

OPTIMAL CONTROL OF PLASTICITY WITH INERTIA

Stephan Walther*

Abstract The paper is concerned with an optimal control problem governed by the equations of elasto plasticity with linear kinematic hardening and the inertia term at small strain. The objective is to optimize the displacement field and plastic strain by controlling volume forces. The idea given in [11] is used to transform the state equation into an evolution variational inequality (EVI) involving a certain maximal monotone operator. Results from [18] are then used to analyze the EVI. A regularization is obtained via the Yosida approximation of the maximal monotone operator, this approximation is smoothed further to derive optimality conditions for the smoothed optimal control problem.

1 INTRODUCTION

We consider the following optimal control problem governed by the equations of elasto plasticity with linear kinematic hardening and the inertia term at small strain:

$$(1.1) \quad \left\{ \begin{array}{l} \min J(u, \dot{u}, z, f) = \Psi(u, \dot{u}, z) + \frac{\alpha}{2} \|f\|_{\mathfrak{X}_c}^2, \\ \text{s.t. } \rho \ddot{u} - \operatorname{div} \mathbb{C}(\nabla^s u - z) = f, \\ \dot{z} \in A(\mathbb{C} \nabla^s u - (\mathbb{C} + \mathbb{B})z), \\ (u, \dot{u}, z)(0) = (u_0, v_0, z_0), \\ u \in H^1(H_D^1(\Omega; \mathbb{R}^d)) \cap H^2(L^2(\Omega; \mathbb{R}^d)), \\ z \in H^1(L^2(\Omega; \mathbb{R}_s^{d \times d})), \\ f \in \mathfrak{X}_c. \end{array} \right.$$

Herein, $\Omega \subset \mathbb{R}^d$ is the body under consideration with density ρ , where $d \in \mathbb{N}$ is the dimension. Its boundary is split into two disjoint parts Γ_D and Γ_N . Furthermore, $u : [0, T] \times \Omega \rightarrow \mathbb{R}^d$ is the displacement field and $z : [0, T] \times \Omega \rightarrow \mathbb{R}_s^{d \times d}$ the plastic strain. The initial data (u_0, v_0, z_0) is given and fixed. The volume force is given by $f : [0, T] \times \Omega \rightarrow \mathbb{R}^d$. The time derivative of, for instance, the plastic strain is denoted by \dot{z} and the symmetric gradient by $\nabla^s = 1/2(\nabla + \nabla^\top)$. Moreover, \mathbb{C} is the elasticity tensor and \mathbb{B} the hardening parameter. The flow rule is represented by the maximal monotone operator A , in Section 5 below we will choose the von-Mises flow rule. The control space \mathfrak{X}_c is a nonempty and closed subspace of $H^1([0, T]; L^2(\Omega; \mathbb{R}^d))$ and $\alpha > 0$ a fixed Tikhonov parameter. Note that higher time regularity of the control was also required when analyzing other problems with a non-smooth hyperbolic evolution structure, e.g. in [28]. The precise definitions and assumptions are presented in Section 2 below. Note that the problem (1.1) is formulated only with the displacement

This research was supported by the German Research Foundation (DFG) under grant number ME 3281/9-1 within the priority program Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization (SPP 1962).

*TU Dortmund, Faculty of Mathematics, Vogelpothsweg 87, 44227 Dortmund, Germany
(stephan.walther@tu-dortmund.de, <http://www.mathematik.tu-dortmund.de/lx>).

and plastic strain as state variables. The stress is normally given by $\sigma = \mathbb{C}(\nabla^s u - z)$ (and can thus be easily integrated in Ψ), however, we eliminated it in (1.1) for convenience. Additionally, the Dirichlet displacement and Neumann boundary forces (see the definition of the div-operator in Section 2) are set to zero, cf. Remark 3.8 below. Regarding a detailed description and derivation of the plasticity model, we refer to [16, 20, 22].

Let us put our work into perspective. Optimal control problems governed by plasticity were considered in [6, 15, 17–19, 23–26]. The articles [15, 17] are concerned with the static case of elasto plasticity, for further articles of the static case we refer to the references therein. For the time dependent quasi-static case we are only aware of [6, 18, 19, 23–26]. The articles [23–26] and the application in [18] are devoted to the case of elasto plasticity with linear kinematic hardening, whereas [6, 19] are concerned with the case of perfect plasticity, that is, with no hardening. In contrast, we consider the case of elasto plasticity with the inertia term (and linear kinematic hardening), that is, the second time derivative of the displacement (the acceleration) multiplied by the density ρ is present in the balance of momentum. Due to this inertia term, the equations are physically more reasonable than the quasi-static case (of course, the solution to both systems might not differ much when the acceleration of the body is small, which is the reasoning when neglecting the inertia term). As said above, the application in [18] investigates quasi-static (homogenized) plasticity with hardening. This application is analyzed by applying results concerned with an abstract optimal control problem governed by a first-order evolution variational inequality (EVI) involving a maximal monotone operator. As we will see below, the state equation in (1.1) can also be transformed into such an abstract EVI. Let us note that the existence of a solution was already proven in [11, Theorem 5.1] by using essentially the same transformation into an EVI as we will do. However, there it was transformed into a second order EVI and the maximal monotonicity of the (slightly different) operator given therein was proven in another way. In contrast, we consider a first order EVI and will provide the concrete form of the resolvent in Proposition 3.11 (which will also be used later in Section 4.2 to provide optimality conditions), the maximal monotonicity of our operator will then follow easily. Having transformed the state equation, we can apply the results from [18]. We only have to heed two differences between our EVI and the EVI analyzed in [18]. First, our maximal monotone operator does not fulfill some properties required in [18], second, the given data are more regular in time than in [18], as we will elaborate on at the end of Section 3.2. However, the better regularity in time will compensate the missing properties of our maximal monotone operator, so that the unique existence of a solution can still be shown (Theorem 3.15) and thus we can apply results from [18]. There is a large list of literature on plasticity with inertia, we only refer to [1–3, 7–9] and the references therein. However, to best of the author's knowledge, there exists no contribution to optimal control of plasticity with the inertia term, except in [27]. We emphasize that this paper is essentially based on [27, Part IV] and on the transformation idea from [11].

The paper is organized as follows. After introducing our notation and standing assumptions in Section 2, we transform the state equation in (1.1) into a first-order EVI, prove the unique existence for given data and provide regularization and convergence results in Section 3. Afterwards, in Section 4, we analyze the optimal control problem (1.1), show the existence of a global solution, provide an approximation result via a regularized problem and finally present optimality conditions.

2 NOTATION AND STANDING ASSUMPTIONS

Notation When X is a normed vector space we denote its norm by $\|\cdot\|_X$. For normed vector spaces X and Y we denote the space of linear and continuous functions on X with values in Y by $\mathcal{L}(X; Y)$. We abbreviate $\mathcal{L}(X) := \mathcal{L}(X; X)$. The dual space of X is denoted by $X^* = \mathcal{L}(X; \mathbb{R})$. The inner product of a Hilbert space H is denoted by $(\cdot, \cdot)_H$. For the whole paper, we fix the final time $T > 0$. For $t > 0$ we denote the Bochner space of square-integrable functions on the time interval $[0, t]$ by $L^2([0, t]; X)$, the Bochner-Sobolev space by $H^1([0, t]; X)$ and the space of continuous functions by $C([0, t]; X)$. We

furthermore abbreviate $L^2(X) := L^2([0, T]; X)$, $H^1(X) := H^1([0, T]; X)$ and $C(X) := C([0, T]; X)$. When $G \in \mathcal{L}(X; Y)$ is a linear and continuous operator, we can define an operator in $\mathcal{L}(L^2(X); L^2(Y))$ by $G(u)(t) := G(u(t))$ for all $u \in L^2(X)$ and for almost all $t \in [0, T]$, we denote this operator also by G , that is, $G \in \mathcal{L}(L^2(X); L^2(Y))$, and analog for Bochner-Sobolev spaces, i.e., $G \in \mathcal{L}(H^1(X); H^1(Y))$. Given a coercive and symmetric operator $G \in \mathcal{L}(H)$ in a real Hilbert space H , we denote its coercivity constant by γ_G , i.e., $(Gh, h)_H \geq \gamma_G \|h\|_H^2$ for all $h \in H$. With this operator we can define a new scalar product, which induces an equivalent norm, by $H \times H \ni (h_1, h_2) \mapsto (Gh_1, h_2)_H \in \mathbb{R}$. We denote the Hilbert space equipped with this scalar product by H_G , that is $(h_1, h_2)_{H_G} = (Gh_1, h_2)_H$ for all $h_1, h_2 \in H$. If $p \in [1, \infty]$, then we denote its conjugate exponent by p' , that is $\frac{1}{p} + \frac{1}{p'} = 1$. Throughout the paper, by $L^p(\Omega; M)$ we denote Lebesgue spaces with values in M , where $p \in [1, \infty]$ and M is a finite dimensional space. By $W^{1,p}(\Omega; M)$ we denote Sobolev spaces and $W_D^{1,p}(\Omega; M)$ is the subspace containing functions which traces are zero on Γ_D . For the dual spaces of $W^{1,p'}(\Omega; M)$ and $W_D^{1,p'}(\Omega; M)$ we write $W^{-1,p}(\Omega; M)$ and $W_D^{-1,p}(\Omega; M)$. We use the usual abbreviations $H^1(\Omega; M) := W^{1,2}(\Omega; M)$, $H_D^1(\Omega; M) := W_D^{1,2}(\Omega; M)$, $H^{-1}(\Omega; M) := W^{-1,2}(\Omega; M)$ and $H_D^{-1}(\Omega; M) := W_D^{-1,2}(\Omega; M)$. Finally, by $\mathbb{R}_s^{d \times d}$, we denote the space of symmetric matrices and $c, C > 0$ are generic constants.

STANDING ASSUMPTIONS

The following standing assumptions are tacitly assumed for the rest of the paper without mentioning them every time.

Domain The domain $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$, is bounded with Lipschitz boundary Γ . The boundary consists of two disjoint measurable parts Γ_N and Γ_D such that $\Gamma = \Gamma_N \cup \Gamma_D$. While Γ_N is a relatively open subset, Γ_D is a relatively closed subset of Γ with positive boundary measure. In addition, the set $\Omega \cup \Gamma_N$ is regular in the sense of Gröger, cf. [12].

Furthermore, the density of Ω is given by $\rho > 0$.

Coefficients The elasticity tensor and the hardening parameter satisfy $\mathbb{C}, \mathbb{B} \in \mathcal{L}(\mathbb{R}_s^{d \times d})$ and are symmetric and coercive, i.e., there exist constants $\underline{c} > 0$ and $\underline{b} > 0$ such that $(\mathbb{C}\sigma, \sigma)_{\mathbb{R}_s^{d \times d}} \geq \underline{c} \|\sigma\|_{\mathbb{R}_s^{d \times d}}^2$ and $(\mathbb{B}\sigma, \sigma)_{\mathbb{R}_s^{d \times d}} \geq \underline{b} \|\sigma\|_{\mathbb{R}_s^{d \times d}}^2$ for all $\sigma \in \mathbb{R}_s^{d \times d}$.

We abbreviate further

$$(2.1) \quad \mathbb{D} := \mathbb{B}(\mathbb{C} + \mathbb{B})^{-1}\mathbb{C} \in \mathcal{L}(\mathbb{R}_s^{d \times d}) \quad \text{and} \quad \mathbb{E} := \mathbb{C}(\mathbb{C} + \mathbb{B})^{-1} \in \mathcal{L}(\mathbb{R}_s^{d \times d})$$

and note that \mathbb{D} is symmetric and coercive, according to [11, Lemma 4.2]. Moreover, for instance, we denote the adjoint of \mathbb{E} by \mathbb{E}^\top .

Initial data We choose $u_0, v_0 \in H_D^1(\Omega; \mathbb{R}^d)$ and $z_0 \in L^2(\Omega; \mathbb{R}_s^{d \times d})$ and define $q_0 := \mathbb{C}\nabla^s u_0 - (\mathbb{C} + \mathbb{B})z_0 \in L^2(\Omega; \mathbb{R}_s^{d \times d})$. Moreover, we assume that (u_0, v_0, q_0) is an element of $D(\mathcal{A})$, where $D(\mathcal{A})$ is given in Definition 3.5.

Operators Throughout the paper, $\nabla^s := \frac{1}{2}(\nabla + \nabla^\top) : W^{1,p}(\Omega; \mathbb{R}^d) \rightarrow L^p(\Omega; \mathbb{R}_s^{d \times d})$ denotes the linearized strain. Its restriction to $W_D^{1,p}(\Omega; \mathbb{R}_s^{d \times d})$ is denoted by the same symbol and, for the adjoint of this restriction, we write $-\text{div} := (\nabla^s)^* : L^{p'}(\Omega; \mathbb{R}_s^{d \times d}) \rightarrow W_D^{-1,p'}(\Omega; \mathbb{R}^d)$.

The operator $A : L^2(\Omega; \mathbb{R}_s^{d \times d}) \rightarrow 2L^2(\Omega; \mathbb{R}_s^{d \times d})$ is maximal monotone with domain $D(A)$. Furthermore, by $A_\lambda : L^2(\Omega; \mathbb{R}_s^{d \times d}) \rightarrow L^2(\Omega; \mathbb{R}_s^{d \times d})$, $\lambda > 0$, we denote the Yosida approximation of A and by

$R_\lambda = (I + \lambda A)^{-1}$ the resolvent of A , so that $A_\lambda = \frac{1}{\lambda}(I - R_\lambda)$. Moreover, for every $\lambda > 0$ the resolvent R_λ can be expressed pointwise, that is, there exists $\tilde{R}_\lambda : \mathbb{R}_s^{d \times d} \rightarrow \mathbb{R}_s^{d \times d}$ such that

$$(2.2) \quad R_\lambda(\tau)(x) = \tilde{R}_\lambda(\tau(x)) \quad \text{f.a.a. } x \in \Omega \text{ and } \forall \tau \in L^2(\Omega; \mathbb{R}_s^{d \times d}).$$

With a slight abuse of notation we denote also \tilde{R}_λ by R_λ . It is to be noted that this is the case for the subdifferential of an indicator function of a pointwise defined set, where the resolvent is simply the projection onto this set, this example will be considered in Section 5 below. For further reference on maximal monotone operators, we refer to [5], [29, Ch. 32], [4, Ch. 55], and [21, Ch. 55].

Optimization problem By $J : L^2(\mathcal{H}) \times \mathfrak{X}_c \rightarrow \mathbb{R}$, $J(u, v, z, f) := \Psi(u, v, z) + \frac{\alpha}{2} \|f\|_{\mathfrak{X}_c}^2$ we denote the objective function, where \mathcal{H} is given in Definition 3.5 and the control space \mathfrak{X}_c is a Hilbert space and embedded into $H^1(L^2(\Omega; \mathbb{R}^d))$. We assume that $\Psi : L^2(\mathcal{H}) \rightarrow \mathbb{R}$ is weakly lower semicontinuous, continuous and bounded from below and that the Tikhonov parameter α is a positive constant.

3 STATE EQUATION

We begin our investigation with the state equation. At first we give the definition of a solution and then transform the state equation into an EVI with a new (maximal monotone) operator \mathcal{A} . In Section 3.2 we prove the existence of a solution by showing that the operator \mathcal{A} is maximal monotone, then we can apply [4, Theorem 55.A]. Finally, in Section 3.3 we can use some results in [18] to obtain convergence results.

The formal strong formulation of the state equation reads

$$\begin{aligned} (3.1a) \quad & \rho \ddot{u} - \nabla \cdot \mathbb{C}(\nabla^s u - z) = f && \text{in } \Omega, \\ (3.1b) \quad & v \cdot \mathbb{C}(\nabla^s u - z) = 0 && \text{on } \Gamma_N, \\ (3.1c) \quad & u = 0 && \text{on } \Gamma_D, \\ (3.1d) \quad & \dot{z} \in A(\mathbb{C}\nabla^s u - (\mathbb{C} + \mathbb{B})z) && \text{in } \Omega, \\ (3.1e) \quad & (u, \dot{u}, z)(0) = (u_0, v_0, z_0) && \text{in } \Omega. \end{aligned}$$

Note that we have assumed in the standing assumptions above that the density is constant in Ω . It is possible to consider a density which has a spatial dependency (that is, a function from Ω to $(0, \infty)$), one has then in particular to verify that the operator Q , given in Definition 3.5, is well defined, that is, the multiplication of ρ (and also $1/\rho$) with a Sobolev function is again a Sobolev function. However, for simplicity we assume that ρ is constant.

We impose the following assumption for the rest of this section.

Assumption 3.1 (Standing assumption for Section 3). Let $f \in H^1(L^2(\Omega; \mathbb{R}_s^{d \times d}))$ be given.

3.1 DEFINITION AND TRANSFORMATION

Let us begin with the definition of a solution to the state equation (3.1).

Definition 3.2 (Solution to plasticity with inertia). We call $u \in H^1(H_D^1(\Omega; \mathbb{R}^d)) \cap H^2(L^2(\Omega; \mathbb{R}^d))$ and $z \in H^1(L^2(\Omega; \mathbb{R}_s^{d \times d}))$ solution of (3.1) if

$$(3.2) \quad \begin{aligned} & \rho \ddot{u} - \operatorname{div} \mathbb{C}(\nabla^s u - z) = f, \\ & \dot{z} \in A(\mathbb{C}\nabla^s u - (\mathbb{C} + \mathbb{B})z), \\ & (u, \dot{u}, z)(0) = (u_0, v_0, z_0) \end{aligned}$$

holds.

Before we can transform the state equation into an EVI we need to reformulate it, to this end we introduce the following

Definition 3.3 (*z to q mapping*). We define

$$\mathfrak{Q} : H^1(\Omega; \mathbb{R}^d) \times L^2(\Omega; \mathbb{R}_s^{d \times d}) \rightarrow L^2(\Omega; \mathbb{R}_s^{d \times d}), \quad (u, z) \mapsto \mathbb{C} \nabla^s u - (\mathbb{C} + \mathbb{B})z$$

and its inverse (for fixed u)

$$\mathfrak{Z} : H^1(\Omega; \mathbb{R}^d) \times L^2(\Omega; \mathbb{R}_s^{d \times d}) \rightarrow L^2(\Omega; \mathbb{R}_s^{d \times d}), \quad (u, q) \mapsto (\mathbb{C} + \mathbb{B})^{-1}(\mathbb{C} \nabla^s u - q).$$

These operators will also be used to transform the optimal control problem in Section 4.1 below.

For the following lemma, we recall the definition of \mathbb{D} and \mathbb{E} given in the standing assumptions above, that is, $\mathbb{D} = \mathbb{B}(\mathbb{C} + \mathbb{B})^{-1}\mathbb{C}$ and $\mathbb{E} = \mathbb{C}(\mathbb{C} + \mathbb{B})^{-1}$.

Lemma 3.4 (*Transformation of z to q*). We consider

$$(3.3) \quad \begin{aligned} \rho \ddot{u} - \operatorname{div}(\mathbb{D} \nabla^s u + \mathbb{E}q) &= f, \\ (\mathbb{C} + \mathbb{B})^{-1} \dot{q} + A(q) - \mathbb{E}^\top \nabla^s \dot{u} &\ni 0, \\ (u, \dot{u}, q)(0) &= (u_0, v_0, q_0) = (u_0, v_0, \mathfrak{Q}(u_0, z_0)) \end{aligned}$$

for functions $u \in H^1(H_D^1(\Omega; \mathbb{R}^d)) \cap H^2(L^2(\Omega; \mathbb{R}^d))$, $q \in H^1(L^2(\Omega; \mathbb{R}_s^{d \times d}))$. Recall that \mathbb{E}^\top is the adjoint of \mathbb{E} . Then the following holds:

When (u, z) is a solution of (3.1), then $(u, q) = (u, \mathfrak{Q}(u, z))$ solves (3.3). Vice versa, when (u, q) solves (3.3), then $(u, z) = (u, \mathfrak{Z}(u, q))$ is a solution of (3.1).

Proof. Both implications can be immediately obtained by using the definition of \mathfrak{Q} and \mathfrak{Z} and inserting z in (3.2) and q in (3.3), respectively (note that $\mathbb{C} - \mathbb{C}(\mathbb{C} + \mathbb{B})^{-1}\mathbb{C} = (I - \mathbb{C}(\mathbb{C} + \mathbb{B})^{-1})\mathbb{C} = \mathbb{B}(\mathbb{C} + \mathbb{B})^{-1}\mathbb{C} = \mathbb{D}$). \square

We are now in the position to introduce the EVI, respectively the operator \mathcal{A} .

Definition 3.5 (*The operator \mathcal{A}*). For $p \in [1, \infty]$ we set

$$\mathcal{Y}_p := W_D^{1,p}(\Omega; \mathbb{R}^d) \times L^2(\Omega; \mathbb{R}^d) \times L^p(\Omega; \mathbb{R}_s^{d \times d}) \quad \text{and} \quad \mathcal{H} := \mathcal{Y}_2.$$

The scalar product on \mathcal{H} is defined by

$$\begin{aligned} ((u_1, v_1, q_1), (u_2, v_2, q_2))_{\mathcal{H}} \\ := (\mathbb{D} \nabla^s u_1, \nabla^s u_2)_{L^2(\Omega; \mathbb{R}_s^{d \times d})} + (v_1, v_2)_{L^2(\Omega; \mathbb{R}^d)} + (q_1, q_2)_{L^2(\Omega; \mathbb{R}_s^{d \times d})} \end{aligned}$$

(recall that \mathbb{D} is symmetric and coercive). We define

$$\mathcal{A} : D(\mathcal{A}) \rightarrow 2^{\mathcal{H}}, \quad (u, v, q) \mapsto \begin{pmatrix} -v \\ -\operatorname{div}(\mathbb{D} \nabla^s u + \mathbb{E}q) \\ A(q) - \mathbb{E}^\top \nabla^s v \end{pmatrix}$$

with the domain

$$D(\mathcal{A}) := \{(u, v, q) \in H_D^1(\Omega; \mathbb{R}^d) \times H_D^1(\Omega; \mathbb{R}^d) \times D(A) : \operatorname{div}(\mathbb{D} \nabla^s u + \mathbb{E}q) \in L^2(\Omega; \mathbb{R}^d)\}.$$

Moreover, we set

$$R : L^2(\Omega; \mathbb{R}^d) \rightarrow \mathcal{Y}_\infty, \quad f \mapsto (0, f, 0)$$

and

$$Q := (I, (1/\rho)I, \mathbb{C} + \mathbb{B}).$$

Lemma 3.6 (Transformation into an EVI). *The tuple (u, q) solves (3.3) if and only if $(u, v, q) = (u, \dot{u}, q) \in H^1(\mathcal{H})$ is a solution of*

$$(3.4) \quad Q^{-1}(\dot{u}, \dot{v}, \dot{q}) + \mathcal{A}(u, v, q) \ni Rf, \quad (u, v, q)(0) = (u_0, v_0, q_0).$$

Proof. This follows immediately from the definition of \mathcal{A} . □

Remark 3.7 (Consequences of the transformation). We note that this transformation has some consequences for the optimal control problem and its regularization discussed in Section 4. A first approach to regularize (3.1) would be to simply regularize the operator A , as we did in the case of elasto plasticity in [27, Part III]. However, our approach is different, due to the transformation into an EVI we can regularize the operator \mathcal{A} , this is our method in Section 3.3 and Section 4. We also mention that the fact that $v = \dot{u}$ will be lost after the regularization (cf. Corollary 4.5 and Definition 4.7) and that we will transform our objective function in Definition 4.1, so that we obtain an optimal control problem with respect to the state (u, v, q) in (4.1). The optimality conditions given in Theorem 4.14 below are then also formulated for this transformed problem.

Remark 3.8 (Neumann surface forces and Dirichlet displacement). Let us shortly discuss some issues with possible surface forces and Dirichlet displacements. Regarding surface forces, they are currently equal to zero and contained in the domain $D(\mathcal{A})$ by the requirement $\operatorname{div}(\mathbb{D}\nabla^s u + \mathbb{E}q) \in L^2(\Omega; \mathbb{R}^d)$. Allowing now surface forces which are time dependent, the domain, and thus \mathcal{A} itself, would also depend on the time.

An approach for Dirichlet displacements would be to exchange the displacement with a “new” displacement minus the Dirichlet displacement, then one could still define the domain $D(\mathcal{A})$ as a subset of $H_D^1(\Omega; \mathbb{R}^d) \times H_D^1(\Omega; \mathbb{R}^d) \times D(A)$. However, this would again make the domain and the operator itself time dependent (the Dirichlet displacement would occur also in the operator).

In both cases one could still show that the arising operator is maximal monotone for a fixed time, but for different points in time the monotonicity would be perturbed by the time dependent functions. Having now a closer look at Theorem 3.15 below, respectively [4, Theorem 55.A], we see that a comparison of two different points in time is used to derive a priori estimates. Following this proof, the time depend functions would occur and a straightforward adaption is not possible.

At this juncture, let us also elaborate on the underlying spaces of the operator \mathcal{A} . One might try to exchange $L^2(\mathbb{R}; \Omega)$ with a negative Sobolev space in the definition of \mathcal{H} to allow surface forces. However, with this definition of \mathcal{H} , for instance, the proof of Lemma 3.12 (which is used to show the monotonicity of \mathcal{A}) would not be valid anymore. Thus, our choice of \mathcal{H} seems reasonable.

3.2 EXISTENCE OF A SOLUTION

We prove now the existence of a solution to (3.1) by using an existence result for EVIs involving a maximal monotone operator given in [4, Theorem 55.A], thus we need to show that \mathcal{A} is maximal monotone. Since the monotonicity of \mathcal{A} can be easily obtained (cf. Lemma 3.12), it remains to prove that the resolvent exists (cf. the proof of Proposition 3.13). For this it is sufficient to show the existence of a solution to (3.9) in the case $p = 2$. However, since the existence and Lipschitz continuity for $p > 2$ is needed to derive optimality conditions in Section 4.2, we already provide the following corollary for later needed results.

Corollary 3.9 (Extended nonlinear elasticity). *Let $\lambda > 0$ and $p \in [2, \bar{p}]$, where \bar{p} is from [14, Theorem 1.1], with $2 - \frac{d}{2} \geq -\frac{d}{p}$. We assume that there exist $m, M, D \in \mathbb{R}, D \geq 0 < m \leq M$, such that the family of*

functions $\{b_\sigma : \Omega \times \mathbb{R}_s^{d \times d} \rightarrow \mathbb{R}_s^{d \times d}\}_{\sigma \in \mathbb{R}_s^{d \times d}}$ has the following properties:

$$(3.5) \quad b_0(\cdot, 0) \in L^\infty(\Omega; \mathbb{R}_s^{d \times d}),$$

$$(3.6) \quad b_\sigma(\cdot, \tau) \text{ is measurable,}$$

$$(3.7) \quad (b_\sigma(x, \tau) - b_{\bar{\sigma}}(x, \bar{\tau})) : (\tau - \bar{\tau}) + D(|\sigma - \bar{\sigma}| + |\tau - \bar{\tau}|)|\sigma - \bar{\sigma}| \geq m|\tau - \bar{\tau}|^2,$$

$$(3.8) \quad |b_\sigma(x, \tau) - b_{\bar{\sigma}}(x, \bar{\tau})| \leq M(|\tau - \bar{\tau}| + |\sigma - \bar{\sigma}|)$$

for almost all $x \in \Omega$ and all $\sigma, \bar{\sigma}, \tau, \bar{\tau} \in \mathbb{R}_s^{d \times d}$.

Then for every $\varphi \in L^p(\Omega; \mathbb{R}_s^{d \times d})$ and $L \in W_D^{-1,p}(\Omega; \mathbb{R}^d)$ there exists a unique solution $u \in W_D^{1,p}(\Omega; \mathbb{R}^d)$ of

$$-\operatorname{div} b_\varphi(\cdot, \nabla^s u) + \frac{u}{\lambda^2} = L.$$

Moreover, there exists a constant C such that the inequality

$$\|u_1 - u_2\|_{W^{1,p}(\Omega; \mathbb{R}^d)} \leq C \left(\|\varphi_1 - \varphi_2\|_{L^p(\Omega; \mathbb{R}_s^{d \times d})} + \|L_1 - L_2\|_{W_D^{-1,p}(\Omega; \mathbb{R}^d)} \right)$$

holds for all $\varphi_1, \varphi_2 \in L^p(\Omega; \mathbb{R}_s^{d \times d})$ and $L_1, L_2 \in W_D^{-1,p}(\Omega; \mathbb{R}^d)$, where u_1 and u_2 are the solutions with respect to (φ_1, L_1) and (φ_2, L_2) .

Proof. Note that $b_\sigma(\cdot, \tau) \in L^p(\Omega; \mathbb{R}_s^{d \times d})$ holds for all $\tau, \sigma \in L^p(\Omega; \mathbb{R}_s^{d \times d})$ (and in fact for all $p \in [1, \infty]$), which follows from (3.5), (3.6) and (3.8) (taking into account that a pointwise limit of measurable functions is also measurable, see [23, Corollary 3.1.5]).

Let us at first consider the case $p = 2$. Then the existence of a solution follows from the Browder-Minty theorem, Korn's inequality and the Poincaré inequality. In order to verify the inequality, let $\varphi_1, \varphi_2 \in L^2(\Omega; \mathbb{R}_s^{d \times d})$, $L_1, L_2 \in H_D^{-1}(\Omega; \mathbb{R}^d)$ and $u_1, u_2 \in H_D^1(\Omega; \mathbb{R}^d)$ the corresponding solutions. Then we obtain

$$\begin{aligned} \langle L_1 - L_2, u_1 - u_2 \rangle &= (b_{\varphi_1(\cdot)}(\cdot, \nabla^s u_1(\cdot)) - b_{\varphi_2(\cdot)}(\cdot, \nabla^s u_2(\cdot)), \nabla^s(u_1 - u_2))_{L^2(\Omega; \mathbb{R}_s^{d \times d})} \\ &\quad + \left\| \frac{u_1 - u_2}{\lambda} \right\|_{L^2(\Omega; \mathbb{R}^d)}^2 \\ &\geq m \|\nabla^s(u_1 - u_2)\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})}^2 - D \|\varphi_1 - \varphi_2\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})}^2 \\ &\quad - D \int_\Omega |\nabla^s(u_1 - u_2)| |\varphi_1 - \varphi_2| + \frac{1}{\lambda^2} \|u_1 - u_2\|_{L^2(\Omega; \mathbb{R}^d)}^2. \end{aligned}$$

Using now

$$\langle L_1 - L_2, u_1 - u_2 \rangle \leq \|L_1 - L_2\|_{W_D^{-1,2}(\Omega; \mathbb{R}^d)} \|u_1 - u_2\|_{W^{1,2}(\Omega; \mathbb{R}^d)}$$

and

$$D \int_\Omega |\nabla^s(u_1 - u_2)| |\varphi_1 - \varphi_2| \leq D \|\nabla^s(u_1 - u_2)\|_{L^2(\Omega; \mathbb{R}^d)} \|\varphi_1 - \varphi_2\|_{L^2(\Omega; \mathbb{R}^d)}$$

and Young's inequality, yields

$$\begin{aligned} \|u_1 - u_2\|_{W^{1,2}(\Omega; \mathbb{R}^d)}^2 &\leq \bar{C} \left(\|\varphi_1 - \varphi_2\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})}^2 + \|L_1 - L_2\|_{W_D^{-1,2}(\Omega; \mathbb{R}^d)}^2 \right) \\ &\leq \bar{C} \left(\|\varphi_1 - \varphi_2\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})} + \|L_1 - L_2\|_{W_D^{-1,2}(\Omega; \mathbb{R}^d)} \right)^2 \end{aligned}$$

for a certain positive constant \bar{C} , hence, the asserted inequality is fulfilled.

For the general case let now $\varphi \in L^\infty(\Omega; \mathbb{R}_s^{d \times d})$ and $L \in W_D^{-1,p}(\Omega; \mathbb{R}^d)$, we define $b : \Omega \times \mathbb{R}_s^{d \times d} \rightarrow \mathbb{R}_s^{d \times d}$ by

$$b(x, \tau) := b_{\varphi(x)}(x, \tau)$$

and $L_u \in W_D^{-1,p}(\Omega; \mathbb{R}^d)$ by

$$\langle L_u, v \rangle := \langle L, v \rangle - \frac{1}{\lambda^2} (u, v)_{L^2(\Omega; \mathbb{R}^d)},$$

where $u \in H_D^1(\Omega; \mathbb{R}^d) \hookrightarrow L^q(\Omega; \mathbb{R}^d)$, with $1 - \frac{d}{2} = -\frac{d}{q}$ when $d > 2$, $q = 2$ when $d = 2$ and $q = \infty$ when $d = 1$, is the solution in the case $p = 2$ and $v \in W^{1,p'}(\Omega; \mathbb{R}^d) \hookrightarrow L^{q'}(\Omega; \mathbb{R}^d)$ (note that $1 - \frac{d}{p'} + \frac{d}{q} = 1 + \frac{d}{p} - \frac{d}{q} = 2 + \frac{d}{p} - \frac{d}{2} \geq 0$ when $d > 2$ and $1 - \frac{d}{p'} + \frac{d}{q} > 0$ otherwise). We can now apply [14, Theorem 1.1] (here we need $\varphi \in L^\infty(\Omega; \mathbb{R}_s^{d \times d})$ to satisfy [14, (1.6a)], the other requirements in [14, Assumption 1.5] are obviously fulfilled due to (3.5)–(3.8)) to obtain $\bar{u} \in W_D^{1,p}(\Omega; \mathbb{R}^d)$ such that

$$(b(\cdot, \nabla^s \bar{u}), \nabla^s v)_{L^2(\Omega; \mathbb{R}_s^{d \times d})} = \langle L_u, v \rangle,$$

that is,

$$(b_{\varphi}(\cdot, \nabla^s \bar{u}), \nabla^s v)_{L^2(\Omega; \mathbb{R}_s^{d \times d})} + \frac{1}{\lambda^2} (u, v)_{L^2(\Omega; \mathbb{R}^d)} = \langle L, v \rangle,$$

holds for all $v \in W_D^{1,p'}(\Omega; \mathbb{R}^d)$, we get in particular $u = \bar{u} \in W_D^{1,p}(\Omega; \mathbb{R}^d)$ since u is the unique solution of the equation above for all $v \in H_D^1(\Omega; \mathbb{R}^d)$.

To prove the asserted inequality let $\varphi_1, \varphi_2 \in L^\infty(\Omega; \mathbb{R}_s^{d \times d})$, $L_1, L_2 \in W_D^{-1,p}(\Omega; \mathbb{R}^d)$ and $u_1, u_2 \in W^{1,p}(\Omega; \mathbb{R}^d)$ the corresponding solutions and define L_{u_1}, L_{u_2} as before. Having a closer look at the proof of [14, Theorem 1.1], respectively [12, Theorem 1], one can see that there exists a constant $c > 0$, depending only on p, m and M (thus not on $L_1, L_2, \varphi_1, \varphi_2$), such that

$$\|u_1 - u_2\|_{W^{1,p}(\Omega; \mathbb{R}^d)} \leq c \|A_1(u_2) - A_2(u_2) - L_{u_1} + L_{u_2}\|_{W_D^{-1,p}(\Omega; \mathbb{R}^d)},$$

where $A_i : W^{1,p}(\Omega; \mathbb{R}^d) \rightarrow W_D^{-1,p}(\Omega; \mathbb{R}^d)$ is defined by

$$\langle A_i(v_1), v_2 \rangle := (b_{\varphi_i}(\cdot, \nabla^s v_1), \nabla^s v_2)_{L^2(\Omega; \mathbb{R}_s^{d \times d})}$$

for all $v_1 \in W^{1,p}(\Omega; \mathbb{R}^d)$, $v_2 \in W^{1,p'}(\Omega; \mathbb{R}^d)$ and for $i \in \{1, 2\}$. We finally obtain

$$\begin{aligned} \|u_1 - u_2\|_{W^{1,p}(\Omega; \mathbb{R}^d)} &\leq c \left(\|A_1(u_2) - A_2(u_2)\|_{W_D^{-1,p}(\Omega; \mathbb{R}^d)} + \|L_{u_1} - L_{u_2}\|_{W_D^{-1,p}(\Omega; \mathbb{R}^d)} \right) \\ &\leq c \left(M \|\varphi_1 - \varphi_2\|_{L^p(\Omega; \mathbb{R}_s^{d \times d})} + \|L_1 - L_2\|_{W_D^{-1,p}(\Omega; \mathbb{R}^d)} \right. \\ &\quad \left. + \frac{C}{\lambda^2} \|u_1 - u_2\|_{H^1(\Omega; \mathbb{R}^d)} \right), \end{aligned}$$

where we have used again the embeddings $H^1(\Omega; \mathbb{R}^d) \hookrightarrow L^q(\Omega; \mathbb{R}^d)$ and $W^{1,p'}(\Omega; \mathbb{R}^d) \hookrightarrow L^{q'}(\Omega; \mathbb{R}^d)$. Taking into account that the assertion is already proven in the case $p = 2$, we see that the desired inequality holds.

One can now obtain the result for all $\varphi_1, \varphi_2 \in L^p(\Omega; \mathbb{R}_s^{d \times d})$ by an approximation (using the just proven inequality to see that the corresponding sequence u_n is a Cauchy sequence). \square

The operator R_0 in the following proposition will later be the resolvent, or a smoothed version of the resolvent, of A and should not be confused with R from [Definition 3.5](#).

Proposition 3.10 (Solution operator \mathcal{T}_{R_0}). *Let $\lambda > 0$ and $p \geq 2$ as in [Corollary 3.9](#) and $h = (h_1, h_2, h_3) \in \mathcal{Y}_p$. Moreover, let $R_0 : \mathbb{R}_s^{d \times d} \rightarrow \mathbb{R}_s^{d \times d}$ be Lipschitz continuous and monotone. Then there exists a unique solution $u \in W_D^{1,p}(\Omega; \mathbb{R}^d)$ of*

$$(3.9) \quad -\operatorname{div}(\mathbb{D}\nabla^s u + \mathbb{E}R_0(\mathbb{E}^\top \nabla^s(u - h_1) + h_3)) = \frac{h_2}{\lambda} + \frac{h_1 - u}{\lambda^2}.$$

We denote the solution operator of this equation by $\mathcal{T}_{R_0} : \mathcal{Y}_p \rightarrow W_D^{1,p}(\Omega; \mathbb{R}^d)$, that is, $\mathcal{T}_{R_0}(h) = u$. Furthermore, \mathcal{T}_{R_0} is Lipschitz continuous. Note that the dependency of \mathcal{T}_{R_0} on λ and p will always be clear from the context.

Proof. For all $\sigma \in \mathbb{R}_s^{d \times d}$ we define $b_\sigma : \Omega \times \mathbb{R}_s^{d \times d} \rightarrow \mathbb{R}_s^{d \times d}$ by

$$b_\sigma(x, \tau) := \mathbb{D}\tau + \mathbb{E}R_0(\mathbb{E}^\top \tau + \sigma),$$

then the assertion follows from [Corollary 3.9](#) (with $\varphi := -\mathbb{E}^\top \nabla^s h_1 + h_3$ for a given $h \in \mathcal{Y}_p$), let us only prove that (3.7) is fulfilled, the other requirements can be easily checked. To this end let $\sigma, \bar{\sigma}, \tau, \bar{\tau} \in \mathbb{R}_s^{d \times d}$, then

$$\begin{aligned} (b_\sigma(x, \tau) - b_{\bar{\sigma}}(x, \bar{\tau})) &: (\tau - \bar{\tau}) \\ &\geq \gamma_{\mathbb{D}}|\tau - \bar{\tau}|^2 + \left(R_0(\mathbb{E}^\top \tau + \sigma) - R_0(\mathbb{E}^\top \bar{\tau} + \bar{\sigma})\right) : \left(\mathbb{E}^\top (\tau - \bar{\tau}) + (\sigma - \bar{\sigma})\right) \\ &\quad - \left(R_0(\mathbb{E}^\top \tau + \sigma) - R_0(\mathbb{E}^\top \bar{\tau} + \bar{\sigma})\right) : (\sigma - \bar{\sigma}) \\ &\geq \gamma_{\mathbb{D}}|\tau - \bar{\tau}|^2 - L_{R_0}|\sigma - \bar{\sigma}|^2 - L_{R_0}\|\mathbb{E}\| |\tau - \bar{\tau}| |\sigma - \bar{\sigma}| \end{aligned}$$

holds, where L_{R_0} is the Lipschitz constant of R_0 . □

Note that $R_\lambda : \mathbb{R}_s^{d \times d} \rightarrow \mathbb{R}_s^{d \times d}$ fulfills the requirements in [Proposition 3.10](#) since $R_\lambda : L^2(\Omega; \mathbb{R}_s^{d \times d}) \rightarrow L^2(\Omega; \mathbb{R}_s^{d \times d})$ is Lipschitz continuous and also monotone (cf. [4, Proposition 55.1 (ii) and Proposition 55.2 (a)]) and due to (2.2) these properties carry over to $R_\lambda : \mathbb{R}_s^{d \times d} \rightarrow \mathbb{R}_s^{d \times d}$.

Let us also mention that R_0 in [Proposition 3.10](#) does not have to be monotone, the inequality

$$(R_0(a) - R_0(b)) : (a - b) \geq -\varepsilon|a - b|^2$$

for $a, b \in \mathbb{R}_s^{d \times d}$ with $\varepsilon < \gamma_{\mathbb{D}}/\|\mathbb{E}^\top\|^2$ would be sufficient.

We can now prove the existence of the resolvent of \mathcal{A} , from which we can then derive the maximal monotonicity of \mathcal{A} in [Proposition 3.13](#) below.

Proposition 3.11 (Existence of the resolvent of \mathcal{A}). *For every $\lambda > 0$ and $h = (h_1, h_2, h_3) \in \mathcal{H}$, the tuple*

$$\begin{pmatrix} u \\ v \\ q \end{pmatrix} = \begin{pmatrix} \mathcal{T}_{R_\lambda}(h) \\ \frac{1}{\lambda}(\mathcal{T}_{R_\lambda}(h) - h_1) \\ R_\lambda(\mathbb{E}^\top \nabla^s(\mathcal{T}_{R_\lambda}(h) - h_1) + h_3) \end{pmatrix}$$

is contained in $D(\mathcal{A})$ and the unique solution of $(u, v, q) + \lambda\mathcal{A}(u, v, q) \ni h$.

Proof. Using the definition of \mathcal{T}_{R_λ} we get

$$-\lambda \operatorname{div}(\mathbb{D}\nabla^s u + \mathbb{E}q) = h_2 - v,$$

which is the second row in $(u, v, q) + \lambda \mathcal{A}(u, v, q) \ni h$ and we also get $(u, v, q) \in D(\mathcal{A})$ (note that $rg(R_\lambda) \subset D(A)$). That the first and last row in $(u, v, q) + \lambda \mathcal{A}(u, v, q) \ni h$ is also fulfilled follows immediately from the definitions of u, v and q .

Furthermore, when (u, v, q) is a solution of $(u, v, q) + \lambda \mathcal{A}(u, v, q) \ni h$, then one verifies analog that (u, v, q) must have the claimed form, therefore the uniqueness follows from the uniqueness of a solution to (3.9). \square

Lemma 3.12 (Monotonicity of \mathcal{A}). *The equation*

$$\begin{aligned} & (\mathcal{A}(u_1, v_1, q_1) - \mathcal{A}(u_2, v_2, q_2), (u_1, v_1, q_1) - (u_2, v_2, q_2))_{\mathcal{H}} \\ &= (A(q_1) - A(q_2), q_1 - q_2)_{L^2(\Omega; \mathbb{R}_s^{d \times d})} \end{aligned}$$

holds for all $(u_1, v_1, q_1), (u_2, v_2, q_2) \in D(\mathcal{A})$.

Proof. Using the definition of \mathcal{A} and the scalar product in \mathcal{H} we obtain

$$\begin{aligned} & (\mathcal{A}(u_1, v_1, q_1), (u_1, v_1, q_1) - (u_2, v_2, q_2))_{\mathcal{H}} \\ &= -(\mathbb{D}\nabla^s v_1, \nabla^s(u_1 - u_2))_{L^2(\Omega; \mathbb{R}_s^{d \times d})} - (\operatorname{div}(\mathbb{D}\nabla^s u_1 + \mathbb{E}q_1), v_1 - v_2)_{L^2(\Omega; \mathbb{R}^d)} \\ &\quad + (A(q_1) - \mathbb{E}^\top, \nabla^s v_1)_{q_1 - q_2} L^2(\Omega; \mathbb{R}_s^{d \times d}) \\ &= (\mathbb{D}\nabla^s v_1, \nabla^s u_2)_{L^2(\Omega; \mathbb{R}_s^{d \times d})} - (\mathbb{D}\nabla^s u_1, \nabla^s v_2)_{L^2(\Omega; \mathbb{R}_s^{d \times d})} - (\mathbb{E}^\top \nabla^s v_2, q_1)_{L^2(\Omega; \mathbb{R}_s^{d \times d})} \\ &\quad + (\mathbb{E}^\top \nabla^s v_1, q_2)_{L^2(\Omega; \mathbb{R}_s^{d \times d})} + (A(q_1), q_1 - q_2)_{L^2(\Omega; \mathbb{R}_s^{d \times d})}, \end{aligned}$$

evaluating now $(\mathcal{A}(u_2, v_2, q_2), (u_1, v_1, q_1) - (u_2, v_2, q_2))_{\mathcal{H}}$ and taking the difference yields the assertion. \square

Proposition 3.13 (\mathcal{A} is maximal monotone). *The operator $\mathcal{A} : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ is maximal monotone.*

Proof. The monotonicity of \mathcal{A} follows immediately from Lemma 3.12 and the monotonicity of A .

To prove that \mathcal{A} is maximal monotone, it is, according to [4, Proposition 55.1 (B)], sufficient that $R(I + \mathcal{A}) = \mathcal{H}$, that is, we have to show that for every $(h_1, h_2, h_3) \in \mathcal{H}$ there exists $(u, v, q) \in D(\mathcal{A})$ such that $(u, v, q) + \mathcal{A}(u, v, q) \ni (h_1, h_2, h_3)$. This follows from Proposition 3.11 with $\lambda = 1$. \square

In what follows it is convenient to give the integration operator a name.

Definition 3.14 (Integration operator). We define $\mathcal{F} : H^1(L^2(\Omega; \mathbb{R}^d)) \rightarrow H^2(L^2(\Omega; \mathbb{R}^d))$ by $(\mathcal{F}f)(t) := \int_0^t f(s)ds$ for all $f \in H^1(L^2(\Omega; \mathbb{R}^d))$. Moreover, we abbreviate $\mathcal{F}_\rho := \mathcal{F}/\rho$. As usual, we denote the operators with different inverse images and ranges with the same symbol, for instance $\mathcal{F} : L^2(L^2(\Omega; \mathbb{R}^d)) \rightarrow H^1(L^2(\Omega; \mathbb{R}^d))$.

Theorem 3.15 (Existence of a solution to the state equation). *There exists a unique solution $(u, v, q) \in H^1(\mathcal{H})$ of (3.4). Moreover, the inequality*

$$\|(\dot{u}, \dot{v}, \dot{q})\|_{L^2(\mathcal{H})} \leq C(1 + \|f\|_{H^1(L^2(\Omega; \mathbb{R}^d))})$$

holds, where the constant C does not depend on f .

Proof. The tuple (u, v, q) is a solution of (3.4) if and only if $w := (u, v, q)$ solves

$$\dot{w} + \tilde{\mathcal{A}}(w) \ni \frac{1}{\rho} Rf, \quad w(0) = (u_0, v_0, q_0),$$

with $\tilde{\mathcal{A}} := Q\mathcal{A}$. One easily verifies that $\tilde{\mathcal{A}}$ is a maximal monotone operator with respect to $\mathcal{H}_{Q^{-1}}$ (that is, the space \mathcal{H} equipped with the scalar product $(Q^{-1}\cdot, \cdot)_{\mathcal{H}_{Q^{-1}}}$), thus we can apply [4, Theorem 55.A] to obtain a solution $w \in H^1(\mathcal{H})$. Moreover, as can be seen in the proof of [4, Theorem 55.A] we get

$$\sqrt{\gamma_{Q^{-1}}}\|\dot{w}_\lambda\|_{C(\mathcal{H})} \leq \|\dot{w}_\lambda\|_{C(\mathcal{H}_{Q^{-1}})} \leq C(1 + \frac{1}{\rho}\|Rf\|_{H^1(\mathcal{H})}) = C(1 + \frac{1}{\rho}\|f\|_{H^1(L^2(\Omega; \mathbb{R}^d))})$$

for all $\lambda > 0$, where w_λ is the solution of

$$\dot{w}_\lambda + \tilde{\mathcal{A}}_\lambda(w_\lambda) = \frac{1}{\rho}Rf, \quad w_\lambda(0) = (u_0, v_0, q_0).$$

Since $w_\lambda \rightarrow w$ in $H^1(\mathcal{H})$, we obtain the desired inequality. □

Remark 3.16 (\mathcal{A} is not a subdifferential). Let us show that the maximal monotone operator $\mathcal{A} : \mathcal{H} \rightarrow \mathcal{H}$ is not a subdifferential, that is, there exists no proper, convex and lower semicontinuous function $\Phi : \mathcal{H} \rightarrow (-\infty, \infty]$ such that

$$\begin{aligned} \mathcal{A}(u, v, q) = \partial\Phi(u, v, q) &= \{(c, d, e) \in \mathcal{H} : \Phi(\hat{u}, \hat{v}, \hat{q}) \geq \Phi(u, v, q) \\ &\quad + ((c, d, e), (\hat{u} - u, \hat{v} - v, \hat{q} - q))_{\mathcal{H}} \forall (\hat{u}, \hat{v}, \hat{q}) \in \mathcal{H}\} \end{aligned}$$

holds for all $(u, v, q) \in \mathcal{H}$. In fact, there exists even not any function $\Phi : \mathcal{H} \rightarrow (-\infty, \infty]$ such that the equation above holds, which can be seen as follows:

Let us assume that such a Φ exists and recall that $(u_0, v_0, q_0) \in D(\mathcal{A})$. Then, using Lemma 3.12 with $(u_1, v_1, q_1) = (u + u_0, v, q_0)$ and $(u_2, v_2, q_2) = (u_0, 0, q_0)$,

$$\begin{aligned} \Phi(u_0, 0, q_0) &\geq \Phi(u + u_0, v, q_0) - (\mathcal{A}(u + u_0, v, q_0), (u, v, 0))_{\mathcal{H}} \\ &= \Phi(u + u_0, v, q_0) - (\mathcal{A}(u_0, 0, q_0), (u, v, 0))_{\mathcal{H}} \\ &\geq \Phi(u_0, 0, q_0) + (\mathcal{A}(u_0, 0, q_0), (u, v, 0))_{\mathcal{H}} - (\mathcal{A}(u_0, 0, q_0), (u, v, 0))_{\mathcal{H}} \\ &= \Phi(u_0, 0, q_0) \end{aligned}$$

holds for all (u, v) such that $(u + u_0, v, q_0) \in D(\mathcal{A})$, hence,

$$\begin{aligned} (\mathcal{A}(u_0, 0, q_0), (\hat{u} - u, \hat{v} - v, 0))_{\mathcal{H}} &= \Phi(\hat{u} + u_0, \hat{v}, q_0) - \Phi(u + u_0, v, q_0) \\ &\geq (\mathcal{A}(u + u_0, v, q_0), (\hat{u} - u, \hat{v} - v, 0))_{\mathcal{H}} \end{aligned}$$

which gives

$$0 \geq (\mathbb{D}\nabla^s u, \nabla^s \hat{v})_{\mathcal{H}} - (\mathbb{D}\nabla^s \hat{u}, \nabla^s v)_{\mathcal{H}}$$

for all $(\hat{u}, \hat{v}), (u, v)$ such that $(u + u_0, v, q_0), (\hat{u} + u_0, \hat{v}, q_0) \in D(\mathcal{A})$. Choosing now an arbitrary $u \in C_c^\infty(\Omega; \mathbb{R}^d)$, $u \neq 0$, $\hat{v} = u$ and $\hat{u} = v = 0$, we obtain the desired contradiction.

In light of Remark 3.16, the case of plasticity with inertia essentially differs from the EVI analyzed in [18] in two aspects. First, we have more regularity in time as explained after Lemma 3.6. Second, we lose a certain boundedness of the maximal monotone operator, which was assumed in [18, Sect. 2], and it is not a subdifferential.

It is also to be noted that Remark 3.16 is independent of the operator A .

3.3 REGULARIZATION AND CONVERGENCE RESULTS

As already pointed out earlier, (3.4) can be transformed into the EVI which was analyzed in [18], namely

$$(3.10) \quad \dot{p} \in \mathcal{A}(R\mathcal{F}_\rho f - Qp), \quad p(0) = p_0,$$

where we set $p_0 := -Q^{-1}(u_0, v_0, q_0)$. We can observe that when (u, v, q) is a solution of (3.4), then $p := R\mathcal{F}f - Q^{-1}(u, v, q)$ is a solution of (3.10) and when p is a solution of (3.10), then $(u, v, q) := R\mathcal{F}_\rho f - Qp$ is a solution of (3.4).

Thanks to this transformation we can use several results from [18] (namely Lemma 3.7, Lemma 3.8 and Proposition 3.5), with which the derivation of convergence results will be an easy task. Let us emphasize that the maximal monotone operator therein was assumed to have a closed domain and that

$$A^0 : D(A) \rightarrow \mathcal{H}, \quad h \mapsto \arg \min_{v \in A(h)} \|v\|_{\mathcal{H}}$$

is bounded on bounded sets. Clearly, both assumptions are not fulfilled for \mathcal{A} , however, they were only needed to prove [18, Theorem 3.3], which can be replaced by Theorem 3.15, cf. also [18, Remark 3.13]. Therefore we can still apply the above mentioned results.

Theorem 3.17 (Weak convergence of the state). *Let $f \in H^1(L^2(\Omega; \mathbb{R}^d))$, $\{f_n\}_{n \in \mathbb{N}} \subset H^1(L^2(\Omega; \mathbb{R}^d))$ such that $f_n \rightharpoonup f$ in $H^1(L^2(\Omega; \mathbb{R}^d))$ and $\mathcal{F}f_n \rightarrow \mathcal{F}f$ in $L^1(L^2(\Omega; \mathbb{R}^d))$. Moreover, let $(u, v, q) \in H^1(\mathcal{H})$ be the solution of (3.4) and $(u_n, v_n, q_n) \in H^1(\mathcal{H})$, for every $n \in \mathbb{N}$, either the solution of*

$$Q^{-1}(\dot{u}_n, \dot{v}_n, \dot{q}_n) + \mathcal{A}(u_n, v_n, q_n) \ni Rf_n, \quad (u_n, v_n, q_n)(0) = (u_0, v_0, q_0)$$

or

$$Q^{-1}(\dot{u}_n, \dot{v}_n, \dot{q}_n) + \mathcal{A}_{\lambda_n}(u_n, v_n, q_n) = Rf_n, \quad (u_n, v_n, q_n)(0) = (u_0, v_0, q_0),$$

where $\{\lambda_n\}_{n \in \mathbb{N}} \subset (0, \infty)$, $\lambda_n \searrow 0$.

Then $(u_n, v_n, q_n) \rightarrow (u, v, q)$ in $H^1(\mathcal{H})$ and $(u_n, v_n, q_n) \rightarrow (u, v, q)$ in $C(H^1(\Omega; \mathbb{R}^d)) \times L^1(L^2(\Omega; \mathbb{R}^d)) \times C(L^2(\Omega; \mathbb{R}_s^{d \times d}))$. If additionally $\mathcal{F}f_n \rightarrow \mathcal{F}f$ in $C(L^2(\Omega; \mathbb{R}_s^{d \times d}))$, then $v_n \rightarrow v$ in $C(L^2(\Omega; \mathbb{R}^d))$.

Proof. The function $p := R\mathcal{F}f - Q^{-1}(u, v, q) \in H^1(\mathcal{H})$ is the unique solution of (3.10) and $p_n := R\mathcal{F}f_n - Q^{-1}(u_n, v_n, q_n) \in H^1(\mathcal{H})$ either the unique solution of

$$\dot{p}_n \in \mathcal{A}(R\mathcal{F}_\rho f_n - Qp_n), \quad p_n(0) = p_0.$$

or

$$\dot{p}_n = \mathcal{A}_{\lambda_n}(R\mathcal{F}_\rho f_n - Qp_n), \quad p_n(0) = p_0.$$

Thanks to Theorem 3.15, $(\dot{u}_n, \dot{v}_n, \dot{q}_n)$ is bounded in $L^2(\mathcal{H})$. We can now apply [18, Lemma 3.7] with $A_n = \mathcal{A}$ and $A_n = \mathcal{A}_{\lambda_n}$ (note that we can choose $A_n = \mathcal{A}_{\lambda_n}$ according to [18, Lemma 3.8]) to obtain the desired result. Note that the convergence in [18, Lemma 3.7] then means $R\mathcal{F}f_n - Q^{-1}(u_n, v_n, q_n) \rightarrow R\mathcal{F}f - Q^{-1}(u, v, q)$ in $C(\mathcal{H})$, so that the convergence $(u_n, v_n, q_n) \rightarrow (u, v, q)$ in $C(H^1(\Omega; \mathbb{R}^d)) \times L^1(L^2(\Omega; \mathbb{R}^d)) \times C(L^2(\Omega; \mathbb{R}_s^{d \times d}))$ follows from the fact that the range of R is a subset of $\{0\} \times L^2(\Omega; \mathbb{R}^d) \times \{0\}$. \square

Proposition 3.18 (Strong convergence for fixed forces). *Let $f \in H^1(L^2(\Omega; \mathbb{R}^d))$ and $(u, v, q) \in H^1(\mathcal{H})$ be the solution of (3.4) and $(u_n, v_n, q_n) \in H^1(\mathcal{H})$, for every $n \in \mathbb{N}$, the solution of*

$$Q^{-1}(\dot{u}_n, \dot{v}_n, \dot{q}_n) + \mathcal{A}_{\lambda_n}(u_n, v_n, q_n) = Rf, \\ (u_n, v_n, q_n)(0) = (u_0, v_0, q_0).$$

where $\{\lambda_n\}_{n \in \mathbb{N}} \subset (0, \infty)$, $\lambda_n \searrow 0$.

Then $(u_n, v_n, q_n) \rightarrow (u, v, q)$ in $H^1(\mathcal{H})$.

Proof. We can argue as in the proof of Theorem 3.17, the assertion follows then directly from [18, Proposition 3.5]. \square

4 OPTIMAL CONTROL

Also in this section we will make use of [18]. Since the smoothed operator \mathcal{A}_s , given in Definition 4.7, possesses the required properties for A_s in [18, Assumption 5.1 (ii)] with $\mathcal{Z} = \mathcal{H}$ (as we will see in Definition 4.7 and Proposition 4.10 below), we can apply the finding concerned with the differentiability of the solution operator associated with the EVI therein. Before we give the details in Section 4.2, we tend to the existence and approximation of optimal controls.

4.1 EXISTENCE AND APPROXIMATION OF OPTIMAL CONTROLS

Let us now consider the optimal control problem (1.1). Note that we assumed in Section 2 that Ψ is defined on $L^2(\mathcal{H})$ and not on $H^1(\mathcal{H})$, which excludes for example evaluations at certain points in time. Similar as in [18], we could also consider an objective function on $H^1(\mathcal{H})$, then we would only obtain a (possible) weak solution of the adjoint state in Theorem 4.14, see also [18, Theorem 5.12]. We decided to define Ψ on $L^2(\mathcal{H})$ only for simplicity and to keep the discussion concise.

Since we have transformed our state equation (3.1) into (3.3) by introducing the new variable q , it is reasonable to do the same with the optimal control problem. To this end, we need the following Definition 4.1 (Transformed objective function). We define

$$\Psi_z : L^2(\mathcal{H}) \rightarrow \mathbb{R}, \quad (u, v, q) \mapsto \Psi(u, v, \mathfrak{Z}(u, q))$$

and the transformed objective function

$$J_z : L^2(\mathcal{H}) \times \mathfrak{X}_c \rightarrow \mathbb{R}, \quad (u, v, q, f) \mapsto \Psi_z(u, v, q) + \frac{\alpha}{2} \|f\|_{\mathfrak{X}_c}^2$$

Using the definition above and the transformation of the state equation into (3.4), we obtain the equivalence of (1.1) and

$$(4.1) \quad \left\{ \begin{array}{l} \min \quad J_z(u, v, q, f) = \Psi_z(u, v, q) + \frac{\alpha}{2} \|f\|_{\mathfrak{X}_c}^2, \\ \text{s.t.} \quad Q^{-1}(\dot{u}, \dot{v}, \dot{q}) + \mathcal{A}(u, v, q) \ni Rf, \quad (u, v, q)(0) = (u_0, v_0, q_0), \\ \quad (u, v, q) \in H^1(H_D^1(\Omega; \mathbb{R}^d) \times L^2(\Omega; \mathbb{R}^d) \times L^2(\Omega; \mathbb{R}^{d \times d})), \\ \quad f \in \mathfrak{X}_c. \end{array} \right.$$

Let us now select a sequence $\{\lambda_n\}_{n \in \mathbb{N}} \subset (0, \infty)$ such that $\lambda_n \searrow 0$. We consider the regularized optimization problem

$$(4.2) \quad \left\{ \begin{array}{l} \min \quad J_z(u, v, q, f) = \Psi_z(u, v, q) + \frac{\alpha}{2} \|f\|_{\mathfrak{X}_c}^2, \\ \text{s.t.} \quad Q^{-1}(\dot{u}, \dot{v}, \dot{q}) + \mathcal{A}_{\lambda_n}(u, v, q) = Rf, \quad (u, v, q)(0) = (u_0, v_0, q_0) \\ \quad (u, v, q) \in H^1(H_D^1(\Omega; \mathbb{R}^d) \times L^2(\Omega; \mathbb{R}^d) \times L^2(\Omega; \mathbb{R}^{d \times d})), \\ \quad f \in \mathfrak{X}_c. \end{array} \right.$$

Theorem 4.2 (Existence and approximation of optimal solutions). *Suppose that the control space \mathfrak{X}_c is such that $\mathcal{F} : H^1(L^2(\Omega; \mathbb{R}^d)) \rightarrow H^2(L^2(\Omega; \mathbb{R}^d))$ with $\mathcal{F}(f)(t) = \int_0^t f(s)ds$ is compact from \mathfrak{X}_c into $L^1(L^2(\Omega; \mathbb{R}^d))$.*

Then there exists a global solution of (4.1) (and thus of (1.1)) and of (4.2) for every $n \in \mathbb{N}$.

Moreover, let $(\bar{u}_n, \bar{v}_n, \bar{q}_n, \bar{f}_n)_{n \in \mathbb{N}}$ be a sequence of global solution of (4.2). Then there exists a weak accumulation point $(\bar{u}, \bar{v}, \bar{q}, \bar{f})$ and every weak accumulation point is a global solution of (4.1). The subsequence of states which converges weakly towards $(\bar{u}, \bar{v}, \bar{q})$ in $H^1(\mathcal{H})$, converges also strongly in $C(H^1(\Omega; \mathbb{R}^d)) \times$

$L^1(L^2(\Omega; \mathbb{R}^d)) \times C(L^2(\Omega; \mathbb{R}_s^{d \times d}))$ and, when \mathcal{F} is compact from \mathfrak{X}_c into $C(L^2(\Omega; \mathbb{R}^d))$, then the subsequence of v_n converges also strongly in $C(L^2(\Omega; \mathbb{R}_s^{d \times d}))$. Moreover, the subsequence of controls converges strongly to f in \mathfrak{X}_c .

Proof. The existence of a global solution to (4.1) follows from the standard direct method of the calculus of variations using Theorem 3.17 and the assumed compactness of \mathcal{F} , the proof is for instance analog to the proof of [18, Theorem 4.2]. The existence of a global solution to (4.2) follows easily using the Lipschitz continuity of \mathcal{A}_{λ_n} (which implies the Lipschitz continuity of the corresponding solution operator).

The convergence result can also be obtained by standard arguments using again Theorem 3.17 and Proposition 3.18, the proof is again analog to [18, Theorem 4.5 & Corollary 4.6]. Note that the strong convergence of the states in $C(H^1(\Omega; \mathbb{R}^d)) \times L^1(L^2(\Omega; \mathbb{R}^d)) \times C(L^2(\Omega; \mathbb{R}_s^{d \times d}))$, and also of v_n in $C(L^2(\Omega; \mathbb{R}_s^{d \times d}))$ when \mathcal{F} is compact from \mathfrak{X}_c into $C(L^2(\Omega; \mathbb{R}^d))$, follows directly from Theorem 3.17. \square

A strong convergence result of the states (in $H^1(\mathcal{H})$) is not provided in the theorem above. In [18, Corollary 4.6] we were able to prove the strong convergence either when the associated maximal monotone operator is a subdifferential, which is here not the case (Remark 3.16), or when it can be deduced from the weak convergence and the convergence of the evaluations of Ψ . Since we supposed that Ψ is defined on $L^2(\mathcal{H})$, this cannot be the case. However, as elaborated on at the beginning of this section, it is possible for instance to consider a different Ψ defined on $H^1(\mathcal{H})$ such that this property holds.

Let us shortly interrupt the discussion and give two examples for the control space \mathfrak{X}_c .

Example 4.3 (Control space). In order to satisfy the assumption on \mathfrak{X}_c in Theorem 4.2, we can use the lemma of Lions-Aubin (cf. [21, III. Proposition 1.3]) and for instance choose $\mathfrak{X}_c = H^1(L^2(\Omega; \mathbb{R}^d)) \cap L^2(H^1(\Omega; \mathbb{R}^d))$ or $\mathfrak{X}_c = \{f \in H^1(L^2(\Omega; \mathbb{R}^d)) : \mathcal{F}f \in L^2(H^1(\Omega; \mathbb{R}^d))\}$ with corresponding norms.

Having dealt with the existence and approximation of optimal solutions we turn to the optimality condition for a further smoothed problem.

4.2 OPTIMALITY CONDITIONS

In order to derive first order optimality conditions we smoothen at first the optimal control problem further. Then we prove the differentiability of the smoothed solution operator and can after that finally present our main result, the optimality conditions for the smoothed optimization problem.

We impose the following assumptions for the rest of this subsection.

Assumption 4.4 (Standing assumptions for Section 4.2). (i) Let $R_s : \mathbb{R}_s^{d \times d} \rightarrow \mathbb{R}_s^{d \times d}$ be monotone, Lipschitz continuous and Fréchet differentiable.

(ii) We fix $2 < \hat{p} < p < \bar{p}$, where \bar{p} is from Corollary 3.9 (respectively [14, Theorem 1.1]), such that $2 - \frac{d}{2} \geq -\frac{d}{\hat{p}}$.

(iii) Let the initial data (u_0, v_0, q_0) be an element of \mathcal{Y}_p , where p is given in (ii).

Thanks to Proposition 3.11, we can give the precise form of the resolvent and Yosida approximation of \mathcal{A} in the following

Corollary 4.5 (Precise form of the resolvent). Let $\lambda > 0$ and denote the resolvent of \mathcal{A} by \mathcal{R}_λ . Then

$$\mathcal{R}_\lambda(h) = \begin{pmatrix} \mathcal{T}_{R_\lambda}(h) \\ \frac{1}{\lambda} \mathcal{T}_{R_\lambda}(h) - \frac{h_1}{\lambda} \\ R_\lambda(\mathbb{E}^\top \nabla^s(\mathcal{T}_{R_\lambda}(h) - h_1) + h_3) \end{pmatrix}$$

so that

$$\mathcal{A}_\lambda(h) = \frac{1}{\lambda} \begin{pmatrix} h_1 - \mathcal{T}_{R_\lambda}(h) \\ h_2 - \frac{1}{\lambda} \mathcal{T}_{R_\lambda}(h) + \frac{h_1}{\lambda} \\ h_3 - R_\lambda(\mathbb{E}^\top \nabla^s(\mathcal{T}_{R_\lambda}(h) - h_1) + h_3) \end{pmatrix}$$

for every $h = (h_1, h_2, h_3) \in \mathcal{H}$.

The Yosida approximation \mathcal{A}_λ is in view of Proposition 3.10 Lipschitz continuous from \mathcal{Y}_p to \mathcal{Y}_p , where p is given in Assumption 4.4 (ii). Therefore the state equation in (4.2) admits a solution in \mathcal{Y}_p (note that R maps into \mathcal{Y}_∞). However, since this regularity is not present in (1.1), we did not use it. In contrast, the same is true for the smoothed Yosida approximation, which is given below in Definition 4.7 (see Definition 4.11), but here this additional regularity will be used to prove the differentiability of the smoothed solution operator in Proposition 4.12.

In order to smoothen the Yosida approximation, respectively the resolvent, of \mathcal{A} , we smoothen the resolvent of A and then define the smoothed resolvent for \mathcal{A} analog to \mathcal{R}_λ . We denote this smoothed resolvent of A by $R_s : \mathbb{R}_s^{d \times d} \rightarrow \mathbb{R}_s^{d \times d}$ (which indicates that the resolvent of A can be expressed pointwise), from the properties given in Assumption 4.4 (i) one can easily derive the following inequalities, which will be useful when proving the differentiability of \mathcal{T}_{R_s} in Lemma 4.9 below.

Lemma 4.6 (Properties of R'_s). *There exists a constant C such that $|R'_s(\sigma)\tau| \leq C|\tau|$ and $0 \leq R'_s(\sigma)\tau : \tau$ holds for all $\sigma, \tau \in \mathbb{R}_s^{d \times d}$. Moreover, the same is true for $R'_s(\cdot)^*$.*

Proof. Let $\sigma, \tau \in \mathbb{R}_s^{d \times d}$ be arbitrary. The Lipschitz continuity and Fréchet differentiability of R_s gives

$$\left| \frac{r(t\tau)}{t} + R'_s(\sigma)\tau \right| = \frac{|R_s(\sigma + t\tau) - R_s(\sigma)|}{t} \leq L|\tau|$$

for all $t \in \mathbb{R} \setminus \{0\}$, where r is the remainder term of R_s . The limit $t \rightarrow 0$ yields the first assertion.

The second claim follows using the monotonicity,

$$0 \leq \frac{R_s(\sigma + t\tau) - R_s(\sigma)}{t} : \tau \rightarrow R'_s(\sigma)\tau : \tau$$

as $0 \neq t \rightarrow 0$.

Now, by definition we have $R'_s(\sigma)\tau : \eta = \tau : R'_s(\sigma)^*\eta$ for all $\sigma, \tau, \eta \in \mathbb{R}_s^{d \times d}$, so that the second assertion also holds for $R'_s(\cdot)^*$. Choosing in particular $\tau = R_s(\sigma)^*\eta$ we get

$$|R'_s(\sigma)^*\eta|^2 = |R'_s(\sigma)R'_s(\sigma)^*\eta : \eta| \leq C|R'_s(\sigma)^*\eta| |\eta|,$$

which yields the first assertion for $R'_s(\cdot)^*$. □

Definition 4.7 (Smoothed resolvent). Let $\lambda_s \in (0, \infty)$. We define

$$\mathcal{R}_s : \mathcal{Y}_p \rightarrow \mathcal{Y}_p, \quad h = (h_1, h_2, h_3) \mapsto \begin{pmatrix} \mathcal{T}_{R_s}(h) \\ \frac{1}{\lambda_s} \mathcal{T}_{R_s}(h) - \frac{h_1}{\lambda_s} \\ R_s(\mathbb{E}^\top \nabla^s(\mathcal{T}_{R_s}(h) - h_1) + h_3) \end{pmatrix}$$

and $\mathcal{A}_s := \frac{1}{\lambda_s}(I - \mathcal{R}_s)$ (see Assumption 4.4 (ii) for p). According to Proposition 3.10 and Assumption 4.4 (i), \mathcal{R}_s and \mathcal{A}_s are well defined and Lipschitz continuous. As usual, with a slight abuse of notation, we denote operators for different p with the same symbol.

Let us now consider the smoothed optimization problem

$$(4.3) \quad \begin{cases} \min & J_z(u, v, q, f) = \Psi_z(u, v, q) + \frac{\alpha}{2} \|f\|_{\mathfrak{X}_c}^2, \\ \text{s.t.} & Q^{-1}(\dot{u}, \dot{v}, \dot{q}) + \mathcal{A}_s(u, v, q) = Rf, \quad (u, v, q)(0) = (u_0, v_0, q_0) \\ & (u, v, q) \in H^1(H_D^1(\Omega; \mathbb{R}^d) \times L^2(\Omega; \mathbb{R}^d) \times L^2(\Omega; \mathbb{R}_s^{d \times d})), \\ & f \in \mathfrak{X}_c. \end{cases}$$

Analog to [Theorem 4.2](#) one can analogously prove that there exists a global solution of (4.3).

As was done in [[18](#), [Theorem 4.5 & Corollary 4.6](#)], when \mathcal{A}_s and \mathcal{A}_{λ_s} are globally “close together”, one can prove a result analog to the convergence result in [Theorem 4.2](#) with a sequence $(\bar{u}_s, \bar{v}_s, \bar{q}_s, \bar{f}_s)_{s>0}$ of global solutions to (4.3) when $\sup_{h \in \mathcal{H}} \|\mathcal{A}_{\lambda_s}(h) - \mathcal{A}_s(h)\|_{\mathcal{H}}$ tends fast enough to zero relative to λ_s . The following lemma shows that this is the case when the same is true for $\frac{1}{\lambda_s} \sup_{\tau \in L^2(\Omega; \mathbb{R}_s^{d \times d})} \|A_{\lambda_s}(\tau) - A_s(\tau)\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})}$ with $A_s = \frac{1}{\lambda_s}(I - R_s)$, which holds in the case of the von-Mises flow rule investigated in [Section 5](#) below for suitable sequences $\{\lambda_s\}_{\lambda_s>0}$ and $\{s\}_{s>0}$, cf. (5.2). Note also that [[18](#), [Lemma 3.15](#)] was used in [[18](#), [Theorem 4.5 & Corollary 4.6](#)], so that it was in particular required that $\sup_{\tau \in L^2(\Omega; \mathbb{R}_s^{d \times d})} \|A_{\lambda_s}(\tau) - A_s(\tau)\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})}$ tends faster to zero than $\exp(\frac{1}{\lambda_s})$, thus the additional factor $\frac{1}{\lambda_s}$ does not play a big role.

Lemma 4.8 (Convergence of the smoothed resolvent). *The inequality*

$$\|\mathcal{A}_{\lambda_s}(h) - \mathcal{A}_s(h)\|_{\mathcal{H}} \leq C \sqrt{1 + \frac{1}{\lambda_s^2} \sup_{\tau \in L^2(\Omega; \mathbb{R}_s^{d \times d})} \|A_{\lambda_s}(\tau) - A_s(\tau)\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})}}$$

holds for all $h \in L^2(\Omega; \mathbb{R}_s^{d \times d})$, where $A_s := \frac{1}{\lambda_s}(I - R_s)$ and the constant does only depend on \mathbb{C} and \mathbb{B} , $C = C(\mathbb{C}, \mathbb{B})$.

Proof. Let us abbreviate

$$M := \sup_{\tau \in L^2(\Omega; \mathbb{R}_s^{d \times d})} \|R_{\lambda_s}(\tau) - R_s(\tau)\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})}.$$

Due to the definitions of \mathcal{A}_s and A_s we only have to prove that

$$(4.4) \quad \|\mathcal{R}_{\lambda_s}(h) - \mathcal{R}_s(h)\|_{\mathcal{H}} \leq C \sqrt{1 + \frac{1}{\lambda_s^2} M}$$

holds for all $h \in \mathcal{H}$. To this end let $h \in \mathcal{H}$ be arbitrary and abbreviate $u := \mathcal{T}_{R_{\lambda_s}}(h)$, $u_s := \mathcal{T}_{R_s}(h) \in H_D^1(\Omega; \mathbb{R}^d)$, hence, u is the solution of (3.9) with respect to R_{λ_s} and u_s with respect to R_s , testing both equations with $u - u_s$ and subtracting the second from the first, we get

$$\begin{aligned} & \|\nabla^s(u - u_s)\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})_{\mathbb{D}}}^2 + \left\| \frac{u - u_s}{\lambda_s} \right\|_{L^2(\Omega; \mathbb{R}^d)}^2 \\ &= -(\mathbb{E}(R_{\lambda_s}(w) - R_s(w_s)), \nabla^s(u - u_s))_{L^2(\Omega; \mathbb{R}_s^{d \times d})} \\ &= -((R_{\lambda_s}(w_s) - R_s(w_s)), w - w_s)_{L^2(\Omega; \mathbb{R}_s^{d \times d})} \\ & \quad - ((R_{\lambda_s}(w) - R_{\lambda_s}(w_s)), w - w_s)_{L^2(\Omega; \mathbb{R}_s^{d \times d})} \\ &\leq -(\mathbb{D}^{-1} \mathbb{E}(R_{\lambda_s}(w_s) - R_s(w_s)), \nabla^s(u - u_s))_{L^2(\Omega; \mathbb{R}_s^{d \times d})_{\mathbb{D}}} \\ &\leq \frac{1}{2} \|\mathbb{D}^{-1} \mathbb{E}(R_{\lambda_s}(w) - R_s(w_s))\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})_{\mathbb{D}}}^2 + \frac{1}{2} \|\nabla^s(u - u_s)\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})_{\mathbb{D}}}^2 \\ &\leq \frac{\|\mathbb{E}^T \mathbb{D}^{-1} \mathbb{E}\|}{2} M^2 + \frac{1}{2} \|\nabla^s(u - u_s)\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})_{\mathbb{D}}}^2 \end{aligned}$$

with $w := \mathbb{E}^\top \nabla^s(u - h_1) + h_3$ and $w_s := \mathbb{E}^\top \nabla^s(u_s - h_1) + h_3$, where we used in particular the monotonicity of R_{λ_s} . Thus we obtain

$$(4.5) \quad \|\nabla^s(u - u_s)\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})_{\mathbb{D}}}^2 + \left\| \frac{u - u_s}{\lambda_s} \right\|_{L^2(\Omega; \mathbb{R}^d)}^2 \leq CM^2.$$

We get further

$$\begin{aligned} & \|R_{\lambda_s}(w) - R_s(w_s)\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})} \\ & \leq \|R_{\lambda_s}(w) - R_{\lambda_s}(w_s)\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})} + \|R_{\lambda_s}(w_s) - R_s(w_s)\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})} \\ & \leq \frac{C}{\lambda_s}M + M \end{aligned}$$

where we have used (4.5). We arrive at

$$\|\mathcal{R}_{\lambda_s}(h) - \mathcal{R}_s(h)\|_{\mathcal{H}}^2 \leq CM^2 + \frac{C}{\lambda_s^2}M^2$$

which implies (4.4). □

Let us now turn to optimality conditions. We first need to prove the Fréchet differentiability of the smoothed solution operator of the constraint in (4.3). To this end, we need two norm gaps in Lemma 4.9 and Proposition 4.10, recall that the corresponding coefficients are fixed in Assumption 4.4 (ii).

Lemma 4.9 (Fréchet differentiability of \mathcal{T}_{R_s}). *The operator \mathcal{T}_{R_s} is from \mathcal{Y}_p into $W_D^{1,\hat{p}}(\Omega; \mathbb{R}^d)$ Fréchet differentiable and, for $h, g \in \mathcal{Y}_p$, $\eta := \mathcal{T}'_{R_s}(h)g$ is of class $W_D^{1,p}(\Omega; \mathbb{R}^d)$ and the unique solution of*

$$(4.6) \quad -\operatorname{div}(\mathbb{D}\nabla^s\eta + \mathbb{E}R'_s(\mathbb{E}^\top \nabla^s(u - h_1) + h_3)(\mathbb{E}^\top \nabla^s(\eta - g_1) + g_3))) = \frac{g_2}{\lambda_s} + \frac{g_1 - \eta}{\lambda_s^2}$$

for all $\varphi \in W_D^{1,p'}(\Omega; \mathbb{R}^d)$, where $u := \mathcal{T}_{R_s}(h)$.

Moreover, there exists a constant C such that the extension of $\mathcal{T}'_{R_s}(h)$ to an element of $\mathcal{L}(\mathcal{H}; H_D^1(\Omega; \mathbb{R}^d))$ fulfills $\|\mathcal{T}'_{R_s}(h)g\|_{H_D^1(\Omega; \mathbb{R}^d)} \leq C\|g\|_{\mathcal{H}}$ for all $h \in \mathcal{Y}_p$ and $g \in \mathcal{H}$.

Proof. Let $h, g \in \mathcal{Y}_p$. At first we prove that (4.6) has a unique solution $\eta \in W_D^{1,p}(\Omega; \mathbb{R}^d)$ with respect to h and g . For $\sigma \in \mathbb{R}_s^{d \times d}$ we define $b_\sigma : \Omega \times \mathbb{R}_s^{d \times d} \rightarrow \mathbb{R}_s^{d \times d}$ by

$$b_\sigma(x, \tau) := \mathbb{D}\tau + \mathbb{E}R'_s(\mathbb{E}^\top \nabla^s(u(x) - h_1(x)) + h_3(x))(\mathbb{E}^\top \tau + \sigma)$$

for almost all $x \in \Omega$ and all $\tau \in \mathbb{R}_s^{d \times d}$. The existence of η follows now from Corollary 3.9 (with $\varphi := -\mathbb{E}^\top \nabla^s g_1 + g_3$), when we have verified the requirements on b_σ therein. Moreover, Corollary 3.9 also shows that the solution operator of (4.6) is continuous with respect to $g \in \mathcal{Y}_p$ (clearly, it is also linear).

Clearly, $b_0(x, 0) = 0 \in L^\infty(\Omega; \mathbb{R}_s^{d \times d})$ and $b_\sigma(\cdot, \tau)$ is measurable as a pointwise limit of measurable functions (see [23, Corollary 3.1.5]), for all $\tau, \sigma \in \mathbb{R}_s^{d \times d}$. Moreover, we have

$$\begin{aligned} & (b_\sigma(x, \tau) - b_{\bar{\sigma}}(x, \bar{\tau})) : (\tau - \bar{\tau}) \\ & \geq \gamma_{\mathbb{D}}|\tau - \bar{\tau}|^2 + R'_s(w(x))(\mathbb{E}^\top(\tau - \bar{\tau}) + (\sigma - \bar{\sigma})) : \mathbb{E}^\top(\tau - \bar{\tau}) \\ & \geq \gamma_{\mathbb{D}}|\tau - \bar{\tau}|^2 - C|\sigma - \bar{\sigma}| |\tau - \bar{\tau}|, \end{aligned}$$

with $w := \mathbb{E}^\top \nabla^s(u - h) + h_3$, and

$$|b_\sigma(x, \tau) - b_{\bar{\sigma}}(x, \bar{\tau})| \leq C(|\tau - \bar{\tau}| + |\sigma - \bar{\sigma}|)$$

for all $\sigma, \bar{\sigma}, \tau, \bar{\tau} \in \mathbb{R}_s^{d \times d}$ and almost all $x \in \Omega$, where we have used Lemma 4.6 in both estimations. Therefore (3.5)–(3.8) are fulfilled.

Considering now the equations for $u_g := \mathcal{T}_{R_s}(h + g)$ and $u := \mathcal{T}_{R_s}(h)$, we see that

$$-\operatorname{div}(\mathbb{D}\nabla^s(u_g - u - \eta)) + \frac{u_g - u - \eta}{\lambda_s^2} = \operatorname{div}(\mathbb{E}(R_s(\mu + v_g) - R_s(\mu) - R'_s(\mu)v_g)) + \operatorname{div}(\mathbb{E}R'_s(\mu)((\mathbb{E}^\top \nabla^s(u_g - u - \eta))),$$

where

$$\begin{aligned} \mu &:= \mathbb{E}^\top \nabla^s(u - h_1) + h_3, \\ v_g &:= \mathbb{E}^\top \nabla^s(u_g - u - g_1) + g_3 \in L^p(\Omega; \mathbb{R}_s^{d \times d}), \end{aligned}$$

hence,

$$-\operatorname{div}(\mathbb{D}\nabla^s(u_g - u - \eta) - \mathbb{E}R'_s(\mu)((\mathbb{E}^\top \nabla^s(u_g - u - \eta)))) + \frac{u_g - u - \eta}{\lambda_s^2} = \operatorname{div} \mathbb{E}r_\mu(v_g),$$

where $r_\mu(v_g)$ is the remainder term of R_s at μ in direction v_g . Applying Corollary 3.9 once again with

$$b_\sigma(x, \tau) := \mathbb{D}\tau + \mathbb{E}R'_s(\mu(x))\mathbb{E}^\top \tau$$

(and $p = \hat{p}$) we obtain

$$\frac{\|u_g - u - \eta\|_{W^{1,\hat{p}}(\Omega; \mathbb{R}^d)}}{\|g\|_{\mathcal{Y}_p}} \leq C \frac{\|r_\mu(v_g)\|_{L^{\hat{p}}(\Omega; \mathbb{R}_s^{d \times d})}}{\|g\|_{\mathcal{Y}_p}} \leq C \frac{\|r_\mu(v_g)\|_{L^{\hat{p}}(\Omega; \mathbb{R}_s^{d \times d})}}{\|v_g\|_{L^p(\Omega; \mathbb{R}_s^{d \times d})}} \rightarrow 0,$$

as $g \rightarrow 0$ in \mathcal{Y}_p , where we also used the Lipschitz continuity of \mathcal{T}_{R_s} and the fact that $R_s : L^p(\Omega; \mathbb{R}_s^{d \times d}) \rightarrow L^{\hat{p}}(\Omega; \mathbb{R}_s^{d \times d})$ is Fréchet differentiable (cf. [10, Theorem 7]).

That the extension of $\mathcal{T}'_{R_s}(h)$ to an element of $L(\mathcal{H}; H^1_D(\Omega; \mathbb{R}^d))$ fulfills the asserted inequality, can be proven as above (one can simply test (4.6) with $\eta \in H^1_D(\Omega; \mathbb{R}^d)$ and use Lemma 4.6). \square

Proposition 4.10 (Fréchet differentiability of \mathcal{A}_s). *The mapping \mathcal{R}_s is from \mathcal{Y}_p to \mathcal{H} Fréchet differentiable and there exists a constant C such that the extension of $\mathcal{R}'_s(h) \in \mathcal{L}(\mathcal{Y}_p; \mathcal{H})$ to an element of $\mathcal{L}(\mathcal{H})$ fulfills $\|\mathcal{R}'_s(h)g\|_{\mathcal{H}} \leq C\|g\|_{\mathcal{H}}$ for all $h \in \mathcal{Y}_p$ and $g \in \mathcal{H}$.*

For $h \in \mathcal{Y}_p$ and $g \in \mathcal{H}$ we have

$$\mathcal{R}'_s(h)g = \begin{pmatrix} \mathcal{T}'_{R_s}(h)g \\ \frac{1}{\lambda_s} \mathcal{T}'_{R_s}(h)g - \frac{g_1}{\lambda_s} \\ R'_s(\mathbb{E}^\top \nabla^s(\mathcal{T}_{R_s}(h) - h_1) + h_3)(\mathbb{E}^\top \nabla^s(\mathcal{T}'_{R_s}(h)g - g_1) + g_3) \end{pmatrix}$$

The same is true for $\mathcal{A}_s = \frac{1}{\lambda_s}(I - \mathcal{R}_s)$ with $\mathcal{A}'_s(h)g = \frac{1}{\lambda_s}(g - \mathcal{R}'_s(h)g)$ for all $h \in \mathcal{Y}_p$ and $g \in \mathcal{H}$.

Proof. The assertion follows from Lemma 4.9, Lemma 4.6 for the estimate of $(\mathcal{R}'_s(h)g)_3$, the fact that $R_s : L^{\hat{p}}(\Omega; \mathbb{R}_s^{d \times d}) \rightarrow L^2(\Omega; \mathbb{R}_s^{d \times d})$ is Fréchet differentiable (cf. [10, Theorem 7]) and the chain rule. \square

Now, we can use [18, Theorem 5.5] to derive the differentiability of the solution operator of the constraint in (4.3) from the differentiability of \mathcal{A}_s . To this end, we first introduce the solution operator in

Definition 4.11 (Smoothed solution operator). We denote the solution operator of

$$(4.7) \quad Q^{-1}(\dot{u}, \dot{v}, \dot{q}) + \mathcal{A}_s(u, v, q) = Rf, \quad (u, v, q)(0) = (u_0, v_0, q_0)$$

by $\mathcal{S}_s : L^2(L^2(\Omega; \mathbb{R}^d)) \rightarrow H^1(\mathcal{Y}_p)$, that is, $\mathcal{S}_s(f) = (u, v, q)$, which existence follows from Banach's contraction principle since \mathcal{A}_s is Lipschitz continuous according to [Definition 4.7](#). Here we use the improved regularity of (u_0, v_0, q_0) , see [Assumption 4.4 \(iii\)](#).

Proposition 4.12 (Fréchet differentiability of the smoothed solution operator). *The solution operator $\mathcal{S}_s : L^2(L^2(\Omega; \mathbb{R}^d)) \rightarrow H^1(\mathcal{Y}_p)$ is Lipschitz continuous, $\mathcal{S}_s : H^1(L^2(\Omega; \mathbb{R}^d)) \rightarrow H^1(\mathcal{H})$ is Fréchet differentiable and, for $f, g \in H^1(L^2(\Omega; \mathbb{R}^d))$, $\eta := \mathcal{S}'_s(f)g \in H^1(\mathcal{H})$ is the unique solution of*

$$(4.8) \quad Q^{-1}\dot{\eta} + \mathcal{A}'_s(w)\eta = Rg, \quad \eta(0) = 0,$$

where $w := \mathcal{S}_s(f)$. Moreover, there exists a constant C , such that $\|\mathcal{S}'_s(f)g\|_{H^1(\mathcal{H})} \leq C\|g\|_{L^2(L^2(\Omega; \mathbb{R}^d))}$ holds for all $f, g \in H^1(L^2(\Omega; \mathbb{R}^d))$.

Proof. Our goal is to use [[18](#), Theorem 5.5], to this end we first consider the transformed equation from [Section 3.3](#). We again set $p_0 := -Q^{-1}(u_0, v_0, q_0)$ and denote the solution operator of

$$(4.9) \quad \dot{p} = \mathcal{A}_s(RF - Qp), \quad p(0) = p_0$$

by $\tilde{\mathcal{S}}_s : L^2(L^2(\Omega; \mathbb{R}^d)) \rightarrow H^1(\mathcal{Y}_p)$, that is, $\tilde{\mathcal{S}}_s(F) = p$. Thus we have $\mathcal{S}_s(f) = R\mathcal{F}_\rho f - Q\tilde{\mathcal{S}}_s(\mathcal{F}_\rho f)$ for all $f \in L^2(L^2(\Omega; \mathbb{R}^d))$.

We can now apply [[18](#), Lemma 5.3 & Theorem 5.5] (with $\mathcal{X} = L^2(\Omega; \mathbb{R}^d)$, $\mathcal{Y} = \mathcal{Y}_p$, $\mathcal{Z} = \mathcal{H}$, $z = p$ and $z_0 = p_0$), note that the assumptions in [[18](#), Assumption 5.1 (ii)] are satisfied thanks to [Proposition 4.10](#). Thus the solution operator $\tilde{\mathcal{S}}_s : L^2(L^2(\Omega; \mathbb{R}^d)) \rightarrow H^1(\mathcal{Y}_p)$ is Lipschitz continuous and $\tilde{\mathcal{S}}_s : H^1(L^2(\Omega; \mathbb{R}^d)) \rightarrow H^1(\mathcal{H})$ is Fréchet differentiable, hence, the desired Lipschitz continuity and Fréchet differentiability also hold for \mathcal{S}_s . Furthermore, the asserted inequality holds and we have $\eta = R\mathcal{F}_\rho g - Q\tilde{\eta}$, where $\eta := \mathcal{S}'_s(f)g$ and $\tilde{\eta} := \tilde{\mathcal{S}}'_s(\mathcal{F}_\rho f)\mathcal{F}_\rho g$. [[18](#), Theorem 5.5] also shows that $\tilde{\eta}$ is the unique solution of

$$\partial_t \tilde{\eta} = \mathcal{A}'_s(R\mathcal{F}_\rho f - Qp)(R\mathcal{F}_\rho g - Q\tilde{\eta}), \quad \tilde{\eta}(0) = 0,$$

where $p := \tilde{\mathcal{S}}_s(\mathcal{F}_\rho f)$. Taking into account that $\tilde{\eta} = R\mathcal{F}g - Q^{-1}\eta$ and $\partial_t \mathcal{F}g = g$, we see that η is the solution of [\(4.8\)](#). □

Remark 4.13 (Control space). As seen in the proposition above, the smoothed solution operator defined on $H^1(L^2(\Omega; \mathbb{R}^d))$ is Fréchet differentiable. The norm gaps, which arise from the exponents in [Assumption 4.4 \(ii\)](#), are only needed for the differentiability of \mathcal{T}_{R_s} but not in the control space. Unfortunately, we still require the compactness property imposed on \mathfrak{X}_c in [Theorem 4.2](#) to use the convergence results in [Section 3.3](#). However, we can avoid taking a subspace of $H^1(L^{\tilde{p}}(\Omega; \mathbb{R}^d))$, for a certain $\tilde{p} > 2$, as the control space.

Let us now consider the following reduced optimization problem

$$(4.10) \quad \min_{f \in \mathfrak{X}_c} F_z(f),$$

where the reduced objective function $F_z : \mathfrak{X}_c \rightarrow \mathbb{R}$ is defined by $F_z(f) := J_z(\mathcal{S}_s(f), f)$. Clearly, [\(4.10\)](#) and [\(4.3\)](#) are equivalent.

We can finally present our main result.

Theorem 4.14 (Optimality conditions for (4.10)). Let $\bar{f} \in \mathfrak{X}_c$ and abbreviate $(\bar{u}, \bar{v}, \bar{q}) := \mathcal{S}_s(\bar{f}) \in H^1(\mathcal{Y}_p)$ and $\bar{w} := \mathcal{T}_{R_s}(\bar{u}, \bar{v}, \bar{q}) \in H^1(W_D^{1,p}(\Omega; \mathbb{R}_s^{d \times d}))$. Then the variational equation

$$(4.11) \quad F'_z(\bar{f})g = \Psi'_z(\mathcal{S}_s(\bar{f}))\mathcal{S}'_s(\bar{f})g + \alpha \left(\bar{f}, g \right)_{\mathfrak{X}_c} = 0$$

holds for all $g \in \mathfrak{X}_c$ if and only if there exists a unique adjoint state $(\varphi, \eta^*) = (\varphi_1, \varphi_2, \varphi_3, \eta^*) \in H^1(\mathcal{H} \times H_D^1(\Omega; \mathbb{R}^d))$ such that the following optimality system is satisfied:

State equation:

$$(4.12a) \quad \begin{pmatrix} \dot{\bar{u}} \\ \dot{\bar{v}} \\ \dot{\bar{q}} \end{pmatrix} = \frac{1}{\lambda_s} \begin{pmatrix} \bar{w} - \bar{u} \\ (\bar{w} - \bar{u})/(\rho\lambda_s) - \bar{v}/\rho \\ (\mathbb{C} + \mathbb{B})(\bar{p} - \bar{q}) \end{pmatrix} + \begin{pmatrix} 0 \\ \bar{f}/\rho \\ 0 \end{pmatrix}$$

$$(4.12b) \quad -\operatorname{div}(\mathbb{D}\nabla^s \bar{w} + \mathbb{E}\bar{p}) = \bar{v}/\lambda_s + (\bar{w} - \bar{u})/\lambda_s$$

$$(4.12c) \quad \bar{p} = R_s(\mathbb{E}\nabla^s(\bar{w} - \bar{u}) + \bar{q})$$

$$(4.12d) \quad (\bar{u}, \bar{v}, \bar{q})(0) = (u_0, v_0, q_0)$$

Adjoint equation:

$$(4.12e) \quad \begin{pmatrix} \dot{\varphi}_1 \\ \dot{\varphi}_2 \\ \dot{\varphi}_3 \end{pmatrix} = \frac{1}{\lambda_s} \begin{pmatrix} \eta^* - \varphi_1 \\ (\eta^* - \varphi_1)/(\rho\lambda_s) - \varphi_2/\rho \\ (\mathbb{C} + \mathbb{B})(r^* - \varphi_3) \end{pmatrix} - Q\Psi'_z(\bar{u}, \bar{v}, \bar{q})$$

$$(4.12f) \quad -\operatorname{div}(\mathbb{D}\nabla^s \eta^* + \mathbb{E}r^*) = \varphi_2/\lambda_s + (\eta^* - \varphi_1)/\lambda_s$$

$$(4.12g) \quad r^* = R'_s(\mathbb{E}^\top \nabla^s(\bar{w} - \bar{u}) + \bar{q})^* (\mathbb{E}^\top \nabla^s(\eta^* - \varphi_1) + \varphi_3)$$

$$(4.12h) \quad (\varphi_1, \varphi_2, \varphi_3)(T) = 0$$

Gradient equation:

$$(4.12i) \quad (\varphi_2, g)_{L^2(L^2(\Omega; \mathbb{R}^d))} = \alpha \left(\bar{f}, g \right)_{\mathfrak{X}_c} \quad \forall g \in \mathfrak{X}_c.$$

In particular, if \bar{f} is locally optimal for (4.10), then there exists a unique adjoint state $(\varphi, \eta^*) \in H^1(\mathcal{H} \times H_D^1(\Omega; \mathbb{R}^d))$ such that (4.12) is fulfilled.

Proof. At first we proof that the assertion holds when we exchange (4.12) with

$$(4.13) \quad \begin{aligned} Q^{-1}(\dot{\bar{u}}, \dot{\bar{v}}, \dot{\bar{q}}) + \mathcal{A}_s(\bar{u}, \bar{v}, \bar{q}) &= R\bar{f}, \quad (\bar{u}, \bar{v}, \bar{q})(0) = (u_0, v_0, q_0), \\ Q^{-1}\dot{\varphi} + \mathcal{A}'_s(\bar{u}, \bar{v}, \bar{q})^* \varphi &= -\Psi'_z(\bar{u}, \bar{v}, \bar{q}), \quad \varphi(T) = 0, \\ (\varphi_2, g)_{L^2(L^2(\Omega; \mathbb{R}^d))} &= \alpha \left(\bar{f}, g \right)_{\mathfrak{X}_c} \quad \forall g \in \mathfrak{X}_c. \end{aligned}$$

To this end, let φ be the solution of the second equation in (4.13) (which unique existence follows as in [18, Lemma 5.11]) and $\eta := \mathcal{S}'_s(\bar{f})g \in H^1(\mathcal{H})$ for an arbitrary $g \in \mathfrak{X}_c$, then

$$\begin{aligned} (\varphi_2, g)_{L^2(L^2(\Omega; \mathbb{R}^d))} &= (\varphi, Rg)_{L^2(\mathcal{H})} = (\varphi, Q^{-1}\dot{\eta})_{L^2(\mathcal{H})} + (\varphi, \mathcal{A}'_s(\bar{u}, \bar{v}, \bar{q})\eta)_{L^2(\mathcal{H})} \\ &= (Q^{-1}\dot{\varphi}, \eta)_{L^2(\mathcal{H})} + (\mathcal{A}'_s(\bar{u}, \bar{v}, \bar{q})^* \varphi, \eta)_{L^2(\mathcal{H})} \\ &= -\left(\Psi'_z(\mathcal{S}_s(\bar{f})), \eta \right)_{L^2(\mathcal{H})} \end{aligned}$$

holds for all $g \in \mathfrak{X}_c$, which implies the equivalence between (4.11) and the last equation in (4.13). Moreover, it is well known that if \bar{f} is locally optimal for (4.3), then (4.11) must hold.

Let us now prove the equivalence between (4.13) and (4.12). We choose $h, \xi \in \mathcal{H}$ and denote by $\eta^* \in H_D^1(\Omega; \mathbb{R}_s^{d \times d})$ the solution of

$$(4.14) \quad -\operatorname{div}(\mathbb{D}\nabla^s \eta^* + \mathbb{E}R'_s(\mathbb{E}^\top \nabla^s(\mathcal{T}_{R_s}(h) - h_1) + h_3)^*(\mathbb{E}^\top \nabla^s(\eta^* - \xi_1) + \xi_3))) = \frac{\xi_2}{\lambda_s} + \frac{\xi_1 - \eta^*}{\lambda_s^2}$$

for all $\phi \in H_D^1(\Omega; \mathbb{R}^d)$ (the existence of η^* follows as in Lemma 4.9, note that the inequalities in Lemma 4.6 hold also for the adjoint operator). Then

$$\mathcal{R}'_s(h)^* \xi = \begin{pmatrix} \eta^* \\ \frac{1}{\lambda_s} \eta^* - \frac{\xi_1}{\lambda_s} \\ R'_s(\mathbb{E}^\top \nabla^s(\mathcal{T}_{R_s}(h) - h_1) + h_3)^*(\mathbb{E}^\top \nabla^s(\eta^* - \xi_1) + \xi_3) \end{pmatrix}$$

holds, which can be seen as follows: Let $g \in \mathcal{H}$ and abbreviate

$$\begin{aligned} \eta &:= \mathcal{T}'_{R_s}(h)g, & \eta_v &:= \frac{\eta - g_1}{\lambda_s}, & \eta_q &:= R'_s(\mathbb{E}^\top \nabla^s(\mathcal{T}_{R_s}(h) - h_1) + h_3)(\mathbb{E}^\top \nabla^s(\eta - g_1) + g_3), \\ \eta_v^* &:= \frac{\eta^* - \xi_1}{\lambda_s}, & \eta_q^* &:= R'_s(\mathbb{E}^\top \nabla^s(\mathcal{T}_{R_s}(h) - h_1) + h_3)^*(\mathbb{E}^\top \nabla^s(\eta^* - \xi_1) + \xi_3). \end{aligned}$$

Testing (4.6) with $\phi = \xi_1 - \eta^*$ gives

$$\begin{aligned} &(\mathbb{D}\nabla^s \eta, \nabla^s(\xi_1 - \eta^*))_{L^2(\Omega; \mathbb{R}_s^{d \times d})} + (\eta_v - g_2, \eta_v^*)_{L^2(\Omega; \mathbb{R}^d)} \\ &= (\mathbb{E}\eta_q, \nabla^s(\eta^* - \xi_1))_{L^2(\Omega; \mathbb{R}_s^{d \times d})} \\ &= (\mathbb{E}^\top \nabla^s(\eta - g_1) + g_3, \eta_q^*)_{L^2(\Omega; \mathbb{R}_s^{d \times d})} - (\eta_q, \xi_3)_{L^2(\Omega; \mathbb{R}_s^{d \times d})}, \end{aligned}$$

and testing (4.14) with $\phi = \eta - g_1$ yields

$$(\mathbb{D}\nabla^s \eta^*, \nabla^s(\eta - g_1))_{L^2(\Omega; \mathbb{R}_s^{d \times d})} + (\xi_2 - \eta_v^*, \eta_v)_{L^2(\Omega; \mathbb{R}^d)} = (\mathbb{E}\eta_q^*, \nabla^s(g_1 - \eta))_{L^2(\Omega; \mathbb{R}_s^{d \times d})},$$

thus, adding both equations together, we arrive at

$$\begin{aligned} &(\mathbb{D}\nabla^s \eta, \nabla^s \xi_1)_{H^1(\Omega; \mathbb{R}^d)} + (\eta_v, \xi_2)_{L^2(\Omega; \mathbb{R}_s^{d \times d})} + (\eta_q, \xi_3)_{L^2(\Omega; \mathbb{R}_s^{d \times d})} \\ &= (\mathbb{D}\nabla^s \eta^*, \nabla^s g_1)_{H^1(\Omega; \mathbb{R}^d)} + (\eta_v^*, g_2)_{L^2(\Omega; \mathbb{R}_s^{d \times d})} + (\eta_q^*, g_3)_{L^2(\Omega; \mathbb{R}_s^{d \times d})}, \end{aligned}$$

which is equivalent to

$$(\mathcal{R}'_s(h)g, \xi)_{\mathcal{H}} = (g, \mathcal{R}'_s(h)^* \xi)_{\mathcal{H}}.$$

Now one only has to use the definitions of \mathcal{A}_s and R to obtain the equivalence between (4.12) and (4.13). □

5 EXAMPLES

Let us conclude with examples about a concrete objective function, the gradient equation in Theorem 4.14 regarding a concrete control space and finally a realization of the maximal monotone operator A (which will be the von-Mises flow rule).

Objective function Let us consider a tracking type objective function, that is,

$$\Psi(u, v, z) = \frac{1}{2} \|(u, v, z) - (u_d, v_d, z_d)\|_{L^2(\mathcal{H})}^2$$

with a desired state $(u_d, v_d, z_d) \in L^2(\mathcal{H})$. Then

$$\Psi_z(u, v, q) = \frac{1}{2} \|(u, v, (\mathbb{C} + \mathbb{B})^{-1}(\mathbb{C}\nabla^s u - q)) - (u_d, v_d, z_d)\|_{L^2(\mathcal{H})}^2$$

and

$$\Psi'_z(u, v, q) = \begin{pmatrix} \hat{u} \\ v - v_d \\ ((\mathbb{C} + \mathbb{B})^{-1}(\mathbb{C}\nabla^s u - q)) - z_d \end{pmatrix}$$

where \hat{u} is such that $-\operatorname{div}(\mathbb{D}\nabla^s(\hat{u} - u + u_d) - ((\mathbb{C} + \mathbb{B})^{-1}(\mathbb{C}\nabla^s u - q) - z_d)) = 0$, hence, in this example the adjoint equation in (4.12) has to be completed by this equation. Note that when one uses a finite element approach to solve (4.12) numerically, then one can eliminate this additional equation after multiplying (4.12e) with a test function, that is, taking the \mathcal{H} -scalar product. When the \mathcal{H} -scalar product of $Q\Psi'_z(u, v, q)$ and a test function (η_1, η_2, η_3) is evaluated, the term $(\mathbb{D}\nabla^s \hat{u}, \eta_1)_{L^2(\Omega; \mathbb{R}_s^{d \times d})}$ arises, then one can use the additional equation to eliminate \hat{u} (respectively the equation).

Control space Let us consider the space

$$\mathfrak{X}_c := \{f \in H^1(L^2(\Omega; \mathbb{R}^d)) \cap L^2(H^1(\Omega; \mathbb{R}^d)) : f(0) = f(T) = 0\}$$

with the scalar product

$$(f, g)_{\mathfrak{X}_c} = (\dot{f}, \dot{g})_{L^2(L^2(\Omega; \mathbb{R}^d))} + (\nabla f, \nabla g)_{L^2 L^2(\Omega; \mathbb{R}^{d \times d})},$$

see Example 4.3. The Gradient equation in (4.12) then becomes

$$\alpha (\dot{f}, \dot{g})_{L^2(L^2(\Omega; \mathbb{R}^d))} + \alpha (\nabla f, \nabla g)_{L^2 L^2(\Omega; \mathbb{R}^{d \times d})} = (\varphi_2, g)_{L^2(L^2(\Omega; \mathbb{R}^d))}$$

for all $g \in \mathfrak{X}_c$, which is the weak formulation of

$$\ddot{f} + \Delta f = -\frac{\varphi_2}{\alpha}.$$

Maximal monotone operator A We consider the case of linear kinematic hardening with the von Mises yield condition, cf. [13] for a detailed description of this model. In this case, A is the convex subdifferential of the indicator functional $I_{\mathcal{K}(\Omega)}$ of the following set of admissible stresses

$$\mathcal{K}(\Omega) := \{\tau \in L^2(\Omega; \mathbb{R}_s^{d \times d}) : |\tau^D(x)| \leq \gamma \text{ f.a.a. } x \in \Omega\},$$

where $\tau^D := \tau - \frac{1}{d} \operatorname{tr}(\tau)I$ is the deviator of $\tau \in \mathbb{R}_s^{d \times d}$ and γ denotes the initial uni-axial yield stress, a given material parameter. The domain of $A = \partial I_{\mathcal{K}(\Omega)}$ is trivially $\mathcal{K}(\Omega)$, which is nonempty, closed and convex. For the Yosida approximation of $\partial I_{\mathcal{K}(\Omega)}$, one obtains by a straightforward calculation

$$(5.1) \quad A_\lambda(\tau) = \frac{1}{\lambda}(\tau - \pi_{\mathcal{K}}(\tau)) = \frac{1}{\lambda} \max \left\{ 0, 1 - \frac{\gamma}{|\tau^D|} \right\} \tau^D,$$

with the resolvent $R_\lambda = \pi_{\mathcal{K}(\Omega)}$, where $\pi_{\mathcal{K}(\Omega)}$ denotes the projection on $\mathcal{K}(\Omega)$ in $L^2(\Omega; \mathbb{R}_s^{d \times d})$, i.e., $\pi_{\mathcal{K}}(\sigma) := \arg \min_{\tau \in \mathcal{K}(\Omega)} \|\tau - \sigma\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})}^2$. We get in particular

$$R_\lambda(\tau) = \pi_{\mathcal{K}(\Omega)}(\tau) = \tau - \max \left\{ 0, 1 - \frac{Y}{|\tau^D|} \right\} \tau^D,$$

so that (2.2) is fulfilled.

To smoothen this function let $s > 0$ and

$$\max_s : \mathbb{R} \rightarrow \mathbb{R} \quad r \mapsto \begin{cases} \max\{0, r\}, & |r| \geq s, \\ \frac{1}{4s}(r+s)^2, & |r| < s, \end{cases}$$

we then set

$$R_s : \mathbb{R}_s^{d \times d} \rightarrow \mathbb{R}_s^{d \times d}, \quad \tau \mapsto \tau - \max_s \left(1 - \frac{Y}{|\tau^D|} \right) \tau^D.$$

One easily checks that $\max_s \in C^1(\mathbb{R})$ and we obtain

$$\begin{aligned} (5.2) \quad & \|\partial I_{\lambda,s}(\tau) - \partial I_\lambda(\tau)\|_{L^2(\Omega; \mathbb{R}_s^{d \times d})} \\ & \leq \frac{1}{\lambda} \left(\int_\Omega \left| \max_s \left(1 - \frac{Y}{|\tau(x)^D|} \right) - \max \left(0, 1 - \frac{Y}{|\tau(x)^D|} \right) \right|^2 |\tau(x)^D|^2 \right)^{1/2} \\ & \leq \frac{|\Omega| \gamma s}{4\lambda(1-s)} \end{aligned}$$

for all $\tau \in L^2(\Omega; \mathbb{R}_s^{d \times d})$. Moreover, it is also easy to verify that $R_s : \mathbb{R}_s^{d \times d} \rightarrow \mathbb{R}_s^{d \times d}$ is Fréchet differentiable with

$$R'_s(\tau)h = h - \max'_s \left(1 - \frac{Y}{|\tau^D|} \right) \frac{Y}{|\tau^D|^3} (\tau^D : h^D) \tau^D - \max_s \left(1 - \frac{Y}{|\tau^D|} \right) h^D$$

and the following lemma shows that R_s is monotone and Lipschitz continuous, thus Assumption 4.4 (i) is satisfied. Note also that $R'_s(\tau)$ is self adjoint for every $\tau \in \mathbb{R}_s^{d \times d}$, hence, $\mathcal{R}'_s(h)$ is also self adjoint for all $h \in \mathcal{H}$ (cf. the proof of Theorem 4.14).

Lemma 5.1. *For every $s \in (0, 1)$, the mapping R_s is monotone and Lipschitz continuous with constant 1.*

Proof. It is well known that, since \max_s is continuously differentiable and convex,

$$(5.3) \quad \max_s(x) - \max_s(y) \geq \max_s(y)'(x - y)$$

holds for all $x, y \in \mathbb{R}$.

Let $\tau, \sigma \in \mathbb{R}_s^{d \times d}$, w.l.o.g. we can assume that $|\sigma^D| \geq |\tau^D| > 0$ and

$$\max_s \left(1 - \frac{Y}{|\sigma^D|} \right) \geq \max_s \left(1 - \frac{Y}{|\tau^D|} \right),$$

then, using (5.3) with $x = 1 - \frac{Y}{|\tau^D|}$ and $y = 1 - \frac{Y}{|\sigma^D|}$, we get

$$\begin{aligned} & \left| \max_s \left(1 - \frac{Y}{|\tau^D|} \right) \tau^D - \max_s \left(1 - \frac{Y}{|\sigma^D|} \right) \sigma^D \right| \\ & \leq \left(\max_s \left(1 - \frac{Y}{|\sigma^D|} \right) - \max_s \left(1 - \frac{Y}{|\tau^D|} \right) \right) |\tau^D| + \max_s \left(1 - \frac{Y}{|\sigma^D|} \right) |\tau^D - \sigma^D| \\ & \leq \gamma \max'_s \left(1 - \frac{Y}{|\sigma^D|} \right) |\tau^D| \left| \frac{1}{|\tau^D|} - \frac{1}{|\sigma^D|} \right| + \max_s \left(1 - \frac{Y}{|\sigma^D|} \right) |\tau^D - \sigma^D|, \end{aligned}$$

taking into account that

$$\gamma |\tau^D| \left| \frac{1}{|\tau^D|} - \frac{1}{|\sigma^D|} \right| = \gamma \left| \frac{|\tau^D| - |\sigma^D|}{|\sigma^D|} \right| \leq \frac{\gamma}{|\sigma^D|} |\tau^D - \sigma^D|$$

we obtain

$$\begin{aligned} & \left| \max_s \left(1 - \frac{\gamma}{|\tau^D|} \right) \tau^D - \max_s \left(1 - \frac{\gamma}{|\sigma^D|} \right) \sigma^D \right| \\ & \leq \left(\max_s' \left(1 - \frac{\gamma}{|\sigma^D|} \right) \frac{\gamma}{|\sigma^D|} + \max_s \left(1 - \frac{\gamma}{|\sigma^D|} \right) \right) |\tau^D - \sigma^D| \\ & \leq \max_s(1) |\tau - \sigma| = |\tau - \sigma|, \end{aligned}$$

where we used (5.3) again with $x = 1$ and $y = 1 - \frac{\gamma}{|\sigma^D|}$, and the fact that $|\tau^D - \sigma^D| \leq |\tau - \sigma|$. This proves the Lipschitz continuity of R_s . We also get

$$\begin{aligned} & (R_s(\tau) - R_s(\sigma)) : (\tau - \sigma) \\ & = |\tau - \sigma|^2 - \left(\max_s \left(1 - \frac{\gamma}{|\tau^D|} \right) \tau^D - \max_s \left(1 - \frac{\gamma}{|\sigma^D|} \right) \sigma^D \right) : (\tau - \sigma) \\ & \geq |\tau - \sigma|^2 - |\tau - \sigma|^2 = 0 \end{aligned}$$

which shows the monotonicity of R_s . □

ACKNOWLEDGEMENTS

I would like to thank Christian Meyer (Technische Universität Dortmund) for fruitful discussions regarding [Corollary 3.9](#).

REFERENCES

- [1] G. Anzellotti and S. Luckhaus, Dynamical evolution of elasto-perfectly plastic bodies, *Applied Mathematics and Optimization* 15 (1987), 121–140, [doi:10.1007/bf01442650](https://doi.org/10.1007/bf01442650).
- [2] J. F. Babadjian and M. G. Mora, Approximation of dynamic and quasi-static evolution problems in elasto-plasticity by cap models, *Quarterly of Applied Mathematics* (2015), 265–316.
- [3] S. Bartels and T. Roubíček, Thermoviscoplasticity at small strains, *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik: Applied Mathematics and Mechanics* 88 (2008), 735–754, [doi:10.1002/zamm.200800042](https://doi.org/10.1002/zamm.200800042).
- [4] L. Boron and E. Zeidler, *Nonlinear Functional Analysis and its Applications: III: Variational Methods and Optimization*, Springer New York, 1984, [doi:10.1007/978-1-4612-5020-3](https://doi.org/10.1007/978-1-4612-5020-3).
- [5] H. Brézis, *Opérateurs Maximaux Monotones*, North-Holland, Amsterdam, 1973.
- [6] S. W. Christian Meyer, Optimal control of perfect plasticity part I: Stress tracking, *Mathematical Control & Related Fields* (2021), [doi:10.3934/mcrf.2021022](https://doi.org/10.3934/mcrf.2021022).
- [7] G. Dal Maso and R. Scala, Quasistatic evolution in perfect plasticity as limit of dynamic processes, *Journal of Dynamics and Differential Equations* 26 (2014), 915–954, [doi:10.1007/s10884-014-9409-7](https://doi.org/10.1007/s10884-014-9409-7).
- [8] E. Davoli, T. Roubíček, and U. Stefanelli, Dynamic perfect plasticity and damage in viscoelastic solids, *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik* 99 (2019), e201800161, [doi:10.1002/zamm.201800161](https://doi.org/10.1002/zamm.201800161).

- [9] E. Davoli and U. Stefanelli, Dynamic perfect plasticity as convex minimization, *SIAM Journal on Mathematical Analysis* 51 (2019), 672–730, doi:10.1137/17m1148864.
- [10] H. Goldberg, W. Kampschky, and F. Tröltzsch, On Nemytskii operators in L^p -spaces of abstract functions, *Mathematische Nachrichten* 155 (1992), 127–140, doi:10.1002/mana.19921550110.
- [11] K. Gröger, *Initial Value Problems for Elastoplastic and Elasto-viscoplastic Systems*, BSB BG Teubner Verlagsgesellschaft, 1979.
- [12] K. Gröger, A $W^{1,p}$ -estimate for solutions to mixed boundary value problems for second order elliptic differential equations, *Math. Ann.* 283 (1989), 679–687, doi:10.1007/bfo1442860.
- [13] W. Han and B.D. Reddy, *Plasticity: Mathematical Theory and Numerical Analysis*, volume 9, Springer Science & Business Media, 2012, doi:10.1007/978-1-4614-5940-8.
- [14] R. Herzog, C. Meyer, and G. Wachsmuth, Integrability of displacement and stresses in linear and nonlinear elasticity with mixed boundary conditions, *Journal of Mathematical Analysis and Applications* 382 (2011), 802–813, doi:10.1016/j.jmaa.2011.04.074.
- [15] R. Herzog, C. Meyer, and G. Wachsmuth, B- and strong stationarity for optimal control of static plasticity with hardening, *SIAM Journal on Optimization* 23 (2013), 321–352, doi:10.1137/110821147.
- [16] J. Lubliner, *Plasticity Theory*, Courier Corporation, 2008.
- [17] A. Maury, G. Allaire, and F. Jouve, Elasto-plastic shape optimization using the level set method, *SIAM Journal on Control and Optimization* 56 (2018), 556–581, doi:10.1137/17m1128940.
- [18] H. Meinschmidt, C. Meyer, and S. Walther, Optimal control of an abstract evolution variational inequality with application to homogenized plasticity, *Journal of Nonsmooth Analysis and Optimization* Volume 1 (2020), doi:10.46298/jnsao-2020-5800, <https://jnsao.episciences.org/6467>.
- [19] C. Meyer and S. Walther, Optimal control of perfect plasticity part II: Displacement tracking, *SIAM Journal on Control and Optimization* 59 (2021), 2498–2523, doi:10.1137/20m1327331.
- [20] N. Ottosen and M. Ristinmaa, *The Mechanics of Constitutive Modeling*, Elsevier, Amsterdam, 2005, doi:10.1016/b978-0-08-044606-6.x5000-0.
- [21] R. E. Showalter, *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*, volume 49, American Mathematical Soc., 2013.
- [22] J. Simo and T. Hughes, *Computational Inelasticity*, Interdisciplinary Applied Mathematics, Springer, New York, 1998, doi:10.1007/b98904.
- [23] G. Wachsmuth, *Optimal Control of Quasistatic Plasticity*, Dr. Hut, 2011.
- [24] G. Wachsmuth, Optimal control of quasi-static plasticity with linear kinematic hardening, Part I: Existence and discretization in time, *SIAM J. Control Optim.* 50 (2012), 2836–2861 + loose erratum, doi:10.1137/110839187.
- [25] G. Wachsmuth, Optimal control of quasistatic plasticity with linear kinematic hardening II: Regularization and differentiability, *Z. Anal. Anwend.* 34 (2015), 391–418, doi:10.4171/zaa/1546.
- [26] G. Wachsmuth, Optimal control of quasistatic plasticity with linear kinematic hardening III: Optimality conditions, *Z. Anal. Anwend.* 35 (2016), 81–118, doi:10.4171/zaa/1556.

- [27] S. Walther, *Optimal Control of Plasticity Systems*, PhD thesis, Universitätsbibliothek Dortmund, 2020, [doi:10.17877/de290r-21905](https://doi.org/10.17877/de290r-21905).
- [28] I. Yousept, Optimal control of non-smooth hyperbolic evolution Maxwell equations in type-II superconductivity, *SIAM Journal on Control and Optimization* 55 (2017), 2305–2332, [doi:10.1137/16m1074229](https://doi.org/10.1137/16m1074229).
- [29] E. Zeidler and L. Boron, *Nonlinear Functional Analysis and Its Applications: II/ A: Linear Monotone Operators*, Monotone operators / transl. by the author and by Leo F. Boron, Springer New York, 1989, [doi:10.1007/978-1-4612-0985-0](https://doi.org/10.1007/978-1-4612-0985-0).