol{\sigma}_2|$ for some constant $c > 0$ independent of $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\sigma}_2$, and such that $\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2 \in \mathcal{P}$, with $\boldsymbol{\Sigma}_2 = (\boldsymbol{\sigma}_2, \boldsymbol{\chi}_2)$.

With the following additional assumption, the problem DUAL is shown in [3] to have a unique solution $(\boldsymbol{u}, \boldsymbol{\Sigma}) \in H^1(0, T; V \times \mathcal{T})$.

ASSUMPTION 2.3. For any $\boldsymbol{\Sigma} \in K$, and any $\kappa \in [0, 1)$, we have $\kappa \boldsymbol{\Sigma} \in K$ and

$$\inf_{x \in \Omega} \operatorname{dist}(\kappa \boldsymbol{\Sigma}(x), \partial K) > 0 .$$

Both of these assumptions are easily verified for materials undergoing combined linear kinematic and isotropic hardening, or linear kinematic or isotropic hardening only.

The next result plays an important role in the analysis of well-posedness, as well as in the convergence analysis to be presented here.

PROPOSITION 2.4. Let V and S be two Hilbert spaces. Let $b : V \times S \to \mathbb{R}$ be a continuous bilinear form. Define two bounded linear operators $B : S \to V'$ and $B' : V \to S'$ by

$$b(v, s) = \langle Bs, v \rangle = \langle B'v, s \rangle \quad \text{for } v \in V, \ s \in S .$$

Then, the following statements are equivalent:

(a) the bilinear form $b(\cdot, \cdot)$ satisfies the Babuška–Brezzi condition

$$\sup_{0 \neq s \in S} \frac{|b(v, s)|}{\|s\|_S} \geq c_0 \|v\|_V \quad \forall \, v \in V ;$$

(b) the operator B is an isomorphism from $(\operatorname{Ker} B)^{\perp}$ *onto* V', *where*

$$\operatorname{Ker} B = \{s \in S : b(v, s) = 0 \ \forall\, v \in V\} ;$$

(c) the operator B′ is an isomorphism from V onto $(\operatorname{Ker} B)^{\circ}$, *where*

$$(\operatorname{Ker} B)^{\circ} = \{f \in S' : \langle f, s \rangle = 0 \ \forall\, s \in \operatorname{Ker} B\} .$$

It has been shown in [3] that

$$\sup_{0 \neq \tau \in S} \frac{|b(\boldsymbol{v}, \boldsymbol{\tau})|}{\|\boldsymbol{\tau}\|_S} \geqslant \beta_b \|\boldsymbol{v}\|_V \quad \forall\, \boldsymbol{v} \in V \tag{2.34}$$

for some constant $\beta_b > 0$, and for V and S as defined in (2.14) and (2.15). Thus, the equivalent statements in Proposition 2.4 hold for the bilinear form (2.28).

The following elementary result will be used repeatedly:

$$a, b, x \geqslant 0 \quad \text{and} \quad x^2 \leqslant ax + b \quad \Rightarrow \quad x^2 \leqslant a^2 + 2b . \tag{2.35}$$

3. Time-discrete approximations of the stress problem

3.1. A family of generalized mid-point schemes

We partition the time interval $[0, T]$ into N equal parts, and let $k = T/N$ denote the stepsize. The partition points are $t_n = nk$, $n = 0, 1, \ldots, N$. Let θ be a real parameter. A family of generalized mid-point time-discrete approximations of the problem STRESS is then given by the following.

PROBLEM STRESS$_{k,\theta}$. Find a sequence $\{\boldsymbol{\Sigma}_n^k = (\boldsymbol{\sigma}_n^k, \boldsymbol{\chi}_n^k)\}_{n=0}^N \subset \mathcal{T}$ with $\boldsymbol{\Sigma}_0^k = \boldsymbol{0}$, such that for $n = 1, 2, \ldots, N$, $\boldsymbol{\Sigma}_{n-1+\theta}^k = \theta \boldsymbol{\Sigma}_n^k + (1-\theta)\boldsymbol{\Sigma}_{n-1}^k \in \mathcal{P}_{n-1+\theta}$, and

$$A(\Delta \boldsymbol{\Sigma}_n^k, \boldsymbol{T} - \boldsymbol{\Sigma}_{n-1+\theta}^k) \geqslant 0 \quad \forall\, \boldsymbol{T} \in \mathcal{P}_{n-1+\theta} , \tag{3.1}$$

where $\Delta \boldsymbol{\Sigma}_n^k = \boldsymbol{\Sigma}_n^k - \boldsymbol{\Sigma}_{n-1}^k$; the constraint set $\mathcal{P}_{n-1+\theta}$ is defined by

$$\mathcal{P}_{n-1+\theta} \equiv \mathcal{P}(t_{n-1+\theta}) = \{\boldsymbol{T} = (\boldsymbol{\tau}, \boldsymbol{\mu}) \in \mathcal{P} : b(\boldsymbol{v}, \boldsymbol{\tau}) = \langle l(t_{n-1+\theta}), \boldsymbol{v} \rangle \ \forall\, \boldsymbol{v} \in V\} .$$

It has been shown in [3] that for $\theta \in [\frac{1}{2}, 1]$, the problem STRESS$_{k,\theta}$ admits a unique solution, which satisfies the inequality

$$\begin{aligned} \max_{0 \leqslant n \leqslant N} \|\boldsymbol{\Sigma}(t_n) - \boldsymbol{\Sigma}_n^k\|_{\mathcal{T}}^2 \leqslant c \Bigg(& \sum_{j=1}^{N-1} \|E_{j,\theta}(\boldsymbol{\Sigma}) - E_{j+1,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} + \|E_{N,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} + k \sum_{j=1}^{N} \|\delta \boldsymbol{\Sigma}(t_j) - \dot{\boldsymbol{\Sigma}}(t_{j-1+\theta})\|_{\mathcal{T}} \Bigg)^2 \\ & + ck \sum_{j=1}^{N} \|\delta \boldsymbol{\Sigma}(t_j) - \dot{\boldsymbol{\Sigma}}(t_{j-1+\theta})\|_{\mathcal{T}} \|E_{j,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} , \end{aligned} \tag{3.2}$$

where

$$E_{n,\theta}(\boldsymbol{\Sigma}) = \theta \boldsymbol{\Sigma}(t_n) + (1-\theta)\boldsymbol{\Sigma}(t_{n-1}) - \boldsymbol{\Sigma}(t_{n-1+\theta}) . \tag{3.3}$$

It is also shown in [3] that when $\theta \not\in [\frac{1}{2}, 1]$, the scheme is divergent. So from now on, we will always assume $\theta \in [\frac{1}{2}, 1]$.

The inequality (3.2) is the basis for optimal order error estimates when the solution is assumed to be sufficiently smooth. However, for the case in which one assumes only the minimum degree of regularity under which well-posedness is proved, that is, $(\boldsymbol{u}, \boldsymbol{\Sigma}) \in H^1(0, T; V \times \mathcal{T})$, (3.2) is not useful for convergence analysis. We accordingly modify the estimate (3.2).

We set $\boldsymbol{e}_n = \boldsymbol{\Sigma}(t_n) - \boldsymbol{\Sigma}_n^k$; then since $\theta \in [\frac{1}{2}, 1]$, we have

$$A(\boldsymbol{e}_n - \boldsymbol{e}_{n-1}, \theta \boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1}) \geqslant \frac{1}{2}(\|\boldsymbol{e}_n\|_A^2 - \|\boldsymbol{e}_{n-1}\|_A^2) , \tag{3.4}$$

where $\|\boldsymbol{T}\|_A = \sqrt{A(\boldsymbol{T}, \boldsymbol{T})}$ defines a norm on $\mathcal{T}$ which is equivalent to the norm $\|\boldsymbol{T}\|_{\mathcal{T}}$.

On the other hand, with $\Delta\boldsymbol{\Sigma}(t_n) = \boldsymbol{\Sigma}(t_n) - \boldsymbol{\Sigma}(t_{n-1})$ and $\Delta\boldsymbol{\Sigma}_n = \boldsymbol{\Sigma}_n^k - \boldsymbol{\Sigma}_{n-1}^k$, we have

$$\begin{aligned} A(\boldsymbol{e}_n - \boldsymbol{e}_{n-1}, \theta\boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1}) &= A(\Delta\boldsymbol{\Sigma}(t_n), \theta\boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1}) - A(\Delta\boldsymbol{\Sigma}_n^k, \theta\boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1}) \\ &= A(\Delta\boldsymbol{\Sigma}(t_n), \theta\boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1}) - A(\Delta\boldsymbol{\Sigma}_n^k, \boldsymbol{\Sigma}(t_{n-1+\theta}) - \boldsymbol{\Sigma}_{n-1+\theta}^k) - A(\Delta\boldsymbol{\Sigma}_n^k, E_{n,\theta}(\boldsymbol{\Sigma}))\,. \end{aligned}$$

Setting $\boldsymbol{T} = \boldsymbol{\Sigma}(t_{n-1+\theta})$ in (3.1), we have

$$-A(\Delta\boldsymbol{\Sigma}_n^k, \boldsymbol{\Sigma}(t_{n-1+\theta}) - \boldsymbol{\Sigma}_{n-1+\theta}^k) \leq 0\,.$$

Thus

$$\begin{aligned} A(\boldsymbol{e}_n - \boldsymbol{e}_{n-1}, \theta\boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1}) &\leq A(\Delta\boldsymbol{\Sigma}(t_n), \theta\boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1}) - A(\Delta\boldsymbol{\Sigma}_n^k, E_{n,\theta}(\boldsymbol{\Sigma})) \\ &= A(\Delta\boldsymbol{\Sigma}(t_n), \theta\boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1}) + A(\Delta\boldsymbol{e}_n, E_{n,\theta}(\boldsymbol{\Sigma})) - A(\Delta\boldsymbol{\Sigma}_n, E_{n,\theta}(\boldsymbol{\Sigma})) \end{aligned} \tag{3.5}$$

Now, for any $t \in I_n = [t_{n-1}, t_n]$, from Proposition 2.4 we have the existence of a unique $\boldsymbol{\tau}_\delta(t) \in (\operatorname{Ker} B)^\perp$ such that

$$b(\boldsymbol{v}, \boldsymbol{\tau}_\delta(t)) = \langle l(t) - l_{n-1+\theta}, \boldsymbol{v}\rangle \quad \forall\, \boldsymbol{v} \in V$$

and

$$\|\boldsymbol{\tau}_\delta(t)\|_S \leq c\|l(t) - l_{n-1+\theta}\|_{V'}\,.$$

Here, and in what follows, a subscript δ indicates that the quantity is associated with a difference. By Assumption 2.2, we can find $\boldsymbol{\mu}_\delta(t) \in M$ such that $\boldsymbol{T}_\delta(t) = (\boldsymbol{\tau}_\delta(t), \boldsymbol{\mu}_\delta(t)) \in \mathscr{P}(t)$ and

$$\|\boldsymbol{T}_\delta(t)\|_{\mathscr{T}} \leq c\|l(t) - l_{n-1+\theta}\|_{V'}\,. \tag{3.6}$$

Set $\boldsymbol{T}(t) = \boldsymbol{\Sigma}_{n-1+\theta}^k + \boldsymbol{T}_\delta(t)$. Obviously, $\boldsymbol{T}(t) \in \mathscr{P}(t)$. We take $\boldsymbol{T} = \boldsymbol{T}(t)$ in (2.32) to obtain

$$A(\dot{\boldsymbol{\Sigma}}(t), \boldsymbol{T}_\delta(t) + \boldsymbol{\Sigma}_{n-1+\theta}^k - \boldsymbol{\Sigma}(t)) \geq 0\,,$$

that is,

$$A(\dot{\boldsymbol{\Sigma}}(t), \theta\boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1}) \leq A(\dot{\boldsymbol{\Sigma}}(t), \boldsymbol{T}_\delta(t)) + A(\dot{\boldsymbol{\Sigma}}(t), \theta\boldsymbol{\Sigma}(t_n) + (1-\theta)\boldsymbol{\Sigma}(t_{n-1}) - \boldsymbol{\Sigma}(t))\,.$$

By integrating this relation over I_n we obtain

$$A(\Delta\boldsymbol{\Sigma}(t_n), \theta\boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1}) \leq \int_{I_n} A(\dot{\boldsymbol{\Sigma}}(t), \boldsymbol{T}_\delta(t))\,dt + \int_{I_n} A(\dot{\boldsymbol{\Sigma}}(t), \theta\boldsymbol{\Sigma}(t_n) + (1-\theta)\boldsymbol{\Sigma}(t_{n-1}) - \boldsymbol{\Sigma}(t))\,dt\,,$$

which may be used in (3.5) to yield

$$\begin{aligned} A(\boldsymbol{e}_n - \boldsymbol{e}_{n-1}, \theta\boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1}) \leq{}& \int_{I_n} A(\dot{\boldsymbol{\Sigma}}(t), \boldsymbol{T}_\delta(t))\,dt + \int_{I_n} A(\dot{\boldsymbol{\Sigma}}(t), \boldsymbol{\Sigma}(t_{n-1+\theta}) - \boldsymbol{\Sigma}(t))\,dt \\ &+ A(\Delta\boldsymbol{e}_n, E_{n,\theta}(\boldsymbol{\Sigma}))\,. \end{aligned} \tag{3.7}$$

Henceforth, to simplify the writing, we introduce the moduli of continuity

$$\omega_k(l) = \sup\{\|l(s) - l(t)\|_{V'} : 0 \leq s, t \leq T, |t-s| \leq k\}\,, \tag{3.8}$$

$$\omega_k(\boldsymbol{\Sigma}) = \sup\{\|\boldsymbol{\Sigma}(s) - \boldsymbol{\Sigma}(t)\|_{\mathscr{T}} : 0 \leq s, t \leq T, |t-s| \leq k\}\,. \tag{3.9}$$

Note that $l \in H^1(0,T;V')$ and $\boldsymbol{\Sigma} \in H^1(0,T;\mathscr{T})$ are uniformly continuous with respect to $t \in [0,T]$. Hence, $\omega_k(l) \to 0$ and $\omega_k(\boldsymbol{\Sigma}) \to 0$ as $k \to 0$.

Combining (3.4) and (3.7), we now have

$$\|\boldsymbol{e}_n\|_A^2 - \|\boldsymbol{e}_{n-1}\|_A^2 \leq c(\omega_k(l) + \omega_k(\boldsymbol{\Sigma})) \int_{I_n} \|\dot{\boldsymbol{\Sigma}}(t)\|_{\mathcal{T}} \,\mathrm{d}t + 2A(\boldsymbol{e}_n - \boldsymbol{e}_{n-1}, E_{n,\theta}(\boldsymbol{\Sigma}))\,.$$

Applying this inequality recursively and recalling that $\boldsymbol{e}_0 = \mathbf{0}$, we see that

$$\begin{aligned}\|\boldsymbol{e}_n\|_A^2 &\leq c(\omega_k(l) + \omega_k(\boldsymbol{\Sigma})) \int_0^{t_n} \|\dot{\boldsymbol{\Sigma}}(t)\|_{\mathcal{T}} \,\mathrm{d}t + 2\sum_{j=1}^{n} A(\boldsymbol{e}_j - \boldsymbol{e}_{j-1}, E_{j,\theta}(\boldsymbol{\Sigma}))\\ &= c(\omega_k(l) + \omega_k(\boldsymbol{\Sigma})) \int_0^{t_n} \|\dot{\boldsymbol{\Sigma}}(t)\|_{\mathcal{T}} \,\mathrm{d}t + 2A(\boldsymbol{e}_n, E_{n,\theta}(\boldsymbol{\Sigma})) + 2\sum_{j=1}^{n} A(\boldsymbol{e}_j, E_{j,\theta}(\boldsymbol{\Sigma}) - E_{j+1,\theta}(\boldsymbol{\Sigma}))\,.\end{aligned}$$

Since $\|\cdot\|_A$ is an equivalent norm on $\mathcal{T}$, with

$$M = \max_{0 \leq n \leq N} \|\boldsymbol{e}_n\|_{\mathcal{T}}\,,$$

we then have

$$M^2 \leq c(\omega_k(l) + \omega_k(\boldsymbol{\Sigma}))\|\dot{\boldsymbol{\Sigma}}\|_{L^1(0,T;\mathcal{T})} + c\left(\|E_{N,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} + \sum_{n=1}^{N-1} \|E_{n,\theta}(\boldsymbol{\Sigma}) - E_{n+1,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}}\right)M\,.$$

Now, using the inequality (2.35), we obtain

$$M \leq c\{(\omega_k(l) + \omega_k(\boldsymbol{\Sigma}))\|\dot{\boldsymbol{\Sigma}}\|_{L^1(0,T;\mathcal{T})}\}^{1/2} + c\left\{\|E_{N,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} + \sum_{n=1}^{N-1} \|E_{n,\theta}(\boldsymbol{\Sigma}) - E_{n+1,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}}\right\}. \tag{3.10}$$

By Theorem 2.1, for any $\varepsilon > 0$ we have $\overline{\boldsymbol{\Sigma}} \in C^\infty([0,T];\mathcal{T})$ such that

$$\|\boldsymbol{\Sigma} - \overline{\boldsymbol{\Sigma}}\|_{H^1(0,T;\mathcal{T})} < \varepsilon\,. \tag{3.11}$$

Furthermore, it is easy to see that

$$\begin{aligned}&\|E_{N,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} + \sum_{n=1}^{N-1} \|E_{n,\theta}(\boldsymbol{\Sigma}) - E_{n+1,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}}\\ &\quad\leq \int_{I_N} \|\dot{\boldsymbol{\Sigma}}(t)\|_{\mathcal{T}} \,\mathrm{d}t + \sum_{n=1}^{N-1} \|E_{n,\theta}(\overline{\boldsymbol{\Sigma}}) - E_{n+1,\theta}(\overline{\boldsymbol{\Sigma}})\|_{\mathcal{T}} + c\int_0^T \|\dot{\boldsymbol{\Sigma}}(t) - \dot{\overline{\boldsymbol{\Sigma}}}(t)\|_{\mathcal{T}} \,\mathrm{d}t\,.\end{aligned}$$

From [4] we also have the estimate

$$\sum_{n=1}^{N-1} \|E_{n,\theta}(\overline{\boldsymbol{\Sigma}}) - E_{n+1,\theta}(\overline{\boldsymbol{\Sigma}})\|_{\mathcal{T}} \leq ck\|\ddot{\overline{\boldsymbol{\Sigma}}}\|_{L^1(0,T;\mathcal{T})}\,.$$

Making use of all these results, from (3.10) we finally obtain the following.

THEOREM 3.1. Let $\boldsymbol{\Sigma} \in H^1(0,T;\mathcal{T})$ *be the solution of the problem* STRESS. *Then, for the solution* $\{\boldsymbol{\Sigma}_n^k\}_{n=0}^N$ *of the discrete problem* $\mathrm{STRESS}_{k,\theta}$, *the following error estimate holds:*

$$\begin{aligned}\max_{0 \leq n \leq N} \|\boldsymbol{\Sigma}(t_n) - \boldsymbol{\Sigma}_n^l\|_{\mathcal{T}} &\leq c\{(\omega_k(l) + \omega_k(\boldsymbol{\Sigma}))\|\dot{\boldsymbol{\Sigma}}\|_{L^1(0,T;\mathcal{T})}\}^{1/2}\\ &\quad + c\{\|\dot{\boldsymbol{\Sigma}}\|_{L^1(t_{N-1},t_N;\mathcal{T})} + k\|\ddot{\overline{\boldsymbol{\Sigma}}}\|_{L^1(0,T;\mathcal{T})} + \|\dot{\boldsymbol{\Sigma}} - \dot{\overline{\boldsymbol{\Sigma}}}\|_{L^1(0,T;\mathcal{T})}\}\,;\end{aligned} \tag{3.12}$$

in particular,

$$\max_{0 \leq n \leq N} \|\boldsymbol{\Sigma}(t_n) - \boldsymbol{\Sigma}_n^k\|_{\mathcal{T}} \to 0 \quad \text{as } k \to 0\,.$$

4. Fully discrete approximations of the dual problem

We now discuss a family of fully discrete approximations to the problem DUAL.

The fully discrete schemes discussed here can also be viewed as mixed approximations to the stress problem STRESS, the term 'mixed' here referring to the fact that a Lagrange multiplier is introduced as a result of the constraint associated with the bilinear form $b(\cdot,\cdot)$.

To begin with, we assume that a uniform partition of the time interval $[0, T]$ into N sub-intervals is given, with step size $k = T/N$. We assume further that a finite element mesh of the spatial domain Ω is constructed in the usual way, with the mesh size defined by $h = \max h_K$, where h_K is the diameter of element K, a general element of the triangulation. The finite element subspace V^h consists of piecewise linear functions in $V = [H_0^1(\Omega)]^d$, while S^h and M^h are defined to be subspaces of S and M, respectively, comprising piecewise constants. We set $\mathcal{T}^h = S^h \times M^h$, and

$$\mathcal{P}^h = \{\boldsymbol{T}^h = (\boldsymbol{\tau}^h, \boldsymbol{\mu}^h) \in \mathcal{T}^h : \boldsymbol{T}^h(x) \in K \text{ a.e. in } \Omega\}.$$

Again, the parameter $\theta \in [\frac{1}{2}, 1]$, and we set $\delta f_n^{hk} = (f_n^{hk} - f_{n-1}^{hk})/k$. Then, the family of fully discrete schemes for the problem DUAL is

PROBLEM DUAL$_{hk}$. Find $(\boldsymbol{w}^{hk}, \boldsymbol{\Sigma}^{hk}) = \{(\boldsymbol{w}_n^{hk}, \boldsymbol{\Sigma}_n^{hk})\}_{n=0}^N \subset V^h \times \mathcal{P}^h$ with $(\boldsymbol{w}_0^{hk}, \boldsymbol{\Sigma}_0^{hk}) = \boldsymbol{0}$, such that for $n = 1, \ldots, N$,

$$b(\boldsymbol{v}^h, \boldsymbol{\sigma}_{n-1+\theta}^{hk}) = \langle l(t_{n-1+\theta}), \boldsymbol{v}^h \rangle \quad \forall \boldsymbol{v}^h \in V^h, \tag{4.1}$$

$$A_h(\delta \boldsymbol{\Sigma}_n^{hk}, \boldsymbol{T}^h - \boldsymbol{\Sigma}_{n-1+\theta}^{hk}) + b(\boldsymbol{w}_{n-1+\theta}^{hk}, \boldsymbol{\tau}^h - \boldsymbol{\sigma}_{n-1+\theta}^{hk}) \geqslant 0 \quad \forall \boldsymbol{T}^h = (\boldsymbol{\tau}^h, \boldsymbol{\mu}^h) \in \mathcal{P}^h. \tag{4.2}$$

Here, as before, we use the notation $\boldsymbol{\Sigma}_{n-1+\theta}^{hk} = \theta \boldsymbol{\Sigma}_n^{hk} + (1-\theta)\boldsymbol{\Sigma}_{n-1}^{hk}$. We also use $\boldsymbol{w}_{n-1+\theta}^{hk} \in V^h$ to denote an approximation of the velocity $\boldsymbol{w}(t) \equiv \dot{\boldsymbol{u}}(t)$ at $t = t_{n-1+\theta}$. The bilinear form $A_h(\cdot,\cdot) : \mathcal{T} \times \mathcal{T} \to \mathbb{R}$ is the approximation to $A(\cdot,\cdot)$, defined by

$$A_h(\boldsymbol{\Sigma}, \boldsymbol{T}) = \int_\Omega \boldsymbol{\sigma} : \boldsymbol{C}_h^{-1} \boldsymbol{\tau} \, dx + \int_\Omega \boldsymbol{\chi} \cdot \boldsymbol{H}_h^{-1} \boldsymbol{\mu} \, dx \tag{4.3}$$

in which the approximate moduli $\boldsymbol{C}_h^{-1}$ and $\boldsymbol{H}_h^{-1}$ are piecewise constant approximations of $\boldsymbol{C}^{-1}$ and $\boldsymbol{H}^{-1}$. They may be defined to be the average values of $\boldsymbol{C}^{-1}$ and $\boldsymbol{H}^{-1}$ on elements, for example. The approximations $\boldsymbol{C}_h^{-1}$ and $\boldsymbol{H}_h^{-1}$ are assumed to satisfy the material properties enjoyed by $\boldsymbol{C}^{-1}$ and $\boldsymbol{H}^{-1}$, with the constants independent of h, and so $A_h(\cdot,\cdot)$ inherits the properties (2.33) of continuity and $\mathcal{T}$-ellipticity possessed by $A(\cdot,\cdot)$, with the constants independent of h. We assume $c^h(\boldsymbol{C}, \boldsymbol{H}) \equiv \max\{\|\boldsymbol{C}_h^{-1} - \boldsymbol{C}\|_{L^\infty(\Omega)} \|\boldsymbol{H}_h^{-1} - \boldsymbol{H}\|_{L^\infty(\Omega)}\} \to 0$ as $h \to 0$.

It is shown in [3] that under suitable assumptions, the discrete problem DUAL$_{hk}$ has a solution. We introduce the projection operator $\Pi^h : \mathcal{T} \to \mathcal{T}^h$, which is orthogonal with respect to the inner product defined by the bilinear form $A_h(\cdot,\cdot)$; that is, for $\boldsymbol{T} \in \mathcal{T}$, $\Pi^h \boldsymbol{T}$ is the unique element in $\mathcal{T}^h$ such that

$$A_h(\boldsymbol{T} - \Pi^h \boldsymbol{T}, \boldsymbol{T}^h) = 0 \quad \forall \boldsymbol{T}^h \in \mathcal{T}^h. \tag{4.4}$$

From the expression (4.3) we see that the projection has the form $\Pi^h \boldsymbol{T} = (\boldsymbol{\tau}_1^h, \boldsymbol{\mu}_1^h)$, with $\boldsymbol{\tau}_1^h \in S^h$ and $\boldsymbol{\mu}_1^h \in M^h$ being orthogonal projections of $\boldsymbol{\tau}$ and $\boldsymbol{\mu}$ onto S^h and M^h in the inner products defined by the bilinear forms $a_h(\cdot,\cdot)$ and $c_h(\cdot,\cdot)$, respectively, where

$$a_h(\boldsymbol{\sigma}, \boldsymbol{\tau}) = \int_\Omega \boldsymbol{\sigma} : \boldsymbol{C}_h^{-1} \boldsymbol{\tau} \, dx,$$

$$c_h(\boldsymbol{\chi}, \boldsymbol{\mu}) = \int_\Omega \boldsymbol{\chi} \cdot \boldsymbol{H}_h^{-1} \boldsymbol{\mu} \, dx.$$

We will use the same symbol Π^h also to denote these two orthogonal projections; that is, we will write $\Pi^h \boldsymbol{T} = (\Pi^h \boldsymbol{\tau}, \Pi^h \boldsymbol{\mu})$. The following result is proved in [3].

LEMMA 4.1. The orthogonal projections $\Pi^h : S \to S^h$ and $\Pi^h : M \to M^h$ are piecewise averaging operators; that is, for $\boldsymbol{T} = (\boldsymbol{\tau}, \boldsymbol{\mu}) \in \mathcal{T}$,

$$\Pi^h \boldsymbol{\tau}|_K = \frac{1}{\text{meas}(K)} \int_K \boldsymbol{\tau}(\boldsymbol{x})\, dx\,, \qquad \Pi^h \boldsymbol{\mu}|_K = \frac{1}{\text{meas}(K)} \int_K \boldsymbol{\mu}(\boldsymbol{x})\, dx\,, \quad \textit{for any element } K\,. \tag{4.5}$$

Consequently, by the convexity of the set $\mathcal{P}$, if $\boldsymbol{T} \in \mathcal{P}$, then $\Pi^h \boldsymbol{T} \in \mathcal{P}^h$. Also, we have

$$b(\boldsymbol{v}^h, \Pi^h \boldsymbol{\tau} - \boldsymbol{\tau}) = 0 \quad \forall\, \boldsymbol{v}^h \in V^h,\ \boldsymbol{\tau} \in S\,. \tag{4.6}$$

Denote by $\|\cdot\|_h$ the norm induced by the discrete bilinear form $A_h(\cdot, \cdot)$; that is,

$$\|\boldsymbol{T}\|_h = A_h(\boldsymbol{T}, \boldsymbol{T})^{1/2}\,.$$

By the assumptions made on $\boldsymbol{C}_h^{-1}$ and $\boldsymbol{H}_h^{-1}$, the norm $\|\cdot\|_h$ is equivalent to $\|\cdot\|_{\mathcal{T}}$ with the equivalence constants independent of h.

Let $\boldsymbol{e}_n = \boldsymbol{\Sigma}(t_n) - \boldsymbol{\Sigma}_n^{hk}$, $n = 0, 1, \ldots, N$, denote the approximation error, with $\boldsymbol{e}_0 = \boldsymbol{0}$, and

$$M = \max_{0 \le n \le N} \|\boldsymbol{e}_n\|_{\mathcal{T}}\,.$$

Since $\theta \in [\frac{1}{2}, 1]$, we have

$$A_h(\delta \boldsymbol{e}_n, \theta \boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1}) \ge \frac{1}{2k}(\|\boldsymbol{e}_n\|_h^2 - \|\boldsymbol{e}_{n-1}\|_h^2)\,. \tag{4.7}$$

To derive an upper bound, we write

$$A_h(\delta \boldsymbol{e}_n, \theta \boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1}) = A_h(\delta \boldsymbol{e}_n, E_{n,\theta}(\boldsymbol{\Sigma})) + A_h(\delta \boldsymbol{e}_n, \boldsymbol{\Sigma}(t_{n-1+\theta}) - \boldsymbol{\Sigma}_{n-1+\theta}^{hk}) \tag{4.8}$$

where $E_{n,\theta}(\boldsymbol{\Sigma})$ is defined in (3.3). Hence

$$\frac{1}{2k}(\|\boldsymbol{e}_n\|_h^2 - \|\boldsymbol{e}_{n-1}\|_h^2) \le A_h(\delta \boldsymbol{e}_n, E_{n,\theta}(\boldsymbol{\Sigma})) + A_h(\delta \boldsymbol{e}_n, \boldsymbol{\Sigma}(t_{n-1+\theta}) - \boldsymbol{\Sigma}_{n-1+\theta}^{hk})\,. \tag{4.9}$$

We now examine the second term on the right-hand side of (4.9).

$$\begin{aligned}
A_h(\delta \boldsymbol{e}_n, \boldsymbol{\Sigma}(t_{n-1+\theta}) - \boldsymbol{\Sigma}_{n-1+\theta}^{hk}) &= A_h(\delta \boldsymbol{\Sigma}_n, \boldsymbol{\Sigma}(t_{n-1+\theta}) - \boldsymbol{\Sigma}_{n-1+\theta}^{hk}) - A_h(\delta \boldsymbol{\Sigma}_n^{hk}, \boldsymbol{\Sigma}(t_{n-1+\theta}) - \boldsymbol{\Sigma}_{n-1+\theta}^{hk}) \\
&= A_h(\delta \boldsymbol{\Sigma}_n, -E_{n,\theta}(\boldsymbol{\Sigma})) + A_h(\delta \boldsymbol{\Sigma}_n, \theta \boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1}) \\
&\quad - A_h(\delta \boldsymbol{\Sigma}_n^{hk}, \Pi^h \boldsymbol{\Sigma}(t_{n-1+\theta}) - \boldsymbol{\Sigma}_{n-1+\theta}^{hk}) - A_h(\delta \boldsymbol{\Sigma}_n^{hk}, \boldsymbol{\Sigma}(t_{n-1+\theta}) \\
&\quad - \Pi^h \boldsymbol{\Sigma}(t_{n-1+\theta}))\,.
\end{aligned} \tag{4.10}$$

Now, take $\boldsymbol{T}^h = \Pi^h \boldsymbol{\Sigma}(t_{n-1+\theta}) \in \mathcal{P}^h$ in (4.2) to obtain

$$-A_h(\delta \boldsymbol{\Sigma}_n^{hk}, \Pi^h \boldsymbol{\Sigma}(t_{n-1+\theta}) - \boldsymbol{\Sigma}_{n-1+\theta}^{hk}) \le b(\boldsymbol{w}_{n-1+\theta}^{hk}, \Pi^h \boldsymbol{\sigma}(t_{n-1+\theta}) - \boldsymbol{\sigma}_{n-1+\theta}^{hk})\,.$$

Setting $t = t_{n-1+\theta}$ and $\boldsymbol{v} = \boldsymbol{v}^h$ in (2.30), and subtracting (4.1) from the resulting equation, we find that

$$b(\boldsymbol{v}^h, \boldsymbol{\sigma}(t_{n-1+\theta}) - \boldsymbol{\sigma}_{n-1+\theta}^{hk}) = 0 \quad \forall\, \boldsymbol{v}^h \in V^h\,. \tag{4.11}$$

Next, applying (4.6), we have

$$b(\boldsymbol{w}_{n-1+\theta}^{hk}, \Pi^h \boldsymbol{\sigma}(t_{n-1+\theta}) - \boldsymbol{\sigma}_{n-1+\theta}^{hk}) = b(\boldsymbol{w}_{n-1+\theta}^{hk}, \Pi^h \boldsymbol{\sigma}(t_{n-1+\theta}) - \boldsymbol{\sigma}(t_{n-1+\theta})) = 0.$$

Therefore,

$$-A_h(\delta \boldsymbol{\Sigma}_n^{hk}, \Pi^h \boldsymbol{\Sigma}(t_{n-1+\theta}) - \boldsymbol{\Sigma}_{n-1+\theta}^{hk}) \le 0\,. \tag{4.12}$$

Because Π^h is the orthogonal projection onto $\mathcal{T}^h$ in the inner product induced by the bilinear form $A_h(\cdot, \cdot)$, we have

$$A_h(\delta \boldsymbol{\Sigma}_n^{hk}, \boldsymbol{\Sigma}(t_{n-1+\theta}) - \Pi^h \boldsymbol{\Sigma}(t_{n-1+\theta})) = 0\,. \tag{4.13}$$

Using (4.12) and (4.13) in (4.10), we see that

$$A_h(\delta \boldsymbol{e}_n, \boldsymbol{\Sigma}(t_{n-1+\theta}) - \boldsymbol{\Sigma}^{hk}_{n-1+\theta}) \le A_h(\delta \boldsymbol{\Sigma}_n, -E_{n,\theta}(\boldsymbol{\Sigma})) + A_h(\delta \boldsymbol{\Sigma}_n, \theta \boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1}) . \tag{4.14}$$

Now, we take $\boldsymbol{T} = \boldsymbol{\Sigma}^{hk}_{n-1+\theta} \in \mathscr{P}$ in (2.31) and integrate the inequality over $I_n = [t_{n-1}, t_n]$ to obtain

$$\int_{I_n} A(\dot{\boldsymbol{\Sigma}}(t), \boldsymbol{\Sigma}^{hk}_{n-1+\theta} - \boldsymbol{\Sigma}(t))\, dt + \int_{I_n} b(\boldsymbol{w}(t), \boldsymbol{\sigma}^{hk}_{n-1+\theta} - \boldsymbol{\sigma}(t))\, dt \ge 0 ,$$

which can be rewritten, after being divided by k, as

$$\begin{aligned} A_h(\delta \boldsymbol{\Sigma}_n, \theta \boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1}) \le \frac{1}{k}\int_{I_n} & A(\dot{\boldsymbol{\Sigma}}(t), \theta\boldsymbol{\Sigma}(t_n) + (1-\theta)\boldsymbol{\Sigma}(t_{n-1}) - \boldsymbol{\Sigma}(t))\, dt \\ & + \frac{1}{k}\int_{I_n} b(\boldsymbol{w}(t), \boldsymbol{\sigma}^{hk}_{n-1+\theta} - \boldsymbol{\sigma}(t))\, dt . \end{aligned} \tag{4.15}$$

Combining (4.9), (4.14) and (4.15), we obtain

$$\begin{aligned} \frac{1}{2k}(\|\boldsymbol{e}_n\|_h^2 - \|\boldsymbol{e}_{n-1}\|_h^2) \le{}& A_h(\delta \boldsymbol{e}_n, E_{n,\theta}(\boldsymbol{\Sigma})) + A_h(\delta \boldsymbol{\Sigma}_n, -E_{n,\theta}(\boldsymbol{\Sigma})) \\ & + \frac{1}{k}\int_{I_n} A_h(\dot{\boldsymbol{\Sigma}}(t), \theta \boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1})\, dt - \frac{1}{k}\int_{I_n} A(\dot{\boldsymbol{\Sigma}}(t), \theta \boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1})\, dt \\ & + \frac{1}{k}\int_{I_n} A(\dot{\boldsymbol{\Sigma}}(t), \theta\boldsymbol{\Sigma}(t_n) + (1-\theta)\boldsymbol{\Sigma}(t_{n-1}) - \boldsymbol{\Sigma}(t))\, dt \\ & + \frac{1}{k}\int_{I_n} b(\boldsymbol{w}(t), \boldsymbol{\sigma}^{hk}_{n-1+\theta} - \boldsymbol{\sigma}(t))\, dt . \end{aligned}$$

Now, multiplying the inequality by k and rearranging some terms, we get

$$\begin{aligned} \frac{1}{2}(\|\boldsymbol{e}_n\|_h^2 - \|\boldsymbol{e}_{n-1}\|_h^2) \le{}& A_h(\boldsymbol{e}_n - \boldsymbol{e}_{n-1}, E_{n,\theta}(\boldsymbol{\Sigma})) + \int_{I_n} A(\dot{\boldsymbol{\Sigma}}(t), E_{n,\theta}(\boldsymbol{\Sigma}))\, dt - \int_{I_n} A_h(\dot{\boldsymbol{\Sigma}}(t), E_{n,\theta}(\boldsymbol{\Sigma}))\, dt \\ & + \int_{I_n} A_h(\dot{\boldsymbol{\Sigma}}(t), \theta \boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1})\, dt - \int_{I_n} A(\dot{\boldsymbol{\Sigma}}(t), \theta \boldsymbol{e}_n + (1-\theta)\boldsymbol{e}_{n-1})\, dt \\ & + \int_{I_n} A_h(\dot{\boldsymbol{\Sigma}}(t), \boldsymbol{\Sigma}(t_{n-1+\theta}) - \boldsymbol{\Sigma}(t))\, dt + \int_{I_n} b(\boldsymbol{w}(t), \boldsymbol{\sigma}^{hk}_{n-1+\theta} - \boldsymbol{\sigma}(t))\, dt . \end{aligned}$$

Using the quantities $c^h(\boldsymbol{C}, \boldsymbol{H})$ and $\omega_k(\boldsymbol{\Sigma})$ defined before, we then derive from the above inequality

$$\begin{aligned} \|\boldsymbol{e}_n\|_h^2 - \|\boldsymbol{e}_{n-1}\|_h^2 \le{}& 2A_h(\boldsymbol{e}_n - \boldsymbol{e}_{n-1}, E_{n,\theta}(\boldsymbol{\Sigma})) + cc^h(\boldsymbol{C}, \boldsymbol{H}) \int_{I_n} \|\dot{\boldsymbol{\Sigma}}(t)\|_{\mathscr{T}}\, dt (\|E_{n,\theta}(\boldsymbol{\Sigma})\|_{\mathscr{T}} + M + \omega_k(\boldsymbol{\Sigma})) \\ & + \int_{I_n} b(\boldsymbol{w}(t), \boldsymbol{\sigma}^{hk}_{n-1+\theta} - \boldsymbol{\sigma}(t))\, dt . \end{aligned} \tag{4.16}$$

Next, we estimate the last term in (4.16). From (4.1) and (2.30), we have

$$b(\boldsymbol{v}^h(t), \boldsymbol{\sigma}(t_{n-1+\theta}) - \boldsymbol{\sigma}^{hk}_{n-1+\theta}) = 0 \quad \forall\, \boldsymbol{v}^h(t) \in V^h$$

and so

$$\begin{aligned}\int_{I_n} b(\boldsymbol{w}(t), \boldsymbol{\sigma}^{hk}_{n-1+\theta} - \boldsymbol{\sigma}(t))\,dt &= \int_{I_n} b(\boldsymbol{w}(t), \boldsymbol{\sigma}^{hk}_{n-1+\theta} - \boldsymbol{\sigma}(t_{n-1+\theta}))\,dt + \int_{I_n} b(\boldsymbol{w}(t), \boldsymbol{\sigma}(t_{n-1+\theta}) - \boldsymbol{\sigma}(t))\,dt \\
&\leq \int_{I_n} b(\boldsymbol{w}(t) - \boldsymbol{v}^h(t), \boldsymbol{\sigma}^{hk}_{n-1+\theta} - \boldsymbol{\sigma}(t_{n-1+\theta}))\,dt + c\omega_k(\boldsymbol{\sigma}) \int_{I_n} \|\boldsymbol{w}(t)\|_V\,dt \\
&= \int_{I_n} b(\boldsymbol{w}(t) - \boldsymbol{v}^h(t), \theta(\boldsymbol{\sigma}_n - \boldsymbol{\sigma}(t_n)) + (1-\theta)(\boldsymbol{\sigma}_{n-1} - \boldsymbol{\sigma}(t_{n-1})))\,dt \\
&\quad + \int_{I_n} b(\boldsymbol{w}(t) - \boldsymbol{v}^h(t), E_{n,\theta}(\boldsymbol{\sigma}))\,dt + c\omega_k(\boldsymbol{\sigma}) \int_{I_n} \|\boldsymbol{w}(t)\|_V\,dt \\
&\leq c \int_{I_n} \|\boldsymbol{w}(t) - \boldsymbol{v}^h(t)\|_V\,dt (\|\boldsymbol{\sigma}_n - \boldsymbol{\sigma}(t_n)\|_S + \|\boldsymbol{\sigma}_{n-1} - \boldsymbol{\sigma}(t_{n-1})\|_S + \|E_{n,\theta}(\boldsymbol{\sigma})\|_S) \\
&\quad + c\omega_k(\boldsymbol{\sigma}) \int_{I_n} \|\boldsymbol{w}(t)\|_V\,dt \\
&\leq c \int_{I_n} \|\boldsymbol{w}(t) - \boldsymbol{v}^h(t)\|_V\,dt (M + \|E_{n,\theta}(\boldsymbol{\Sigma})\|_S) + c\omega_k(\boldsymbol{\sigma}) \int_{I_n} \|\boldsymbol{w}(t)\|_V\,dt\,.\end{aligned}$$

Applying the inequality recursively, and using the fact that $\boldsymbol{e}_0 = \boldsymbol{0}$ we have

$$\begin{aligned}\|\boldsymbol{e}_n\|_h^2 &\leq 2 \sum_{j=1}^{n} A_h(\boldsymbol{e}_j - \boldsymbol{e}_{j-1}, E_{j,\theta}(\boldsymbol{\Sigma})) \\
&\quad + cc^h(\boldsymbol{C}, \boldsymbol{H}) \int_0^{t_n} \|\dot{\boldsymbol{\Sigma}}(t)\|_{\mathcal{T}}\,dt \Big(\max_{0\leq m\leq N} \|E_{m,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} + M + \omega_k(\boldsymbol{\Sigma}) \Big) \\
&\quad + c\Big(\max_{0\leq m\leq N} \|E_{m,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} + M \Big) \int_0^{t_n} \|\boldsymbol{w}(t) - \boldsymbol{v}^h(t)\|_V\,dt + c\omega_k(\boldsymbol{\sigma}) \int_0^{t_n} \|\boldsymbol{w}(t)\|_V\,dt \\
&= 2A_h(\boldsymbol{e}_n, E_{n,\theta}(\boldsymbol{\Sigma})) + 2 \sum_{j=1}^{n-1} A_h(\boldsymbol{e}_j, E_{j,\theta}(\boldsymbol{\Sigma}) - E_{j+1,\theta}(\boldsymbol{\Sigma})) \\
&\quad + cc^h(\boldsymbol{C}, \boldsymbol{H}) \int_0^{t_n} \|\dot{\boldsymbol{\Sigma}}(t)\|_{\mathcal{T}}\,dt \Big(\max_{0\leq m\leq N} \|E_{m,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} + M + \omega_k(\boldsymbol{\Sigma}) \Big) \\
&\quad + c\Big(\max_{0\leq m\leq N} \|E_{m,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} + M \Big) \int_0^{t_n} \|\boldsymbol{w}(t) - \boldsymbol{v}^h(t)\|_V\,dt + c\omega_k(\boldsymbol{\sigma}) \int_0^{t_n} \|\boldsymbol{w}(t)\|_V\,dt\,.\end{aligned}$$

Therefore,

$$\begin{aligned}M^2 &\leq cM \Big(\|E_{N,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} + \sum_{n=1}^{N-1} \|E_{n,\theta}(\boldsymbol{\Sigma}) - E_{n+1,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} + c^h(\boldsymbol{C}, \boldsymbol{H}) \|\dot{\boldsymbol{\Sigma}}\|_{L^1(0,T;\mathcal{T})} + \|\boldsymbol{w} - \boldsymbol{v}^h\|_{L^1(0,T;V)} \Big) \\
&\quad + cc^h(\boldsymbol{C}, \boldsymbol{H}) \|\dot{\boldsymbol{\Sigma}}\|_{L^1(0,T;\mathcal{T})} \Big(\max_{0\leq n\leq N} \|E_{n,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} + \omega_k(\boldsymbol{\Sigma}) \Big) \\
&\quad + c \max_{0\leq n\geq N} \|E_{n,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} \|\boldsymbol{w} - \boldsymbol{v}^h\|_{L^1(0,T;V)} + c\omega_k(\boldsymbol{\sigma}) \|\dot{\boldsymbol{w}}\|_{L^1(0,T;V)}\,.\end{aligned}$$

Next, apply the inequality (2.35) to obtain

$$\begin{aligned}M &\leq c \Big(\|E_{N,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} + \sum_{n=1}^{N-1} \|E_{n,\theta}(\boldsymbol{\Sigma}) - E_{n+1,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} + c^h(\boldsymbol{C}, \boldsymbol{H}) \|\dot{\boldsymbol{\Sigma}}\|_{L^1(0,T;\mathcal{T})} + \|\boldsymbol{w} - \boldsymbol{v}^h\|_{L^1(0,T;V)} \Big) \\
&\quad + c \Big\{ c^h(\boldsymbol{C}, \boldsymbol{H}) \|\dot{\boldsymbol{\Sigma}}\|_{L^1(0,T;\mathcal{T})} \Big(\max_{0\leq n\leq N} \|E_{n,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} + \omega_k(\boldsymbol{\Sigma}) \Big) \\
&\quad + \max_{0\leq n\leq N} \|E_{n,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} \|\boldsymbol{w} - \boldsymbol{v}^h\|_{L^1(0,T;V)} + \omega_k(\boldsymbol{\sigma}) \|\dot{\boldsymbol{w}}\|_{L^1(0,T;)} \Big\}^{1/2}\,.\end{aligned}$$

Since $\boldsymbol{v}^h \in L^1(0, T; V^h)$ is arbitrary, we then get the estimate

$$\max_{0\le n\le N}\|\boldsymbol{e}_n\|_h \le c\Big(\|E_{N,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} + \sum_{n=1}^{N-1}\|E_{n,\theta}(\boldsymbol{\Sigma}) - E_{n+1,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}}$$
$$+ c^h(\boldsymbol{C},\boldsymbol{H})\|\dot{\boldsymbol{\Sigma}}\|_{L^1(0,T;\mathcal{T})} + \inf_{\boldsymbol{v}^h\in L^1(0,T;V^h)}\|\boldsymbol{w}-\boldsymbol{v}^h\|_{L^1(0,T;V)}\Big)$$
$$+ c\{c^h(\boldsymbol{C},\boldsymbol{H})\|\dot{\boldsymbol{\Sigma}}\|_{L^1(0,T;\mathcal{T})}\Big(\max_{0\le n\le N}\|E_{n,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}} + \omega_k(\boldsymbol{\Sigma})\Big)$$
$$+ \max_{0\le n\le N}\|E_{n,\theta}(\boldsymbol{\Sigma})\|_{\mathcal{T}}\inf_{\boldsymbol{v}^h\in L^1(0,T;V^h)}\|\boldsymbol{w}-\boldsymbol{v}^h\|_{L^1(0,T;V)} + \omega_k(\boldsymbol{\sigma})\|\dot{\boldsymbol{w}}\|_{L^1(0,T;V)}\}^{1/2}. \tag{4.17}$$

Now for any $\varepsilon > 0$, by Theorem 2.1, there exists $\bar{\boldsymbol{w}} \in L^2(0,T;(H^2(\Omega))^d)$ such that

$$\|\boldsymbol{w} - \bar{\boldsymbol{w}}\|_{L^2(0,T;V)} < \varepsilon .$$

Furthermore, by a standard result in finite element interpolation theory (cf. e.g. [2]),

$$\inf_{\boldsymbol{v}^h(t)\in V^h}\|\bar{\boldsymbol{w}}(t) - \boldsymbol{v}^h(t)\|_V \le ch\|\bar{\boldsymbol{w}}(t)\|_{(H^2(\Omega))^d} .$$

Then

$$\inf_{\boldsymbol{v}^h\in L^1(0,T;V^h)}\|\boldsymbol{w}-\boldsymbol{v}^h\|_{L^1(0,T;V)} \le \|\boldsymbol{w}-\bar{\boldsymbol{w}}\|_{L^1(0,T;V)} + \inf_{\boldsymbol{v}^h\in L^1(0,T;V^h)}\|\bar{\boldsymbol{w}}-\boldsymbol{v}^h\|_{L^1(0,T;V)}$$
$$\le c\varepsilon + ch\|\bar{\boldsymbol{w}}\|_{L^1(0,T;(H^2(\Omega))^d)} .$$

The other terms on the right-hand side of (4.17) can be estimated as in the case of the time-discrete schemes in the previous section. Recalling also that $\|\cdot\|_h$ is a norm equivalent to $\|\cdot\|_{\mathcal{T}}$ with the equivalence constants independent of h, we have the following result.

THEOREM 4.2. Let $(\boldsymbol{u},\boldsymbol{\Sigma}) \in H^1(0,T;V\times\mathcal{T})$ *be the solution to the problem* DUAL. *Then, the discrete solution of the problem* DUAL_{hk} *converges:*

$$\max_{0\le n\le N}\|\boldsymbol{\Sigma}(t_n) - \boldsymbol{\Sigma}_n^{hk}\|_{\mathcal{T}} \to 0 \quad \text{as } k,h\to 0 .$$

References

[1] G. Anzellotti and S. Luckhaus, Dynamical evolution of elasto-perfectly plastic bodies, Appl. Math. Optimiz. 15 (1987) 121–140.
[2] P.G. Ciarlet, The Finite Element Method for Elliptic Problems (North Holland, Amsterdam, 1978).
[3] W. Han and B.D. Reddy, Computational plasticity: the variational basis and numerical analysis, Comput. Mech. Adv. 2 (1995) 283–400.
[4] W. Han and B.D. Reddy, Convergence analysis of discrete approximations of the primal problem in hardening plasticity, submitted.
[5] C. Johnson, A mixed finite element method for plasticity problems with hardening, SIAM J. Numer. Anal. 14 (1977) 575–583.
[6] H. Matthies, Existence theorems in thermoplasticity, Journal de Méecanique 18 (1979) 675–712.
[7] J.C. Simo and T.J.R. Hughes, Computational Inelasticity (Springer-Verlag, New York, 1997).