Available online at www.sciencedirect.com

ScienceDirect

Comput. Methods Appl. Mech. Engrg. 196 (2007) 3881–3897

Computer methods in applied mechanics and engineering

www.elsevier.com/locate/cma

A discontinuous Galerkin formulation for classical and gradient plasticity – Part 1: Formulation and analysis

J.K. Djoko [a], F. Ebobisse [a], A.T. McBride [b], B.D. Reddy [a,b,*]

[a] Department of Mathematics and Applied Mathematics, University of Cape Town, 7701 Rondebosch, South Africa
[b] Centre for Research in Computational and Applied Mechanics, University of Cape Town, 7701 Rondebosch, South Africa

Received 1 November 2005; accepted 30 October 2006
Available online 27 March 2007

Abstract

A discontinuous Galerkin formulation is developed and analyzed for the cases of classical and gradient plasticity. The model of gradient plasticity is based on the von Mises yield function, in which dependence is on the isotropic hardening parameter and its Laplacian. The problem takes the form of a variational inequality of the second kind. The discontinuous Galerkin formulation is shown to be consistent and convergent. Error estimates are obtained for the cases of semi- and fully discrete formulations; these mimic the error estimates obtained for classical plasticity with the conventional Galerkin formulation.

Keywords: Discontinuous Galerkin; Elastoplasticity; Gradient plasticity; Variational inequality; Convergence

1. Introduction

The theoretical underpinnings of the classical theory of elastoplasticity have been vigorously developed during the last half-century, and constitutive models for infinitesimal and large strains, and for a range of material types that include both ductile and brittle behaviour, are now firmly established. Computational methods for the solution of complex problems involving inelastic behaviour have undergone a parallel development, and there is now a good understanding of the associated algorithms and, particularly for small strains, of the convergence theory [1,2].

Motivated in large part by the inability of classical theories to model material behaviour at the meso-scale level, various plasticity theories that incorporate size-dependence via the inclusion of strain gradients have been developed. These theories include in a natural way a length scale, and permit phenomena such as shear bands to be captured. For example, in the early works [3,4], the yield function is augmented by a term involving the Laplacian of the equivalent plastic strain, and possibly higher-order terms. The associated variational structure has been discussed in subsequent work [5,6], in which numerical approximations have also been carried out. A general survey of the theory of strain gradient plasticity is provided in [7], while Needleman [8] has presented a survey of computational aspects of meso-scale mechanics. In the works of Gurtin [9] and Gudmundson [10], it is assumed that plastic flow is governed not necessarily by the magnitude of the stress deviator, but more generally by microstress tensors that also satisfy a balance law.

Recent contributions to numerical aspects of strain gradient plasticity at small and large strains have been made by Svedberg and Runesson [11], Liebe and Steinmann [12], and Liebe et al. [13]. A feature of theories of gradient plasticity

* Corresponding author. Address: Department of Mathematics and Applied Mathematics, University of Cape Town, 7701 Rondebosch, South Africa. Tel.: +27 21 650 3787; fax: +27 21 685 2281.

E-mail addresses: jules.djokokamdem@up.ac.za (J.K. Djoko), ebobisse@maths.uct.ac.za (F. Ebobisse), andrew.mcbride@uct.ac.za (A.T. McBride), daya.reddy@uct.ac.za (B.D. Reddy).

0045-7825/$ - see front matter
doi:10.1016/j.cma.2006.10.045

that requires careful attention in the development of computational approaches to approximate solutions is the introduction of terms involving higher-order derivatives of quantities such as the equivalent plastic strain. In classical theories these and other internal variables may be approximated by piecewise discontinuous functions since there is no requirement of continuity. Furthermore, in finite element approximations these quantities may be condensed out at element level, or computations may be carried out locally at integration points, in either case with the consequence that the predictor part of the solution algorithm involves only the displacement degrees of freedom. Such approaches have either to be modified in gradient theories, or it becomes essential to assume continuity of the relevant internal variables, with a consequently significant increase in the size of the discrete problem to be solved.

The need to develop a computational procedure for gradient plasticity that retains the simplicity of the classical algorithms is therefore self-evident. A candidate approach is the class of discontinuous Galerkin (DG) methods, in which inter-element continuity is relaxed in a framework in which the discrete problem remains consistent. DG methods were developed in the 1970s and 1980s [14,15], but it is only in recent years that they have been exploited in a wide range of problems. The collection [16] provides an excellent overview of the key approaches for elliptic and hyperbolic problems. Within the context of linear elasticity there have been important contributions by Rivière and Wheeler [17] and Wheeler [18], the latter considering the case of non-convex domains and vanishing compressibility. A DG method has recently been developed for strain gradient dependent damage models [19,20], while the work by Engel et al. [21] treats continuous/discontinuous Galerkin methods for fourth-order problems by reducing the classical requirement of C^1 continuity of the unknown variable to one of continuity.

The goal of this work is to explore the use of DG methods for the solution of plane problems in elastoplasticity. Both the classical and gradient plasticity formulations are considered, though the emphasis is on the latter, for which case DG methods carry significant advantages, as discussed earlier. For the case of gradient plasticity a simple candidate model, first developed by Aifantis and co-workers [3–5] is used. In this model the von Mises yield function with isotropic hardening is augmented by a term involving the Laplacian of the hardening parameter.

Careful attention is paid to the appropriate variational form of the problem for gradient plasticity. The focus is on the primal formulation in which the kinematic quantities such as displacement, plastic strain, and equivalent plastic strain, are the unknown variables. The primal formulation, which has received a detailed treatment in [1], is less popular as a basis for computational treatments of the problem than the dual form of the problem, which uses the flow law in its traditional form of the normality law with the Kuhn-Tucker conditions. Nevertheless, as will be seen, the primal formulation is particularly well suited to problems such as that in gradient plasticity in which higher-order derivatives of the kinematic quantities appear.

Both the classical and gradient plasticity problems are formulated as variational inequalities of the second kind, and semi- and fully discrete DG approximations are considered. Full convergence analyses are carried out of the well-posedness of the problems. In a subsequent work [22] the classical algorithms for the solution of the discrete problems are extended and analyzed, and a series of numerical results will be presented and discussed.

The structure of the rest of this work is as follows. In Section 2 the governing equations and inequalities for the classical and gradient problems are presented, and well-posedness of the variational formulations is discussed. A simple symmetric interior penalty DG method is introduced in Section 3, and the consistency of the discrete formulation is shown. Section 4 is devoted to an analysis of continuous-in-time approximations, and here it is shown that the displacement approximation satisfies the same error estimate as that which is valid for conventional Galerkin approximations of the classical problem [1]. Finally, in Section 5 fully discrete DG approximations are analyzed. Here too the same estimate associated with the classical approach, at least in the absence of greater regularity of the solution, is obtained. The work concludes with a description of computational aspects, which are to be presented in the companion paper.

2. The governing equations for the problem

Let Ω be a bounded convex Lipschitz domain in $\mathbb{R}^2$, which is occupied by an elastoplastic body in its undeformed configuration. A material point in Ω is denoted by $\boldsymbol{x}$ and the time domain under consideration is the interval $[0, T]$. The boundary of Ω is denoted by $\partial\Omega$. The body is assumed to undergo infinitesimal deformations. Its behaviour is governed by the equation of equilibrium

$$\mathrm{div}\boldsymbol{\sigma} + \boldsymbol{f} = \mathbf{0} \tag{1}$$

in which $\boldsymbol{\sigma}$ is the symmetric stress tensor, and $\boldsymbol{f}(\boldsymbol{x}, t) : \Omega \times [0, T] \to \mathbb{R}^2$ is the body force. The elastic part of the constitutive relations is given by

$$\boldsymbol{\sigma} = \mathscr{C}\boldsymbol{e} = \mathscr{C}(\boldsymbol{\epsilon} - \boldsymbol{p}) \tag{2}$$

in which $\boldsymbol{e}$ is the elastic strain, defined to be the difference between the total strain $\boldsymbol{\epsilon}$ and the plastic strain $\boldsymbol{p}$; that is,

$$\boldsymbol{e} = \boldsymbol{\epsilon} - \boldsymbol{p}. \tag{3}$$

All strain quantities are symmetric. In addition the total strain is given in terms of the displacement by

$$\boldsymbol{\epsilon}(\boldsymbol{u}) = \frac{1}{2}(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^{\mathrm{T}}) \tag{4}$$

and plastic behavior assumed to be incompressible, so that

$$\mathrm{tr}\,\boldsymbol{p} = 0. \tag{5}$$

Elastic behavior is assumed to be isotropic and homogeneous, so that the elasticity tensor $\mathscr{C}$ is given in component form by $C_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$; that is,

$$\mathscr{C}\boldsymbol{e} = \lambda(\mathrm{tr}\boldsymbol{e})\boldsymbol{I} + 2\mu\boldsymbol{e}$$

for any symmetric tensor $\boldsymbol{e}$. Here $\lambda > 0$ and $\mu > 0$ are the Lamé constants which are related to Young's modulus E and Poisson's ratio v by

$$\lambda = \frac{Ev}{(1+v)(1-2v)}, \quad \mu = \frac{E}{2(1+v)}. \tag{6}$$

The material is assumed to be homogeneous so that λ and μ are constant.

It follows from the positivity of the Lamé moduli that

$$\boldsymbol{\eta} : \mathscr{C}\boldsymbol{\eta} \geqslant 2\mu|\boldsymbol{\eta}|^2 \quad \text{for any tensor } \boldsymbol{\eta}. \tag{7}$$

We also define

$$|\mathscr{C}|_\infty := \max_{i,j,k,l} |C_{ijkl}| = \lambda + 2\mu. \tag{8}$$

2.1. Classical plasticity

Two forms of plastic behavior are modelled; the first is the classical model based on the assumption, pointwise a.e., of a convex elastic domain $\mathscr{E}$ with boundary $\partial\mathscr{E}$, the yield surface, and a generalized normality law. For definiteness $\mathscr{E}$ is assumed to be defined by the von Mises condition, and both kinematic and isotropic hardening are adopted, so that the region of admissible generalized stresses becomes the set $(\boldsymbol{\sigma}, \boldsymbol{\alpha}, \bar{g})$ that satisfies

$$\varphi(\boldsymbol{\sigma}, \boldsymbol{\alpha}, \bar{g}) = |\boldsymbol{s} + \boldsymbol{\alpha}| + \bar{g} - \kappa \leqslant 0. \tag{9}$$

Here κ is a constant related to the initial yield stress in uniaxial tension, and

$$\begin{aligned} &\boldsymbol{s} = \mathrm{dev}\boldsymbol{\sigma} \text{ (the stress deviator)}, \\ &\boldsymbol{\alpha} = -k_1\boldsymbol{p} \text{ (the back-stress)}, \\ &\bar{g} = -k_2\gamma \text{ (the internal force conjugate to the isotropic hardening parameter } \gamma). \end{aligned} \tag{10}$$

The flow law then takes the form

$$\begin{aligned} &\dot{\boldsymbol{p}} = \Lambda\frac{\partial\varphi}{\partial\boldsymbol{\sigma}}, \\ &\dot{\gamma} = \Lambda, \\ &\Lambda \geqslant 0, \ \varphi \leqslant 0, \ \Lambda\varphi = 0. \end{aligned} \tag{11}$$

It follows that γ, the isotropic hardening parameter, is the equivalent plastic strain; that is, $\Lambda = \dot{\gamma} = |\dot{\boldsymbol{p}}|$.

We will make use of the flow law in its *primal* form [1], in which the generalized stresses are related to the rates of change of the conjugate kinematic quantities through the dissipation function D. For the case at hand we have

$$D(\boldsymbol{q}, \mu) = \begin{cases} \kappa|\boldsymbol{q}| & \text{if } |\boldsymbol{q}| \leqslant \eta, \\ +\infty & \text{otherwise}, \end{cases} \tag{12}$$

where $\boldsymbol{q}$ and η are an arbitrary plastic strain and hardening parameter. Using arguments of convex analysis as in [1], the flow law (11) is shown to be equivalent to

$$D(\boldsymbol{q}, \eta) \geqslant D(\dot{\boldsymbol{p}}, \dot{\gamma}) + (\boldsymbol{\sigma} - k_1\boldsymbol{p}) : (\boldsymbol{q} - \dot{\boldsymbol{p}}) - k_2\gamma(\eta - \dot{\gamma}) \quad \forall(\boldsymbol{q}, \eta). \tag{13}$$

For the sake of convenience, homogenous Dirichlet boundary conditions are assumed; that is,

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{on } \partial\Omega. \tag{14}$$

In addition, the body is assumed initially undeformed and unstressed with the initial conditions

$$\boldsymbol{u}(\boldsymbol{x}, 0) = \boldsymbol{0} \quad \text{and} \quad \boldsymbol{p}(\boldsymbol{x}, 0) = \boldsymbol{0}, \quad \boldsymbol{x} \in \Omega. \tag{15}$$

The variational problem [1]. We define the function spaces of displacements V, plastic strains Q, and hardening variables $\overline{M}$ by

$$\begin{aligned} &V = H_0^1(\Omega)^2, \\ &Q = \{\boldsymbol{q} = (q_{ij}) | q_{ji} = q_{ij}, q_{ij} \in L^2(\Omega),\ \operatorname{tr} \boldsymbol{q} = 0 \quad \text{a.e. in } \Omega\}, \\ &\overline{M} = L^2(\Omega) \end{aligned} \tag{16}$$

and the product space $\overline{Z}$ and the convex set $\overline{W}$ by

$$\begin{aligned} &\overline{Z} = V \times Q \times \overline{M} \\ &\overline{W} = \{(\boldsymbol{v}, \boldsymbol{q}, \eta) \in \overline{Z} : |\boldsymbol{q}| \leqslant \eta \text{ a.e. in } \Omega\}. \end{aligned} \tag{17}$$

For any function $\phi : \Omega \times [0, T] \to \mathbb{R}$, we will use the notation $\phi(t)$ for the function $\boldsymbol{x} \in \Omega \to \phi(\boldsymbol{x}, t)$.

Let $1 \leqslant p \leqslant \infty$. For any Banach space X, we define the spaces

$$\begin{aligned} &L^p(0, T; X) = \left\{ v : [0, T] \to X \text{ measurable:} \quad \int_0^T \|v(t)\|_X^p \,\mathrm{d}t < \infty \right\} \quad 1 \leqslant p < \infty, \\ &L^\infty(0, T; X) = \left\{ v : [0, T] \to\ X \text{ measurable:} \quad \exists C > 0 \quad \|v(t)\|_X \leqslant C \text{ a.e. } t \in [0, T] \right\}. \end{aligned}$$

These are Banach spaces equipped respectively with the norms

$$\begin{aligned} &\|v\|_{L^p(0,T;X)} = \left(\int_0^T \|v(t)\|_X^p \,\mathrm{d}t \right)^{1/p} \quad \text{for } 1 \leqslant p < \infty, \\ &\|v\|_{L^\infty(0,T;X)} = \operatorname*{ess\,sup}_{0 \leqslant t \leqslant T} \|v(t)\|_X. \end{aligned}$$

We also define the space

$$H^1(0, T; X) = \{v \in L^2(0, T : X) | \dot{v} \in L^2(0, T; X)\},$$

in which the time derivative $\dot{v}$ is defined in a weak sense. This is a Hilbert space with the inner product and norm

$$\begin{aligned} &(u, v)_{H^1(0,T;X)} = \int_0^T [(u(t), v(t))_X + (\dot{u}(t), \dot{v}(t))_X] \,\mathrm{d}x, \\ &\|v\|_{H^1(0,T;X)} = (v, v)_{H^1(0,T;X)}^{1/2}. \end{aligned}$$

The variational inequality corresponding to the primal formulation in classical plasticity (see [1]) is that of finding $\boldsymbol{w} := (\boldsymbol{u}, \boldsymbol{p}, \gamma) : [0, T] \to \overline{Z}$ such that $\boldsymbol{w}(0) = \boldsymbol{0}$, $\dot{\boldsymbol{w}}(t) \in \overline{W}$ for almost every $t \in [0, T]$ and

$$\bar{a}(\boldsymbol{w}(t), \boldsymbol{z} - \dot{\boldsymbol{w}}(t)) + j(\boldsymbol{z}) - j(\dot{\boldsymbol{w}}) \geqslant \langle \ell(t), \boldsymbol{z} - \dot{\boldsymbol{w}} \rangle \quad \forall \boldsymbol{z} \in \overline{Z} \tag{18}$$

in which the bilinear form and functionals are given by

$$\begin{aligned} &\bar{a} : \overline{Z} \times \overline{Z} \to \mathbb{R}, \bar{a}(\boldsymbol{w}, \boldsymbol{z}) = \int_\Omega (\mathscr{C}(\boldsymbol{\epsilon}(\boldsymbol{u}) - \boldsymbol{p}) : (\boldsymbol{\epsilon}(\boldsymbol{v}) - \boldsymbol{q}) + k_1 \boldsymbol{p} : \boldsymbol{q} + k_2 \gamma \eta) \,\mathrm{d}x \quad \forall \boldsymbol{w} = (\boldsymbol{u}, \boldsymbol{p}, \gamma), \boldsymbol{z} = (\boldsymbol{v}, \boldsymbol{q}, \eta) \in \overline{Z}, \\ &j : \overline{Z} \to \mathbb{R},\ j(\boldsymbol{z}) = \begin{cases} \int_\Omega \kappa |\boldsymbol{q}| \mathrm{d}x & \text{if } \boldsymbol{z} \in \overline{W} \quad \forall \boldsymbol{z} = (\boldsymbol{v}, \boldsymbol{q}, \eta) \in \overline{Z}, \\ \infty & \text{otherwise} \end{cases} \\ &\ell : \overline{Z} \to \mathbb{R},\ \langle \ell, \boldsymbol{z} \rangle = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} \,\mathrm{d}x \quad \forall \boldsymbol{z} = (\boldsymbol{v}, \boldsymbol{q}, \eta) \in \overline{Z}. \end{aligned} \tag{19}$$

Note that by simple integration, the condition $\dot{\boldsymbol{w}}(t) \in \overline{W}$ implies also that $\boldsymbol{w}(t) \in \overline{W}$.

The following result is proved in [1].

Theorem 2.1. *Assume that* $\boldsymbol{f} \in H^1(0,T;H^1(\Omega)^2)$ *and* $\boldsymbol{f}(0) = \boldsymbol{0}$. *Then the variational inequality* (18) *has a unique solution* $(\boldsymbol{u}, \boldsymbol{p}, \gamma) \in H^1(0,T;V) \times H^1(0,T;Q) \times H^1(0,T;\overline{M})$.

Remarks

1. The proof makes use of the continuity of the bilinear form and functionals, the coercivity of $\bar{a}(\cdot,\cdot)$ (for which one of $k_1 > 0$ or $k_2 > 0$ is essential), and the weak lower semi-continuity of $j(\cdot)$, which follows from its convexity and its strong lower semi-continuity (the latter follows from the closedness of the set $\overline{W}$ and Fatou's Lemma).
2. The problem is posed on the whole space $\overline{Z}$ rather than on $\overline{W}$. This is legitimate since $\dot{\boldsymbol{w}}(t)$ is sought in $\overline{W}$, and this implies by integration that $\boldsymbol{w}(t) \in \overline{W}$. The functional j is extended by $+\infty$ to all $\overline{Z} \setminus \overline{W}$, consistent with the definition $(19)_2$, while it is also clear (set $\boldsymbol{z} = \boldsymbol{0}$ in (18)) that $j(\dot{\boldsymbol{w}})$ is bounded (see also [1] for further details).

2.2. Gradient plasticity

We consider next a simple strain gradient plasticity model, first introduced in [5] and studied computationally in [6,12], in which the classical yield condition (9) is replaced by one in which the yield condition depends also on the Laplacian of the scalar hardening parameter or equivalent plastic strain. That is, we still have (9), but the conjugate force g ($\bar{g}$ in $(10)_3$) is now given by

$$g = -k_2\gamma + k_3\nabla^2\gamma \tag{20}$$

in which k_3 is a positive constant and ∇^2 is the Laplacian operator. The dissipation inequality for this problem is formally given by

$$D(\boldsymbol{q},\eta) \geqslant D(\dot{\boldsymbol{p}},\dot{\gamma}) + (\boldsymbol{\sigma} - k_1\boldsymbol{p}) : (\boldsymbol{q} - \dot{\boldsymbol{p}}) - k_2\gamma(\eta - \dot{\gamma}) + k_3\nabla^2\gamma(\eta - \dot{\gamma}) \quad \forall(\boldsymbol{q},\eta). \tag{21}$$

In order to construct the relevant weak formulation of the problem corresponding to gradient plasticity we define the space M of scalar hardening parameters by

$$M = H_0^1(\Omega) \tag{22}$$

and we set

$$\begin{aligned} &Z = V \times Q \times M \\ &W = \{(\boldsymbol{v},\boldsymbol{q},\eta) \in Z : |\boldsymbol{q}| \leqslant \eta \text{ a.e. in } \Omega\}. \end{aligned} \tag{23}$$

We integrate (21) over Ω and perform integration by parts on the term involving k_3. We then obtain the dissipation inequality in weak form

$$j(\boldsymbol{z}) \geqslant j(\dot{\boldsymbol{w}}) + \int_\Omega (\boldsymbol{\sigma} - k_1\boldsymbol{p}) : (\boldsymbol{q} - \dot{\boldsymbol{p}})\,\mathrm{d}x - \int_\Omega k_2\gamma(\eta - \dot{\gamma})\,\mathrm{d}x - \int_\Omega k_3\nabla\gamma \cdot \nabla(\eta - \dot{\gamma})\,\mathrm{d}x \quad \forall \boldsymbol{z} \in Z \tag{24}$$

in which the definition of $j(\cdot)$ is unchanged. By combining the inequality (24) with the weak form of the equilibrium equation we arrive at the problem of finding $\boldsymbol{w} := (\boldsymbol{u},\boldsymbol{p},\gamma) : [0,T] \to Z$ which satisfies: $\boldsymbol{w}(0) = \boldsymbol{0}$, $\dot{\boldsymbol{w}}(t) \in W$ for almost every $t \in [0,T]$ and

$$a(\boldsymbol{w}(t), \boldsymbol{z} - \dot{\boldsymbol{w}}(t)) + j(\boldsymbol{z}) - j(\dot{\boldsymbol{w}}) \geqslant \langle \ell(t), \boldsymbol{z} - \dot{\boldsymbol{w}} \rangle \quad \forall \boldsymbol{z} \in Z \tag{25}$$

in which the bilinear form $a(\cdot,\cdot)$ is given by

$$a : Z \times Z \to \mathbb{R}, \quad a(\boldsymbol{w},\boldsymbol{z}) = \bar{a}(\boldsymbol{w},\boldsymbol{z}) + \int_\Omega k_3\nabla\gamma \cdot \nabla\eta\,\mathrm{d}x \tag{26}$$

with $\bar{a}$ defined in (19).

Remark. In contrast to the classical problem, the presence in the yield condition of a term involving a Laplacian leads to the requirement that γ be sought in the space $H^1(\Omega)$ rather than $L^2(\Omega)$. Furthermore, it is necessary to impose either a Dirichlet or Neumann boundary condition on γ, and here it is assumed that it is the homogeneous Dirichlet condition that is appropriate. Further discussion on the physical implications of boundary conditions on the equivalent plastic strain may be found in [23,24].

Again, following the corresponding proof for the classical problem in [1], it can be shown that for $\boldsymbol{f} \in H^1(0,T;L^2(\Omega)^2)$, the variational inequality (25) has a unique solution

$$\boldsymbol{w} = (\boldsymbol{u}, \boldsymbol{p}, \gamma) \in H^1(0, T; V) \times H^1(0, T; Q) \times H^1(0, T; M).$$

2.3. Regularity of solutions

It is important, both when dealing with finite element interpolation error estimates and when determining the consistency of DG formulations, to set out in clear terms the regularity of the exact solution to the problem. The question of the regularity of solutions to problems in plasticity is a much more subtle matter than the corresponding question for linear elasticity, say, and complete results are not available. We summarise some key regularity results for classical hardening plasticity, and indicate the assumptions that need to be made for the case of gradient plasticity.

First, we assume that

$$\boldsymbol{f} \in H^1(0, T; [H^1(\Omega)]^2). \tag{27}$$

Then it has been shown by Repin [25] for the Hencky problem (essentially one step in a family of incremental problems arrived at by discretization in time), that for a Lipschitz domain Ω and *hardening* plasticity,

$$\boldsymbol{u} \in [H^2_{\mathrm{loc}}(\Omega)]^2 \quad \text{and} \quad \boldsymbol{\sigma} \in [H^1_{\mathrm{loc}}(\Omega)]^{2\times 2}. \tag{28}$$

Additional smoothness of the domain leads to greater regularity. In particular, as shown in [25], if the boundary of Ω is C^3, then

$$\boldsymbol{u} \in [H^2(\Omega)]^2 \quad \text{and} \quad \boldsymbol{\sigma} \in [H^1(\Omega)]^{2\times 2}. \tag{29}$$

There are no corresponding results for the case of gradient plasticity.

We will assume in all cases that

$$\begin{aligned} &\boldsymbol{u} \in H^1(0, T; [H^2(\Omega)]^2),\\ &\boldsymbol{p} \in H^1(0, T; [H^1(\Omega)]^{2\times 2}),\\ &\gamma \in H^1(0, T; H^2(\Omega)). \end{aligned} \tag{30}$$

These assumptions are somewhat stronger than what is known for the case of classical plasticity, at least for Lipschitz domains, while in the case of gradient plasticity it is not unreasonable to assume that the regularizing effect of the gradient terms imparts a greater degree of regularity to the solution, when compared with the classical case.

3. A discontinuous Galerkin formulation

We denote by $\mathscr{P}_k(K)$ the space of polynomials of degree at most $k \geqslant 0$ on K. Let $\mathscr{T}_h = \{K\}$ be a shape-regular subdivision of Ω, where K are triangles. Set $h_K = \mathrm{diam}(K)$ and $h = \max\{h_K, K \in \mathscr{T}_h\}$.

Let $\mathscr{E}_h = \{e\}$ denote the set of the edges of $\mathscr{T}_h$, and $\mathscr{E}_h^o = \mathscr{E}_h \setminus \partial\Omega$ the set of all interior edges. We associate with each edge of an element K_i the outward unit normal vector $\boldsymbol{n}_i$. For an edge that lies on the boundary, $\boldsymbol{n}_i$ is defined to be the outward normal to $\partial\Omega$. The length of an edge $e \in \mathscr{E}_h$ is h_e.

For a positive integer m set

$$\begin{aligned} &H^m(\mathscr{T}_h) = \{v \in L^2(\Omega), v|_K \in H^m(K) \quad \forall K \in \mathscr{T}_h\},\\ &T(\mathscr{E}_h) = \prod_{K\in\mathscr{T}_h} L^2(\partial K). \end{aligned} \tag{31}$$

The jumps and the averages of $\eta \in L^2(\mathscr{E}_h)$, $\boldsymbol{v} \in L^2(\mathscr{E}_h)^2$ and $\boldsymbol{\tau} \in L^2(\mathscr{E}_h)^{2\times 2}$ across an edge e that is common to elements K_1 and K_2 are defined by

$$\begin{aligned} &[\![\eta]\!] = \eta_1\boldsymbol{n}_1 + \eta_2\boldsymbol{n}_2, \quad \{\eta\} = \frac{1}{2}(\eta_1 + \eta_2),\\ &[\![\boldsymbol{v}]\!] = \boldsymbol{v}_1 \otimes \boldsymbol{n}_1 + \boldsymbol{v}_2 \otimes \boldsymbol{n}_2, \quad \{\boldsymbol{v}\} = \frac{1}{2}(\boldsymbol{v}_1 + \boldsymbol{v}_2),\\ &[\![\boldsymbol{\tau}]\!] = \boldsymbol{\tau}_1\boldsymbol{n}_1 + \boldsymbol{\tau}_2\boldsymbol{n}_2, \quad \{\boldsymbol{\tau}\} = \frac{1}{2}(\boldsymbol{\tau}_1 + \boldsymbol{\tau}_2) \end{aligned} \tag{32}$$

in which η_i, $\boldsymbol{v}_i$ and $\boldsymbol{\tau}_i$ are the one-sided values of the quantities concerned along an edge $e \in \partial K_i$, while $\boldsymbol{n}_i$ is the outward unit normal vector to edge e on element K_i.

If e is an edge of element K_1 that lies on $\partial\Omega$, then the jumps and averages are defined by

$$\begin{aligned}
&[\![\eta]\!] = \eta_1 \boldsymbol{n}_1, \quad \{\eta\} = \eta_1,\\
&[\![\boldsymbol{v}]\!] = \boldsymbol{v}_1 \otimes \boldsymbol{n}_1, \quad \{\boldsymbol{v}\} = \boldsymbol{v}_1,\\
&[\![\boldsymbol{\tau}]\!] = \boldsymbol{\tau}_1 \boldsymbol{n}_1, \quad \{\boldsymbol{\tau}\} = \boldsymbol{\tau}_1.
\end{aligned} \tag{33}$$

The following identity relates the scalar product of two quantities to the products of their jumps and averages [26]:

$$\sum_{K \in \mathscr{T}_h} \int_{\partial K} \boldsymbol{v} \cdot \boldsymbol{\tau n}\,\mathrm{d}s = \sum_{e \in \mathscr{E}_h} \int_e [\![\boldsymbol{v}]\!] : \{\boldsymbol{\tau}\}\,\mathrm{d}s + \sum_{e \in \mathscr{E}_h^0} \int_e \{\boldsymbol{v}\} \cdot [\![\boldsymbol{\tau}]\!]\,\mathrm{d}s. \tag{34}$$

In what follows we will make use of Young's inequality

$$ab \leqslant \frac{\epsilon}{2} a^2 + \frac{1}{2\epsilon} b^2, \quad \text{for } a, b \in \mathbb{R}^+ \quad \text{and} \quad \epsilon > 0 \tag{35}$$

and the inequalities (see [15])

$$\begin{aligned}
&\|v_h\|_e^2 \leqslant c_1 h_K^{-1} \|v_h\|_K^2 \quad \text{for } v_h \in \mathscr{P}_k(K),\\
&\|v\|_e^2 \leqslant c_2 (h_K^{-1} \|v\|_K^2 + h_K \|\nabla v\|_K^2) \quad \text{for } v \in H^1(\mathscr{T}_h),
\end{aligned} \tag{36}$$

where c_1 and c_2 are positive constants independent of h_K and e is an edge of K. Here and henceforth $\|\cdot\|_K$ and $\|\cdot\|_e$ will denote respectively the L^2-norms on an element K and edge e.

The following finite-dimensional spaces and subsets will be required:

$$\begin{aligned}
&V_h = \{\boldsymbol{v}_h \in L^2(\Omega)^2;\ \boldsymbol{v}_h|_K \in \mathscr{P}_1(K)^2 \quad \forall K \in \mathscr{T}_h\},\\
&Q_h = \{\boldsymbol{q}_h \in L^2(\Omega)^{2\times 2};\ q_h|_K \in \mathscr{P}_0(K)^{2\times 2} \quad \forall K \in \mathscr{T}_h\},\\
&\overline{M}_h = \{\eta_h \in L^2(\Omega);\ \eta_h|_K \in \mathscr{P}_0(K) \quad \forall K \in \mathscr{T}_h\},\\
&\overline{Z}_h = V_h \times Q_h \times \overline{M}_h,\\
&\overline{W}_h = \{\boldsymbol{z}_h = (\boldsymbol{v}_h, \boldsymbol{q}_h, \eta_h) \in \overline{Z}_h;\ |\boldsymbol{q}_h| \leqslant \eta_h \quad \text{in each } K \in \mathscr{T}_h\},\\
&M_h = \{\eta_h \in L^2(\Omega);\ \eta_h|_K \in \mathscr{P}_1(K) \quad \forall K \in \mathscr{T}_h\},\\
&Z_h = V_h \times Q_h \times M_h,\\
&W_h = \{\boldsymbol{z}_h = (\boldsymbol{v}_h, \boldsymbol{q}_h, \eta_h) \in Z_h;\ |\boldsymbol{q}_h| \leqslant \eta_h \quad \text{in each } K \in \mathscr{T}_h\}.
\end{aligned}$$

We introduce on $\overline{Z}_h$ and Z_h the norms $\|\cdot\|_{\bar{h}}$ and $\|\cdot\|_h$, defined for $\boldsymbol{z}_h = (\boldsymbol{v}_h, \boldsymbol{q}_h, \eta_h)$ by

$$\|\boldsymbol{z}_h\|_{\bar{h}}^2 = \sum_K (\|\boldsymbol{\epsilon}(\boldsymbol{v}_h)\|_K^2 + \|\boldsymbol{q}_h\|_K^2 + \|\eta_h\|_K^2) + \sum_e \frac{1}{h_e} \|[\![\boldsymbol{v}_h]\!]\|_e^2, \tag{37}$$

$$\|\boldsymbol{z}_h\|_h^2 = \|\boldsymbol{z}_h\|_{\bar{h}}^2 + \sum_K \|\nabla \eta_h\|_K^2 + \sum_e \frac{1}{h_e} \|[\![\eta_h]\!]\|_e^2. \tag{38}$$

To verify the positive-definiteness of $\|\cdot\|_{\bar{h}}$ we note that for $\boldsymbol{z} = (\boldsymbol{v}, \boldsymbol{q}, \eta) \in Z_h$ and $\|\boldsymbol{z}\|_{\bar{h}} = 0$, $\boldsymbol{q} = \boldsymbol{0}$, $\eta = 0$ and $\|\boldsymbol{\epsilon}(\boldsymbol{v})\|_K = 0$. Thus $\boldsymbol{v}|_K$ is a rigid body motion on K, that is $\boldsymbol{v} = \boldsymbol{a}_K + \boldsymbol{b}_K \times \boldsymbol{x}$, where $\boldsymbol{a}_K$ and $\boldsymbol{b}_K$ are constant in each element K. But $\|[\![\boldsymbol{v}]\!]\|_e = 0$ for each interior edge e, so that $\boldsymbol{a}$ and $\boldsymbol{b}$ are independent of K. Finally, on the boundary edges $\|[\![\boldsymbol{v}]\!]\|_e = \|\boldsymbol{v}\|_e = 0$, so that $\boldsymbol{a} = \boldsymbol{b} = \boldsymbol{0}$. Thus $\|\boldsymbol{z}\|_{\bar{h}} = 0$ implies $\boldsymbol{v} = \boldsymbol{0}$, $\boldsymbol{q} = \boldsymbol{0}$ and $\eta = 0$. The positive-definiteness of $\|\cdot\|_h$ follows then from (38).

We consider a symmetric interior penalty formulation and introduce the bilinear forms and functionals

$$\begin{aligned}
&\bar{a}_h((\boldsymbol{u}_h, \boldsymbol{p}_h, \gamma_h), (\boldsymbol{v}_h, \boldsymbol{q}_h, \eta_h))\\
&\quad = \sum_K \int_K (\mathscr{C}(\boldsymbol{\epsilon}(\boldsymbol{u}_h) - \boldsymbol{p}_h) : (\boldsymbol{\epsilon}(\boldsymbol{v}_h) - \boldsymbol{q}_h) + k_1 \boldsymbol{p}_h : \boldsymbol{q}_h + k_2 \gamma_h \eta_h)\,\mathrm{d}x\\
&\qquad - \sum_e \int_e (\{\mathscr{C}(\boldsymbol{\epsilon}(\boldsymbol{u}_h) - \boldsymbol{p}_h)\} : [\![\boldsymbol{v}_h]\!] + \{\mathscr{C}(\boldsymbol{\epsilon}(\boldsymbol{v}_h) - \boldsymbol{q}_h)\} : [\![\boldsymbol{u}_h]\!])\,\mathrm{d}s\\
&\qquad + \sum_e \frac{\beta_1}{h_e} \int_e [\![\boldsymbol{u}_h]\!] : [\![\boldsymbol{v}_h]\!]\,\mathrm{d}s,
\end{aligned} \tag{39}$$

$$
\begin{aligned}
&a_h((\boldsymbol{u}_h,\boldsymbol{p}_h,\gamma_h),(\boldsymbol{v}_h,\boldsymbol{q}_h,\eta_h))\\
&\quad=\bar{a}_h((\boldsymbol{u}_h,\boldsymbol{p}_h,\gamma_h),(\boldsymbol{v}_h,\boldsymbol{q}_h,\eta_h))+\sum_K\int_K k_3\nabla\gamma_h\cdot\nabla\eta_h\,\mathrm{d}x\\
&\qquad-\sum_e\int_e k_3(\{\nabla\gamma_h\}\cdot[\![\eta_h]\!]+\{\nabla\eta_h\}\cdot[\![\gamma_h]\!])\,\mathrm{d}s+\sum_e\frac{\beta_2}{h_e}\int_e[\![\gamma_h]\!]\cdot[\![\eta_h]\!]\,\mathrm{d}s,
\end{aligned}
\tag{40}
$$

$$
j(\boldsymbol{v}_h,\boldsymbol{q}_h,\gamma_h)=\begin{cases}\sum_K\int_K\kappa|\boldsymbol{q}_h|\,\mathrm{d}x & \text{if } \boldsymbol{w}_h\in W_h,\\ +\infty & \text{otherwise,}\end{cases}
\tag{41}
$$

where β_1 and β_2 are positive penalty parameters.

The semi-discrete DG approximations corresponding to problems (18) and (25) are then as follows: given that $\boldsymbol{w}_h(0)=\boldsymbol{0}$, find $\boldsymbol{w}_h:[0,T]\to\overline{Z}_h$ such that for almost all $t\in(0,T)$, $\dot{\boldsymbol{w}}_h(t)\in\overline{W}_h$ and

$$
\bar{a}_h(\boldsymbol{w}_h(t),\boldsymbol{z}_h-\dot{\boldsymbol{w}}_h(t))+j(\boldsymbol{z}_h)-j(\dot{\boldsymbol{w}}_h(t))\geqslant\langle\ell(t),\boldsymbol{z}_h-\dot{\boldsymbol{w}}_h(t)\rangle\quad\forall\boldsymbol{z}_h\in\overline{Z}_h,
\tag{42}
$$

and given that $\boldsymbol{w}_h(0)=\boldsymbol{0}$, find $\boldsymbol{w}_h:[0,T]\to Z_h$ such that for almost all $t\in(0,T)$, $\dot{\boldsymbol{w}}_h(t)\in W_h$ and

$$
a_h(\boldsymbol{w}_h(t),\boldsymbol{z}_h-\dot{\boldsymbol{w}}_h(t))+j(\boldsymbol{z}_h)-j(\dot{\boldsymbol{w}}_h(t))\geqslant\langle\ell(t),\boldsymbol{z}_h-\dot{\boldsymbol{w}}_h(t)\rangle\quad\forall\boldsymbol{z}_h\in Z_h.
\tag{43}
$$

Lemma 3.1 (Consistency). *Let $\bar{\boldsymbol{w}}$ and $\boldsymbol{w}$ be respectively the solutions of (18) and (25). Then*

$$
\begin{aligned}
&\bar{a}_h(\bar{\boldsymbol{w}}(t),\boldsymbol{z}-\dot{\bar{\boldsymbol{w}}}(t))+j(\boldsymbol{z})-j(\dot{\bar{\boldsymbol{w}}}(t))\geqslant\langle\ell(t),\boldsymbol{z}-\dot{\bar{\boldsymbol{w}}}(t)\rangle,\quad\forall\boldsymbol{z}\in\overline{Z}_h,\\
&a_h(\boldsymbol{w}(t),\boldsymbol{z}-\dot{\boldsymbol{w}}(t))+j(\boldsymbol{z})-j(\dot{\boldsymbol{w}}(t))\geqslant\langle\ell(t),\boldsymbol{z}-\dot{\boldsymbol{w}}(t)\rangle,\quad\forall\boldsymbol{z}\in Z_h.
\end{aligned}
\tag{44}
$$

Proof. We present the proof for gradient plasticity only; the proof for classical plasticity follows along similar lines.

Let $\boldsymbol{z}_h=(\boldsymbol{v}_h,\boldsymbol{q}_h,\eta_h)\in Z_h$. From the regularity assumption (30) the equilibrium equation holds pointwise. Multiplying this equation by $\boldsymbol{v}_h-\dot{\boldsymbol{u}}(t)$, integrating by parts on $K\in\mathscr{T}_h$, summing over $K\in\mathscr{T}_h$, and using (34), we obtain

$$
\begin{aligned}
&\sum_K\int_K\boldsymbol{\sigma}(t):\boldsymbol{\epsilon}(\boldsymbol{v}_h(t)-\dot{\boldsymbol{u}}(t))\,\mathrm{d}x-\sum_{e\in\mathscr{E}_h^0}\int_e[\![\boldsymbol{\sigma}(t)]\!]:\{\boldsymbol{v}_h(t)-\dot{\boldsymbol{u}}(t)\}\,\mathrm{d}s-\sum_{e\in\mathscr{E}_h}\int_e\{\boldsymbol{\sigma}(t)\}:[\![\boldsymbol{v}_h(t)-\dot{\boldsymbol{u}}(t)]\!]\,\mathrm{d}s\\
&\quad=\sum_K\int_K\boldsymbol{f}\cdot(\boldsymbol{v}_h(t)-\dot{\boldsymbol{u}}(t))\,\mathrm{d}x.
\end{aligned}
\tag{45}
$$

Similarly, from the dissipation inequality (21), we get

$$
\begin{aligned}
j(\boldsymbol{z}_h)-j(\dot{\boldsymbol{w}}(t))\geqslant{}&\sum_K\int_K\boldsymbol{\sigma}(t):(\boldsymbol{q}_h-\dot{\boldsymbol{p}}(t))\,\mathrm{d}x-k_1\sum_K\int_K\boldsymbol{p}(t):(\boldsymbol{q}_h-\dot{\boldsymbol{p}}(t))-k_2\sum_K\int_K\gamma(t)(\eta_h-\dot{\gamma}(t))\\
&-k_3\sum_K\int_K\nabla\gamma(t)\cdot\nabla(\eta_h-\dot{\gamma}(t))\,\mathrm{d}x+k_3\sum_{e\in\mathscr{E}_h^0}\int_e\mathrm{tr}([\![\nabla\gamma(t)]\!])\cdot\{\eta_h-\dot{\gamma}(t)\}\,\mathrm{d}s\\
&+k_3\sum_{e\in\mathscr{E}_h}\int_e\{\nabla\gamma(t)\}\cdot[\![\eta_h-\dot{\gamma}(t)]\!]\,\mathrm{d}s,
\end{aligned}
\tag{46}
$$

where for any $\boldsymbol{v}\in[L^2(\mathscr{E}_h)]^2$, we have that $\mathrm{tr}([\![\boldsymbol{v}]\!])=\boldsymbol{v}_1\cdot\boldsymbol{n}_1+\boldsymbol{v}_2\cdot\boldsymbol{n}_2$ is the trace of the matrix $[\![\boldsymbol{v}]\!]$ defined in (32).

From (45) and (46) and the definition of a_h we easily obtain $(44)_2$ provided that the jump terms involving $\boldsymbol{\sigma}$ and $\nabla\gamma$ vanish. From (30) both of these quantities are in $H^1(\Omega)$, and hence in $H^1_{\mathrm{loc}}(\Omega)$. We use a result due to Evans and Gariepy ([27], Section 4.9.2, Theorem 2), according to which functions in $H^1_{\mathrm{loc}}(\Omega)$ are continuous across interior edges in Ω. Thus the jump terms involving $\boldsymbol{\sigma}$ and $\nabla\gamma$ vanish. This completes the proof. □

We have the following.

Lemma 3.2 (Well-posedness). *The problems (42) and (43) each have exactly one solution.*

Proof. It suffices to show that ℓ is continuous, j is convex, and lower semi-continuous and the bilinear forms $\bar{a}_h(\cdot,\cdot)$ and $a_h(\cdot,\cdot)$ are continuous and coercive with respect to the norms in (37) and (38). □

As in the previous lemma, we present the proof for the case of gradient plasticity. We first observe [15] that there exists a positive constant α, independent of the mesh size h, such that

$$\|\boldsymbol{v}_h\|_{L^2(\Omega)} \leqslant \alpha\left(\sum_K \|\boldsymbol{\epsilon}(\boldsymbol{v}_h)\|_K^2 + \sum_e \frac{1}{h_e}\|[\![\boldsymbol{v}_h]\!]\|_e^2\right)^{1/2}. \tag{47}$$

Hence

$$|\langle \ell, \boldsymbol{z}_h\rangle| \leqslant \left(\sum_K \|\boldsymbol{f}\|_K^2\right)^{1/2}\left(\sum_K \|\boldsymbol{v}_h\|_K^2\right)^{1/2} \leqslant c\left(\sum_K \|\boldsymbol{f}\|_K^2\right)^{1/2}\|\boldsymbol{z}_h\|_h.$$

Next, we note that j is proper and convex, and is easily shown to be lower semi-continuous on Z_h, with respect to the norm $\|\cdot\|_h$.

Now, we show that a_h is continuous with respect to the norm $\|\cdot\|_h$.

Take $\boldsymbol{w}_h = (\boldsymbol{u}_h, \boldsymbol{p}_h, \gamma_h) \in Z_h$ and $\boldsymbol{z}_h = (\boldsymbol{v}_h, \boldsymbol{q}_h, \eta_h) \in Z_h$. We first rewrite a_h as

$$\begin{aligned}
a_h(\boldsymbol{w}_h, \boldsymbol{z}_h) = {} & \left.\begin{aligned}&\sum_K \int_K (\mathscr{C}(\boldsymbol{\epsilon}(\boldsymbol{u}_h) - \boldsymbol{p}_h) : (\boldsymbol{\epsilon}(\boldsymbol{v}_h) - \boldsymbol{q}_h)\,\mathrm{d}x + k_1 \boldsymbol{p}_h : \boldsymbol{q}_h)\,\mathrm{d}x \\ &+ \sum_K \int_K (k_2 \gamma_h \eta_h + k_3 \nabla\gamma_h \cdot \nabla\eta_h)\,\mathrm{d}x\end{aligned}\right\} \quad (Q_1) \\
& - \sum_e \int_e (\{\mathscr{C}(\boldsymbol{\epsilon}(\boldsymbol{u}_h) - \boldsymbol{p}_h)\} : [\![\boldsymbol{v}_h]\!] + \{\mathscr{C}(\boldsymbol{\epsilon}(\boldsymbol{v}_h) - \boldsymbol{q}_h)\} : [\![\boldsymbol{u}_h]\!])\,\mathrm{d}s \quad (Q_2) \\
& - \sum_e \int_e k_3(\{\nabla\gamma_h\} \cdot [\![\eta_h]\!] + \{\nabla\eta_h\} \cdot [\![\gamma_h]\!])\,\mathrm{d}s \quad (Q_3) \\
& + \sum_e \frac{\beta_2}{h_e}\int_e [\![\gamma_h]\!] \cdot [\![\eta_h]\!]\,\mathrm{d}s + \frac{\beta_1}{h_e}\int_e [\![\boldsymbol{u}_h]\!] : [\![\boldsymbol{v}_h]\!]\,\mathrm{d}s. \quad (Q_4)
\end{aligned} \tag{48}$$

We now need to estimate Q_1, Q_2, Q_3 and Q_4. First, using the norm of $\mathscr{C}$ and the triangle and Minkowski's inequalities, we have

$$\begin{aligned}
|Q_1| &\leqslant \max(|\mathscr{C}|_\infty, k_1, k_2, k_3)\sum_K((\|\boldsymbol{\epsilon}(\boldsymbol{u}_h)\|_K + \|\boldsymbol{p}_h\|_K)(\|\boldsymbol{\epsilon}(\boldsymbol{v}_h)\|_K + \|\boldsymbol{q}_h\|_K) + \|\boldsymbol{p}_h\|_K\|\boldsymbol{q}_h\|_K + \|\gamma_h\|_{1,K}\|\eta_h\|_{1,K}) \\
&\leqslant c\max(\lambda + 2\mu, k_1, k_2, k_3)\|\boldsymbol{w}_h\|_h\|\boldsymbol{z}_h\|_h.
\end{aligned} \tag{49}$$

Next,

$$\begin{aligned}
|Q_2| \leqslant {} & \left(\sum_e h_e\|\{\mathscr{C}(\boldsymbol{\epsilon}(\boldsymbol{u}_h) - \boldsymbol{p}_h)\}\|_e^2\right)^{1/2}\left(\sum_e h_e^{-1}\|[\![\boldsymbol{v}_h]\!]\|_e^2\right)^{1/2} \\
& + \left(\sum_e h_e\|\{\mathscr{C}(\boldsymbol{\epsilon}(\boldsymbol{v}_h) - \boldsymbol{q}_h)\}\|_e^2\right)^{1/2}\left(\sum_e h_e^{-1}\|[\![\boldsymbol{u}_h]\!]\|_e^2\right)^{1/2}.
\end{aligned} \tag{50}$$

To bound the first term in (50), we use the triangle inequality on the averaging operator $\{\ \}$, and $(36)_1$. For $e \in \mathscr{E}_h^0$, take $e = \partial K_1 \cap \partial K_2$, and set $K_{12} = K_1 \cup K_2$. We have

$$h_e\|\{\mathscr{C}(\boldsymbol{\epsilon}(\boldsymbol{u}_h) - \boldsymbol{p}_h)\}\|_e^2 = h_e\left\|\mathscr{C}\left(\boldsymbol{\epsilon}\left(\frac{\boldsymbol{u}_{1h} + \boldsymbol{u}_{2h}}{2}\right) - \left(\frac{\boldsymbol{p}_{1h} + \boldsymbol{p}_{2h}}{2}\right)\right)\right\|_e^2 \leqslant c_1|\mathscr{C}|_\infty^2(\|\boldsymbol{\epsilon}(\boldsymbol{u}_h)\|_{K_{12}}^2 + \|\boldsymbol{p}_h\|_{K_{12}}^2).$$

Summing over all edges $e \in \mathscr{E}_h$, we have

$$\sum_e h_e\|\{\mathscr{C}(\boldsymbol{\epsilon}(\boldsymbol{u}_h) - \boldsymbol{p}_h)\}\|_e^2 \leqslant c_1|\mathscr{C}|_\infty^2\|\boldsymbol{w}_h\|_h^2.$$

Therefore,

$$|Q_2| \leqslant c_1^{1/2}(\lambda + 2\mu)\|\boldsymbol{w}_h\|_h\|\boldsymbol{z}_h\|_h. \tag{51}$$

Similarly,

$$|Q_3| \leqslant c_1^{1/2}(\lambda + 2\mu)\|\boldsymbol{w}_h\|_h\|\boldsymbol{z}_h\|_h. \tag{52}$$

We also have

$$
\begin{aligned}
|Q_4| &= \left|\sum_e \frac{\beta_2}{h_e}\int_e [\![\gamma_h]\!]\cdot[\![\eta_h]\!]\,\mathrm{d}s + \frac{\beta_1}{h_e}\int_e [\![\boldsymbol{u}_h]\!] : [\![\boldsymbol{v}_h]\!]\,\mathrm{d}s\right| \\
&\leqslant \beta_1\left(\sum_e h_e^{-1}\|[\![\boldsymbol{u}_h]\!]\|_e^2\right)^{1/2}\left(\sum_e h_e^{-1}\|[\![\boldsymbol{v}_h]\!]\|_e^2\right)^{1/2} + \beta_2\left(\sum_e h_e^{-1}\|\gamma_h\|_e^2\right)^{1/2}\left(\sum_e h_e^{-1}\|\eta_h\|_e^2\right)^{1/2} \\
&\leqslant (\beta_1+\beta_2)\|\boldsymbol{w}_h\|_h\|\boldsymbol{z}_h\|_h. \qquad (53)
\end{aligned}
$$

Adding (49)–(53), we find a constant $c = c(k_1, k_2, k_3, \lambda, \mu, \beta_1, \Omega) > 0$ such that

$$
|a_h(\boldsymbol{w}_h, \boldsymbol{z}_h)| \leqslant c(k_1, k_2, k_3, \lambda, \mu, \beta_1, \Omega)\|\boldsymbol{w}_h\|_h\|\boldsymbol{z}_h\|_h \quad \forall \boldsymbol{w}_h, \boldsymbol{z}_h \in Z_h. \qquad (54)
$$

Thus a_h is continuous.

Concerning the coercivity of a_h we write, for every $\boldsymbol{w}_h = (\boldsymbol{u}_h, \boldsymbol{p}_h, \gamma_h)$ and $\boldsymbol{z}_h = (\boldsymbol{v}_h, \boldsymbol{q}_h, \eta_h) \in Z_h$,

$$
a_h(\boldsymbol{w}_h, \boldsymbol{z}_h) = b(\boldsymbol{u}_h, \boldsymbol{v}_h) - c(\boldsymbol{p}_h, \boldsymbol{v}_h) - c(\boldsymbol{q}_h, \boldsymbol{u}_h) + d((\boldsymbol{p}_h, \gamma_h), (\boldsymbol{q}_h, \eta_h)), \qquad (55)
$$

where

$$
\begin{aligned}
b(\boldsymbol{u}_h, \boldsymbol{v}_h) &= \sum_K \int_K \mathscr{C}\boldsymbol{\epsilon}(\boldsymbol{u}_h) : \boldsymbol{\epsilon}(\boldsymbol{v}_h)\,\mathrm{d}x - \sum_e \int_e \{\mathscr{C}\boldsymbol{\epsilon}(\boldsymbol{u}_h)\} : [\![\boldsymbol{v}_h]\!]\,\mathrm{d}s \\
&\quad - \sum_e \int_e \{\mathscr{C}\boldsymbol{\epsilon}(\boldsymbol{v}_h)\} : [\![\boldsymbol{u}_h]\!]\,\mathrm{d}s + \sum_e \frac{\beta_1}{h_e}\int_e [\![\boldsymbol{v}_h]\!] : [\![\boldsymbol{v}_h]\!]\,\mathrm{d}s, \qquad (56)
\end{aligned}
$$

$$
c(\boldsymbol{p}_h, \boldsymbol{v}_h) = \sum_K \int_K \mathscr{C}\boldsymbol{p}_h : \boldsymbol{\epsilon}(\boldsymbol{v}_h)\,\mathrm{d}x - \sum_e \int_e \{\mathscr{C}\boldsymbol{p}_h\} : [\![\boldsymbol{v}_h]\!]\,\mathrm{d}s \qquad (57)
$$

$$
\begin{aligned}
d((\boldsymbol{p}_h, \gamma_h), (\boldsymbol{q}_h, \eta_h)) &= \sum_K \int_K (\mathscr{C}\boldsymbol{p}_h : \boldsymbol{q}_h + k_1\boldsymbol{p}_h : \boldsymbol{q}_h)\,\mathrm{d}x \\
&\quad - \sum_e \int_e k_3(\{\nabla\gamma_h\}\cdot[\![\eta_h]\!] + \{\nabla\eta_h\}\cdot[\![\gamma_h]\!])\,\mathrm{d}s \\
&\quad + \sum_e \frac{\beta_2}{h_e}\int_e [\![\gamma_h]\!]\cdot[\![\eta_h]\!]\,\mathrm{d}s \\
&\quad + \sum_K \int_K (k_2\gamma_h\eta_h + k_3\nabla\gamma_h\cdot\nabla\eta_h)\,\mathrm{d}x. \qquad (58)
\end{aligned}
$$

The coercivity of a_h is obtained from the following lemma.

Lemma 3.3. *Let $c_1 > 0$ be the constant in (36)$_1$. For some positive constants η_1 and η_2 suitably chosen so that*

$$
r_1 := \min\left(2\mu - \eta_1, \beta_1 - \frac{c_1(\lambda+2\mu)^2}{\eta_1}\right) > 0, \qquad (59)
$$

$$
r_2 := \min\left(2\mu + k_1, \beta_2 - \frac{k_3}{\eta_2}, \min(k_2, k_3(1 - c_1\eta_2))\right) > 0, \qquad (60)
$$

we have

$$
\begin{aligned}
&b(\boldsymbol{v}_h, \boldsymbol{v}_h) \geqslant r_1\|\boldsymbol{v}_h\|_h^2, \\
&d((\boldsymbol{q}_h, \eta_h), (\boldsymbol{q}_h, \eta_h)) \geqslant r_2\|(\boldsymbol{q}_h, \eta_h)\|_h^2, \qquad (61) \\
&c(\boldsymbol{q}_h, \boldsymbol{v}_h) \leqslant (\lambda + 2\mu)(1 + c_1^{1/2})\|\boldsymbol{q}_h\|_0\|\boldsymbol{v}_h\|_h,
\end{aligned}
$$

where

$$
\|\boldsymbol{v}_h\|_h^2 = \sum_K \|\boldsymbol{\epsilon}(\boldsymbol{v}_h)\|_K^2 + \sum_e h_e^{-1}\|[\![\boldsymbol{v}_h]\!]\|_e^2, \qquad (62)
$$

$$
\|(\boldsymbol{q}_h, \eta_h)\|_h^2 = \sum_K (\|\boldsymbol{q}_h\|_K^2 + \|\eta_h\|_{1,K}^2) + \sum_e h_e^{-1}\|[\![\eta_h]\!]\|_e^2. \qquad (63)
$$

Proof. We have

$$
\begin{aligned}
b(\boldsymbol{v}_h,\boldsymbol{v}_h) &= \sum_K \int_K \mathscr{C}\boldsymbol{\epsilon}(\boldsymbol{v}_h):\boldsymbol{\epsilon}(\boldsymbol{v}_h)\,\mathrm{d}x - 2\sum_e \int_e \{\mathscr{C}\boldsymbol{\epsilon}(\boldsymbol{v}_h)\}:[\![\boldsymbol{v}_h]\!]\,\mathrm{d}s + \sum_e \frac{\beta_1}{h_e}\int_e [\![\boldsymbol{v}_h]\!]:[\![\boldsymbol{v}_h]\!]\,\mathrm{d}s \\
&\geqslant 2\mu\sum_K \|\boldsymbol{\epsilon}(\boldsymbol{v}_h)\|_K^2 + \beta_1\sum_e h_e^{-1}\|[\![\boldsymbol{v}_h]\!]\|_e^2 - 2\left(\sum_e h_e\|\{\mathscr{C}\boldsymbol{\epsilon}(\boldsymbol{v}_h)\}\|_e^2\right)^{1/2}\left(\sum_e h_e^{-1}\|[\![\boldsymbol{v}_h]\!]\|_e^2\right)^{1/2} \\
&\geqslant (2\mu-\eta_1)\sum_K \|\boldsymbol{\epsilon}(\boldsymbol{v}_h)\|_K^2 + \left(\beta_1 - \frac{c_1(\lambda+2\mu)^2}{\eta_1}\right)\sum_e h_e^{-1}\|[\![\boldsymbol{v}_h]\!]\|_e^2 \quad \text{from } (36)_1 \text{ and } (35) \\
&\geqslant \min\left(2\mu-\eta_1, \beta_1 - \frac{c_1(\lambda+2\mu)^2}{\eta_1}\right)\|\boldsymbol{v}_h\|_h^2 = r_1\|\boldsymbol{v}_h\|^2,
\end{aligned} \tag{64}
$$

provided that we choose η_1 and β_1 such that

$$
\beta_1 > \frac{c_1(\lambda+2\mu)^2}{2\mu} \quad \text{and} \quad \frac{c_1(\lambda+2\mu)^2}{\beta_1} < \eta_1 < 2\mu. \tag{65}
$$

Next,

$$
\begin{aligned}
|c(\boldsymbol{p}_h,\boldsymbol{v}_h)| &= \left|\sum_K \int_K \mathscr{C}\boldsymbol{p}_h:\boldsymbol{\epsilon}(\boldsymbol{v}_h)\mathrm{d}x - \sum_e\int_e\{\mathscr{C}\boldsymbol{p}_h\}:[\![\boldsymbol{v}_h]\!]\,\mathrm{d}s\right| \\
&\leqslant |\mathscr{C}|_\infty \sum_K \|\boldsymbol{q}_h\|_K\|\boldsymbol{\epsilon}(\boldsymbol{v}_h)\|_K + \sum_e \|\{\mathscr{C}\boldsymbol{q}_h\}\|_e\|[\![\boldsymbol{v}_h]\!]\|_e \\
&\leqslant |\mathscr{C}|_\infty\left(\sum_K\|\boldsymbol{q}_h\|_K^2\right)^{1/2}\left(\sum_K\|\boldsymbol{\epsilon}(\boldsymbol{v}_h)\|_K^2\right)^{1/2} + c|\mathscr{C}|_\infty\left(\sum_K\|\boldsymbol{q}_h\|_K^2\right)^{1/2}\left(\sum_e h_e^{-1}\|[\![\boldsymbol{v}_h]\!]\|_e^2\right)^{1/2} \quad \text{from } (36)_1 \\
&\leqslant (\lambda+2\mu)(1+c_1^{1/2})\|\boldsymbol{q}_h\|_0\|\boldsymbol{v}_h\|_h.
\end{aligned} \tag{66}
$$

Finally, we have

$$
\begin{aligned}
d((\boldsymbol{q}_h,\eta_h),(\boldsymbol{q}_h,\eta_h)) &= \sum_K\int_K(\mathscr{C}\boldsymbol{q}_h:\boldsymbol{q}_h + k_1\boldsymbol{q}_h:\boldsymbol{q}_h)\,\mathrm{d}x - 2\sum_e\int_e k_3\{\nabla\eta_h\}\cdot[\![\eta_h]\!]\,\mathrm{d}s + \sum_e\frac{\beta_2}{h_e}\int_e[\![\eta_h]\!]\cdot[\![\eta_h]\!]\,\mathrm{d}s \\
&\quad + \sum_K\int_K(k_2\eta_h\eta_h + k_3\nabla\eta_h\cdot\nabla\eta_h)\,\mathrm{d}x \\
&\geqslant \sum_K(2\mu+k_1)\|\boldsymbol{q}_h\|_K^2 - 2k_3\left(\sum_e h_e\|\{\nabla\eta_h\}\|_e^2\right)^{1/2}\left(\sum_e h_e^{-1}\|[\![\eta_h]\!]\|_e^2\right)^{1/2} \\
&\quad + \sum_e\frac{\beta_2}{h_e}\|[\![\eta_h]\!]\|_e^2 + \sum_K(k_2\|\eta_h\|_K^2 + k_3\|\nabla\eta_h\|_K^2) \\
&\geqslant (2\mu+k_1)\sum_K\|\boldsymbol{q}_h\|_K^2 + \left(\beta_2 - \frac{k_3}{\eta_2}\right)\sum_e h_e^{-1}\|[\![\eta_h]\!]\|_e^2 + \min(k_2, k_3 - c_1k_3\eta_2)\sum_K\|\eta_h\|_{1,K}^2 \\
&\quad \text{from } (36)_1 \text{ and } (35) \geqslant \min\left(2\mu+k_1, \beta_2 - \frac{k_3}{\eta_2}, \min(k_2, k_3-c_1k_3\eta_2)\right)\|(\boldsymbol{q}_h,\eta_h)\|_h^2 = r_2\|(\boldsymbol{q}_h,\eta_h)\|_h^2,
\end{aligned} \tag{67}
$$

provided that we choose η_2 and β_2 such that

$$
\beta_2 > c_1k_3 \quad \text{and} \quad \frac{k_3}{\beta_2} < \eta_2 < \frac{1}{c_1}. \tag{68}
$$

Now with η_1, η_2, β_1 and β_2 as in (65), (68) and (35) we obtain for every $\boldsymbol{z}_h = (\boldsymbol{v}_h, \boldsymbol{q}_h, \eta_h) \in Z_h$

$$
\begin{aligned}
a_h(\boldsymbol{z}_h,\boldsymbol{z}_h) &= b(\boldsymbol{v}_h,\boldsymbol{v}_h) - 2c(\boldsymbol{q}_h,\boldsymbol{v}_h) + d((\boldsymbol{q}_h,\eta_h),(\boldsymbol{q}_h,\eta_h)) \\
&\geqslant r_1\|\boldsymbol{v}_h\|_h^2 - (\lambda+2\mu)(1+c_1^{1/2})\eta_3\|\boldsymbol{q}_h\|_0^2 - \frac{(\lambda+2\mu)(1+c_1^{1/2})}{\eta_3}\|\boldsymbol{v}_h\|_h^2 + r_2(\|\boldsymbol{q}_h\|_0^2 + \|\eta_h\|_h^2) \\
&= \left[r_1 - \frac{(\lambda+2\mu)(1+c_1^{1/2})}{\eta_3}\right]\|\boldsymbol{v}_h\|_h^2 + r_2\|\eta_h\|_h^2 + [r_2 - (\lambda+2\mu)(1+c_1^{1/2})\eta_3]\|\boldsymbol{q}_h\|_0^2.
\end{aligned}
$$

Thus by taking η_3 such that

$$\frac{(\lambda+2\mu)(1+c_1^{1/2})}{r_1} < \eta_3 < \frac{r_2}{(\lambda+2\mu)(1+c_1^{1/2})}, \tag{69}$$

we find for some constant $c = c(k_1, k_2, k_3, \mu, \lambda, \beta_1, \beta_2, \Omega) > 0$ that

$$a_h(z_h, z_h) \geqslant c\|z_h\|_h^2. \qquad \square$$

Remark. The choice of the positive constants η_1, η_2 and η_3 is subject to the constraint that one of the hardening parameters k_1 or k_2 is sufficiently large. Therefore, the coercivity of the bilinear form a_h is obtained under the same condition.

4. Continuous-in-time a priori error estimate

In this section an a priori error estimate for the continuous-in-time problem is derived. We first collect some useful results.

For a scalar-valued function $\eta \in H^2(\Omega)$ let $\Pi_K : H^2(\Omega) \to P_1(K)$ denote the usual interpolation operator [28], which satisfies the error estimate

$$\|\eta - \Pi_K\eta\|_{H^1(K)} \leqslant ch_K|\eta|_{H^2(K)}. \tag{70}$$

The estimate is extended in a straightforward way to be valid for vector- and matrix-valued functions.

For $\boldsymbol{q} \in [H^1(\Omega)]^{2\times 2}$ let Π_K be the local L^2-orthogonal projection operator onto $P_0(K)$ (in fact on each element $K \in \mathscr{T}_h$, and $\boldsymbol{q} \in Q$, $\Pi_K\boldsymbol{q}$ is the average value of $\boldsymbol{q}$ on K). We then have

$$\|\Pi_K\boldsymbol{q} - \boldsymbol{q}\|_{L^2(K)} \leqslant ch_K|\boldsymbol{q}|_{H^1(K)}. \tag{71}$$

Next, we have the following interpolation result.

Lemma 4.1. *Let $\boldsymbol{w}(t) \in Z$. Then there exists $\boldsymbol{w}_I(t) \in Z_h$ with $\dot{\boldsymbol{w}}_I(t) \in Z_h$ and*

$$\begin{aligned} &\|\boldsymbol{w}_I(t) - \boldsymbol{w}(t)\|_h \leqslant ch, \\ &\|\dot{\boldsymbol{w}}_I(t) - \dot{\boldsymbol{w}}(t)\|_h \leqslant ch. \end{aligned} \tag{72}$$

Remark. We note that, for $\boldsymbol{w}_I(t) \in Z_h$, $\dot{\boldsymbol{w}}_I(t) \in Z_h$ from the definition of the time derivative and the fact that Z_h is a closed subspace.

Proof. This is done in two steps: first, we construct $\boldsymbol{w}_I(t)$, and secondly we derive the error estimates (72).

For $\boldsymbol{w}(t) = (\boldsymbol{u}(t), \boldsymbol{p}(t), \gamma(t)) \in Z$ we define $\boldsymbol{w}_I(t) \in Z_h$ by

$$a_h(\boldsymbol{w}_I(t) - \boldsymbol{w}(t), z_h) = 0, \quad z_h \in Z_h. \tag{73}$$

Since the bilinear form $a_h(\cdot,\cdot)$ is continuous and coercive on Z_h, $\boldsymbol{w}_I(t)$ is well defined.

Let $\Pi\boldsymbol{w}(t) \in Z_h$ be defined by

$$\Pi\boldsymbol{w}(t)|_K = (\Pi_K\boldsymbol{u}(t), \Pi_K\boldsymbol{p}(t), \Pi_K\gamma(t));$$

setting $z_h = \boldsymbol{w}_I(t) - \Pi\boldsymbol{w}(t)$ in (73) we obtain

$$c\|\boldsymbol{w}_I(t) - \rho_h\boldsymbol{w}(t)\|_{h1}^2 \leqslant a_h(\boldsymbol{w}_I(t) - \rho_h\boldsymbol{w}(t), \boldsymbol{w}_I(t) - \rho_h\boldsymbol{w}(t)) = R \tag{74}$$

with

$$
\begin{aligned}
R &:= a_h(\boldsymbol{w}(t) - \rho_h \boldsymbol{w}(t), \boldsymbol{w}_I(t) - \rho_h \boldsymbol{w}(t)) \\
&= \left.\begin{aligned}
&\sum_K \int_K \{\mathscr{C}[\boldsymbol{\epsilon}(\boldsymbol{u}(t) - \Pi_K \boldsymbol{u}(t)) - (\boldsymbol{p}(t) - \Pi_K \boldsymbol{p}(t))] \\
&\quad : [\boldsymbol{\epsilon}(\boldsymbol{u}_I(t) - \Pi_K \boldsymbol{u}(t)) - (\boldsymbol{p}_I(t) - \Pi_K \boldsymbol{p}(t))] \\
&\quad + k_1(\boldsymbol{p}(t) - \Pi_K \boldsymbol{p}(t)) : (\boldsymbol{p}_I(t) - \Pi_K \boldsymbol{p}(t)) \\
&\quad + k_2(\gamma(t) - \Pi_K \gamma(t))(\gamma_I(t) - \Pi_K \gamma(t)) \\
&\quad + k_3 \nabla(\gamma(t) - \Pi_K \gamma(t)) \cdot \nabla(\gamma_I(t) - \Pi_K \gamma(t))\} \,\mathrm{d}x
\end{aligned}\right\} \quad (Q_5) \\
&\quad \left.\begin{aligned}
&- \sum_e \int_e \mathscr{C}[\boldsymbol{\epsilon}(\boldsymbol{u}(t) - \Pi_K \boldsymbol{u}(t)) - (\boldsymbol{p}(t) - \Pi_K \boldsymbol{p}(t))] : [\![\boldsymbol{u}_I(t) - \Pi_K \boldsymbol{u}(t)]\!] \,\mathrm{d}s \\
&- \sum_e \int_e k_3 \{\nabla(\gamma(t) - \Pi_K \gamma(t))\} \cdot [\![\gamma_I(t) - \Pi_K \gamma(t)]\!] \,\mathrm{d}s
\end{aligned}\right\} \quad (Q_6) \\
&\quad \left.\begin{aligned}
&- \sum_e \int_e \mathscr{C}[\boldsymbol{\epsilon}(\boldsymbol{u}_I(t) - \Pi_K \boldsymbol{u}(t)) - (\boldsymbol{p}_I(t) - \Pi_K \boldsymbol{p}(t))] : [\![\boldsymbol{u}(t) - \Pi_K \boldsymbol{u}(t)]\!] \,\mathrm{d}s \\
&- \sum_e \int_e k_3 \{\nabla(\gamma_I(t) - \Pi_K \gamma(t))\} \cdot [\![\gamma(t) - \Pi_K \gamma(t)]\!] \,\mathrm{d}s
\end{aligned}\right\} \quad (Q_7) \\
&\quad \left.\begin{aligned}
&+ \sum_e \int_e \frac{\beta_1}{h_e} [\![\boldsymbol{u}(t) - \Pi_K \boldsymbol{u}(t)]\!] : [\![\boldsymbol{u}_I(t) - \Pi_K \boldsymbol{u}(t)]\!] \,\mathrm{d}s \\
&+ \frac{\beta_2}{h_e} [\![\gamma(t) - \Pi_K \gamma(t)]\!] \cdot [\![\gamma_I(t) - \Pi_K \gamma(t)]\!] \,\mathrm{d}s
\end{aligned}\right\} \quad (Q_8).
\end{aligned}
$$

We now estimate $Q_5 - Q_8$ using the Cauchy–Schwarz, Minkowski and Young inequalities, (70) and (71). First, Q_6 is treated thanks to $(36)_2$, to obtain

$$
\begin{aligned}
|Q_6| &\leqslant \left(\sum_e h_e \|\{\mathscr{C}\boldsymbol{\epsilon}(\boldsymbol{u}(t) - \Pi_K \boldsymbol{u}(t)) - (\boldsymbol{p}(t) - \Pi_K \boldsymbol{p}(t))\}\|_e^2\right)^{1/2} \times \left(\sum_e h_e^{-1} \|[\![\boldsymbol{u}_I(t) - \Pi_K \boldsymbol{u}(t)]\!]\|_e^2\right)^{1/2} \\
&\quad + k_3 \left(\sum_e h_e \|\{\nabla(\gamma(t) - \Pi_K \gamma(t))\}\|_e^2\right)^{1/2} \left(\sum_e h_e^{-1} \|[\![\gamma_I(t) - \Pi_K \gamma(t)]\!]\|_e^2\right)^{1/2} \\
&\leqslant \frac{1}{2\varepsilon} \sum_e h_e \|\{\mathscr{C}(\boldsymbol{\epsilon}(\boldsymbol{u}(t) - \Pi_K \boldsymbol{u}(t)) - (\boldsymbol{p}(t) - \Pi_K \boldsymbol{p}(t)))\}\|_e^2 + k_3 h_K^2 \|\nabla^2(\gamma(t) - \Pi_K \gamma(t))\|_K^2 + \varepsilon \|\boldsymbol{w}_I(t) - \Pi \boldsymbol{w}(t)\|_h^2 \\
&\leqslant c \sum_K \left(h_K^2 + h_K^2 \|\mathrm{div}(\boldsymbol{\sigma}(t) - \mathscr{C}(\boldsymbol{\epsilon}(\Pi_K \boldsymbol{u}(t)) - \Pi_K \boldsymbol{p}(t)))\|_K^2\right) + k_3 \sum_K h_K^2 \|\nabla^2(\gamma(t) - \Pi_K \gamma(t))\|_K^2 + \varepsilon \|\boldsymbol{w}_I(t) - \Pi \boldsymbol{w}(t)\|_h^2 \\
&\leqslant ch^2 + \varepsilon \|\boldsymbol{w}_I(t) - \Pi \boldsymbol{w}(t)\|_h^2.
\end{aligned} \tag{75}
$$

Next,

$$
\begin{aligned}
|Q_5| &\leqslant c \left(\sum_K |\boldsymbol{u}(t) - \Pi_K \boldsymbol{u}(t)|_{1,K}^2 + \|\boldsymbol{p}(t) - \Pi_K \boldsymbol{p}(t)\|_K^2 + \|\gamma(t) - \Pi_K \gamma(t)\|_{1,K}^2\right)^{1/2} \\
&\quad \times \left(\sum_K \|\boldsymbol{\epsilon}(\boldsymbol{u}_I(t) - \Pi_K \boldsymbol{u}(t))\|_K^2 + \|\boldsymbol{p}_I(t) - \Pi_K \boldsymbol{p}(t)\|_K^2 + \|\gamma_I(t) - \Pi_K \gamma(t)\|_{1,K}^2\right)^{1/2} \\
&\leqslant ch^2 + \frac{\varepsilon}{2} \|\boldsymbol{w}_I(t) - \Pi \boldsymbol{w}(t)\|_h^2.
\end{aligned} \tag{76}
$$

The term Q_7 is treated by using $(36)_1$, to get

$$
\begin{aligned}
|Q_7| &\leqslant c \left(\sum_K \|\mathscr{C}(\boldsymbol{\epsilon}(\boldsymbol{u}_I(t) - \Pi_K \boldsymbol{u}(t)) - (\boldsymbol{p}_I(t) - \Pi_K \boldsymbol{p}(t)))\|_K^2\right)^{1/2} \times \left(\sum_K \|\nabla(\boldsymbol{u}(t) - \Pi_K \boldsymbol{u}(t))\|_K^2\right)^{1/2} \\
&\quad + ck_3 \left(\sum_e \|\nabla(\gamma_I(t) - \Pi_K \gamma(t))\|_e^2\right)^{1/2} \left(\sum_K \|\nabla(\gamma(t) - \Pi_K \gamma(t))\|_K^2\right)^{1/2} \\
&\leqslant ch^2 + \varepsilon \|\boldsymbol{w}_I(t) - \Pi \boldsymbol{w}(t)\|_h^2.
\end{aligned} \tag{77}
$$

Finally,

$$|Q_8| \leqslant \frac{\beta_1^2+\beta_2^2}{2\varepsilon}\sum_e h_e^{-1}\|[\![\boldsymbol{u}(t)-\Pi_K\boldsymbol{u}(t)]\!]\|_e^2 + h_e^{-1}\|[\![\gamma(t)-\Pi_K\gamma(t)]\!]\|_e^2 + \varepsilon\|\boldsymbol{w}_I(t)-\Pi\boldsymbol{w}(t)\|_h^2$$
$$\leqslant c(\beta_1^2+\beta_2^2)\|\nabla(\boldsymbol{u}(t)-\Pi_K\boldsymbol{u}(t))\|_0^2 + \|\nabla(\gamma(t)-\Pi_K\gamma(t))\|_0^2 + \varepsilon\|\boldsymbol{w}_I(t)-\Pi\boldsymbol{w}(t)\|_h^2 \leqslant ch^2 + \varepsilon\|\boldsymbol{w}_I(t)-\Pi\boldsymbol{w}(t)\|_h^2. \quad (78)$$

Replacing (77) and (78) in (74) and taking ε sufficiently small, we obtain

$$\|\boldsymbol{w}_I(t)-\Pi\boldsymbol{w}(t)\|_h \leqslant ch. \quad (79)$$

The first of (72) is obtained using the triangle inequality, (79), and the interpolation estimate

$$\|\Pi\boldsymbol{w}(t)-\boldsymbol{w}(t)\|_h \leqslant ch. \quad (80)$$

The second estimate is obtained by first differentiating (73) with respect to time to get

$$a_h(\dot{\boldsymbol{w}}_I(t)-\dot{\boldsymbol{w}}(t),\boldsymbol{z}_h)=0 \quad (81)$$

and by repeating the analysis. □

With these preliminary results in place, we then have the following.

Theorem 4.2. *Let* $\boldsymbol{w}_I(t)\in Z_h$ *defined by* (73), *and let* $\boldsymbol{w}_h(t)$ *be the solution of* (43). *Assume that* $\boldsymbol{f}(t)\in H^1(\Omega)^2$, *and that* $\boldsymbol{w}=(\boldsymbol{u},\boldsymbol{p},\gamma)$ *is the solution of* (25). *Then there exists a positive constant* c, *independent of* h, *such that*

$$\|\boldsymbol{w}-\boldsymbol{w}_h\|_{L^\infty(0,T;Z_h+Z)} \leqslant \|\boldsymbol{w}-\boldsymbol{w}_I\|_{L^\infty(0,T;Z_h+Z)} + c\|\dot{\boldsymbol{w}}_I-\dot{\boldsymbol{w}}\|_{L^2(0,T;Z_h+Z)}$$
$$+c\left(\inf_{\boldsymbol{z}_h\in L^2(0,T;Z_h)}\|\boldsymbol{z}_h-\dot{\boldsymbol{w}}\|_{L^2(0,T;Z_h+Z)}^{1/2} + \inf_{\boldsymbol{v}_h\in L^2(0,T;V_h)}\|\boldsymbol{v}_h-\dot{\boldsymbol{u}}\|_{L^2(0,T;V_h+V)}^{1/2} + \inf_{\boldsymbol{q}_h\in L^2(0,T;Q_h)}\|\boldsymbol{q}_h-\dot{\boldsymbol{p}}\|_{L^2(0,T;Q)}^{1/2}\right). \quad (82)$$

Proof. Taking $\boldsymbol{z}_h=\dot{\boldsymbol{w}}_h(t)$ in $(44)_2$, and adding (43), we get for all $\boldsymbol{z}_h\in Z_h$,

$$-a_h(\boldsymbol{w}_h(t),\boldsymbol{z}_h-\dot{\boldsymbol{w}}_h(t)) \leqslant a_h(\boldsymbol{w}(t),\dot{\boldsymbol{w}}_h(t)-\dot{\boldsymbol{w}}(t)) + j(\boldsymbol{z}_h) - j(\dot{\boldsymbol{w}}(t)) - \langle \ell(t),\boldsymbol{z}_h-\dot{\boldsymbol{w}}(t)\rangle. \quad (83)$$

Now a_h is continuous and coercive on Z_h, so that $\|\boldsymbol{z}_h\|_{a_h}^2=a_h(\boldsymbol{z}_h,\boldsymbol{z}_h)$ is a norm on Z_h, equivalent to $\|\boldsymbol{z}_h\|_h$. Using (83) and (73) we obtain, for $\boldsymbol{z}_h\in Z_h$,

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{w}_I(t)-\boldsymbol{w}_h(t)\|_{a_h}^2 = a_h(\boldsymbol{w}_I(t)-\boldsymbol{w}_h(t),\dot{\boldsymbol{w}}_I(t)-\boldsymbol{z}_h) + a_h(\boldsymbol{w}_I(t),\boldsymbol{z}_h-\dot{\boldsymbol{w}}_h(t)) - a_h(\boldsymbol{w}_h(t),\boldsymbol{z}_h-\dot{\boldsymbol{w}}_h(t))$$
$$\leqslant a_h(\boldsymbol{w}_I(t)-\boldsymbol{w}_h(t),\dot{\boldsymbol{w}}_I(t)-\dot{\boldsymbol{w}}(t)) + a_h(\boldsymbol{w}_I(t)-\boldsymbol{w}_h(t),\dot{\boldsymbol{w}}(t)-\boldsymbol{z}_h) + a_h(\boldsymbol{w}(t),\boldsymbol{z}_h-\dot{\boldsymbol{w}}(t))$$
$$+ j(\boldsymbol{z}_h) - j(\dot{\boldsymbol{w}}(t)) - \langle \ell(t),\boldsymbol{z}_h-\dot{\boldsymbol{w}}(t)\rangle. \quad (84)$$

We now have to estimate each term on the right hand side of (84).

Using Lemma 4.1 we have

$$a_h(\boldsymbol{w}_I(t)-\boldsymbol{w}_h(t),\dot{\boldsymbol{w}}_I(t)-\dot{\boldsymbol{w}}(t)) \leqslant c\Big(\|\boldsymbol{w}_I(t)-\boldsymbol{w}_h(t)\|_{a_h}^2 + \|\dot{\boldsymbol{w}}_I(t)-\dot{\boldsymbol{w}}(t)\|_h^2\Big),$$
$$a_h(\boldsymbol{w}_I(t)-\boldsymbol{w}_h(t),\dot{\boldsymbol{w}}(t)-\boldsymbol{z}_h) \leqslant c\Big(\|\boldsymbol{w}_I(t)-\boldsymbol{w}_h(t)\|_{a_h}^2 + c\|\dot{\boldsymbol{w}}(t)-\boldsymbol{z}_h\|_h^2\Big). \quad (85)$$

By definition, and using the regularity of $\boldsymbol{w} = (\boldsymbol{u}, \boldsymbol{p}, \gamma)$, we have for $\boldsymbol{z}_h = (\boldsymbol{v}_h, \boldsymbol{q}_h, \eta_h)$,

$$\begin{aligned} a_h(\boldsymbol{w}(t), \boldsymbol{z}_h - \dot{\boldsymbol{w}}(t)) &= \sum_K \int_K (\mathscr{C}(\boldsymbol{\epsilon}(\boldsymbol{u}(t)) - \boldsymbol{p}(t)) : (\boldsymbol{\epsilon}(\boldsymbol{v}_h(t) - \dot{\boldsymbol{u}}(t)) - (\boldsymbol{q}_h - \dot{\boldsymbol{p}}(t)))\,\mathrm{d}x \\ &\quad + \sum_K \int_K [k_1 \boldsymbol{p}(t) : (\boldsymbol{q}_h(t) - \dot{\boldsymbol{p}}(t)) + k_2 \gamma(t)(\eta_h(t) - \dot{\gamma}(t))]\,\mathrm{d}x + k_3 \sum_K \int_K \nabla\gamma(t) \cdot \nabla(\eta_h(t) - \dot{\gamma}(t)))\,\mathrm{d}x \\ &\quad - \sum_e \int_e (\boldsymbol{\sigma}(t) : [\![\boldsymbol{v}_h(t) - \dot{\boldsymbol{u}}(t)]\!] + k_3 \nabla\gamma(t) \cdot [\![\eta_h - \dot{\gamma}(t)]\!])\,\mathrm{d}s \\ &\leqslant \sum_K [|\mathscr{C}|_\infty \|\boldsymbol{\epsilon}(\boldsymbol{u}(t)) - \boldsymbol{p}(t)\|_K (\|\boldsymbol{\epsilon}(\boldsymbol{v}_h(t) - \dot{\boldsymbol{u}}(t))\|_K + \|\boldsymbol{q}_h - \dot{\boldsymbol{p}}(t)\|_K) + ck_1 \|\boldsymbol{q}_h(t) - \dot{\boldsymbol{p}}(t)\|_K + ck_2 \|\eta_h(t) \\ &\quad - \dot{\gamma}(t)\|_K + ck_3 \|\nabla(\eta_h(t) - \dot{\gamma}(t))\|_K] + \left(\sum_e h_e \|\boldsymbol{\sigma}(t)\|_e^2\right)^{1/2} \left(\sum_e h_e^{-1} \|[\![\boldsymbol{v}_h(t) - \dot{\boldsymbol{u}}(t)]\!]\|_e^2\right)^{1/2} \\ &\quad + k_3 \left(\sum_e h_e \|\nabla\gamma(t)\|_e^2\right)^{1/2} \left(\sum_e h_e^{-1} \|[\![\eta_h - \dot{\gamma}(t)]\!]\|_e^2\right)^{1/2} \leqslant c(1 + h^{1/2}) \|\boldsymbol{z}_h - \dot{\boldsymbol{w}}(t)\|_h. \end{aligned} \tag{86}$$

This leaves the terms in (84) involving j and ℓ. Bearing in mind the definition (19) of j we have, for any $\hat{\boldsymbol{z}}_h = (\boldsymbol{v}_h, \boldsymbol{q}_h, |\boldsymbol{q}_h|) \in W_h$,

$$\begin{aligned} j(\hat{\boldsymbol{z}}_h) - j(\dot{\boldsymbol{w}}(t)) - \langle \ell(t), \hat{\boldsymbol{z}}_h - \dot{\boldsymbol{w}}(t) \rangle &\leqslant \sum_K \int_K |\boldsymbol{q}_h(t) - \dot{\boldsymbol{p}}(t)| + \boldsymbol{f}(t) \cdot (\boldsymbol{v}_h(t) - \dot{\boldsymbol{u}}(t)) \\ &\leqslant c\|\boldsymbol{q}_h(t) - \dot{\boldsymbol{p}}(t)\|_0 + c\|\boldsymbol{v}_h(t) - \dot{\boldsymbol{u}}(t)\|_0. \end{aligned} \tag{87}$$

Combining (84)–(87), we obtain, for $\boldsymbol{z}_h = (\boldsymbol{v}_h, \boldsymbol{q}_h, \eta_h) \in Z_h$,

$$\frac{\mathrm{d}}{\mathrm{d}t} \|\boldsymbol{w}_I(t) - \boldsymbol{w}_h(t)\|_{a_h}^2 \leqslant c(\|\boldsymbol{w}_I(t) - \boldsymbol{w}_h(t)\|_{a_h}^2 + \|\dot{\boldsymbol{w}}_I(t) - \dot{\boldsymbol{w}}(t)\|_h^2 + \|\boldsymbol{z}_h - \dot{\boldsymbol{w}}(t)\|_h + \|\boldsymbol{q}_h - \dot{\boldsymbol{p}}(t)\|_0 + \|\boldsymbol{v}_h - \dot{\boldsymbol{u}}(t)\|_0). \tag{88}$$

Applying Gronwall's Lemma with $\boldsymbol{w}_h(\boldsymbol{0}) = \boldsymbol{0}$, the Sobolev embedding theorem, and the equivalence between $\|\cdot\|_{a_h}$ and $\|\cdot\|_h$, we obtain

$$\|\boldsymbol{w}_I(t) - \boldsymbol{w}_h(t)\|_h \leqslant c(\|\dot{\boldsymbol{w}}_I - \dot{\boldsymbol{w}}\|_{L^2(0,T;Z_h+Z)} + \|\boldsymbol{z}_h - \dot{\boldsymbol{w}}\|_{L^2(0,T;Z_h+Z)}^{1/2} + \|\boldsymbol{q}_h - \dot{\boldsymbol{p}}(t)\|_{L^2(0,T;Q)} + \|\boldsymbol{v}_h - \dot{\boldsymbol{u}}(t)\|_{L^2(0,T;V_h+V)}). \tag{89}$$

The theorem follows after application of the triangle inequality. □

Remark. Choosing $\boldsymbol{z}_h$ such that $\boldsymbol{z}_h|_K = \Pi_K \dot{\boldsymbol{w}}_I(t)$, we obtain

$$\|\boldsymbol{w} - \boldsymbol{w}_h\|_{L^\infty(0,T;Z_h+Z)} \leqslant ch^{1/2}.$$

This is the same rate of convergence as that obtained for classical plasticity using the conventional Galerkin method [1].

5. Fully discrete discontinuous Galerkin approximations

In this section we discretize (25) in time using the backward Euler scheme. We denote the function ψ evaluated at time t_n by ψ^n. We first discretize the time interval $[0, T]$ into N subintervals with node points $t_n = nk$, $0 \leqslant n \leqslant N$, where $k = t_{n+1} - t_n = T/N$ is the step-size. We set $\delta \boldsymbol{w}^n = \Delta \boldsymbol{w}^n / k$ where $\Delta \boldsymbol{w}^n = \boldsymbol{w}^n - \boldsymbol{w}^{n-1}$.

The reader should consult [1] for the details of the technique used in the proof of the main result in this section.

Consider the following fully discrete DG approximation of Problem (25):given $\ell \in H^1(0, T; Z')$, and $\boldsymbol{w}_h^0 = \boldsymbol{0}$, find a sequence $(\boldsymbol{w}_h^n)_{n=1}^N$ in W_h, with $\delta w_h^n \in W_h$ satisfying

$$a_h(\boldsymbol{w}_h^n, \boldsymbol{z}_h - \delta \boldsymbol{w}_h^n) + j(\boldsymbol{z}_h) - j(\delta \boldsymbol{w}_h^n) \geqslant \langle \ell^n, \boldsymbol{z}_h - \delta \boldsymbol{w}_h^n \rangle, \quad \forall \boldsymbol{z}_h \in Z_h. \tag{90}$$

The existence of a unique solution to Problem (90) follows from the arguments presented in Lemma 3.2.

We have the following stability result, which follows directly from ([1], Lemma 7.2).

Lemma 5.1. *The solution $(\boldsymbol{w}_h^n)_{n=1}^N$ is stable in the sense there exist positive constants c_1, c_2 independent of k, such that*

$$\sum_{n=1}^N \|\delta \boldsymbol{w}_h^n\|_h^2 k \leqslant c_1 \|\dot{\ell}\|_{L^2(0,T;Z')}^2 \quad \text{and} \quad \max_{1 \leqslant n \leqslant N} \|\boldsymbol{w}_h^n\|_h \leqslant c_2 \|\dot{\ell}\|_{L^1(0,T;Z')}. \tag{91}$$

The following result concerning the time derivative of the interpolant $\boldsymbol{w}_I(t)$ of $\boldsymbol{w}(t)$ will be used in the proof of Theorem 5.3.

Lemma 5.2. *If $\ddot{\boldsymbol{w}} \in L^1(0,T;Z)$, and $\dot{\boldsymbol{w}}_I(t)$ is the interpolant of $\dot{\boldsymbol{w}}(t)$, then there exists a positive constant c, independent of h and k, such that*

$$\left\|\frac{\boldsymbol{w}_I(t_{m+1}) - \boldsymbol{w}_I(t_m)}{k} - \dot{\boldsymbol{w}}_I(t_m)\right\|_h \leqslant c\|\ddot{\boldsymbol{w}}\|_{L^1(t_m,t_{m+1};Z+Z_h)}. \tag{92}$$

The proof is obtained by using (73) to show that $\ddot{\boldsymbol{w}}_I \in L^1(0,T;Z_h)$, followed by a Taylor expansion of $\boldsymbol{w}_I$ about t_m with integral remainder (see Lemma 11.4 in [1]).

Next we state the main result regarding the accuracy of the solution $\boldsymbol{w}_h^n$ of problem (90).

Theorem 5.3. *Let $(\boldsymbol{w}_h^n)_{n=1}^N$ be a sequence of solutions of (90). If $\boldsymbol{f}(t) \in L^2(\Omega)^2$ and the solution $\boldsymbol{w} = (\boldsymbol{u}, \boldsymbol{p}, \gamma)$ of (25) satisfies $\ddot{\boldsymbol{w}} \in L^1(0,T;Z)$, then there exists a positive constant c, independent of the mesh size h and k, such that*

$$\max_{1\leqslant n\leqslant N} \|\boldsymbol{w}^n - \boldsymbol{w}_h^n\|_h \leqslant ck + c\left[k\sum_{i=1}^N \inf_{z_h\in Z_h} \|\dot{\boldsymbol{w}}^i - \boldsymbol{z}_h\|_h\right]^{\frac{1}{2}}. \tag{93}$$

Proof. We follow closely the proof of Theorem 11.7 in [1]. Let $\boldsymbol{e}^n = \boldsymbol{w}^n - \boldsymbol{w}_h^n = \boldsymbol{\eta}^n + \boldsymbol{e}_h^n$, for $1 \leqslant n \leqslant N$ where $\boldsymbol{\eta}^n = \boldsymbol{w}^n - \boldsymbol{w}_I^n$, $\boldsymbol{e}_h^n = \boldsymbol{w}_I^n - \boldsymbol{w}_h^n$, and $\boldsymbol{w}_I^n = \boldsymbol{w}_I(t_n)$. Set

$$A_n = a_h(\boldsymbol{e}_h^n, \delta\boldsymbol{e}_h^n) = \frac{1}{k}a_h(\boldsymbol{e}_h^n, \boldsymbol{e}_h^n) - \frac{1}{k}a_h(\boldsymbol{e}_h^n, \boldsymbol{e}_h^{n-1}). \tag{94}$$

Using the Cauchy–Schwarz and Young's inequalities we obtain

$$A_n \geqslant \frac{1}{2k}(\|\boldsymbol{e}_h^n\|_{a_h}^2 - \|\boldsymbol{e}_h^{n-1}\|_{a_h}^2). \tag{95}$$

Next, we have to find an upper bound for A_n. From the linearity of $a_h(\cdot,\cdot)$,

$$A_n = a_h(\boldsymbol{w}_I^n, \delta\boldsymbol{w}_I^n - \delta\boldsymbol{w}_h^n) - a_h(\boldsymbol{w}_h^n, \delta\boldsymbol{w}_I^n - \boldsymbol{z}_h) - a_h(\boldsymbol{w}_h^n, \boldsymbol{z}_h - \delta\boldsymbol{w}_h^n). \tag{96}$$

Now combining (90), and $(44)_2$ with $t = t_n$ and $\boldsymbol{z}_h = \delta\boldsymbol{w}_h^n$, we obtain

$$-a_h(\boldsymbol{w}_h^n, \boldsymbol{z}_h - \delta\boldsymbol{w}_h^n) \leqslant j(\boldsymbol{z}_h) - j(\dot{\boldsymbol{w}}^n) - \langle \ell^n, \boldsymbol{z}_h - \dot{\boldsymbol{w}}^n\rangle + a_h(\boldsymbol{w}^n, \delta\boldsymbol{w}_h^n - \dot{\boldsymbol{w}}^n). \tag{97}$$

Combining (97), (96), (73), and using $\delta\boldsymbol{e}_h^n = \delta\boldsymbol{w}_I^n - \delta\boldsymbol{w}_h^n = (\boldsymbol{e}_h^n - \boldsymbol{e}_h^{n-1})/k$, we get

$$\begin{aligned} A_n &\leqslant a_h(\boldsymbol{w}_I^n - \boldsymbol{w}^n, \delta\boldsymbol{w}_I^n - \delta\boldsymbol{w}_h^n) + a_h(\boldsymbol{w}^n, \delta\boldsymbol{w}_I^n - \delta\boldsymbol{w}_h^n) - a_h(\boldsymbol{w}_h^n, \delta\boldsymbol{w}_I^n - \boldsymbol{z}_h) + a_h(\boldsymbol{w}^n, \delta\boldsymbol{w}_h^n - \dot{\boldsymbol{w}}^n) + j(\boldsymbol{z}_h) - j(\dot{\boldsymbol{w}}^n) - \langle \ell^n, \boldsymbol{z}_h - \dot{\boldsymbol{w}}^n\rangle \\ &\leqslant a_h(\boldsymbol{w}^n, \delta\boldsymbol{w}_I^n - \delta\boldsymbol{w}_h^n) - a_h(\boldsymbol{w}_h^n, \delta\boldsymbol{w}_I^n - \boldsymbol{z}_h) + a_h(\boldsymbol{w}^n, \delta\boldsymbol{w}_h^n - \dot{\boldsymbol{w}}^n) + j(\boldsymbol{z}_h) - j(\dot{\boldsymbol{w}}^n) - \langle \ell^n, \boldsymbol{z}_h - \dot{\boldsymbol{w}}^n\rangle \\ &\leqslant \underbrace{a_h(\boldsymbol{e}_h^n, \delta\boldsymbol{w}_I^n - \boldsymbol{z}_h) + a_h(\boldsymbol{w}^n, \boldsymbol{z}_h - \dot{\boldsymbol{w}}^n)}_{Q_9} + \underbrace{j(\boldsymbol{z}_h) - j(\dot{\boldsymbol{w}}^n) - \langle \ell^n, \boldsymbol{z}_h - \dot{\boldsymbol{w}}^n\rangle}_{Q_{10}}. \end{aligned} \tag{98}$$

From (95) and (98) we obtain, choosing $\boldsymbol{z}_h = (\boldsymbol{v}_h, \boldsymbol{q}_h, \eta_h) \in Z_h$ in the term Q_9 and $\boldsymbol{z}_h = \hat{\boldsymbol{z}}_h := (\boldsymbol{v}_h, \boldsymbol{q}_h, |\boldsymbol{q}_h|)$ in Q_{10},

$$\frac{1}{2k}\left(\|\boldsymbol{e}_h^n\|_{a_h}^2 - \|\boldsymbol{e}_h^{n-1}\|_{a_h}^2\right) \leqslant Q_9 + Q_{10} \leqslant c(\|\boldsymbol{e}_h^n\|_{a_h}\|\delta\boldsymbol{w}_I^n - \boldsymbol{z}_h\|_{a_h} + \|\boldsymbol{w}^n\|_{a_h}\|\boldsymbol{z}_h^n - \dot{\boldsymbol{w}}^n\|_{a_h} + \|\boldsymbol{q}_h - \dot{\boldsymbol{p}}^n\|_0 + \|\boldsymbol{v}_h - \dot{\boldsymbol{u}}^n\|_0). \tag{99}$$

Now replacing n with i, summing over i with $1 \leqslant i \leqslant n$ and with $\boldsymbol{e}_h^0 = \boldsymbol{0}$, we obtain

$$\|\boldsymbol{e}_h^n\|_{a_h}^2 \leqslant ckM\sum_{i=1}^n \|\delta\boldsymbol{w}_I^i - \boldsymbol{z}_h\|_{a_h} + ck\sum_{i=1}^n\left(\|\boldsymbol{z}_h - \dot{\boldsymbol{w}}^i\|_{a_h} + \|\boldsymbol{q}_h^i - \dot{\boldsymbol{p}}^i\|_0 + \|\boldsymbol{v}_h^i - \dot{\boldsymbol{u}}^i\|_0\right),$$

where $M = \max_i \boldsymbol{e}_h^i\|_{a_h}$ and we have also used the property $\max_n\|\boldsymbol{w}^n\|_Z \leqslant c\|\dot{\ell}\|_{Z'}$, which follows from ([1], Lemma 7.2). Finally,

$$M^2 \leqslant ckM\sum_{i=1}^N \|\delta\boldsymbol{w}_I^i - \boldsymbol{z}_h\|_{a_h} + ck\sum_{i=1}^N \|\boldsymbol{z}_h - \dot{\boldsymbol{w}}^i\|_{a_h}. \tag{100}$$

Noting that $a, b, x \geqslant 0$ and $x^2 \leqslant ax + b$ imply that $x \leqslant a + \sqrt{2b}$, we obtain

$$M \leqslant ck\sum_{i=1}^N \|\delta\boldsymbol{w}_I^i - \boldsymbol{z}_h\|_{a_h} + ck^{1/2}\left(\sum_{i=1}^N \|\boldsymbol{z}_h - \dot{\boldsymbol{w}}^i\|_{a_h}\right)^{1/2}. \tag{101}$$

The result follows by choosing $\boldsymbol{z}_h = \dot{\boldsymbol{w}}_I(t)$ in (101) and using Lemma 5.2. □

Remarks

1. Again, as with the case of semi-discrete approximations, by using the interpolation error estimates (70) and (71) and the regularity of the solution, it is seen that the order of spatial convergence is the same as that obtained for the classical problem with the conventional Galerkin method.
2. It is possible, using the approach in ([1], Theorem 11.6), to obtain $O(k^2)$ convergence with the use of a Crank–Nicolson as opposed to a backward Euler approximation in time.

6. Conclusion

A discontinuous Galerkin formulation has been constructed and analyzed for problems involving both classical and gradient plasticity. Error estimates have been derived for semi- and fully discrete formulations. There remains the issue of computation, and in this context a key question concerns the extension of the well-known predictor–corrector algorithms [2,1] to the problems considered here. This will be the subject of a subsequent work [22] in which a detailed analysis will be presented of the algorithms, and in which various numerical examples will serve to illustrate the theoretical results presented in this work.

Acknowledgement

The authors thank the National Research Foundation for support that has made this work possible.

References

[1] W. Han, B.D. Reddy, Plasticity: Mathematical Theory and Numerical Analysis, Springer, 1999.
[2] J.C. Simo, T.J.R. Hughes, Computational Inelasticity, Springer-Verlag, 1998.
[3] E.C. Aifantis, On the microstructural origin of certain inelastic models, J. Engrg. Mater. Technol. 106 (1984) 326–330.
[4] E.C. Aifantis, The physics of plastic deformation, Int. J. Plasticity 3 (1987) 211–247.
[5] H.-B. Mühlhaus, E. Aifantis, A variational principle for gradient plasticity, Int. J. Solids Struct. 28 (7) (1991) 845–857.
[6] R. de Borst, H.-B. Mühlhaus, Gradient-dependent plasticity: formulation and algorithmic aspects, Int. J. Numer. Methods Engrg. 35 (1992) 521–539.
[7] N.A. Fleck, J.W. Hutchinson, Strain gradient plasticity, Adv. Appl. Mech. 33 (1997) 295–361.
[8] A. Needleman, Computational mechanics at the mesoscale, Acta Mater. 48 (2000) 105–124.
[9] M.E. Gurtin, A gradient theory of small-deformation isotropic plasticity that accounts for the burgers vector and for dissipation due to plastic spin, J. Mech. Phys. Solids 52 (2004) 2545–2568.
[10] P. Gudmundson, A unified treatment of strain gradient plasticity, J. Mech. Phys. Solids 52 (2004) 1379–1406.
[11] T. Svedberg, K. Runesson, An algorithm for gradient-regularized plasticity coupled to damage based on a dual mixed fe-formulation, Comput. Methods Appl. Mech. Engrg. 161 (1998) 49–65.
[12] T. Liebe, P. Steinmann, Theory and numerics of a thermodynamically consistent framework for geometrically linear gradient plasticity, Int. J. Numer. Methods Eng. 51 (2001) 1437–1467.
[13] T. Liebe, A. Menzel, P. Steinmann, Theory and numerics of geometrically non-linear gradient plasticity, Int. J. Engrg. Sci. 41 (2003) 1603–1629.
[14] J.A. Nitsche, Über ein Variationspringzip zur Lösung von Dirichlet-problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind, Abhandlungen der mathematischer Seminar der Universität Hamburg 36 (1971) 9–15.
[15] D.N. Arnold, An interior penalty finite element method with discontinuous elements, SIAM J. Numer. Anal. 19 (1982) 742–760.
[16] B. Cockburn, G. Karniadakis, C.-W. Shu, Discontinuous Galerkin Methods: Theory, Computation and Applications, Springer-Verlag, 2000.
[17] B. Rivière, M.F. Wheeler, Optimal error estimates for discontinuous Galerkin methods applied to linear elasticity problems, Technical Report 00-30, Texas Institute for Computational and Applied Mathematics.
[18] M.F. Wheeler, An elliptic collocation-finite element method with interior penalty, SIAM J. Numer. Anal. 15 (1978) 152–161.
[19] G.N. Wells, K. Garikipati, L. Molari, A discontinuous Galerkin formulation for a strain gradient-dependent damage model, Comput. Methods Appl. Mech. Engrg. 193 (2004) 3633–3645.
[20] L. Molari, G.N. Wells, K. Garikipati, F. Ubertini, A discontinuous Galerkin method for strain gradient-dependent damage: study of interpolations and convergence, Comput. Methods Appl. Mech. Engrg. 195 (2006) 1480–1498.
[21] G. Engel, K. Garikipati, T.J.R. Hughes, M.G. Larson, L. Mazzei, R.L. Taylor, Continuous/discontinuous finite element approximations of fourth-order elliptic problems in structural and continuum mechanics with applications to thin beams and plates, and strain gradient elasticity, Comput. Methods Appl. Mech. Engrg. 191 (2002) 3669–3750.
[22] J.K. Djoko, F. Ebobisse, A.T. McBride, B.D. Reddy, A discontinuous Galerkin formulation for classical and gradient plasticity. Part 2: Solution algorithms and numerical examples, submitted for publication.
[23] M.E. Gurtin, On the plasticity of single crystals: free energy, microforces, plastic strain gradients, J. Mech. Phys. Solids 48 (2000) 989–1036.
[24] M.E. Gurtin, A gradient theory of single-crystal viscoplasticty that accounts for geometrically necessary dislocations, J. Mech. Phys. Solids 50 (2002) 5–32.
[25] S. Repin, Errors of finite element method for perfectly elastoplastic problems, Math. Mod. Meth. Appl. Sci. 6 (1996) 587–604.
[26] D.N. Arnold, F. Brezzi, B. Cockburn, L.D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, SIAM J. Numer. Anal. 39 (2002) 1749–1779.
[27] L.C. Evans, R.F. Gariepy, Measure Theory and Fine Properties of Functions, CRC Press, 1992.
[28] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, 1978.