

A DUALITY-BASED PATH-FOLLOWING SEMISMOOTH NEWTON METHOD FOR ELASTO-PLASTIC CONTACT PROBLEMS

M. HINTERMÜLLER AND S. RÖSEL

ABSTRACT. A Fenchel dualization scheme for the one-step time-discretized contact problem of quasi-static elasto-plasticity with kinematic or isotropic hardening is considered. The associated path is induced by a combined Moreau-Yosida / Tichonov regularization of the dual problem. The sequence of solutions to the regularized problems is shown to converge strongly to the optimal displacement-stress-strain triple of the original elasto-plastic contact problem in the space-continuous setting. This property relies on the density of the intersection of certain convex sets which is shown as well. It is also argued that the mappings associated with the resulting problems are Newton- or slantly differentiable. Consequently, each regularized subsystem can be solved mesh-independently at a local superlinear rate of convergence. For efficiency purposes, an inexact path-following approach is proposed and a numerical validation of the theoretical results is given.

1. INTRODUCTION

In this paper we consider the quasi-static elasto-plasticity model with an associative flow law (sometimes called Prandtl-Reuss normality law) and von Mises hardening under the small strain assumption set forth in [21]. First investigations of the elasto-plastic problem from a mathematical point of view can be found in [15, 31], where [31] includes existence for the fully continuous case. Numerical analysis of the semi-discrete and fully-discrete versions can be found, for example, in [2, 21]. Appropriate discretization schemes for plasticity problems with hardening have been investigated extensively in the recent past. Here we only mention [3, 9, 10] for adaptive finite element methods. Concerning numerical solution methods, we refer to the multigrid approach in [45], various generalized Newton methods in finite dimensions [12, 19, 40, 45, 46], the algorithm in [42] as well as interior point strategies, cf. e.g. [35].

A general introduction to elastic contact problems including corresponding numerical approaches can be found in the monographs [29, 39], and multigrid methods for elastic contact are analyzed, e.g., in [33] and [34, 36], where the latter references are devoted to two-body contact. For the treatment of elastic friction problems we refer to [13, 36] as well as to the efficient active set algorithm proposed in [30], to mention only two. Subspace correction methods for variational inequalities of

2010 *Mathematics Subject Classification.* 49M15, 49M29, 74C05, 74S05.

Key words and phrases. Elasto-plastic contact, Variational inequality of the 2nd kind, Fenchel duality, Moreau-Yosida/Tichonov regularization, Path-following, Semismooth Newton .

The authors gratefully acknowledge the support of the DFG research center MATHEON within project C28 "Optimal control of phase separation phenomena" as well as the START program Y305 "Interfaces and Free Boundaries" funded by the Austrian Science Fund FWF.

the second kind with application to frictional contact have been investigated in [6]. In [12, 20] plastic material behavior is incorporated in addition to the contact constraints. In the latter references the elasto-plastic friction problem is reformulated utilizing a nonlinear complementarity problem (NCP) function yielding a nonsmooth system which can be solved efficiently by applying a generalized Newton method in a discrete framework.

While some attention has been paid to infinite-dimensional methods in linear elasticity with (frictional) contact [37, 43], elasto-plastic problems are still less researched. Among the few available references we mention [8] for domain decomposition methods leading to a linear rate of convergence. The approach to plasticity problems without contact constraints in [19], however, turns out to be problematic as far as function space convergence of the employed semismooth Newton (SSN) solver is concerned. In fact, due to the lack of a sufficient norm gap between domain and image space of the mapping involved in the underlying nonsmooth system, generalized differentiability in the sense of [24] does not hold true. The resulting lack of a well-defined infinite-dimensional generalized Newton iteration usually results in a mesh-dependent solver.

In the present paper, we introduce a path-following semismooth Newton method which admits a rigorous convergence analysis in the continuous setting. For this purpose, we study a regularized version of the Fenchel-dual problem of the underlying elasto-plastic contact problem with the regularization parameter inducing a dual path to the solution of the original problem. Each path-problem can be solved at a local superlinear rate and in a mesh-independent way upon discretization.

2. PROBLEM FORMULATION

The starting point of our analysis is a time-discretized elasto-plastic contact problem in the displacement u , the plastic strain p and the hardening parameter α (sometimes called conjugate force) which model the evolution of a body subject to given applied forces. Further, in the elasto-plasticity literature the pair (σ, α) is called generalized stress. The body is represented by a bounded domain $\Omega \subset \mathbb{R}^N$, $N = 2, 3$, with $N^{0,1}$ -property [47] and it adheres to a fixed part $\Gamma_d \subset \partial\Omega$ with positive surface measure. We further denote by $\Gamma_n \subset \partial\Omega \setminus \Gamma_d$ some relatively open part of the boundary where a given surface load $g \in L^2(\Gamma_n)$ is applied. A given volume force density is denoted by $f \in L^2(\Omega)$. The elasto-plastic behavior at a material point $x \in \Omega$ is determined by the von Mises yield criterion leading to a dissipation functional which typically is nonsmooth, lower semicontinuous (l.s.c.) and convex [21].

Often, the displacement of the body is restricted by a given rigid obstacle giving rise to an elasto-plastic contact problem. Therefore we fix a set $\Gamma_c \subset \partial\Omega$ which potentially contains the contact region with the obstacle. We emphasize here that the approach presented in this work does not hinge on $\Gamma_c \neq \emptyset$. To measure the gap between Ω and the obstacle we use a given function

$$\psi \in Z := H^{1/2}(\Gamma_c) \text{ with } \psi \geq 0 \text{ almost everywhere (a.e.) on } \Gamma_c;$$

see [39]. For the time being we neglect frictional forces such that in terms of the variational formulation, we incorporate the contact constraint by a kinematic non-penetration condition on the displacement u :

$$(2.1) \quad \tau_n u \leq \psi \text{ on } \Gamma_c,$$

where $\tau_n : [H_{0,\Gamma_d}^1(\Omega)]^N \rightarrow Z, u \mapsto (\tau|_{\Gamma_c}(u)) \cdot n$ denotes the normal trace mapping restricted to Γ_c . For analytical reasons we assume that Γ_c is relatively open with $N^{1,1}$ -property and C^∞ -boundary $\partial\Gamma_c$.

For simplicity and without loss of generality we further stipulate

$$(2.2) \quad \bar{\Gamma}_c \subset \Sigma,$$

where Σ denotes the interior of $\partial\Omega \setminus \Gamma_d$ in $\partial\Omega$, to avoid working with the space $H_{00}^{1/2}(\Gamma_c)$. Concerning the splitting of the boundary we further assume

$$\partial\Omega = \bar{\Gamma}_c \cup \bar{\Gamma}_n \cup \bar{\Gamma}_d, \quad \Gamma_c \cap \Gamma_n \cap \Gamma_d = \emptyset, \quad \partial\Sigma \in C^\infty.$$

Denoting a given time step by $\delta t = t_1 - t_0 > 0$, the time-discrete elasto-plastic contact problem in variational form [8, 39] reads

$$(2.3) \quad \begin{cases} \min & \tilde{J}(u, p, \alpha) \quad \text{over } (u, p, \alpha) \in X \times Q \times L^2(\Omega)^m \\ \text{subject to (s.t.)} & \tau_n u \leq \psi \text{ on } \Gamma_c, \end{cases}$$

with

$$\begin{aligned} \tilde{J}(u, p, \alpha) := & \frac{1}{2} \int_{\Omega} \mathbb{C}(\varepsilon(u) - p) : (\varepsilon(u) - p) + \alpha : \mathbb{H}^{-1} \alpha \, dx \\ & + \delta t \int_{\Omega} \varphi^*((p - p_0)/\delta t, \mathbb{H}^{-1}(\alpha_0 - \alpha)/\delta t) \, dx \\ & - \int_{\Omega} f \cdot u \, dx - \int_{\Gamma_n} g \cdot u \, dx, \end{aligned}$$

where φ^* denotes the convex conjugate of a given l.s.c., convex and proper function φ which is defined as the characteristic function χ_K of the convex set K which in turn is defined by a given yield function ϕ , cf. (2.8) below. Here, $\alpha_0, p_0 \in L^2(\Omega)^m$ denote the given hardening and plastic strain at the preceding time instance, respectively. We endow the Hilbert spaces

$$X := [H_{0,\Gamma_d}^1(\Omega)]^N, \quad Q := [L^2(\Omega)]_{\text{sym}}^{N \times N}$$

with the usual scalar products. In this context, $\mathbb{C}(x) \in \mathbb{R}^{N \times N \times N \times N}, \mathbb{C}_{ijkl} \in L^\infty(\Omega)$, denotes the fourth-order elasticity tensor which is assumed to be symmetric, i.e. $\mathbb{C}_{ijkl} = \mathbb{C}_{klij} = \mathbb{C}_{jikl}$ and pointwise stable, i.e. $\exists C > 0$ with

$$\mathbb{C}(x)\sigma : \sigma \geq C|\sigma|_F^2 \quad \forall \sigma \in \mathbb{R}_{\text{sym}}^{N \times N} \text{ and a.e. } x \in \Omega,$$

where $A : B = \sum_{i,j=1 \dots N} a_{ij} \cdot b_{ij}$ for $A, B \in \mathbb{R}^{N \times N}$. Analogous properties are supposed to be fulfilled by the hardening modulus \mathbb{H} . The symmetric part of the displacement gradient is denoted by $\varepsilon(u)$, i.e.,

$$\varepsilon(u)(x) = \frac{1}{2}(\nabla u(x) + \nabla u(x)^\top).$$

Further, $\text{tr}(\sigma) := \sum_{i=1}^N \sigma_{ii}$ stands for the matrix trace operator.

Time-continuous problem. Given some material-dependent l.s.c., convex and proper yield functional $\phi : \mathbb{R}_{\text{sym}}^{N \times N} \times \mathbb{R}^m \rightarrow \mathbb{R} \cup \{+\infty\}$, the underlying elasto-plastic time-continuous problem consists of seeking $(u, p, \sigma, \alpha)(t) \in X \times Q \times Q \times L^2(\Omega)^m, t \in (0, T)$, with $(u, p, \sigma, \alpha)(0) = 0$ such that

$$\begin{aligned}
(2.4) \quad & u = 0 \text{ on } \Gamma_d, \\
(2.5) \quad & \sigma n = g \text{ on } \Gamma_n, \\
(2.6) \quad & \operatorname{div} \sigma = -f, \\
(2.7) \quad & \varepsilon(u) = \mathbb{C}^{-1} \sigma + p, \\
(2.8) \quad & (\sigma, \alpha) \in K := \{(\tau, \beta) : \phi(\tau, \beta) \leq 0\}, \\
(2.9) \quad & (\dot{p}, -\mathbb{H}^{-1} \dot{\alpha}) \in N_K(\sigma, \alpha), \\
(2.10) \quad & \tau_T \sigma = 0, \\
(2.11) \quad & \tau_{nn} \sigma \leq 0, \\
(2.12) \quad & \tau_{nn} \sigma (\tau_n u - \psi) = 0, \\
(2.13) \quad & \tau_n u \leq \psi
\end{aligned}$$

for a.e. $t \in (0, T)$, where $N_K(\sigma, \alpha)$ denotes the normal cone to the convex set K at (σ, α) . Furthermore, $\tau_{nn} \sigma := (\tau_n \sigma)^\top n$, and $\tau_T \sigma := \tau_n \sigma - (\tau_{nn} \sigma) n$ denotes the tangential trace on Γ_c ; cf. [39]. Note that (2.7)-(2.9) determine the plasticity behavior and (2.10)-(2.13) represent the complementarity conditions of contact. Moreover, (2.3) represents the weak formulation of one time step of a time-discrete version of (2.4)-(2.13) obtained from an implicit Euler time-discretization of the associative flow law (2.9). We refer to [8, 39] for details.

Elimination of α . Employing the von Mises yield criterion for combined isotropic-kinematic hardening to model the plastic straining, i.e.,

$$\phi(\sigma, \alpha) = |\operatorname{dev} \sigma - \operatorname{dev} \alpha_2|_F - \sigma_y (1 + H \alpha_1) + \chi_{\mathbb{R}_0^+}(\alpha_1), \quad \alpha = (\alpha_1, \alpha_2),$$

where $\operatorname{dev} \sigma := \sigma - \frac{\operatorname{tr}(\sigma)}{N} \cdot I$, one may compute the convex conjugate φ^* of φ for relevant hardening profiles including isotropic or kinematic hardening. In the latter cases, the hardening parameter α can be eliminated from the optimization problem [8]. The resulting trace condition on p will be incorporated into the space Q which gives

$$Q_0 := \{q \in [L^2(\Omega)]_{\operatorname{sym}}^{N \times N} : \operatorname{tr}(q) = 0 \text{ a.e. in } \Omega\}.$$

We note that Q_0 inherits the scalar product of Q , such that (2.3) may be recast into the form

$$(2.14) \quad \begin{cases} \min & J(u, p) \quad \text{over } (u, p) \in X \times Q_0 \\ \text{s.t.} & \tau_n u \leq \psi \text{ on } \Gamma_c \end{cases}$$

with

$$J(u, p) := \frac{1}{2} \int_{\Omega} \mathbb{C}(\varepsilon(u) - p) : (\varepsilon(u) - p) + \tilde{\mathbb{H}} p : p \, dx + \int_{\Omega} \beta |p|_F \, dx + l(u, p),$$

where $\beta \in L^2(\Omega)$ with $\beta \geq \sigma_y$ a.e. in Ω , $l \in (X \times Q_0)^*$, the topological dual space to $X \times Q_0$, and the fourth-order tensor $\tilde{\mathbb{H}}$ can be derived depending on the given hardening law; see (2.15) and (2.16) below. Here the material-dependent yield stress is denoted by $\sigma_y > 0$ and the modulus of hardening by $H > 0$. In the case of isotropic hardening, i.e.,

$$\phi(\sigma, \alpha) = |\operatorname{dev} \sigma|_F - \sigma_y (1 + H \alpha) + \chi_{\mathbb{R}_0^+}(\alpha),$$

we obtain $m = 1$, $\mathbb{H} = (k_1)$, $k_1 > 0$, $\tilde{\mathbb{H}} = (H\sigma_y)^2 k_1 I$, $\beta = \sigma_y(1 + H\alpha_0)$ and

$$(2.15) \quad l(u, p) = - \int_{\Gamma_n} g u \, ds - \int_{\Omega} f u \, dx - \int_{\Omega} \mathbb{C}(\varepsilon(u) - p) : p_0.$$

In the kinematic hardening case, i.e.,

$$\phi(\sigma, \alpha) = |\operatorname{dev} \sigma - \operatorname{dev} \alpha|_F - \sigma_y,$$

it holds that $m = \frac{n(n+1)}{2}$, $\tilde{\mathbb{H}} = \mathbb{H}$, $\beta = \sigma_y$ and

$$(2.16) \quad l(u, p) = - \int_{\Gamma_n} g u \, ds - \int_{\Omega} f u \, dx + \int_{\Omega} \alpha_0 : p \, dx - \int_{\Omega} \mathbb{C}(\varepsilon(u) - p) : p_0.$$

Note that (2.14) is equivalent to an elliptic variational inequality problem of the mixed (i.e. first and second) kind. Writing

$$\begin{aligned} y &:= (u, p) \in Y := X \times Q_0, \\ p &:= \Pi_{Q_0}(u, p), \quad \Pi_{Q_0} \in \mathcal{L}(Y, Q_0), \\ D([u, p], [v, r]) &:= \int_{\Omega} \mathbb{C}(\varepsilon(u) - p) : (\varepsilon(v) - r) + \tilde{\mathbb{H}}p : r \, dx, \end{aligned}$$

yields a more compact form of $J : Y \rightarrow \mathbb{R}$:

$$(EP) \quad J(y) = \frac{1}{2} \langle Dy, y \rangle_{(Y^*, Y)} + l(y) + \int_{\Omega} \beta \cdot |\Pi_{Q_0} y|_F \, dx,$$

where $D \in \mathcal{L}(Y, Y^*)$, i.e. it is a linear and continuous operator from Y to its topological dual Y^* . We note that the bilinear form associated with D is Y -elliptic under mild assumptions on \mathbb{H} which are fulfilled for isotropic or kinematic hardening [21]. Consequently, F is strongly convex.

Existence. Owing to the coercivity of F , standard arguments show that (2.14) admits a unique solution $\bar{y} = (\bar{u}, \bar{p}) \in Y$.

Remark. Using Moreau's theorem, (EP) can be further reduced to a (Fréchet) differentiable problem in the displacement only, cf. [19]. However, the resulting optimality condition is not eligible to Newton differentiation (in the sense of [24]) in infinite dimensions which may result in mesh-dependent convergence of an associated generalized Newton scheme. While the Newton differentiability of the stationarity system is always given in finite dimensions, the spatially continuous case requires a certain norm gap which is indispensable for the Newton differentiation of the involved composed max-function, cf. [26, 27] or section 6 for related issues. Such an integrability gap can never be achieved without further regularization.

3. FENCHEL DUALITY FOR THE ELASTO-PLASTIC CONTACT PROBLEM

For numerical purposes it turns out that the Fenchel dual problem to (2.14) is favorable in the sense that, upon regularization, it can be solved efficiently by semismooth Newton techniques.

In order to establish a compact set-up for the application of the Fenchel duality theory, the elasto-plastic contact problem (2.14) will be rewritten in the form

$$(EPC) \quad \min F(y) + G(\Lambda y), \quad \text{over } y \in Y,$$

with a Gâteaux-differentiable function F , a l.s.c., proper, and convex function G and a linear and continuous operator Λ . In fact, we define $F : Y \rightarrow \mathbb{R}$ by

$$F(y) := \frac{1}{2} \langle Dy, y \rangle_{(Y^*, Y)} + l(y).$$

Further, we denote the convex cone associated to the constraint (2.1) by

$$K_1 := \{z \in Z : z \leq 0 \text{ a.e. on } \Gamma_c\}$$

and define $G : Z \times L^2(\Omega)^d \rightarrow \mathbb{R} \cup \{+\infty\}$ by

$$G(z, q) := G_1(z) + G_2(q) := \chi_{\psi + K_1}(z) + \int_{\Omega} \beta |q|_2 \, dx.$$

Moreover, we set

$$\Lambda := \begin{bmatrix} \tau_n & 0 \\ 0 & M^{1/2} P^{-1} \end{bmatrix} \in \mathcal{L}(Y, Z \times L^2(\Omega)^d),$$

where $\chi_{\psi + K_1}$ is the indicator function of the set $\psi + K_1$, and

$$P : (L^2(\Omega)^d, \|\cdot\|_{L^2(\Omega)^d}) \rightarrow (Q_0, \|\cdot\|_{Q_0})$$

denotes the canonical parametrization

$$(3.1) \quad [q_1, \dots, q_d] \xrightarrow{P} \begin{bmatrix} q_1 & q_N & \dots & q_{d-(N-1)} \\ q_N & q_2 & \ddots & \vdots \\ \vdots & & \ddots & \\ q_{d-(N-1)} & \dots & q_{N-1} & q_d \\ & & q_d & -\sum_{i=1}^{N-1} q_i \end{bmatrix}$$

with $d = \frac{N(N+1)}{2} - 1$. The symmetric positive definite matrix M is defined by

$$Pp : Pq = Mp \cdot q \quad \forall p, q \in \mathbb{R}^d.$$

This setting differs from the one presented in [43] mainly by the choice of the operator Λ which entails a slightly different interpretation of the dual variable q , cf. (3.9).

We next compute and analyze the dual problem to (EPC).

Computation of the Fenchel conjugates. The convex conjugate $F^* : Y^* \rightarrow \mathbb{R}$ of $F : Y \rightarrow \mathbb{R}$ is given by

$$F^*(y^*) = \frac{1}{2} \langle y^* - l, D^{-1}(y^* - l) \rangle_{(Y^*, Y)}.$$

For the nondifferentiable part G we obtain

$$G^* : Z^* \times L^2(\Omega)^d \rightarrow \mathbb{R} \cup \{+\infty\}, \quad G^*(z^*, q) = G_1^*(z^*) + G_2^*(q),$$

with

$$G_2^* : L^2(\Omega)^d \rightarrow \mathbb{R} \cup \{+\infty\}, \quad G_2^*(q) = \chi_{K_2}(q),$$

where $K_2 := \{q \in L^2(\Omega)^d : |q|_2 \leq \beta \text{ a.e. in } \Omega\}$, and

$$G_1^* : Z^* \rightarrow \mathbb{R} \cup \{+\infty\}, \quad G_1^*(z^*) = \sup_{z \in K_1 + \psi} \langle z^*, z \rangle = \chi_{K_1^*}(z^*) + \langle z^*, \psi \rangle,$$

where it is understood that

$$(3.2) \quad \begin{aligned} K_1^* &:= Z_+^* = \{z^* \in Z^* : z^* \geq 0\} \\ &= \{z^* \in Z^* : \langle z^*, z \rangle \geq 0 \quad \forall z \in Z \text{ with } z \geq 0 \text{ a.e. on } \Gamma_c\}. \end{aligned}$$

The dual problem to (EPC) is given by

$$(D) \quad \sup -F^*(-\Lambda^*[z^*, q]) - G^*(z^*, q) \quad \text{over } [z^*, q] \in Z^* \times L^2(\Omega)^d,$$

which can be equivalently expressed as

$$\begin{cases} -\inf & F^*(\Lambda^*[z^*, q]) - \langle z^*, \psi \rangle & \text{over } [z^*, q] \in Z^* \times L^2(\Omega)^d \\ \text{s.t.} & z^* \leq 0, \\ & |q|_2 \leq \beta \text{ a.e. in } \Omega. \end{cases}$$

Note the sign change in the dual variables and that the first inequality constraint has to be understood in the sense of (3.2).

Since $K_1 + \psi$ has empty interior, a generalized Slater condition fails to hold. Hence the Fenchel duality theorem in its usual version [16] is not applicable. However, in our special situation we are still able to preclude the presence of a duality gap.

Proposition 3.1 (Duality). *There is no duality gap, i.e. it holds that*

$$\inf (\text{EPC}) = \sup (\text{D}).$$

Moreover, there exists a unique solution $(\bar{z}, \bar{q}) \in Z^* \times L^2(\Omega)^d$ to the dual problem.

Proof. We make use of [4, Theorem 1, Chapter 4.6], and need to show that

$$(3.3) \quad 0 \in \text{int}(\Lambda^* \text{dom } G^* + \text{dom } F^*).$$

As F^* is finite everywhere, we have $\text{dom } F^* = Y^*$. Further, $\text{dom } G^* \neq \emptyset$ implies $\Lambda^* \text{dom } G^* + \text{dom } F^* = Y^*$ such that (3.3) is always satisfied. It follows that no duality gap occurs.

Regarding existence and uniqueness of a solution to (D) we notice that the objective function is continuous and strictly convex since F^* is strongly convex and Λ^* is injective by the surjectivity of τ_n , cf. (2.2). Moreover, coercivity of the objective function follows from ellipticity of the bilinear form associated to D^{-1} . Indeed, with $\kappa > 0$ denoting the corresponding ellipticity constant, it follows that

$$\begin{aligned} & F^*(\Lambda^*[z^*, q]) - \langle z^*, \psi \rangle \\ &= \frac{1}{2} \langle \Lambda^*[z^*, q] - l, D^{-1}(\Lambda^*[z^*, q] - l) \rangle_{(Y^*, Y)} - \langle z^*, \psi \rangle \\ &\geq \frac{\kappa}{2} \|\Lambda^*[z^*, q]\|_{Z^* \times Q}^2 - \|\Lambda D^{-1}l + [\psi, 0]\| \| [z^*, q] \|_{Z^* \times Q} + \frac{\kappa}{2} \|l\|^2 \\ &\geq \frac{\kappa}{2\|\Lambda^{-*}\|} \| [z^*, q] \|_{Z^* \times Q}^2 - \|\Lambda D^{-1}l + [\psi, 0]\| \| [z^*, q] \|_{Z^* \times Q} + \frac{\kappa}{2} \|l\|^2, \end{aligned}$$

where the last estimate follows from the fact that Λ^* has a bounded inverse on its (closed) range owing to the closed range theorem. \square

Optimality conditions. By the absence of a duality gap (Proposition 3.1), the solution $\bar{y} = [\bar{u}, \bar{p}]$ of the primal problem (EPC) can be recovered from the solution $[\bar{z}, \bar{q}]$ of (D) from

$$(P-D) \quad \begin{aligned} \Lambda^*[\bar{z}, \bar{q}] &= D\bar{y} + l, \\ -[\bar{z}, \bar{q}] &\in \partial G(\Lambda\bar{y}). \end{aligned}$$

Due to (3.3), we may characterize the solution $[\bar{z}, \bar{q}] \in Z^* \times L^2(\Omega)^d$ by the existence of $\bar{\lambda} = [\bar{\mu}, \bar{\nu}] \in Z \times L^2(\Omega)^d$ satisfying

$$(OC1) \quad \Lambda D^{-1} \Lambda^* [\bar{z}, \bar{q}] - \Lambda D^{-1} l - [\psi, 0] + \bar{\lambda} = 0,$$

$$(OC2) \quad \bar{z} \leq 0, |\bar{q}|_2 \leq \beta \text{ a.e. in } \Omega,$$

$$(OC3) \quad \langle \bar{\mu}, z^* - \bar{z} \rangle \leq 0, (\bar{\nu}, q - \bar{q}) \leq 0 \quad \forall z^* \leq 0, \forall |q|_2 \leq \beta \text{ a.e. in } \Omega,$$

where the (OC3) expresses that $\bar{\lambda}$ is an element of the normal cone to $-K_1^* \times K_2$ at $[\bar{z}, \bar{q}]$. Equivalently, there exists $[\bar{\mu}, \bar{\zeta}] \in Z \times L^2(\Omega)$ with

$$(3.4) \quad \Lambda D^{-1} \Lambda^* [\bar{z}, \bar{q}] - \Lambda D^{-1} l - [\psi, 0] + [\bar{\mu}, \bar{\zeta}] = 0,$$

$$(3.5) \quad \bar{\zeta} - \max(0, \bar{\zeta} + c(|\bar{q}|_2 - \beta)) = 0, \quad c > 0,$$

$$(3.6) \quad \bar{z} \leq 0,$$

$$(3.7) \quad \langle \bar{\mu}, z^* - \bar{z} \rangle \leq 0 \quad \forall z^* \leq 0.$$

In general, these conditions are not directly eligible to the semismooth Newton method in the sense of [24]: Firstly, for generalized differentiation of the mapping associated with the left hand side of (3.5) in infinite dimensions, the setting lacks a suitable norm gap, see [26, 27] and section 6. Note that these issues are absent if a direct discretization is applied which may, however, be at the cost of mesh dependent convergence rates.

Secondly, (3.7) cannot be reformulated with the help of a pointwise NCP-function, i.e., a function $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ with the property

$$a \geq 0, b \geq 0, ab = 0 \iff \phi(a, b) = 0.$$

This is due to the fact that elements of Z^* in general do not allow for a pointwise interpretation. For these reasons we employ a penalization-regularization approach in the next sections.

Interpretation of the dual variables. Considering the second component in (P-D) and using $P^* = MP^{-1}$, we obtain a direct relation between \bar{q} and the optimal stress $\bar{\sigma} := \mathbb{C}(\varepsilon(\bar{u}) - \bar{p})$:

$$P(M^{-1/2}\bar{q}) = -\bar{\sigma} + \mathbb{H}\bar{p} \quad \text{in } Q_0^*,$$

assuming $\alpha_0 = p_0 \equiv 0$. Thus it holds that

$$(3.8) \quad P(M^{-1/2}\bar{q}) = \text{dev } \mathbb{H}\bar{p} - \text{dev } \bar{\sigma} \quad \text{in } Q_0.$$

In the case of the kinematic hardening law it can be shown that $\bar{\alpha} = \mathbb{H}\bar{p}$, cf. [8]. Together with (3.8), this induces

$$(3.9) \quad |\bar{q}|_2 - \sigma_y = |\text{dev } \bar{\sigma} - \text{dev } \bar{\alpha}|_F - \sigma_y,$$

which corresponds to the von Mises yield function value. In this way, the norm of \bar{q} determines the elasto-plastic material behavior. An analogous relation can be shown in the case of isotropic hardening. Moreover, by multiplying (P-D) by $[u, 0], u \in X$, it may be shown, analogously to the elastic case [39, 43], that \bar{z} corresponds to the normal stress $(\bar{\sigma}n) \cdot n \in Z^*$ at the contact boundary.

4. REGULARIZATION

In order to render the optimality conditions (OC1-3) amenable to the semismooth Newton method we now choose a Hilbert subspace $H = H_1 \times H_2 \subset L^2(\Gamma_c) \times L^2(\Omega)^d$ with dense embedding

$$H = H_1 \times H_2 \hookrightarrow Z^* \times L^2(\Omega)^d.$$

To obtain a consistent regularization, H_1 and H_2 are required to satisfy the following properties.

Assumption 4.1 (Density of convex intersections). *The following density assertions are supposed to hold:*

$$\overline{\iota_1^* \{z \in H_1 : z \leq 0 \text{ a.e. on } \Gamma_c\}}^{Z^*} = Z_-^*,$$

$$\overline{\{q \in H_2 : |q|_2 \leq \beta \text{ a.e. in } \Omega\}}^{L^2(\Omega)^d} = \{q \in L^2(\Omega)^d : |q|_2 \leq \beta \text{ a.e. in } \Omega\},$$

where $Z_-^* := \{z^* \in Z^* : \langle z^*, z \rangle_{(Z^*, Z)} \leq 0 \ \forall z \geq 0\}$ and ι_1^* is given by (4.1).

We further define $W := L^2(\Gamma_c) \times L^2(\Omega)^d$ and denote by

$$\iota^* = [\iota_1^*, \iota_2^*] : W \hookrightarrow [Z \times L^2(\Omega)^d]^* \simeq Z^* \times L^2(\Omega)^d$$

the canonical injection

$$(4.1) \quad [z, q] \mapsto [(z, \cdot)_{L^2(\Gamma_c)} |z, q] \in [Z \times L^2(\Omega)^d]^*.$$

Moreover, in the following illustration (see Figure 1) of the embedding framework including two Gelfand triples, we also specify the canonical injection

$$\tilde{\iota} : H \rightarrow W, [z, q] \xrightarrow{\tilde{\iota}} [z, q].$$

In this section only ι and ι^* will be mentioned explicitly whereas the other injections

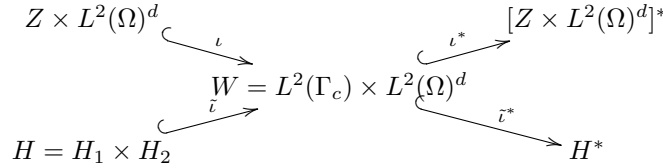


Figure 1. Gelfand triple framework for the regularization

are employed tacitly.

Suitable choices for H_1 and H_2 with regard to Assumption 4.1, possibly depending on the smoothness of Γ_c , can be made using Lemmas 5.3 and 5.4 as well as Lemma 5.5 of the subsequent section. For specific examples, we refer to section 7 below.

For algorithmic reasons it may be advantageous to include a non-negative shift parameter

$$(\hat{\mu}, \hat{\nu}) \in H_+^{1/2}(\Gamma_c) \times L_+^\infty(\Omega),$$

see [25]. Finally we replace β by an L^∞ -approximation β_γ which shall satisfy

$$\sigma_y \leq \beta_\gamma \leq \beta \text{ a.e.}, \quad \|\beta_\gamma - \beta\|_{L^2(\Omega)} \leq \frac{1}{\gamma}$$

for all γ .

Regularized problem. Following [14] we consider the regularized problem:

$$(D_\gamma) \quad \min J_\gamma^*(z, q) \quad \text{over } [z, q] \in H$$

with

$$J_\gamma^*(z, q) := F^*(\Lambda^* \iota^*[z, q]) - (z, \psi)_{L^2(\Gamma_c)} + M_\gamma^1(z) + M_\gamma^2(q) + T_\gamma(z, q),$$

where we employ the following Moreau-Yosida-type regularizations of the indicator function associated with the inequality constraints in (D):

$$\begin{aligned} M_\gamma^1(z) &:= \frac{1}{2\gamma} \|\hat{\mu} + \gamma z\|_{L^2(\Gamma_c)}^2, \\ M_\gamma^2(q) &:= \frac{1}{2\gamma} \|\hat{\nu} + \gamma(|q|_2 - \beta_\gamma)\|_{L^2(\Omega)}^2, \end{aligned}$$

as well as a regularization term of Tichonov type:

$$(4.2) \quad T_\gamma([z, q]) := \frac{1}{2\gamma} b([z, q], [z, q]),$$

where $b : H \times H \rightarrow \mathbb{R}$ is a continuous and coercive bilinear form represented by the operator $B \in \mathcal{L}(H, H^*)$ with ellipticity constant $\kappa_b > 0$.

Optimality condition. Note that (D $_\gamma$) has a unique solution $v_\gamma = [z_\gamma, q_\gamma] \in H$ which is characterized by

$$(OC1_\gamma) \quad 0 = N_\gamma v_\gamma - \iota \hat{w} + ([\mu_\gamma, \nu_\gamma], \cdot)_W \quad \text{in } H^*$$

with

$$(OC2_\gamma) \quad \begin{cases} \hat{w} = [\hat{z}, \hat{q}] = \Lambda D^{-1} l + [\psi, 0], \\ \mu_\gamma = [\hat{\mu} + \gamma z_\gamma]^+ \in L^2(\Gamma_c), \\ \nu_\gamma = [\hat{\nu} + \gamma(|q_\gamma|_2 - \beta_\gamma)]^+ \mathbf{q}(q_\gamma) \in L^2(\Omega)^d, \end{cases}$$

where we define $\mathbf{q}(\cdot) : L^2(\Omega)^d \rightarrow L^2(\Omega)^d$ by

$$\mathbf{q}(v) := \begin{cases} \frac{v}{|v|_2} & \text{if } |v|_2 > 0, \\ 0 & \text{else.} \end{cases}$$

Furthermore, the homeomorphism $N_\gamma \in \mathcal{L}(H, H^*)$ is defined as

$$N_\gamma := \iota \Lambda D^{-1} \Lambda^* \iota^* + \frac{1}{\gamma} B.$$

We close this section with an important consistency result concerning $\gamma \rightarrow +\infty$. This result suggests a path-following-type approach, where the associated primal-dual-path is induced by a sequence (γ_k) with $\gamma_k > 0$.

Theorem 4.2 (Convergence of regularized dual solutions). *Let $(\gamma) \subset \mathbb{R}^+$, $\gamma \rightarrow \infty$. Under Assumption 4.1 it holds that*

- (i) $v_\gamma = [z_\gamma, q_\gamma] \rightharpoonup [\bar{z}, \bar{q}]$ in $Z^* \times L^2(\Omega)^d$,
- (ii) $\lambda_\gamma = [\mu_\gamma, \nu_\gamma] \rightharpoonup [\bar{\mu}, \bar{\nu}]$ in $H_1^* \times H_2^*$,

and

$$\Lambda^* \iota^* v_\gamma \rightarrow \Lambda^* [\bar{z}, \bar{q}] \text{ in } Y^*.$$

Proof. See appendix B. □

As a simple consequence of the previous theorem, the sequence of approximations of the optimal displacement-strain pair and the sequence of trial stresses converge strongly to the corresponding solution of the original elasto-plastic contact problem (EPC). It may further be inferred that the sequence (q_γ) converges even with respect to the norm topology in $L^2(\Omega)^d$.

Corollary 4.3 (Convergence of primal solutions). *Under Assumption 4.1, the following assertions hold true:*

- (i) For $y_\gamma := D^{-1}(\Lambda^* \iota^*[z_\gamma, q_\gamma] - l)$ it holds that $y_\gamma \rightarrow \bar{y}$ in Y .
- (ii) For $\sigma_\gamma := \mathbb{C}(\varepsilon(u_\gamma) - p_\gamma)$ it holds that $\sigma_\gamma \rightarrow \bar{\sigma}$ in Q .
- (iii) It holds that $q_\gamma \rightarrow \bar{q}$ in $L^2(\Omega)^d$.

Proof.

- (i) The statement follows from the continuity of the operator D .
- (ii) The assertion follows from (i).
- (iii) The assertion follows from (i) and the fact that Λ_2^* is a topological isomorphism.

□

5. AUXILIARY RESULTS ON DENSITY-INVARIANT CONVEX INTERSECTIONS

In this section we discuss several conditions which lead to suitable options for the regularization space H with regard to Assumption 4.1. In general, for a Banach space V , an arbitrary dense subset $U \subset V$ as well as a convex and closed subset $K \subset V$ the inclusion

$$(5.1) \quad K \cap U \subset K \cap V$$

is not necessarily dense even for linear subspaces K and U . Therefore we investigate several situations relevant for our application in which the density of inclusion (5.1) is guaranteed. Readers who are merely interested in numerical aspects may as well directly consult the options for H specified in section 7 and take the Assumption 4.1 for granted.

Lemma 5.1 (intersection-invariant dense embedding). *Let V be a Hilbert space and U a dense subset $U \subset V$. Let $K \subset V$ be nonempty, convex and closed. If the projection mapping $P_K : V \rightarrow K$ is U -invariant, i.e.*

$$P_K(U) \subset U,$$

then $\overline{U \cap K}^V = K$, i.e. $U \cap K$ is dense in K with respect to the norm in V .

Proof. For $v \in K$ there exists a sequence $(u_n) \subset U$ with $u_n \rightarrow v$. Now, $P_K(u_n) \in U$ for all n by assumption, such that

$$\|P_K(u_n) - v\|_V = \|P_K(u_n) - P_K(v)\|_V \leq \|u_n - v\|_V \rightarrow 0.$$

□

Lemma 5.2 (superposition and trace). *Let $\theta : \mathbb{R} \rightarrow \mathbb{R}$ be Lipschitz continuous and assume that $\theta'(t)$ exists except for finitely many points $t \in \mathbb{R}$. Further let Ω be a Lipschitz domain. Assume that $\mu(\Omega) < +\infty$ or $\theta(0) = 0$. For the trace operator $\tau : H^1(\Omega) \rightarrow H^{1/2}(\partial\Omega)$ it holds that*

$$(5.2) \quad (\theta \circ \tau)(u) = (\tau \circ \theta)(u) \quad \text{in } L^2(\partial\Omega)$$

for all $u \in H^1(\Omega)$.

Proof. Under the above conditions, the superposition

$$\theta_1 = \theta : L^2(\partial\Omega) \rightarrow L^2(\partial\Omega)$$

is well-defined and continuous. Further, it is well known that the superposition

$$\theta_2 = \theta : H^1(\Omega) \rightarrow H^1(\Omega)$$

is also well-defined [32] and continuous, cf. [38]. Since (5.2) holds for any $u \in C(\bar{\Omega}) \cap H^1(\Omega)$, a density argument completes the proof. \square

Lemma 5.3. *For $L_-^2(\Gamma_c) := \{z \in L^2(\Gamma_c) : z \leq 0 \text{ a.e. on } \Gamma_c\}$ it holds that*

$$\overline{\iota_1^*(L_-^2(\Gamma_c))}^{Z^*} = Z_-^*.$$

Proof. Define the closed, convex and nonempty set $M \subset Z^*$ by

$$M := \overline{\iota_1^*(L_-^2(\Gamma_c))}^{Z^*} \subset Z_-^*$$

and assume $Z_-^* \setminus M \neq \emptyset$. For $0 \neq z^* \in Z_-^* \setminus M$ it holds that $\alpha z^* \in Z_-^* \setminus M$ for all $\alpha > 0$. Furthermore, there exists a sequence $(v_n) \subset L^2(\Gamma_c)$ with

$$(5.3) \quad v_n \rightarrow z^* \text{ in } Z^*.$$

We first assume that $\|v_n\|_{L^2(\Gamma_c)} \rightarrow +\infty$ as $n \rightarrow +\infty$. The Hahn-Banach Separation Theorem implies that for all $n \in \mathbb{N}$ there exists $z_n \in Z$ with

$$(5.4) \quad \langle z_n, \frac{1}{\|v_n\|_{L^2(\Gamma_c)}^2} z^* \rangle_{(Z, Z^*)} > 1 \quad \text{and}$$

$$(5.5) \quad \langle z_n, v \rangle_{(Z, Z^*)} \leq 1 \quad \text{for all } v \in M.$$

We decompose $z_n = z_n^+ + z_n^-$ into a positive part $z_n^+ := \max(0, z)$ and a negative part $z_n^- := \min(0, z)$, where it is easy to see that $\{z_n^+, z_n^-\} \subset Z$. Indeed, recall (cf. e.g. [18, p.20]) that $Z = H^{1/2}(\Gamma_c)$ is defined by the set of all $z \in L^2(\Gamma_c)$ with finite s -seminorm, $s \in (0, 1)$, i.e.

$$|z|_{\Gamma_c, s}^2 := \int_{\Gamma_c} \int_{\Gamma_c} \frac{|z(x) - z(y)|^2}{|x - y|^{(n-1)+2s}} dx dy < +\infty.$$

Further observe that $\max(0, z) \in L^2(\Gamma_c)$ and superposition with Lipschitz functions preserves the finiteness of the s -seminorm. Alternatively one may invoke Lemma 5.2. From (5.4) and $z^* \in Z_-^*$ it follows that

$$\langle z_n^-, \frac{1}{\|v_n\|_{L^2(\Gamma_c)}^2} z^* \rangle_{(Z, Z^*)} > 1,$$

in particular $z_n^- \neq 0$. Setting $v = \iota_1^*(z_n^-) \|z_n^-\|_Z \|z_n^-\|_{L^2(\Gamma_c)}^{-2}$ in (5.5), where $v \in M$ by definition, one obtains

$$(5.6) \quad \langle z_n^-, v \rangle_{(Z, Z^*)} = \|z_n^-\|_Z \leq 1.$$

On the other hand, for v_n according to (5.3) and for sufficiently large $n \in \mathbb{N}$ it holds that

$$\begin{aligned} \|z_n^-\|_{L^2(\Gamma_c)} &= \sup_{v \in L^2(\Gamma_c)} \frac{(z_n^-, v)_{L^2(\Gamma_c)}}{\|v\|_{L^2(\Gamma_c)}} \\ &\geq \frac{(z_n^-, v_n)_{L^2(\Gamma_c)}}{\|v_n\|_{L^2(\Gamma_c)}} = \frac{(z_n^-, z^*)_{(Z, Z^*)}}{\|v_n\|_{L^2(\Gamma_c)}} - \frac{(z_n^-, z^* - \iota_1^* v_n)_{(Z, Z^*)}}{\|v_n\|_{L^2(\Gamma_c)}} \\ &\geq \|v_n\|_{L^2(\Gamma_c)} - \frac{\|z_n^-\|_Z \|z^* - \iota_1^* v_n\|_{Z^*}}{\|v_n\|_{L^2(\Gamma_c)}} \\ &\rightarrow +\infty \quad \text{for } n \rightarrow +\infty, \end{aligned}$$

due to (5.3) and (5.6). This clearly contradicts (5.6).

If (v_n) is bounded in $L^2(\Gamma_c)$, it converges weakly (along a subsequence) in $L^2(\Gamma_c)$ to an element $u \in L^2(\Gamma_c)$, such that $z^* = u$ by (5.3). This in turn implies $z^* \in L^2(\Gamma_c) \cap Z_-^* = L_-^2(\Gamma_c)$. The latter equation relies on the density of Z_+ in $L_+^2(\Gamma_c)$ with respect to the norm topology in $L^2(\Gamma_c)$, which holds as a consequence of Lemma 5.1. Thus it holds that $z^* \in M$, which contradicts the initial hypothesis. \square

Lemma 5.4. *Let $\Omega \subset \mathbb{R}^n$ be open, $\mu(\Omega) < +\infty$, $d \in \mathbb{N}$ and $\beta : \Omega \rightarrow \mathbb{R}$ measurable with $\beta(x) \geq \sigma > 0$ a.e. in Ω . For*

$$K := \{u \in L^2(\Omega)^d : |u|_2 \leq \beta \text{ a.e. in } \Omega\}$$

it holds that $K_\infty := K \cap [C_0^\infty(\Omega)]^d$ is densely contained in K , i.e. $\overline{K_\infty}^{L^2(\Omega)^d} = K$.

Proof. Let $u \in K$ and $\varepsilon > 0$.

Part I. We first choose a function $g \in C_0^0(\Omega)^d$, $g = [g_1, \dots, g_d]$, with the following properties:

$$(5.7) \quad \begin{cases} |g_j(x)| \leq |u_j(x)| & \text{a.e. in } \Omega, \\ \|g_j - u_j\|_{L^2(\Omega)} & < \frac{\varepsilon}{\sqrt{d}}, \end{cases}$$

for $j = 1, \dots, d$. A suitable choice can be made using Lusin's Theorem: In fact, there exist for all $\delta > 0$

$$K_j \subset \Omega \text{ compact, } \mu(\Omega \setminus K_j) < \delta, \quad j = 1, \dots, d,$$

with $u_j|_{K_j}$ continuous. We define the $C_0^0(\Omega)$ -functions

$$g_j(x) := \frac{\min(\delta, \text{dist}(x, \Omega \setminus K_j))}{\delta} u_j(x)$$

which fulfill (5.7) for sufficiently small δ .

Therefore $g \in C_0^0(\Omega)^d$ is an element of K . Moreover,

$$\|g - u\|_{L^2(\Omega)^d}^2 = \sum_{j=1}^d \|g_j - u_j\|_{L^2(\Omega)}^2 < \varepsilon^2.$$

We thus have shown that $K_0 := K \cap C_0^0(\Omega)^d$ is densely contained in K .

Part II. To conclude the proof we take an arbitrary sequence $(u_n)_{n \in \mathbb{N}} \subset K_0$ which fulfills

$$\|u_n - u\|_{L^2(\Omega)^d} \rightarrow 0.$$

Further set $\tilde{u}_n := \frac{n}{n+1} u_n \in C_0^0(\Omega)^d$.

By continuity and the hypothesis $\beta(x) \geq \sigma$ a.e. in Ω there exists $\delta_n > 0$ with

$$(5.8) \quad |\tilde{u}_n|_2(x) \leq \beta(x) - \delta_n \quad \text{for a.e. } x \in \Omega.$$

Moreover, for every n a suitable mollification yields a sequence $(v_n^k)_{k \in \mathbb{N}} \subset [C_0^\infty(\Omega)]^d$ with

$$(5.9) \quad \lim_k v_n^k \rightarrow \tilde{u}_n \text{ in } C(\bar{\Omega}).$$

Combining (5.8) and the uniform convergence property (5.9) one obtains that for each n there exists $k(n)$ such that $v_n^k \in K_\infty$ and $\|v_n^k - \tilde{u}_n\|_{L^2(\Omega)^d} < \frac{\varepsilon}{3}$ for all $k \geq k(n)$.

Finally choose n sufficiently large such that

$$\|u - u_n\|_{L^2(\Omega)^d} < \frac{\varepsilon}{3}, \quad \|u_n - \tilde{u}_n\|_{L^2(\Omega)^d} < \frac{\varepsilon}{3}.$$

Applying the triangle inequality shows that $(v_n^{k(n)}) \subset K_\infty$ satisfies

$$\|v_n^{k(n)} - u\|_{L^2(\Omega)^d} < \varepsilon,$$

for sufficiently large n , which accomplishes the proof. \square

The contact boundary as a Riemannian manifold. In order to allow for a distribution theory on the manifold Γ_c similar to the Euclidean case, we need to define the space of test functions $C_0^\infty(\Gamma_c)$ on a manifold Γ_c which requires a smooth structure. In connection with an associated Riemannian measure this leads to the definition of Sobolev spaces on manifolds allowing for a complete calculus theory, cf. [17]. For the alternative approach via the completion of smooth functions w.r.t. the $W^{k,p}$ -norm see [22].

In the remaining part of this section we therefore assume that the contact boundary Γ_c is smooth, i.e., a C^∞ -submanifold of \mathbb{R}^n . More precisely, since $\partial\Omega$ is assumed to have the $N^{0,1}$ -property [47], $\partial\Omega$ (possibly after an appropriate orthogonal coordinate transformation) is given locally by the graph of functions $\alpha_i \in C_B^{0,1}$, $i = 1, \dots, m$. We assume that those α_i whose graph has nonempty intersection with Γ_c , are not only in $C_B^{1,1}$ but in $C^\infty \cap C_B^{1,1}$ on an appropriate bounded domain in \mathbb{R}^{N-1} . Here, the space $C_B^{k,\kappa}$ is defined as the set of k -times continuously differentiable functions with bounded derivatives of order less than or equal k and κ -Hölder-continuous k -th derivative [47].

In this way, Γ_c becomes an $(N-1)$ -dimensional C^∞ -submanifold of \mathbb{R}^N . We further endow the Cartesian product of the tangent spaces of Γ_c with the usual scalar product in \mathbb{R}^N . This canonical construction yields a Riemannian manifold $(\Gamma_c, \langle \cdot, \cdot \rangle_{\mathbb{R}^N})$.

Lemma 5.5. *Let Γ_c be a C^∞ -submanifold of \mathbb{R}^N and consider (Γ_c, g) , $g = \langle \cdot, \cdot \rangle_{\mathbb{R}^N}$, as a Riemannian manifold with associated Riemannian measure $\mu = \mu(g)$. Then for $L_-^2(\Gamma_c) := \{u \in L^2(\Gamma_c) : u \leq 0 \text{ } \mu\text{-a.e. on } \Gamma_c\}$,*

$$K_\infty := L_-^2(\Gamma_c) \cap [C_0^\infty(\Gamma_c)]$$

is densely contained in $L_-^2(\Gamma_c)$.

Proof. Let $u \in L_-^2(\Gamma_c)$. Since $C_0^\infty(\Gamma_c)$ is dense in $L^2(\Gamma_c)$ [17] there exists a sequence $(v_k) \subset C_0^\infty(\Gamma_c)$, such that $v_k \rightarrow u$ in $L^2(\Gamma_c)$. We further denote by

$$\psi_k \in C^{0,1}(\mathbb{R}) \cap C^\infty(\mathbb{R}), k \in \mathbb{N},$$

non-positive functions with uniformly bounded Lipschitz modules L_k , i.e. $\sup_k L_k < +\infty$, which satisfy

$$\psi_k(t) \rightarrow \min(0, t) \quad (\text{pointwise}).$$

Such a function can be easily constructed [17, Example 5.3]. Using the triangle inequality we infer

$$\begin{aligned} & \|u - \psi_k(v_k)\|_{L^2(\Gamma_c)} \\ & \leq \underbrace{\|\min(0, u) - \psi_k(u)\|_{L^2(\Gamma_c)}}_{\rightarrow 0} + \underbrace{\|\psi_k(u) - \psi_k(v_k)\|_{L^2(\Gamma_c)}}_{\leq L_k \|u - v_k\|_{L^2(\Gamma_c)}} \end{aligned}$$

where the convergence of the left summand follows from the Dominated Convergence Theorem. This completes the proof. \square

6. SEMISMOOTH NEWTON METHOD

Considering the necessary and sufficient optimality conditions (OC1 $_{\gamma}$) - (OC2 $_{\gamma}$) of the regularized problem, the goal of this section is the application of the semismooth Newton method applied to a suitable operator equation which equivalently characterizes the optimality conditions. The notion of Newton differentiability which is applied here can be found in [11, 24] and reads as follows.

Definition 6.1 (Newton differentiability). Let X, Y be Banach spaces and $U \subset X$ be an open set. A mapping $F : U \rightarrow Y$ is called Newton differentiable in U if there exists a family of mappings $G_F : U \rightarrow \mathcal{L}(X, Y)$ which satisfy

$$\|F(x+h) - F(x) - G_F(x+h)h\|_Y = o(\|h\|_X), \quad \|h\|_X \rightarrow 0,$$

for all $x \in U$.

The corresponding generalized Newton method converges locally at a superlinear rate [11]. Further, mesh independence results [23, 28] are available. We emphasize that the semismooth Newton method has found considerable attention throughout the last decade as it has proved to be a remarkably efficient method, notably for the solution of various problems in PDE-constrained optimization [24, 25, 26] and variational inequalities [14, 27, 37], to mention only a few.

We further rely on the following calculus rules related to the Newton differentiability of several nonsmooth functions which can be found in [24] and [27].

For measurable subsets $\tilde{\Omega} \subset \Omega$ or $\tilde{\Omega} \subset \partial\Omega$ and $1 \leq q \leq p \leq \infty$, the operator $[\cdot]^+$ defined by

$$\begin{aligned} [\cdot]^+ &: L^p(\tilde{\Omega})^d \rightarrow L^q(\tilde{\Omega})^d, \\ v &\mapsto (x \mapsto \max(0, v(x))), \end{aligned}$$

from now on always denotes the pointwise max-operator.

Lemma 6.2 (Newton differentiability of the pointwise maximum). *The pointwise maximum function $F(\cdot) := [\cdot]^+$*

$$F : L^p(\tilde{\Omega}) \rightarrow L^q(\tilde{\Omega}),$$

is Newton differentiable for $1 \leq q < p \leq +\infty$. A corresponding Newton derivative is given by

$$G_F(u)h := \begin{cases} 0, & \text{on } \mathcal{I}(u), \\ h, & \text{on } \mathcal{A}(u), \end{cases}$$

where $\mathcal{A}(u) := \{x \in \tilde{\Omega} : u(x) > 0\}$ and $\mathcal{I}(u) := \tilde{\Omega} \setminus \mathcal{A}(u)$.

Lemma 6.3 (Newton differentiability of a generalized maximum function). *Let $\beta \in L^\infty(\Omega)$ with $\beta(x) \geq c > 0$ a.e. in Ω . Then the mapping*

$$\mathfrak{m} : u \mapsto [|u|_2 - \beta]^+ \mathfrak{q}(u)$$

is Newton differentiable as a mapping from $L^p(\Omega)^d \rightarrow L^s(\Omega)^d$ for $3 \leq 3s \leq p < +\infty$. A corresponding Newton derivative is given by

$$G_{\mathfrak{m}}(u) := \chi_{\mathcal{A}(u)} \cdot \mathfrak{M}(u)$$

where

$$(6.1) \quad \begin{cases} \rho(u) & := [|u|_2 - \beta]^+ \frac{1}{|u|_2}, \\ \mathfrak{M}(u)(\cdot) & := \rho(u)(\cdot) + (1 - \rho(u)) \frac{uu^\top(\cdot)}{|u|_2^2}, \\ \mathcal{A}(u) & := \{x \in \Omega : |u|_2(x) > \beta(x)\}. \end{cases}$$

Reformulation. We equivalently reformulate the optimality condition (OC1 $_\gamma$) for v_γ by the nonsmooth equation

$$(6.2) \quad \Psi_\gamma(\lambda_\gamma) = 0$$

using the operator $\Psi_\gamma : H^* \rightarrow H^*$ defined by

$$\Psi_\gamma \begin{bmatrix} \mu \\ \nu \end{bmatrix} := \begin{bmatrix} \mu \\ \nu \end{bmatrix} - \tilde{t}^* \begin{bmatrix} [\hat{\mu} + \gamma z(\lambda)]^+ \\ [\hat{\nu} + \gamma(|q(\lambda)|_2 - \beta_\gamma)]^+ \mathfrak{q}(q(\lambda)) \end{bmatrix},$$

where $v(\lambda) := (z(\lambda), q(\lambda)) := N_\gamma^{-1}(u\hat{w} - \lambda) \in H$ denotes the solution to (OC1 $_\gamma$) given some candidate λ for λ_γ . The superlinear convergence of the generalized Newton method

$$(6.3) \quad \lambda^{j+1} = \lambda^j - G_{\Psi_\gamma}(\lambda^j)^{-1} \Psi_\gamma(\lambda^j)$$

to solve (6.2) hinges, among others, on the Newton differentiability of Ψ_γ in the sense of Definition 6.1. In view of the preceding calculus rules the latter relies on the following assumption.

Assumption 6.4 (Norm gap). *With regard to Lemma 6.2 and Lemma 6.3, the Newton differentiability of Ψ_γ requires additional restrictions on the choice of the spaces H_1 and H_2 . For this purpose, imposing the conditions*

$$H_1 \subset L^{2+\varepsilon}(\Gamma_c), \quad \varepsilon > 0, \quad \text{and} \quad H_2 \subset [L^6(\Omega)]^d$$

is sufficient.

From now on, we assume that the regularization space H is chosen in such a way that Assumption 6.4 is fulfilled. Thus, the operator $\Psi_\gamma : H^* \rightarrow H^*$ is Newton differentiable. We proceed by computing a particular Newton derivative.

Using the chain rule for Newton derivatives with affine continuous functions, $G_{\Psi_\gamma}(\lambda) \in \mathcal{L}(H^*, H^*)$ is given by

$$G_{\Psi_\gamma}(\lambda)(\cdot) = \text{id}_{H^*}(\cdot) + \gamma \tilde{t}^* \begin{bmatrix} \chi_{\mathcal{Z}_\gamma(z(\lambda))} & 0 \\ 0 & \chi_{\mathcal{Q}_\gamma(q(\lambda))} \mathfrak{M}(q(\lambda)) \end{bmatrix} \circ N_\gamma^{-1}(\cdot),$$

which includes the following quantities:

$$\begin{aligned} \rho(q) & := [|q|_2 + \frac{\hat{\nu}}{\gamma} - \beta_\gamma]^+ \frac{1}{|q|_2}, \\ \mathfrak{M}(q(\lambda))(\cdot) & = \rho(q(\lambda))(\cdot) + (1 - \rho(q(\lambda))) \frac{q(\lambda)q(\lambda)^\top(\cdot)}{|q(\lambda)|_2^2}, \\ \mathcal{Z}_\gamma(z) & := \{x \in \Gamma_c : (z + \frac{\hat{\mu}}{\gamma})(x) > 0\}, \\ \mathcal{Q}_\gamma(q) & := \{x \in \Omega : (|q|_2 + \frac{\hat{\nu}}{\gamma} - \beta_\gamma)(x) > 0\}. \end{aligned}$$

We start the analysis of the generalized Newton iteration (6.3) by the following lemma.

Lemma 6.5 (Uniform invertibility). *The operator*

$$G_{\Psi_\gamma}(\lambda) \in \mathcal{L}(H^*, H^*)$$

is uniformly invertible, i.e., for all $\delta \in H^$ we have*

$$\|\delta\|_{H^*} \leq c(\gamma) \|G_{\Psi_\gamma}(\lambda)\delta\|_{H^*}, \quad \text{with } c(\gamma) > 0.$$

Proof. Similarly to [14] we decompose

$$G_{\Psi_\gamma}(\lambda) = \tilde{N}_\gamma(\lambda) \circ N_\gamma^{-1}$$

with

$$\tilde{N}_\gamma(\lambda) = \left(N_\gamma + \gamma \tilde{t}^* \begin{bmatrix} \chi_{Z_\gamma(z(\lambda))} & 0 \\ 0 & \chi_{\mathcal{Q}_\gamma(q(\lambda))} \mathfrak{M}(q(\lambda)) \end{bmatrix} \right).$$

The operator $\tilde{N}_\gamma(\lambda) \in \mathcal{L}(H, H^*)$ is uniformly invertible, i.e., independently of λ , since

$$\begin{aligned} & \langle \tilde{t}^* \begin{bmatrix} \chi_{Z_\gamma(z(\lambda))} & 0 \\ 0 & \chi_{\mathcal{Q}_\gamma(q(\lambda))} \mathfrak{M}(q(\lambda)) \end{bmatrix} \begin{bmatrix} r \\ p \end{bmatrix}, \begin{bmatrix} r \\ p \end{bmatrix} \rangle_{(H^*, H)} \\ &= (\chi_{Z_\gamma(z(\lambda))} r, r)_{L^2(\Gamma_c)} + (\chi_{\mathcal{Q}_\gamma(q(\lambda))} \mathfrak{M}(q(\lambda)) p, p)_{L^2(\Omega)^d} \\ &\geq \int_{Q_\gamma(q(\lambda))} \rho(q(\lambda)) \left(|p|_2^2 - \frac{(q(\lambda):p)^2}{|q(\lambda)|_2^2} \right) \geq 0. \end{aligned}$$

The assertion follows from the ellipticity of the bilinear form associated to N_γ . \square

Lemma 6.5 guarantees that step (ii) of the following semismooth Newton algorithm is well-defined.

Algorithm 6.6 (SSN algorithm in λ).

- (i) Initialize $\lambda^0 := (\mu^0, \nu^0) \in H^* = H_1^* \times H_2^*$; set $j := 0$.
- (ii) Unless some stopping rule is satisfied, compute the solution $\delta_\lambda^j \in H^*$ of

$$(6.4) \quad G_{\Psi_\gamma}(\lambda^j) \delta_\lambda^j = -\Psi_\gamma(\lambda^j).$$

- (iii) Set $\lambda^{j+1} := \lambda^j + \delta^j$, $j := j + 1$ and return to (ii).

We immediately infer local superlinear convergence.

Corollary 6.7 (Semismooth Newton algorithm). *If $\lambda^0 \in H^*$ is sufficiently close to λ_γ , then the following assertions hold true:*

- (i) *The Newton iterates $(\lambda^j) \subset H^*$ generated by Algorithm 6.6 converge superlinearly to $\lambda_\gamma \in W$.*
- (ii) *The Newton iterates $(v^j) \subset H$ defined by $v^j = N_\gamma^{-1}(\iota(\hat{w}) - \lambda^j)$ generated by the preceding algorithm converge superlinearly to v_γ in H .*
- (iii) *If $\lambda^0 \in W$, then $(\lambda^j)_{j \in \mathbb{N}} \subset W$.*

Proof.

- (i) The assertion follows directly from [24, Theorem 1.1].
- (ii) The assertion is a consequence of the fact that superlinear convergence is preserved by the topological isomorphism N_γ .

- (iii) If $\lambda^j \in W$, then we have $\Psi_\gamma(\lambda^j) \in W$.
The definition of the Newton step (6.4) yields

$$G_{\Psi_\gamma}(\lambda^j)\delta_\lambda^j = -\Psi_\gamma(\lambda^j) \iff \delta_\lambda^j + \gamma \tilde{\ell}^* \underbrace{\begin{bmatrix} \chi_{\mathcal{Z}_\gamma(z(\lambda^j))} & 0 \\ 0 & \chi_{\mathcal{Q}_\gamma(q(\lambda^j))} \mathfrak{M}(q(\lambda^j)) \end{bmatrix}}_{\in W} \circ N_\gamma^{-1} \delta_\lambda^j = \underbrace{-\Psi_\gamma(\lambda^j)}_{\in W}$$

which proves the assertion. \square

For completeness we also specify the SSN algorithm in v (rather than in λ) which is analyzed in Corollary 6.7.

Algorithm 6.8 (SSN algorithm in v).

- (i) Initialize $\lambda^0 \in W$, $v^0 := N_\gamma^{-1}(\iota(\hat{w}) - \lambda^0)$; set $j := 0$.
(ii) Unless some stopping rule is satisfied, compute the solution $\delta_v^j \in H$ of
(6.5)
$$\tilde{N}_\gamma(\lambda^j)(-\delta_v^j) = -\Psi_\gamma(\lambda^j).$$

(iii) Set $v^{j+1} := v^j + \delta_v^j$, $j := j + 1$ and return to (ii).

For the globalization of our Newton-scheme one may use a line search procedure [14]. For this purpose, we need to check whether the update direction, say δ_v^j in Algorithm 6.8, is related to the gradient of J_γ^* . This is the content of the following result.

Proposition 6.9 (Gradient-relatedness). *The search directions (δ_v^j) generated by Algorithm 6.8 satisfy*

$$\langle J_\gamma^{*'}(v^j), \delta_v^j \rangle_{(H^*, H)} \leq -\frac{\kappa_b}{\gamma C(\gamma)^2} \|J_\gamma^{*'}(v^j)\|_{H^*}^2,$$

where $C(\gamma) = \sup_\lambda \|\tilde{N}_\gamma(\lambda)\| \in (0, +\infty)$.

Proof. Note that $J_\gamma^{*'}(v^j) = -\Psi_\gamma(\lambda^j)$. Using (6.5) we conclude that

$$\begin{aligned} \langle J_\gamma^{*'}(v^j), \delta_v^j \rangle_{(H^*, H)} &= \langle J_\gamma^{*'}(v^j), -\tilde{N}_\gamma(\lambda^j)^{-1}(J_\gamma^{*'}(v^j)) \rangle_{(H^*, H)} \\ &\leq -\frac{\kappa_b}{\gamma \|\tilde{N}_\gamma(\lambda^j)\|} \|J_\gamma^{*'}(v^j)\|_{H^*}^2, \end{aligned}$$

since it holds for arbitrary $g = \tilde{N}_\gamma(\lambda)v \in H^*$ that

$$\begin{aligned} \langle \tilde{N}_\gamma(\lambda)^{-1}g, g \rangle &= \langle \tilde{N}_\gamma(\lambda)v, v \rangle \\ &\geq \frac{\kappa_b}{\gamma} \|v\|_H^2 \geq \frac{\kappa_b}{\gamma} \frac{1}{\|\tilde{N}_\gamma(\lambda)\|^2} \|g\|_{H^*}^2. \end{aligned}$$

The definition of \mathfrak{M} , cf. (6.1), yields for $v = [r, p] \in H$ that

$$\begin{aligned} \|\tilde{N}_\gamma(\lambda)v\|_{H^*} &\leq \|N_\gamma v\|_{H^*} + \gamma \left\| \tilde{\ell}^* \begin{bmatrix} \chi_{\mathcal{Z}_\gamma(z(\lambda))} r \\ \chi_{\mathcal{Q}_\gamma(q(\lambda))} \mathfrak{M}(q(\lambda)) p \end{bmatrix} \right\|_{H^*} \\ &\leq \|N_\gamma\| \|v\|_H + \gamma C \|v\|_W \leq (\|N_\gamma\| + \gamma C) \|v\|_H, \end{aligned}$$

where $C > 0$ may take different values on different occasions. This ends the proof. \square

Remark 6.10 (Global convergence). We immediately infer that endowing the search directions (δ_v^j) with a line search method fulfilling the Armijo condition yields global convergence of the generalized Newton method [7].

7. NUMERICAL VALIDATION

In this section we validate the theoretical algorithmic framework by numerical tests. For this purpose, we specify the Tichonov regularization as well as the precise discrete setting.

Regularization. We propose two choices for the Tichonov regularization pair $[H, b]$.

(R1) Setting $H := H^{1/2}(\Gamma_c) \times H_0^1(\Omega)^d$, we define

$$b_1([z_1, z_2], [q_1, q_2]) := (z_1, z_2)_{H^{1/2}(\Gamma_c)} + \sum_{i=1}^d (\nabla q_{1,i}, \nabla q_{2,i})_{L^2(\Omega)^N}.$$

(R2) If Γ_c is a C^∞ -submanifold of \mathbb{R}^N , we set $H := H_0^1(\Gamma_c) \times H_0^1(\Omega)^d$ and

$$b_2([z_1, z_2], [q_1, q_2]) := (\nabla z_1, \nabla z_2)_{\overrightarrow{L^2}(\Gamma_c)} + \sum_{i=1}^d (\nabla q_{1,i}, \nabla q_{2,i})_{L^2(\Omega)^N},$$

where the Hilbert space $\overrightarrow{L^2}(\Gamma_c)$ is defined by the set of (equivalence classes of) vector fields $u : \Gamma_c \rightarrow T\Gamma_c$, i.e. $u(x) \in T_x\Gamma_c$ for all $x \in \Gamma_c$, with integrable Riemannian product $\langle u, u \rangle_{\mathbb{R}^N}$ on Γ_c equipped with the canonical surface measure. Here, $T\Gamma_c := \cup_x T_x(\Gamma_c)$ denotes the tangent bundle to Γ_c . For details see section 5.

Recalling the discussion in section 5, both choices fulfill Assumption 4.1. Moreover, the Sobolev Imbedding Theorem ensures that Assumption 6.4 is satisfied [1]. For the choice of the bilinear form on the boundary, ellipticity of b_2 is ensured by Theorem A.1. We emphasize that other choices like the $[H^1(\Omega)]^d$ -inner product for the domain variable q are equally conceivable and may be preferred depending on the application and the approximations of physical quantities derived from the dual variables, cf. (3.9).

In view of Theorem 4.2 and Corollary 4.3, Algorithm 6.8 is embedded into an update scheme for γ , i.e. once Algorithm 6.8 terminates successfully for a given γ , the penalty parameter is increased and Algorithm 6.8 is restarted.

In order to avoid the inverse D^{-1} in (6.5) we explicitly involve the primal variable y and solve the coupled elliptic second-order system

$$(7.1) \quad \begin{bmatrix} D & -\Lambda^* \iota^* \\ \iota \Lambda & \frac{1}{\gamma} B + \gamma \tilde{\iota}^* G_M(v) \end{bmatrix} \begin{bmatrix} \delta_y \\ \delta_v \end{bmatrix} = \begin{bmatrix} 0 \\ -\iota \Lambda y + \iota[\psi, 0] - \frac{1}{\gamma} Bv - \tilde{\iota}^* M(v) \end{bmatrix},$$

where

$$M(z, q) = \begin{bmatrix} [\hat{\mu} + \gamma z]^+ \\ [\hat{\nu} + \gamma(|q|_2 - \beta_\gamma)]^+ \mathbf{q}(q) \end{bmatrix},$$

$$G_M(z, q) = \begin{bmatrix} \chi_{z_\gamma(z)} & 0 \\ 0 & \chi_{\mathcal{Q}_\gamma(q)} \mathfrak{M}(q) \end{bmatrix}.$$

In the following numerical examples $\Omega \subset \mathbb{R}^2$ is polygonal and Γ_c is a line segment. We are thus able to choose option (R2) for the Tichonov regularization.

Discretization. We employ a conforming finite element method to solve (7.1) numerically. Let (\mathcal{T}_h) be a regular triangulation of Ω with mesh width h and $(\mathcal{S}_{h,c})$

a partition of Γ_c into line segments with maximal length h_c . Define for $\Gamma \subset \partial\Omega$ the discrete function spaces

$$\begin{aligned} P_{1,h}^\Gamma(\Omega) &= \{u \in L^\infty(\Omega) : u|_T \in P_1 \forall T \in \mathcal{T}_h, u|_\Gamma = 0 \text{ a.e.}\} \cap C^0(\bar{\Omega}) \\ P_{0,h}(\Omega) &= \{u \in L^\infty(\Omega) : u|_T \in P_0 \forall T \in \mathcal{T}_h\} \\ P_{1,h_c}^{\partial\Gamma_c}(\Gamma_c) &= \{u \in L^\infty(\Gamma_c) : u|_S \in P_1 \forall S \in \mathcal{S}_{h_c}, u|_{\partial\Gamma_c} = 0\} \cap C^0(\bar{\Gamma}_c), \end{aligned}$$

where P_k denotes the set of polynomials of total degree less than or equal k . We omit the superscript Γ whenever Γ has vanishing surface measure.

Note that equation (7.1) is posed in $[Y \times (H_1 \times H_2)]^*$. In the discrete setting we test (7.1) in the space

$$Y_h \times (H_{1,h_c} \times H_{2,h}) := ([P_{1,h}^\Gamma(\Omega)]^2 \times [P_{0,h}(\Omega)]^2) \times (P_{1,h_c}^{\partial\Gamma_c}(\Gamma_c) \times [P_{1,h}^{\partial\Omega}(\Omega)]^2)$$

and arrive at the following discrete algorithm:

Algorithm 7.1 (Globalized SSN).

- (i) Fix $\gamma, h, h_c > 0$ and $\text{tol} > 0$, set $j := 0$.
- (ii) Initialize dual variables: $v^0 \in H_{1,h_c} \times H_{2,h}$.
- (iii) Initialize primal variables: $y^0 \in Y_h$ by solving $Dy^0 = \Lambda^* \iota^* v^0 - l$.
- (iv) While the stopping criterion $\|\Psi_\gamma(\lambda^j)\|_{H_{1,h_c}^* \times H_{2,h}^*} < \text{tol}$ not fulfilled
 - (a) solve (7.1) in $[Y_h \times (H_{1,h_c} \times H_{2,h})]^*$,
 - (b) compute step size $\alpha^j > 0$ with Armijo line search based on
$$\alpha \mapsto J_\gamma^*(v^j + \alpha \delta_v^j),$$
 - (c) update $[y^{j+1}, v^{j+1}] := [y^j + \alpha^j \delta_y^j, v^j + \alpha^j \delta_v^j]$.
- (v) Set $j := j + 1$ and return to step (iv).

In virtue of the definition (R2) we compute the discrete norm $\|\cdot\|_{H_{1,h_c}^* \times H_{2,h}^*}$ by solving the corresponding Poisson problems. For the implementation of the operator D we incorporate the zero-trace condition in the definition of the space Q_0 using the parametrization P defined in (3.1).

Algorithmic details. The L^2 -scalar products involving the nonsmooth terms $\mathbf{m}(q)$ and $G_{\mathbf{m}}(q)$ for $q \in H_{2,h}$ for the discrete version of the Newton system as well as the right hand side of (7.1) are computed by using Gaussian quadrature. In our numerical tests we observed that the iteration count typically decreases as the number of quadrature points increases. The results described below are obtained for a fixed number of 400 quadrature points per triangle regardless of the mesh width, where numerical integration over the set of triangles is computed in parallel. The stopping criterion for the semismooth Newton iteration in Algorithm 7.1 is set to $\text{tol} = 10^{-10}$.

We further verify the theoretical superlinear rate of convergence experimentally by investigating the convergence quotients Q_j ,

$$Q_j = \frac{\|v_\gamma^{(\omega-5+j)} - v_\gamma^*\|_H}{\|v_\gamma^{(\omega-6+j)} - v_\gamma^*\|_H},$$

for $j = 1 \dots, 5$, where ω denotes the iteration count for Algorithm 7.1 and v_γ^* denotes the solution obtained by applying the same algorithm with higher precision threshold $\text{tol} = 10^{-14}$.

Example (a) - L-shape. The geometry and material parameters of the first example are taken from [41]. We consider an L-shaped domain $\Omega = (0, 0.5] \times (0.5, 1) \cup (0.5, 1) \times (0, 1)$ and assume that the elastic behavior of the material is described by $\mathbb{C}p = \mu_1 \operatorname{tr}(p)I + 2\mu_2 p$ with $\mu_1 = \mu_2 = 1.0\text{e}03$. It is further assumed that the material obeys the kinematic hardening law. The plastic material parameters are given as follows: $\sigma_y = 1.25$, $\mathbb{H} = 100 \cdot I$. The body shall be fixed at $\Gamma_d = (0.5, 1) \times \{0\}$. We set $\psi \equiv 1.0\text{e}-01$ on $\Gamma_c = (0, 1) \times \{1\}$ and apply a pressure $g(x) = -1.0\text{e}01 \cdot n(x)$ on $\Gamma_n = (0, 0.5) \times \{0.5\}$ which leads to a nonempty contact region at the solution.

We further admit zero initial conditions: $\alpha_0 \equiv 0, p_0 \equiv 0$ and vanishing volume force $f \equiv 0$. For each discretization level and fixed γ we use as starting point the (prolongated) solution of the corresponding problem on a coarse uniform mesh with 417 nodes. To verify mesh-independent convergence of Algorithm 7.1, we compute the solution for each fixed γ on meshes with decreasing mesh width, cf. Table 1, using uniform mesh refinement. Thereby the partition \mathcal{S}_{h_c} of Γ_c is defined by those mesh nodes that lie on the contact boundary Γ_c . Considering Table 1 we clearly observe a mesh-independent convergence behavior for fixed γ as the mesh width tends to zero. Slight variations in the iterations count may occur due to errors arising from numerical integration, the ill-conditioning of the system matrix for high γ -values and a too coarse refinement of Γ_c . Mesh-independent convergence can further be verified considering the convergence quotients Q_j which stabilize with increasing mesh width, cf. Figure 4. We further stress that in this example the iterations count can be decreased considerably by applying nested iterations using the solution as a starting point for the algorithm on the next finer mesh, cf. the lower section of Table 1. As a result of applying a heuristic inexact path-following approach with regard to the penalty parameter γ (cf. below), we display in Figure 2 the approximate optimal plastic strain as well as the regions of extensive plastic straining in the deformed configuration. Employing relation (3.9), we also plot the approximate yield function in Figure 3.

Table 1. Algorithm 7.1, Example (a), $\text{tol}=1.0\text{e}-10$: no. of iterations w.r.t. mesh size and γ , * nested iterations

γ / no. of nodes	833	1601	3201	6273	12545	24833
1.0e00	1	1	1	1	1	1
1.0e01	1	1	1	1	1	1
1.0e02	1	1	1	1	1	1
1.0e03	8	7	9	9	10	11
1.0e04	17	21	19	22	23	26
1.0e05	73	80	86	77	80	83
1.0e05*	73	86	65	31	24	25
1.0e06*	193	223	224	272	226	180

Example (b) - Screw wrench. In this example we consider an elasto-plastic screw wrench whose geometry can be extracted from Figure 5. The elastic behavior is described by $\mathbb{C}p = \mu_1 \operatorname{tr}(p)I + 2\mu_2 p$ with $\mu_1 = 1.15\text{e}01, \mu_2 = 7.69\text{e}00$. The material is assumed to satisfy the isotropic hardening law and the plasticity parameters are given as follows: $\sigma_y = 2\text{e}-01, H = 1.0\text{e}-03, k_1 = 1.0\text{e}07$. Moreover, $\Gamma_d := (0, 1) \times \{2\} \cup (0, 1) \times \{3\}$, and $\Gamma_c := (0, 1) \times \{4\}$ with $\psi \equiv 1.0\text{e}00$, such that the contact constraint can be expected to be inactive at the solution. We apply

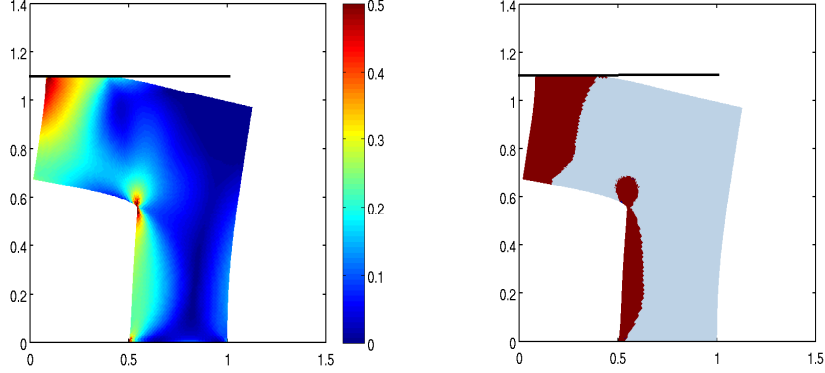


Figure 2. Example (a): $|p_\gamma|_F$ for $\gamma \approx 1.2e09$ (left), plastic zones (dark), i.e. $|p_\gamma|_F > 0.1$ (right)

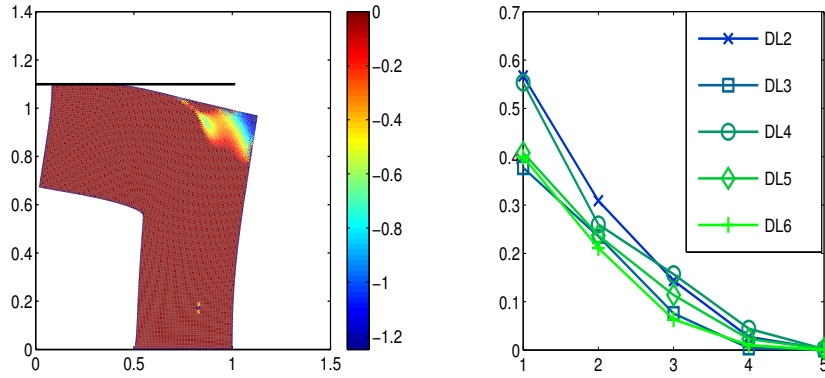


Figure 3. Example (a): yield functional for $\gamma \approx 1.2e09$

Figure 4. Example (a): $Q_j, j = 1, \dots, 5$, for $\gamma = 1.0e05$

a pressure $g(x) := -6.0e-03 \cdot n(x)$ on $\Gamma_n = \text{conv}(\{(5, 2.6), (8, 2)\})$. Further, we admit zero initial conditions: $\alpha_0 \equiv 0, p_0 \equiv 0$ and a vanishing volume force $f \equiv 0$. The results obtained by Algorithm 7.1 are summarized in Table 2 and have been computed by initializing trivially with $v^0 \equiv 0$ and using uniform meshes as in Example (a). As the iterations count stabilizes with decreasing mesh width we observe a mesh-independent convergence behavior for each fixed γ as well as local superlinear convergence.

Inexact Path-Following. In order to study convergence with regard to the regularization-penalization-parameter γ we implement a heuristic version of the inexact path-following (IPF) approach designed for the obstacle problem [25] and test it on Example (b). Starting from an increasing sequence (γ_k) , each subproblem (D_{γ_k}) is only solved approximately with increasing precision. The computed

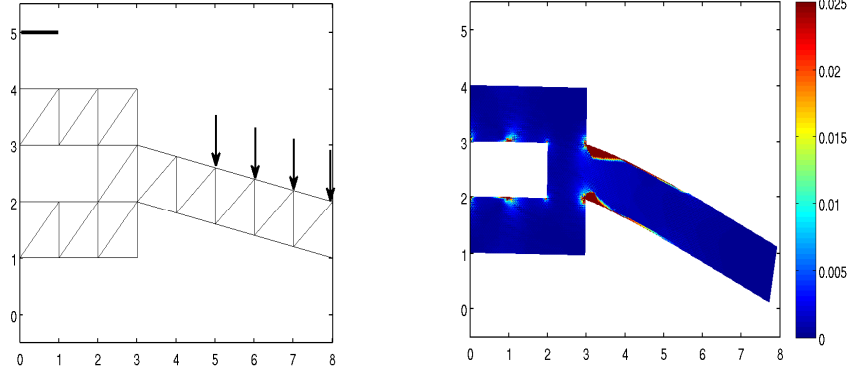


Figure 5. Example (b): initial configuration (*left*), $|p_\gamma|_F$ for $\gamma \approx 8.24e06$ (*right*)

γ / no. of nodes	873	1641	3281	6353	12705	24927
1.0e01	2	2	2	2	2	2
1.0e02	7	7	7	7	7	7
1.0e03	12	12	13	11	13	14
1.0e04	25	22	26	22	20	21
1.0e05	43	41	49	39	36	34
1.0e06	56	61	75	69	68	70

Table 2. Algorithm 7.1, Example (b), $\text{tol}=1.0e-10$: no. of iterations w.r.t. mesh size and γ

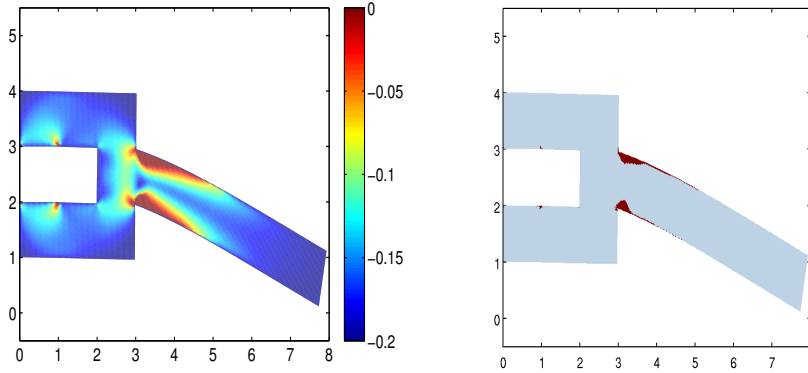


Figure 6. Example (b): yield functional for $\gamma \approx 8.24e06$ (*left*), plastic zones (*dark*), i.e. $|p_\gamma|_F > 1e-02$ (*right*)

approximate solution $\tilde{v}_{\gamma_k} \approx v_{\gamma_k}$ is used as a starting point for the subsequent problem ($D_{\gamma_{k+1}}$). In this way the effort of approximately solving the subproblems can be expected to be kept rather low. Differently from [25] we are testing a constant augmentation of γ on fixed meshes. In order to account for the discretization error in the residual $\|\Psi_\gamma(\lambda^j)\|_{H_{1,h_c}^* \times H_{2,h}^*}$ for approximating the corresponding function space quantity, a mesh-dependent stopping criterion for the inner iteration is used. To design a heuristic outer break criterion we keep track of the change in the

primal-dual-path value functional

$$V_\gamma = J_\gamma^*(\tilde{z}_\gamma, \tilde{q}_\gamma)$$

and stop the outer iteration whenever the relative change in V_γ drops below a certain threshold.

Algorithm 7.2 (IPF).

- (i) Fix $\gamma_0, h, h_c, \tau_{\text{out}}, \tau_{\text{in}} > 0, \theta > 1$; set $k := 0$.
- (ii) Initialize dual variables: $v^0 \in H_{1,h_c} \times H_{2,h}$.
- (iii) Initialize primal variables: $y^0 \in Y_h$ by solving $Dy^0 = \Lambda^* l^* v^0 - l$.
- (iv) While $\frac{|V_{\gamma_{k-1}} - V_{\gamma_k}|}{|V_{\gamma_{k-2}}|} < \tau_{\text{out}} \cdot h$ is not fulfilled
 - (a) apply Algorithm 7.1 to find $\tilde{\lambda}_{\gamma_k}$ (and \tilde{v}_{γ_k}) with

$$\|\Psi_{z,\gamma_k}(\tilde{\lambda}_{\gamma_k})\|_{H_{1,h_c}^*} < \max(h_c, \frac{\tau_{\text{in}}}{\sqrt{\gamma_k}}) \quad \text{and}$$

$$\|\Psi_{q,\gamma_k}(\tilde{\lambda}_{\gamma_k})\|_{H_{2,h}^*} < \max(h, \frac{\tau_{\text{in}}}{\sqrt{\gamma_k}}),$$
 where $\Psi_{\gamma_k} = [\Psi_{z,\gamma_k}, \Psi_{q,\gamma_k}]$,
 - (b) update $\gamma_{k+1} := \gamma_k \cdot \theta$.
- (v) Set $k := k + 1$ and return to step (ii) with $v^0 = \tilde{v}_{\gamma_k}$.

Table 3 shows the results for the application of Algorithm 7.2 to Example (b). The initial value v^0 for problem (D_{γ_0}) is simply $v^0 \equiv 0$. We observe that the number of outer iterations increases only slightly as the number of nodes increases, whereas the average number of inner iterations for each fixed discretization level appears to stay within reasonable bounds as γ_k tends to infinity. Algorithm 7.2 with parameters specified in Table 3 applied to the finest discretization level is stopped at $\gamma_{\text{end}} \approx 8.24\text{e}06$. It should be pointed out that a straightforward application of Algorithm 7.1 to $(D_{\gamma_{\text{end}}})$, requiring the same precision as in Algorithm 7.2 step (iv),(a), requires more than double the iterations which shows the advantage of our path-following approach. In Figure 5 (right), the approximate optimal plastic strain in the deformed configuration is plotted. The corresponding yield function as well as the regions of extensive plastic straining are shown in Figure 6.

no. of nodes	873	1641	3281	6353	12705	24927
iterations	6(21)	6(23)	7(25)	7(41)	8(48)	8(72)

Table 3. No. of outer(total inner) iterations for the Inexact Path-Following Algorithm 7.2, $\gamma_0 = 1.0\text{e}01$, $\theta = 7.0\text{e}00$, $\tau_{\text{out}} = 7.0\text{e}01$ and $\tau_{\text{in}} = 1.0\text{e}01$

While the stopping rule in the while-loop of step (iv) of Algorithm 7.2 is critical in general, it seems to be consistent in that other choices of τ_{in} in conjunction with a factor θ of the same order of magnitude leads to similar values for γ_{end} . The choice of the different tolerances in Algorithm 7.2 was following suggestions from solving variational inequalities of the first kind by inexact path-following strategies, cf. [25, 43]. In general, one may safeguard this stopping rule by checking the residual upon termination of the while-loop. In case an appropriate norm of the residual is considered too large, then one reduces τ_{out} , e.g. by a multiplicative factor in $(0, 1)$,

and return to step (iv) of Algorithm 7.2. This is repeated until the residual norm drops below a given tolerance.

Outlook. In both examples the number of inner iterations increases considerably with the order of magnitude of γ . A suitable path-following strategy leading to an automated regularization-discretization update procedure promises a higher efficiency compared to the heuristic used in Algorithm 7.2. For variational inequalities of the first kind these methods are already well established and prove to be remarkably efficient; see e.g. [25]. In this regard alternative choices for the coupling of the parameter γ for both Moreau-Yosida regularizations and the Tichonov regularization may be preferable.

It should be pointed out that the approach presented in this paper can be extended to contact problems with Tresca friction. These problems are characterized by an additional weighted $L^1(\Gamma_c)$ -norm functional resulting in an additional inequality in the dual problem.

APPENDIX A. POINCARÉ-FRIEDRICHS-INEQUALITY ON RIEMANNIAN MANIFOLDS

Theorem A.1. *Let (\bar{M}, g) be a compact smooth manifold with smooth boundary $\partial\bar{M}$ and $M = \bar{M} \setminus \partial\bar{M}$ a (boundaryless) smooth manifold with corresponding Riemannian metric g and Riemannian measure $\mu = \mu(g)$. Let $p \geq 1$. Then there exists a $C > 0$ such that:*

$$\int_M |u|^p \, d\mu \leq C \int_M |\nabla u|^p \, d\mu \quad \forall u \in W_0^{1,p}(M),$$

where $|\nabla u| := g(\nabla u, \nabla u)^{1/2}$ on the tangent bundle of M .

Proof. We propose a slight variation of [22, Theorem 2.10]. The result can be proven analogously to the Euclidean case. By a scaling argument the proof reduces to showing

$$\inf_{u \in U} \int_M |\nabla u|^p \, d\mu > 0,$$

where $U := \{u \in W_0^{1,p}(M) : \|u\|_{L^p(M)} = 1\}$ is a closed subspace of $W_0^{1,p}(M)$. Now assume that there exists a sequence $(u_k) \subset U$ with

$$(A.1) \quad \lim_k \int_M |\nabla u_k|^p \, d\mu \rightarrow 0.$$

Consequently, the sequence (u_k) is bounded in $W_0^{1,p}(M)$. The Rellich-Kondrakhov Theorem for Sobolev spaces on Riemannian manifolds, cf. [5, chap. 2, Theorem 2.34] or [44, chap. 4, Proposition 4.4], ensures that (w.l.o.g. the entire sequence) (u_k) converges strongly (in $L^p(M)$) to an element $u_0 \in L^p(M)$. By (A.1) we even have $u_0 \in W_0^{1,p}(M)$, where $\nabla u_0 \equiv 0$. This implies $\nabla(u_0 \circ \varphi_\alpha^{-1}) \equiv 0$ on $\varphi_\alpha(\Omega_\alpha)$ where $(\Omega_\alpha, \varphi_\alpha)_\alpha$ is an appropriate atlas of $M = \cup_\alpha \Omega_\alpha$. This in turn implies $u_0 \equiv 0$ μ -a.e. on M . But since U is closed, it holds $\|u_0\|_{L^p(M)} = 1$, a contradiction. \square

APPENDIX B. PROOF OF THEOREM 4.2

Proof. Throughout the proof we use $K > 0$ as a constant which may take different values on different occasions.

Step 1: $(q_\gamma) \subset L^2(\Omega)^d$ is bounded.

Multiplying (OC1 $_{\gamma}$) by v_{γ} yields

$$\langle \Lambda^* \iota^* v_{\gamma}, D^{-1} \Lambda^* \iota^* v_{\gamma} \rangle_{(Y^*, Y)} + \frac{1}{\gamma} b(v_{\gamma}, v_{\gamma}) - (\hat{w}, v_{\gamma})_W + (\lambda_{\gamma}, v_{\gamma})_W = 0.$$

This induces

$$\langle \Lambda^* \iota^* v_{\gamma}, D^{-1} \Lambda^* \iota^* v_{\gamma} \rangle_{(Y^*, Y)} + (\lambda_{\gamma}, v_{\gamma})_W \leq \|\hat{q}\|_{L^2(\Omega)^d} \|q_{\gamma}\|_{L^2(\Omega)^d} + \iota_1^*(z_{\gamma})(\hat{z}),$$

and thus

$$(B.1) \quad \kappa \|\Lambda^* \iota^* v_{\gamma}\|_{Y^*}^2 + (\mu_{\gamma}, z_{\gamma})_{L^2(\Gamma_c)} + (\nu_{\gamma}, q_{\gamma})_{L^2(\Omega)^d} \leq K(\|z_{\gamma}\|_{Z^*} + \|q_{\gamma}\|_{L^2(\Omega)^d}).$$

Using the domain decomposition approach for Ω from [14] we get

$$\begin{aligned} \|q_{\gamma}\|_{L^2(\Omega)^d}^2 + \frac{1}{\gamma} \kappa \|\Lambda^* \iota^* v_{\gamma}\|_{Y^*}^2 + \frac{1}{\gamma} (\mu_{\gamma}, z_{\gamma})_{L^2(\Gamma_c)} \\ \leq \frac{K}{\gamma} (\|z_{\gamma}\|_{Z^*} + \|q_{\gamma}\|_{L^2(\Omega)^d}) + K \|q_{\gamma}\|_{L^2(\Omega)^d}, \end{aligned}$$

which implies

$$(B.2) \quad \|q_{\gamma}\|_{L^2(\Omega)^d}^2 + \frac{\kappa}{\gamma \|\Lambda_1^{-*}\|} \|z_{\gamma}\|_{Z^*}^2 + \frac{1}{\gamma} (\mu_{\gamma}, z_{\gamma})_{L^2(\Gamma_c)} \leq \frac{K}{\gamma} \|z_{\gamma}\|_{Z^*} + K \|q_{\gamma}\|_{L^2(\Omega)^d}.$$

Next we consider the boundary term $(\mu_{\gamma}, z_{\gamma})_{L^2(\Gamma_c)}$:

$$\begin{aligned} (\mu_{\gamma}, z_{\gamma})_{L^2(\Gamma_c)} &= (\mu_{\gamma}, \frac{1}{\gamma} \hat{\mu} + z_{\gamma} - \frac{1}{\gamma} \hat{\mu})_{L^2(\Gamma_c)} \\ &= \frac{1}{\gamma} \|\mu_{\gamma}\|_{L^2(\Gamma_c)}^2 - \frac{1}{\gamma} (\mu_{\gamma}, \hat{\mu})_{L^2(\Gamma_c)} \\ &= \frac{1}{2\gamma} \|\mu_{\gamma} - \hat{\mu}\|_{L^2(\Gamma_c)}^2 + \frac{1}{2\gamma} \|\mu_{\gamma}\|_{L^2(\Gamma_c)}^2 - \frac{1}{2\gamma} \|\hat{\mu}\|_{L^2(\Gamma_c)}^2 \\ (B.3) \quad &\geq -\frac{1}{2\gamma} \|\hat{\mu}\|_{L^2(\Gamma_c)}^2. \end{aligned}$$

Consequently we obtain from (B.2)

$$\|q_{\gamma}\|_{L^2(\Omega)^d}^2 + \frac{1}{\gamma K} \|z_{\gamma}\|_{Z^*}^2 \leq K + \frac{K}{\gamma} (\|z_{\gamma}\|_{Z^*}) + K \|q_{\gamma}\|_{L^2(\Omega)^d},$$

from which we conclude that (q_{γ}) is bounded in $L^2(\Omega)^d$, and $(\frac{1}{\sqrt{\gamma}} z_{\gamma})$ is bounded in Z^* .

Step 2: $(z_{\gamma}) \subset Z^*$ is bounded.

Reconsidering (B.1) we have

$$\kappa \|\Lambda^* \iota^* v_{\gamma}\|_{Y^*}^2 + (\mu_{\gamma}, z_{\gamma})_{L^2(\Gamma_c)} + \underbrace{(\nu_{\gamma}, q_{\gamma})_{L^2(\Omega)^d}}_{\geq 0} \leq K(\|z_{\gamma}\|_{Z^*} + \|q_{\gamma}\|_{L^2(\Omega)^d}).$$

Similarly to the above estimates, we get

$$\frac{1}{K} \|z_{\gamma}\|_{Z^*}^2 \leq K + K \|z_{\gamma}\|_{Z^*} + K \|q_{\gamma}\|_{L^2(\Omega)^d},$$

which yields that (z_{γ}) is bounded in Z^* according to step 1.

Step 3: $[\mu_{\gamma}, \nu_{\gamma}] \subset H_1^* \times H_2^*$ is bounded.

Using the results from step 1, we again multiply (OC1 $_{\gamma}$) by v_{γ} to obtain

$$\frac{1}{\gamma} b(v_{\gamma}, v_{\gamma}) + \underbrace{([\mu_{\gamma}, \nu_{\gamma}], v_{\gamma})_W}_{\geq -K} \leq \underbrace{K \|q_{\gamma}\|_{L^2(\Omega)^d} + K \|z_{\gamma}\|_{Z^*}}_{\leq K},$$

and thus

$$(B.4) \quad \left\| \frac{1}{\sqrt{\gamma}} v_{\gamma} \right\|_H \leq K.$$

Taking the $\|\cdot\|_{H^*}$ -norm in (OC1 $_{\gamma}$) yields

$$\begin{aligned} \|[\mu_{\gamma}, \nu_{\gamma}]\|_{H^*} &\leq \|\iota \hat{w}\|_{H^*} + \|\iota \Lambda D^{-1} \Lambda^* \iota^* v_{\gamma}\|_{H^*} + \frac{1}{\gamma} \|B v_{\gamma}\|_{H^*} \\ &\leq K + \|\iota \Lambda D^{-1} \Lambda^*\| \|\iota^* v_{\gamma}\|_{Z^* \times L^2(\Omega)^d} + \frac{1}{\gamma} \|B\| \|v_{\gamma}\|_H. \end{aligned}$$

Taking account of step 1 and 2 as well as (B.4), this proves the assertion.

We thus have

$$[z_\gamma, q_\gamma, \mu_\gamma, \nu_\gamma] \rightharpoonup [\tilde{z}, \tilde{q}, \tilde{\mu}, \tilde{\nu}] \in Z^* \times L^2(\Omega)^d \times H_1^* \times H_2^*,$$

for an appropriate subsequence $[z_\gamma, q_\gamma, \mu_\gamma, \nu_\gamma] \subset H_1 \times H_2 \times L^2(\Gamma_c) \times L^2(\Omega)^d$, sharing the same indices by abuse of notation.

Step 4: $\tilde{v} := [\tilde{z}, \tilde{q}]$ is feasible, i.e. (OC2) holds.

With step 1 and 2 it is easily seen that

$$F^*(\Lambda^* \iota^*[z_\gamma, q_\gamma]) + \underbrace{T_\gamma(q_\gamma)}_{\geq 0}$$

is bounded from below. Moreover, we have

$$\begin{aligned} J_\gamma^*(z_\gamma, q_\gamma) &\leq J_\gamma^*(z, q) \\ &= F^*(\Lambda^* \iota^*[z, q]) + \frac{1}{2\gamma} b([z, q], [z, q]) < K \end{aligned}$$

for all $[z, q] \in Z \times L^2(\Omega)^d$ with $z \leq -\frac{\hat{\mu}}{\gamma}$ a.e. in Γ_c , and $|q|_2 \leq \sigma_y - \frac{\hat{\nu}}{\gamma}$ a.e. in Ω , with γ sufficiently large. Consequently, $M_\gamma^1(z_\gamma) + M_\gamma^2(q_\gamma)$ is bounded.

In a similar fashion as in [14], we exploit the weak lower semicontinuity of

$$L^2(\Omega)^d \ni q \mapsto \| [|q|_2 - \beta]^+ \|_{L^2(\Gamma_c)}^2 \in \mathbb{R}$$

to conclude that

$$\begin{aligned} 0 &= \int_\Omega ([|q|_2 - \beta]^+)^2 dx \leq \liminf_\gamma \int_\Omega ([|q_\gamma|_2 - \beta]^+)^2 dx \\ &\leq \liminf_\gamma \int_\Omega ([|q_\gamma|_2 - (\beta_\gamma - \frac{\hat{\nu}}{\gamma})]^+)^2 dx \longrightarrow 0, \end{aligned}$$

since $M_\gamma^2(q_\gamma) = \frac{\gamma}{2} \| [|q_\gamma|_2 - (\beta_\gamma - \frac{\hat{\nu}}{\gamma})]^+ \|_{L^2(\Omega)}^2$ is bounded.

Furthermore for $z \in Z, z \geq 0$, we obtain

$$\begin{aligned} \langle [z_\gamma + \frac{\hat{\mu}}{\gamma}]^+, z \rangle_{(Z^*, Z)} &= \langle [z_\gamma + \frac{\hat{\mu}}{\gamma}]^+, z \rangle_{L^2(\Gamma_c)} \\ &\geq \langle z_\gamma, z \rangle_{L^2(\Gamma_c)} + \frac{1}{\gamma} \langle \hat{\mu}, z \rangle_{L^2(\Gamma_c)} \rightarrow \langle \tilde{z}, z \rangle_{(Z^*, Z)}. \end{aligned}$$

On the other hand

$$\begin{aligned} \langle [z_\gamma + \frac{\hat{\mu}}{\gamma}]^+, z \rangle_{(Z^*, Z)} &\leq \| [z_\gamma + \frac{\hat{\mu}}{\gamma}]^+ \|_{Z^*} \| z \|_Z \\ &\leq K \| [z_\gamma + \frac{\hat{\mu}}{\gamma}]^+ \|_{L^2(\Gamma_c)} \| z \|_Z \rightarrow 0, \end{aligned}$$

as $L^2(\Gamma_c) \xrightarrow{\iota_1^*} Z^*$ and by the boundedness of $M_\gamma^1(z_\gamma)$. This accomplishes step 4.

Step 5: (OC1) is satisfied.

For $v \in H$, (OC1 $_\gamma$) reads

$$\begin{aligned} 0 &= \langle \iota^* v_\gamma, \Lambda D^{-1} \Lambda^* \iota^* v \rangle_{(Z^* \times L^2(\Omega)^d, Z \times L^2(\Omega)^d)} + \frac{1}{\gamma} b(v_\gamma, v) \\ &\quad - (\iota \hat{w}, v)_W + ([\mu_\gamma, \nu_\gamma], v)_W. \end{aligned}$$

Passing to the limit as $\gamma \rightarrow +\infty$ yields for $v \in H$

$$\begin{aligned} 0 &= \langle \Lambda D^{-1} \Lambda^* [\tilde{z}, \tilde{q}], \iota^* v \rangle_{(Z \times L^2(\Omega)^d, Z^* \times L^2(\Omega)^d)} - (\iota \hat{w}, v)_W + \langle [\tilde{\mu}, \tilde{\nu}], v \rangle_{(H^*, H)} \\ &= \langle \Lambda D^{-1} \Lambda^* [\tilde{z}, \tilde{q}], v \rangle_W - (\iota \hat{w}, v)_W + \langle [\tilde{\mu}, \tilde{\nu}], v \rangle_{(H^*, H)}. \end{aligned}$$

From the density of H in W we infer

$$-[\tilde{\mu}, \tilde{\nu}] = \Lambda D^{-1} \Lambda^* [\tilde{z}, \tilde{q}] - \Lambda D^{-1} l - [\psi, 0],$$

and thus (OC1).

Step 6: It holds that $\Lambda^* \iota^* v_\gamma \rightarrow \Lambda^* [\tilde{z}, \tilde{q}]$ in Y^* .

By the weak lower semicontinuity of $F^*(\Lambda^* \cdot)$ we have

$$\liminf_{\gamma \rightarrow +\infty} F^*(\Lambda^* \iota^* [z_\gamma, q_\gamma]) \geq F^*(\Lambda^* [\tilde{z}, \tilde{q}]).$$

On the other hand, exploiting the minimality of $[z_\gamma, q_\gamma]$, we obtain for any $[z, q] \in H$ with $z \leq 0, |q|_2 \leq \beta$,

$$\begin{aligned} F^*(\Lambda^* \iota^* [z_\gamma, q_\gamma]) &\leq F^*(\Lambda^* \iota^* [z, q]) + \frac{1}{2\gamma} \|\hat{\mu} + \gamma z\|_{L^2(\Gamma_c)}^2 \\ &\quad + \frac{1}{2\gamma} \|\hat{\nu} + \gamma(|q|_2 - \beta_\gamma)\|_{L^2(\Omega)}^2 + \frac{1}{2\gamma} b([z, q], [z, q]), \end{aligned}$$

such that

$$\begin{aligned} &\limsup_{\gamma \rightarrow +\infty} F^*(\Lambda^* \iota^* [z_\gamma, q_\gamma]) \\ &\leq F^*(\Lambda^* \iota^* [z, q]) + \limsup_{\gamma \rightarrow +\infty} \frac{1}{2\gamma} \|\hat{\mu}\|_{L^2(\Gamma_c)}^2 + \limsup_{\gamma \rightarrow +\infty} \frac{1}{2\gamma} \|\hat{\nu} + \gamma(\beta - \beta_\gamma)\|_{L^2(\Omega)}^2 \\ &= F^*(\Lambda^* \iota^* [z, q]) + \limsup_{\gamma \rightarrow +\infty} \frac{1}{2\gamma} \|\hat{\nu}\|_{L^2(\Omega)}^2 + \limsup_{\gamma \rightarrow +\infty} \langle \hat{\nu}, \beta - \beta_\gamma \rangle_{L^2(\Omega)} + \limsup_{\gamma \rightarrow +\infty} \frac{\gamma}{2} \|(\beta - \beta_\gamma)\|_{L^2(\Omega)}^2 \\ &= F^*(\Lambda^* \iota^* [z, q]), \end{aligned}$$

since $\|\beta - \beta_\gamma\| \leq \frac{1}{\gamma}$. By Assumption 4.1 we conclude

$$\limsup_{\gamma \rightarrow +\infty} F^*(\Lambda^* \iota^* [z_\gamma, q_\gamma]) \leq F^*(\Lambda^* [\tilde{z}, \tilde{q}]),$$

and thus

$$\lim_{\gamma \rightarrow +\infty} F^*(\Lambda^* \iota^* [z_\gamma, q_\gamma]) = F^*(\Lambda^* [\tilde{z}, \tilde{q}]).$$

The weak convergence $\iota^* [z_\gamma, q_\gamma] \rightharpoonup [\tilde{z}, \tilde{q}]$ and the ellipticity of the bilinear form associated to D^{-1} yield the assertion.

Step 7: The normal cone property (OC3) is satisfied.

Owing to (OC1) and the results of the preceding steps we obtain for all $v = [z, q] \in H$ with $z \leq 0, |q|_2 \leq \beta$,

$$\begin{aligned} \langle \tilde{\lambda}, \iota^* v - \tilde{v} \rangle &= -\langle \Lambda D^{-1} \Lambda^* \tilde{v}, \iota^* v - \tilde{v} \rangle + \langle \hat{w}, \iota^* v - \tilde{v} \rangle \\ &= \lim_{\gamma \rightarrow +\infty} \left(-\langle \iota \Lambda D^{-1} \Lambda^* \iota^* v_\gamma, v - v_\gamma \rangle + \langle \iota \hat{w}, v - v_\gamma \rangle \right) \\ &= \lim_{\gamma \rightarrow +\infty} \left(\frac{1}{\gamma} \langle B v_\gamma, v - v_\gamma \rangle + ([\mu_\gamma, \nu_\gamma], v - v_\gamma)_w \right) \\ &\leq \limsup_{\gamma \rightarrow +\infty} \frac{1}{\gamma} \|v_\gamma\|_H \|B\| \|v\| + \limsup_{\gamma \rightarrow +\infty} ([\mu_\gamma, \nu_\gamma], v - v_\gamma)_w \\ &= \limsup_{\gamma \rightarrow +\infty} ([\mu_\gamma, \nu_\gamma], v - v_\gamma)_w, \end{aligned}$$

where the last equality follows from (B.4). We further verify that

$$\begin{aligned}
(\nu_\gamma, q - q_\gamma)_{L^2(\Omega)^d} &= \gamma \int_{\Omega} [|q_\gamma|_2 - \beta_\gamma + \frac{\hat{\nu}}{\gamma}]^+ \mathbf{q}(q_\gamma)(q - q_\gamma) \, dx \\
&\leq \gamma \int_{\Omega} [|q_\gamma|_2 - (\beta_\gamma - \frac{\hat{\nu}}{\gamma})]^+ (\beta - |q_\gamma|_2) \, dx \\
&\leq \gamma \int_{\Omega} [|q_\gamma|_2 - (\beta_\gamma - \frac{\hat{\nu}}{\gamma})]^+ (\beta - (\beta_\gamma - \frac{\hat{\nu}}{\gamma})) \, dx \\
&\leq \| [|q_\gamma|_2 - (\beta_\gamma - \frac{\hat{\nu}}{\gamma})]^+ \|_{L^2(\Omega)} (\gamma \| \beta - \beta_\gamma \|_{L^2(\Omega)} + \| \hat{\nu} \|_{L^2(\Omega)}) \\
&\leq K \| [|q_\gamma|_2 - (\beta_\gamma - \frac{\hat{\nu}}{\gamma})]^+ \|_{L^2(\Omega)}.
\end{aligned}$$

Now, the boundedness of $M_\gamma^2(q_\gamma)$ implies that

$$\lim_{\gamma \rightarrow +\infty} \| [|q_\gamma|_2 - (\beta_\gamma - \frac{\hat{\nu}}{\gamma})]^+ \|_{L^2(\Omega)} \rightarrow 0.$$

Likewise, it holds that

$$\begin{aligned}
(\mu_\gamma, z - z_\gamma)_{L^2(\Gamma_c)} &\leq -(\mu_\gamma, z_\gamma)_{L^2(\Gamma_c)} = -\gamma ([z_\gamma + \frac{\hat{\mu}}{\gamma}]^+, z_\gamma + \frac{\hat{\mu}}{\gamma})_{L^2(\Gamma_c)} + \frac{1}{\gamma} (\mu_\gamma, \hat{\mu})_{L^2(\Gamma_c)} \\
&\leq \frac{1}{\gamma} \| \mu_\gamma \|_{L^2(\Gamma_c)} \| \hat{\mu} \|_{L^2(\Gamma_c)} \rightarrow 0 \text{ for } \gamma \rightarrow +\infty,
\end{aligned}$$

by the boundedness of $M_\gamma^1(z_\gamma)$. Consequently we obtain $\langle \tilde{\lambda}, \iota^* v - \tilde{v} \rangle \leq 0$ for all $v = [z, q] \in H$ with $z \leq 0, |q|_2 \leq \beta$. In virtue of Assumption 4.1 a density argument completes the proof. \square

REFERENCES

- [1] R.A. Adams and J.J.F. Fournier, *Sobolev Spaces*, Pure and Applied Mathematics, Elsevier Science, 2003.
- [2] J. Albery and C. Carstensen, *Numerical analysis of time-dependent primal elastoplasticity with hardening*, SIAM Journal on Numerical Analysis **37** (2000), no. 4, 1271–1294.
- [3] J. Albery, C. Carstensen, and D. Zarrabi, *Adaptive numerical analysis in primal elastoplasticity with hardening*, Computer Methods in Applied Mechanics and Engineering **171** (1999), 175–204.
- [4] J.P. Aubin and I. Ekeland, *Applied Nonlinear Analysis*, Dover Books on Mathematics Series, Dover Publications, 2006.
- [5] T. Aubin, *Some Nonlinear Problems in Riemannian Geometry*, Springer Berlin Heidelberg, 1998.
- [6] L. Badaea and R. Krause, *One- and two-level Schwarz methods for variational inequalities of the second kind and their application to frictional contact*, Numerische Mathematik **120** (2012), no. 4, 573–599.
- [7] J.-F. Bonnans, J.C. Gilbert, C. Lemaréchal, and C.A. Sagastizábal, *Numerical Optimization*, 2nd ed., Springer, 2006.
- [8] C. Carstensen, *Domain decomposition for a non-smooth convex minimization problem and its application to plasticity*, Numerical Linear Algebra with Applications **4** (1997), no. 3, 177–190.

- [9] C. Carstensen, R. Klose, and A. Orlando, *Reliable and efficient equilibrated a posteriori error finite element error control in elastoplasticity and elastoviscoplasticity with hardening*, *Comput. Methods Appl. Mech. Engrg.* **195** (2006), no. 19-22, 2574–2598.
- [10] C. Carstensen, A. Orlando, and J. Valdman, *A convergent adaptive finite element method for the primal problem of elastoplasticity*, *J. Numer. Meth. Engrg* (2005), pp. 1851–1887.
- [11] X. Chen, Z. Nashed, and L. Qi, *Smoothing methods and semismooth methods for nondifferentiable operator equations*, *SIAM Journal on Numerical Analysis* **38** (2001), no. 4, pp. 1200–1216.
- [12] P. W. Christensen, *A semi-smooth Newton method for elasto-plastic contact problems*, *International Journal of Solids and Structures* **39** (2002), no. 8, 2323 – 2341.
- [13] M. Cocu, E. Pratt, and M. Raous, *Formulation and approximation of quasistatic frictional contact*, *International Journal of Engineering Science* **34** (1996), no. 7, 783 – 798.
- [14] J.C. De Los Reyes and M. Hintermüller, *A duality based semismooth Newton framework for solving variational inequalities of the second kind*, *Interfaces and Free Boundaries* **13** (2011), 437–462.
- [15] G. Duvaut and J.-L. Lions, *Inequalities in Mechanics and Physics*, Springer, Berlin, 1976.
- [16] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, *Classics in Applied Mathematics*, Society for Industrial and Applied Mathematics, 1987.
- [17] A. Grigor’yan, *Heat Kernel and Analysis on Manifolds*, American Mathematical Society - International Press, 2009.
- [18] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman Publishing, 1985.
- [19] P. Gruber and J. Valdman, *Solution of one-time-step problems in elastoplasticity by a slant Newton method*, *SIAM Journal on Scientific Computing* **31** (2009), no. 2, 1558–1580.
- [20] C. Hager and B. Wohlmuth, *Nonlinear complementarity functions for plasticity problems with frictional contact*, *Comput. Methods Appl. Mech. Engrg.* **198** (2009), 3411–3427.
- [21] W. Han and B.D. Reddy, *Computational plasticity: The variational basis and numerical analysis*, *Computational Mechanics Advances* **2** (1995), 283–400.
- [22] E. Hebey, *Nonlinear Analysis on Manifolds: Sobolev Spaces and Inequalities*, American Mathematical Society, 1999.
- [23] M. Hintermüller, *Mesh independence and fast local convergence of a primal-dual active-set method for mixed control-state constrained elliptic control problems*, *ANZIAM J.* **49** (2007), no. 1, 1–38.
- [24] M. Hintermüller, K. Ito, and K. Kunisch, *The primal-dual active set strategy as a semi-smooth Newton method*, *SIAM J. Optimization* **13** (2003), no. 3, 865–888.
- [25] M. Hintermüller and K. Kunisch, *Path-following methods for a class of constrained minimization problems in function space*, *SIAM Journal on Optimization* **17** (2006), no. 1, 159–187.

- [26] ———, *PDE-constrained optimization subject to pointwise constraints on the control, the state and its derivative*, SIAM Journal on Optimization **20** (2009), no. 3, 1133–1156.
- [27] M. Hintermüller and C. Rautenberg, *A sequential minimization technique for elliptic quasi-variational inequalities with gradient constraints*, SIAM Journal on Optimization **22** (2012), no. 4, 1224–1257.
- [28] M. Hintermüller and M. Ulbrich, *A mesh-independence result for semismooth Newton methods*, Math. Program., Ser. B **171** (2004), 151–184.
- [29] I. Hlaváček, J. Haslinger, J. Nečas, and J. Lovíšek, *Solution of Variational Inequalities in Mechanics*, Springer, New York, 1988.
- [30] S. Hübner, G. Stadler, and B. Wohlmuth, *A primal-dual active set algorithm for three-dimensional contact problems with Coulomb friction*, SIAM J. Sci. Comput. **30** (2008), 572–596.
- [31] C. Johnson, *Existence theorems for plasticity problems*, J. Math. pures et appl. **55** (1976), 431–444.
- [32] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and their Applications*, Academic Press, Inc., New York, 1980.
- [33] R. Kornhuber and R. Krause, *Adaptive multigrid methods for Signorini’s problem in linear elasticity*, Computing and Visualization in Science **4** (2001), no. 1, 9–20.
- [34] R. Kornhuber, R. Krause, O. Sander, P. Deuffhard, and S. Ertel, *A monotone multigrid solver for two body contact problems in biomechanics*, Computing and Visualization in Science **11** (2008), no. 1, 3–15.
- [35] K. Krabbenhoft, A. V. Lyamin, S. W. Sloan, and P. Wriggers, *An interior-point algorithm for elastoplasticity*, International Journal for Numerical Methods in Engineering **69** (2007), no. 3, 592–626.
- [36] R. Krause, *A nonsmooth multiscale method for solving frictional two-body contact problems in 2D and 3D with multigrid efficiency*, SIAM J. Sci. Comput. **31** (2009), no. 2, 1399–1423.
- [37] K. Kunisch and G. Stadler, *Generalized Newton methods for the 2D-Signorini contact problem with friction in function space*, ESAIM: Mathematical Modelling and Numerical Analysis **39** (2005), 827–854.
- [38] M. Marcus and V.J. Mizel, *Every superposition operator mapping one Sobolev space into another is continuous*, Journal of Functional Analysis **33** (1979), no. 2, 217 – 229.
- [39] J. T. Oden and N. Kikuchi, *Contact Problems in Elasticity*, Society for Industrial and Applied Mathematics, 1988.
- [40] M. Sauter and C. Wieners, *On the superlinear convergence in computational elasto-plasticity*, Computer Methods in Applied Mechanics and Engineering **200** (2011), no. 49-52, 3646 – 3658.
- [41] A. Schröder and S. Wiedemann, *Error estimates in elastoplasticity using a mixed method*, Appl. Numer. Math. **61** (2011), no. 10, 1031–1045.
- [42] J.C. Simo and T.J.R. Hughes, *Computational Inelasticity*, Springer-Verlag Berlin, 1998.
- [43] G. Stadler, *Path-following and augmented Lagrangian methods for contact problems in linear elasticity*, J. Comput. Appl. Math. **203** (2007), 533–547.
- [44] M.E. Taylor, *Partial Differential Equations I: Basic Theory*, Springer-Verlag New York, 1996.

- [45] C. Wieners, *Multigrid methods for Prandtl-Reuss plasticity*, Numerical Linear Algebra with Applications **6** (1999), no. 6, 457–478.
- [46] ———, *SQP methods for incremental plasticity with kinematic hardening*, IUTAM Symposium on Theoretical, Computational and Modelling Aspects of Inelastic Media (B. Daya Reddy, ed.), IUTAM BookSeries, vol. 11, Springer Netherlands, 2008, pp. 143–153.
- [47] J. Wloka, *Partielle Differentialgleichungen*, B.G. Teubner, Stuttgart, 1982.

MICHAEL HINTERMÜLLER, DEPARTMENT OF MATHEMATICS, HUMBOLDT-UNIVERSITÄT ZU BERLIN,
UNTER D. LINDEN 6, D-10099 BERLIN, GERMANY

E-mail address: `hint@math.hu-berlin.de`

SIMON RÖSEL, DEPARTMENT OF MATHEMATICS, HUMBOLDT-UNIVERSITÄT ZU BERLIN, UNTER
D. LINDEN 6, D-10099 BERLIN, GERMANY

E-mail address: `roesel@math.hu-berlin.de`