

B- AND STRONG STATIONARITY FOR OPTIMAL CONTROL OF STATIC PLASTICITY WITH HARDENING

ROLAND HERZOG, CHRISTIAN MEYER, AND GERD WACHSMUTH

ABSTRACT. Optimal control problems for the variational inequality of static elastoplasticity with linear kinematic hardening are considered. The control-to-state map is shown to be weakly directionally differentiable, and local optimal controls are proved to verify an optimality system of B-stationary type. For a modified problem, local minimizers are shown to even satisfy an optimality system of strongly stationary type.

1 Introduction

In this paper we continue the investigation of first-order necessary optimality conditions for optimal control problems in static elastoplasticity. The forward system in the stress-based (so-called dual) form is represented by a variational inequality (VI) of mixed type: find generalized stresses $\Sigma \in S^2$ and displacements $\mathbf{u} \in V$ which satisfy $\Sigma \in \mathcal{K}$ and

$$\left. \begin{aligned} \langle A\Sigma, \mathbf{T} - \Sigma \rangle + \langle B^*\mathbf{u}, \mathbf{T} - \Sigma \rangle &\geq 0 \quad \text{for all } \mathbf{T} \in \mathcal{K} \\ B\Sigma &= \ell \quad \text{in } V'. \end{aligned} \right\} \quad (\text{VI})$$

where A and B are linear operators. The closed, convex set $\mathcal{K} \subset S^2$ of admissible stresses is determined by the von Mises yield condition. The details require a certain amount of notation and are made precise below.

The optimization of elastoplastic systems is of significant importance for industrial deformation processes. We emphasize that, in spite of its limited physical importance itself, the static problem (VI) appears as a time step of its quasi-static variant, which will be investigated elsewhere. We will consider primarily the following prototypical optimal control problem

$$\left. \begin{aligned} \text{Minimize } & J(\mathbf{u}, \mathbf{g}) \\ \text{s.t. } & \text{the plasticity problem (VI) with } \ell \in V' \text{ defined by} \\ & \langle \ell, \mathbf{v} \rangle = - \int_{\Gamma_N} \mathbf{g} \cdot \mathbf{v} \, ds, \quad \mathbf{v} \in V \\ \text{and } & \mathbf{g} \in U_{\text{ad}} \end{aligned} \right\} \quad (\text{P})$$

in which the boundary loads \mathbf{g} appear as control variables. The details are made precise below.

The optimal control of (VI) leads to an infinite dimensional MPEC (mathematical program with equilibrium constraints). The derivation of necessary optimality conditions is challenging due to the lack of Fréchet differentiability of the associated control-to-state map $\ell \mapsto (\Sigma, \mathbf{u})$. The same is true for the re-formulation of (VI) as a complementarity system involving the so-called plastic multiplier. It is well

known that for the resulting MPCC (mathematical program with complementarity constraints) classical constraint qualifications fail to hold.

To overcome these difficulties, several competing stationarity concepts for MPCCs have been developed, see for instance [Scheel and Scholtes \[2000\]](#) for an overview in the finite dimensional case up to the year 2000 and [Flegel and Kanzow \[2005\]](#), [Kanzow and Schwartz \[2010\]](#), [Steffensen and Ulbrich \[2010\]](#) and the references therein for recent further developments. It was shown recently in [Herzog et al. \[2012\]](#) that local optima of (\mathbf{P}) satisfy first-order optimality conditions of C-(Clarke)-stationary type. This was achieved by approximating them by sequences of solutions to regularized problems. In these regularized problems, (\mathbf{VI}) is replaced by a smooth equation. We exemplarily refer to [Barbu \[1984\]](#), [Hintermüller \[2001\]](#) for related results for optimal control of the obstacle problem.

In the present paper, we pursue a different approach. We prove that local optima of (\mathbf{P}) satisfy first-order optimality conditions of B-(Bouligand)-stationary type. This optimality concept is based solely on primal variables, and the main step in the proof is to establish the weak directional differentiability of the control-to-state map.

Moreover, we show that for a modified problem, local optima are even strongly stationary. In order to obtain this result, we suppose that the modified problem has so-called “ample” controls, i.e., distributed control functions which act on both right-hand sides of (\mathbf{VI}) . In addition, we dispose of control constraints in the modified problem. These modifications are in accordance with previous strong stationarity results for the optimal control of the obstacle problem, see [Mignot and Puel \[1984\]](#). However, we present a different and more elementary technique of proof.

Let us put our work into perspective. In contrast to the multitude of papers concerning regularization and *C-stationarity* conditions for infinite dimensional MPECs, see e.g., [Friedman \[1986\]](#), [Bonnans and Tiba \[1991\]](#), [Bergounioux \[1997\]](#), [Bergounioux \[1998\]](#), [Bergounioux and Zidani \[1999\]](#), [Ito and Kunisch \[2000, 2010\]](#), [Hintermüller \[2001, 2008\]](#), [Hintermüller and Kopacka \[2009\]](#), [Hintermüller et al. \[2009\]](#), [Zhu \[2006\]](#), [Farshbaf-Shaker \[2011\]](#), there are fewer contributions which address the question of stricter optimality conditions. We refer to the classical paper of [Mignot and Puel \[1984\]](#), where the obstacle problem is discussed. In [Hintermüller and Kopacka \[2009\]](#) conditions are derived which guarantee the convergence of stationary points of regularized problems to *strongly stationary* points in case of optimal control of the obstacle problem. It is to be noted that these conditions depend on the regularized sequence itself and cannot be guaranteed a priori. Recently, [Outrata et al. \[2011\]](#) confirmed the strong stationarity result of [Mignot and Puel \[1984\]](#) for the obstacle problem by using a completely different technique based on results of [Jarušek and Outrata \[2007\]](#). To the authors’ knowledge, B- and strong stationarity results for optimal control problems governed by variational inequalities other than of obstacle type (as for instance (\mathbf{VI})) have not been discussed so far.

The paper is organized as follows. In the remainder of this section we introduce the notation and state our generic assumptions. In [Section 2](#), we review some results about the forward problem (\mathbf{VI}) . The proof of B-stationarity is achieved in [Section 3](#). [Section 4](#) is devoted to the investigation of strong stationarity for the modified problem.

Notation and Assumptions. Our notation for the forward problem follows [Han and Reddy \[1999\]](#) and [Herzog and Meyer \[2011\]](#). We restrict the discussion to the case of linear kinematic hardening.

Function Spaces. Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with Lipschitz boundary Γ in dimension $d \in \{2, 3\}$. This assumption is made more precise in [Assumption 1.1](#) (1). We point out that the presented analysis is not restricted to the case $d \leq 3$, but for reasons of physical interpretation we focus on the two and three dimensional case. The boundary consists of two disjoint parts Γ_N and Γ_D , on which boundary loads and zero displacement conditions are imposed, respectively. We denote by $\mathbb{S} := \mathbb{R}_{\text{sym}}^{d \times d}$ the space of symmetric d -by- d matrices, endowed with the inner product $\mathbf{A} : \mathbf{B} = \sum_{i,j=1}^d A_{ij} B_{ij}$, and we define

$$V = H_D^1(\Omega; \mathbb{R}^d) = \{\mathbf{u} \in H^1(\Omega; \mathbb{R}^d) : \mathbf{u} = \mathbf{0} \text{ on } \Gamma_D\}, \quad S = L^2(\Omega; \mathbb{S}).$$

Here, V is the space for the displacement \mathbf{u} and S is the space for both, the stress $\boldsymbol{\sigma}$, and the back stress $\boldsymbol{\chi}$. We refer to $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi}) \in S^2$ as the generalized stress. The boundary control \mathbf{g} belongs to the space $L^2(\Gamma_N; \mathbb{R}^d)$.

Yield Function and Admissible Stresses. We restrict our discussion to the von Mises yield function. In the context of linear kinematic hardening, it reads

$$\phi(\boldsymbol{\Sigma}) = (|\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D|^2 - \bar{\sigma}_0^2)/2 \quad (1.1)$$

for $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi}) \in S^2$, where $|\cdot|$ denotes the pointwise Frobenius norm of matrices and

$$\boldsymbol{\sigma}^D = \boldsymbol{\sigma} - (1/d) (\text{trace } \boldsymbol{\sigma}) \mathbf{I}$$

is the deviatoric part of $\boldsymbol{\sigma}$. The yield function gives rise to the set of admissible generalized stresses

$$\mathcal{K} = \{\boldsymbol{\Sigma} \in S^2 : \phi(\boldsymbol{\Sigma}) \leq 0 \text{ a.e. in } \Omega\}. \quad (1.2)$$

Due to the structure of the yield function, $\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D$ appears frequently and we abbreviate it and its adjoint by

$$\mathcal{D}\boldsymbol{\Sigma} = \boldsymbol{\sigma}^D + \boldsymbol{\chi}^D \quad \text{and} \quad \mathcal{D}^*\boldsymbol{\sigma} = \begin{pmatrix} \boldsymbol{\sigma}^D \\ \boldsymbol{\sigma}^D \end{pmatrix}$$

for matrices $\boldsymbol{\Sigma} \in S^2$ as well as for functions $\boldsymbol{\Sigma} \in S^2$. When considered as an operator in function space, \mathcal{D} maps $S^2 \rightarrow S$. For later reference, we also remark that

$$\mathcal{D}^*\mathcal{D}\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma}^D + \boldsymbol{\chi}^D \\ \boldsymbol{\sigma}^D + \boldsymbol{\chi}^D \end{pmatrix} \quad \text{and} \quad (\mathcal{D}^*\mathcal{D})^2 = 2\mathcal{D}^*\mathcal{D}$$

holds.

Operators. The linear operators $A : S^2 \rightarrow S^2$ and $B : S^2 \rightarrow V'$ appearing in [\(VI\)](#) are defined as follows. For $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi}) \in S^2$ and $\mathbf{T} = (\boldsymbol{\tau}, \boldsymbol{\mu}) \in S^2$, let $A\boldsymbol{\Sigma}$ be defined through

$$\langle \mathbf{T}, A\boldsymbol{\Sigma} \rangle = \int_{\Omega} \boldsymbol{\tau} : \mathbb{C}^{-1} \boldsymbol{\sigma} \, dx + \int_{\Omega} \boldsymbol{\mu} : \mathbb{H}^{-1} \boldsymbol{\chi} \, dx. \quad (1.3)$$

The term $(1/2) \langle A\boldsymbol{\Sigma}, \boldsymbol{\Sigma} \rangle$ corresponds to the energy associated with the stress state $\boldsymbol{\Sigma}$. Here $\mathbb{C}^{-1}(x)$ and $\mathbb{H}^{-1}(x)$ are linear maps from \mathbb{S} to \mathbb{S} (i.e., they are fourth order tensors) which may depend on the spatial variable x . For $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi}) \in S^2$ and $\mathbf{v} \in V$, let

$$\langle B\boldsymbol{\Sigma}, \mathbf{v} \rangle = - \int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{\varepsilon}(\mathbf{v}) \, dx. \quad (1.4)$$

We recall that $\boldsymbol{\varepsilon}(\mathbf{v}) = \frac{1}{2}(\nabla \mathbf{v} + (\nabla \mathbf{v})^\top)$ denotes the (linearized) strain tensor.

Here and throughout, $\langle \cdot, \cdot \rangle$ denotes the dual pairing between V and its dual V' , or the scalar products in S or S^2 , respectively. Moreover, $(\cdot, \cdot)_E$ refers to the scalar product of $L^2(E)$ where $E \subset \Omega$ or $E = \Gamma_N$.

Assumption 1.1 (Standing Assumption).

- (1) The boundary Γ of the domain $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$ is Lipschitz, i.e., the boundary consists of a finite number of local graphs of Lipschitz maps, see, e.g., [Grisvard, 1985, Definition 1.2.1.1]. Moreover, the boundary is assumed to consist 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 . Furthermore Γ_D is assumed to have positive measure.
- (2) The yield stress $\tilde{\sigma}_0$ is assumed to be a positive constant. It equals $\sqrt{2/3} \sigma_0$, where σ_0 is the uni-axial yield stress, a given material parameter.
- (3) \mathbb{C}^{-1} and \mathbb{H}^{-1} are elements of $L^\infty(\Omega; \mathcal{L}(\mathbb{S}, \mathbb{S}))$, where $\mathcal{L}(\mathbb{S}, \mathbb{S})$ denotes the space of linear operators $\mathbb{S} \rightarrow \mathbb{S}$. Both $\mathbb{C}^{-1}(x)$ and $\mathbb{H}^{-1}(x)$ are assumed to be uniformly coercive. Moreover, we assume that \mathbb{C}^{-1} and \mathbb{H}^{-1} are symmetric, i.e., $\boldsymbol{\tau} : \mathbb{C}^{-1}(x) \boldsymbol{\sigma} = \boldsymbol{\sigma} : \mathbb{C}^{-1}(x) \boldsymbol{\tau}$ and a similar relation for \mathbb{H}^{-1} holds for all $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathbb{S}$. Using index notation, the symmetry assumptions can be expressed as

$$(\mathbb{C}^{-1})_{ijkl} = (\mathbb{C}^{-1})_{jikl} = (\mathbb{C}^{-1})_{klij},$$

and similarly for \mathbb{H}^{-1} .

- (4) The objective $J : V \times L^2(\Gamma_N; \mathbb{R}^d) \rightarrow \mathbb{R}$ is Fréchet differentiable. Moreover, $U_{\text{ad}} \subset L^2(\Gamma_N; \mathbb{R}^d)$ is a nonempty, closed, and convex set.

Assumption (1) implies that Korn's inequality holds on Ω , i.e.,

$$\|\mathbf{u}\|_{H^1(\Omega; \mathbb{R}^d)}^2 \leq c_K (\|\mathbf{u}\|_{L^2(\Gamma_D; \mathbb{R}^d)}^2 + \|\boldsymbol{\varepsilon}(\mathbf{u})\|_S^2) \quad (1.5)$$

for all $\mathbf{u} \in H^1(\Omega; \mathbb{R}^d)$. Note that (1.5) entails in particular that $\|\boldsymbol{\varepsilon}(\mathbf{u})\|_S$ is a norm on $H_D^1(\Omega; \mathbb{R}^d)$ equivalent to the $H^1(\Omega; \mathbb{R}^d)$ norm. We remark that Assumption (1) could be relaxed to allow more general domains, as long as Korn's inequality continues to hold, and the trace map $\tau_N : V \rightarrow L^2(\Gamma_N; \mathbb{R}^d)$ onto the Neumann part Γ_N of the boundary remains well defined.

Assumption (3) is satisfied, e.g., for isotropic and homogeneous materials, for which

$$\mathbb{C}^{-1} \boldsymbol{\sigma} = \frac{1}{2\mu} \boldsymbol{\sigma} - \frac{\lambda}{2\mu(2\mu + d\lambda)} \text{trace}(\boldsymbol{\sigma}) \mathbf{I}$$

with Lamé constants μ and λ , provided that $\mu > 0$ and $d\lambda + 2\mu > 0$ hold. These constants appear only here and there is no risk of confusion with the plastic multiplier λ or the Lagrange multiplier μ , which are introduced in Section 2 and Section 4, respectively. A common example for the hardening modulus is given by $\mathbb{H}^{-1} \boldsymbol{\chi} = \boldsymbol{\chi}/k_1$ with hardening constant $k_1 > 0$, see [Han and Reddy, 1999, Section 3.4]. Assumption (3) implies that $\langle A\boldsymbol{\Sigma}, \boldsymbol{\Sigma} \rangle \geq \underline{\alpha} \|\boldsymbol{\Sigma}\|_{S^2}^2$ for some $\underline{\alpha} > 0$, i.e. A is a coercive operator.

As an example for the objective, we mention

$$J(\mathbf{u}, \mathbf{g}) = \frac{\nu_{\mathbf{u}}}{2} \|\mathbf{u} - \mathbf{u}_d\|_{L^2(\Omega; \mathbb{R}^d)}^2 + \frac{\nu_{\boldsymbol{\varepsilon}}}{2} \|\boldsymbol{\varepsilon}(\mathbf{u}) - \boldsymbol{\varepsilon}_d\|_S^2 + \frac{\nu_{\mathbf{g}}}{2} \|\mathbf{g}\|_{L^2(\Gamma_N; \mathbb{R}^d)}^2,$$

where $\mathbf{u}_d \in L^2(\Omega; \mathbb{R}^d)$, $\boldsymbol{\varepsilon}_d \in S$ and $\nu_{\mathbf{u}}, \nu_{\boldsymbol{\varepsilon}}, \nu_{\mathbf{g}} \geq 0$ are given parameters. We emphasize that we could also consider more general objectives which depend in addition on the generalized stresses $\boldsymbol{\Sigma} \in S^2$. For simplicity of the presentation, we do not elaborate on this straightforward extension. As an example for U_{ad} , we mention the set of boundary stresses with modulus bounded by $\rho > 0$, i.e.,

$$U_{\text{ad}} = \{\mathbf{g} \in L^2(\Gamma_N; \mathbb{R}^d) : |\mathbf{g}(x)|_{\mathbb{R}^d} \leq \rho \text{ for almost all } x \in \Gamma_N\}.$$

2 Known Results Concerning the Forward Problem

In this section we collect some results concerning **(VI)**. Given $\ell \in V'$, the existence and uniqueness of a solution to **(VI)** is well known, see for instance [Han and Reddy, 1999, Section 8.1]. In particular, the admissible set

$$\{\boldsymbol{\Sigma} \in \mathcal{K} : B\boldsymbol{\Sigma} = \ell\}$$

is non-empty, see [Herzog and Meyer, 2011, Proposition 3.1]. As a consequence, we may introduce the control-to-state map

$$G : V' \rightarrow S^2 \times V, \quad \text{mapping } \ell \mapsto (G^\Sigma, G^u)(\ell) = (\boldsymbol{\Sigma}, \mathbf{u}).$$

The following result can be found in Herzog and Meyer [2011].

Theorem 2.1. *The solution operator $G : V' \rightarrow S^2 \times V$ is Lipschitz continuous, i.e.*

$$\|G(\ell_1) - G(\ell_2)\|_{S^2 \times V} = \|\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2\|_{S^2} + \|\mathbf{u}_1 - \mathbf{u}_2\|_V \leq L \|\ell_1 - \ell_2\|_{V'}$$

holds with a Lipschitz constant $L > 0$.

In our subsequent analysis we will frequently make use of an equivalent formulation of **(VI)** which involves a Lagrange multiplier for the yield condition $\phi(\boldsymbol{\Sigma}) \leq 0$, termed the plastic multiplier. We refer to Herzog et al. [2012] and Herzog et al. [2011] for the following result.

Theorem 2.2. *Let $\ell \in V'$ be given. The pair $(\boldsymbol{\Sigma}, \mathbf{u}) \in S^2 \times V$ is the unique solution of **(VI)** if and only if there exists a plastic multiplier $\lambda \in L^2(\Omega)$ such that*

$$A\boldsymbol{\Sigma} + B^*\mathbf{u} + \lambda \mathcal{D}^*\mathcal{D}\boldsymbol{\Sigma} = \mathbf{0} \quad \text{in } S^2, \quad (2.1a)$$

$$B\boldsymbol{\Sigma} = \ell \quad \text{in } V', \quad (2.1b)$$

$$0 \leq \lambda(x) \perp \phi(\boldsymbol{\Sigma}(x)) \leq 0 \quad \text{a.e. in } \Omega \quad (2.1c)$$

holds. Moreover, λ is unique.

Note that $\lambda(x) \perp \phi(\boldsymbol{\Sigma}(x))$ is a shorthand notation for $\lambda(x)\phi(\boldsymbol{\Sigma}(x)) = 0$. Using **Theorem 2.2**, problem **(P)** can be stated equivalently as an MPCC, in which **(VI)** is replaced by **(2.1)**.

The complementarity condition **(2.1c)** gives rise to the following definition of subsets of Ω :

$$\mathcal{A}(\ell) := \{x \in \Omega : \phi(\boldsymbol{\Sigma}(x)) = 0\}, \quad (\text{active set}) \quad (2.2a)$$

$$\mathcal{A}_s(\ell) := \{x \in \Omega : \lambda(x) > 0\}, \quad (\text{strongly active set}) \quad (2.2b)$$

$$\mathcal{B}(\ell) := \{x \in \Omega : \phi(\boldsymbol{\Sigma}(x)) = \lambda(x) = 0\}, \quad (\text{biactive set}) \quad (2.2c)$$

$$\mathcal{I}(\ell) := \{x \in \Omega : \phi(\boldsymbol{\Sigma}(x)) < 0\}, \quad (\text{inactive set}) \quad (2.2d)$$

where $\boldsymbol{\Sigma}$ and λ are given by the solution of **(2.1)**. The notation for the sets in **(2.2)** is driven by the point of view that **(VI)** and equivalently **(2.1)** are the necessary and sufficient optimality conditions for the lower-level problem

$$\text{Minimize } \frac{1}{2} \langle A\boldsymbol{\Sigma}, \boldsymbol{\Sigma} \rangle \quad \text{s.t. } B\boldsymbol{\Sigma} = \ell \quad \text{and} \quad \phi(\boldsymbol{\Sigma}) \leq 0, \quad (2.3)$$

in which $\phi(\boldsymbol{\Sigma}) \leq 0$ appears as a constraint with Lagrange multiplier λ . We remark that $\mathcal{A}_s(\ell)$, $\mathcal{B}(\ell)$, and $\mathcal{I}(\ell)$ are pairwise disjoint sets. Furthermore $\mathcal{A}(\ell) = \mathcal{A}_s(\ell) \cup \mathcal{B}(\ell)$ and $\Omega = \mathcal{A}(\ell) \cup \mathcal{I}(\ell)$.

Remark 2.3. *The component-wise evaluation of (2.1a) yields*

$$\mathbb{C}^{-1} \boldsymbol{\sigma} - \boldsymbol{\varepsilon}(\mathbf{u}) + \lambda \mathcal{D}\boldsymbol{\Sigma} = \mathbf{0}, \quad (2.4a)$$

$$\mathbb{H}^{-1} \boldsymbol{\chi} + \lambda \mathcal{D}\boldsymbol{\Sigma} = \mathbf{0}. \quad (2.4b)$$

Combining both equations, we find

$$\mathbb{C}^{-1} \boldsymbol{\sigma} - \boldsymbol{\varepsilon}(\mathbf{u}) - \mathbb{H}^{-1} \boldsymbol{\chi} = \mathbf{0}. \quad (2.5)$$

3 Bouligand Stationarity

In virtue of the control-to-state map G with components $(G^\boldsymbol{\Sigma}, G^{\mathbf{u}})$, we may reduce problem **(P)** to the control variable \mathbf{g} :

$$\begin{aligned} \text{Minimize } & j(\mathbf{g}) := J(G^{\mathbf{u}}(-\tau_N^* \mathbf{g}), \mathbf{g}) \\ \text{s.t. } & \mathbf{g} \in U_{\text{ad}}. \end{aligned}$$

The term $-\tau_N^* \mathbf{g}$ denotes the load ℓ induced by the boundary stresses \mathbf{g} , i.e.,

$$\tau_N^* : L^2(\Gamma_N; \mathbb{R}^d) \rightarrow V', \quad \langle \tau_N^* \mathbf{g}, \mathbf{v} \rangle := \int_{\Gamma_N} \mathbf{g} \cdot \mathbf{v} \, ds, \quad \mathbf{v} \in V,$$

compare **(P)**. Note that the bounded linear and compact operator τ_N^* is the adjoint of the boundary trace map $\tau_N : V \rightarrow L^2(\Gamma_N; \mathbb{R}^d)$ onto the Neumann part Γ_N of the boundary.

The aim of this section is to prove that j is directionally differentiable so that local minimizers $\bar{\mathbf{g}}$ necessarily satisfy

$$\delta j(\bar{\mathbf{g}}; \mathbf{g} - \bar{\mathbf{g}}) \geq 0 \quad \text{for all } \mathbf{g} \in U_{\text{ad}}, \quad (3.1)$$

see **Theorem 3.10**. In fact, (3.1) can be extended to tangential directions of U_{ad} , see **Corollary 3.12**. Note that this optimality condition involves only primal variables. We will show later on (see **Remark 3.16**) that (3.1) is equivalent to the notion of Bouligand, or B-stationarity for MPCCs, as defined in **Scheel and Scholtes [2000]** for finite dimensional problems.

We derive (3.1) by the *implicit programming approach* (see, e.g., **[Luo et al., 1996, Section 4.2]** for MPECs), in which the lower-level problem is replaced by its solution map. Alternative approaches to derive B-stationarity conditions involve (i) concepts based on the evaluation of the *tangent cone* of the set of feasible $(\mathbf{g}, \boldsymbol{\Sigma}, \mathbf{u})$, see **[Luo et al., 1996, Section 3.3]**, or (ii) based on the directional derivative of a nonsmooth *exact penalty problem*. We refer to **Kořvara and Outrata [2004]** and the references therein for an overview.

In order to establish (3.1), the main step is to prove the weak directional differentiability of G . This is achieved in the following subsection and it is a result which could also be of independent interest. The variational inequality (3.1) then follows easily by a chain rule argument, see **Section 3.2**. In **Section 3.3** we confirm that (3.1) is indeed equivalent to the concept of B-stationarity. The purpose of **Section 3.4** is to point out parallels of our implicit programming approach with the tangent cone technique in **[Luo et al., 1996, Section 3.3]**.

Remark 3.1.

- (1) *It can be easily shown that **(P)** possesses at least one global optimal solution, see **[Herzog and Meyer, 2011, Proposition 3.6]**. Notice however that one cannot expect the solution to be unique due to the nonlinearity of G .*
- (2) *To keep the presentation concise, we restrict the discussion to the control of boundary stresses only. There would be no difficulty in including additional volume forces as control variables as in **Herzog and Meyer [2011]**.*

3.1. Weak Directional Differentiability of the Control-to-State Map. In this subsection, we will show that $G : V' \rightarrow S^2 \times V$ is directionally differentiable in a weak sense in all directions

$$\frac{G(\ell + t\delta\ell) - G(\ell)}{t} \rightharpoonup \delta_w G(\ell; \delta\ell) \quad \text{in } S^2 \times V \quad \text{as } t \searrow 0.$$

The weak limit $\delta_w G(\ell; \delta\ell) = (\boldsymbol{\Sigma}', \mathbf{u}')$ is given by the unique solution $(\boldsymbol{\Sigma}', \mathbf{u}') \in \mathcal{S}_\ell \times V$ of the following variational inequality:

$$\langle A\boldsymbol{\Sigma}', \mathbf{T} - \boldsymbol{\Sigma}' \rangle + \langle B^* \mathbf{u}', \mathbf{T} - \boldsymbol{\Sigma}' \rangle + (\lambda, \mathcal{D}\boldsymbol{\Sigma}' : \mathcal{D}(\mathbf{T} - \boldsymbol{\Sigma}'))_{\Omega} \geq 0 \quad \text{for all } \mathbf{T} \in \mathcal{S}_\ell, \quad (3.2a)$$

$$B\boldsymbol{\Sigma}' = \delta\ell, \quad (3.2b)$$

where the convex cone \mathcal{S}_ℓ is defined by

$$\mathcal{S}_\ell := \{ \mathbf{T} \in S^2 : \sqrt{\lambda} \mathcal{D}\mathbf{T} \in S, \quad \mathcal{D}\boldsymbol{\Sigma}(x) : \mathcal{D}\mathbf{T}(x) \leq 0 \text{ a.e. in } \mathcal{B}(\ell), \\ \mathcal{D}\boldsymbol{\Sigma}(x) : \mathcal{D}\mathbf{T}(x) = 0 \text{ a.e. in } \mathcal{A}_s(\ell) \}. \quad (3.3)$$

Here, $(\boldsymbol{\Sigma}, \mathbf{u}, \lambda) \in S^2 \times V \times L^2(\Omega)$ is the unique solution of (2.1), i.e. $(\boldsymbol{\Sigma}, \mathbf{u}) = G(\ell)$ and $\lambda \in L^2(\Omega)$ is the associated plastic multiplier. The structure of \mathcal{S}_ℓ is typical for directional derivatives of solutions to optimization problems such as (2.3). Concerning the linearization of inequality constraints, one needs to distinguish three cases. Inactive constraints impose no restrictions on the derivative, while (strongly active) active constraints have to remain (active) feasible to first order, see [Jittorntrum \[1984\]](#). The additional condition $\sqrt{\lambda} \mathcal{D}\mathbf{T} \in S$ serves to make the third term in (3.2a) a priori well defined. It will be shown below that this condition is indeed satisfied for the weak directional derivative of G and therefore it does not introduce an artificial restriction. Note that in case $\lambda \notin L^\infty(\Omega)$, the set \mathcal{S}_ℓ is not closed in S^2 . Nevertheless, as shown in [Theorem 3.2](#), there exists a unique solution of (3.2).

We also refer to [Mignot \[1976\]](#), where conical differentiability of the solution operator of the elliptic obstacle problem is proven. Here the structure of the respective counterparts to (3.2) and (3.3) is simpler due to the linearity of the inequality constraint. Concerning the parabolic obstacle problem, the solution operator is proven to be conically differentiable in [Jarušek et al. \[2003\]](#). To our best knowledge, these are the only references concerning the differentiability of solution operators of VIs. In particular, variational inequalities which are not of obstacle type, such as (VI), have not been discussed so far in this respect. Finally, we point out that an equivalent characterization of the weak derivative involving a derivative of the plastic multiplier is given in (3.32).

The main result of this subsection is the following.

Theorem 3.2. *The control-to-state map $G : V' \rightarrow S^2 \times V$ is weakly directionally differentiable at every $\ell \in V'$ in all directions $\delta\ell \in V'$. The weak directional derivative is given by the unique solution of (3.2). Moreover, for fixed ℓ , the weak directional derivative depends (globally) Lipschitz continuously on the direction $\delta\ell$.*

Before we are in the position to prove [Theorem 3.2](#), we need several auxiliary results. Let us consider a fixed but arbitrary sequence of positive real numbers $\{t_n\}$ tending to zero as $n \rightarrow \infty$. We introduce a perturbed problem associated with t_n by

$$\langle A\boldsymbol{\Sigma}_n, \mathbf{T} - \boldsymbol{\Sigma}_n \rangle + \langle B^* \mathbf{u}_n, \mathbf{T} - \boldsymbol{\Sigma}_n \rangle \geq 0 \quad \text{for all } \mathbf{T} \in \mathcal{K}, \quad (3.4a)$$

$$B\boldsymbol{\Sigma}_n = \ell + t_n \delta\ell. \quad (3.4b)$$

with $\Sigma_n \in \mathcal{K}$. Clearly, (3.4) admits a unique solution and, in view of [Theorem 2.1](#), we have

$$\left\| \frac{\Sigma_n - \Sigma}{t_n} \right\|_{S^2} + \left\| \frac{\mathbf{u}_n - \mathbf{u}}{t_n} \right\|_V \leq L \|\delta\ell\|_{V'} < \infty. \quad (3.5)$$

Therefore, the sequence $\left\{ \left(\frac{\Sigma_n - \Sigma}{t_n}, \frac{\mathbf{u}_n - \mathbf{u}}{t_n} \right) \right\}$ is bounded in $S^2 \times V$ and there exists a weakly convergent subsequence, which is denoted by the same symbol to simplify the notation. At the end of the proof of [Theorem 3.2](#), we shall see that *every* weakly convergent subsequence of difference quotients has the same limit so that a well known argument yields the weak convergence of the whole sequence. This justifies the simplification of notation. The weak limit is denoted by $(\tilde{\Sigma}, \tilde{\mathbf{u}})$, i.e.

$$\left(\frac{\Sigma_n - \Sigma}{t_n}, \frac{\mathbf{u}_n - \mathbf{u}}{t_n} \right) \rightharpoonup (\tilde{\Sigma}, \tilde{\mathbf{u}}) \quad \text{in } S^2 \times V \text{ for } n \rightarrow \infty. \quad (3.6)$$

Our goal is to show that $(\tilde{\Sigma}, \tilde{\mathbf{u}})$ satisfies (3.2). To this end, we proceed as follows.

- (1) We verify that $\tilde{\Sigma}$ satisfies the sign conditions in (3.3) ([Proposition 3.3](#)).
- (2) We introduce the plastic multipliers for the perturbed problems, see (3.12).
- (3) We show the weak directional differentiability of the plastic multiplier ([Proposition 3.4](#)).
- (4) We establish $\tilde{\Sigma} \in \mathcal{S}_\ell$ as well as a complementarity relation for $\tilde{\Sigma}$ and the weak directional derivative of the plastic multiplier ([Proposition 3.8](#)).
- (5) In the proof of [Theorem 3.2](#), this complementarity relation is used to show that $(\tilde{\Sigma}, \tilde{\mathbf{u}})$ satisfies (3.2). Moreover, we prove the uniqueness of the solution of (3.2).

Proposition 3.3. *The weak limit $(\tilde{\Sigma}, \tilde{\mathbf{u}})$ satisfies $\mathcal{D}\Sigma : \mathcal{D}\tilde{\Sigma} \leq 0$ a.e. in $\mathcal{B}(\ell)$ and $\mathcal{D}\Sigma : \mathcal{D}\tilde{\Sigma} = 0$ a.e. in $\mathcal{A}_s(\ell)$.*

Proof. Due to $\Sigma, \Sigma_n \in \mathcal{K}$ for every $n \in \mathbb{N}$, one finds

$$\mathcal{D}\Sigma(x) : (\mathcal{D}\Sigma_n(x) - \mathcal{D}\Sigma(x)) \leq |\mathcal{D}\Sigma(x)| |\mathcal{D}\Sigma_n(x)| - \tilde{\sigma}_0^2 \leq 0$$

on the active set $\mathcal{A}(\ell)$ and thus

$$\mathcal{D}\Sigma(x) : \frac{\mathcal{D}\Sigma_n(x) - \mathcal{D}\Sigma(x)}{t_n} \leq 0 \quad \text{a.e. in } \mathcal{A}(\ell) \quad (3.7)$$

for all $n \in \mathbb{N}$. We note that the set $\{\mathbf{T} \in S^2 : \mathcal{D}\Sigma(x) : \mathcal{D}\mathbf{T}(x) \leq 0 \text{ a.e. in } \mathcal{A}(\ell)\}$ is convex and closed, thus weakly closed. This implies

$$\mathcal{D}\Sigma : \mathcal{D}\tilde{\Sigma} \leq 0 \quad \text{a.e. in } \mathcal{A}(\ell). \quad (3.8)$$

Since $\lambda(x) = 0$ a.e. in $\mathcal{I}(\ell)$ and $\lambda(x) \geq 0$ a.e. in $\mathcal{A}(\ell)$, (3.7) yields

$$\left(\lambda, \mathcal{D}\Sigma : \frac{\mathcal{D}\Sigma_n - \mathcal{D}\Sigma}{t_n} \right)_\Omega \leq 0 \quad \text{for all } n \in \mathbb{N}. \quad (3.9)$$

Now we test (3.4a) with $\mathbf{T} = \Sigma$ which is clearly feasible since $\Sigma \in \mathcal{K}$. Then (3.9) together with (2.1a) implies

$$\begin{aligned} 0 &\leq - \left(\lambda, \mathcal{D}\Sigma : \frac{\mathcal{D}\Sigma_n - \mathcal{D}\Sigma}{t_n} \right)_\Omega \\ &\leq t_n \left[- \left\langle A \left(\frac{\Sigma_n - \Sigma}{t_n} \right), \left(\frac{\Sigma_n - \Sigma}{t_n} \right) \right\rangle - \left\langle B^* \left(\frac{\mathbf{u}_n - \mathbf{u}}{t_n} \right), \left(\frac{\Sigma_n - \Sigma}{t_n} \right) \right\rangle \right] \\ &\leq c t_n \left\| \frac{\Sigma_n - \Sigma}{t_n} \right\|_{S^2} \left\| \frac{\mathbf{u}_n - \mathbf{u}}{t_n} \right\|_V \leq c t_n \|\delta\ell\|_{V'}^2 \rightarrow 0 \quad \text{as } n \rightarrow \infty \end{aligned} \quad (3.10)$$

because of (3.5). Furthermore, (3.6) implies the weak convergence of $(\mathcal{D}\Sigma_n - \mathcal{D}\Sigma)/t_n$ in S to $\mathcal{D}\tilde{\Sigma}$. Hence, we obtain

$$\left(\lambda, \mathcal{D}\Sigma : \frac{\mathcal{D}\Sigma_n - \mathcal{D}\Sigma}{t_n}\right)_\Omega \rightarrow (\lambda, \mathcal{D}\Sigma : \mathcal{D}\tilde{\Sigma})_\Omega$$

and therefore, (3.10) gives

$$(\lambda, \mathcal{D}\Sigma : \mathcal{D}\tilde{\Sigma})_\Omega = 0. \quad (3.11)$$

Hence, since $\lambda > 0$ on $\mathcal{A}_s(\ell)$ holds, (3.8) implies $\mathcal{D}\Sigma : \mathcal{D}\tilde{\Sigma} = 0$ a.e. in $\mathcal{A}_s(\ell)$. \square

In order to prove the existence of the weak directional derivative of the plastic multiplier, we reformulate the perturbed problem (3.4) by introducing the plastic multiplier λ_n , see **Theorem 2.2**:

$$A\Sigma_n + B^* \mathbf{u}_n + \lambda_n \mathcal{D}^* \mathcal{D}\Sigma_n = \mathbf{0} \quad (3.12a)$$

$$B\Sigma_n = \ell + t_n \delta \ell \quad (3.12b)$$

$$0 \leq \lambda_n(x) \perp \phi(\Sigma_n(x)) \leq 0 \quad \text{a.e. in } \Omega. \quad (3.12c)$$

Arguing as in **Remark 2.3** yields

$$\begin{aligned} \lambda \mathcal{D}\Sigma &= -\mathbb{H}^{-1} \chi, \\ \lambda_n \mathcal{D}\Sigma_n &= -\mathbb{H}^{-1} \chi_n. \end{aligned} \quad (3.13)$$

These relations define the starting point for the proof of convergence of the plastic multipliers in the following proposition.

Proposition 3.4. *We have convergence of the plastic multipliers $\lambda_n \rightarrow \lambda$ in $L^2(\Omega)$. Their difference quotient $(\lambda_n - \lambda)/t_n \rightharpoonup \tilde{\lambda}$ converges weakly in $L^1(\Omega)$. Moreover, $\tilde{\lambda} \in L^2(\Omega)$ and $\tilde{\lambda}$ is uniquely determined by the weak limit $\tilde{\Sigma}$ of the stresses. In particular, $\tilde{\lambda} = 0$ and $\tilde{\chi} = \mathbf{0}$ hold on $\mathcal{I}(\ell)$.*

Proof. We address $\lambda_n \rightarrow \lambda$ first. Since $\lambda = 0$ on $\mathcal{I}(\ell)$ and $\lambda_n = 0$ on $\mathcal{I}(\ell + t_n \delta \ell)$, we get by (3.13) the following characterization of λ and λ_n ,

$$\begin{aligned} \tilde{\sigma}_0^2 \lambda &= \lambda \mathcal{D}\Sigma : \mathcal{D}\Sigma = -\mathbb{H}^{-1} \chi : \mathcal{D}\Sigma \quad \text{a.e. in } \Omega, \\ \tilde{\sigma}_0^2 \lambda_n &= \lambda_n \mathcal{D}\Sigma_n : \mathcal{D}\Sigma_n = -\mathbb{H}^{-1} \chi_n : \mathcal{D}\Sigma_n \quad \text{a.e. in } \Omega. \end{aligned}$$

Taking the difference yields

$$\begin{aligned} \lambda_n - \lambda &= \frac{1}{\tilde{\sigma}_0^2} \left(-\mathbb{H}^{-1} \chi_n : \mathcal{D}\Sigma_n + \mathbb{H}^{-1} \chi : \mathcal{D}\Sigma \right) \\ &= \frac{1}{\tilde{\sigma}_0^2} \left(\mathbb{H}^{-1} \chi : (\mathcal{D}\Sigma - \mathcal{D}\Sigma_n) + \mathbb{H}^{-1} (\chi - \chi_n) : \mathcal{D}\Sigma_n \right). \end{aligned} \quad (3.14)$$

Clearly, the right-hand side is bounded in $L^2(\Omega)$ since $\mathcal{D}\Sigma, \mathcal{D}\Sigma_n \in L^\infty(\Omega; \mathbb{S})$. Thus there exists a subsequence λ_{n_k} of λ_n which converges weakly in $L^2(\Omega)$ to some λ° . We will show $\lambda^\circ = \lambda$, therefore the weak limit is independent of the chosen subsequence, and hence the whole sequence $\{\lambda_n\}$ converges weakly in $L^2(\Omega)$, i.e. $\lambda_n \rightharpoonup \lambda$. Using again (3.13) and the convergence of $\chi_n \rightarrow \chi$ in S gives

$$\lambda_n \mathcal{D}\Sigma_n = -\mathbb{H}^{-1} \chi_n \rightarrow -\mathbb{H}^{-1} \chi = \lambda \mathcal{D}\Sigma \quad \text{in } S. \quad (3.15)$$

On the other hand, the weak convergence $\lambda_{n_k} \rightharpoonup \lambda^\circ$ in $L^2(\Omega)$ implies

$$\lambda_{n_k} \mathcal{D}\Sigma_{n_k} \rightharpoonup \lambda^\circ \mathcal{D}\Sigma \quad \text{in } L^1(\Omega; \mathbb{S}).$$

Therefore $\lambda = \lambda^\circ$ holds on $\mathcal{A}(\ell)$. Due to the complementarity condition (2.1c), $\lambda = 0$ holds on $\mathcal{I}(\ell)$ and this implies $\|\lambda\|_{L^2(\Omega)} \leq \|\lambda^\circ\|_{L^2(\Omega)}$.

The complementarity relations (2.1c) and (3.12c) together with (3.15) imply

$$\|\lambda\|_{L^2(\Omega)} = \frac{1}{\tilde{\sigma}_0} \|\lambda \mathcal{D}\Sigma\|_S = \lim_{n \rightarrow \infty} \frac{1}{\tilde{\sigma}_0} \|\lambda_n \mathcal{D}\Sigma_n\|_S = \lim_{n \rightarrow \infty} \|\lambda_n\|_{L^2(\Omega)}.$$

Since the norm is weakly lower semicontinuous this implies

$$\|\lambda\|_{L^2(\Omega)} = \lim_{n \rightarrow \infty} \|\lambda_n\|_{L^2(\Omega)} = \liminf_{k \rightarrow \infty} \|\lambda_{n_k}\|_{L^2(\Omega)} \geq \|\lambda^\circ\|_{L^2(\Omega)}.$$

We conclude $\|\lambda^\circ\|_{L^2(\Omega)} = \|\lambda\|_{L^2(\Omega)}$. Due to $\lambda^\circ = \lambda$ on $\mathcal{A}(\ell)$ and $\lambda = 0$ on $\mathcal{I}(\ell)$, $\lambda^\circ = \lambda$ is satisfied. This shows the independence of the weak limit from the chosen subsequence, thus the whole sequence $\{\lambda_n\}$ converges weakly to λ . Additionally, in view of the convergence of norms, $\lambda_n \rightarrow \lambda$ strongly in $L^2(\Omega)$.

We proceed by showing the weak convergence of the difference quotients for λ . Using (3.6) and $\Sigma_n \rightarrow \Sigma$ in S^2 , we obtain from (3.14)

$$\begin{aligned} \frac{\lambda_n - \lambda}{t_n} &= \frac{1}{\tilde{\sigma}_0^2} (\mathbb{H}^{-1}\chi : \frac{\mathcal{D}\Sigma - \mathcal{D}\Sigma_n}{t_n} + \mathbb{H}^{-1}\chi - \chi_n : \mathcal{D}\Sigma_n) \\ &\rightharpoonup \tilde{\lambda} := \frac{-1}{\tilde{\sigma}_0^2} (\mathbb{H}^{-1}\chi : \mathcal{D}\tilde{\Sigma} + \mathbb{H}^{-1}\tilde{\chi} : \mathcal{D}\Sigma) \quad \text{in } L^1(\Omega). \end{aligned}$$

Finally, by (3.13) and Proposition 3.3,

$$\tilde{\lambda} = \frac{-1}{\sigma_0^2} \mathbb{H}^{-1}\tilde{\chi} : \mathcal{D}\Sigma \in L^2(\Omega).$$

Next we show that $\tilde{\lambda} = 0$ holds on $\mathcal{I}(\ell)$. For convenience, let us abbreviate

$$\mathcal{A}_{s,k} := \mathcal{A}_s(\ell + t_{n_k} \delta \ell).$$

We consider the sets $\mathcal{I}(\ell) \cap \mathcal{A}_{s,k}$, on which $|\mathcal{D}\Sigma| < \tilde{\sigma}_0$ and $|\mathcal{D}\Sigma_{n_k}| = \tilde{\sigma}_0$ hold. From $\Sigma_{n_k} \rightarrow \Sigma$ in S^2 we infer $|\mathcal{D}\Sigma_{n_k}| \rightarrow |\mathcal{D}\Sigma|$ in $L^2(\Omega)$. Lemma A.2 with $M = \mathcal{I}(\ell)$, $f = \tilde{\sigma}_0 - |\mathcal{D}\Sigma|$, $f_k = \tilde{\sigma}_0 - |\mathcal{D}\Sigma_{n_k}|$ yields $|\mathcal{I}(\ell) \cap \mathcal{A}_{s,k}| \rightarrow 0$ as $k \rightarrow \infty$. We may now invoke Lemma A.3 with $f_{n_k} := (\lambda_{n_k} - \lambda)/t_{n_k} \geq 0$ and $f := \tilde{\lambda}$, both restricted to $\mathcal{I}(\ell)$, to conclude that

$$\tilde{\lambda}(x) = 0 \quad \text{a.e. in } \mathcal{I}(\ell). \quad (3.16)$$

It remains to show $\tilde{\chi} = \mathbf{0}$ on $\mathcal{I}(\ell)$. Using the above convergence results, we obtain

$$\begin{aligned} \frac{\lambda_{n_k} \mathcal{D}\Sigma_{n_k} - \lambda \mathcal{D}\Sigma}{t_{n_k}} &= \frac{\lambda_{n_k} - \lambda}{t_{n_k}} \mathcal{D}\Sigma + \lambda_{n_k} \frac{\mathcal{D}\Sigma_{n_k} - \mathcal{D}\Sigma}{t_{n_k}} \\ &\rightharpoonup \tilde{\lambda} \mathcal{D}\Sigma + \lambda \mathcal{D}\tilde{\Sigma} \quad \text{in } L^1(\Omega; \mathbb{S}). \end{aligned} \quad (3.17)$$

Due to (3.13) we find

$$\frac{\lambda_n \mathcal{D}\Sigma_n - \lambda \mathcal{D}\Sigma}{t_n} = \frac{\mathbb{H}^{-1}\chi - \mathbb{H}^{-1}\chi_n}{t_n} \rightharpoonup -\mathbb{H}^{-1}\tilde{\chi} \quad \text{in } S.$$

Consequently,

$$-\mathbb{H}^{-1}\tilde{\chi} = \tilde{\lambda} \mathcal{D}\Sigma + \lambda \mathcal{D}\tilde{\Sigma}, \quad (3.18)$$

which implies in particular that $\tilde{\chi} = \mathbf{0}$ on $\mathcal{I}(\ell)$. \square

We are now in the position to state an equation relating the terms for the directional derivative problem.

Lemma 3.5. *The weak limit $(\tilde{\Sigma}, \tilde{\mathbf{u}}, \tilde{\lambda})$ satisfies*

$$A\tilde{\Sigma} + B^*\tilde{\mathbf{u}} + \lambda \mathcal{D}^* \mathcal{D}\tilde{\Sigma} + \tilde{\lambda} \mathcal{D}^* \mathcal{D}\Sigma = \mathbf{0} \quad \text{in } S^2. \quad (3.19)$$

Proof. Let us start with an arbitrary $\mathbf{T} \in L^\infty(\Omega; \mathbb{S}^2)$. In view of (2.1a) and (3.12a) we have

$$\left\langle A \frac{\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}}{t_n}, \mathbf{T} \right\rangle + \left\langle B^* \frac{\mathbf{u}_n - \mathbf{u}}{t_n}, \mathbf{T} \right\rangle + \left(\frac{\lambda_n \mathcal{D}\boldsymbol{\Sigma}_n - \lambda \mathcal{D}\boldsymbol{\Sigma}}{t_n}, \mathcal{D}\mathbf{T} \right)_\Omega = 0.$$

Due to (3.17) we can pass to the limit to obtain

$$\langle A\tilde{\boldsymbol{\Sigma}}, \mathbf{T} \rangle + \langle B^*\tilde{\mathbf{u}}, \mathbf{T} \rangle + (\lambda, \mathcal{D}\tilde{\boldsymbol{\Sigma}} : \mathcal{D}\mathbf{T})_\Omega + (\tilde{\lambda}, \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\mathbf{T})_\Omega = 0 \quad \text{for all } \mathbf{T} \in L^\infty(\Omega; \mathbb{S}^2). \quad (3.20)$$

By (3.18), we know that $\lambda \mathcal{D}\tilde{\boldsymbol{\Sigma}}$ belongs to S^2 and thus the left hand side of (3.20) is a continuous mapping from S^2 to \mathbb{R} with respect to \mathbf{T} . Therefore the density of $L^\infty(\Omega; \mathbb{S}^2)$ in S^2 implies that (3.20) holds for all $\mathbf{T} \in S^2$. \square

In order to pass from (3.19) to the variational inequality (3.2a), we need to verify a certain complementarity condition, see Proposition 3.8 below. To this end, we recall a particular case of [Herzog et al., 2012, Proposition 3.15]:

Lemma 3.6. *Let $\varphi \in C_0^\infty(\Omega)$ with $\varphi \geq 0$ be arbitrary and suppose that*

$$\boldsymbol{\Phi}_k \rightharpoonup \boldsymbol{\Phi} \quad \text{in } S^2, \quad \mathbf{z}_k \rightharpoonup \mathbf{z} \quad \text{in } V, \quad B\boldsymbol{\Phi}_k \equiv h \quad \text{in } V'.$$

Assume further that

$$\langle A\boldsymbol{\Phi}_k, \varphi \boldsymbol{\Phi}_k \rangle + \langle B^*\mathbf{z}_k, \varphi \boldsymbol{\Phi}_k \rangle \leq 0 \quad \text{for all } k \in \mathbb{N}.$$

Then $\langle A\boldsymbol{\Phi}, \varphi \boldsymbol{\Phi} \rangle + \langle B^*\mathbf{z}, \varphi \boldsymbol{\Phi} \rangle \leq 0$.

Lemma 3.7. *We have*

$$(\tilde{\lambda} \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\mathbf{T})(x) \begin{cases} \leq 0 & \text{a.e. in } \mathcal{B}(\ell) \\ = 0 & \text{a.e. in } \Omega \setminus \mathcal{B}(\ell). \end{cases} \quad (3.21)$$

for all $\mathbf{T} \in \mathcal{S}_\ell$.

Proof. Recall that $\tilde{\lambda} = 0$ a.e. in $\mathcal{I}(\ell)$. Moreover we have $\tilde{\lambda} \geq 0$ a.e. in $\Omega \setminus \mathcal{A}_s(\ell) = \{x \in \Omega : \lambda(x) = 0\}$, which follows from

$$0 \leq \frac{\lambda_n}{t_n} = \frac{\lambda_n - \lambda}{t_n} \rightharpoonup \tilde{\lambda} \quad \text{in } L^1(\Omega \setminus \mathcal{A}_s(\ell)) \quad (3.22)$$

and the weak closedness in L^1 of the set of nonnegative functions. Now the assertion follows from $\mathbf{T} \in \mathcal{S}_\ell$, i.e., $\mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\mathbf{T} = 0$ on $\mathcal{A}_s(\ell)$ and $\mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\mathbf{T} \leq 0$ on $\mathcal{B}(\ell)$. \square

Proposition 3.8. *The weak limit $\tilde{\boldsymbol{\Sigma}}$ belongs to \mathcal{S}_ℓ . Moreover, the relation $\tilde{\lambda} \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\tilde{\boldsymbol{\Sigma}} = 0$ holds a.e. in Ω .*

Proof. We have already verified in Proposition 3.3 that $\mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\tilde{\boldsymbol{\Sigma}} \leq 0$ in $\mathcal{B}(\ell)$ and $\mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\tilde{\boldsymbol{\Sigma}} = 0$ in $\mathcal{A}_s(\ell)$. To prove $\tilde{\boldsymbol{\Sigma}} \in \mathcal{S}_\ell$, it remains to show $\sqrt{\tilde{\lambda}} \mathcal{D}\tilde{\boldsymbol{\Sigma}} \in S$. This follows by testing (3.18) with $\mathcal{D}\tilde{\boldsymbol{\Sigma}}$ and using $\mathcal{D}\boldsymbol{\Sigma} \in L^\infty(\Omega; \mathbb{S})$.

To show the slackness condition, let $\varphi \in C_0^\infty(\Omega)$ with $0 \leq \varphi \leq 1$ be given. Since $\boldsymbol{\Sigma}, \boldsymbol{\Sigma}_n \in \mathcal{K}$ for all $n \in \mathbb{N}$, we may test (VI) with $\mathbf{T} = \varphi \boldsymbol{\Sigma}_n + (1 - \varphi) \boldsymbol{\Sigma} \in \mathcal{K}$ and (3.4a) with $\mathbf{T} = \varphi \boldsymbol{\Sigma} + (1 - \varphi) \boldsymbol{\Sigma}_n \in \mathcal{K}$. Adding the arising inequalities implies

$$\left\langle A \frac{\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}}{t_n}, \varphi \frac{\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}}{t_n} \right\rangle + \left\langle B^* \frac{\mathbf{u}_n - \mathbf{u}}{t_n}, \varphi \frac{\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}}{t_n} \right\rangle \leq 0.$$

Now we apply Lemma 3.6 with

$$\boldsymbol{\Phi}_n := (\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma})/t_n \rightharpoonup \tilde{\boldsymbol{\Sigma}} \quad \text{in } S^2, \quad \mathbf{z}_n := (\mathbf{u}_n - \mathbf{u})/t_n \rightharpoonup \tilde{\mathbf{u}} \quad \text{in } V.$$

This yields

$$\langle A\tilde{\boldsymbol{\Sigma}}, \varphi \tilde{\boldsymbol{\Sigma}} \rangle + \langle B^*\tilde{\mathbf{u}}, \varphi \tilde{\boldsymbol{\Sigma}} \rangle \leq 0 \quad \text{for all } \varphi \in C_0^\infty(\Omega) \text{ with values in } [0, 1].$$

A simple scaling argument shows the same inequality for all nonnegative $\varphi \in C_0^\infty(\Omega)$, and thus

$$\tilde{\Sigma}(x) : (A\tilde{\Sigma})(x) + \tilde{\Sigma}(x) : (B^*\tilde{\mathbf{u}})(x) \leq 0 \quad \text{a.e. in } \Omega.$$

Since $\lambda = 0$ on $\mathcal{B}(\ell)$, (3.19) implies

$$\tilde{\Sigma}(x) : (A\tilde{\Sigma})(x) + \tilde{\Sigma}(x) : (B^*\tilde{\mathbf{u}})(x) + \tilde{\lambda}(x) (\mathcal{D}\Sigma)(x) : (\mathcal{D}\tilde{\Sigma})(x) = 0 \quad \text{a.e. on } \mathcal{B}(\ell).$$

Therefore $\tilde{\lambda} \mathcal{D}\Sigma : \mathcal{D}\tilde{\Sigma} \geq 0$ on the biactive set $\mathcal{B}(\ell)$ and due to $\tilde{\Sigma} \in \mathcal{S}_\ell$ and (3.21) we have

$$\tilde{\lambda} \mathcal{D}\Sigma : \mathcal{D}\tilde{\Sigma} = 0 \quad \text{a.e. in } \Omega. \quad (3.23)$$

□

Finally we are in the position to prove **Theorem 3.2**.

Proof of Theorem 3.2. Let $\ell, \delta\ell \in V'$ be given. Let $\mathbf{T} \in \mathcal{S}_\ell$ be arbitrary. We test (3.19) with $\mathbf{T} - \tilde{\Sigma}$ which leads to

$$\begin{aligned} \langle A\tilde{\Sigma}, \mathbf{T} - \tilde{\Sigma} \rangle + \langle B^*\tilde{\mathbf{u}}, \mathbf{T} - \tilde{\Sigma} \rangle + (\lambda, \mathcal{D}\tilde{\Sigma} : \mathcal{D}(\mathbf{T} - \tilde{\Sigma}))_\Omega \\ = (\tilde{\lambda}, \mathcal{D}\Sigma : \mathcal{D}\tilde{\Sigma})_\Omega - (\tilde{\lambda}, \mathcal{D}\Sigma : \mathcal{D}\mathbf{T})_\Omega. \end{aligned}$$

The first addend on the right-hand side vanishes due to **Proposition 3.8**. In view of **Lemma 3.7**, we conclude

$$\langle A\tilde{\Sigma}, \mathbf{T} - \tilde{\Sigma} \rangle + \langle B^*\tilde{\mathbf{u}}, \mathbf{T} - \tilde{\Sigma} \rangle + (\lambda, \mathcal{D}\tilde{\Sigma} : \mathcal{D}(\mathbf{T} - \tilde{\Sigma}))_\Omega \geq 0,$$

which is the claimed variational inequality for the derivative (3.2a).

The equation in (VI) and (3.4b) imply

$$B \frac{\Sigma_n - \Sigma}{t_n} = \delta\ell \quad \text{in } V'$$

and the weak convergence immediately gives $B\tilde{\Sigma} = \delta\ell$.

We have shown that the weak limit of the difference quotients $(\tilde{\Sigma}, \tilde{\mathbf{u}})$ satisfies (3.2).

It remains to verify that (3.2) does not admit other solutions. Suppose on the contrary that (Σ', \mathbf{u}') and (Σ'', \mathbf{u}'') are two solutions, then a simple testing argument using $\lambda \geq 0$ and the coercivity of A , thanks to **Assumption 1.1** (3), shows that Σ' and Σ'' must coincide. To verify the uniqueness of the displacement field, we define $\boldsymbol{\tau} = \boldsymbol{\varepsilon}(\mathbf{u}' - \mathbf{u}'')$ and $\mathbf{T}' = (\boldsymbol{\tau}, -\boldsymbol{\tau}) + \Sigma'$ and $\mathbf{T}'' = \Sigma''$. These are feasible as test functions in (3.2a) due to the structure of \mathcal{K} . This implies

$$0 \leq \langle B^*(\mathbf{u}' - \mathbf{u}''), \mathbf{T}' - \mathbf{T}'' \rangle = - \int_\Omega \boldsymbol{\varepsilon}(\mathbf{u}' - \mathbf{u}'') : \boldsymbol{\varepsilon}(\mathbf{u}' - \mathbf{u}'') \, dx \leq 0.$$

From here, Korn's inequality (1.5) shows $\mathbf{u}' = \mathbf{u}''$.

Thus the weak limit is unique and a well known argument implies the convergence of the whole sequence, i.e.

$$\frac{G(\ell + t_n \delta\ell) - G(\ell)}{t_n} = \left(\frac{\Sigma_n - \Sigma}{t_n}, \frac{\mathbf{u}_n - \mathbf{u}}{t_n} \right) \rightharpoonup (\Sigma', \mathbf{u}') = \delta_w G(\ell; \delta\ell),$$

which is the first assertion of **Theorem 3.2**.

It remains to show that, for fixed ℓ , the weak directional derivative depends Lipschitz continuously on the direction $\delta\ell \in V'$. Similarly to (2.5), we can show that testing (3.2a) with $\mathbf{T} = \Sigma' + (\boldsymbol{\tau}, -\boldsymbol{\tau})$ for $\boldsymbol{\tau} \in S$ implies

$$\mathbb{C}^{-1} \boldsymbol{\sigma}' - \boldsymbol{\varepsilon}(\mathbf{u}') - \mathbb{H}^{-1} \boldsymbol{\chi}' = \mathbf{0}. \quad (3.24)$$

Now let $\delta\ell_1, \delta\ell_2 \in V'$ be two directions and let $(\boldsymbol{\Sigma}'_i, \mathbf{u}'_i) = \delta_w G(\ell; \delta\ell_i)$. We insert $\mathbf{T} = \boldsymbol{\Sigma}'_2$ into (3.2a) in case of $\delta\ell_1$ and vice versa. Then, taking differences, we obtain by the non-negativity of λ

$$\langle A(\boldsymbol{\Sigma}'_1 - \boldsymbol{\Sigma}'_2), \boldsymbol{\Sigma}'_1 - \boldsymbol{\Sigma}'_2 \rangle + \langle B^*(\mathbf{u}'_1 - \mathbf{u}'_2), \boldsymbol{\Sigma}'_1 - \boldsymbol{\Sigma}'_2 \rangle \leq 0.$$

Using (3.2b) shows

$$\langle A(\boldsymbol{\Sigma}'_1 - \boldsymbol{\Sigma}'_2), \boldsymbol{\Sigma}'_1 - \boldsymbol{\Sigma}'_2 \rangle + \langle \mathbf{u}'_1 - \mathbf{u}'_2, \delta\ell_1 - \delta\ell_2 \rangle \leq 0.$$

The Lipschitz continuity now follows from (3.24), the coercivity of A and Korn's inequality (1.5). \square

3.2. Optimality Conditions. We start the discussion of (P) with a chain rule result for general differentiable functionals:

Lemma 3.9. *Let W, H be normed linear spaces and $\mathcal{G} : W \rightarrow H$ weakly directionally differentiable at $w \in W$, i.e.*

$$\frac{\mathcal{G}(w + t\delta w) - \mathcal{G}(w)}{t} \rightarrow \delta_w \mathcal{G}(w; \delta w) \quad \text{in } H \quad \text{as } t \searrow 0 \quad (3.25)$$

for every $\delta w \in W$. Let $J : H \times W \rightarrow \mathbb{R}$ be Fréchet differentiable. Then the functional $j : W \rightarrow \mathbb{R}$, defined by

$$j(w) = J(\mathcal{G}(w), w)$$

is directionally differentiable at w , and its directional derivative in the direction $\delta w \in W$ is given by

$$\delta j(w; \delta w) = J'(\mathcal{G}(w), w)(\delta_w \mathcal{G}(w; \delta w), \delta w). \quad (3.26)$$

Proof. Let $w, \delta w \in W$ be given, $\delta w \neq 0$. Using the Fréchet differentiability of J , we have for $h, \delta h \in H$

$$J(h + \delta h, w + \delta w) - J(h, w) - J'(h, w)(\delta h, \delta w) + r(h, w; \delta h, \delta w) = 0,$$

where the remainder $r : H \times W \times H \times W \rightarrow \mathbb{R}$ satisfies

$$\frac{|r(h, w; \delta h, \delta w)|}{\|(\delta h, \delta w)\|_{H \times W}} \rightarrow 0 \quad \text{as } \|(\delta h, \delta w)\|_{H \times W} \searrow 0. \quad (3.27)$$

We have

$$\begin{aligned} & j(w + t\delta w) - j(w) - tJ'(\mathcal{G}(w), w)(\delta_w \mathcal{G}(w), \delta w) \\ &= J(\mathcal{G}(w + t\delta w), w + t\delta w) - J(\mathcal{G}(w), w) - tJ'(\mathcal{G}(w), w)(\delta_w \mathcal{G}(w), \delta w) \\ &= J'(\mathcal{G}(w), w)(\mathcal{G}(w + t\delta w) - \mathcal{G}(w), t\delta w) - tJ'(\mathcal{G}(w), w)(\delta_w \mathcal{G}(w), \delta w) \\ &\quad + r(\mathcal{G}(w), w; \mathcal{G}(w + t\delta w) - \mathcal{G}(w), t\delta w) \\ &= J'(\mathcal{G}(w), w)(\mathcal{G}(w + t\delta w) - \mathcal{G}(w) - t\delta_w \mathcal{G}(w), 0) \\ &\quad + r(\mathcal{G}(w), w; \mathcal{G}(w + t\delta w) - \mathcal{G}(w), t\delta w). \end{aligned}$$

The operator $J'(\mathcal{G}(w), w) : H \times W \rightarrow \mathbb{R}$ is linear and continuous, hence weakly continuous. Using (3.25), we find

$$\frac{1}{t} J'(\mathcal{G}(w), w)(\mathcal{G}(w + t\delta w) - \mathcal{G}(w) - t\delta_w \mathcal{G}(w), 0) \rightarrow 0 \quad \text{in } \mathbb{R} \quad \text{as } t \searrow 0,$$

where we exploit that weak convergence in \mathbb{R} is equivalent to strong convergence in \mathbb{R} . Moreover, we have

$$\frac{|r(\mathcal{G}(w), w; \mathcal{G}(w + t\delta w) - \mathcal{G}(w), t\delta w)|}{\|(\mathcal{G}(w + t\delta w) - \mathcal{G}(w), t\delta w)\|_{H \times W}} \frac{\|(\mathcal{G}(w + t\delta w) - \mathcal{G}(w), t\delta w)\|_{H \times W}}{t} \rightarrow 0$$

as $t \searrow 0$, since the first factor converges to zero due to (3.27) and since the second factor is bounded due to (3.25). This shows

$$\frac{1}{t} |j(w + t \delta w) - j(w) - t J'(\mathcal{G}(w), w)(\delta_w \mathcal{G}(w), \delta w)| \rightarrow 0$$

as $t \rightarrow 0$. \square

In order to apply the general setting of Lemma 3.9 to our optimal control problem (P), we set

$$\begin{aligned} W &= L^2(\Gamma_N; \mathbb{R}^d), & w &= \mathbf{g}, \\ H &= V, & \mathcal{G} &= G^{(\mathbf{u})} \circ (-\tau_N^*). \end{aligned}$$

The weak directional differentiability of \mathcal{G} follows from Theorem 3.2.

We conclude that local minimizers $\bar{\mathbf{g}}$ of (P) necessarily satisfy (3.1). The following theorem states this in more explicit terms.

Theorem 3.10. *Let $\bar{\mathbf{g}} \in U_{\text{ad}}$ be a local optimal solution of (P) with associated state $(\bar{\Sigma}, \bar{\mathbf{u}}) = G(\bar{\ell})$ for $\bar{\ell} := -\tau_N^* \bar{\mathbf{g}}$. Then the following variational inequality is satisfied:*

$$\delta j(\bar{\mathbf{g}}; \mathbf{g} - \bar{\mathbf{g}}) = J'(\bar{\mathbf{u}}, \bar{\mathbf{g}})(\mathbf{u}', \mathbf{g} - \bar{\mathbf{g}}) \geq 0 \quad \text{for all } \mathbf{g} \in U_{\text{ad}}, \quad (3.28)$$

where (Σ', \mathbf{u}') solves the derivative problem (3.2) with $\delta \ell := -\tau_N^*(\mathbf{g} - \bar{\mathbf{g}})$ as right-hand side, i.e., $(\Sigma', \mathbf{u}') = \delta_w G(-\tau_N^* \bar{\mathbf{g}}; -\tau_N^*(\mathbf{g} - \bar{\mathbf{g}}))$.

The optimality condition (3.28) is a statement about the directional derivatives of the reduced objective in directions from the cone of feasible directions of U_{ad} . In many situations, one may use continuity of the directional derivative w.r.t. the direction to extend the inequality (3.28) to the tangent cone. (In view of the convexity of U_{ad} , the tangent cone is the closure of the cone of feasible directions.) In the present situation, $J'(\bar{\mathbf{u}}, \bar{\mathbf{g}})$ is a bounded linear (hence weakly continuous) map. Hence any additional properties of the directional derivatives of the reduced objective j hinge upon properties of the control-to-state map.

We show that the control-to-state map $G \circ (-\tau_N^*)$ indeed possesses a stronger differentiability property (compare [Sachs, 1978, Definition 2.2]) than merely the weak directional differentiability shown in Theorem 3.2. The compactness of τ_N^* and the Lipschitz continuity of G are crucial here. To simplify notation, we introduce the feasible set of (P)

$$\mathcal{F} := \{(\mathbf{g}, \Sigma, \mathbf{u}) \in L^2(\Gamma_N; \mathbb{R}^d) \times S^2 \times V : \mathbf{g} \in U_{\text{ad}}, (\Sigma, \mathbf{u}) = G(-\tau_N^* \mathbf{g})\}. \quad (3.29)$$

Lemma 3.11. *Let $(\mathbf{g}, \Sigma, \mathbf{u}) \in \mathcal{F}$ and $(\mathbf{g}_n, \Sigma_n, \mathbf{u}_n) \in \mathcal{F}$, and $t_n \searrow 0$ be given such that*

$$\frac{\mathbf{g}_n - \mathbf{g}}{t_n} \rightharpoonup \delta \mathbf{g} \text{ in } L^2(\Gamma_N; \mathbb{R}^d).$$

Then

$$\frac{(\Sigma_n, \mathbf{u}_n) - (\Sigma, \mathbf{u})}{t_n} \rightharpoonup \delta_w G(-\tau_N^* \mathbf{g}; -\tau_N^* \delta \mathbf{g}) \text{ in } S^2 \times V.$$

Proof. Note that $(\mathbf{g}_n, \Sigma_n, \mathbf{u}_n) \in \mathcal{F}$ implies $(\Sigma_n, \mathbf{u}_n) = G(-\tau_N^* \mathbf{g}_n)$. Due to the Lipschitz continuity of G (see Theorem 2.1) and the compactness of τ_N^* , we have

$$\left\| \frac{G(-\tau_N^* \mathbf{g}_n) - G(-\tau_N^*(\mathbf{g} + t_n \delta \mathbf{g}))}{t_n} \right\|_{S^2 \times V} \leq \frac{L}{t_n} \|\tau_N^*(\mathbf{g}_n - (\mathbf{g} + t_n \delta \mathbf{g}))\|_{V'} \rightarrow 0.$$

This implies

$$\begin{aligned} \frac{\Sigma_n - \Sigma}{t_n} &= \frac{\Sigma_n - G^\Sigma(-\tau_N^*(\mathbf{g} + t_n \delta \mathbf{g}))}{t_n} + \frac{G^\Sigma(-\tau_N^*(\mathbf{g} + t_n \delta \mathbf{g})) - \Sigma}{t_n} \\ &\rightarrow \mathbf{0} + \delta_w G^\Sigma(-\tau_N^* \mathbf{g}; -\tau_N^* \delta \mathbf{g}). \end{aligned}$$

Using the same argumentation for the displacements \mathbf{u} yields the claim. \square

This continuity result allows us to extend the inequality (3.28) to the closure of the cone of feasible directions w.r.t. the *weak* topology, i.e., to the tangent cone

$$\mathcal{T}(\mathbf{g}, U_{\text{ad}}) := \left\{ \begin{array}{l} \delta \mathbf{g} \in L^2(\Gamma_N; \mathbb{R}^d) : \text{there exist } \mathbf{g}_n \in U_{\text{ad}} \text{ and } t_n \searrow 0 \text{ s.t.} \\ \frac{\mathbf{g}_n - \mathbf{g}}{t_n} \rightharpoonup \delta \mathbf{g} \text{ in } L^2(\Gamma_N; \mathbb{R}^d) \end{array} \right\} \quad (3.30)$$

for $\mathbf{g} \in U_{\text{ad}}$. The following corollary is now a straightforward consequence of [Theorem 3.10](#), [Lemma 3.11](#) and the weak continuity of $J'(\bar{\mathbf{u}}, \bar{\mathbf{g}})$.

Corollary 3.12. *Let $\bar{\mathbf{g}} \in U_{\text{ad}}$ be a local optimal solution of [\(P\)](#) with associated state $(\bar{\Sigma}, \bar{\mathbf{u}}) = G(\ell)$ for $\ell := -\tau_N^* \bar{\mathbf{g}}$. Then the following variational inequality is satisfied:*

$$\delta j(\bar{\mathbf{g}}; \delta \mathbf{g}) = J'(\bar{\mathbf{u}}, \bar{\mathbf{g}})(\mathbf{u}', \delta \mathbf{g}) \geq 0 \quad \text{for all } \delta \mathbf{g} \in \mathcal{T}(\bar{\mathbf{g}}, U_{\text{ad}}), \quad (3.31)$$

where $(\Sigma', \mathbf{u}') = \delta_w G(-\tau_N^* \bar{\mathbf{g}}; -\tau_N^* \delta \mathbf{g})$.

3.3. Equivalence to B-Stationarity. In this section, we briefly reformulate the optimality condition of [Theorem 3.10](#) and [Corollary 3.12](#) in order to allow a comparison with the B-stationarity conditions known for finite dimensional MPCCs, see [Scheel and Scholtes \[2000\]](#). We start with an equivalent formulation of the variational inequality (3.2) for the derivative, which involves the derivative λ' of the plastic multiplier.

Proposition 3.13. *Let $\ell, \delta \ell \in V'$ be given. Let $(\Sigma, \mathbf{u}, \lambda)$ be the state and plastic multiplier associated with ℓ . A pair $(\Sigma', \mathbf{u}') \in S^2 \times V$ is the unique solution of (3.2) if and only if there exists a multiplier $\lambda' \in L^2(\Omega)$ such that*

$$A\Sigma' + B^*\mathbf{u}' + \lambda \mathcal{D}^* \mathcal{D}\Sigma' + \lambda' \mathcal{D}^* \mathcal{D}\Sigma = \mathbf{0} \quad \text{in } S^2, \quad (3.32a)$$

$$B\Sigma' = \delta \ell \quad \text{in } V', \quad (3.32b)$$

$$\mathbb{R} \ni \lambda'(x) \perp \mathcal{D}\Sigma : \mathcal{D}\Sigma'(x) = 0 \quad \text{a.e. in } \mathcal{A}_s(\ell), \quad (3.32c)$$

$$0 \leq \lambda'(x) \perp \mathcal{D}\Sigma : \mathcal{D}\Sigma'(x) \leq 0 \quad \text{a.e. in } \mathcal{B}(\ell), \quad (3.32d)$$

$$0 = \lambda'(x) \perp \mathcal{D}\Sigma : \mathcal{D}\Sigma'(x) \in \mathbb{R} \quad \text{a.e. in } \mathcal{I}(\ell). \quad (3.32e)$$

Moreover, λ' is unique.

Remark 3.14. *By setting $F = (F_1, F_2) : \mathbb{S} \times \mathbb{R} \rightarrow \mathbb{R}^2$, $F_1(\Sigma, \lambda) := -\phi(\Sigma)$, $F_2(\Sigma, \lambda) := \lambda$, we find that the complementarity relations (3.32c)–(3.32e) are equivalent to*

$$\min \{ F'_i(\Sigma, \lambda)(\Sigma', \lambda') : i \in \{1, 2\} \text{ such that } F_i(\Sigma, \lambda) = 0 \} = 0 \quad \text{a.e. in } \Omega, \quad (3.33)$$

which parallels to the notation of [[Scheel and Scholtes, 2000](#), Section 2.1].

Proof of Proposition 3.13. Suppose that $(\Sigma', \mathbf{u}') \in \mathcal{S}_\ell \times V$ is the unique solution of (3.2), then (3.32b) follows immediately. As seen in [Section 3.1](#), (Σ', \mathbf{u}') equals the weak limit $(\tilde{\Sigma}, \tilde{\mathbf{u}})$ of the difference quotient in (3.6). [Proposition 3.4](#) and [Lemma 3.5](#) imply that there exists a unique $\lambda' = \tilde{\lambda} \in L^2(\Omega)$ such that (3.32a) holds true. Moreover, by (3.16) we have $\lambda' = 0$ on $\mathcal{I}(\ell)$, which is (3.32e). Equation (3.32c) follows from $\Sigma' \in \mathcal{S}_\ell$. The relations on $\mathcal{B}(\ell)$ in (3.32d) follow from $\lambda' \geq 0$ on $\mathcal{B}(\ell)$ by (3.22), $\Sigma' \in \mathcal{S}_\ell$, and from $\lambda' \mathcal{D}\Sigma : \mathcal{D}\tilde{\Sigma} = 0$ by [Proposition 3.8](#). Thus (Σ', \mathbf{u}') , together with λ' , indeed solves (3.32).

If on the other hand (Σ', \mathbf{u}') is a solution of (3.32), then the same arguments as in the proof of [Lemma 3.7](#) yield $\lambda' \mathcal{D}\Sigma : \mathcal{D}\mathbf{T} \leq 0$ a.e. in Ω for all $\mathbf{T} \in \mathcal{S}_\ell$.

Furthermore, the complementarity relations in (3.32c)–(3.32e) immediately imply $\lambda' \mathcal{D}\Sigma : \mathcal{D}\Sigma' = 0$, hence the variational inequality (3.2a) follows from (3.32a) tested with $\mathbf{T} - \Sigma'$. Finally, $\Sigma' \in \mathcal{S}_\ell$ is readily obtained from (3.32c) and (3.32d), and by using that (3.32a), tested with Σ' , implies $\sqrt{\lambda} \mathcal{D}\Sigma' \in S$. \square

Thus we have found the following equivalent presentation of the optimality conditions in Corollary 3.12.

Corollary 3.15. *Let $\bar{\mathbf{g}} \in U_{\text{ad}}$ be a local optimal solution of (P) with associated state $(\bar{\Sigma}, \bar{\mathbf{u}}) = G(\ell)$ for $\ell := -\tau_N^* \bar{\mathbf{g}}$. Then the following variational inequality is satisfied:*

$$J'(\bar{\mathbf{u}}, \bar{\mathbf{g}})(\mathbf{u}', \delta \mathbf{g}) \geq 0 \quad \text{for all directions } (\delta \mathbf{g}, \Sigma', \mathbf{u}', \lambda')$$

satisfying (3.32a), (3.32b) with $\delta \ell = -\tau_N^* \delta \mathbf{g}$, (3.33), and $\delta \mathbf{g} \in \mathcal{T}(\bar{\mathbf{g}}, U_{\text{ad}})$.

Remark 3.16. *This optimality condition is the infinite dimensional version of the B-stationary concept given in [Scheel and Scholtes, 2000, Section 2.1]. We recall that, as in the finite dimensional setting, B-stationarity is a purely primal concept which does not involve any dual variables.*

Already in the case of a finite dimensional MPEC, the combinatorial structure of (3.33) implies that the verification of B-stationarity conditions requires the evaluation of a possibly large number of inequality systems. In the infinite dimensional setting, however, one has to deal even with infinitely many inequalities. Thus, B-stationarity is in general not useful for numerical computations.

3.4. Optimality Conditions Involving the Tangent Cone. In this section, we point out parallels of our implicit programming approach with the tangent cone technique in [Luo et al., 1996, Section 3.3] for the finite dimensional setting. To be precise, we will show that the statement of Corollary 3.12, namely the optimality condition (3.31), is equivalent to the first-order stationarity condition [Luo et al., 1996, p.115]

$$J'(\bar{\mathbf{u}}, \bar{\mathbf{g}})(\delta \mathbf{u}, \delta \mathbf{g}) \geq 0 \quad \text{for all } (\delta \mathbf{g}, \delta \Sigma, \delta \mathbf{u}) \in \mathcal{T}((\bar{\mathbf{g}}, \bar{\Sigma}, \bar{\mathbf{u}}), \mathcal{F}), \quad (3.34)$$

where \mathcal{F} is the feasible set of (P), see (3.29), and $\mathcal{T}((\bar{\mathbf{g}}, \bar{\Sigma}, \bar{\mathbf{u}}), \mathcal{F})$ is its tangent cone. Our notation concerning \mathcal{F} , \mathcal{T} and \mathcal{L} follows Luo et al. [1996].

The elements of a tangent cone are defined via a limit process. In contrast to the finite dimensional setting, the question of topology for these limits arises. In order to derive first-order necessary optimality conditions, the topology used for the definition of the tangent cone has to be chosen compatible with the differentiability properties of the objective, see [Sachs, 1978, Theorem 3.2]. In view of Lemma 3.11, we work with the *weak* topologies of $L^2(\Gamma_N; \mathbb{R}^d)$, S^2 and V .

The tangent cone to \mathcal{F} at $(\mathbf{g}, \Sigma, \mathbf{u}) \in \mathcal{F}$ is thus defined as

$$\mathcal{T}((\mathbf{g}, \Sigma, \mathbf{u}), \mathcal{F}) := \left\{ \begin{array}{l} (\delta \mathbf{g}, \delta \Sigma, \delta \mathbf{u}) \in L^2(\Gamma_N; \mathbb{R}^d) \times S^2 \times V : \\ \text{there exist } (\mathbf{g}_n, \Sigma_n, \mathbf{u}_n) \in \mathcal{F} \text{ and } t_n \searrow 0 \text{ s.t.} \\ \frac{\mathbf{g}_n - \mathbf{g}}{t_n} \rightharpoonup \delta \mathbf{g} \text{ in } L^2(\Gamma_N; \mathbb{R}^d), \\ \frac{(\Sigma_n, \mathbf{u}_n) - (\Sigma, \mathbf{u})}{t_n} \rightharpoonup (\delta \Sigma, \delta \mathbf{u}) \text{ in } S^2 \times V \end{array} \right\}. \quad (3.35)$$

Moreover, we recall the (weak) tangent cone to U_{ad} at a point $\mathbf{g} \in U_{\text{ad}}$, see (3.30),

$$\mathcal{T}(\mathbf{g}, U_{\text{ad}}) := \left\{ \begin{array}{l} \delta \mathbf{g} \in L^2(\Gamma_N; \mathbb{R}^d) : \text{there exist } \mathbf{g}_n \in U_{\text{ad}} \text{ and } t_n \searrow 0 \text{ s.t.} \\ \frac{\mathbf{g}_n - \mathbf{g}}{t_n} \rightharpoonup \delta \mathbf{g} \text{ in } L^2(\Gamma_N; \mathbb{R}^d) \end{array} \right\}.$$

Following [Luo et al., 1996, eq. (3.2.9), p.123], we define the linearized cone

$$\mathcal{L}((\mathbf{g}, \boldsymbol{\Sigma}, \mathbf{u}), \mathcal{F}) := \left\{ \begin{array}{l} (\delta \mathbf{g}, \delta \boldsymbol{\Sigma}, \delta \mathbf{u}) \in L^2(\Gamma_N; \mathbb{R}^d) \times S^2 \times V : \delta \mathbf{g} \in \mathcal{T}(\mathbf{g}, U_{\text{ad}}), \\ (\delta \boldsymbol{\Sigma}, \delta \mathbf{u}) \text{ solves the AVI (3.2) with } \delta \ell = -\tau_N^* \delta \mathbf{g} \\ \text{and some } \lambda \text{ satisfying (2.1)} \end{array} \right\}. \quad (3.36)$$

We can now exploit the uniqueness of λ (Theorem 2.2) and the unique solvability of the AVI (affine variational inequality) (3.2), whose solution is denoted by $\delta_w G(-\tau_N^* \mathbf{g}; -\tau_N^* \delta \mathbf{g})$. We thus conclude that the linearized cone is indeed

$$\mathcal{L}((\mathbf{g}, \boldsymbol{\Sigma}, \mathbf{u}), \mathcal{F}) = \left\{ \begin{array}{l} (\delta \mathbf{g}, \delta \boldsymbol{\Sigma}, \delta \mathbf{u}) \in L^2(\Gamma_N; \mathbb{R}^d) \times S^2 \times V : \delta \mathbf{g} \in \mathcal{T}(\mathbf{g}, U_{\text{ad}}), \\ (\delta \boldsymbol{\Sigma}, \delta \mathbf{u}) = \delta_w G(-\tau_N^* \mathbf{g}; -\tau_N^* \delta \mathbf{g}) \end{array} \right\}. \quad (3.37)$$

Note that the admissible set \mathcal{S}_ℓ (with $\ell = -\tau_N^* \mathbf{g}$) for the AVI corresponds to the *directional critical set along* $\delta \mathbf{g} \in \mathcal{T}(\mathbf{g}, U_{\text{ad}})$ in [Luo et al., 1996, Lemma 3.2.1]. In the present setting, the set \mathcal{S}_ℓ is independent of the particular direction $\delta \mathbf{g}$.

In view of these definitions, our optimality condition (3.31) reads

$$J'(\bar{\mathbf{u}}, \bar{\mathbf{g}})(\delta \mathbf{u}, \delta \mathbf{g}) \geq 0 \quad \text{for all } (\delta \mathbf{g}, \delta \boldsymbol{\Sigma}, \delta \mathbf{u}) \in \mathcal{L}((\bar{\mathbf{g}}, \bar{\boldsymbol{\Sigma}}, \bar{\mathbf{u}}), \mathcal{F}), \quad (3.31')$$

which is (3.34) with \mathcal{T} replaced by \mathcal{L} . In order to show the desired equivalence of (3.34) and (3.31'), it remains to verify $\mathcal{T} = \mathcal{L}$.

The equality $\mathcal{T} = \mathcal{L}$ is termed the *full constraint qualification* in [Luo et al., 1996, Section 3.3], and it coincides with the *extreme CQ* as well as with the *basic CQ* since the plastic multiplier λ is unique. In practice, one often invokes stronger constraint qualifications which are more manageable, see [Luo et al., 1996, Chapter 4].

One of these stronger constraint qualifications is the BIF (Bouligand differentiable implicit function condition), see [Luo et al., 1996, Section 4.2.2], which means that the set \mathcal{F} is (locally) solvable for $(\boldsymbol{\Sigma}, \mathbf{u})$ as a function of \mathbf{g} , and that this function is directionally differentiable. Under some additional assumptions, Theorem 4.2.31 of Luo et al. [1996] then implies that $\mathcal{T} = \mathcal{L}$. It is not straightforward to adapt this line of reasoning to a general infinite dimensional setting. Therefore, we prove that $\mathcal{L}((\mathbf{g}, \boldsymbol{\Sigma}, \mathbf{u}), \mathcal{F}) = \mathcal{T}((\mathbf{g}, \boldsymbol{\Sigma}, \mathbf{u}), \mathcal{F})$ holds in all feasible points in a direct way. The argument uses again the additional differentiability property shown in Lemma 3.11.

Lemma 3.17. *Let $(\mathbf{g}, \boldsymbol{\Sigma}, \mathbf{u}) \in \mathcal{F}$. Then*

$$\mathcal{T}((\mathbf{g}, \boldsymbol{\Sigma}, \mathbf{u}), \mathcal{F}) = \mathcal{L}((\mathbf{g}, \boldsymbol{\Sigma}, \mathbf{u}), \mathcal{F}).$$

Proof. We first show that $\mathcal{T}((\mathbf{g}, \boldsymbol{\Sigma}, \mathbf{u}), \mathcal{F}) \subset \mathcal{L}((\mathbf{g}, \boldsymbol{\Sigma}, \mathbf{u}), \mathcal{F})$ holds. To this end, let $(\delta \mathbf{g}, \delta \boldsymbol{\Sigma}, \delta \mathbf{u}) \in \mathcal{T}((\mathbf{g}, \boldsymbol{\Sigma}, \mathbf{u}), \mathcal{F})$ be given. By definition, there exist a sequence $(\mathbf{g}_n, \boldsymbol{\Sigma}_n, \mathbf{u}_n) \in \mathcal{F}$ and $t_n \searrow 0$ such that

$$\frac{\mathbf{g}_n - \mathbf{g}}{t_n} \rightharpoonup \delta \mathbf{g} \text{ in } L^2(\Gamma_N; \mathbb{R}^d) \quad \text{and} \quad \frac{(\boldsymbol{\Sigma}_n, \mathbf{u}_n) - (\boldsymbol{\Sigma}, \mathbf{u})}{t_n} \rightharpoonup (\delta \boldsymbol{\Sigma}, \delta \mathbf{u}) \text{ in } S^2 \times V.$$

In particular, this implies $\delta \mathbf{g} \in \mathcal{T}(\mathbf{g}, U_{\text{ad}})$. Moreover, due to Lemma 3.11, we have

$$(\delta \boldsymbol{\Sigma}, \delta \mathbf{u}) = \delta_w G(-\tau_N^* \mathbf{g}; -\tau_N^* \delta \mathbf{g}).$$

This shows $(\delta \mathbf{g}, \delta \boldsymbol{\Sigma}, \delta \mathbf{u}) \in \mathcal{L}((\mathbf{g}, \boldsymbol{\Sigma}, \mathbf{u}), \mathcal{F})$.

Now, for the converse inclusion, let $(\delta \mathbf{g}, \delta \boldsymbol{\Sigma}, \delta \mathbf{u}) \in \mathcal{L}((\mathbf{g}, \boldsymbol{\Sigma}, \mathbf{u}), \mathcal{F})$ be given. Since $\delta \mathbf{g} \in \mathcal{T}(\mathbf{g}, U_{\text{ad}})$, there exist sequences $\mathbf{g}_n \in U_{\text{ad}}$ and $t_n \searrow 0$ such that

$$\frac{\mathbf{g}_n - \mathbf{g}}{t_n} \rightharpoonup \delta \mathbf{g} \quad \text{in } L^2(\Gamma_N; \mathbb{R}^d).$$

Let us denote by $(\boldsymbol{\Sigma}_n, \mathbf{u}_n) = G(-\tau_N^*(\mathbf{g}_n))$ the associated stresses and displacements. In particular, we have $(\mathbf{g}_n, \boldsymbol{\Sigma}_n, \mathbf{u}_n) \in \mathcal{F}$. Using [Lemma 3.11](#), we conclude that

$$\frac{(\boldsymbol{\Sigma}_n, \mathbf{u}_n) - (\boldsymbol{\Sigma}, \mathbf{u})}{t_n} \rightarrow \delta_w G(-\tau_N^* \mathbf{g}; -\tau_N^* \delta \mathbf{g}) \quad \text{in } S^2 \times V.$$

Since $(\delta \boldsymbol{\Sigma}, \delta \mathbf{u}) = \delta_w G(-\tau_N^* \mathbf{g}; -\tau_N^* \delta \mathbf{g})$ by definition of $\mathcal{L}((\mathbf{g}, \boldsymbol{\Sigma}, \mathbf{u}), \mathcal{F})$, we infer $(\delta \mathbf{g}, \delta \boldsymbol{\Sigma}, \delta \mathbf{u}) \in \mathcal{T}((\mathbf{g}, \boldsymbol{\Sigma}, \mathbf{u}), \mathcal{F})$. \square

4 Strong Stationarity

The main part of this section is devoted to the derivation of strong stationarity conditions for a modified problem in [Section 4.1](#). We obtain a system consistent with the notion of strong stationarity for finite dimensional MPCCs as in [Scheel and Scholtes \[2000\]](#). Subsequently, we derive in [Section 4.2](#) an equivalent formulation which coincides with the optimality conditions given in [Mignot and Puel \[1984\]](#) for the obstacle problem, but which were not termed ‘strong stationarity conditions’ at the time. Compared to [Mignot and Puel \[1984\]](#), we present a different and more elementary technique of proof which avoids the concept of conical derivatives. In [Section 4.3](#), we state strong stationarity conditions for the original problem [\(P\)](#) (without proving their necessity) and show that they imply the B-stationarity conditions.

We suppose [Assumption 1.1](#) to hold throughout this section.

4.1. Strong Stationarity for a Modified Problem. We consider the following modification of the optimal control problem [\(P\)](#):

$$\left. \begin{array}{l} \text{Minimize } \tilde{J}(\mathbf{u}, \ell, \mathcal{L}) \\ \text{s.t. } \langle A\boldsymbol{\Sigma}, \mathbf{T} - \boldsymbol{\Sigma} \rangle + \langle B^* \mathbf{u}, \mathbf{T} - \boldsymbol{\Sigma} \rangle \geq \langle \mathcal{L}, \mathbf{T} - \boldsymbol{\Sigma} \rangle \quad \text{for all } \mathbf{T} \in \mathcal{K} \\ B\boldsymbol{\Sigma} = \ell \quad \text{and} \quad \boldsymbol{\Sigma} \in \mathcal{K}, \end{array} \right\} (\tilde{\mathbf{P}})$$

where $\tilde{J}: V \times V' \times S^2 \rightarrow \mathbb{R}$ is Fréchet differentiable.

Problem [\(P̃\)](#) differs from [\(P\)](#) in the following ways. First of all, the term $\ell \in V'$ is used directly as a control variable, whereas in [\(P\)](#), only loads $\ell = -\tau_N^* \mathbf{g}$ induced by boundary control functions \mathbf{g} were used. Secondly, $\mathcal{L} \in S^2$ appears as an additional control variable on the right-hand side of the variational inequality. Finally, no control constraints are present. These modifications are in accordance with previous strong stationarity results for control of the obstacle problem, see [Mignot and Puel \[1984\]](#), where it was also required that the set of feasible controls maps onto the range space of the variational inequality.

In the elastic case, i.e. $\mathcal{K} = S^2$ and $\boldsymbol{\chi} \equiv \mathbf{0}$, the control \mathcal{L} can be interpreted as a prestress applied to the work piece Ω . In the context of elastoplasticity however, problem [\(P̃\)](#) is of rather academic nature. Nevertheless, it is worth to be investigated since the particular structure of [\(P̃\)](#) allows to establish first-order optimality conditions in strongly stationary form. In contrast to the results in [Section 3](#), strongly stationary optimality systems involve adjoint states and Lagrange multipliers. Our technique for the derivation of strong stationarity conditions differs completely from the one used in [Mignot and Puel \[1984\]](#) for optimal control of the obstacle problem. Instead of transforming the B-stationarity conditions as done in [Mignot and Puel \[1984\]](#), we introduce two auxiliary problems, which turn out to be ‘standard’ control problems that allow the application of the generalized Karush-Kuhn-Tucker (KKT) theory. As a local optimal control of [\(P̃\)](#) is also locally optimal for these auxiliary problems, the associated KKT systems apply and strong stationarity is a consequence of these KKT systems.

Let us consider a fixed local optimum $(\bar{\ell}, \bar{\mathcal{L}}) \in V' \times S^2$. We do not address the existence of a global solution to $(\tilde{\mathbf{P}})$. This is a delicate question since the solution operator $(\ell, \mathcal{L}) \mapsto (\boldsymbol{\Sigma}, \mathbf{u})$ associated with the variational inequality in $(\tilde{\mathbf{P}})$ is not completely continuous from $S^2 \times V'$ to $S^2 \times V$. Thus, the discussion of global existence for $(\tilde{\mathbf{P}})$ would go beyond the scope of this paper.

The results for (2.1) readily transfer to the variational inequality in $(\tilde{\mathbf{P}})$ since the underlying analysis is not affected by the additional inhomogeneity $\mathcal{L} \in S^2$. Therefore, we find the following result analogous to [Theorem 2.2](#):

Lemma 4.1. *For every $(\ell, \mathcal{L}) \in V' \times S^2$, there is a unique solution $(\boldsymbol{\Sigma}, \mathbf{u}) \in S^2 \times V$ of the variational inequality in $(\tilde{\mathbf{P}})$. The pair $(\boldsymbol{\Sigma}, \mathbf{u}) \in S^2 \times V$ is the unique solution if and only if there exists a plastic multiplier $\lambda \in L^2(\Omega)$ such that*

$$A\boldsymbol{\Sigma} + B^*\mathbf{u} + \lambda \mathcal{D}^*\mathcal{D}\boldsymbol{\Sigma} = \mathcal{L} \quad \text{in } S^2, \quad (4.1a)$$

$$B\boldsymbol{\Sigma} = \ell \quad \text{in } V', \quad (4.1b)$$

$$0 \leq \lambda(x) \perp \phi(\boldsymbol{\Sigma}(x)) \leq 0 \quad \text{a.e. in } \Omega \quad (4.1c)$$

holds. Moreover, λ is unique.

In view of this lemma, $(\tilde{\mathbf{P}})$ is equivalent to

$$\text{Minimize } \tilde{J}(\mathbf{u}, \ell, \mathcal{L}) \quad \text{s.t. } (4.1).$$

In finite dimensions, strong stationarity can be proved by the local decomposition approach, see for instance [[Scheel and Scholtes, 2000](#), Theorem 2.2]. To this end, we have to fix some notation. Let us denote by $(\bar{\boldsymbol{\Sigma}}, \bar{\mathbf{u}}, \bar{\lambda})$ the solution of (4.1) for controls $(\bar{\ell}, \bar{\mathcal{L}})$. Moreover, similarly to (2.2), we define up to sets of measure zero

$$\bar{\mathcal{A}} := \{x \in \Omega : \phi(\bar{\boldsymbol{\Sigma}}(x)) = 0\} \quad (\text{active set}) \quad (4.2a)$$

$$\bar{\mathcal{A}}_s := \{x \in \Omega : \bar{\lambda}(x) > 0\} \quad (\text{strongly active set}) \quad (4.2b)$$

$$\bar{\mathcal{B}} := \{x \in \Omega : \phi(\bar{\boldsymbol{\Sigma}}(x)) = \bar{\lambda}(x) = 0\} = \bar{\mathcal{A}} \setminus \bar{\mathcal{A}}_s \quad (\text{biactive set}) \quad (4.2c)$$

$$\bar{\mathcal{I}} := \{x \in \Omega : \phi(\bar{\boldsymbol{\Sigma}}(x)) < 0\} = \Omega \setminus \bar{\mathcal{A}} \quad (\text{inactive set}). \quad (4.2d)$$

The local decomposition approach introduces auxiliary problems which avoid the complementarity condition (4.1c). Similar to [[Scheel and Scholtes, 2000](#), eq. (2)], we consider measurable (not necessarily disjoint) decompositions $\Omega = \tilde{\mathcal{A}} \cup \tilde{\mathcal{I}}$ satisfying

$$\begin{aligned} \bar{\mathcal{A}} &= \bar{\mathcal{A}}_s \cup \bar{\mathcal{B}} \supset \tilde{\mathcal{A}} \supset \bar{\mathcal{A}}_s, \\ \bar{\mathcal{I}} \cup \bar{\mathcal{B}} &\supset \tilde{\mathcal{I}} \supset \bar{\mathcal{I}}. \end{aligned} \quad (4.3)$$

Note that this decomposition is unique if $\bar{\mathcal{B}}$ is a set of measure zero, i.e., if strict complementarity holds. We tighten the complementarity condition (4.1c) and replace it by

$$\begin{aligned} \lambda &\geq 0 \quad \text{a.e. on } \Omega, & \phi(\boldsymbol{\Sigma}) &\leq 0 \quad \text{a.e. on } \Omega, \\ \lambda &= 0 \quad \text{a.e. on } \tilde{\mathcal{I}}, & \phi(\boldsymbol{\Sigma}) &= 0 \quad \text{a.e. on } \tilde{\mathcal{A}}. \end{aligned} \quad (4.4)$$

We can now define the auxiliary problems (without a complementarity condition)

$$\text{Minimize } \tilde{J}(\mathbf{u}, \ell, \mathcal{L}) \quad \text{s.t. } (4.1a), (4.1b), (4.4). \quad (\tilde{\mathbf{P}}_{\tilde{\mathcal{A}}, \tilde{\mathcal{I}}})$$

In order to adapt the proof of strong stationarity in [[Scheel and Scholtes, 2000](#), Theorem 2.2] to the present setting, we would need to show the following steps. (i) A minimizer of $(\tilde{\mathbf{P}})$ satisfies a system of weak stationarity. (ii) If strong stationarity is violated, there exists a decomposition $\tilde{\mathcal{A}}, \tilde{\mathcal{I}}$ satisfying (4.3) such that this minimizer is *not* a KKT point of $(\tilde{\mathbf{P}}_{\tilde{\mathcal{A}}, \tilde{\mathcal{I}}})$. This would lead to a contradiction,

provided that $(\tilde{\mathbf{P}}_{\tilde{\mathcal{A}}, \tilde{\mathcal{I}}})$ satisfies a constraint qualification. The latter is ensured in finite dimensions by the assumption of SMFCQ for $(\tilde{\mathbf{P}})$.

It is not straightforward to transfer this technique of proof to our infinite dimensional problem for two reasons. (i) — the weak stationarity for minimizers of $(\tilde{\mathbf{P}})$ is not immediate. (ii) — an infinite dimensional version of SMFCQ is missing.

We therefore give here a direct proof of strong stationarity consisting of the following main steps.

- (1) We show a constraint qualification for the auxiliary problems ([Lemma 4.3](#)).
- (2) Using the KKT conditions for the auxiliary problems, we can prove that the system of strong stationarity is satisfied ([Theorem 4.5](#)).

Moreover, by inspecting the proof of [[Scheel and Scholtes, 2000](#), Theorem 2.2], we find that it is sufficient to consider only *two* particular auxiliary problems, corresponding to the two extremal choices in ([4.3](#)), i.e.,

$$\begin{aligned} \tilde{\mathcal{A}}_1 &= \bar{\mathcal{A}}_s, & \tilde{\mathcal{A}}_2 &= \bar{\mathcal{A}}, \\ \tilde{\mathcal{I}}_1 &= \Omega \setminus \tilde{\mathcal{A}}_1 = \bar{\mathcal{B}} \cup \bar{\mathcal{I}}, & \tilde{\mathcal{I}}_2 &= \Omega \setminus \tilde{\mathcal{A}}_2 = \bar{\mathcal{I}}. \end{aligned}$$

These choices lead to the definitions of the convex feasible sets

$$\begin{aligned} Z_i &:= \{\Sigma \in S^2 : \phi(\Sigma(x)) \leq 0 \text{ a.e. in } \tilde{\mathcal{I}}_i\}, \\ M_i &:= \{\lambda \in L^2(\Omega) : \lambda(x) \geq 0 \text{ a.e. in } \tilde{\mathcal{A}}_i, \quad \lambda(x) = 0 \text{ a.e. in } \tilde{\mathcal{I}}_i\} \end{aligned}$$

with $i = 1, 2$, cf. ([4.4](#)). While the convex constraints in ([4.4](#)) have been incorporated into the definition of the sets Z_i and M_i (which are therefore convex), the non-convex constraint $\phi(\Sigma(x)) = 0$ will be treated in an explicit way. The two auxiliary problems under consideration can now be stated as

$$\left. \begin{aligned} &\text{Minimize} && \tilde{J}(\mathbf{u}, \ell, \mathcal{L}) \\ &\text{s.t.} && A\Sigma + B^*\mathbf{u} + \lambda \mathcal{D}^* \mathcal{D}\Sigma = \mathcal{L} \quad \text{in } S^2, \\ &&& B\Sigma = \ell \quad \text{in } V', \\ &&& \phi(\Sigma(x)) = 0 \quad \text{a.e. in } \tilde{\mathcal{A}}_i, \\ &\text{and} && \Sigma \in Z_i, \quad \lambda \in M_i, \end{aligned} \right\} \quad (\tilde{\mathbf{P}}_i)$$

with $i = 1, 2$.

Since every feasible point of $(\tilde{\mathbf{P}}_1)$ and $(\tilde{\mathbf{P}}_2)$ is feasible for $(\tilde{\mathbf{P}})$ as well, we have the following result.

Lemma 4.2. *Let $(\bar{\ell}, \bar{\mathcal{L}}) \in V' \times S^2$ be a local optimal solution to $(\tilde{\mathbf{P}})$ with associated state $(\bar{\Sigma}, \bar{\mathbf{u}}, \bar{\lambda})$. Then $(\bar{\ell}, \bar{\mathcal{L}})$ is also locally optimal for both auxiliary problems $(\tilde{\mathbf{P}}_1)$ and $(\tilde{\mathbf{P}}_2)$.*

In order to apply the KKT theory in Banach spaces to $(\tilde{\mathbf{P}}_1)$ and $(\tilde{\mathbf{P}}_2)$, we verify the constraint qualification of [Zowe and Kurcyusz \[1979\]](#) which is frequently also termed regular point condition. To this end, let us introduce the space S_∞^2 by

$$S_\infty^2 := \{\mathbf{T} \in S^2 : \mathcal{D}\mathbf{T} \in L^\infty(\Omega; \mathbb{S})\}.$$

Endowed with the norm $\|\mathbf{T}\|_{S^2} + \|\mathcal{D}\mathbf{T}\|_{L^\infty(\Omega; \mathbb{S})}$, S_∞^2 becomes a Banach space. Note that every Σ satisfying the constraints in $(\tilde{\mathbf{P}})$, $(\tilde{\mathbf{P}}_1)$, or $(\tilde{\mathbf{P}}_2)$, respectively, is an element of S_∞^2 . This is due to the structure of ϕ , see ([1.1](#)).

Let us abbreviate $\mathbf{x} := (\boldsymbol{\Sigma}, \mathbf{u}, \lambda, \ell, \mathcal{L})$ and define $e_i : S_\infty^2 \times V \times L^2(\Omega) \times V' \times S^2 \rightarrow S^2 \times V' \times L^\infty(\tilde{\mathcal{A}}_i)$, $i = 1, 2$, by

$$e_i(\mathbf{x}) := \begin{pmatrix} A\boldsymbol{\Sigma} + B^*\mathbf{u} + \lambda \mathcal{D}^* \mathcal{D}\boldsymbol{\Sigma} - \mathcal{L} \\ B\boldsymbol{\Sigma} - \ell \\ \phi(\boldsymbol{\Sigma})|_{\tilde{\mathcal{A}}_i} \end{pmatrix}$$

where $\phi(\boldsymbol{\Sigma})|_{\tilde{\mathcal{A}}_i}$ denotes the restriction of $\phi(\boldsymbol{\Sigma})$ to $\tilde{\mathcal{A}}_i$. Note that the equality constraints in $(\tilde{\mathbf{P}}_1)$ and $(\tilde{\mathbf{P}}_2)$ are equivalent to $e_i(\mathbf{x}) = \mathbf{0}$. It is easy to see that e_i is of class C^1 since the nonlinear terms are differentiable with respect to $\boldsymbol{\Sigma} \in S_\infty^2$ and $\lambda \in L^2(\Omega)$. Finally, we define the cones

$$C_i(\bar{\mathbf{x}}) := \{t(\mathbf{x} - \bar{\mathbf{x}}) : t \geq 0, \mathbf{x} = (\boldsymbol{\Sigma}, \mathbf{u}, \lambda, \ell, \mathcal{L}) \in S_\infty^2 \times V \times L^2(\Omega) \times V' \times S^2, \boldsymbol{\Sigma} \in Z_i, \lambda \in M_i\}$$

for $i = 1, 2$.

Lemma 4.3. *Let $\bar{\mathbf{x}} \in S_\infty^2 \times V \times L^2(\Omega) \times V' \times S^2$ be feasible for $(\tilde{\mathbf{P}})$. Then there holds*

$$e'_i(\bar{\mathbf{x}}) C_i(\bar{\mathbf{x}}) = S^2 \times V' \times L^\infty(\tilde{\mathcal{A}}_i)$$

for $i = 1, 2$. Consequently, the regular point condition is fulfilled for $(\tilde{\mathbf{P}}_1)$ and $(\tilde{\mathbf{P}}_2)$.

Proof. The derivative of e_i at $\bar{\mathbf{x}}$ in the direction $\delta\mathbf{x} = (\delta\boldsymbol{\Sigma}, \delta\mathbf{u}, \delta\lambda, \delta\ell, \delta\mathcal{L})$ is given by

$$e'_i(\bar{\mathbf{x}}) \delta\mathbf{x} = \begin{pmatrix} A\delta\boldsymbol{\Sigma} + B^*\delta\mathbf{u} + \bar{\lambda} \mathcal{D}^* \mathcal{D} \delta\boldsymbol{\Sigma} + \delta\lambda \mathcal{D}^* \mathcal{D}\bar{\boldsymbol{\Sigma}} - \delta\mathcal{L} \\ B\delta\boldsymbol{\Sigma} - \delta\ell \\ (\mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D} \delta\boldsymbol{\Sigma})|_{\tilde{\mathcal{A}}_i} \end{pmatrix},$$

with $i = 1, 2$. Now let $(\mathcal{L}, \ell, f) \in S^2 \times V' \times L^\infty(\tilde{\mathcal{A}}_i)$ be arbitrary. By construction of $C_i(\bar{\mathbf{x}})$, we see that

$$\delta\mathbf{x}_i = \begin{pmatrix} \delta\boldsymbol{\Sigma}_i \\ \delta\mathbf{u}_i \\ \delta\lambda_i \\ \delta\ell_i \\ \delta\mathcal{L}_i \end{pmatrix} := \begin{pmatrix} \chi_{\tilde{\mathcal{A}}_i} \bar{\boldsymbol{\Sigma}} / \bar{\sigma}_0^2 f \\ \mathbf{0} \\ 0 \\ -\ell + B(\chi_{\tilde{\mathcal{A}}_i} \bar{\boldsymbol{\Sigma}} / \bar{\sigma}_0^2 f) \\ -\mathcal{L} + (A + \bar{\lambda} \mathcal{D}^* \mathcal{D}) \chi_{\tilde{\mathcal{A}}_i} \bar{\boldsymbol{\Sigma}} / \bar{\sigma}_0^2 f \end{pmatrix} \in C_i(\bar{\mathbf{x}}) \quad (4.5)$$

belongs to $C_i(\bar{\mathbf{x}})$ for $i = 1, 2$. Here $\chi_{\tilde{\mathcal{A}}_i}$ denotes the characteristic function on $\tilde{\mathcal{A}}_i$. Note that $C_i(\bar{\mathbf{x}})$ does not contain any restriction on the $\boldsymbol{\Sigma}$ component on the set $\tilde{\mathcal{A}}_i$. In view of $|\mathcal{D}\bar{\boldsymbol{\Sigma}}|^2 = \bar{\sigma}_0^2$ on $\tilde{\mathcal{A}}_i$ for $i = 1, 2$, we have $(\mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D} \delta\boldsymbol{\Sigma}_i)|_{\tilde{\mathcal{A}}_i} = f$ and hence

$$e'_i(\bar{\mathbf{x}}) \delta\mathbf{x}_i = (\mathcal{L}, \ell, f),$$

which establishes the case. \square

Remark 4.4. *We point out that the presence of “ample” controls which cover the entire range space $S^2 \times V'$ of the variational inequality, is essential for the verification of the regular point condition. To our best knowledge, it is an open question how to verify a suitable constraint qualification if additional restrictions on the control are present, as for instance in the case of (\mathbf{P}) , where $\mathcal{L} = \mathbf{0}$ and ℓ is induced by a boundary control. This is the main reason why the following analysis does not apply to (\mathbf{P}) .*

Similarly to our result, the technique of Mignot and Puel [1984] for the derivation of strong stationarity conditions for optimal control of the obstacle problem also requires “ample” controls, i.e. distributed controls that are not restricted by additional control constraints, see [Mignot and Puel, 1984, Section 4]. It is straightforward to see that the upcoming analysis can be adapted to optimal control of the obstacle

problem and delivers the same result as in [Mignot and Puel \[1984\]](#) but by a different technique, provided that the obstacle problem under consideration is H^2 -regular.

We are now in the position to prove a first-order necessary optimality system of strongly stationary type for $(\tilde{\mathbf{P}})$. For convenience, we summarize our notation for primal and dual quantities in [Table 4.1](#). Note that — as is usual in the study of MPCCs — there is no Lagrange multiplier associated with the complementarity constraint $\lambda \phi(\boldsymbol{\Sigma}) = 0$.

	state variable	adjoint variable
generalized stresses	$\boldsymbol{\Sigma}$	$\boldsymbol{\Upsilon}$
displacement field	\mathbf{u}	\mathbf{w}
	constraint	associated multiplier
plastic multiplier	$\lambda \geq 0$	μ
yield condition	$\phi(\boldsymbol{\Sigma}) \leq 0$	θ

TABLE 4.1. Summary of primal and dual variable names

Theorem 4.5. *Let $(\bar{\mathcal{L}}, \bar{\ell}) \in S^2 \times V'$ be a locally optimal solution to $(\tilde{\mathbf{P}})$ with associated optimal state $(\bar{\boldsymbol{\Sigma}}, \bar{\mathbf{u}}, \bar{\lambda}) \in S_\infty^2 \times V \times L^2(\Omega)$. Then there exists an adjoint state $(\boldsymbol{\Upsilon}, \mathbf{w}) \in S^2 \times V$ and Lagrange multipliers $\mu \in L^2(\Omega)$ and $\theta \in L^2(\Omega)$ such that the following optimality system is fulfilled:*

$$A\bar{\boldsymbol{\Sigma}} + \bar{\lambda} \mathcal{D}^* \mathcal{D} \bar{\boldsymbol{\Sigma}} + B^* \bar{\mathbf{u}} = \bar{\mathcal{L}} \quad (4.6a)$$

$$B\bar{\boldsymbol{\Sigma}} = \bar{\ell} \quad (4.6b)$$

$$0 \leq \bar{\lambda} \perp \phi(\bar{\boldsymbol{\Sigma}}) \leq 0 \quad \text{a.e. in } \Omega \quad (4.6c)$$

$$A\boldsymbol{\Upsilon} + B^* \mathbf{w} + \bar{\lambda} \mathcal{D}^* \mathcal{D} \boldsymbol{\Upsilon} + \theta \mathcal{D}^* \mathcal{D} \bar{\boldsymbol{\Sigma}} = \mathbf{0} \quad (4.7a)$$

$$B\boldsymbol{\Upsilon} = -\partial_{\mathbf{u}} \tilde{J}(\bar{\mathbf{u}}, \bar{\ell}, \bar{\mathcal{L}}) \quad (4.7b)$$

$$\partial_{\mathcal{L}} \tilde{J}(\bar{\mathbf{u}}, \bar{\ell}, \bar{\mathcal{L}}) - \boldsymbol{\Upsilon} = \mathbf{0} \quad (4.8a)$$

$$\partial_{\ell} \tilde{J}(\bar{\mathbf{u}}, \bar{\ell}, \bar{\mathcal{L}}) - \mathbf{w} = \mathbf{0} \quad (4.8b)$$

$$\mathcal{D} \bar{\boldsymbol{\Sigma}} : \mathcal{D} \boldsymbol{\Upsilon} - \mu = 0 \quad (4.9a)$$

$$\mu \bar{\lambda} = 0 \quad \text{a.e. in } \Omega \quad (4.9b)$$

$$\theta \phi(\bar{\boldsymbol{\Sigma}}) = 0 \quad \text{a.e. in } \Omega \quad (4.9c)$$

$$\theta \geq 0, \quad \mu \geq 0 \quad \text{a.e. in } \bar{\mathcal{B}} \quad (4.9d)$$

Moreover, the adjoint states $\boldsymbol{\Upsilon}$ and \mathbf{w} and Lagrange multipliers μ and θ are unique. Here, $(\partial_{\mathbf{u}} \tilde{J}(\cdot), \partial_{\ell} \tilde{J}(\cdot), \partial_{\mathcal{L}} \tilde{J}(\cdot)) \in V' \times V \times S^2$ denotes the partial derivatives of \tilde{J} w.r.t. $\{\mathbf{u}, \ell, \mathcal{L}\}$.

Remark 4.6. *In contrast to the C-stationarity conditions discussed in [Herzog et al. \[2012\]](#), the strong stationarity conditions provide a sign on the biactive set not only for the product of the multipliers, but even for each multiplier individually, cf. [\(4.9d\)](#). This is the essential difference between C- and strong stationarity, see also [Scheel and Scholtes \[2000\]](#).*

Proof of Theorem 4.5. We start by defining the Lagrangian $\mathcal{L}_i : S_\infty^2 \times V \times L^2(\Omega) \times V' \times S^2 \times S^2 \times V \times L^\infty(\tilde{\mathcal{A}}_i)' \rightarrow \mathbb{R}$ associated with $(\tilde{\mathbf{P}}_i)$, $i = 1, 2$, by

$$\begin{aligned} \mathcal{L}_i(\boldsymbol{\Sigma}, \mathbf{u}, \lambda, \ell, \mathcal{L}, \boldsymbol{\Upsilon}, \mathbf{w}, \theta) &:= \tilde{J}(\mathbf{u}, \ell, \mathcal{L}) \\ &+ \langle A\boldsymbol{\Sigma}, \boldsymbol{\Upsilon} \rangle + \langle B^* \mathbf{u}, \boldsymbol{\Upsilon} \rangle + \langle \lambda, \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\boldsymbol{\Upsilon} \rangle_\Omega - \langle \mathcal{L}, \boldsymbol{\Upsilon} \rangle \\ &+ \langle B\boldsymbol{\Sigma}, \mathbf{w} \rangle - \langle \ell, \mathbf{w} \rangle + \langle \phi(\boldsymbol{\Sigma}), \theta \rangle_{L^\infty(\tilde{\mathcal{A}}_i), L^\infty(\tilde{\mathcal{A}}_i)'}. \end{aligned}$$

As was mentioned before, the convex constraints $\boldsymbol{\Sigma} \in Z_i$ and $\lambda \in M_i$ are treated as abstract constraints and hence they are not included in the definition of \mathcal{L}_i . According to Lemma 4.3, the constraint qualification of Zowe and Kurcyusz holds for $(\tilde{\mathbf{P}}_i)$, $i = 1, 2$.

Let us abbreviate $\bar{\mathbf{x}} := (\bar{\boldsymbol{\Sigma}}, \bar{\mathbf{u}}, \bar{\lambda}, \bar{\ell}, \bar{\mathcal{L}})$. Since $(\bar{\mathcal{L}}, \bar{\ell})$ is locally optimal for $(\tilde{\mathbf{P}}_i)$, $i = 1, 2$, there exist Lagrange multipliers $(\boldsymbol{\Upsilon}_i, \mathbf{w}_i, \theta_i) \in S^2 \times V \times L^\infty(\tilde{\mathcal{A}}_i)'$ such that the following optimality systems are satisfied:

$$\partial_{\mathcal{L}} \mathcal{L}_i(\bar{\mathbf{x}}, \boldsymbol{\Upsilon}_i, \mathbf{w}_i, \theta_i) = 0 \quad (4.10a)$$

$$\partial_{\ell} \mathcal{L}_i(\bar{\mathbf{x}}, \boldsymbol{\Upsilon}_i, \mathbf{w}_i, \theta_i) = 0 \quad (4.10b)$$

$$\partial_{\mathbf{u}} \mathcal{L}_i(\bar{\mathbf{x}}, \boldsymbol{\Upsilon}_i, \mathbf{w}_i, \theta_i) = 0 \quad (4.10c)$$

$$\partial_{\boldsymbol{\Sigma}} \mathcal{L}_i(\bar{\mathbf{x}}, \boldsymbol{\Upsilon}_i, \mathbf{w}_i, \theta_i)(\mathbf{T} - \bar{\boldsymbol{\Sigma}}) \geq 0 \quad \text{for all } \mathbf{T} \in S_\infty^2 \cap Z_i \quad (4.10d)$$

$$\partial_{\lambda} \mathcal{L}_i(\bar{\mathbf{x}}, \boldsymbol{\Upsilon}_i, \mathbf{w}_i, \theta_i)(\xi - \bar{\lambda}) \geq 0 \quad \text{for all } \xi \in M_i \quad (4.10e)$$

with $i = 1, 2$. It remains to verify that (4.10) implies (4.7)–(4.9) and that the dual variables are unique as claimed. This is done in the following four steps.

Step (1): We begin by verifying (4.8) and (4.7b).

The evaluation (4.10a) and (4.10b) gives

$$\partial_{\mathcal{L}} \tilde{J}(\bar{\mathbf{u}}, \bar{\ell}, \bar{\mathcal{L}}) - \boldsymbol{\Upsilon}_i = \mathbf{0} \quad \text{and} \quad \partial_{\ell} \tilde{J}(\bar{\mathbf{u}}, \bar{\ell}, \bar{\mathcal{L}}) - \mathbf{w}_i = \mathbf{0}$$

for $i = 1, 2$. Since $(\bar{\ell}, \bar{\mathcal{L}})$ is fixed, this yields $\boldsymbol{\Upsilon}_1 = \boldsymbol{\Upsilon}_2 =: \boldsymbol{\Upsilon}$ and $\mathbf{w}_1 = \mathbf{w}_2 =: \mathbf{w}$, which proves (4.8a) and (4.8b) and the uniqueness of $\boldsymbol{\Upsilon}$ and \mathbf{w} . Furthermore, (4.7b) immediately follows from (4.10c).

Step (2): Next we confirm (4.9a), (4.9b) and the second part of (4.9d).

We set $\mu := \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Upsilon}$ to fulfill (4.9a). Note that $\mu \in L^2(\Omega)$ holds since $\bar{\boldsymbol{\Sigma}} \in S_\infty^2$. Now let us use (4.10e) with $i = 2$, which is equivalent to

$$\int_{\Omega} \mu (\xi - \bar{\lambda}) \, dx = \int_{\Omega} (\mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Upsilon})(\xi - \bar{\lambda}) \, dx \geq 0 \quad \text{for all } \xi \in M_2. \quad (4.11)$$

By construction of M_2 , we have $0 \in M_2$ and $2\bar{\lambda} \in M_2$. Inserting these as test functions into (4.11) yields $(\mu, \bar{\lambda})_{\Omega} = 0$ so that (4.11) results in

$$\int_{\Omega} \mu \xi \, dx \geq 0 \quad \text{for all } \xi \in M_2. \quad (4.12)$$

To evaluate this inequality pointwise, let E be an arbitrary measurable subset of $\bar{\mathcal{A}}$ and choose $\xi = \chi_E$ as test function in (4.12), where χ_E denotes the characteristic function of E . This is clearly feasible since $\chi_E(x) = 0$ a.e. in \mathcal{I} . Then we obtain $\int_E \mu \, dx \geq 0$ for all $E \subset \bar{\mathcal{A}}$, giving in turn $\mu(x) \geq 0$ a.e. in $\bar{\mathcal{A}}$ and thus $\mu(x) \geq 0$ a.e. in $\bar{\mathcal{B}}$ as claimed in (4.9d). Moreover, since $\bar{\lambda} = 0$ in $\bar{\mathcal{I}}$, $\mu \geq 0$ in $\bar{\mathcal{A}}$, and $\bar{\lambda} \geq 0$, the equation $(\bar{\lambda}, \mu)_{\Omega} = 0$ implies $\mu(x) \bar{\lambda}(x) = 0$ a.e. in Ω , which is (4.9b).

Step (3): We proceed to prove (4.7a) and (4.9c).

Next let us consider (4.10d) with $i = 2$ which reads (in view of $\tilde{\mathcal{A}}_2 = \bar{\mathcal{A}}$)

$$\begin{aligned} & \langle A\Upsilon, \mathbf{T} - \bar{\Sigma} \rangle + \langle B^* \mathbf{w}, \mathbf{T} - \bar{\Sigma} \rangle + (\bar{\lambda}, \mathcal{D}\Upsilon : (\mathcal{D}\mathbf{T} - \mathcal{D}\bar{\Sigma}))_{\Omega} \\ & + \langle \phi'(\bar{\Sigma})(\mathbf{T} - \bar{\Sigma}), \theta_2 \rangle_{L^\infty(\bar{\mathcal{A}}), L^\infty(\bar{\mathcal{A}})'} \geq 0 \quad \text{for all } \mathbf{T} \in S_\infty^2 \cap Z_2. \end{aligned} \quad (4.13)$$

Now let $\varphi \in L^\infty(\bar{\mathcal{A}})$ be arbitrary and choose

$$\mathbf{T}(x) = \begin{cases} \varphi(x) \mathcal{D}^* \mathcal{D} \bar{\Sigma}(x) + \bar{\Sigma}(x), & x \in \bar{\mathcal{A}} \\ \bar{\Sigma}(x), & x \in \bar{\mathcal{I}} \end{cases}$$

as test function in (4.13) which belongs to Z_2 due to the feasibility of $\bar{\Sigma}$. Since $\varphi \in L^\infty(\bar{\mathcal{A}})$ is arbitrary, one obtains

$$\begin{aligned} & \int_{\bar{\mathcal{A}}} \varphi A\Upsilon : \mathcal{D}^* \mathcal{D} \bar{\Sigma} \, dx + \int_{\bar{\mathcal{A}}} \varphi (B^* \mathbf{w}) : \mathcal{D}^* \mathcal{D} \bar{\Sigma} \, dx + 2 \int_{\bar{\mathcal{A}}} \varphi \bar{\lambda} \mathcal{D}\Upsilon : \mathcal{D} \bar{\Sigma} \, dx \\ & + 2 \langle \varphi \mathcal{D} \bar{\Sigma} : \mathcal{D} \bar{\Sigma}, \theta_2 \rangle_{L^\infty(\bar{\mathcal{A}}), L^\infty(\bar{\mathcal{A}})'} = 0 \quad \text{for all } \varphi \in L^\infty(\bar{\mathcal{A}}) \end{aligned} \quad (4.14)$$

where we used $\mathcal{D}\mathcal{D}^*\mathcal{D} = 2\mathcal{D}$. Thanks to (4.9a), (4.9b), and $|\mathcal{D}\bar{\Sigma}|^2 = \tilde{\sigma}_0^2$ on $\bar{\mathcal{A}}$, (4.14) results in

$$\langle \varphi, \theta_2 \rangle = -\frac{1}{2\tilde{\sigma}_0^2} \int_{\bar{\mathcal{A}}} \varphi (A\Upsilon + B^* \mathbf{w}) : \mathcal{D}^* \mathcal{D} \bar{\Sigma} \, dx \quad \text{for all } \varphi \in L^\infty(\bar{\mathcal{A}}). \quad (4.15)$$

Due to $\bar{\Sigma} \in S_\infty^2$, $\Upsilon \in S^2$, and $\mathbf{w} \in V$, we have $(A\Upsilon + B^* \mathbf{w}) : \mathcal{D}^* \mathcal{D} \bar{\Sigma} \in L^2(\bar{\mathcal{A}})$. This implies

$$|\langle \varphi, \theta_2 \rangle| \leq \frac{1}{2\tilde{\sigma}_0} \|\varphi\|_{L^2(\Omega)} \|(A\Upsilon + B^* \mathbf{w}) : \mathcal{D}^* \mathcal{D} \bar{\Sigma}\|_{L^2(\Omega)} \quad \text{for all } \varphi \in L^\infty(\bar{\mathcal{A}}).$$

Since $L^\infty(\bar{\mathcal{A}})$ is dense in $L^2(\bar{\mathcal{A}})$ this implies $\theta_2 \in L^2(\bar{\mathcal{A}})$. We define

$$\theta(x) := \begin{cases} \theta_2(x), & x \in \bar{\mathcal{A}} \\ 0, & x \notin \bar{\mathcal{A}}. \end{cases}$$

Then, due to $\theta \in L^2(\Omega)$ and due to the density of S_∞^2 in S^2 , (4.13) implies

$$\begin{aligned} & \langle A\Upsilon, \mathbf{T} - \bar{\Sigma} \rangle + \langle B^* \mathbf{w}, \mathbf{T} - \bar{\Sigma} \rangle + (\bar{\lambda}, \mathcal{D}\Upsilon : (\mathcal{D}\mathbf{T} - \mathcal{D}\bar{\Sigma}))_{\Omega} \\ & + (\theta, \mathcal{D}\bar{\Sigma} : (\mathcal{D}\mathbf{T} - \mathcal{D}\bar{\Sigma}))_{\Omega} \geq 0 \quad \text{for all } \mathbf{T} \in Z_2. \end{aligned} \quad (4.16)$$

Since Z_2 does not involve a condition on the set $\bar{\mathcal{A}}$, the above inequality readily yields

$$A\Upsilon + B^* \mathbf{w} + \bar{\lambda} \mathcal{D}^* \mathcal{D} \Upsilon + \theta \mathcal{D}^* \mathcal{D} \bar{\Sigma} = \mathbf{0} \quad \text{a.e. in } \bar{\mathcal{A}}. \quad (4.17)$$

To derive a pointwise version of (4.16) on the inactive set $\bar{\mathcal{I}}$, we observe that almost every $x_0 \in \bar{\mathcal{I}}$ is a common Lebesgue point of $A\Upsilon + B^* \mathbf{w} + \bar{\lambda} \mathcal{D}^* \mathcal{D} \Upsilon + \theta \mathcal{D}^* \mathcal{D} \bar{\Sigma}$ and $(A\Upsilon + B^* \mathbf{w} + \bar{\lambda} \mathcal{D}^* \mathcal{D} \Upsilon + \theta \mathcal{D}^* \mathcal{D} \bar{\Sigma}) : \bar{\Sigma}$. Fix any such x_0 and $\mathbf{T} \in \mathbb{S}^2$ with $\phi(\mathbf{T}) \leq 0$, and define

$$\mathbf{T}_r(x) = \begin{cases} \mathbf{T}, & x \in B_r(x_0) \\ \bar{\Sigma}(x), & \text{otherwise,} \end{cases}$$

where $r > 0$ is sufficient small such that $B_r(x_0) \subset \Omega$. Note that $\mathbf{T}_r \in Z_2$ so that it is feasible for (4.16). Inserting this test function into (4.16) and taking the limit $r \searrow 0$ gives the pointwise form of (4.16) on $\bar{\mathcal{I}}$, i.e.

$$\begin{aligned} & (A\Upsilon + B^* \mathbf{w})(x) : (\mathbf{T} - \bar{\Sigma}(x)) \\ & + (\bar{\lambda} \mathcal{D}\Upsilon + \theta \mathcal{D}\bar{\Sigma})(x) : (\mathcal{D}\mathbf{T} - \mathcal{D}\bar{\Sigma}(x)) \geq 0 \quad \text{a.e. in } \bar{\mathcal{I}} \end{aligned} \quad (4.18)$$

for all $\mathbf{T} \in \mathbb{S}^2$ satisfying $\phi(\mathbf{T}) \leq 0$. For almost all $x \in \bar{\mathcal{I}}$, we have $\phi(\bar{\boldsymbol{\Sigma}}(x)) < 0$. Therefore, for $\rho > 0$ sufficiently small (depending on x), there holds

$$\phi(\bar{\boldsymbol{\Sigma}}(x) + \mathbf{T}) \leq 0 \quad \text{for all } \mathbf{T} \in \mathbb{S}^2 \text{ with } |\mathbf{T}| \leq \rho.$$

Thus, by (4.18), one deduces

$$(A\boldsymbol{\Upsilon} + B^*\mathbf{w})(x) : \mathbf{T} + (\bar{\lambda} \mathcal{D}\boldsymbol{\Upsilon} + \theta \mathcal{D}\bar{\boldsymbol{\Sigma}})(x) : \mathcal{D}\mathbf{T} \geq 0 \quad \text{for all } \mathbf{T} \in \mathbb{S}^2 \text{ with } |\mathbf{T}| \leq \rho.$$

for almost all $x \in \bar{\mathcal{I}}$. Together with (4.17) this implies

$$A\boldsymbol{\Upsilon} + B^*\mathbf{w} + \bar{\lambda} \mathcal{D}^* \mathcal{D}\boldsymbol{\Upsilon} + \theta \mathcal{D}^* \mathcal{D}\bar{\boldsymbol{\Sigma}} = \mathbf{0} \quad \text{a.e. in } \Omega,$$

i.e. (4.7a).

Since we extended θ on $\bar{\mathcal{I}}$ by zero and $\phi(\bar{\boldsymbol{\Sigma}}(x)) = 0$ holds a.e. in $\bar{\mathcal{A}}$, we obtain $\theta(x)\phi(\bar{\boldsymbol{\Sigma}}(x)) = 0$ a.e. in Ω , which coincides with (4.9c). We observe that θ is uniquely defined since it is determined by (4.7a) on $\bar{\mathcal{A}}$ and necessarily $\theta = 0$ on $\bar{\mathcal{I}}$ by (4.9c).

Step (4): It remains to prove the sign condition for θ in (4.9d).

To this end, consider (4.10d) with $i = 1$, i.e. (using $\tilde{\mathcal{A}}_1 = \bar{\mathcal{A}}_s$)

$$\begin{aligned} \langle A\boldsymbol{\Upsilon}, \mathbf{T} - \bar{\boldsymbol{\Sigma}} \rangle + \langle B^*\mathbf{w}, \mathbf{T} - \bar{\boldsymbol{\Sigma}} \rangle + (\bar{\lambda} \mathcal{D}\boldsymbol{\Upsilon} : (\mathcal{D}\mathbf{T} - \mathcal{D}\bar{\boldsymbol{\Sigma}}))_{\Omega} \\ + \langle \phi'(\bar{\boldsymbol{\Sigma}})(\mathbf{T} - \bar{\boldsymbol{\Sigma}}), \theta_1 \rangle_{L^\infty(\bar{\mathcal{A}}_s), L^\infty(\bar{\mathcal{A}}_s)'} \geq 0 \quad \text{for all } \mathbf{T} \in S_\infty^2 \cap Z_1. \end{aligned} \quad (4.19)$$

Now let $\varphi \in L^\infty(\bar{\mathcal{B}})$ be arbitrary with $\varphi(x) \in [0, 1]$ a.e. in $\bar{\mathcal{B}}$ and test (4.19) with

$$\mathbf{T}(x) = \begin{cases} \bar{\boldsymbol{\Sigma}}(x), & x \in \bar{\mathcal{I}} \cup \bar{\mathcal{A}}_s \\ (1 - \varphi(x))\bar{\boldsymbol{\Sigma}}(x), & x \in \bar{\mathcal{B}}. \end{cases}$$

Note that this test function is feasible since \mathbf{T} is a convex combination on $\bar{\mathcal{B}}$ of the two functions $\mathbf{0}$ and $\bar{\boldsymbol{\Sigma}}$ which belong to $S_\infty^2 \cap Z_1$. In this way, we obtain

$$0 \stackrel{(4.19)}{\leq} - \int_{\bar{\mathcal{B}}} \varphi (A\boldsymbol{\Upsilon} : \bar{\boldsymbol{\Sigma}} + (B^*\mathbf{w}) : \bar{\boldsymbol{\Sigma}} + \bar{\lambda} \mathcal{D}\boldsymbol{\Upsilon} : \mathcal{D}\bar{\boldsymbol{\Sigma}}) dx \stackrel{(4.7a)}{=} \int_{\bar{\mathcal{B}}} \varphi \theta \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\bar{\boldsymbol{\Sigma}} dx.$$

Due to $|\mathcal{D}\bar{\boldsymbol{\Sigma}}| = \bar{\sigma}_0$ on $\bar{\mathcal{B}}$, we arrive at

$$\int_{\bar{\mathcal{B}}} \varphi \theta dx \geq 0 \quad \text{for all } \varphi \in L^\infty(\bar{\mathcal{B}}) \text{ satisfying } \varphi \in [0, 1] \text{ a.e. in } \bar{\mathcal{B}}, \quad (4.20)$$

which proves the nonnegativity of θ on $\bar{\mathcal{B}}$. \square

4.2. Discussion of Strong Stationarity. In the following we reformulate the strong stationarity conditions (4.6)–(4.9) in order to allow a comparison to the optimality conditions given in Mignot and Puel [1984] for optimal control of the obstacle problem.

Proposition 4.7. *The strong stationarity system (4.6)–(4.9) is equivalent to the following set of conditions*

$$A\bar{\Sigma} + \bar{\lambda}\mathcal{D}^*\mathcal{D}\bar{\Sigma} + B^*\bar{\mathbf{u}} = \bar{\mathcal{L}} \quad (4.21a)$$

$$B\bar{\Sigma} = \bar{\ell} \quad (4.21b)$$

$$0 \leq \bar{\lambda} \perp \phi(\bar{\Sigma}) \leq 0 \quad \text{a.e. in } \Omega \quad (4.21c)$$

$$\langle A\Upsilon, \mathbf{T} \rangle + \langle B^*\mathbf{w}, \mathbf{T} \rangle + (\bar{\lambda}, \mathcal{D}\Upsilon : \mathcal{D}\mathbf{T})_\Omega \geq 0 \quad \text{for all } \mathbf{T} \in \bar{\mathcal{S}} \quad (4.22a)$$

$$B\Upsilon = -\partial_{\bar{\mathbf{u}}}\tilde{J}(\bar{\mathbf{u}}, \bar{\ell}, \bar{\mathcal{L}}) \quad (4.22b)$$

$$-\Upsilon \in \bar{\mathcal{S}} \quad (4.22c)$$

$$\partial_{\bar{\mathcal{L}}}\tilde{J}(\bar{\mathbf{u}}, \bar{\ell}, \bar{\mathcal{L}}) - \Upsilon = \mathbf{0} \quad (4.23a)$$

$$\partial_{\bar{\ell}}\tilde{J}(\bar{\mathbf{u}}, \bar{\ell}, \bar{\mathcal{L}}) - \mathbf{w} = \mathbf{0}, \quad (4.23b)$$

where $\bar{\mathcal{S}}$ is defined similarly to (3.3) by

$$\begin{aligned} \bar{\mathcal{S}} := \{ \mathbf{T} \in S^2 : \sqrt{\bar{\lambda}}\mathcal{D}\mathbf{T} \in S, \quad \mathcal{D}\bar{\Sigma}(x) : \mathcal{D}\mathbf{T}(x) \leq 0 \quad \text{a.e. in } \bar{\mathcal{B}}, \\ \mathcal{D}\bar{\Sigma}(x) : \mathcal{D}\mathbf{T}(x) = 0 \quad \text{a.e. in } \bar{\mathcal{A}}_s \} \end{aligned}$$

with $\bar{\mathcal{B}}$ and $\bar{\mathcal{A}}_s$ as in (4.2).

Remark 4.8. *The above notion of strong stationarity is equivalent to the one introduced by Mignot and Puel in case of the obstacle problem, cf. [Mignot and Puel, 1984, Theorem 2.2]. We point out that the adjoint system (4.22a)–(4.22c) cannot be written in form of a variational inequality.*

Proof of Proposition 4.7. We only have to show that (4.22) is equivalent to (4.7) and (4.9).

Let us start by assuming that (4.7) and (4.9) are fulfilled. By (4.9a), (4.9b), and (4.9d) we know that $\mathcal{D}\bar{\Sigma} : \mathcal{D}\Upsilon \geq 0$ a.e. in $\bar{\mathcal{B}}$ and $\bar{\lambda}(\mathcal{D}\bar{\Sigma} : \mathcal{D}\Upsilon) = 0$ a.e. in Ω . The latter equality immediately implies $\mathcal{D}\bar{\Sigma} : \mathcal{D}\Upsilon = 0$ a.e. in $\bar{\mathcal{A}}_s$. Moreover, if we test (4.7a) with Υ and use $\theta \in L^2(\Omega)$ and $\mathcal{D}\bar{\Sigma} \in L^\infty(\Omega, \mathbb{S})$, then $\sqrt{\bar{\lambda}}\mathcal{D}\Upsilon \in S$ is obtained, giving in turn $-\Upsilon \in \bar{\mathcal{S}}$. To verify (4.22a), let $\mathbf{T} \in \bar{\mathcal{S}}$ be arbitrary. Note that (4.9c) yields $\theta = 0$ a.e. in $\bar{\mathcal{I}}$ and (4.9d) implies $\theta \geq 0$ a.e. in $\bar{\mathcal{B}}$. Using the sign conditions on $\mathcal{D}\bar{\Sigma} : \mathcal{D}\mathbf{T}$ implied by $\mathbf{T} \in \bar{\mathcal{S}}$, we get

$$(\theta, \mathcal{D}\bar{\Sigma} : \mathcal{D}\mathbf{T})_\Omega \leq 0.$$

Inserting this into (4.7a) results in (4.22a) so that (4.22) is indeed verified.

The opposite direction is shown in three steps. To this end assume that $(\Upsilon, \mathbf{w}) \in S^2 \times V$ satisfies (4.22).

Step (1): We confirm the conditions on μ in (4.9a), (4.9b), and (4.9d).

First let us define μ according to (4.9a) by $\mu := \mathcal{D}\bar{\Sigma} : \mathcal{D}\Upsilon$. Due to $\mathcal{D}\bar{\Sigma} \in L^\infty(\Omega, \mathbb{S})$ in virtue of $\phi(\bar{\Sigma}) \leq 0$, we have $\mu \in L^2(\Omega)$. Moreover, because of $-\Upsilon \in \bar{\mathcal{S}}$ by (4.22c), one obtains $\mu = 0$ a.e. in $\bar{\mathcal{A}}_s = \{x \in \Omega : \bar{\lambda}(x) \neq 0\}$ so that $\mu\bar{\lambda} = 0$, i.e. (4.9b) is satisfied. The sign condition on μ on the biactive set $\bar{\mathcal{B}}$ finally follows directly from $-\Upsilon \in \bar{\mathcal{S}}$.

Step (2): We verify the first adjoint equation (4.7a).

Let us start by defining

$$\mathbf{M} := A\Upsilon + B^*\mathbf{w} + \bar{\lambda}\mathcal{D}^*\mathcal{D}\Upsilon$$

so that

$$\int_{\Omega} \mathbf{M} : \mathbf{T} \, dx \geq 0 \quad \text{for all } \mathbf{T} \in \bar{\mathcal{S}}. \quad (4.24)$$

due to (4.22a). Let $x \in \Omega$ be arbitrary and define $\theta(x) \in \mathbb{R}$ and $\mathbf{M}_0(x) \in \mathbb{S}^2$ with $\mathbf{M}_0(x) : \mathcal{D}^* \mathcal{D} \bar{\boldsymbol{\Sigma}}(x) = 0$, i.e. $\mathbf{M}_0(x)$ is in the orthogonal complement of $\text{span}(\mathcal{D}^* \mathcal{D} \bar{\boldsymbol{\Sigma}}(x))$, by

$$\mathbf{M}(x) = -\theta(x) \mathcal{D}^* \mathcal{D} \bar{\boldsymbol{\Sigma}}(x) + \mathbf{M}_0(x). \quad (4.25)$$

Since $\bar{\mathcal{S}}$ does not involve any condition on the inactive set $\bar{\mathcal{I}}$ (apart from the regularity condition $\sqrt{\bar{\lambda}} \mathcal{D} \mathbf{T} \in \mathcal{S}$), we are allowed to insert $\chi_{\bar{\mathcal{I}}} \mathbf{T}$ with arbitrary $\mathbf{T} \in L^\infty(\Omega, \mathbb{S})^2$ as test function, giving in turn $\mathbf{M} = \mathbf{0}$ a.e. in $\bar{\mathcal{I}}$. In particular, this implies $\theta(x) = 0$ and $\mathbf{M}_0(x) = \mathbf{0}$ for almost all $x \in \bar{\mathcal{I}}$. Next we investigate the regularity of θ . By multiplying (4.25) with $\mathcal{D}^* \mathcal{D} \bar{\boldsymbol{\Sigma}}(x)$ and taking $\phi(\bar{\boldsymbol{\Sigma}}) = 0$ a.e. in $\bar{\mathcal{A}}$ into account, we obtain f.a.a. $x \in \bar{\mathcal{A}}$

$$\begin{aligned} -2 \tilde{\sigma}_0 \theta(x) &= -\theta(x) \mathcal{D}^* \mathcal{D} \bar{\boldsymbol{\Sigma}}(x) : \mathcal{D}^* \mathcal{D} \bar{\boldsymbol{\Sigma}}(x) = \mathbf{M}(x) : \mathcal{D}^* \mathcal{D} \bar{\boldsymbol{\Sigma}}(x) \\ &= (\mathbf{A} \boldsymbol{\Upsilon} + \mathbf{B}^* \mathbf{w})(x) : \mathcal{D}^* \mathcal{D} \bar{\boldsymbol{\Sigma}}(x) + \bar{\lambda}(x) \mathcal{D}^* \mathcal{D} \boldsymbol{\Upsilon}(x) : \mathcal{D}^* \mathcal{D} \bar{\boldsymbol{\Sigma}}(x) \\ &= (\mathbf{A} \boldsymbol{\Upsilon} + \mathbf{B}^* \mathbf{w})(x) : \mathcal{D}^* \mathcal{D} \bar{\boldsymbol{\Sigma}}(x) + 2 \bar{\lambda}(x) \mathcal{D} \boldsymbol{\Upsilon}(x) : \mathcal{D} \bar{\boldsymbol{\Sigma}}(x). \end{aligned}$$

Because of (4.22c) we have $\mathcal{D} \boldsymbol{\Upsilon}(x) : \mathcal{D} \bar{\boldsymbol{\Sigma}}(x) = 0$ a.e. in $\bar{\mathcal{A}}_s = \{x \in \Omega : \bar{\lambda}(x) \neq 0\}$ and thus

$$\theta(x) = \frac{-1}{2 \tilde{\sigma}_0} (\mathbf{A} \boldsymbol{\Upsilon} + \mathbf{B}^* \mathbf{w})(x) : \mathcal{D}^* \mathcal{D} \bar{\boldsymbol{\Sigma}}(x) \in L^2(\bar{\mathcal{A}})$$

due to $\mathcal{D}^* \mathcal{D} \bar{\boldsymbol{\Sigma}} \in L^\infty(\Omega, \mathbb{S})$. The sign conditions in $\bar{\mathcal{S}}$ are satisfied by definition of \mathbf{M}_0 . We define $\Lambda_s = \{x \in \bar{\mathcal{A}} : \bar{\lambda} \leq s\}$ for $s > 0$. Due to $\theta \mathcal{D}^* \mathcal{D} \bar{\boldsymbol{\Sigma}} \in \mathcal{S}^2$, (4.25), the definition of \mathbf{M} and $\bar{\lambda} \leq s$ on Λ_s , we have $-\chi_{\Lambda_s} \mathbf{M}_0 \in \bar{\mathcal{S}}$. Therefore, (4.24) together with (4.25) implies $\mathbf{M}_0 = \mathbf{0}$ on Λ_s . Since $s > 0$ was arbitrary, $\mathbf{M}_0 = \mathbf{0}$ a.e. in Ω . Now we extend θ on $\bar{\mathcal{I}}$ by zero. Therefore, $\theta \in L^2(\Omega)$ holds. Using $\mathbf{M} = \mathbf{0}$ a.e. on $\bar{\mathcal{I}}$, $\mathbf{M}_0 = \mathbf{0}$ and (4.25) we obtain

$$-\theta \mathcal{D}^* \mathcal{D} \bar{\boldsymbol{\Sigma}} = \mathbf{M} = \mathbf{A} \boldsymbol{\Upsilon} + \mathbf{B}^* \mathbf{w} + \bar{\lambda} \mathcal{D}^* \mathcal{D} \boldsymbol{\Upsilon},$$

which is (4.7a).

Step (3): It remains to prove the complementarity relation (4.9c) and the sign condition on θ in (4.9d).

Since $\theta = 0$ in $\bar{\mathcal{I}} = \{x \in \Omega : \phi(\bar{\boldsymbol{\Sigma}}(x)) \neq 0\}$ by construction, $\theta \phi(\bar{\boldsymbol{\Sigma}}) = 0$ a.e. in Ω , i.e. (4.9c) is trivially fulfilled. To verify the sign condition on the biactive set, let $E \subset \bar{\mathcal{B}}$ be some measurable subset. If we insert $-\chi_E \bar{\boldsymbol{\Sigma}} \in \bar{\mathcal{S}}$ as test function in (4.24), then, in view of $\mathbf{M} = -\theta \mathcal{D}^* \mathcal{D} \bar{\boldsymbol{\Sigma}}$ on $\bar{\mathcal{B}}$, we obtain

$$0 \leq \int_E \theta \mathcal{D} \bar{\boldsymbol{\Sigma}} : \mathcal{D} \bar{\boldsymbol{\Sigma}} \, dx = \tilde{\sigma}_0 \int_E \theta \, dx.$$

Since $E \subset \bar{\mathcal{B}}$ was arbitrary, we have $\theta \geq 0$ on $\bar{\mathcal{B}}$. \square

4.3. Strong Stationarity implies B-Stationarity. In this subsection, we state strong stationarity conditions for the original problem (P) (without proving their necessity) and show that they imply the B-stationarity conditions from Theorem 3.10. In view of (4.6)–(4.9), the strong stationarity conditions for the original problem are defined as follows.

Definition 4.9. We say that an optimal control $\bar{\mathbf{g}} \in U_{\text{ad}}$ with associated state $(\bar{\boldsymbol{\Sigma}}, \bar{\mathbf{u}}, \bar{\lambda}) \in S^2 \times V \times L^2(\Omega)$ satisfies the strong stationarity condition for **(P)** if an adjoint state $(\boldsymbol{\Upsilon}, \mathbf{w}) \in S^2 \times V$ and Lagrange multipliers $\mu, \theta \in L^2(\Omega)$ exist such that

$$A\bar{\boldsymbol{\Sigma}} + \bar{\lambda} \mathcal{D}^* \mathcal{D}\bar{\boldsymbol{\Sigma}} + B^* \bar{\mathbf{u}} = \mathbf{0} \quad (4.26a)$$

$$B\bar{\boldsymbol{\Sigma}} = -\tau_N^* \bar{\mathbf{g}} \quad (4.26b)$$

$$0 \leq \bar{\lambda} \perp \phi(\bar{\boldsymbol{\Sigma}}) \leq 0 \quad \text{a.e. in } \Omega \quad (4.26c)$$

$$A\boldsymbol{\Upsilon} + B^* \mathbf{w} + \bar{\lambda} \mathcal{D}^* \mathcal{D}\boldsymbol{\Upsilon} + \theta \mathcal{D}^* \mathcal{D}\bar{\boldsymbol{\Sigma}} = \mathbf{0} \quad (4.27a)$$

$$B\boldsymbol{\Upsilon} = -\partial_{\mathbf{u}} J(\bar{\mathbf{u}}, \bar{\mathbf{g}}) \quad (4.27b)$$

$$\langle \partial_{\mathbf{g}} J(\bar{\mathbf{u}}, \bar{\mathbf{g}}), \mathbf{g} - \bar{\mathbf{g}} \rangle + \int_{\Gamma_N} \mathbf{w} \cdot (\mathbf{g} - \bar{\mathbf{g}}) \, ds \geq 0 \quad \text{for all } \mathbf{g} \in U_{\text{ad}} \quad (4.28a)$$

$$\mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Upsilon} - \mu = 0 \quad (4.29a)$$

$$\mu \bar{\lambda} = 0 \quad \text{a.e. in } \Omega \quad (4.29b)$$

$$\theta \phi(\bar{\boldsymbol{\Sigma}}) = 0 \quad \text{a.e. in } \Omega \quad (4.29c)$$

$$\theta \geq 0, \quad \mu \geq 0 \quad \text{a.e. in } \bar{\mathcal{B}} \quad (4.29d)$$

holds true.

Remark 4.10. As already mentioned in [Remark 4.4](#), we cannot prove that [\(4.26\)](#)–[\(4.29\)](#) are necessary for the local optimality of $\bar{\mathbf{g}}$. The reason is that a verification of the regular point condition for the auxiliary problems $(\tilde{\mathbf{P}}_1)$ and $(\tilde{\mathbf{P}}_2)$ does not seem to be possible in case of **(P)**. Note that a regularization technique would yield C -stationarity conditions that coincide with [\(4.26\)](#)–[\(4.29\)](#) except that [\(4.29d\)](#) has to be replaced by $\theta \mu \geq 0$ in $\bar{\mathcal{B}}$, cf. [[Herzog et al., 2012, Section 3.3](#)].

The following proposition shows that strong stationarity implies B-stationarity.

Proposition 4.11. Assume that $\bar{\mathbf{g}} \in U_{\text{ad}}$ with associated state $(\bar{\boldsymbol{\Sigma}}, \bar{\mathbf{u}}, \bar{\lambda}) \in S^2 \times V \times L^2(\Omega)$ fulfills the strong stationarity condition [\(4.26\)](#)–[\(4.29\)](#). Then $\bar{\mathbf{g}}$ satisfies the B-stationarity condition [\(3.28\)](#), i.e. the variational inequality

$$J'(\bar{\mathbf{u}}, \bar{\mathbf{g}})(\mathbf{u}', \mathbf{g} - \bar{\mathbf{g}}) \geq 0 \quad \text{for all } \mathbf{g} \in U_{\text{ad}},$$

where $(\boldsymbol{\Sigma}', \mathbf{u}')$ solves the derivative problem [\(3.2\)](#) with $\delta \ell := -\tau_N^*(\mathbf{g} - \bar{\mathbf{g}})$ as right-hand side.

Proof. According to [Proposition 3.13](#) the variational inequality for $(\boldsymbol{\Sigma}', \mathbf{u}')$ can equivalently be expressed in terms of [\(3.32\)](#) (with $(\bar{\boldsymbol{\Sigma}}, \bar{\mathbf{u}})$ instead of $(\boldsymbol{\Sigma}, \mathbf{u})$), which involves a multiplier $\lambda' \in L^2(\Omega)$. If we test [\(3.32b\)](#) with \mathbf{w} and set $\delta \ell = -\tau_N^*(\mathbf{g} - \bar{\mathbf{g}})$, we arrive at

$$\begin{aligned} (\mathbf{w}, \mathbf{g} - \bar{\mathbf{g}})_{\Gamma_N} &= -\langle B^* \mathbf{w}, \boldsymbol{\Sigma}' \rangle \\ &= \langle A\boldsymbol{\Upsilon}, \boldsymbol{\Sigma}' \rangle + (\bar{\lambda}, \mathcal{D}\boldsymbol{\Upsilon} : \mathcal{D}\boldsymbol{\Sigma}')_{\Omega} + (\theta, \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Sigma}')_{\Omega} \quad \text{by (4.27a)} \\ &= -\langle B^* \mathbf{u}', \boldsymbol{\Upsilon} \rangle - (\lambda', \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Upsilon})_{\Omega} + (\theta, \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Sigma}')_{\Omega} \quad \text{by (3.32a)} \\ &= \langle \partial_{\mathbf{u}} J(\bar{\mathbf{u}}, \bar{\mathbf{g}}), \mathbf{u}' \rangle - (\lambda', \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Upsilon})_{\Omega} + (\theta, \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Sigma}')_{\Omega} \quad \text{by (4.27b)}. \end{aligned}$$

For the last two addends in the previous equation, we obtain the following sign conditions:

$$\theta \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Sigma}' \begin{cases} = 0 \text{ in } \bar{\mathcal{I}}, & \text{since } \theta = 0 \text{ a.e. in } \bar{\mathcal{I}} \text{ by (4.29c)} \\ \leq 0 \text{ in } \bar{\mathcal{B}}, & \text{since } \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Sigma}' \leq 0 \text{ by (3.32d)} \\ & \text{and } \theta \geq 0 \text{ a.e. in } \bar{\mathcal{B}} \text{ by (4.29d)} \\ = 0 \text{ in } \bar{\mathcal{A}}_s, & \text{since } \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Sigma}' = 0 \text{ a.e. in } \bar{\mathcal{A}}_s \text{ by (3.32c)} \end{cases}$$

and

$$\lambda' \mathcal{D}\bar{\Sigma} : \mathcal{D}\Upsilon \begin{cases} = 0 \text{ in } \bar{\mathcal{I}}, & \text{since } \lambda' = 0 \text{ a.e. in } \bar{\mathcal{I}} \text{ by (3.32e)} \\ \geq 0 \text{ in } \bar{\mathcal{B}}, & \text{since } \mathcal{D}\bar{\Sigma} : \mathcal{D}\Upsilon = \mu \geq 0 \text{ by (4.29a) and (4.29d)} \\ & \text{and } \lambda' \geq 0 \text{ a.e. in } \bar{\mathcal{B}} \text{ by (3.32d)} \\ = 0 \text{ in } \bar{\mathcal{A}}_s, & \text{since } \mathcal{D}\bar{\Sigma} : \mathcal{D}\Upsilon = \mu = 0 \text{ a.e. in } \bar{\mathcal{A}}_s \text{ by (4.29b)}. \end{cases}$$

These sign conditions imply

$$(\mathbf{w}, \mathbf{g} - \bar{\mathbf{g}})_{\Gamma_N} \leq \langle \partial_{\mathbf{u}} J(\bar{\mathbf{u}}, \bar{\mathbf{g}}), \mathbf{u}' \rangle.$$

Inserting this into (4.28a) yields the desired result. \square

A Auxiliary results

In this section, $|B|$ denotes the Lebesgue measure of a set $B \subset \Omega$.

Lemma A.1. *Let $M \subset \Omega$ be some measurable set and let $b \in L^1(M)$ with $b > 0$ a.e. in M be given. Then, for all $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$\int_B b \, dx \geq \delta \quad \text{for all } B \subset M \text{ s.t. } |B| \geq \varepsilon$$

holds.

Proof. Let us define $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, $g(\gamma) := |\{x \in M : b \leq \gamma\}|$. We obtain that g is monotone increasing and $g(0) = 0$. Therefore $\lim_{\gamma \searrow 0} g(\gamma)$ exists and we have

$$\begin{aligned} \lim_{\gamma \searrow 0} g(\gamma) &= \lim_{n \rightarrow \infty} g(2^{-n}) = \lim_{n \rightarrow \infty} |\{b \leq 2^{-n}\}| \\ &= \left| \bigcap_{n=1}^{\infty} \{b \leq 2^{-n}\} \right| = |\{b = 0\}| = 0. \end{aligned}$$

This shows that g is continuous at 0. Hence there is $\delta_0 > 0$ with $g(\delta_0) \leq \varepsilon/2$. For $G := \{b \leq \delta_0\}$ we obtain $|G| \leq \varepsilon/2$. Let $B \subset M$ with $|B| \geq \varepsilon$ be arbitrary. We have

$$\int_B b \, dx \geq \int_{B \setminus G} b \, dx \geq \int_{B \setminus G} \delta_0 \, dx = \delta_0 |B \setminus G| \geq \delta_0 \varepsilon/2.$$

With $\delta := \delta_0 \varepsilon/2$ the lemma is proved. \square

Lemma A.2. *Let $M \subset \Omega$ be measurable and $\{f_n\} \subset L^1(M)$ a sequence with $f_n \rightarrow f \in L^1(\Omega)$. If $f > 0$ a.e. in M , then $|\{x \in M : f_n = 0\}| \rightarrow 0$.*

Proof. We prove this lemma by contradiction. Let us assume that there is $\varepsilon > 0$ and a subsequence n_k such that

$$|\{f_{n_k} = 0\}| \geq \varepsilon \quad \text{for all } k \in \mathbb{N}.$$

According to Lemma A.1 there exists $\delta > 0$ with

$$\int_B f \, dx \geq \delta \quad \text{for all } B \subset M \text{ with } |B| \geq \varepsilon.$$

Now we have

$$\|f - f_{n_k}\|_{L^1(\Omega)} \geq \int_{\{f_{n_k}=0\}} |f - f_{n_k}| \, dx = \int_{\{f_{n_k}=0\}} f \, dx \geq \delta \quad \text{for all } k \in \mathbb{N}.$$

This is a contradiction to $f_n \rightarrow f$ in $L^1(\Omega)$. \square

Lemma A.3. *Let $\{f_n\} \subset L^1(\Omega)$ with $f_n \geq 0$ a.e. in Ω and $f_n \rightarrow f$ in $L^1(\Omega)$ be given. If $|\{x \in \Omega : f_n(x) > 0\}| \rightarrow 0$ as $n \rightarrow \infty$ holds, then $f \equiv 0$.*

Proof. Let us abbreviate $A_n = \{x \in \Omega : f_n(x) > 0\}$. Since $|A_n| \rightarrow 0$, there is a subsequence $\{n_k\} \subset \mathbb{N}$ with $\sum_{k=1}^{\infty} |A_{n_k}| < \infty$. For the sets $B_j := \bigcup_{k=j}^{\infty} A_{n_k}$ we obtain $|B_j| \rightarrow 0$ as $j \rightarrow \infty$. By construction, $f_{n_i} \equiv 0$ holds on $\Omega \setminus A_{n_i}$ and hence

$$\int_{\Omega \setminus B_j} f_{n_i} \, dx = 0 \quad \text{if } i > j.$$

The weak convergence $f_n \rightharpoonup f$ implies

$$0 = \int_{\Omega \setminus B_j} f_{n_i} \, dx \rightarrow \int_{\Omega \setminus B_j} f \, dx \quad \text{as } i \rightarrow \infty$$

and thus

$$\int_{\Omega \setminus B_j} f \, dx = 0 \quad \text{for all } j \in \mathbb{N}.$$

The set of nonnegative functions is weakly closed in $L^1(\Omega)$, hence $f \geq 0$ holds and we obtain $f \equiv 0$ in $\Omega \setminus B_j$ for all $j \in \mathbb{N}$. Using De Morgan's Law, $|B_j| \rightarrow 0$ yields $\bigcup_{j=1}^{\infty} \Omega \setminus B_j = \Omega \setminus \bigcap_{j=1}^{\infty} B_j = \Omega$ and thus we conclude $f \equiv 0$ on Ω . \square

Acknowledgment. This work was supported by a DFG grant within the **Priority Program SPP 1253** (*Optimization with Partial Differential Equations*), which is gratefully acknowledged.

The authors would like to thank three anonymous referees for their careful reading of the manuscript and their suggestions. We are particularly thankful for the question raised by one of the referees concerning alternative approaches to the derivation of B-stationarity conditions, which led to **Section 3.4**.

References

- V. Barbu. *Optimal Control of Variational Inequalities*, volume 100 of *Research Notes in Mathematics*. Pitman, Boston, 1984.
- M. Bergounioux. Optimal control of an obstacle problem. *Applied Mathematics and Optimization*, 36:147–172, 1997.
- M. Bergounioux. Optimal control problems governed by abstract elliptic variational inequalities with state constraints. *SIAM Journal on Control and Optimization*, 36(1):273–289, 1998.
- M. Bergounioux and H. Zidani. Pontryagin maximum principle for optimal control of variational inequalities. *SIAM Journal on Control and Optimization*, 37(4):1273–1290, 1999.
- F. Bonnans and D. Tiba. Pontryagin's principle in the control of semilinear elliptic variational inequalities. *Applied Mathematics and Optimization*, 23:299–312, 1991.
- M. H. Farshbaf-Shaker. A penalty approach to optimal control of Allen-Cahn variational inequalities: MPEC-view. *submitted*, 2011.
- M. Flegel and C. Kanzow. On the Guignard constraint qualification for mathematical programs with equilibrium constraints. *Optimization*, 54(6):517–534, 2005.
- A. Friedman. Optimal control for variational inequalities. *SIAM Journal on Control and Optimization*, 24(3):439–451, 1986.
- P. Grisvard. *Elliptic Problems in Nonsmooth Domains*. Pitman, Boston, 1985.
- W. Han and B. D. Reddy. *Plasticity*. Springer, New York, 1999.
- R. Herzog and C. Meyer. Optimal control of static plasticity with linear kinematic hardening. *Journal of Applied Mathematics and Mechanics*, 91(10):777–794, 2011. doi: [10.1002/zamm.200900378](https://doi.org/10.1002/zamm.200900378).

- R. Herzog, C. Meyer, and G. Wachsmuth. Existence and regularity of the plastic multiplier in static and quasistatic plasticity. *GAMM Reports*, 34(1):39–44, 2011. doi: [10.1002/gamm.201110006](https://doi.org/10.1002/gamm.201110006).
- R. Herzog, C. Meyer, and G. Wachsmuth. C-stationarity for optimal control of static plasticity with linear kinematic hardening. *in revision*, 2012.
- M. Hintermüller. Inverse coefficient problems for variational inequalities: Optimality conditions and numerical realization. *ESAIM Mathematical Modelling and Numerical Analysis*, 35(1):129–152, 2001.
- M. Hintermüller. An active-set equality constrained Newton solver with feasibility restoration for inverse coefficient problems in elliptic variational inequalities. *Inverse Problems*, 24(3):034017, 23, 2008.
- M. Hintermüller and I. Kopacka. Mathematical programs with complementarity constraints in function space: C- and strong stationarity and a path-following algorithm. *SIAM Journal on Optimization*, 20(2):868–902, 2009. ISSN 1052-6234. doi: [10.1137/080720681](https://doi.org/10.1137/080720681).
- M. Hintermüller, I. Kopacka, and M.H. Tber. Recent advances in the numerical solution of MPECs in function space. In *Numerical Techniques for Optimization Problems with PDE Constraints*, volume 6 of *Oberwolfach Report No. 4/2009*, pages 36–40, Zurich, 2009. European Mathematical Society Publishing House.
- K. Ito and K. Kunisch. Optimal control of elliptic variational inequalities. *Applied Mathematics and Optimization*, 41:343–364, 2000.
- K. Ito and K. Kunisch. Optimal control of parabolic variational inequalities. *Journal de Mathématiques Pures et Appliquées*, 93(4):329–360, 2010. doi: [10.1016/j.matpur.2009.10.005](https://doi.org/10.1016/j.matpur.2009.10.005).
- J. Jarušek, M. Krbeč, M. Rao, and J. Sokołowski. Conical differentiability for evolution variational inequalities. *Journal of Differential Equations*, 193:131–146, 2003.
- Jiří Jarušek and Jiří Outrata. On sharp necessary optimality conditions in control of contact problems with strings. *Nonlinear Analysis*, 67:1117–1128, 2007.
- K. Jittorntrum. Solution point differentiability without strict complementarity in nonlinear programming. *Mathematical Programming Study*, 21:127–138, 1984.
- C. Kanzow and A. Schwartz. Mathematical programs with equilibrium constraints: enhanced Fritz John-conditions, new constraint qualifications, and improved exact penalty results. *SIAM Journal on Optimization*, 20(5):2730–2753, 2010. ISSN 1052-6234. doi: [10.1137/090774975](https://doi.org/10.1137/090774975).
- M. Kočvara and J. V. Outrata. Optimization problems with equilibrium constraints and their numerical solution. *Mathematical Programming, Series B*, 101:119–149, 2004.
- Z.-Q. Luo, J.-S. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, 1996.
- F. Mignot. Contrôle dans les inéquations variationnelles elliptiques. *Journal of Functional Analysis*, 22(2):130–185, 1976.
- F. Mignot and J.-P. Puel. Optimal control in some variational inequalities. *SIAM Journal on Control and Optimization*, 22(3):466–476, 1984.
- Jiří Outrata, Jiří Jarušek, and Jana Stará. On optimality conditions in control of elliptic variational inequalities. *Set-Valued and Variational Analysis*, 19(1): 23–42, 2011. ISSN 1877-0533. doi: [10.1007/s11228-010-0158-4](https://doi.org/10.1007/s11228-010-0158-4).
- E. Sachs. Differentiability in optimization theory. *Mathematische Operationsforschung und Statistik Series Optimization*, 9(4):497–513, 1978. ISSN 0323-3898. doi: [10.1080/02331937808842516](https://doi.org/10.1080/02331937808842516).
- H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. *Mathematics of Operations*

- Research*, 25(1):1–22, 2000.
- S. Steffensen and M. Ulbrich. A new relaxation scheme for mathematical programs with equilibrium constraints. *SIAM Journal on Optimization*, 20(5):2504–2539, 2010. doi: [10.1137/090748883](https://doi.org/10.1137/090748883).
- S. W. Zhu. Optimal control of variational inequalities with delays in the highest order spatial derivatives. *Acta Mathematica Sinica, English Series*, 22(2):607–624, 2006.
- J. Zowe and S. Kurcyusz. Regularity and stability for the mathematical programming problem in Banach spaces. *Applied Mathematics and Optimization*, 5(1):49–62, 1979.

CHEMNITZ UNIVERSITY OF TECHNOLOGY, FACULTY OF MATHEMATICS, D-09107 CHEMNITZ, GERMANY

E-mail address: roland.herzog@mathematik.tu-chemnitz.de

URL: <http://www.tu-chemnitz.de/herzog>

TU DARMSTADT, GRADUATE SCHOOL CE, DOLIVOSTR. 15, D-64293 DARMSTADT, GERMANY

E-mail address: cmeyer@gsc.tu-darmstadt.ce

URL: <http://www.graduate-school-ce.de/index.php?id=128>

CHEMNITZ UNIVERSITY OF TECHNOLOGY, FACULTY OF MATHEMATICS, D-09107 CHEMNITZ, GERMANY

E-mail address: gerd.wachsmuth@mathematik.tu-chemnitz.de

URL: http://www.tu-chemnitz.de/mathematik/part_dgl/people/wachsmuth