CrossMark

Convergence of dG(1) in elastoplastic evolution

Carsten Carstensen[1] · Dongjie Liu[2] · Jochen Alberty[3]

Abstract

The discontinuous Galerkin (dG) methodology provides a hierarchy of time discretization schemes for evolutionary problems such as elastoplasticity with the Prandtl-Reuß flow rule. A dG time discretization has been proposed for a variational inequality in the context of rate-independent inelastic material behaviour in Alberty and Carstensen in (CMME 191:4949–4968, 2002) with the help of duality in convex analysis to justify certain jump terms. This paper establishes the first a priori error analysis for the dG(1) scheme with discontinuous piecewise linear polynomials in the temporal and lowest-order finite elements for the spatial discretization. Compared to a generalized mid-point rule, the dG(1) formulation distributes the action of the material law in the form of the variational inequality in time and so it introduces an error in the material law. This may result in a suboptimal convergence rate for the dG(1) scheme and this paper shows that the stress error in the $L^\infty(L^2)$ norm is merely $O(h + k^{3/2})$ based on a seemingly sharp error analysis. The numerical investigation for a benchmark problem with known analytic solution provides empirical evidence of a higher convergence rate of the dG(1) scheme compared to dG(0).

Dedicated to Alexander Mielke on the occasion of his 60th birthday.

✉ Dongjie Liu
 liudj@shu.edu.cn

 Carsten Carstensen
 cc@math.hu-berlin.de

 Jochen Alberty
 jochen@fuh-alberty.de

[1] Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

[2] Department of Mathematics, College of Sciences, Shanghai University, Shanghai 200444, People's Republic of China

[3] Diesterwegstr.4, 60594 Frankfurt, Germany

Springer

1 Introduction and overview

The discontinuous Galerkin time discretization, abbreviated as dG(k) with polynomial degree at most $k \in \mathbb{N}_0$, of elastoplastic evolution problems has been proposed, implemented, and validated in [3]. One main motivation was the usage of a hierarchy of time discretization schemes in adaptive time-step control with the view that a higher-order scheme is more accurate. The latter is supported by numerical experiments in computational mechanics and Fig. 1 displays the exact and approximate stress (in the tangential component) in the numerical benchmark of Sect. 5 as a function of time t. The Crank-Nicolson (CN) scheme is shown to be of quadratic order in time [2] under high (and possibly unrealistic) smoothness assumptions. The oscillations in Fig. 1 clearly question the premise that CN performs better than the first-order backward Euler (bE) scheme. In comparison, dG(1) appears more accurate than dG(0). Notice that bE, dG(0), and dG(1) lead to admissible generalized stress approximations pointwise in time-space, while the higher-order in CN leads to nonphysical oscillations.

A further motivation is the lower regularity required in the dG methodology, at least for a linear model problem in [20, p 209], which is desirable in elastoplasticity.
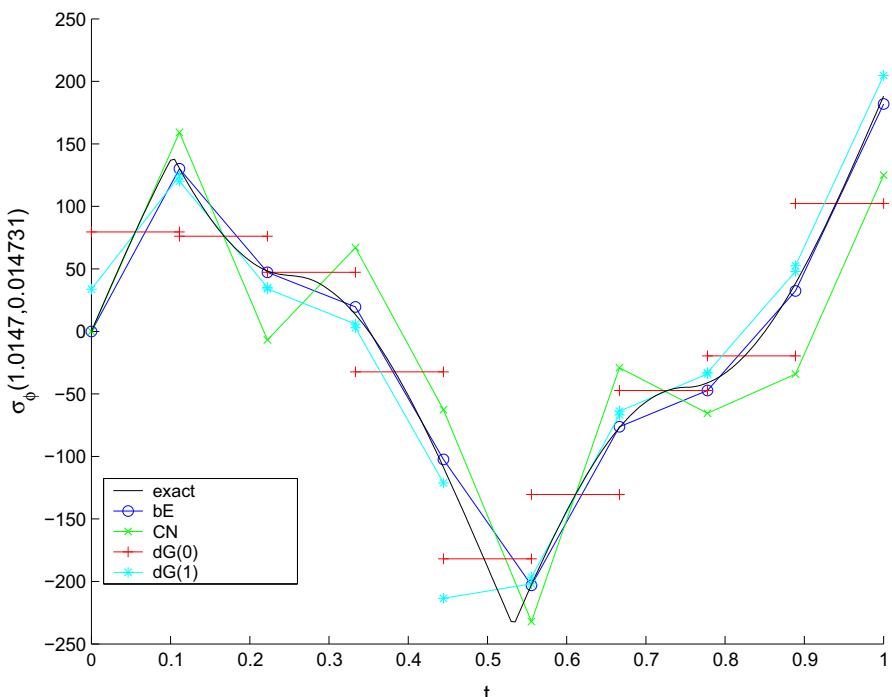
Fig. 1 Time evolution for one stress component $\sigma_\phi(1.0147, 0.0147, t)$ in the computational benchmark of Sect. 5 for exact values and for backward Euler (bE), Crank–Nicolson (CN), dG(0), and dG(1) schemes

This paper justifies the superiority of dG(1) over dG(0) for smooth solutions and presents an error analysis for the dG(1) scheme for elastoplasticity with hardening in the weak primal and dual formulation [12]. The analysis of time difference schemes (bE and CN) is straightforward in the evaluation of the time-derivatives on the exact and on the discrete level at the *same time* (in a generalized midpoint rule). In the context of variational inequalities, this leads to one inequality with discrete and another one with continuous quantities; their sum establishes a difference identity bounded by approximation error terms as outlined, e.g., in [2] for elastoplasticity with sharp error estimates. One aspect is that the Prandtl-Reuß flow is modeled exactly (at the generalized midpoint in time and on each simplex in space) and the original hardening material laws act as a regularization and control over certain error terms. In contrast with this, the discontinuous Galerkin time approximation smears out the elastoplastic evolution law in a weak formulation and strongly requires the discrete stresses to be admissible a.e. in the time-space cylinder Q.

On the continuous level, the Prandtl-Reuß flow rule links the exact generalized stress Σ and the rate $\dot{P} = \partial P / \partial t$ of the generalized plastic strain P through a variational inequality in the dual formulation of elastoplasticity (a.e. in Q)

$$\dot{P} \star (\overline{T} - \Sigma) \leq 0 \quad \text{for all admissible generalized stresses } \overline{T}. \tag{1.1}$$

The concept of admissible generalized stresses means that Σ belongs to a closed, convex, and nonvoid set K a.e. in Q and so do all test-functions, like the above $\overline{T}$.

The discrete analog approximates the above pointwise structure through time-space integrals and so it leads to an approximation error in the material law in the dG(1) scheme (dG(0) is almost bE). A textbook error analysis of dG time discretization [20, Chap 12] (of linear parabolic model problems) starts with a splitting of the error and the introduction of an intermediate quantity Σ_{dG} with the same integral mean as Σ over the time interval (t_{j-1}, t_j) and exact interpolation $\Sigma_{\mathrm{dG},j}^- = \Sigma_j$ at the right end-point of this interval. This amounts to an *extrapolation* of the values in a convex set and, in general, to a violation of the restriction $\Sigma_{\mathrm{dG}} \in K$ a.e. in Q. This paper examines the approximation of Σ by a piecewise linear polynomial $\widetilde{\Sigma}$ in time with respect to a time discretisation with time points $t_j := jk$ for all $j = 0, 1, \ldots, J$ (with $J = t_J/k \in \mathbb{N}$) and the uniform time increment k. In contrast to *extrapolation*, the *interpolation* $\widetilde{\Sigma}$ meets the exact stress Σ in all time points t_j and is globally continuous in time and admissible a.e. in Q. The space-averages of $\widetilde{\Sigma}$ are admissible in the discrete variational inequality and result in a first contribution for an estimation of the procedural error $\Theta := \widetilde{\Sigma} - \Sigma_{hk}$ for the above interpolation $\widetilde{\Sigma}$ of the true stress and the discrete stress Σ_{hk} from a full discretization with the dG(1) method in time and the standard lowest-order conforming approximation in space.

An optimal error analysis of dG time discretization in the textbook [20] derives an evolution relation for the error part Θ from the continuous problem. In linear problems, the test functions $\overline{T}$ belong to a linear space, whereas they are restricted to the convex set K in the Prandtl-Reuß flow rule at hand. The naive choice $\overline{T} = \Sigma_{hk} \in K$ is feasible, but leads to the integral $\int_{Q_j} \dot{P} \star (\Sigma_{hk} - \Sigma) dQ \leq 0$ over the time-space sub-cylinder

$Q_j := (t_{j-1}, t_j) \times \Omega$ for $j = 1, \dots, J$. In order to reformulate this material law in terms of Θ, the inequality reads

$$- \int_{Q_j} \dot{P} \star \Theta \, dQ \leq \int_{Q_j} \dot{P} \star (\Sigma - \widetilde{\Sigma}) \, dQ. \qquad (1.2)$$

The left-hand side contributes as a second ingredient for an evolution relation for Θ, while the right-hand side is a perturbation term on its own and requires an estimation from above. Since $\widetilde{\Sigma} \in K$ is admissible, the integrand $\dot{P} \star (\Sigma - \widetilde{\Sigma}) \geq 0$ is nonnegative a.e. in Q and the interpolation error estimate leads to a bound of order $O(k^2)$, which merely proves linear convergence for dG(1) as outlined below in Sect. 4.7.1.

This paper derives a convergence rate $O(h + k^{3/2})$ under the assumption of smooth exact solutions by a test with the linear interpolation $\widetilde{\Sigma}$ of Σ and the evaluation of the Prandtl-Reuß flow rule at the two time interval endpoints. This cannot circumvent the application of a trace inequality for the control of jump terms and eventually leads to an upper error bound $O(h + k^{3/2})$ for the error of the (smooth) exact stress Σ and their fully discrete approximation Σ_{hk} in the maximum norm in time and the L^2 norm in space, abbreviated as $L^\infty(L^2)$. It also bounds the sum of all jumps $[\Sigma_{hk}]_{j-1}$ at t_{j-1} in time,

$$\|\Sigma - \Sigma_{hk}\|_{L^\infty(L^2)}^2 + \sum_{j=1}^{J} \|[\Sigma_{hk}]_{j-1}\|_{L^2(\Omega)}^2 = O(h^2 + k^3),$$

for the maximal mesh-size h under sufficient smoothness assumptions of all continuous variables. The approximation order 1.5 appears to be suboptimal for the norm $L^\infty(L^2)$ compared to the Crank-Nicolson scheme which is second order in k, but the above error estimate also controls the sum of the jump errors and is admissible a.e. in Q. The latter summands may be compared to the jumps of the aforementioned dG approximation Σ_{dG} and they are merely $O(k^{3/2} \|\dot{\Sigma}\|_{L^2(Q_j)})$. An Aubin-Nitsche duality argument even shows superconvergence for the nodal values in linear model problems [20], but a corresponding argument has not been applied in the literature on elastoplasticity.

The remaining parts of the paper are organized as follows. The mathematical modeling of elastoplasticity is briefly recalled from the literature in Sect. 2. The resulting fully discrete scheme is a variational inequality at each time-step and it is given in Sect. 3. There are two discrete formulations, the dual and the primal one, and Theorem 3.2 establishes their equivalence. The error analysis in Sect. 4 explains and verifies the aforementioned asymptotic convergence result. A computational benchmark in Sect. 5 behind the Fig. 1 confirms the superiority of dG(1) over dG(0) even for an exact solution with reduced regularity.

Throughout this paper, the following notation applies for a time-space cylinder $Q = (0, t_J) \times \Omega$ and a typical part $Q_j := (t_{j-1}, t_j) \times \Omega$, for $j = 1, \dots, J$, with respect to a uniform time discretisation with $t_j := jk$ for $j = 0, \dots, J$ for a time increment $k > 0$. A lower index refers to the time-space discretization, e.g., Σ_{hk} for the discrete solution and $\Sigma_{hk,j}^\pm$ refers to the one-sided limit of the function Σ_{hk} at time t_j from the left (−) and the right (+). A tilde indicates a linear interpolation in

time with respect to the time points $t_0, \ldots, t_J$. There are various inner products like $\cdot$ for vectors, $:$ for matrices, and $\star$ for certain finite-dimensional product spaces in (2.2) below, which can all be seen as Euclidian scalar products if the objects are re-labelled as a one-dimensional list. The time-derivative of P is denoted $\dot{P} := \partial P / \partial t$ and the piecewise version of P_{hk} as $P_{hk,\tau}$. For any $d \times d$ matrix $\sigma \in \mathbb{R}_{sym}^{d \times d}$, the deviatoric part $\mathrm{dev}\,\sigma := \sigma - \frac{\mathrm{tr}(\sigma)}{d} I_{d \times d}$ for the $d \times d$ unit matrix $I_{d \times d}$ and the trace $\mathrm{tr}(\sigma) := \sigma_{11} + \sigma_{22} + \cdots + \sigma_{dd}$.

2 Mathematical modeling of elastoplasticity

This section is devoted to the strong form of a model example in elastoplasticity with hardening and the weak primal and dual form [12] for almost every point $(t, x) \in Q$ in time and space. This section is a short summary of an elastoplastic model problem for ease of reading and more details and relevant subcases are found in [2–4,7,8,12,17]. The generalized stress and generalized plastic strains are given as

$$\Sigma = (\sigma, \chi) \quad \text{and} \quad P = (p, \xi).$$

The stress variable σ and the total (linear Green) strain (with $u_{j,k} = \partial u_j / \partial x_k$)

$$\varepsilon(u) := \mathrm{sym}\, Du = ((u_{j,k} + u_{k,j})/2)_{j,k=1,\ldots,d}$$

are linked with the irreversible plastic strain p through an additive split

$$\varepsilon(u) = \mathbb{C}^{-1}\sigma + p$$

of small strain plasticity. The fourth-order elasticity tensor $\mathbb{C}$ acts as

$$\mathbb{C}q = \lambda\, \mathrm{tr}(q)\, I_{d \times d} + 2\mu\, q \quad \text{for all} \quad q \in \mathbb{R}_{sym}^{d \times d}$$

with trace $\mathrm{tr}(q) := q_{11} + \cdots + q_{dd}$, the unit matrix $I_{d \times d}$, and the Lamé constants $\lambda, \mu > 0$. The displacement field u is supposed to satisfy a Dirichlet boundary condition in the form

$$u = 0 \quad \text{on } \Gamma_D$$

for a fixed closed part Γ_D of $\partial\Omega = \Gamma$ of positive $d-1$ dimensional (surface) measure and almost every time t. The equilibrium model reads in local form

$$\sigma = \sigma^T \quad \text{and} \quad \mathrm{div}\,\sigma + f = 0 \quad \text{in } \Omega$$

together with a Neumann boundary condition on the remaining part of the boundary

$$\sigma n = g \quad \text{on} \quad \Gamma_N := \Gamma \backslash \Gamma_D$$

for almost every time t. In terms of the elastic strain $e := \mathbb{C}^{-1}\sigma$, the internal energy assumes the form

$$F(e, \xi) := 1/2(e : \mathbb{C}e + \xi \cdot \mathbb{H}\xi) \tag{2.1}$$

for the aforementioned fourth-order elasticity tensor $\mathbb{C}$ and a symmetric and positive definite hardening tensor $\mathbb{H}$. The internal (hardening) variables ξ and their dual variables χ are written (symbolically) as m dimensional vectors (e.g., the $m = d(d+1)/2$ components of a symmetric $d \times d$ matrix for kinematic hardening or scalar, $m = 1$, for linear isotropic hardening) with $m = 1 + d(d+1)/2$ for combined hardening with $\xi = (\alpha, \beta)$, $\chi = (a, b) \in \mathbb{R} \times \mathbb{R}_{sym}^{d \times d} \equiv \mathbb{R}^m$. Hence, $\xi, \chi \in \mathbb{R}^m$ and $\mathbb{H} \in \mathbb{R}_{sym}^{m \times m}$ is identified with an $m \times m$ SPD matrix in (2.1). Recall that $\varepsilon(u) = e + p$ and notice that

$$\sigma = \partial F(e, \xi)/\partial e = \mathbb{C}e \quad \text{and} \quad \chi = -\partial F(e, \xi)/\partial \xi = -\mathbb{H}\xi.$$

The Prandtl-Reuß flow rule reads (recall that $\dot{P}$ denotes the time derivative of P)

$$\dot{P} \in N_K(\Sigma) := \left\{ Q \in \mathbb{R}_{sym}^{d \times d} \times \mathbb{R}^m : \forall \overline{T} \in K, \ Q \star (\overline{T} - \Sigma) \leq 0 \right\}$$

for the set of admissible generalized stresses $K \subset \mathbb{R}_{sym}^{d \times d} \times \mathbb{R}^m$ determined by the yield function (e.g., the von-Mises yield function) $\Phi : \mathbb{R}_{sym}^{d \times d} \times \mathbb{R}^m \to \mathbb{R}$ via

$$K := \left\{ \overline{T} \in \mathbb{R}_{sym}^{d \times d} \times \mathbb{R}^m : \ \Phi(\overline{T}) \leq 0 \right\}.$$

Throughout this paper, we distinguish between the scalar products $\cdot, :, \star$, defined for vectors $u, v, d \times d$ matrices p, q, and generalized stresses or strains $P = (p, \xi)$, $Q = (q, \chi)$ by $u \cdot v = u_1 v_1 + \cdots + u_d v_d$, or $\xi \cdot \chi = \xi_1 \chi_1 + \cdots + \xi_m \chi_m$,

$$p : q := \sum_{j,k=1}^{d} p_{jk} q_{jk}, \text{ and } P \star Q := (p, \xi) \star (q, \chi) = p : q + \xi \cdot \chi. \tag{2.2}$$

The strong form of the continuous model problem of elastoplasticity with combined isotropic and kinematic hardening reads: Given data f and g as functions in time $[0, t_J]$ and space, given consistent homogeneous initial conditions (i.e., $f = g = 0$ for $t = 0$) the elastoplastic time-evolution determines $u, \sigma, \chi, p,$ and ξ as functions on $[0, t_J] \times \Omega$ with

$$\sigma = \sigma^T = \mathbb{C}(\varepsilon(u) - p), \text{ div } \sigma + f = 0, \ (\dot{p}, \dot{\xi}) \in N_K(\sigma, \chi) \text{ in } [0, t_J] \times \Omega$$

and the boundary conditions

$$u = 0 \text{ on } [0, t_J] \times \Gamma_D \quad \text{and} \quad \sigma n = g \text{ on } [0, t_J] \times \Gamma_N.$$

Following [7,12] the primal and dual formulations differ in the treatment of the elasto-plastic evolution law. The duality relation of subgradients of convex functions in convex analysis [10,21] leads to an equivalent reformulation

$$\dot{P} \in N_K(\Sigma) \Leftrightarrow \Sigma \in \partial \operatorname{supp}_K(\dot{P}).$$

The first inclusion is defined above and, given K via a yield function Φ, reads

$$\Phi(\Sigma) \leq 0 \quad \text{and} \quad \dot{P} \star (\overline{T} - \Sigma) \leq 0 \quad \text{for all } \overline{T} \in \mathbb{R}^{d \times d}_{sym} \times \mathbb{R}^m \quad \text{with } \Phi(\overline{T}) \leq 0.$$

The second inclusion involves the support function

$$\operatorname{supp}_K(Q) := \sup_{\overline{T} \in K} Q \star \overline{T} = \sup_{\Phi(\overline{T}) \leq 0} Q \star \overline{T}$$

and its subdifferential $\partial \operatorname{supp}_K$. Indeed, $\Sigma \in \partial \operatorname{supp}_K(\dot{P})$ reads

$$\Sigma \star (Q - \dot{P}) \leq \operatorname{supp}_K(Q) - \operatorname{supp}_K(\dot{P}) \quad \text{for all} \quad Q \in \mathbb{R}^{d \times d}_{sym} \times \mathbb{R}^m.$$

Throughout this paper, the combined kinematic and isotropic hardening are considered in the von Mises yield function, where $\Sigma = (\sigma, \chi)$ with $\chi = (a, b) = -\mathbb{H}\xi \in \mathbb{R}^m$ for $a \in \mathbb{R}$ and $b \in \mathbb{R}^{d \times d}_{sym} \equiv \mathbb{R}^{m-1}$ and

$$\Phi(\Sigma) \equiv \Phi(\sigma, a, b) := |\operatorname{dev}\sigma - \operatorname{dev} b| - \sigma_y(1 + Ha) \text{ for } \Sigma \equiv (\sigma, a, b) \in \mathbb{R}^{d \times d}_{sym} \times \mathbb{R}^m.$$

Here and throughout this paper, $\mathbb{R}^m \equiv \mathbb{R} \times \mathbb{R}^{d \times d}_{sym}$ and

$$\mathbb{H} = \operatorname{diag}(A, B) \in \mathbb{R}^{m \times m} \text{ for } A > 0 \text{ and positive definite } B \in \mathbb{R}^{(m-1) \times (m-1)}_{sym}.$$

The material parameters $\sigma_y > 0$ and $H \geq 0$ are fixed and constant in time and space; particular cases are discussed in more details in [2,4,7,8,12].

The hardening allows a control of the Green strain in terms of the generalized stresses.

Proposition 2.1 [8] *The functional* supp_K *is the dual of the characteristic functional* χ_K *of K (zero on K and elsewhere infinity) and reads*

$$\operatorname{supp}_K(p, \alpha, \beta) = \begin{cases} \sigma_y|p| & \text{if } \operatorname{tr} p = 0, \ p = -\beta, \text{ and } A^{-1}\alpha + \sigma_y H|p| \leq 0, \\ \infty & \text{otherwise.} \end{cases}$$

Moreover, the inclusion $(\sigma, \chi) \equiv (\sigma, a, b) \in \partial \operatorname{supp}_K(\dot{p}, \dot{\xi})$ *in case* $\dot{p} \neq 0$ *is equivalent to*

$$\frac{\dot{p}}{|\dot{p}|} = \frac{\operatorname{dev}(\sigma - b)}{\sigma_y(1 + Ha)} \quad \text{and} \quad \dot{\alpha} = -A\sigma_y H |\dot{p}|.$$

$\square$

The corresponding weak formulations of the complete model are derived by the principle of virtual displacements or by testing with a test function. The two resulting variational inequalities are summarized below; we refer to [4,12] for further details.

The weak formulation involves test and trial functions in standard Lebesgue and Sobolev spaces, namely,

$$L^2(\Omega) := \left\{ v : \Omega \to \mathbb{R} : v \text{ measurable with } \int_\Omega |v|^2 \, dx < \infty \right\},$$
$$H^1(\Omega) := \left\{ v \in L^2(\Omega) : \partial v / \partial x_1, \ldots, \partial v / \partial x_d \in L^2(\Omega) \right\}$$

and powers thereof (i.e., all components belong to the respective space); $\partial v / \partial x_j$ is a weak derivative [11,21]. Moreover,

$$V := \left\{ v \in H^1(\Omega; \mathbb{R}^d) : v = 0 \quad \text{on} \quad \Gamma_D \right\} \quad \text{and} \quad L := L^2\left(\Omega; \mathbb{R}^{d \times d}_{sym} \times \mathbb{R}^m\right).$$

The weak *primal formulation* seeks $(u, p, \xi) : [0, t_J] \times \Omega \to \mathbb{R}^d \times \mathbb{R}^{d \times d}_{sym} \times \mathbb{R}^m$ with homogeneous initial values and a Dirichlet boundary condition on Γ_D such that, for a.e. time $t \in (0, t_J)$ and for all $v \in V$ and all $(q, \zeta) \in L$, one has

$$\int_\Omega \mathbb{C}(\varepsilon(u(t)) - p(t)) : (\varepsilon(v) - \dot{p}(t) + q) \, dx - \int_\Omega \xi(t) \star \mathbb{H}(\zeta - \dot{\xi}(t)) \, dx$$
$$\leq \int_\Omega f(t) \cdot v \, dx + \int_{\Gamma_N} g(t) \cdot v \, ds + \int_\Omega \mathrm{supp}_K(q, \zeta) \, dx$$
$$- \int_\Omega \mathrm{supp}_K(\dot{p}(t), \dot{\xi}(t)) \, dx.$$

The *dual formulation* seeks $(u, \sigma, \chi) : [0, t_J] \times \Omega \to \mathbb{R}^d \times \mathbb{R}^{d \times d}_{sym} \times \mathbb{R}^m$ with homogeneous initial values and a Dirichlet boundary condition on Γ_D, such that, for a.e. time $t \in (0, t_J)$ and all $v \in V$, one has

$$\int_\Omega \sigma(t) : \varepsilon(v) \, dx = \int_\Omega f(t) \cdot v \, dx + \int_{\Gamma_N} g(t) \cdot v \, ds \tag{2.3}$$

and $\Phi(\sigma(t), \chi(t)) \leq 0$ such that, for all $(\tau, \psi) \in L^2(\Omega; \mathbb{R}^{d \times d}_{sym} \times \mathbb{R}^m)$ with $\Phi(\tau, \psi) \leq 0$, one has

$$\int_\Omega \left(\varepsilon(\dot{u}(t)) - \mathbb{C}^{-1} \dot{\sigma}(t) \right) : (\tau - \sigma(t)) \, dx - \int_\Omega \dot{\chi}(t) \star \mathbb{H}^{-1}(\psi - \chi(t)) \, dx \leq 0. \tag{2.4}$$

On the continuous level the dual and primal formulation are equivalent [4,12]. The role of the time-derivative and the choice of variables are different. The unique existence of solutions is well established [12–14,18,19] as well as a certain regularity result for the stress variables [16]. The work [15] on a related evolution problem presents comprehensive regularity results and illustrates what smoothness can be expected for rate-independent problems.

The time in the rate-independent quasi-static model is a process time and the loads may even allow for a separation $f(t, x) = f_1(t) f_2(x)$ for a.e. $(t, x) \in Q$. Computational examples even consider $f_1(t) = t$ and may lead to rather complicated stress-strain interactions.

3 Discretization

This section is devoted to the motivation of discontinuous Galerkin time discretization schemes for elastoplastic evolution problems. The point of departure is the definition of a distributional derivate for discontinuous, but $\mathcal{I}$-piecewise smooth test functions. The time interval $[0, t_J]$ is partitioned in J subintervals $I_j = (t_{j-1}, t_j]$ (open at the left and closed at the right) of length k, $t_j := jk$ for $j = 0, \dots, J$. Define the set of $\mathcal{I}$-piecewise smooth functions by

$$
\mathcal{C}^1(\mathcal{I}) := \left\{ v \in L^\infty(-\infty, t_J) : v|_{(-\infty, 0]} := 0 \text{ and } v|_{I_j} \in \mathcal{C}^1[t_{j-1}, t_j] \text{ for } j = 1, 2, \dots, J \right\}.
$$

Piecewise uniformly continuous functions u allow for one-sided limits and the definition of the jump

$$
[u]_j := u(t_j^+) - u(t_j^-) \quad \text{with} \quad u_j^\pm := u(t_j^\pm) := \lim_{t \to t_j^\pm} u(t) \quad \text{for} \quad j = 0, \dots, J - 1.
$$

(Throughout, $u_0^- := 0$ owing to the homogeneous initial conditions, whence $[u]_0 := u_0^+$.) Since $u|_{(t_{j-1}, t_j)}$ is $\mathcal{C}^1(t_{j-1}, t_j)$, the time derivative $u_\tau := \partial u / \partial t$ exists on each open interval (t_{j-1}, t_j) in the classical sense as limits of difference quotients. The distributional derivative $\dot{u}$ is defined through

$$
\int_{-\infty}^{t_J} \dot{u}(t) v(t) dt = - \int_{-\infty}^{t_J} u(t) \dot{v}(t) dt
$$

for all test functions $v \in \mathcal{C}_c^\infty(-\infty, t_J)$. With the delta distribution δ_{t_j} at the point t_j, it reads

$$
\dot{u} = u_\tau + \sum_{j=0}^{J-1} [u]_j \delta_{t_j}.
$$

This explains the action of $\dot{u}$ on differentiable test functions. An extension to discontinuous test functions starts with globally continuous and piecewise $\mathcal{C}^1$ test functions v_ϵ which vanish outside some fixed interval I_j. The functions v_ϵ are defined by multiplication of $v \in \mathcal{C}^1(\mathbb{R})$ with the step functions χ_ϵ^j of Fig. 2,

$$
v_\epsilon(t) := \chi_\epsilon^j(t) v(t).
$$
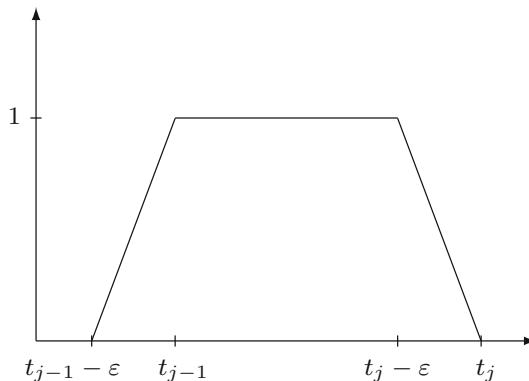
Fig. 2 Test function χ_ε^j

A formulation of a distributional derivative for the fixed interval I_j follows from the equation

$$\int_{-\infty}^{t_J} \dot{u}(t) v_\epsilon(t)\, dt = \int_{t_{j-1}-\epsilon}^{t_j} u_\tau(t) v_\epsilon(t)\, dt + [u]_{j-1} v_\epsilon(t_{j-1})$$

in the limit $\epsilon \to 0$, which is equal to $\int_{I_j} u_\tau(t) v(t)\, dt + [u]_{j-1} v(t_{j-1}^+)$. This argument applies to a general $v \in C^1(\mathcal{I})$ and leads to the formula

$$\lim_{\epsilon \to 0} \int_{\mathbb{R}} \dot{u} \left(\sum_{j=1}^{J} v_\epsilon^j \right) dt = \int_0^{t_J} u_\tau(t) v(t)\, dt + \sum_{j=1}^{J} [u]_{j-1} v\left(t_{j-1}^+\right).$$

(For a proof, extend the function $v_{I_j} \in C^1(I_j)$ to $\hat{v}_j \in C^1(\mathbb{R})$ and multiply by χ_ϵ^j to define $v_\epsilon^j := \chi_\epsilon^j \hat{v}_j$ for $j = 1, \ldots, J$). The spline space of discontinuous Galerkin functions of order ℓ is defined through

$$P_\ell(\mathcal{I}; X) = \left\{ u \in L^\infty([0, t_J], X) : \forall\, j = 1, \ldots, J,\ u|_{I_j} \in P_\ell(I_j; X) \right\}.$$

The homogeneous initial data are reflected in the convention that $u(0^-) = 0$ for all discrete values below. The discrete loads $f_k \in P_1(\mathcal{I}; L^2(\Omega; \mathbb{R}^d))$ and $g_k \in P_1(\mathcal{I}; L^2(\Gamma_N; \mathbb{R}^d))$ are seen as approximations to the respective data f and g, as in Example 3.1 below.

The domain Ω is partitioned into triangles and parallelograms for $2D$ and tetrahedra for $3D$. The resulting triangulation $\mathcal{T}$ is supposed to be regular in the sense of Ciarlet [6,9]. For each $T \in P_\ell(\mathcal{T})$ denotes the algebraic polynomials on T of degree $\leq \ell$ and $P_\ell(\mathcal{T}; \mathbb{R}^d) \equiv P_\ell(\mathcal{T})^d$. The required finite element function spaces read

$$P_\ell(\mathcal{T}; \mathbb{R}^d) := \left\{ v \in L^2(\Omega) : \forall T \in \mathcal{T},\ v|_T \in P_\ell(T; \mathbb{R}^d) \right\},$$

$$V_h := P_1(\mathcal{T}; \mathbb{R}^d) \cap V,$$

$$L_h := P_0\left(\mathcal{T}; \mathbb{R}^{d \times d}_{sym} \times \mathbb{R}^m\right),$$

$$K_h := P_0(\mathcal{T}; K) = \{\Sigma_h \in L_h : \Sigma_h \in K \ \ a.e.\ in\ \Omega\}.$$

The finite element approximation is denoted by subindices k and h (neglected for its continuous counterpart) as the underlying discretization is based on a partition $\mathcal{I}$ in time and a regular triangulation $\mathcal{T}$ in space. Recall the following abbreviations used throughout this paper

$$Q_j := I_j \times \Omega \ \ \text{and} \ \ \int_{I_j} \int_\Omega \ldots dx\, dt \ \ \text{reads} \ \ \int_{Q_j} \ldots dQ.$$

The *discrete primal problem* seeks $(u_{hk}, P_{hk}) \in P_1(\mathcal{I}; V_h \times L_h)$ such that $\Sigma_{hk} = (\sigma_{hk}, \chi_{hk}) := (\mathbb{C}(\varepsilon(u_{hk}) - p_{hk}), -\mathbb{H}\xi_{hk}) \in P_1(\mathcal{I}; L_h)$ and $P_{hk} := (p_{hk}, \xi_{hk})$ satisfy, for all $j = 1, \ldots, J$ and for all $v_{hk} \in P_1(I_j; V_h)$ and all $Q_{hk} \in P_1(I_j; L_h)$, that

$$\int_{Q_j} \sigma_{hk} : \varepsilon(v_{hk})\, dQ = \int_{Q_j} f_k \cdot v_{hk}\, dQ + \int_{I_j} \int_{\Gamma_N} g_k \cdot v_{hk}\, ds\, dt; \qquad (3.1)$$

$$\int_{Q_j} \Sigma_{hk} \star (Q_{hk} - P_{hk,\tau}) dQ - \sum_{j=1}^{J} \int_\Omega (\Sigma_{hk})^+_{j-1} \star [P_{hk}]_{j-1} dx$$

$$\leq \sup_{R_{hk} \in P_1(I_j; K_h)} \int_{Q_j} R_{hk} \star Q_{hk} dQ$$

$$- \sup_{S_{hk} \in P_1(I_j; K_h)} \left(\int_{Q_j} S_{hk} \star P_{hk,\tau} dQ + \int_\Omega (S_{hk})^+_{j-1} \star [P_{hk}]_{j-1}\, dx \right). \quad (3.2)$$

The *discrete dual problem* seeks $(u_{hk}, \Sigma_{hk}) \in P_1(\mathcal{I}; V_h \times K_h)$ such that $\Sigma_{hk} = (\sigma_{hk}, \chi_{hk})$ and $P_{hk} := (p_{hk}, \xi_{hk}) = (\varepsilon(u_{hk}) - \mathbb{C}^{-1}\sigma_{hk}, -\mathbb{H}^{-1}\chi_{hk}) \in P_1(\mathcal{I}; L_h)$ satisfy (3.1) for $j = 1, \ldots, J$ and for all $v_{hk} \in P_1(I_j; V_h)$ and $T_{hk} \in P_1(I_j; K_h)$ that

$$\int_{Q_j} P_{hk,\tau} \star (T_{hk} - \Sigma_{hk})\, dQ + \int_\Omega [P_{hk}]_{j-1} \star (T_{hk} - \Sigma_{hk})^+_{j-1} dx \leq 0. \qquad (3.3)$$

The jump terms in this inequality reflect the aforementioned construction of the distributional derivative for the discontinuous test functions. The implementation of the dG(1) discretization is described in [3].

Lemma 3.1 (pointwise equilibrium) *Given any discrete loads $f_k \in P_1(\mathcal{I}; L^2(\Omega; \mathbb{R}^d))$ and $g_k \in P_1(\mathcal{I}; L^2(\Gamma_N; \mathbb{R}^d))$ and let $\sigma_{hk} \in P_1(\mathcal{I}; P_0(\mathcal{T}; \mathbb{R}^{d \times d}))$ satisfy (3.1). Then any $t \in I_j$, $j = 1, \ldots, J$, and any $v_h \in V_h$ satisfy*

$$\int_\Omega f_k(t) \cdot v_h \, dx + \int_{\Gamma_N} g_k(t) \cdot v_h \, ds = \int_\Omega \sigma_{hk}(t) : \varepsilon(v_h) \, dx. \tag{3.4}$$

Proof Given any $t \in I_j$, $j = 1, \ldots, J$, let $F(t) \in V_h$ denote the Riesz representation of the functional on the right-hand side

$$\int_\Omega F(t) \cdot v_h \, dx = \int_\Omega f_k(t) \cdot v_h \, dx + \int_{\Gamma_N} g_k(t) \cdot v_h \, ds - \int_\Omega \sigma_{hk}(t) : \varepsilon(v_h) \, dx.$$

in the Hilbert space V_h with respect to the scalar product of $L^2(\Omega; \mathbb{R}^n)$ (and so the above equality holds for all $v_h \in V_h$). Then $F(t)$ is affine in $t \in I_j$, written $F \in P_1(I_j; V_h)$. Endow $P_1(I_j; V_h)$ with the scalar product of $L^2(Q_j)$. The equilibrium (3.1) shows that the scalar product $\int_{Q_j} F(t) \cdot v_{hk} \, dx = 0$ for all $v_{hk} \in P_1(I_j; V_h)$. Consequently, $F \in P_1(I_j; V_h)$ vanishes in $L^2(Q_j)$ and so vanishes pointwise almost everywhere. This implies (3.4). This identity follows for the one-sided limits $t_j^\pm$ as well (notice that (f_k, g_k) may have jumps). $\qquad\square$

Example 3.1 (discrete loads) (a) *Linear interpolation* For continuous data (f, g), let $(f_k, g_k) = (\tilde{f}, \tilde{g})$ denote the nodal interpolation, also denoted for other quantities throughout this paper by a tilde,

$$(f_k, g_k)(t) = (t_j - t)/k \, (f, g)(t_{j-1}) + (t - t_{j-1})/k \, (f, g)(t_j) \text{ for } t_{j-1} \leq t \leq t_j.$$

(b) *Natural choice* The discretisation with an approximation (f_k, g_k) to (f, g) in (3.1) may appear artificial at the first glance and one may substitute (f_k, g_k) in (3.1) by (f, g). This is equivalent to the choice of (f_k, g_k) as orthogonal projections of (f, g) onto $P_1(\mathcal{I}; L^2(\Omega \times \Gamma_N; \mathbb{R}^n))$ with respect to the scalar product of $L^2((0, t_J) \times \Omega \times \Gamma_N; \mathbb{R}^n)$.

Remark 3.2 (approximation of loads) All the data approximation terms of Example 3.1 satisfy, for $j = 1, \ldots, J$, the approximation property

$$\begin{aligned}
\|(f - f_k, g - g_k)\|_{L^\infty(I_j; V^* \times V^*)} &\leq C_I k \|(\ddot{f}, \ddot{g})\|_{L^1(I_j; V^* \times V^*)} \\
&\leq C_I k^{3/2} \|(\ddot{f}, \ddot{g})\|_{L^2(I_j; V^* \times V^*)}
\end{aligned} \tag{3.5}$$

with some universal constant C_I and all $(f, g) \in H^2(0, t_J; L^2(\Omega \times \Gamma_N; \mathbb{R}^n))$.

Further error analysis is contained in the next section, while a few general properties of the two discrete problems conclude this one.

Theorem 3.2 *There exist unique solutions to the discrete primal problem and the discrete dual problem. The discrete primal problem and the discrete dual problem are equivalent.*

Proof The existence and uniqueness is proven for the primal formulation in [3] and that of the dual formulation follows from the equivalence to be established here. Let

$(u_{hk}, \Sigma_{hk}, P_{hk})$ solve the discrete primal problem. Since $P_1(I_j; L_h)$ is a Hilbert space (with respect to the scalar product of $L^2(Q_j)$), the linear functional associated with $[P_{hk}]_{j-1}$ and $P_{hk,\tau}|_{I_j}$ has the Riesz representation $\dot{P}_{hk} \in P_1(I_j; L_h)$ defined by

$$\int_{Q_j} \dot{P}_{hk} \star T_{hk} dQ = \int_{Q_j} P_{hk,\tau} \star T_{hk} dQ$$
$$+ \int_{\Omega} [P_{hk}]_{j-1} \star (T_{hk})_{j-1}^+ dx \quad \text{for all } T_{hk} \in P_1(I_j; L_h).$$

Then (3.2) means that all $Q_{hk} \in P_1(I_j; L_h)$ and all $S_{hk} \in P_1(I_j; K_h)$ satisfy

$$\int_{Q_j} (Q_{hk} - \dot{P}_{hk}) \star \Sigma_{hk} dQ \leq \sup_{R_{hk} \in P_1(I_j; K_h)} \int_{Q_j} Q_{hk} \star R_{hk} dQ - \int_{Q_j} \dot{P}_{hk} \star S_{hk} dQ.$$

Some reformulations result in

$$\int_{Q_j} (Q_{hk} - \dot{P}_{hk}) \star (\Sigma_{hk} - S_{hk}) dQ \leq \sup_{R_{hk} \in P_1(I_j; K_h)} \int_{Q_j} Q_{hk} \star (R_{hk} - S_{hk}) dQ. \quad (3.6)$$

The substitution of $Q_{hk} \in P_1(I_j; L_h)$ by $\overline{Q}_{hk} := \lambda Q_{hk}$ with $\lambda \to \infty$ in (3.6) leads to

$$0 \leq \sup_{R_{hk} \in P_1(I_j; K_h)} \int_{Q_j} \overline{Q}_{hk} \star (R_{hk} - S_{hk}) dQ \quad \text{for all } \overline{Q}_{hk} \in P_1(I_j; L_h). \quad (3.7)$$

Given almost every $x \in \Omega$ and $\Sigma_{hk}(\cdot, x) \in P_1(I_j; \mathbb{R}_{sym}^{d \times d} \times \mathbb{R}^m)$ let $\hat{\Sigma}_{hk}(\cdot, x)$ denote its projection onto $P_1(I_j; K_h)$ in the Hilbert space $P_1(I_j; \mathbb{R}_{sym}^{d \times d} \times \mathbb{R}^m)$ with respect to the scalar product of $L^2(I_j; \mathbb{R}_{sym}^{d \times d} \times \mathbb{R}^m)$. Then, $\overline{Q}_{hk} := \Sigma_{hk} - \hat{\Sigma}_{hk}$ in (3.7) implies that

$$\|\hat{\Sigma}_{hk} - \Sigma_{hk}\|_{L^2(Q_j)}^2 \leq \sup_{R_{hk} \in P_1(I_j; K_h)} \int_{Q_j} (R_{hk} - \hat{\Sigma}_{hk}) \star (\Sigma_{hk} - \hat{\Sigma}_{hk}) dQ.$$

Since R_{hk} belongs to $P_1(I_j; K_h)$ and $\hat{\Sigma}_{hk}$ is the projection onto this, the scalar product on the right-hand side is non-positive. Hence $\Sigma_{hk} = \hat{\Sigma}_{hk}$ and $\Sigma_{hk} \in P_1(I_j; K_h)$ is admissible. For all $S_{hk} \in P_1(I_j; K_h)$, the choice $Q_{hk} = 0$ in (3.6) leads to

$$\int_{Q_j} \dot{P}_{hk} \star (S_{hk} - \Sigma_{hk}) dQ \leq 0. \quad (3.8)$$

Hence $\dot{P}_{hk} \in N_{K_h}(\Sigma_{hk})$ and (3.3) of the dual formulation is verified.

To prove the converse implication, suppose $\Phi(\Sigma_{hk}) \leq 0$ and (3.3) for all $T_{hk} \in P_1(I_j; K_h)$. This means that

$$\sup_{T_{hk} \in P_1(I_j;K_h)} \int_{Q_j} \dot{P}_{hk} \star T_{hk} dQ_j = \int_{Q_j} \dot{P}_{hk} \star \Sigma_{hk} dQ_j.$$

This implies the discrete variational inequality (3.2) for $Q_{hk} = 0$. Finally, since $\Sigma_{hk} \in K_h$,

$$\int_{Q_j} Q_{hk} \star \Sigma_{hk} dQ \leq \sup_{R_h \in P_1(I_j;K_h)} \int_{Q_j} Q_{hk} \star R_{hk} dQ$$

for any $Q_{hk} \in P_1(I_j; L_h)$. The sum of the two displayed inequalities is (3.2). $\qquad\square$

The discrete dual formulation reveals explicitly that $\Sigma_{hk} = (\sigma_{hk}, \chi_{hk}) \in P_1(\mathcal{I}; K_h)$ is admissible a.e. in Q. With an appropriate interpretation of $N_{P_1(I_j;K)}$, the inequality (3.8) in the equivalence proof reads

$$\dot{P}_{hk|I_j \times T} \in N_{P_1(I_j;K)}(\Sigma_{hk|I_j \times T}).$$

This is *not* a pointwise version of Proposition 2.1: The dG(1) discretization merely ensures an averaged discrete material evolution law. This causes a new difficulty in the error analysis in comparison to the backward Euler and the Crank–Nicolson schemes. Nevertheless, the discrete kinematic hardening rule allows to deduce a certain discrete material evolution law pointwise in time-space.

Lemma 3.3 (discrete hardening) *The discrete primal solution $P_{hk} \equiv (p_{hk}, \alpha_{hk}, \beta_{hk})$ satisfies*

$$\operatorname{tr}(p_{hk}) = 0 \quad and \quad p_{hk} = -\beta_{hk} \quad a.e. \ in \ Q.$$

Proof Let $\dot{P}_{hk} \equiv (\dot{p}_{hk}, \dot{\alpha}_{hk}, \dot{\beta}_{hk})$ denote the Riesz representation of the discrete plastic strain rate from the proof of Theorem 3.2 and fix a time interval for $j = 1, \ldots, J$. The variational inequality (3.2) implies that the last supremum $M := \sup_{S_{hk} \in P_1(I_j;K_h)} \int_{Q_j} S_{hk} \star \dot{P}_{hk} dQ < \infty$ is finite. For any positive $r > 0$, the test functions $S_{hk} := (\sigma_h, a_h, b_h) := r \operatorname{tr}(\dot{p}_{hk})(1_{d \times d}, 0, 0)$ and $S_h = r(R, 0, R)$ for $R := \dot{p}_{hk} + \dot{\beta}_{hk} \in \mathcal{P}_1(I_j; P_0(\mathcal{T}; \mathbb{R}_{sym}^{d \times d}))$ satisfy $\Phi(\sigma_{hk}, a_{hk}, b_{hk}) \leq 0$ a.e. in Q_j. Hence the supremum M is a global upper bound for the integral $\int_{Q_j} S_{hk} \star \dot{P}_{hk} dQ$ evaluated at the two particular functions $S_{hk} \in P_1(I_j; K_h)$. This leads to

$$r \, \|\operatorname{tr}(\dot{p}_{hk})\|_{L^2(Q_j)}^2 \leq M \quad and \quad r \, \|\dot{p}_{hk} + \dot{\beta}_{hk}\|_{L^2(Q_j)}^2 \leq M.$$

Since $M < \infty$ is fixed, $r \to \infty$ implies

$$\operatorname{tr}(\dot{p}_{hk}) = 0 \quad and \quad \dot{p}_{hk} = -\dot{\beta}_{hk} \quad a.e. \ in \ Q_j.$$

This holds for the Riesz representation $\dot{p}_{hk}$ of the distributional time derivative of p_{hk}. Consequently $\mathrm{tr}([p_{hk}]_{j-1}) = 0$ in Ω and $\mathrm{tr}(p_{hk,\tau}) = 0$ in Q_j. Together with the initial conditions $\mathrm{tr}(p_{hk}) = 0$, one concludes $\mathrm{tr}(p_{hk}) = 0$ a.e. in Q. The same arguments show $p_{hk} = -\beta_{hk} = B^{-1}b_{hk}$ a.e. in Q. ☐

4 A priori convergence analysis

The main results are stated in the first subsection followed by their verifications.

4.1 A priori error estimate

Given the elasticity and hardening tensors $\mathbb{C}$ and $\mathbb{H}$, let $\mathbb{A} := \mathrm{diag}(\mathbb{C}^{-1}, \mathbb{H}^{-1})$ and define the associated norm $|||\bullet|||$ by

$$|||(\tau, \psi)|||^2 := \int_\Omega (\tau, \psi) \star \mathbb{A}(\tau, \psi) dx \quad \text{for all} \quad (\tau, \psi) \in L^2\left(\Omega; \mathbb{R}^{d \times d}_{sym} \times \mathbb{R}^m\right). \tag{4.1}$$

The (in general discontinuous) errors $\Delta := \widetilde{\Sigma} - \Sigma_{hk}$ allow for a right and left limit $\Delta_j^\pm$ at $t_j = jk$; more generally $\bullet_j^+$ (resp.$\bullet_j^-$) denote the one-sided limits $\lim_{t \searrow t_j} \bullet(t)$ (resp. $\lim_{t \nearrow t_j} \bullet(t)$) at t_j. The local mesh-sizes $h_{\mathcal{T}} \in P_0(\mathcal{T})$ are defined by $h_{\mathcal{T}}|_T := \mathrm{diam}(T)$ for all $T \in \mathcal{T}$ and the time intervals are uniformly distributed with the time increment $k > 0$. The following result implies that the stress error converges like $O(h + k^{3/2})$ in terms of the maximal mesh-size $h := \max h_{\mathcal{T}}$.

Theorem 4.1 (main result) *Suppose that* $u \in W^{1,1}(0, t_J; H^2(\Omega; \mathbb{R}^d))$, $P \in W^{3,1}(0, t_J; L)$, *and in (b)* $(f, g) \in W^{2,1}(0, t_J; V^* \times V^*)$. *Then there exist constants* $C_1(\mathbb{A})$ *and* $C_2(\mathbb{A})$, *which depend on all material constants as well as the shapes of the element domains, but not on their sizes, such that the stress error* $\Delta := \widetilde{\Sigma} - \Sigma_{hk}$ *satisfies (a) and (b) for the choice of the discrete data* (f_k, g_k) *in Example* 3.1.

(a) For Example 3.1.a

$$Err := \max\{|||\Delta_0^+|||^2, |||\Delta_1^-|||^2, |||\Delta_1^+|||^2, \dots, |||\Delta_J^-|||^2\} + \sum_{j=1}^J |||[\Delta]_{j-1}|||^2$$

$$\leq C_1(\mathbb{A}) \left(\| \left(k^2 \dddot{P}, h_{\mathcal{T}} D^2 \dot{u}\right) \|_{L^1(L^2)}^2 + \|k^{3/2} \ddot{P}\|_{L^2(L^2)}^2 + \|k^2 \dddot{P}\|_{L^\infty(L^2)}^2 \right).$$

(b) For Example 3.1.b, *the lower bound* Err *is bounded by*

$$Err \leq C_2(\mathbb{A}) \left(\| \left(k^2 \dddot{P}, h_{\mathcal{T}} D^2 \dot{u}\right) \|_{L^1(L^2)}^2 \right.$$

$$\left. + \| \left(k \ddot{f}, k \ddot{g}, k^{3/2} \ddot{P}\right) \|_{L^2(V^* \times V^* \times L)}^2 + \|k^2 \dddot{P}\|_{L^\infty(L^2)}^2 \right).$$

Remark 4.1 (alternative discrete loads) The proof of the theorem shows that certain critical jump terms disappear if the discrete loads $(f_k, g_k) \equiv (\widetilde{f}, \widetilde{g})$ are nodal interpolations of (f, g) in (3.1) as in Example 3.1.a. This leads to an improved convergence

rate $O(h + k^{3/2})$ in time in (a) in comparison to $O(h + k)$ for the seemingly natural choice of Example 3.1.b. One conclusion is the choice of h and k in practice with $h = k^{3/2}$ or $h = k$. A discussion on the sharpness of the mathematical analysis follows in Sect. 4.7.

Remark 4.2 (the role of hardening) The hardening parameters allow for the control of the Green strain error $\varepsilon(u - u_{hk})$ and the plastic strain $P - P_{hk}$ in terms of the generalized stress error in Sect. 4.5. The proof below shows that

$$\|(\Sigma - \Sigma_{hk}, P - P_{hk}, u - u_{hk})\|_{L^\infty(L \times L \times V)} = O(h + k^{3/2})$$

is controlled by the upper bound in the assertion (a) (resp. $O(h+k)$ in (b)) as well (with multiplicative constants which tend to infinity as the hardening parameters become smaller and smaller).

Remark 4.3 (perfect plasticity) In case of no hardening ($m = 0$) but appropriate smoothness of the exact solutions, the assertion of Theorem 4.1.a remains true. In that case (a), the constant $C(\mathbb{A})$ depends exclusively on the elastic material parameters in $\mathbb{C}$. This is very different in case (b), where the constant $C(\mathbb{A})$ tends to infinity as the hardening parameters become smaller and smaller and the estimate (b) fails in the case of perfect plasticity. The work [5] analises the limit for vanishing hardening.

Remark 4.4 (constant coefficients) For notational simplicity, all material constants $\mathbb{A}$, A, B, $\mathbb{C}$, $\mathbb{H}$, H, σ_y are constant in space and time. We refer to [8] for variable material parameters in space and a corresponding perturbation analysis.

4.2 Discrete approximations

Given the exact solution Σ and P, define $\widetilde{\Sigma}$ and $\widetilde{P}$ as the linear interpolation of Σ and P in time

$$\widetilde{\Sigma}(t)|_{I_j} = \frac{t - t_{j-1}}{t_j - t_{j-1}} \Sigma(t_j) + \frac{t_j - t}{t_j - t_{j-1}} \Sigma(t_{j-1}) \quad \text{for} \quad t_{j-1} \le t \le t_j.$$

$$\widetilde{P}(t)|_{I_j} = \frac{t - t_{j-1}}{t_j - t_{j-1}} P(t_j) + \frac{t_j - t}{t_j - t_{j-1}} P(t_{j-1}) \quad \text{for} \quad t_{j-1} \le t \le t_j.$$

Define Σ_{hk}^* as piecewise integral mean (with the volume $|T|$ of $T \in \mathcal{T}$) of $\widetilde{\Sigma}$ by

$$\Sigma_{hk}^*(t)|_T := \frac{1}{|T|} \int_T \widetilde{\Sigma}(t, x) \, dx \in K_h \quad \text{for all} \quad T \in \mathcal{T}, \, 0 \le t \le t_J.$$

Abbreviate the time difference $[\bullet]_{j-1}^j := (\bullet|_{(t_{j-1}, t_j)})(t_j) - (\bullet|_{(t_{j-1}, t_j)})(t_{j-1})$ and the average value $\langle \bullet \rangle_j = (\bullet_j^+ + \bullet_j^-)/2$ for the one-sided limits indicated by an upper index $\pm$ at the time point t_j, e.g., $\bullet_j^- := (\bullet)(t_j^-) = \lim_{t \nearrow t_j^-}(\bullet)(t)$.

Lemma 4.2 (approximate flow rule) *It holds for all $j = 1, \ldots, J$ that*

$$-\int_{Q_j} \tilde{\dot{P}} \star (\tilde{\Sigma} - \Sigma_{hk}) dQ \leq \frac{k}{4} \int_{\Omega} [\dot{P}]_{j-1}^j \star [\tilde{\Sigma} - \Sigma_{hk}]_{j-1}^j dx$$

$$+ \frac{k^2}{8} \|\dddot{P}\|_{L^1(I_j; L^2(\Omega))} \|(\tilde{\Sigma} - \Sigma_{hk})(t_{j-1/2})\|_{L^2(\Omega)}.$$

Proof The Prandtl-Reuß flow rule (2.4) and $p = \varepsilon(u) - \mathbb{C}^{-1}\sigma$ and $\chi = -\mathbb{H}\xi$ from Sect. 2 show that

$$\int_{\Omega} \dot{P}_j \star (\Sigma_{hk,j}^- - \Sigma_j) dx \leq 0 \quad \text{and} \quad \int_{\Omega} \dot{P}_{j-1} \star (\Sigma_{hk,j-1}^+ - \Sigma_{j-1}) dx \leq 0.$$

The sum of the two inequalities leads to

$$\int_{\Omega} (\dot{P}_j + \dot{P}_{j-1}) \star \left(\Sigma_{hk,j}^- + \Sigma_{hk,j-1}^+ - (\Sigma_j + \Sigma_{j-1}) \right) dx$$

$$+ \int_{\Omega} (\dot{P}_j - \dot{P}_{j-1}) \star \left(\Sigma_{hk,j}^- - \Sigma_j - (\Sigma_{hk,j-1}^+ - \Sigma_{j-1}) \right) dx \leq 0.$$

A direct calculation shows that this is equivalent to

$$-\frac{1}{2} \int_{Q_j} (\dot{P}_j + \dot{P}_{j-1}) \star (\tilde{\Sigma} - \Sigma_{hk}) dQ \leq \frac{k}{4} \int_{\Omega} [\dot{P}]_{j-1}^j \star [\tilde{\Sigma} - \Sigma_{hk}]_{j-1}^j dx. \quad (4.2)$$

The (negative of the) left-hand side of (4.2) is split into two parts,

$$\int_{Q_j} \frac{P_j - P_{j-1}}{k} \star (\tilde{\Sigma} - \Sigma_{hk}) dQ$$

$$+ \int_{Q_j} \left(\frac{\dot{P}_j + \dot{P}_{j-1}}{2} - \frac{P_j - P_{j-1}}{k} \right) \star (\tilde{\Sigma} - \Sigma_{hk}) dQ,$$

to obtain the first term of the lemma and a remaining term. A direct calculation leads to the following identities (cf., e.g., (5.10)–(5.11) in [2])

$$2k \left(\frac{\dot{P}_j + \dot{P}_{j-1}}{2} - \frac{P_j - P_{j-1}}{k} \right) = k(\dot{P}_j - 2\dot{P}_{j-1/2} + \dot{P}_{j-1})$$

$$- 2(P_j - k\dot{P}_{j-1/2} - P_{j-1})$$

$$= k \int_0^{k/2} \int_{-t}^t \ddot{P}(t_{j-1/2} + s) ds \, dt - 2 \int_0^{k/2} \int_0^t \int_{-s}^s \dddot{P}(t_{j-1/2} + r) dr \, ds \, dt.$$

With the temporary abbreviation $g(t) := \int_0^t \int_{-s}^s \dddot{P}(t_{j-1/2}+r)dr ds$, this term is equal to

$$kg(k/2) - 2\int_0^{k/2} g(t)dt = 2\int_0^{k/2}\int_t^{k/2} g'(s)ds dt.$$

The combination with (4.2) (recall the notation $dQ = dx dt$ and $\tilde{P} = (P_j - P_{j-1})/k$ etc.) shows

$$-\int_{Q_j} \tilde{\dot{P}} \star (\tilde{\Sigma} - \Sigma_{hk})dQ - \frac{k}{4}\int_\Omega [\dot{P}]_{j-1}^j \star [\tilde{\Sigma} - \Sigma_{hk}]_{j-1}^j dx$$

$$\leq \frac{1}{k}\int_{Q_j}\left(\int_0^{k/2}\int_t^{k/2} g'(s)ds dt\right)\star(\tilde{\Sigma} - \Sigma_{hk})dQ$$

$$\leq \frac{k^2}{8}\|\dddot{P}\|_{L^1(I_j;L^2(\Omega))}\|(\tilde{\Sigma} - \Sigma_{hk})(t_{j-1/2})\|_{L^2(\Omega)}.$$

$$\square$$

The following lemma requires uniform time steps and identifies one term that causes the order $O(k^{3/2})$ in the final result.

Lemma 4.3 (time approximation) *It holds for all $\ell = 1, \ldots, J$ that*

$$\sum_{j=1}^\ell k \int_\Omega [\dot{P}]_{j-1}^j \star [\tilde{\Sigma} - \Sigma_{hk}]_{j-1}^j\, dx$$

$$\leq k^{3/2}\|\ddot{P}\|_{L^2(0,t_\ell;L^2(\Omega))}\sqrt{\sum_{j=1}^{\ell-1}\|[\Sigma_{hk}]_j\|_{L^2(\Omega)}^2}$$

$$+ \frac{k^2}{2}\|\dddot{P}\|_{L^1(0,t_\ell;L^2(\Omega))}\max_{j=1}^{\ell-1}\|\Sigma_j - \langle\Sigma_{hk}\rangle_j\|_{L^2(\Omega)}$$

$$+ k^2\|\dddot{P}\|_{L^\infty(0,t_\ell;L^2(\Omega))}\left(\|(\Sigma - \Sigma_{hk})_\ell^-\|_{L^2(\Omega)} + \|(\Sigma - \Sigma_{hk})_0^+\|_{L^2(\Omega)}\right).$$

Proof Some elementary algebra shows, for $\ell = 1, \ldots, J$ (note $[\tilde{\Sigma}]_{j-1} = 0$), that

$$\sum_{j=1}^\ell \int_\Omega [\dot{P}]_{j-1}^j \star [\tilde{\Sigma} - \Sigma_{hk}]_{j-1}^j\, dx$$

$$= \frac{1}{2}\sum_{j=1}^{\ell-1}\int_\Omega [\Sigma_{hk}]_j \star (\dot{P}_{j+1} - \dot{P}_{j-1})dx$$

$$+ \sum_{j=1}^{\ell-1} \int_\Omega (\Sigma_j - \langle \Sigma_{hk} \rangle_j) \star (-\dot{P}_{j+1} + 2\dot{P}_j - \dot{P}_{j-1}) dx$$

$$+ \int_\Omega (\dot{P}_\ell - \dot{P}_{\ell-1}) \star (\Sigma_\ell - \Sigma_{hk,\ell}^-) dx - \int_\Omega (\dot{P}_1 - \dot{P}_0) \star \left(\Sigma_0 - \Sigma_{hk,0}^+\right) dx.$$

The integral $\dot{P}_{j+1} - \dot{P}_{j-1} = \int_{t_{j-1}}^{t_{j+1}} \ddot{P} dt$ and the Cauchy inequality lead to

$$\frac{1}{2} \sum_{j=1}^{\ell-1} \left| \int_\Omega [\Sigma_{hk}]_j \star (\dot{P}_{j+1} - \dot{P}_{j-1}) dx \right| \leq k^{1/2} \|\ddot{P}\|_{L^2(0,t_\ell;L^2(\Omega))} \sqrt{\sum_{j=1}^{\ell-1} \|[\Sigma_{hk}]_j\|_{L^2(\Omega)}^2}.$$

$$(4.3)$$

The arguments at the end of the proof of the previous lemma show that

$$\sum_{j=1}^{\ell-1} \left| \int_\Omega (\Sigma_j - \langle \Sigma_{hk} \rangle_j) \star (-\dot{P}_{j+1} + 2\dot{P}_j - \dot{P}_{j-1}) dx \right|$$

$$\leq \frac{k}{2} \|\dddot{P}\|_{L^1(0,t_\ell;L^2(\Omega))} \max_{j=1}^{\ell-1} \|\Sigma_j - \langle \Sigma_{hk} \rangle_j\|_{L^2(\Omega)}.$$

The initial and final remaining terms are bounded as

$$\left| \int_\Omega (\dot{P}_1 - \dot{P}_0) \star (\Sigma_0 - \Sigma_{hk,0}^+) dx \right| + \left| \int_\Omega (\dot{P}_\ell - \dot{P}_{\ell-1}) \star (\Sigma_\ell - \Sigma_{hk,\ell}^-) dx \right|$$

$$\leq k \|\ddot{P}\|_{L^\infty(0,t_\ell;L^2(\Omega))} \left(\|(\Sigma - \Sigma_{hk})_\ell^-\|_{L^2(\Omega)} + \|(\Sigma - \Sigma_{hk})_0^+\|_{L^2(\Omega)} \right).$$

The combination of all the above estimates concludes the proof. $\qquad\square$

4.3 Combination of variational inequalities

The Prandtl-Reuß flow rule has been approximated in Lemma 4.2. The discrete counterpart takes $T_{hk} = \Sigma_{hk}^*$ in (3.3). The sum of the two inequalities is summed up for $j = 1, \ldots, \ell, \ell \leq J$, and with the estimate of Lemma 4.3 leads to the point of departure

$$\int_{\Omega \times (0,t_\ell)} (P_{hk,\tau} - \tilde{\dot{P}}) \star (\widetilde{\Sigma} - \Sigma_{hk}) \, dQ + \sum_{j=1}^\ell \int_\Omega [P_{hk}]_{j-1} \star (\widetilde{\Sigma} - \Sigma_{hk})_{j-1}^+ dx$$

$$\leq RHS_1(\ell) := \frac{k^{3/2}}{4} \|\ddot{P}\|_{L^2(0,t_\ell;L^2(\Omega))} \sqrt{\sum_{j=1}^{\ell-1} \|[\Sigma_{hk}]_j\|_{L^2(\Omega)}^2}$$

$$+ \frac{k^2}{8} \|\dddot{P}\|_{L^1(0,t_\ell;L^2(\Omega))} \left(\max_{j=1}^\ell \|(\widetilde{\Sigma} - \Sigma_{hk})(t_{j-1/2})\|_{L^2(\Omega)} + \max_{j=1}^{\ell-1} \|\Sigma_j - \langle \Sigma_{hk} \rangle_j\|_{L^2(\Omega)} \right)$$

$$+ \frac{k^2}{4} \|\ddot{P}\|_{L^\infty(0,t_\ell;L^2(\Omega))} \left(\|(\Sigma - \Sigma_{hk})_\ell^-\|_{L^2(\Omega)} + \|(\Sigma - \Sigma_{hk})_0^+\|_{L^2(\Omega)} \right).$$

Recall that $P = (\varepsilon(u), 0) - \mathbb{A}\Sigma$, $P_{hk} = (\varepsilon(u_{hk}), 0) - \mathbb{A}\Sigma_{hk}$, and with $\widetilde{\sigma}$ and $\widetilde{u}$ denoting the linear interpolants of σ and u in time, set

$$\Delta := \widetilde{\Sigma} - \Sigma_{hk}, \qquad \delta := \widetilde{\sigma} - \sigma_{hk}, \quad \text{and} \quad e := \widetilde{u} - u_{hk}.$$

Then the left-hand side with the piecewise time derivative $\Delta_\tau := \partial\Delta/\partial t$ reads

$$\int_{\Omega\times(0,t_\ell)} \Delta_\tau \star \mathbb{A}\Delta \, dQ + \sum_{j=1}^{\ell} \int_\Omega [\Delta]_{j-1} \star \mathbb{A}\Delta_{j-1}^+ dx$$

$$- \int_{\Omega\times(0,t_\ell)} \delta : \varepsilon(e_\tau) dQ + \sum_{j=1}^{\ell} \int_\Omega [\varepsilon(u_{hk})]_{j-1} : \delta_{j-1}^+ dx.$$

With the homogeneous initial conditions and the convention $\Delta_0^- := 0$, the term $\Delta_0^+ = [\Delta]_0$ is written as the summand for the index zero below. Recall that $\|\!|\bullet|\!\| := \|\mathbb{A}^{1/2}\bullet\|_L$ abbreviate the norm in L with the weight $\mathbb{A}$. Then the first two of the preceding terms with a telescoping sum combine to

$$LHS(\ell) = \frac{1}{2}\|\!|\Delta_\ell^-|\!\|^2 + \frac{1}{2}\sum_{j=0}^{\ell-1}\|\!|[\Delta]_j|\!\|^2.$$

The remaining two terms with the displacement contributions will be placed in the upper bound and lead to

$$LHS(\ell) - RHS_1(\ell) \le RHS_2(\ell) := \int_{\Omega\times(0,t_\ell)} \delta : \varepsilon(e_\tau) dQ$$

$$- \sum_{j=1}^{\ell} \int_\Omega [\varepsilon(u_{hk})]_{j-1} : \delta_{j-1}^+ dx. \quad (4.4)$$

To bound $RHS_2(\ell)$, let G denote the Galerkin projection (optimal in the space of H^1 functions) as a linear and bounded map $G : V \to V_h$ and split the volume contribution with $\varepsilon(e_\tau)$ into

$$\int_{(0,t_\ell)\times\Omega} \delta : \varepsilon(e_\tau) dQ = \int_{(0,t_\ell)\times\Omega} \delta : \varepsilon(\dot{\widetilde{u}} - G(\dot{\widetilde{u}})) dQ + \int_{(0,t_\ell)\times\Omega} \delta : \varepsilon(G(e_\tau)) dQ.$$

The first term on the right-hand side is bounded by $C_1\|\delta\|_{L^\infty(L^2)} \|h_{\mathcal{T}} D^2\dot{u}\|_{L^1(L^2)}$ with the Hessian D^2 of all spatial partial derivatives of order 2 and an interpolation constant C_1, which depends exclusively on the shape-regularity of the triangulation $\mathcal{T}$. Set $v_{hk} := G(e_\tau) \in P_0(\mathcal{I}; V_h)$ and observe that $\varepsilon(v_{hk})$ is constant in time (and piecewise constant in space) in each Q_j. The contribution of Q_j to the second term reads

$$\int_{Q_j} \delta : \varepsilon(v_{hk}) dQ = \frac{k}{2} \int_\Omega \delta_j^- : \varepsilon(v_{hk}) dx + \frac{k}{2} \int_\Omega \delta_{j-1}^+ : \varepsilon(v_{hk}) dx.$$

The last term in (4.4) can be rewritten with $[\varepsilon(u_{hk})]_{j-1} = -[\varepsilon(G(e))]_{j-1}$ and leads to

$$RHS_2(\ell) \le C_1 \|\delta\|_{L^\infty(0,t_\ell;L^2)} \|h_{\mathcal{T}} D^2 \dot{u}\|_{L^1(L^2)} + \sum_{j=1}^\ell I(j) \quad \text{with the abbreviation}$$

$$I(j) := \int_\Omega \left(\frac{k}{2} \delta_j^- : \varepsilon(v_{hk}|_{I_j}) + \frac{k}{2} \delta_{j-1}^+ : \varepsilon(v_{hk}|_{I_j}) + \delta_{j-1}^+ : [\varepsilon(G(e))]_{j-1} \right) dx.$$
(4.5)

4.4 Finish of the proof of (a)

This subsection finalizes the proof in Example 3.1.a, which is special in that $I(j) = 0$ in (4.5). In fact, the equilibrium (2.3) and the discrete equilibrium (3.4) (for the one-sided limits) and, in the case of linear interpolation of Example 3.1.b, $f_k(t_j) = \tilde{f}(t_j) = f(t_j) =: f_j$ etc. show that each summand in $I(j)$ of (4.5) vanishes, e.g.,

$$\int_\Omega \delta_{j-1}^+ : \varepsilon(v_{hk}) dx = \int_\Omega (f_{j-1} - f_{k,j-1}^+) \cdot v_{hk} dx + \int_{\Gamma_N} (g_{j-1} - g_{k,j-1}^+) \cdot v_{hk} ds = 0.$$

This and (4.4)–(4.5) read $LHS(\ell) \le RHS_1(\ell) + C_1 \|\delta\|_{L^\infty(0,t_\ell;L^2)} \|h_{\mathcal{T}} D^2 \dot{u}\|_{L^1(L^2)}$. The fourth-order tensors behind $\mathbb{A}$ and its inverses are bounded and so $\|\Delta_j^\pm\|_L \le C_2 \|\mathbb{A}^{1/2} \Delta_j^\pm\|_L$ for any j etc. With the maximum $M := \max\{|||\Delta_0^+|||, |||\Delta_1^-|||, |||\Delta_1^+|||, \ldots, |||\Delta_J^-|||\}$, this leads to

$$2|||\Delta_\ell^-|||^2 + \sum_{j=0}^{\ell-1} |||[\Delta]_j|||^2$$

$$\le C_2 \left(2k^2 \|\ddot{P}\|_{L^\infty(L^2)} + k^2 \|\ddot{P}\|_{L^1(L^2)} + C_1 \|h_{\mathcal{T}} D^2 \dot{u}\|_{L^1(L^2)} \right) M$$

$$+ \frac{1}{2} \sum_{j=1}^{\ell-1} |||[\Delta]_j|||^2 + \frac{k^3}{2} C_2^2 \|\ddot{P}\|_{L^2(L^2)}^2 \quad \text{for any } \ell = 0, \ldots, J.$$

Consequently,

$$LHS := \max_{\ell=1}^J |||\Delta_\ell^-|||^2 + \frac{1}{2} \sum_{j=0}^{J-1} |||[\Delta]_j|||^2$$

$$\le M^2/8 + C_3 \left(h^2 \|\dot{u}\|_{L^2(H^2)}^2 + k^3 \|\ddot{P}\|_{L^2(L^2)}^2 + k^4 \|\ddot{P}\|_{L^1(L^2)}^2 + k^4 \|\ddot{P}\|_{L^\infty(L^2)}^2 \right).$$

A triangle inequality shows that $M^2/4 \leq LHS$ and that concludes the proof for Example 3.1.a. □

4.5 Displacement control through hardening

The presence of hardening acts as a regularization and allows for solutions with displacements in H^1. It follows with Proposition 2.1 for the generalized stress $\Sigma = (\sigma, -A\alpha, -B\beta)$ and the plastic deformation $P = (\varepsilon(u) - \mathbb{C}^{-1}\sigma, \alpha, \beta)$ that $p := \varepsilon(u) - \mathbb{C}^{-1}\sigma = -\beta$. Lemma 3.3 shows shat all those aforementioned identities hold a.e. in Q for the discrete quantities (with lower index hk) as well. This, but not more, is inherited by the kinematic hardening in the discrete situation with a discrete flow-rule. The error Δ in the generalized stresses is hence equal to

$$\Delta = \widetilde{\Sigma} - \Sigma_{hk} = \left(\widetilde{\sigma} - \sigma_{hk}, -A(\widetilde{\alpha} - \alpha_{hk}), B(\widetilde{p} - p_{hk})\right) \quad \text{a.e. in } Q.$$

This and the additive split $p = \varepsilon(u) - \mathbb{C}^{-1}\sigma$ and its discrete version (in terms of the Frobenius matrix norm $|\bullet|$ inherited from the scalar product :) lead a.e. in Q to

$$|\varepsilon(\widetilde{u} - u_{hk})|^2 = |\widetilde{p} - p_{hk} + \mathbb{C}^{-1}(\widetilde{\sigma} - \sigma_{hk})|^2 \leq 2|\widetilde{p} - p_{hk}|^2$$
$$+ 2|\mathbb{C}^{-1}(\widetilde{\sigma} - \sigma_{hk})|^2 \leq C_4 \, \Delta \star \mathbb{A} \Delta$$

with a material-dependent constant $C_4 = C_4(\mathbb{C}^{-1}, B^{-1})$ with $\mathbb{A} := \text{diag}(\mathbb{C}^{-1}, \mathbb{H}^{-1})$ and $\mathbb{H} = \text{diag}(A, B)$. This and the Korn inequality lead to displacement control in the sense that $\|e(t)\|_{H^1(\Omega)} \leq C_5 \|\|\Delta(t)\|\|$ holds for a.e. time t (and for the one-sided limits at $t_j^{\pm}$) with an universal constant C_5. Since the Galerkin projection is H^1 stable, this leads to

$$k\|v_{hk}\|_{L^\infty(V)} + \|G(e)\|_{L^\infty(V)} \leq C_6 \max_{0 \leq t \leq t_J} \|\|\Delta(t)\|\| = C_6 M. \tag{4.6}$$

4.6 Finish of the proof (b)

The proof requires the hardening control (4.6) and exclusively utilizes the approximation property of Remark 3.2 to control the extra terms $I(j)$ in (4.5) for Example 3.1.b. The equilibrium (2.3) and Lemma 3.1 show that

$$I(j) = \int_\Omega \left(\frac{k}{2}(f - f_k)_j^- \cdot v_{hk}|_{I_j} + \frac{k}{2}(f - f_k)_{j-1}^+ \cdot v_{hk}|_{I_j} + (f - f_k)_{j-1}^+ \cdot [G(e)]_{j-1} \right) dx$$

plus corresponding terms with $g - g_k$ and their integrals along Γ_N. Remark 3.2 shows

$$I(j) \leq k \, C_I \, C_6 \|(\ddot{f}, \ddot{g})\|_{L^1(I_j; V^* \times V^*)} \, M.$$

This bounds the sum in (4.5) and the arguments of Sect. 4.4 conclude the proof. □

4.7 Discussion and refined analysis

The optimal convergence rate $O(h + k^2)$ for the dG(1) scheme is *not* established in this paper although this is understood to hold for the Crank–Nicolson scheme [2]. This subsection discusses the sub-optimality of the mathematical arguments and enlightens the two convergence rates in Theorem 4.1.

4.7.1 Consequences of the flow rule I

The introduction explains that the textbook analysis of the dG schemes cannot be performed and motivates the split of the error with one contribution Θ and a remainder on the right-hand side in (1.2) that needs to be bounded from above. Since $\widetilde{\Sigma}$ is admissible, the flow rule (1.1) states that the integrand $\dot{P} \star (\Sigma - \widetilde{\Sigma}) \geq 0$ is non-negative pointwise a.e. in the time-space cylinder Q_j. Assuming that the continuous variables are all smooth, the Simpson quadrature rule in the time interval (t_{j-1}, t_j) gives

$$\int_{Q_j} \dot{P} \star (\Sigma - \widetilde{\Sigma}) dQ = \int_{Q_j} |\dot{P} \star (\Sigma - \widetilde{\Sigma})| dQ$$
$$= 2k/3 \int_{\Omega} \dot{P}_{j-1/2} \star (\Sigma - \widetilde{\Sigma})_{j-1/2} dx + O(k^5).$$

The central difference scheme leads to $2(\Sigma - \widetilde{\Sigma})_{j-1/2} = \Sigma_j - 2\Sigma_{j-1/2} + \Sigma_{j-1} = k^2 \ddot{\Sigma}_{j-1/2}$ up to the error $O(k^4)$. The conclusion is that $\int_{Q_j} \dot{P} \star (\Sigma - \widetilde{\Sigma}) dQ = O(k^3)$ and the sum over all $j = 1, \ldots J$ is $O(k^2)$. This shows that there is only linear convergence in the end and so, in order to deduce a better convergence rate, this paper has adopted a pointwise test of the continuous flow rule.

4.7.2 Consequences of the flow rule II

The comments of this subsections refine the main result of Theorem 4.1 in that there is an additional term

$$\sum_{j=1}^{J} k \int_{\Omega} |\dot{P}_{j-1} \star [\Sigma_{hk}]_{j-1}| dx \tag{4.7}$$

in the lower bound. The above analysis tests the continuous flow rule (1.1) at the endpoints t_{j-1} and t_j of a typical time interval with a particular choice of an admissible $\overline{T}$ from the discrete generalized stress Σ_{hk}. Since the latter is discontinuous in general at t_{j-1}, the choice of $\overline{T}$ in the proof of Lemma 4.2 will be revisited. The continuous flow rule (for continuous quantities) at the fixed time t_{j-1} holds pointwise in space. Rather than $\Sigma_{hk,j-1}^+$ we choose the admissible stress $\Sigma_{hk,j-1}^-$ and that leads to

$$\int_\Omega \dot{P}_{j-1} \star (\Sigma^+_{hk,j-1} - \Sigma_{j-1})dx - \int_\Omega \dot{P}_{j-1} \star [\Sigma_{hk}]_{j-1}dx$$

$$= \int_\Omega \dot{P}_{j-1} \star (\Sigma^-_{hk,j-1} - \Sigma_{j-1})dx \le 0.$$

The use of the first term in the above displayed inequality leads to the analysis as above but then the extra jump term $-k\int_\Omega \dot{P}_{j-1} \star [\Sigma_{hk}]_{j-1}dx$ appears in the lower bound of Lemma 4.2. The key observation is that the two choices for $\overline{T}$, namely $\Sigma^+_{hk,j-1}$ or $\Sigma^-_{hk,j-1}$, in (1.1) can be made separately for a.e. point $x \in \Omega$. Hence if the term $\dot{P}_{j-1}(x) \star [\Sigma_{hk}]_{j-1}(x)$ is positive, we select $\overline{T} = \Sigma^+_{hk,j-1}(x)$ in (1.1) and otherwise $\overline{T} = \Sigma^-_{hk,j-1}(x)$. The integral of all those (measurable) selections generates the additional term $k\int_\Omega \max\{0, -\dot{P}_{j-1} \star [\Sigma_{hk}]_{j-1}\}dx$ in the lower bound of Lemma 4.2. The analogous selection procedure is possible at the point t_j (for $j < J$) and then leads to the additional term $k\int_\Omega \max\{0, \dot{P}_j \star [\Sigma_{hk}]_j\}dx$ in the lower bound of Lemma 4.2. The refined test leads to an additional term

$$k\int_\Omega \max\{0, -\dot{P}_{j-1} \star [\Sigma_{hk}]_{j-1}\}dx + k\int_\Omega \max\{0, \dot{P}_j \star [\Sigma_{hk}]_j\}dx$$

in the lower bound of Lemma 4.2. The sum of those terms in the flow of the arguments in the above proof results in the additional term (4.7) in the lower bound of the two assertions in Theorem 4.1.

4.7.3 Comments on Theorem 4.1.a

In the absence of further arguments and the presence of plasticity with $\|\dot{P}\|_{L^2(L^2)} > 0$, the extra term (4.7) in the lower bound of the Theorem 4.1.a may be expected to be of the same order as $\sqrt{\sum_{j=1}^J k\|[\Sigma_{hk}]_{j-1}\|_L^2}$; at least the non-negative sign of the integrand gives no reason to believe that a strengthened Cauchy inequality can be applied. However, if not only (4.7) but also the latter term were controlled by the right-hand side $O(h^2 + k^3)$ of Theorem 4.1.a, then $\sum_{j=1}^J \|[\Sigma_{hk}]_{j-1}\|_L^2 = O(h^4/k + k^5)$. With the practical choice $h = k^{3/2}$ of Remark 4.1, this gives $O(k^5)$ and hence a superconvergence result.

This line of thoughts may (i) indicate why any improvement of the theoretical convergence rate beyond the statement of Theorem 4.1 is tricky if not impossible, and (ii) warn against naive attempts to utilize the computed jumps $\|[\Sigma_{hk}]_{j-1}\|_L$ in an adaptive time-stepping (suggested in the engineering literature).

4.7.4 Comments on Theorem 4.1.b

At the first glance it might be surprising that the natural choice with (f_k, g_k) as L^2 projections in the discrete equilibrium (3.1) of Example 3.1.b lead to seemingly sub-optimal rates as highlighted in Remark 4.1. Moreover, the numerical benchmark in the following section adopts the natural choice and shows the higher convergence rates in time within the computational range (and the linear interpolant looks identical).

On the right of the figure:

$E = 70000$

$\nu = 0.33$

$\mu = \dfrac{E}{2(1+\nu)}$

$\lambda = \dfrac{\nu E}{(1+\nu)(1-2\nu)}$

$\sigma_y = 243$

$\mathbb{H}_2 = 1$

$g_1(r, \phi, t) = 240 \sin(2\pi t)(\cos \phi, \sin \phi)$

$g_2(r, \phi, t) = -60 \sin(2\pi t)(\cos \phi, \sin \phi)$
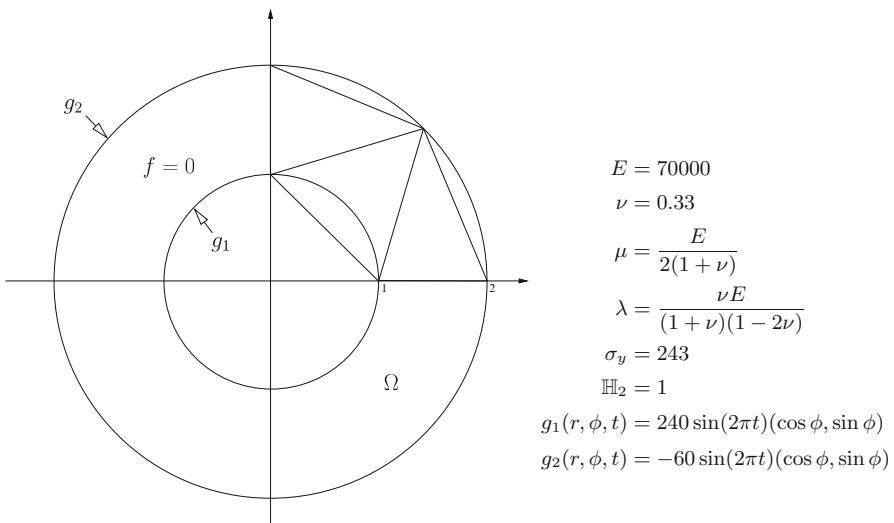
Fig. 3 Geometry and parameters for the computational benchmark

On the other hand, when the discrete loads are different from the linear interpolation, the approximation property (3.5) is supposed to be sharp and shows that $\|[(f - f_k, g - g_k)]_{j-1}\|_{L^2(\Omega \times \Gamma_N)} = O(k^2)$. This and the pointwise equilibrium of Lemma 3.1 enforces a jump $\left\|\left\|[\Sigma_{hk}]_{j-1}\right\|\right\| = O(k^2)$ and the sum over all those jumps gives $\sqrt{\sum_{j=1}^{J} \left\|\left\|[\sigma_{hk}]_{j-1}\right\|\right\|^2} = O(k^{3/2})$.

The discussion in Sect. 4.7.3 about the extra term (4.7) and the same convergence rate for the upper bound suggests that the extended lower bound in Theorem 4.1.b is $O(k^2)$ and can not be improved.

5 Numerical experiments

The numerical experiments for the elastoplastic time evolution that lead to Fig. 1 are considered for the axisymmetric ring Ω of Fig. 3. While the volume force f vanishes, the applied surface loads g_1 and g_2 correspond to a pure Neumann problem with a known analytical solution given in closed form in [1,2] for kinematic hardening ($H_1 = A = 0$), where further details can be found.

The algorithmic details of the implementation are included in [3] and lead to the stress evaluation of Fig. 1. Some snapshots of the elastoplastic evolution computed with the backward Euler (bE), Crank–Nicolson (CN), dG(0), and dG(1) are displayed in Fig. 4. To discuss the convergence rates, the relative error

$$e_Q^2 := \frac{\int_0^1 \int_\Omega \left(\|\mathbb{C}^{-1/2}(\sigma(x, t) - \sigma_{hk}(x, t))\|^2 + \|\mathbb{H}^{-1/2}(\chi(x, t) - \chi_{hk}(x, t))\|^2 \right) dx \, dt}{\int_0^1 \int_\Omega \left(\|\mathbb{C}^{-1/2}\sigma(x, t)\|^2 + \|\mathbb{H}^{-1/2}\chi(x, t)\|^2 \right) dx \, dt},$$

(5.1)

(a) bE at $t = \frac{1}{3}$　　**(b)** CN at $t = \frac{1}{3}$　　**(c)** dG(0) at $t = \frac{1}{3}$　　**(d)** dG(1) at $t = \frac{1}{3}$

(e) bE at $t = \frac{61}{81}$　　**(f)** CN at $t = \frac{61}{81}$　　**(g)** dG(0) at $t = \frac{61}{81}$　　**(h)** dG(1) at $t = \frac{61}{81}$
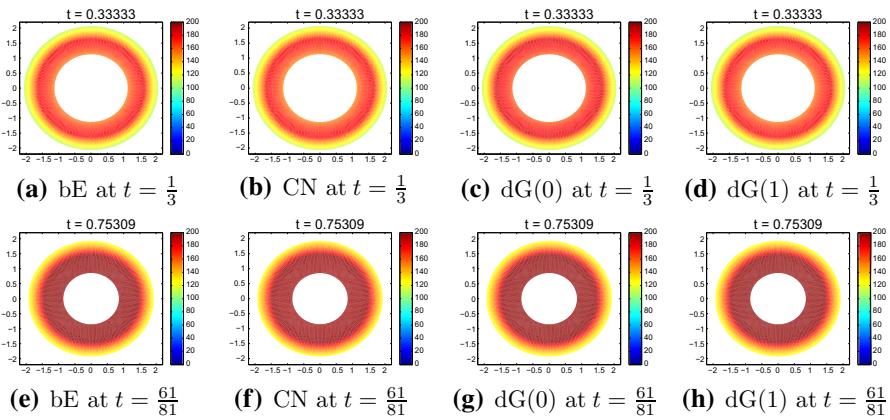
Fig. 4 Elastoplastic evolution for time step size, $k = 3^{-4}$ at various times

and the stress error

$$e_\Omega^2 := \int_\Omega \|\mathbb{C}^{-1/2}(\sigma(x, 1) - \sigma_{hk}(x, 1))\|^2 \, dx \tag{5.2}$$

are computed at $t = T = 1$. The convergence history for the error e_Q (left) and e_Ω (right) is displayed in Fig. 5 for dG(0) and dG(1). Several uniform spatial discretizations are performed and give rise to different curves. Each curve displays the error as a function of the (uniform) time-step size $k = 3^{-n}$ for $n = 2, 3, 4, 5$ with both axes with a logarithmic scaling. Notice that the convergence in time is from right to left. For a fine mesh, the empirical convergence rate in time is close to 1 for dG(0) and in fact better than 1.7 for dG(1). The observed value 2 (resp. 1) for dG(1) (resp. dG(0)) and e_Ω is not a proof for a higher order of convergence, but an indication that the pre-asymptotic range is very large. This might be an overall impression for dG(1): The time discretisation error is so small that the spatial error dominates the convergence.

In this (non-academic) mechanical problem, the solution is globally Lipschitz in space-time and piecewise smooth (cf. [2] for the some explicit formulas): There are up to two rings in space, which evolve in time, where at $x \in \Omega$ with $|x| = R(t)$ the plastic rate $\dot{P}(x, t)$ has a finite jump with respect to t. Hence the high regularity assumptions of our theoretical results are not met. Nevertheless, we observe the predicted maximal convergence rates and give a heuristic explanation in the sequel. The leading term (4.3) in the analysis has to be modified in a time interval I_j for all x with $|x| = R(t)$ for some t with $t_{j-1} \leq t \leq t_{j+1}$, but remains valid elsewhere. For those points x, written $x \in \omega(t_{j-1}, t_{j+1}) \subset \Omega$, the difference $\dot{p}_{j+1}(x) - \dot{p}_{j-1}(x)$ is the L^1 time integral of the piecewise second derivative $\ddot{p}(x, t)$ with respect to time t plus a finite jump of $\dot{p}(x, t)$ at the critical time t. The Cauchy inequality implies the extra term $k^{1/2}$ in the analysis of the L^2 norm, but this is not available for the jump terms. Since the jumps are bounded in $L^\infty(L^\infty)$ and the area $\omega(t_{j-1}, t_{j+1})$ has a volume $O(k)$ in space, the extra term (with $O(k^{-1})$ time steps) reduces the overall convergence rate to $O(k + h)$. Although those arguments could be made rigorous they do not explain
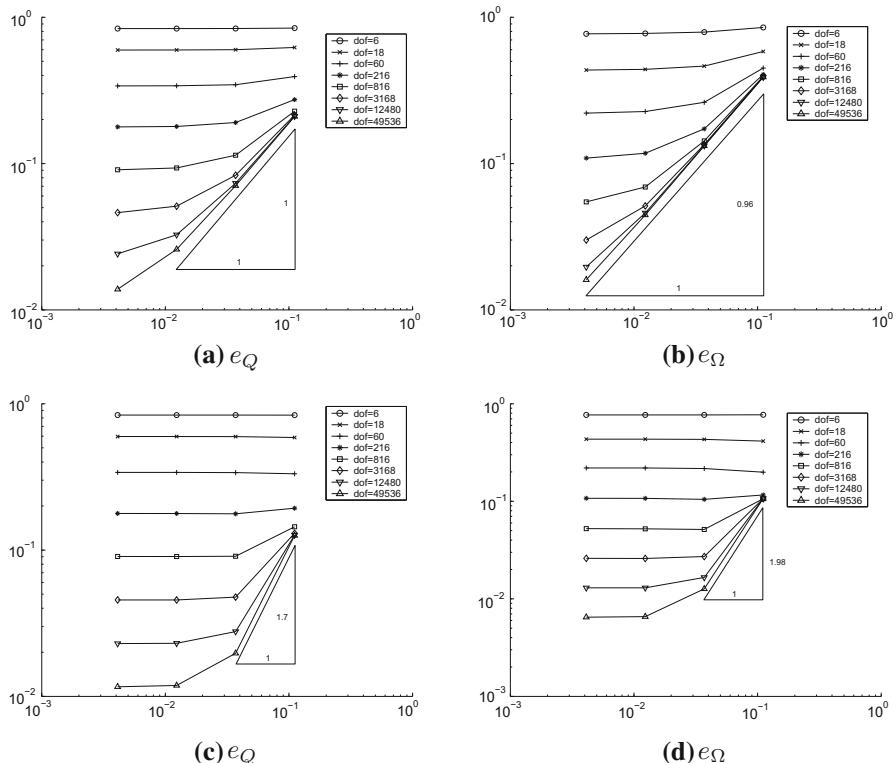
Fig. 5 Convergence for the error e_Q (left) and e_Ω (right) from (5.1)–(5.2) as a function of the time-step size $k = 3^{-2}, \ldots, 3^{-5}$ with dG(0) in **a**–**b** and dG(1) in **c**–**d** for various fixed uniform space discretizations with dof $= 6, \ldots, 49536$ degrees of freedom for the displacements

the better empirical convergence rate for dG(1) compared to dG(0). However, it is the nature of elastoplasticity to diminish the stress peaks compared to a linear elastic material behaviour and there is no reason to believe that the stresses concentrate near the interfaces at $|x| = R(t)$. In the optimistic (although heuristic) hypothesis that the discrete stress differences in (4.3) do not concentrate, one may hope that the L^2 norm of contribution $[\Sigma_{hk}]_j$ over $\omega(t_{j-1}, t_{j+1})$ in the extra term is much smaller than the full L^2 norm over Ω. For instance, if the term $[\Sigma_{hk}]_j$ is equidistributed in space, the relevant term over $\omega(t_{j-1}, t_{j+1})$ scales like $O(k^{1/2})$ times its full L^2 norm over Ω. That extra scaling factor could explain the observed improved convergence rate in time.

Hock for a collaboration at the early stage of this research and the anonymous referees for valuable remarks that improved the presentation.

References

1. Alberty, J.: Zeitdiskretisierungsverfahren für elastoplastische Probleme der Kontinuumsmechanik. Ph.D. thesis, University of Kiel, FRG, Tenea Verlag (2001)
2. Alberty, J., Carstensen, C.: Numerical analysis of time-depending primal elastoplastic with hardening. SIAM J. Numer. Anal. **37**, 1271–1294 (2000)
3. Alberty, J., Carstensen, C.: Discontinuous Galerkin time discretization in elastoplasticity: motivation, numerical algorithms and applications. Comp. Methods Appl. Mech. Eng. **191**, 4949–4968 (2002)
4. Alberty, J., Carstensen, C., Zarrabi, D.: Adaptive numerical analysis in primal elastoplasticity with hardening. Comput. Methods Appl. Mech. Eng. **171**, 175–204 (1999)
5. Bartels, Sören, Mielke, Alexander, Roubíček, Tomáš: Quasi-static small-strain plasticity in the limit of vanishing hardening and its numerical approximation. SIAM J. Numer. Anal. **50**, 951–976 (2012)
6. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods, 15, Texts in Applied Mathematics. Springer, New York (1994)
7. Carstensen, C.: Domain decomposition for a non-smooth convex minimisation problem and its application to plasticity. NLAA **4**, 1–13 (1997)
8. Carstensen, C.: Numerical analysis of the primal problem of elastoplasticity with hardening. Numer. Math. **82**, 577–597 (1999)
9. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam (1978)
10. Ekeland, I., Temam, R.: Convex Analysis and Variational Problems. North-Holland, Amsterdam (1976)
11. Evans, L.C.: Partial Differential Equations, 19, Graduate Studies in Mathematics. American Mathematical Society, Providence (1998)
12. Han, W., Reddy, B.D.: Plasticity, 2nd edn. IAM, Springer, New York (2013)
13. Johnson, C.: Existence theorems for plasticity problems. J. Math. pures et appl. **55**, 431–444 (1976)
14. Johnson, C.: On plasticity with hardening. J. Math. Anal. Appl. **62**, 325–336 (1978)
15. Rindler, Filip, Schwarzacher, Sebastian, Süli, Endre: Regularity and approximation of strong solutions to rate-independent systems. Math. Models Methods Appl. Sci. **27**, 2511–2556 (2017)
16. Seregin, G.A.: On the regularity of the minimizers of some variational problems of plasticity theory. St. Petersb. Math. J. **4**, 989–1020 (1993)
17. Simo, J.C., Hughes, T.J.R.: Computational Inelasticity. Springer, New York (1998)
18. Suquet, P.M.: Discontinuities and plasticity, Nonsmooth Mechanics and Applications, CISM Courses, 302, pp. 279–341. Springer, New York (1988)
19. Temam, R.: Mathematical Problems in Plasticity. Gauthier-Villars, Paris (1985)
20. Thomee, V.: Galerkin Finite Element Methods for Parabolic Problems. Springer, Berlin (2006)
21. Zeidler E.: Nonlinear Functional Analysis and Its Applications iii and iv. Springer, New York (1985,1988)