SECOND-ORDER SUFFICIENT OPTIMALITY CONDITIONS FOR OPTIMAL CONTROL OF STATIC ELASTOPLASTICITY WITH HARDENING

THOMAS BETZ AND CHRISTIAN MEYER

ABSTRACT. The paper is concerned with the optimal control of static elasto-plasticity with linear kinematic hardening. This leads to an optimal control problem governed by an elliptic variational inequality (VI) of first kind in mixed form. Based on L^p-regularity results for the state equation, it is shown that the control-to-state operator is Bouligand differentiable. This enables to establish second-order sufficient optimality conditions by means of a Taylor expansion of a particularly chosen Lagrange function.

Keywords: Second-order sufficient conditions, optimal control of variational inequalities, Bouligand differentiability

AMS subject classification: 49K20, 74C05, 74P10, 35R45

1 Introduction

In this paper we establish second-order sufficient optimality conditions for an optimal control problem governed by an elliptic variational inequality (VI) of first kind in mixed form. This VI models the problem of static elastoplasticity with linear kinematic hardening. The precise statements of the VI and of the optimal control problem will be given at the end of the introduction. The static VI has only limited physical meaning, but can be regarded as time discretization of a corresponding quasi-static counterpart. The latter one models elastoplastic deformation processes at small strain and thus appears in various industrial applications. Therefore, if an instantaneous control strategy is applied to optimize or control such an application, then the static optimal control problem considered in this paper will arise.

Optimal control problems subject to VIs represent mathematical programs with equilibrium constraints (MPECs) in function space. Problems of this type are known to be difficult to handle, even in the finite dimensional case, cf. e.g. [18, 23] and the references therein. These difficulties are induced by the non-smoothness of the control-to-state mapping, which prevents the derivation of necessary optimality conditions in terms of the Karush-Kuhn-Tucker (KKT) conditions. Instead several alternative stationarity concepts such as e.g. Clarke(C)-, Bouligand(B)-, and strong stationarity have been introduced as necessary optimality conditions. MPECs in function space have been considered by many authors in various aspects, in particular concerning the derivation of first-order necessary optimality conditions and regularization and relaxation methods, respectively. We only refer to [1–3,15–17,21,22]. In [13,14] C- and B-stationarity conditions for optimal control of static elastoplasticity are derived. Moreover, the necessity of strong stationarity for local optimality is established in [14] for an academic problem involving an additional, physically meaningless control function.

While second-order sufficient conditions for optimal control of PDEs are well investigated, see e.g. [4–6] and the references therein, the literature on sufficient optimality conditions for optimal control of VIs is rather rare. To the best of our knowledge the only contribution in this field is a paper by Kunisch and Wachsmuth [19], where sufficient conditions for optimal control of the obstacle problem were presented. In a follow-up paper [20] these conditions are used to design an efficient path-following algorithm based on a Moreau-Yosida regularization. Sufficient optimality conditions for optimal control of VIs that are not of obstacle type, such as static elastoplasticity, have not been discussed so far. In particular the nonlinearity appearing in the constraint of the VI substantially complicates the analysis in comparison to the obstacle problem. Therefore it is not straight forward to adapt the method of proof from [19] to our case. Instead we use a different technique and argue by means of a Taylor expansion of a particularly chosen Lagrange function, see Section 4. In order to compensate for higher-order terms in the Taylor expansion we have to prove that the control-to-state mapping is Bouligand differentiable. To our best knowledge such a result has not yet been shown for VIs which are not of obstacle type, so that this finding is of its own interest. We expect that our method to verify sufficient second-order conditions can be adapted to problems of obstacle type and will yield the same results as in [19].

The outline of the paper is as follows: After fixing the notation and stating the precise problem and the standing ssumptions in the remaining part of this section, we will present some known and preliminary results in Section 2. The Bouligand differentiability is shown in Section 3. Finally, Section 4 is devoted to the derivation of the second-order sufficient conditions.

Notation. In all what follows $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, is a bounded Lipschitz domain with boundary Γ. The boundary consists of two disjoint parts Γ_N and Γ_D. Throughout the paper vectors and tensors are denoted by bold-face letters. We denote by $\mathbb{S} := \mathbb{R}^{d \times d}_{\text{sym}}$ the space of symmetric $d \times d$ matrices endowed with the Frobenius norm. For $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathbb{S}$ the associated scalar product is denoted by $\boldsymbol{\sigma} : \boldsymbol{\tau} = \sum_{ij} \sigma_{ij} \tau_{ij}$. Given a regularity exponent $p \in (1, \infty)$, the conjugate exponent is denoted by p', i.e., $1/p + 1/p' = 1$. The symbol X' is used for the dual space of a normed space X. The space of linear and continuous operators from a normed space X into a normed space Y is denoted by $\mathcal{L}(X, Y)$. If $X = Y$, then we sometimes simply write $\mathcal{L}(X)$. Moreover, we define the following abbreviations

$$W_D^{1,p}(\Omega; \mathbb{R}^d) := \{\boldsymbol{u} \in W^{1,p}(\Omega; \mathbb{R}^d) : \boldsymbol{u} = \boldsymbol{0} \text{ on } \Gamma_D\}, \quad V := W_D^{1,2}(\Omega; \mathbb{R}^d),$$

$$W_D^{-1,p}(\Omega; \mathbb{R}^d) := (W_D^{1,p'}(\Omega; \mathbb{R}^d))', \quad U := L^2(\Gamma_N; \mathbb{R}^d), \quad S := L^2(\Omega; \mathbb{S}).$$

The dual pairing between V and V' is denoted by $\langle \cdot, \cdot \rangle$ and the scalar product in L^2-type spaces such as $L^2(\Omega)$, S, and S^2 is always denoted by $(\cdot, \cdot)$. Furthermore, throughout the whole paper, $c > 0$ represents a generic constant.

Statement of the optimal control problem. Let us state the VI of first kind associated to static elastoplasticity with linear kinematic hardening: Given an inhomogeneity $\ell \in V'$ find generalized stress $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\tau}) \in S^2$ and displacement $\boldsymbol{u} \in V$ so that $\boldsymbol{\Sigma} \in \mathcal{K}$ and

$$\left. \begin{array}{r} (A\boldsymbol{\Sigma}, \boldsymbol{T} - \boldsymbol{\Sigma}) + (B^*\boldsymbol{u}, \boldsymbol{T} - \boldsymbol{\Sigma}) \geq 0 \quad \text{for all } \boldsymbol{T} \in \mathcal{K} \\ B\boldsymbol{\Sigma} = \ell \quad \text{in } V' \end{array} \right\} \quad (\textbf{VI})$$

is satisfied. The quantities in (**VI**) are defined as follows: For $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi}), \boldsymbol{T} = (\boldsymbol{\tau}, \boldsymbol{\mu}) \in S^2$ and $\boldsymbol{v} \in V$ the linear operators $A \colon S^2 \to S^2$ and $B \colon S^2 \to V'$ are

defined by

$$(A\boldsymbol{\Sigma}, \boldsymbol{T}) = \int_\Omega \boldsymbol{\tau} : \mathbb{C}^{-1}\boldsymbol{\sigma} \, \mathrm{d}x + \int_\Omega \boldsymbol{\mu} : \mathbb{H}^{-1}\boldsymbol{\chi} \, \mathrm{d}x \quad \text{and} \quad \langle B\boldsymbol{\Sigma}, \boldsymbol{v} \rangle = -\int_\Omega \boldsymbol{\sigma} : \boldsymbol{\varepsilon}(\boldsymbol{v}) \, \mathrm{d}x.$$

Herein $\mathbb{C}^{-1}(x)$ and $\mathbb{H}^{-1}(x)$ are linear maps from $\mathbb{S}$ to $\mathbb{S}$, which may depend on the spatial variable x, and $\boldsymbol{\varepsilon}(\boldsymbol{v}) = 1/2\left(\nabla\boldsymbol{v} + (\nabla\boldsymbol{v})^T\right)$ is the linearized strain tensor. The closed and convex set $\mathcal{K} \subset S^2$ of admissible stresses is determined by the von Mises yield condition, i.e.,

$$\mathcal{K} := \{\boldsymbol{\Sigma} \in S^2 : \phi(\boldsymbol{\Sigma}) \le 0 \text{ a.e. in } \Omega\} \quad \text{with} \quad \phi(\boldsymbol{\Sigma}) = \frac{\left\|\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D\right\|_\mathbb{S}^2 - \sigma_0^2}{2}. \quad (1.1)$$

Here $\boldsymbol{\sigma}^D = \boldsymbol{\sigma} - 1/d(\text{trace}\,\boldsymbol{\sigma})\mathbb{I}$ is the deviatoric part of $\boldsymbol{\sigma}$. It will be convenient in the following to abbreviate $\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D$ by

$$\mathcal{D}\boldsymbol{\Sigma} = \boldsymbol{\sigma}^D + \boldsymbol{\chi}^D.$$

For a detailed introduction into this and other common plasticity models we refer to [10].

With (**VI**) at hand the optimal control under consideration reads

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

where $J \colon V \times U \to \mathbb{R}$ denotes a given objective functional.

Standing assumption. The following assumption is supposed to hold throughout this paper:

Assumption 1.1.

(1) *The domain $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, is a bounded Lipschitz domain in the sense of [7, Chapter 1.2]. The boundary of Ω, denoted by Γ, consists of two disjoint measurable parts Γ_N and Γ_D such that $\Gamma = \Gamma_N \cup \Gamma_D$. While Γ_N is a relatively open subset, Γ_D is a relatively closed subset of Γ with positive measure. Furthermore we suppose that the set $\Omega \cup \Gamma_N$ is regular in the sense of Gröger, cf. [8, Definition 2].*

(2) *The yield stress σ_0 is assumed to be a positive constant.*

(3) *$\mathbb{C}^{-1}$ and $\mathbb{H}^{-1}$ are elements of $L^\infty(\Omega; \mathcal{L}(\mathbb{S}, \mathbb{S}))$. Both $\mathbb{C}^{-1}(x)$ and $\mathbb{H}^{-1}(x)$ are assumed to be uniformly coercive on $\mathbb{S}$. 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 an analogous relation holds for $\mathbb{H}^{-1}$ for all $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathbb{S}$.*

Some words concerning this assumption are in order. If $\Omega \subset \mathbb{R}^2$ is a bounded Lipschitz domain. Then $\Omega \cup \Gamma_N$ is regular in the sense of Gröger, iff $\overline{\partial\Omega \setminus \Gamma_D} \cap \Gamma_D$ is finite and no connected component of Γ_D consists of a single point, cf. [9]. For $d = 3$ the set $\Omega \cup \Gamma_N$ is regular in the sense of Gröger, if the boundary $\partial\Gamma_D$ within $\partial\Omega$ is locally bi-Lipschitz diffeomorphic to the unit interval $(0, 1)$. Sufficient conditions for this are given in [9]. We point out that a broad class of non-smooth domains is regular in the sense of Gröger.

The conditions in Assumption 1.1(3) are for instance fulfilled by isotropic and homogeneous materials, where the compliance tensor is given by

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

with Lamé constants μ and λ. In this case $\mathbb{C}^{-1}$ is coercive provided that $\mu > 0$ and $d\,\lambda + 2\,\mu > 0$. A common example for the hardening modulus is given by $\mathbb{H}^{-1}\boldsymbol{\chi} = \boldsymbol{\chi}/k_1$ with the hardening constant $k_1 > 0$, see [10, Section 3.4].

2   Known and preliminary results

In the following we recall two well-known results concerning (**VI**) which will be useful in the sequel. Furthermore we establish a new regularity result for the solution of (**VI**), see Theorem 2.4 below.

Proposition 2.1. *[11, Propositions 3.1, 3.2 and Lemma 3.3] For every $\ell \in V'$, problem (**VI**) possesses a unique solution $(\boldsymbol{\Sigma}, \boldsymbol{u}) \in S^2 \times V$.*

Based on this we can introduce the solution operator associated to (**VI**):

Definition 2.2. *The control-to-state map $V' \ni \ell \mapsto (\boldsymbol{\Sigma}, \boldsymbol{u}) \in S^2 \times V$ is denoted by G. We sometimes consider G with different domains and ranges. For simplicity these operators are also denoted by G.*

By introducing a slack variable the variational inequality in (**VI**) can equivalently be reformulated by means of a complementarity system:

Theorem 2.3. *[13, Theorem 2.2] Let $\ell \in V'$ be given. The pair $(\boldsymbol{\Sigma}, \boldsymbol{u}) \in S^2 \times V$ is the unique solution of (**VI**) if and only if there exists a plastic multiplier $\lambda \in L^2(\Omega)$ such that*

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

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

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

holds. Moreover, λ is unique.

If the inhomogeneity ℓ in (**VI**) is slightly more regular, then one can improve the integrability of the solution of (**VI**), as shown in the following. This result is essential to prove the Bouligand differentiability of G in Section 3.

Theorem 2.4. *There exists $\hat{p} > 2$ such that for all $p \in [2, \hat{p}]$ and for any $\ell \in W_D^{-1,p}(\Omega; \mathbb{R}^d)$, the unique solution $(\boldsymbol{\Sigma}, \boldsymbol{u})$ of (**VI**) belongs to $L^p(\Omega; \mathbb{S}^2) \times W_D^{1,p}(\Omega; \mathbb{R}^d)$. There exists $L > 0$ such that*

$$\|\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2\|_{L^p(\Omega;\mathbb{S}^2)} + \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_{W_D^{1,p}(\Omega;\mathbb{R}^d)} \le L\,\|\ell_1 - \ell_2\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)},$$

i.e., G is Lipschitz continuous from $W_D^{-1,p}(\Omega; \mathbb{R}^d)$ to $L^p(\Omega; \mathbb{S}^2) \times W_D^{1,p}(\Omega; \mathbb{R}^d)$.

Proof. The arguments are similar to [24, Theorem 4.4.4 and Proposition 4.4.5]. We aim to apply [12, Theorem 1.1]. To this end, we will reformulate (**VI**) and (2.1), respectively, as a single nonlinear PDE in $\boldsymbol{u}$, cf. (2.7) below, and verify [12, Assumption 1.5(2)] for the underlying nonlinearity.

Testing (2.1a) with $(\boldsymbol{\tau}, \boldsymbol{0})$, $\boldsymbol{\tau} \in S$, we find $\mathbb{C}^{-1}\boldsymbol{\sigma} - \boldsymbol{\varepsilon}(\boldsymbol{u}) + \lambda\mathcal{D}\boldsymbol{\Sigma} = 0$ a.e. in Ω. If we test with $(\boldsymbol{0}, \boldsymbol{\mu})$, $\boldsymbol{\mu} \in S$, we arrive at

$$\mathbb{H}^{-1}\boldsymbol{\chi} + \lambda\mathcal{D}\boldsymbol{\Sigma} = 0 \quad a.e.\ in\ \Omega. \tag{2.2}$$

Combining both equations yields

$$\mathbb{C}^{-1}\boldsymbol{\sigma} - \boldsymbol{\varepsilon}(\boldsymbol{u}) - \mathbb{H}^{-1}\boldsymbol{\chi} = 0 \quad a.e.\ in\ \Omega. \tag{2.3}$$

Next we derive a pointwise form of (**VI**). The arguments are standard. For convenience of the reader we shortly recall them. Let x_0 be an arbitrary Lebesgue

point of $\mathbb{C}^{-1}$, $\boldsymbol{\sigma}$, $\mathbb{H}^{-1}$, $\boldsymbol{\chi}$, $\boldsymbol{\varepsilon}(\boldsymbol{u})$ and their products arising in the sequel. Moreover let $(\boldsymbol{\tau}, \boldsymbol{\mu}) \in K$ be given, where $K := \{\boldsymbol{T} \in \mathbb{S}^2 : \phi(\boldsymbol{T}) \leq 0\}$. With $\rho > 0$ such that $B(x_0, \rho) \subset \Omega$, we define

$$(\tilde{\boldsymbol{\tau}}, \tilde{\boldsymbol{\mu}}) := \begin{cases} (\boldsymbol{\tau}, \boldsymbol{\mu}), & x \in B(x_0, \rho) \\ (\boldsymbol{\sigma}, \boldsymbol{\chi}), & x \in \Omega \setminus B(x_0, \rho). \end{cases}$$

Obviously, $(\tilde{\boldsymbol{\tau}}, \tilde{\boldsymbol{\mu}})$ is an element of $\mathcal{K}$. Testing (**VI**) with $(\tilde{\boldsymbol{\tau}}, \tilde{\boldsymbol{\mu}})$ yields

$$0 \leq \frac{1}{|B(x_0, \rho)|} \int_{B(x_0, \rho)} \mathbb{C}^{-1} \boldsymbol{\sigma} : (\boldsymbol{\tau} - \boldsymbol{\sigma}) + \mathbb{H}^{-1} \boldsymbol{\chi} : (\boldsymbol{\mu} - \boldsymbol{\chi}) - \boldsymbol{\varepsilon}(\boldsymbol{u}) : (\boldsymbol{\tau} - \boldsymbol{\sigma}) \, \mathrm{d}x.$$

We take the limit $\rho \searrow 0$ and obtain

$$\mathbb{C}^{-1}(x_0)\boldsymbol{\sigma}(x_0) : (\boldsymbol{\tau} - \boldsymbol{\sigma}(x_0)) + \mathbb{H}^{-1}(x_0)\boldsymbol{\chi}(x_0) : (\boldsymbol{\mu} - \boldsymbol{\chi}(x_0))$$
$$-\boldsymbol{\varepsilon}(\boldsymbol{u}(x_0)) : (\boldsymbol{\tau} - \boldsymbol{\sigma}(x_0)) \geq 0.$$

Since almost every point in Ω is a common Lebesgue point of $\mathbb{C}^{-1}$, $\boldsymbol{\sigma}$, $\mathbb{H}^{-1}$, $\boldsymbol{\chi}$, and $\boldsymbol{\varepsilon}(\boldsymbol{u})$, there holds f.a.a. $x \in \Omega$

$$\mathbb{C}^{-1}(x)\boldsymbol{\sigma}(x) : (\boldsymbol{\tau} - \boldsymbol{\sigma}(x)) + \mathbb{H}^{-1}(x)\boldsymbol{\chi}(x) : (\boldsymbol{\mu} - \boldsymbol{\chi}(x))$$
$$-\boldsymbol{\varepsilon}(\boldsymbol{u}(x)) : (\boldsymbol{\tau} - \boldsymbol{\sigma}(x)) \geq 0 \quad \forall(\boldsymbol{\tau}, \boldsymbol{\mu}) \in K. \tag{2.4}$$

Now we insert $(\boldsymbol{\sigma}(x), \boldsymbol{\mu})$ into (2.4), which results in

$$\mathbb{H}^{-1}(x)\boldsymbol{\chi}(x) : (\boldsymbol{\mu} - \boldsymbol{\chi}(x)) \geq 0 \quad \text{for all } \boldsymbol{\mu} \in \mathbb{S} \text{ such that } (\boldsymbol{\sigma}(x), \boldsymbol{\mu}) \in K.$$

This is the necessary and sufficient optimality condition for the convex problem

$$\min_{\boldsymbol{\mu} \in \bar{K} - \boldsymbol{\sigma}(x)} \frac{1}{2} \|\boldsymbol{\mu}\|^2_{\mathbb{H}^{-1}(x)}$$

with $\bar{K} := \{\boldsymbol{\tau} \in \mathbb{S} : (\boldsymbol{\tau}, 0) \in K\}$. Herein the norm induced by $\mathbb{H}^{-1}(x)$ is defined by $\|\boldsymbol{\mu}\|_{\mathbb{H}^{-1}(x)} = (\mathbb{H}^{-1}(x)\boldsymbol{\mu} : \boldsymbol{\mu})^{\frac{1}{2}}$. Note that $\boldsymbol{\mu} \in \bar{K} - \boldsymbol{\sigma}(x)$ is equivalent to $(\boldsymbol{\sigma}(x), \boldsymbol{\mu}) \in K$. Therefore we have f.a.a. $x \in \Omega$

$$\boldsymbol{\chi}(x) = \mathrm{Proj}^{\mathbb{H}^{-1}(x)}_{\bar{K} - \boldsymbol{\sigma}(x)}(0) = \mathrm{Proj}^{\mathbb{H}^{-1}(x)}_{\bar{K}}(\boldsymbol{\sigma}(x)) - \boldsymbol{\sigma}(x), \tag{2.5}$$

where, for a given closed and convex set $E \subset \mathbb{S}$, $\mathrm{Proj}^{\mathbb{H}^{-1}(x)}_E$ denotes the orthogonal projection on E with respect to the norm induced by $\mathbb{H}^{-1}(x)$. Inserting (2.5) in (2.3) yields

$$\mathbb{C}^{-1}(x)\boldsymbol{\sigma}(x) + \mathbb{H}^{-1}(x)(\boldsymbol{\sigma}(x) - \mathrm{Proj}^{\mathbb{H}^{-1}(x)}_{\bar{K}}(\boldsymbol{\sigma}(x))) = \boldsymbol{\varepsilon}(\boldsymbol{u})(x) \quad \text{a.e. in } \Omega. \tag{2.6}$$

We define $M_x : \mathbb{S} \to \mathbb{S}$ by $\boldsymbol{\sigma} \mapsto \mathbb{C}^{-1}(x)\boldsymbol{\sigma} + \mathbb{H}^{-1}(x)(\boldsymbol{\sigma} - \mathrm{Proj}^{\mathbb{H}^{-1}(x)}_{\bar{K}}(\boldsymbol{\sigma}))$. In view of the monotonicity of $\mathbb{H}^{-1}(x)(\boldsymbol{\sigma} - \mathrm{Proj}^{\mathbb{H}^{-1}(x)}_{\bar{K}}(\boldsymbol{\sigma}))$, see, e.g., [11, Lemma 4.1], we have for every $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathbb{S}$

$$(M_x(\boldsymbol{\sigma}) - M_x(\boldsymbol{\tau}), \, \boldsymbol{\sigma} - \boldsymbol{\tau})_{\mathbb{S}} \geq \left(\mathbb{C}^{-1}(x)(\boldsymbol{\sigma} - \boldsymbol{\tau}), \, \boldsymbol{\sigma} - \boldsymbol{\tau}\right)_{\mathbb{S}} \geq \underline{m} \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|^2_{\mathbb{S}},$$

where $\underline{m}$ is the coercivity constant of $\mathbb{C}^{-1}$. Thus $M_x(\cdot)$ is strongly monotone and coercive because of $M_x(\boldsymbol{0}) = \boldsymbol{0}$. Due to the boundedness of $\mathbb{C}^{-1}$ and $\mathbb{H}^{-1}$ and the non-expansiveness of $\mathrm{Proj}^{\mathbb{H}^{-1}(x)}_{\bar{K}}$ with respect to the norm induced by $\mathbb{H}^{-1}(x)$, there exists $\overline{m} > 0$, independent of $\boldsymbol{\sigma}, \boldsymbol{\tau}$ and x, so that

$$\|M_x(\boldsymbol{\sigma}) - M_x(\boldsymbol{\tau})\|_{\mathbb{S}} \leq \overline{m} \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{\mathbb{S}}.$$

Thus, thanks to the Browder-Minty theorem the inverse $M_x^{-1}(\cdot)$ w.r.t. $\boldsymbol{\sigma}$ exists f.a.a. $x \in \Omega$. Let us define $M^{-1}(x, \boldsymbol{\sigma}) := M_x^{-1}(\boldsymbol{\sigma})$ f.a.a. $x \in \Omega$. Then (2.6) is equivalent to $\boldsymbol{\sigma} = M^{-1}(\cdot, \boldsymbol{\varepsilon}(\boldsymbol{u}))$ and hence, due to (2.1b), $\boldsymbol{u}$ is a solution of

$$\int_{\Omega} M^{-1}(\cdot, \boldsymbol{\varepsilon}(\boldsymbol{u})) : \boldsymbol{\varepsilon}(\boldsymbol{v}) \, \mathrm{d}x = -\langle \ell, \, \boldsymbol{v} \rangle \quad \forall \boldsymbol{v} \in V, \tag{2.7}$$

which is the desired nonlinear PDE in the displacement field only. In order to apply [12, Theorem 1.1] we have to verify that M^{-1} satisfies [12, Assumption 1.5(2)], i.e., $M^{-1}(\cdot, \mathbf{0}) \in L^\infty(\Omega; \mathbb{S})$, $M^{-1}(\cdot, \boldsymbol{\sigma})$ is measurable, and M^{-1} is Lipschitz continuous and strongly monotone w.r.t. $\boldsymbol{\sigma}$ f.a.a. $x \in \Omega$.

The strong monotonicity of $M_x(\cdot)$ implies that $M_x^{-1}(\cdot)$ is Lipschitz continuous with Lipschitz constant $1/\underline{m}$, thus independent of x. Due to $\left(M_x^{-1}(\boldsymbol{\sigma}) - M_x^{-1}(\boldsymbol{\tau}), \boldsymbol{\sigma} - \boldsymbol{\tau}\right)_{\mathbb{S}} \geq \underline{m}/\overline{m}^2 \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{\mathbb{S}}^2$, M_x^{-1} is also strongly monotone. Moreover, $M^{-1}(\cdot, \mathbf{0}) = \mathbf{0} \in L^\infty(\Omega; \mathbb{S})$. Hence it remains to show that $M^{-1}(x, \boldsymbol{\sigma})$ is measurable with respect to x. As $\mathbb{C}^{-1}, \mathbb{H}^{-1}$ are measurable, there exist simple functions $\mathbb{C}_n^{-1}, \mathbb{H}_n^{-1} \in L^\infty(\Omega; \mathcal{L}(\mathbb{S}))$, $n \in \mathbb{N}$, with $\left\|\mathbb{C}^{-1}(x) - \mathbb{C}_n^{-1}(x)\right\|_{\mathcal{L}(\mathbb{S})} \to 0$ and $\left\|\mathbb{H}^{-1}(x) - \mathbb{H}_n^{-1}(x)\right\|_{\mathcal{L}(\mathbb{S})} \to 0$ f.a.a. $x \in \Omega$. Thus, there exists $N_x^{\mathbb{C}} \in \mathbb{N}$, depending on x, such that $\boldsymbol{\tau} : \mathbb{C}_n^{-1}(x)\boldsymbol{\tau} \geq \underline{m}/2 \|\boldsymbol{\tau}\|_{\mathbb{S}}^2$ for all $n \geq N_x^{\mathbb{C}}$. The same holds true for $\mathbb{H}_n^{-1}$. We define

$$\tilde{\mathbb{C}}_n^{-1}(x) := \begin{cases} \mathbb{C}_n^{-1}(x), & n \geq N_x^{\mathbb{C}} \\ \mathbb{I}_{\mathbb{S}}, & \text{else} \end{cases} \quad \text{and} \quad \tilde{\mathbb{H}}_n^{-1}(x) := \begin{cases} \mathbb{H}_n^{-1}(x), & n \geq N_x^{\mathbb{H}} \\ \mathbb{I}_{\mathbb{S}}, & \text{else}, \end{cases}$$

where $\mathbb{I}_{\mathbb{S}}: \mathbb{S} \to \mathbb{S}$ denotes the identity. Obviously $\tilde{\mathbb{C}}_n^{-1}: \Omega \to \mathbb{S}$, $\tilde{\mathbb{H}}_n^{-1}: \Omega \to \mathbb{S}$ are simple functions. Moreover $\tilde{\mathbb{C}}_n^{-1}(x), \tilde{\mathbb{H}}_n^{-1}(x) \in \mathbb{S}$ are coercive and $\tilde{\mathbb{C}}_n^{-1}(x) \to \mathbb{C}^{-1}(x)$ and $\tilde{\mathbb{H}}_n^{-1}(x) \to \mathbb{H}^{-1}(x)$ f.a.a. $x \in \Omega$. Analogous arguments as for M_x show that $M_x^n: \mathbb{S} \to \mathbb{S}$, defined by $\boldsymbol{\sigma} \mapsto \tilde{\mathbb{C}}_n^{-1}(x)\boldsymbol{\sigma} + \tilde{\mathbb{H}}_n^{-1}(x)(\boldsymbol{\sigma} - \text{Proj}_{\tilde{K}}^{\tilde{\mathbb{H}}_n^{-1}(x)}(\boldsymbol{\sigma})) \in \mathbb{S}$, is invertible f.a.a. $x \in \Omega$ and the inverse $(M_x^n)^{-1}(\cdot)$ is Lipschitz continuous with Lipschitz constant $L_x^n = \max(1, 2/\underline{m})$, independent of n. By construction $M^n(x, \boldsymbol{\sigma}) := M_x^n(\boldsymbol{\sigma})$ and $(M^n)^{-1}(x, \boldsymbol{\sigma}) := (M_x^n)^{-1}(\boldsymbol{\sigma})$ are simple functions with respect to x. For $\boldsymbol{\chi} := \text{Proj}_{\tilde{K}}^{\mathbb{H}^{-1}(x)}(\boldsymbol{\sigma})$ and $\boldsymbol{\chi}_n := \text{Proj}_{\tilde{K}}^{\tilde{\mathbb{H}}_n^{-1}(x)}(\boldsymbol{\sigma})$ we find

$$0 \leq \mathbb{H}^{-1}(x)(\boldsymbol{\chi} - \boldsymbol{\sigma}) : (\boldsymbol{\chi}_n - \boldsymbol{\chi}) + \tilde{\mathbb{H}}_n^{-1}(x)(\boldsymbol{\chi}_n - \boldsymbol{\sigma}) : (\boldsymbol{\chi} - \boldsymbol{\chi}_n)$$
$$= (\mathbb{H}^{-1}(x) - \tilde{\mathbb{H}}_n^{-1}(x))(\boldsymbol{\chi} - \boldsymbol{\sigma}) : (\boldsymbol{\chi}_n - \boldsymbol{\chi}) - \tilde{\mathbb{H}}_n^{-1}(x)(\boldsymbol{\chi}_n - \boldsymbol{\chi}) : (\boldsymbol{\chi}_n - \boldsymbol{\chi}).$$

The uniform coercivity of $\tilde{\mathbb{H}}_n^{-1}(x)$ thus implies $\|\mathbb{H}^{-1}(x) - \tilde{\mathbb{H}}_n^{-1}(x)\|_{\mathcal{L}(\mathbb{S})} \|\boldsymbol{\chi} - \boldsymbol{\sigma}\|_{\mathbb{S}} \geq c \|\boldsymbol{\chi}_n - \boldsymbol{\chi}\|_{\mathbb{S}}$ and therefore $\|\boldsymbol{\chi}_n - \boldsymbol{\chi}\|_{\mathbb{S}} \to 0$. Hence $\{\boldsymbol{\chi}_n(x)\}_{n \in \mathbb{N}}$ is bounded in $\mathbb{S}$. Consequently it holds

$$\|M_x^n(\boldsymbol{\sigma}) - M_x(\boldsymbol{\sigma})\|_{\mathbb{S}} =$$
$$= \left\|(\tilde{\mathbb{C}}_n^{-1}(x) - \mathbb{C}^{-1}(x))\boldsymbol{\sigma} + (\tilde{\mathbb{H}}_n^{-1}(x) - \mathbb{H}^{-1}(x))\boldsymbol{\sigma} + \mathbb{H}^{-1}(x)\boldsymbol{\chi} - \tilde{\mathbb{H}}_n^{-1}(x)\boldsymbol{\chi}_n\right\|_{\mathbb{S}}$$
$$\leq \left(\left\|\tilde{\mathbb{C}}_n^{-1}(x) - \mathbb{C}^{-1}(x)\right\|_{\mathcal{L}(\mathbb{S})} + \left\|\tilde{\mathbb{H}}_n^{-1}(x) - \mathbb{H}^{-1}(x)\right\|_{\mathcal{L}(\mathbb{S})}\right) \|\boldsymbol{\sigma}\|_{\mathbb{S}}$$
$$+ \left\|\mathbb{H}^{-1}(x)\right\|_{\mathcal{L}(\mathbb{S})} \|\boldsymbol{\chi} - \boldsymbol{\chi}_n\|_{\mathbb{S}} + \left\|\mathbb{H}^{-1}(x) - \tilde{\mathbb{H}}_n^{-1}(x)\right\|_{\mathcal{L}(\mathbb{S})} \|\boldsymbol{\chi}_n\|_{\mathbb{S}} \xrightarrow{n \to \infty} 0$$

and thus f.a.a. $x \in \Omega$

$$\left\|(M^n)^{-1}(x, \boldsymbol{\sigma}) - M^{-1}(x, \boldsymbol{\sigma})\right\|_{\mathbb{S}} =$$
$$= \left\|(M_x^n)^{-1}(M_x^n((M_x^n)^{-1}(\boldsymbol{\sigma}))) - (M_x^n)^{-1}(M_x^n(M_x^{-1}(\boldsymbol{\sigma})))\right\|_{\mathbb{S}}$$
$$\leq L_x^n \left\|M_x^n((M_x^n)^{-1}(\boldsymbol{\sigma})) - M_x^n(M_x^{-1}(\boldsymbol{\sigma}))\right\|_{\mathbb{S}}$$
$$= \max(1, 2/\underline{m}) \left\|M_x(M_x^{-1}(\boldsymbol{\sigma})) - M_x^n(M_x^{-1}(\boldsymbol{\sigma}))\right\|_{\mathbb{S}} \xrightarrow{n \to \infty} 0,$$

so that $M^{-1}(\cdot, \sigma)$ is indeed measurable. Alltogether we have shown that

- $M^{-1}(\cdot, \mathbf{0}) \in L^\infty(\Omega; \mathbb{S})$,
- $M^{-1}(\cdot, \boldsymbol{\sigma})$ is measurable,
- $\left(M^{-1}(x, \boldsymbol{\sigma}) - M^{-1}(x, \boldsymbol{\tau}), \boldsymbol{\sigma} - \boldsymbol{\tau}\right)_{\mathbb{S}} \geq \frac{\underline{m}}{\overline{m}^2} \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{\mathbb{S}}^2$,
- $\left\|M^{-1}(x, \boldsymbol{\sigma}) - M^{-1}(x, \boldsymbol{\tau})\right\|_{\mathbb{S}} \leq \frac{1}{\underline{m}} \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{\mathbb{S}}$

f.a.a. $x \in \Omega$ and all $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathbb{S}$. Therefore M^{-1} satisfies Assumption 1.5(2) in [12], and [12, Theorem 1.1] implies the existence of $\hat{p} > 2$ such that, for all $p \in [2, \hat{p}]$ and every $\ell \in W_D^{-1,p}(\Omega; \mathbb{R}^d)$, (2.7) admits a unique solution $\boldsymbol{u} \in W_D^{1,p}(\Omega; \mathbb{R}^d)$. Moreover, the associated solution map $W_D^{-1,p}(\Omega; \mathbb{R}^d) \ni \ell \mapsto \boldsymbol{u} \in W_D^{1,p}(\Omega; \mathbb{R}^d)$ is Lipschitz continuous for all $p \in [2, \hat{p}]$. As M_x^{-1} maps $L^p(\Omega; \mathbb{S})$ Lipschitz continuously into itself with Lipschitz constant $1/\underline{m}$, we conclude that $\boldsymbol{\sigma} = M^{-1}(\cdot, \boldsymbol{\varepsilon}(\boldsymbol{u})) \in L^p(\Omega; \mathbb{S})$. Since $\mathbf{0} \in \bar{K}$ and the projection is non-expansive, (2.5) implies $\boldsymbol{\chi} \in L^p(\Omega; \mathbb{S})$, and both $\boldsymbol{\sigma}$ and $\boldsymbol{\chi}$ depend Lipschitz continuously on $\boldsymbol{u}$ and thus on ℓ. $\qquad\square$

Remark 2.5. *We underline that the proof of [12, Theorem 1.1] adapts a technique introduced by Gröger in [8] for scalar second-order differential operators to the case of nonlinear elasticity. For this purpose Assumption 1.1(1) is required.*

Let us shortly comment on the existence of globally optimal controls for (**P**). Based on the Lipschitz continuity of $G: U \to V$ the following existence result can be proven by standard arguments:

Proposition 2.6. *Let $J: V \times U \to \mathbb{R}$ be weakly lower semicontinuous and let $R > 0$, $\hat{\boldsymbol{g}} \in U$ exist such that*

$$J(G(\boldsymbol{g}), \boldsymbol{g}) \geq J(G(\hat{\boldsymbol{g}}), \hat{\boldsymbol{g}}) \quad \forall \boldsymbol{g} \in U \ \text{with} \ \|\boldsymbol{g} - \hat{\boldsymbol{g}}\|_U > R,$$

then there exists a globally optimal solution of (**P**).

Note that we cannot expect the solution to be unique due to the nonlinearity of G.

3 Bouligand differentiability

In this section we establish the Bouligand differentiability of $G: \ell \to (\boldsymbol{\Sigma}, \boldsymbol{u})$ from $W_D^{-1,p}(\Omega; \mathbb{R}^d)$ to $S^2 \times V$. This result will be the essential tool to verify the sufficiency of our second-order conditions in Section 4. Before we are in the position to prove the Bouligand differentiability, we have to recall a directional differentiability result from [14] and derive some auxiliary results for the directional derivative.

Throughout this section let $\bar{\ell} \in W_D^{-1,p}(\Omega; \mathbb{R}^d)$ be fixed but arbitrary and denote the associated state by $(\bar{\boldsymbol{\Sigma}}, \bar{\boldsymbol{u}}, \bar{\lambda})$, i.e., $(\bar{\boldsymbol{\Sigma}}, \bar{\boldsymbol{u}}, \bar{\lambda}) \in S \times V \times L^2(\Omega)$ solves

$$A\bar{\boldsymbol{\Sigma}} + B^*\bar{\boldsymbol{u}} + \bar{\lambda}\mathcal{D}^*\mathcal{D}\bar{\boldsymbol{\Sigma}} = 0 \quad \text{in } S^2, \tag{3.1a}$$

$$B\bar{\boldsymbol{\Sigma}} = \bar{\ell} \quad \text{in } V', \tag{3.1b}$$

$$0 \leq \bar{\lambda}(x) \perp \phi(\bar{\boldsymbol{\Sigma}}(x)) \leq 0 \quad \text{a.e. in } \Omega. \tag{3.1c}$$

Then we define the following subsets of Ω up to sets of zero measure:

$$\bar{\mathcal{A}}_s := \{x \in \Omega : \bar{\lambda}(x) > 0\}, \tag{3.2a}$$

$$\bar{\mathcal{B}} := \{x \in \Omega : \phi(\bar{\boldsymbol{\Sigma}}(x)) = \bar{\lambda}(x) = 0\}, \tag{3.2b}$$

$$\bar{\mathcal{I}} := \{x \in \Omega : \phi(\bar{\boldsymbol{\Sigma}}(x)) < 0\}. \tag{3.2c}$$

The following theorem covers the directional differentiability of G in a weak sense:

Theorem 3.1. *[14, Theorem 3.2] For every $\bar{\ell} \in V'$ and every $\delta\ell \in V'$, the control-to-state map $G: V' \to S^2 \times V$ is directionally differentiable at $\bar{\ell}$ in direction $\delta\ell$ in a weak sense, i.e., there exists $\delta_w G(\bar{\ell}; \delta\ell) \in S^2 \times V$ such that*

$$\frac{G(\bar{\ell} + t\delta\ell) - G(\bar{\ell})}{t} \rightharpoonup \delta_w G(\bar{\ell}; \delta\ell) \quad \text{as } t \searrow 0.$$

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

$$\left(A\mathbf{\Sigma}', \boldsymbol{T} - \mathbf{\Sigma}'\right) + \left(B^*\boldsymbol{u}', \boldsymbol{T} - \mathbf{\Sigma}'\right)$$
$$+ \left(\bar{\lambda}, \mathcal{D}\mathbf{\Sigma}' : \mathcal{D}(\boldsymbol{T} - \mathbf{\Sigma}')\right) \geq 0 \quad \text{for all } \boldsymbol{T} \in \mathcal{S}_\ell, \tag{3.3a}$$

$$B\mathbf{\Sigma}' = \delta\ell, \tag{3.3b}$$

where the convex cone $\mathcal{S}_\ell$ is defined by

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

Again the VI in (3.3a) can be reformulated by means of a slack variable:

Theorem 3.2. *[14, Proposition 3.13] Let $\bar{\ell} \in V'$ and $\delta\ell \in V'$ be given and let $(\bar{\mathbf{\Sigma}}, \bar{\boldsymbol{u}}, \bar{\lambda})$ be the state and plastic multiplier associated with $\bar{\ell}$. A pair $(\mathbf{\Sigma}', \boldsymbol{u}') \in S^2 \times V$ is the unique solution of (3.3) if and only if there exists a multiplier $\lambda' \in L^2(\Omega)$ such that*

$$A\mathbf{\Sigma}' + B^*\boldsymbol{u}' + \bar{\lambda}\mathcal{D}^*\mathcal{D}\mathbf{\Sigma}' + \lambda'\mathcal{D}^*\mathcal{D}\bar{\mathbf{\Sigma}} = 0 \quad \text{in } S^2, \tag{3.4a}$$

$$B\mathbf{\Sigma}' = \delta\ell \quad \text{in } V', \tag{3.4b}$$

$$\mathbb{R} \ni \lambda'(x) \perp \mathcal{D}\bar{\mathbf{\Sigma}} : \mathcal{D}\mathbf{\Sigma}'(x) = 0 \quad \text{a.e. in } \bar{\mathcal{A}}_s, \tag{3.4c}$$

$$0 \leq \lambda'(x) \perp \mathcal{D}\bar{\mathbf{\Sigma}} : \mathcal{D}\mathbf{\Sigma}'(x) \leq 0 \quad \text{a.e. in } \bar{\mathcal{B}}, \tag{3.4d}$$

$$0 = \lambda'(x) \perp \mathcal{D}\bar{\mathbf{\Sigma}} : \mathcal{D}\mathbf{\Sigma}'(x) \in \mathbb{R} \quad \text{a.e. in } \bar{\mathcal{I}} \tag{3.4e}$$

holds. Moreover, λ' is unique.

Under more restrictive assumptions we will sharpen the assertion of Theorem 3.1. To be more precise we require

Assumption 3.3.

(i) *Let $\bar{\ell}, \delta\ell \in W_D^{-1,p}(\Omega; \mathbb{R}^d)$ with $p \in (2, \hat{p}]$ and $\hat{p}$ given by Theorem 2.4.*
(ii) *The solution of (3.1) satisfies $\bar{\boldsymbol{\chi}} \in L^s(\Omega; \mathbb{S})$ with*

$$s > \frac{2p}{p-2}. \tag{3.5}$$

Moreover we set

$$q = \frac{sp}{s+p}. \tag{3.6}$$

Note that $q > 2$ and $p' < q < p$ due to (3.5).

For the rest of this section we suppose Assumption 3.3 to hold.

Next let us consider the following perturbed problem:

$$A\mathbf{\Sigma} + B^*\boldsymbol{u} + \lambda\mathcal{D}^*\mathcal{D}\mathbf{\Sigma} = 0 \quad \text{in } S^2, \tag{3.7a}$$

$$B\mathbf{\Sigma} = \bar{\ell} + \delta\ell \quad \text{in } V', \tag{3.7b}$$

$$0 \leq \lambda(x) \perp \phi(\mathbf{\Sigma}(x)) \leq 0 \quad \text{a.e. in } \Omega. \tag{3.7c}$$

Clearly, (3.7) admits a unique solution. For the difference to the solution of (3.1) we find:

Lemma 3.4. *Let $(\bar{\mathbf{\Sigma}}, \bar{\lambda})$ and $(\mathbf{\Sigma}, \lambda)$ be given by the solution of (3.1) and (3.7), respectively. Then it holds*

(i) $\left\|\mathbf{\Sigma} - \bar{\mathbf{\Sigma}}\right\|_{L^p(\Omega; \mathbb{S}^2)} \leq L \left\|\delta\ell\right\|_{W^{-1,p}(\Omega; \mathbb{R}^d)},$
(ii) $\mathcal{D}\mathbf{\Sigma} - \mathcal{D}\bar{\mathbf{\Sigma}} \to 0$ in $L^m(\Omega; \mathbb{S})$ for all $1 \leq m < \infty$, if $\delta\ell \to 0$ in $W_D^{-1,p}(\Omega; \mathbb{R}^d)$,

(iii) $\left\|\lambda - \bar{\lambda}\right\|_{L^q(\Omega)} \leq c \left\|\delta\ell\right\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)}$.

Proof. Let $(\delta\ell_n)_{n\in\mathbb{N}} \subset W_D^{-1,p}(\Omega;\mathbb{R}^d)$ be an arbitrary sequence with $\delta\ell_n \to 0$ for $n \to 0$ and let $(\boldsymbol{\Sigma}_n, \lambda_n)$ be given by the solution of (3.7) with right hand side $\bar{\ell} + \delta\ell_n$. Assertion (i) follows directly from Theorem 2.4. To prove (ii), observe that $\left\|\mathcal{D}\boldsymbol{\Sigma}_n(x) - \mathcal{D}\bar{\boldsymbol{\Sigma}}(x)\right\|_{\mathbb{S}} \leq 2\sigma_0$ a.e. in Ω, i.e., $\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}}$ is bounded in $L^m(\Omega;\mathbb{S})$. In addition we know $\mathcal{D}\boldsymbol{\Sigma}_n(x) - \mathcal{D}\bar{\boldsymbol{\Sigma}}(x) \to 0$ a.e. on Ω because of (i). Consequently Lebesgue's theorem of dominated convergence implies (ii). From (2.2) and (3.1c) one deduces the following characterization of $\bar{\lambda}$

$$\sigma_0^2 \bar{\lambda} = \bar{\lambda}\mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\bar{\boldsymbol{\Sigma}} = -\mathbb{H}^{-1}\bar{\boldsymbol{\chi}} : \mathcal{D}\bar{\boldsymbol{\Sigma}} \quad \text{a.e. in } \Omega. \tag{3.8}$$

Completely analogously one shows $\sigma_0^2 \lambda_n = \lambda_n \mathcal{D}\boldsymbol{\Sigma}_n : \mathcal{D}\boldsymbol{\Sigma}_n = -\mathbb{H}^{-1}\boldsymbol{\chi}_n : \mathcal{D}\boldsymbol{\Sigma}_n$ a.e. in Ω. Hence it holds

$$\lambda_n - \bar{\lambda} = \frac{1}{\sigma_0^2} \left(-\mathbb{H}^{-1}\bar{\boldsymbol{\chi}} : (\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}}) - \mathbb{H}^{-1}(\boldsymbol{\chi}_n - \bar{\boldsymbol{\chi}}) : \mathcal{D}\boldsymbol{\Sigma}_n \right). \tag{3.9}$$

Thus, due to (3.6) and (i) we obtain

$$\begin{aligned}\left\|\lambda_n - \bar{\lambda}\right\|_{L^q(\Omega)} &\leq c \left(\|\bar{\boldsymbol{\chi}}\|_{L^s(\Omega;\mathbb{S})} \left\|\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}}\right\|_{L^p(\Omega;\mathbb{S})} + \sigma_0 \left\|\boldsymbol{\chi}_n - \bar{\boldsymbol{\chi}}\right\|_{L^q(\Omega;\mathbb{S})} \right) \\ &\leq c \left\|\delta\ell_n\right\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)},\end{aligned}$$

which establishes (iii). $\qquad\square$

To establish a regularity result for the directional derivative $(\boldsymbol{\Sigma}', \boldsymbol{u}', \lambda')$ let us now consider another perturbed problem:

$$A\boldsymbol{\Sigma}_t + B^*\boldsymbol{u}_t + \lambda_t \mathcal{D}^*\mathcal{D}\boldsymbol{\Sigma}_t = 0 \quad \text{in } S^2, \tag{3.10a}$$

$$B\boldsymbol{\Sigma}_t = \bar{\ell} + t\delta\ell \quad \text{in } V', \tag{3.10b}$$

$$0 \leq \lambda_t(x) \perp \phi(\boldsymbol{\Sigma}_t(x)) \leq 0 \quad \text{a.e. in } \Omega \tag{3.10c}$$

with $t > 0$ given.

Lemma 3.5. *Let $(\bar{\boldsymbol{\Sigma}}, \bar{\lambda})$ be given by the solution of (3.1), $(\boldsymbol{\Sigma}_t, \lambda_t)$ by the solution of (3.10) and $(\boldsymbol{\Sigma}', \lambda')$ by the solution of (3.4). Then it holds*

(i) $\frac{\boldsymbol{\Sigma}_t - \bar{\boldsymbol{\Sigma}}}{t} \rightharpoonup \boldsymbol{\Sigma}'$ *in* $L^p(\Omega;\mathbb{S}^2)$ *as* $t \searrow 0$,
(ii) $\frac{\lambda_t - \bar{\lambda}}{t} \rightharpoonup \lambda'$ *in* $L^q(\Omega)$ *as* $t \searrow 0$,
(iii) $\|\lambda'\|_{L^q(\Omega)} \leq c \|\delta\ell\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)}$ *and* $\|\boldsymbol{\Sigma}'\|_{L^p(\Omega;\mathbb{S}^2)} \leq L \|\delta\ell\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)}$.

Proof. Let $(t_n)_{n\in\mathbb{N}} \subset \mathbb{R}^+$ be an arbitrary sequence of positive real numbers converging to zero and let $(\boldsymbol{\Sigma}_{t_n}, \lambda_{t_n})$ be given by the solution of (3.10) with right hand side $\bar{\ell} + t_n\delta\ell$.
(i): According to Theorem 3.1 we know $(\boldsymbol{\Sigma}_{t_n} - \bar{\boldsymbol{\Sigma}})/t_n \rightharpoonup \boldsymbol{\Sigma}'$ in S^2. Moreover it holds

$$\left\|\frac{\boldsymbol{\Sigma}_{t_n} - \bar{\boldsymbol{\Sigma}}}{t_n}\right\|_{L^p(\Omega;\mathbb{S}^2)} \leq L \|\delta\ell\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)} \tag{3.11}$$

because of Theorem 2.4. Thus there exists a subsequence converging weakly in $L^p(\Omega,\mathbb{S}^2)$. The uniqueness of the weak limit therefore implies (i).
(ii): Analogously to (3.9), one shows that

$$\frac{\lambda_{t_n} - \bar{\lambda}}{t_n} = \frac{1}{\sigma_0^2} \left(-\mathbb{H}^{-1}\bar{\boldsymbol{\chi}} : \frac{\mathcal{D}\boldsymbol{\Sigma}_{t_n} - \mathcal{D}\bar{\boldsymbol{\Sigma}}}{t_n} - \mathbb{H}^{-1}\frac{\boldsymbol{\chi}_{t_n} - \bar{\boldsymbol{\chi}}}{t_n} : \mathcal{D}\boldsymbol{\Sigma}_{t_n} \right). \tag{3.12}$$

Due to (i) and (3.6) it holds $\mathbb{H}^{-1}\bar{\boldsymbol{\chi}} : (\mathcal{D}\boldsymbol{\Sigma}_{t_n} - \mathcal{D}\bar{\boldsymbol{\Sigma}})/t_n \rightharpoonup \mathbb{H}^{-1}\bar{\boldsymbol{\chi}} : \mathcal{D}\boldsymbol{\Sigma}'$ in $L^q(\Omega)$. If we choose $m = s$, Lemma 3.4 (ii) implies $\mathbb{H}^{-1}(\boldsymbol{\chi}_{t_n} - \bar{\boldsymbol{\chi}})/t_n : \mathcal{D}\boldsymbol{\Sigma}_{t_n} \rightharpoonup \mathbb{H}^{-1}\boldsymbol{\chi}' : \mathcal{D}\bar{\boldsymbol{\Sigma}}$ in $L^q(\Omega)$. Consequently, $(\lambda_{t_n} - \bar{\lambda})/t_n \rightharpoonup 1/\sigma_0^2 \left(-\mathbb{H}^{-1}\bar{\boldsymbol{\chi}}\mathcal{D}\boldsymbol{\Sigma}' - \mathbb{H}^{-1}\boldsymbol{\chi}' : \mathcal{D}\bar{\boldsymbol{\Sigma}} \right)$ in $L^q(\Omega)$.

Since $\mathcal{D}\boldsymbol{\Sigma}' : \mathcal{D}\bar{\boldsymbol{\Sigma}} = 0$ a.e. in $\bar{\mathcal{A}}_s$, $\bar{\lambda} = 0$ a.e. in $\bar{\mathcal{B}}$ and $\bar{\mathcal{I}}$, cf. (3.4c), (3.2), and (3.1c), we have $\bar{\lambda}\mathcal{D}\boldsymbol{\Sigma}' : \mathcal{D}\bar{\boldsymbol{\Sigma}} = 0$ a.e. in Ω. In view of (2.2) we thus know

$$\mathbb{H}^{-1}\bar{\boldsymbol{\chi}} : \mathcal{D}\boldsymbol{\Sigma}' = -\bar{\lambda}\mathcal{D}\boldsymbol{\Sigma}' : \mathcal{D}\bar{\boldsymbol{\Sigma}} = 0 \quad \text{a.e. in } \Omega. \tag{3.13}$$

Hence

$$\frac{\lambda_{t_n} - \bar{\lambda}}{t_n} \rightharpoonup -\frac{1}{\sigma_0^2}\mathbb{H}^{-1}\boldsymbol{\chi}' : \mathcal{D}\bar{\boldsymbol{\Sigma}} \quad \text{in } L^q(\Omega). \tag{3.14}$$

Testing (3.4a) with $(\mathbf{0}, \boldsymbol{\mu})$, $\boldsymbol{\mu} \in S$, yields $\mathbb{H}^{-1}\boldsymbol{\chi}' + \bar{\lambda}\mathcal{D}\boldsymbol{\Sigma}' + \lambda'\mathcal{D}\bar{\boldsymbol{\Sigma}} = 0$ a.e. in Ω. Because of (3.2), (3.1c), and (3.4c)–(3.4e) one deduces $\sigma_0^2\lambda' = \lambda'\mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\bar{\boldsymbol{\Sigma}}$ a.e. in Ω and therefore

$$\sigma_0^2\lambda' = -\mathbb{H}^{-1}\boldsymbol{\chi}' : \mathcal{D}\bar{\boldsymbol{\Sigma}} - \bar{\lambda}\mathcal{D}\boldsymbol{\Sigma}' : \mathcal{D}\bar{\boldsymbol{\Sigma}}. \tag{3.15}$$

Thus, (3.13) yields $\lambda' = -1/\sigma_0^2\mathbb{H}^{-1}\boldsymbol{\chi}' : \mathcal{D}\bar{\boldsymbol{\Sigma}}$, which together with (3.14) implies (ii).

(iii): In view of (i) and (3.11) it holds $\|\boldsymbol{\Sigma}'\|_{L^p(\Omega;\mathbb{S}^2)} \leq L \|\delta\ell\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)}$. From (3.6) and (3.12) we furthermore deduce

$$\left\|\frac{\lambda_{t_n} - \bar{\lambda}}{t_n}\right\|_{L^q(\Omega)} \leq c \left(\|\bar{\boldsymbol{\chi}}\|_{L^s(\Omega;\mathbb{S})} \left\|\frac{\mathcal{D}\boldsymbol{\Sigma}_{t_n} - \mathcal{D}\bar{\boldsymbol{\Sigma}}}{t_n}\right\|_{L^p(\Omega;\mathbb{S})} + \sigma_0 \left\|\frac{\boldsymbol{\chi}_{t_n} - \bar{\boldsymbol{\chi}}}{t_n}\right\|_{L^q(\Omega;\mathbb{S})}\right)$$

$$\leq c \|\delta\ell\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)}$$

and hence (ii) gives $\|\lambda'\|_{L^q(\Omega)} \leq c \|\delta\ell\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)}$. $\square$

Corollary 3.6. *For $\boldsymbol{u}'$ given by the solution of (3.3) there exists a constant $c > 0$ such that*

$$\|\boldsymbol{u}'\|_{W_D^{1,p}(\Omega;\mathbb{R}^d)} \leq c \|\delta\ell\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)}.$$

Proof. Let $\boldsymbol{\tau} \in S$ be arbitrary and define $\hat{\boldsymbol{T}} \in S^2$ by $\hat{\boldsymbol{T}} := (\boldsymbol{\tau}, -\boldsymbol{\tau})$. Due to $\mathcal{D}\hat{\boldsymbol{T}} = 0$ it then follows $\hat{\boldsymbol{T}} + \boldsymbol{\Sigma}' \in \mathcal{S}_\ell$. Thus we are allowed to test (3.3a) with $\hat{\boldsymbol{T}} + \boldsymbol{\Sigma}'$, which together with Lemma 3.5 (iii) yields

$$\int_\Omega \boldsymbol{\varepsilon}(\boldsymbol{u}') : \boldsymbol{\tau} \, \mathrm{d}x \leq (A\boldsymbol{\Sigma}', \hat{\boldsymbol{T}}) \leq c \|\boldsymbol{\Sigma}'\|_{L^p(\Omega;\mathbb{S}^2)} \|\boldsymbol{\tau}\|_{L^{p'}(\Omega;\mathbb{S})}$$

$$\leq c \|\delta\ell\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)} \|\boldsymbol{\tau}\|_{L^{p'}(\Omega;\mathbb{S})} \quad \forall \boldsymbol{\tau} \in S.$$

By the Hahn-Banach theorem $\boldsymbol{\varepsilon}(\boldsymbol{u}')$ can thus be extended to a functional on $L^{p'}(\Omega;\mathbb{S})$ and Korn's inequality gives the claim. $\square$

Now we are ready to prove the main result of this section.

Theorem 3.7. *Let $(\bar{\boldsymbol{\Sigma}}, \bar{\boldsymbol{u}}, \bar{\lambda})$ be the solution of (3.1), $(\boldsymbol{\Sigma}, \boldsymbol{u}, \lambda)$ the solution of (3.7) and $(\boldsymbol{\Sigma}', \boldsymbol{u}', \lambda')$ the solution of (3.4). Then it holds*

$$\|\boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'\|_{S^2} + \|\boldsymbol{u} - \bar{\boldsymbol{u}} - \boldsymbol{u}'\|_V = o\left(\|\delta\ell\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)}\right), \tag{3.16}$$

i.e., G is Bouligand-differentiable from $W_D^{-1,p}(\Omega;\mathbb{R}^d)$ to $S^2 \times V$.

Remark 3.8. *We point out that a norm gap is needed for the proof of Theorem 3.7, i.e., we were not able to show Bouligand-differentiability from V' to $S^2 \times V$, but we need $\bar{\ell}, \delta\ell \in W_D^{-1,p}(\Omega;\mathbb{R}^d)$. However this is not surprising since such norm gaps are commonly needed for the differentiability of nonlinear operators and also appear in the proof of Fréchet-differentiability for quasi-linear equations, see e.g. [25].*

Proof of Theorem 3.7. Let $(\delta\ell_n)_{n\in\mathbb{N}} \subset W_D^{-1,p}(\Omega;\mathbb{R}^d)$ be an arbitrary sequence with $\delta\ell_n \to 0$ for $n \to 0$. Furthermore by $(\boldsymbol{\Sigma}_n, \boldsymbol{u}_n, \lambda_n)$ we denote the solution of (3.7) with right hand side $\ell + \delta\ell_n$ and by $(\boldsymbol{\Sigma}'_n, \boldsymbol{u}'_n, \lambda'_n)$ the solution of (3.4) with right hand

side $\delta\ell_n$. By subtracting (3.1a) and (3.4a) from (3.7a) and testing with $\boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n$ we arrive at

$$
\begin{aligned}
&\big(A(\boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n),\, \boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n\big) + \underbrace{\big(B^*(\boldsymbol{u}_n - \bar{\boldsymbol{u}} - \boldsymbol{u}'_n),\, \boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n\big)}_{=:I_1} \\
&+ \big(\lambda_n \mathcal{D}\boldsymbol{\Sigma}_n - \bar{\lambda}\mathcal{D}\bar{\boldsymbol{\Sigma}} - \bar{\lambda}\mathcal{D}\boldsymbol{\Sigma}'_n - \lambda'_n \mathcal{D}\bar{\boldsymbol{\Sigma}},\, \mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}'_n\big) = 0.
\end{aligned}
\tag{3.17}
$$

Thanks to (3.1b), (3.7b) and (3.4b) we know $B(\boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n) = 0$ such that $I_1 = 0$. Since $\mathbb{C}^{-1}$ and $\mathbb{H}^{-1}$ are uniformly coercive, the linear operator A induces an equivalent norm on S^2 so that (3.17) implies

$$
\begin{aligned}
&c\,\big\|\boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n\big\|^2_{S^2} \\
&\leq \big(A(\boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n),\, \boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n\big) \\
&= -\int_\Omega \underbrace{\bar{\lambda}}_{\geq 0} \underbrace{\big\|\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}'_n\big\|^2_{\mathbb{S}}}_{\geq 0}\,\mathrm{d}x \\
&\quad -\int_\Omega (\lambda_n - \bar{\lambda} - \lambda'_n)\mathcal{D}\boldsymbol{\Sigma}_n : (\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}'_n)\,\mathrm{d}x \\
&\quad -\int_\Omega \lambda'_n(\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}}) : (\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}'_n)\,\mathrm{d}x \leq I_s + I_b + I_i + I_\Omega,
\end{aligned}
\tag{3.18}
$$

where

$$
I_\Omega := -\int_\Omega \lambda'_n(\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}}) : (\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}'_n)\,\mathrm{d}x.
$$

Moreover, I_s is given by

$$
I_s := -\int_{\bar{\mathcal{A}}_s} (\lambda_n - \bar{\lambda} - \lambda'_n)\mathcal{D}\boldsymbol{\Sigma}_n : (\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}'_n)\,\mathrm{d}x,
$$

and I_b and I_i are defined by the analogous integrals on $\bar{\mathcal{B}}$ and $\bar{\mathcal{I}}$, respectively, where $\bar{\mathcal{A}}_s, \bar{\mathcal{B}}$ and $\bar{\mathcal{I}}$ are as defined in (3.2). Note that $\bar{\mathcal{A}}_s \cup \bar{\mathcal{B}} \cup \bar{\mathcal{I}} = \Omega$. By Lemma 3.4 there exist subsequences λ_{n_k} and $\boldsymbol{\Sigma}_{n_k}$ with $\lambda_{n_k}(x) \to \bar{\lambda}(x)$ and $\boldsymbol{\Sigma}_{n_k}(x) \to \bar{\boldsymbol{\Sigma}}(x)$ a.e. on Ω. For the sake of convenience we denote these subsequences again by λ_n and $\boldsymbol{\Sigma}_n$. Next we will estimate I_s, I_b, I_i and I_Ω separately.

Estimation of I_Ω:
According to (3.6) there exists $m > 1$ with $\frac{1}{q} + \frac{1}{m} + \frac{1}{2} = 1$ such that Lemma 3.5 (iii) implies

$$
\begin{aligned}
I_\Omega &\leq \|\lambda'_n\|_{L^q(\Omega)} \big\|\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}}\big\|_{L^m(\Omega;\mathbb{S})} \big\|\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}'_n\big\|_S \\
&\leq c\,\|\delta\ell_n\|_{W^{-1,p}(\Omega;\mathbb{R}^d)} \big\|\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}}\big\|_{L^m(\Omega;\mathbb{S})} \big\|\boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n\big\|_{S^2}.
\end{aligned}
\tag{3.19}
$$

Estimation of I_s:
If $x \in \bar{\mathcal{A}}_s$, then it holds $\bar{\lambda}(x) > 0$ and (3.1c) gives $\phi(\bar{\boldsymbol{\Sigma}}(x)) = 0$. Due to pointwise convergence and the complementarity (3.7c) there exists $N_x \in \mathbb{N}$, depending on x, such that $\lambda_n(x) > 0$ and $\phi(\boldsymbol{\Sigma}_n(x)) = 0$, i.e., $\|\mathcal{D}\boldsymbol{\Sigma}_n(x)\|_{\mathbb{S}} = \sigma_0$, for all $n \geq N_x$ and f.a.a. $x \in \bar{\mathcal{A}}_s$. From Lemma A.1 we conclude

$$
z_n(x) := \frac{\mathcal{D}\bar{\boldsymbol{\Sigma}}(x) : (\mathcal{D}\bar{\boldsymbol{\Sigma}}(x) - \mathcal{D}\boldsymbol{\Sigma}_n(x))}{\big\|\mathcal{D}\bar{\boldsymbol{\Sigma}}(x)\big\|_{\mathbb{S}} \big\|\mathcal{D}\bar{\boldsymbol{\Sigma}}(x) - \mathcal{D}\boldsymbol{\Sigma}_n(x)\big\|_{\mathbb{S}}} \to 0 \quad \text{a.e. on } \bar{\mathcal{A}}_s.
$$

Obviously it holds $|z_n(x)| \leq 1$ and hence Lebesgue's dominated convergence theorem yields

$$
z_n \to 0 \quad \text{in } L^\xi(\bar{\mathcal{A}}_s) \quad \text{for all } 1 \leq \xi < \infty.
\tag{3.20}
$$

Since $\mathcal{D}\bar{\Sigma} : \mathcal{D}\Sigma'_n = 0$ a.e. in $\bar{\mathcal{A}}_s$, cf. (3.4c), one obtains

$$
\begin{aligned}
I_s = &-\int_{\bar{\mathcal{A}}_s} (\lambda_n - \bar{\lambda} - \lambda'_n)(\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma}) : (\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma'_n) \, \mathrm{d}x \\
&+ \int_{\bar{\mathcal{A}}_s} (\lambda_n - \bar{\lambda} - \lambda'_n) z_n \left\| \mathcal{D}\bar{\Sigma} \right\|_{\mathbb{S}} \left\| \mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma} \right\|_{\mathbb{S}} \, \mathrm{d}x.
\end{aligned}
$$

Again, by (3.6) there exist $m > 1$ and $\xi > 1$ with $\frac{1}{q} + \frac{1}{m} + \frac{1}{2} = 1$ and $\frac{1}{q} + \frac{1}{\xi} + \frac{1}{p} = 1$. Thanks to Lemma 3.4 and Lemma 3.5 (iii) we thus arrive at

$$
\begin{aligned}
I_s \leq\; & \left\| \lambda_n - \bar{\lambda} - \lambda'_n \right\|_{L^q(\Omega)} \left\| \mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma} \right\|_{L^m(\Omega;\mathbb{S})} \left\| \mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma'_n \right\|_{S^2} \\
& + \sigma_0 \left\| \lambda_n - \bar{\lambda} - \lambda'_n \right\|_{L^q(\Omega)} \left\| z_n \right\|_{L^\xi(\bar{\mathcal{A}}_s)} \left\| \Sigma_n - \bar{\Sigma} \right\|_{L^p(\Omega;\mathbb{S}^2)} \\
\leq\; & c \left\| \delta\ell_n \right\|_{W^{-1,p}(\Omega;\mathbb{R}^d)} \left\| \mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma} \right\|_{L^m(\Omega;\mathbb{S})} \left\| \Sigma_n - \bar{\Sigma} - \Sigma'_n \right\|_{S^2} \\
& + c \left\| \delta\ell_n \right\|^2_{W^{-1,p}(\Omega;\mathbb{R}^d)} \left\| z_n \right\|_{L^\xi(\bar{\mathcal{A}}_s)}.
\end{aligned}
\tag{3.21}
$$

Estimation of I_b:
For $x \in \bar{\mathcal{B}}$, one finds $\phi(\bar{\Sigma}(x)) = \bar{\lambda}(x) = 0$. Hence we have either $\lambda_n(x) = 0$ and consequently $\lambda_n(x) - \bar{\lambda}(x) = 0$ or $\lambda_n(x) > 0$ and thus $\phi(\Sigma_n(x)) = 0$, i.e., $\|\mathcal{D}\Sigma_n(x)\|_{\mathbb{S}} = \sigma_0$. Then Lemma A.1 yields that

$$
(\lambda_n - \bar{\lambda})\mathcal{D}\bar{\Sigma} : (\mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma_n) = (\lambda_n - \bar{\lambda})\frac{1}{2} \left\| \mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma_n \right\|^2_{\mathbb{S}}
\tag{3.22}
$$

in both cases. Moreover we know that $\bar{\lambda}\mathcal{D}\bar{\Sigma} : \mathcal{D}\Sigma'_n = \lambda'_n \mathcal{D}\bar{\Sigma} : \mathcal{D}\Sigma'_n = 0$ a.e. in Ω and $\mathcal{D}\bar{\Sigma} : \mathcal{D}\Sigma'_n \leq 0$ and $0 \leq \lambda'_n$ a.e. in $\bar{\mathcal{B}}$, cf. (3.2) and (3.4c)–(3.4e). Therefore we obtain

$$
\begin{aligned}
I_b = &-\int_{\bar{\mathcal{B}}} (\lambda_n - \bar{\lambda} - \lambda'_n)\mathcal{D}\Sigma_n : (\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma'_n) \, \mathrm{d}x \\
= &-\int_{\bar{\mathcal{B}}} (\lambda_n - \bar{\lambda} - \lambda'_n)(\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma}) : (\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma'_n) \, \mathrm{d}x \\
&+ \int_{\bar{\mathcal{B}}} [(\lambda_n - \bar{\lambda})\mathcal{D}\bar{\Sigma} : (\mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma_n) + \underbrace{\lambda_n}_{\geq 0} \underbrace{\mathcal{D}\bar{\Sigma} : \mathcal{D}\Sigma'_n}_{\leq 0} - \underbrace{\bar{\lambda}\mathcal{D}\bar{\Sigma} : \Sigma'_n}_{=0}] \, \mathrm{d}x \\
&+ \int_{\bar{\mathcal{B}}} [\lambda'_n(\mathcal{D}\bar{\Sigma} : \mathcal{D}\Sigma_n - \underbrace{\left\| \mathcal{D}\bar{\Sigma} \right\|^2_{\mathbb{S}}}_{=\sigma_0^2}) - \underbrace{\lambda'_n \mathcal{D}\bar{\Sigma} : \mathcal{D}\Sigma'_n}_{=0}] \, \mathrm{d}x \\
\leq &-\int_{\bar{\mathcal{B}}} (\lambda_n - \bar{\lambda} - \lambda'_n)(\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma}) : (\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma'_n) \, \mathrm{d}x \\
&+ \int_{\bar{\mathcal{B}}} (\lambda_n - \bar{\lambda})\mathcal{D}\bar{\Sigma} : (\mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma_n) \, \mathrm{d}x + \int_{\bar{\mathcal{B}}} \underbrace{\lambda'_n}_{\geq 0} (\sigma_0 \underbrace{\left\| \mathcal{D}\Sigma_n \right\|_{\mathbb{S}}}_{\leq \sigma_0} - \sigma_0^2) \, \mathrm{d}x \\
\leq &-\int_{\bar{\mathcal{B}}} (\lambda_n - \bar{\lambda} - \lambda'_n)(\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma}) : (\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma'_n) \, \mathrm{d}x \\
&+ \int_{\bar{\mathcal{B}}} (\lambda_n - \bar{\lambda})\frac{1}{2} \left\| \mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma_n \right\|^2_{\mathbb{S}} \, \mathrm{d}x,
\end{aligned}
$$

where we used (3.22) for the last estimate. Thus, as in (3.21) there exist m and ξ with $\frac{1}{q} + \frac{1}{m} + \frac{1}{2} = 1$ and $\frac{1}{q} + \frac{1}{\xi} + \frac{1}{p} = 1$, so that

$$
\begin{aligned}
I_b &\le \left\| \lambda_n - \bar{\lambda} - \lambda_n' \right\|_{L^q(\Omega)} \left\| \mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} \right\|_{L^m(\Omega;\mathbb{S})} \left\| \mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}_n' \right\|_{S^2} \\
&\quad + \frac{1}{2} \left\| \lambda_n - \bar{\lambda} \right\|_{L^q(\Omega)} \left\| \mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} \right\|_{L^\xi(\Omega;\mathbb{S})} \left\| \boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} \right\|_{L^p(\Omega;\mathbb{S}^2)} \\
&\le c \left\| \delta\ell_n \right\|_{W^{-1,p}(\Omega;\mathbb{R}^d)} \left\| \mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} \right\|_{L^m(\Omega;\mathbb{S})} \left\| \boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_n' \right\|_{S^2} \\
&\quad + c \left\| \delta\ell_n \right\|_{W^{-1,p}(\Omega;\mathbb{R}^d)}^2 \left\| \mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} \right\|_{L^\xi(\Omega;\mathbb{S})}.
\end{aligned}
\tag{3.23}
$$

Estimation of I_i:
For $x \in \bar{\mathcal{I}}$ there holds $\phi(\bar{\boldsymbol{\Sigma}}(x)) < 0$ and $\bar{\lambda}(x) = 0$ in view of (3.1c). Due to the continuity of ϕ, the pointwise convergence and the complementarity (3.7c) there exists $\tilde{N}_x \in \mathbb{N}$ such that $\phi(\boldsymbol{\Sigma}_n(x)) < 0$ and $\lambda_n(x) = 0$ for all $n \ge \tilde{N}_x$ and for a.a. $x \in \bar{\mathcal{I}}$. In particular this implies $(\lambda_n - \bar{\lambda})/\|\delta\ell_n\|_{W^{-1,p}(\Omega;\mathbb{R}^d)} \to 0$ a.e. in $\bar{\mathcal{I}}$. Furthermore we know $\left\| \lambda_n - \bar{\lambda} \right\|_{L^q(\mathcal{I}(\ell))} / \|\delta\ell_n\|_{W^{-1,p}(\Omega;\mathbb{R}^d)} \le c \; \forall n \in \mathbb{N}$ with $2 < q$, cf. Lemma 3.4 (iii) and (3.6). Thus, Lemma A.2 implies

$$
\omega_n := \frac{\lambda_n - \bar{\lambda}}{\|\delta\ell_n\|_{W^{-1,p}(\Omega;\mathbb{R}^d)}} \to 0 \quad \text{in } L^2(\bar{\mathcal{I}}).
\tag{3.24}
$$

Due to (3.4e), it follows

$$
\begin{aligned}
I_i &= -\int_{\bar{\mathcal{I}}} (\lambda_n - \bar{\lambda} - \lambda_n') \mathcal{D}\boldsymbol{\Sigma}_n : (\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}_n') \, \mathrm{d}x \\
&= \int_{\bar{\mathcal{I}}} \omega_n \|\delta\ell_n\|_{W^{-1,p}(\Omega;\mathbb{R}^d)} \mathcal{D}\boldsymbol{\Sigma}_n : (\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}_n') \, \mathrm{d}x \\
&\le \sigma_0 \|\omega_n\|_{L^2(\bar{\mathcal{I}})} \|\delta\ell_n\|_{W^{-1,p}(\Omega;\mathbb{R}^d)} \left\| \boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_n' \right\|_{S^2}.
\end{aligned}
\tag{3.25}
$$

In summary, (3.19), (3.21), (3.23), and (3.25) together with (3.18), Lemma 3.5 (iii), and Young's inequality yield

$$
\begin{aligned}
\frac{\left\| \boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_n' \right\|_{S^2}^2}{\|\delta\ell_n\|_{W^{-1,p}(\Omega;\mathbb{R}^d)}^2} &\le c \left(\left\| \mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} \right\|_{L^m(\Omega;\mathbb{S})} + \|\omega_n\|_{L^2(\bar{\mathcal{I}})} \right)^2 \\
&\quad + c \left(\|z_n\|_{L^\xi(\bar{\mathcal{A}}_s)} + \left\| \mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} \right\|_{L^\xi(\Omega;\mathbb{S})} \right) \xrightarrow{n\to\infty} 0,
\end{aligned}
\tag{3.26}
$$

where we used Lemma 3.4 (ii), (3.20) and (3.24) to pass to the limit.

To prove the remainder term property for $\boldsymbol{u}$, we subtract (3.1a) and (3.4a) from (3.7a) and test the arising equation with

$$
\tilde{\boldsymbol{T}} := (\boldsymbol{\varepsilon}(\boldsymbol{u}_n) - \boldsymbol{\varepsilon}(\bar{\boldsymbol{u}}) - \boldsymbol{\varepsilon}(\boldsymbol{u}_n'), -\boldsymbol{\varepsilon}(\boldsymbol{u}_n) + \boldsymbol{\varepsilon}(\bar{\boldsymbol{u}}) + \boldsymbol{\varepsilon}(\boldsymbol{u}_n')) \in S^2
$$

so that

$$
\begin{aligned}
&(A(\boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_n'), \tilde{\boldsymbol{T}}) + (B^*(\boldsymbol{u}_n - \bar{\boldsymbol{u}} - \boldsymbol{u}_n'), \tilde{\boldsymbol{T}}) \\
&\quad + \underbrace{(\lambda_n \mathcal{D}\boldsymbol{\Sigma}_n - \bar{\lambda}\mathcal{D}\bar{\boldsymbol{\Sigma}} - \bar{\lambda}\mathcal{D}\boldsymbol{\Sigma}_n' - \lambda_n'\mathcal{D}\bar{\boldsymbol{\Sigma}}, \mathcal{D}\tilde{\boldsymbol{T}})}_{=:I_2} = 0.
\end{aligned}
\tag{3.27}
$$

is obtained. As $\mathcal{D}\tilde{\boldsymbol{T}} = 0$, it holds $I_2 = 0$, and Korn's inequality implies

$$
\begin{aligned}
\|\boldsymbol{u}_n - \bar{\boldsymbol{u}} - \boldsymbol{u}_n'\|_V^2 &\le c \, (B^*(\boldsymbol{u}_n - \bar{\boldsymbol{u}} - \boldsymbol{u}_n'), \tilde{\boldsymbol{T}}) \le c \, (A(\boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_n'), \tilde{\boldsymbol{T}}) \\
&\le c \, \|\boldsymbol{u}_n - \bar{\boldsymbol{u}} - \boldsymbol{u}_n'\|_V \left\| \boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_n' \right\|_{S^2}.
\end{aligned}
$$

Hence (3.26) gives $\|\boldsymbol{u}_n - \bar{\boldsymbol{u}} - \boldsymbol{u}_n'\|_V / \|\delta\ell_n\|_{W^{-1,p}(\Omega;\mathbb{R}^d)} \to 0$.

Since the above arguments hold for every subsequence of $(\boldsymbol{\Sigma}_n, \boldsymbol{u}_n, \lambda_n)$, a standard argument implies (3.16) for the whole sequence. $\qquad\square$

Theorem 3.7 entails two consequences stated in the following corollaries:

Corollary 3.9. *The control-to-state map* $G\colon W_D^{-1,p}(\Omega;\mathbb{R}^d) \to S^2 \times V$ *is direction-ally differentiable at* $\bar{\ell}$ *in direction* $\delta\ell$.

Proof. Since $\mathcal{S}_\ell$ is a cone, $\delta\ell \mapsto \delta G(\bar{\ell},\delta\ell)$ is positively homogeneous so that $(t\boldsymbol{\Sigma}', t\boldsymbol{u}')$ is the solution of (3.4) with right hand side $t\delta\ell$. Consequently, (3.16) yields $\left(\left\| \boldsymbol{\Sigma}_t - \bar{\boldsymbol{\Sigma}} - t\boldsymbol{\Sigma}' \right\|_{S^2} + \left\| \boldsymbol{u}_t - \bar{\boldsymbol{u}} - t\boldsymbol{u}' \right\|_V \right)/t \xrightarrow{t\searrow 0} 0.$ □

Remark 3.10. *If the solution of* (3.1) *satisfies the weaker condition* $\bar{\chi} \in L^s(\Omega;\mathbb{S})$ *with* $s > \frac{p}{p-2}$, *cf.* (3.5), *then it can be shown by similar arguments as in the proof of Theorem 3.7 that the control-to-state map is only directionally differentiable without the remainder term property* (3.16).

Corollary 3.11. *Let the multipliers* λ, $\bar{\lambda}$ *and* λ' *be given by the solutions of* (3.7), (3.1) *and* (3.4), *respectively. Then it holds*

$$\left\| \lambda - \bar{\lambda} - \lambda' \right\|_{L^\alpha(\Omega)} = o\left(\left\| \delta\ell \right\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)} \right) \tag{3.28}$$

with $\frac{1}{\alpha} = \frac{1}{s} + \frac{1}{2}$.

Proof. In view of (3.8), (3.15) and (3.13) it holds

$$\sigma_0^2(\lambda - \bar{\lambda} - \lambda') = -\mathbb{H}^{-1}\boldsymbol{\chi} : \mathcal{D}\boldsymbol{\Sigma} + \mathbb{H}^{-1}\bar{\boldsymbol{\chi}} : \mathcal{D}\bar{\boldsymbol{\Sigma}} + \mathbb{H}^{-1}\boldsymbol{\chi}' : \mathcal{D}\bar{\boldsymbol{\Sigma}}$$
$$= -\mathbb{H}^{-1}(\boldsymbol{\chi} - \bar{\boldsymbol{\chi}}) : (\mathcal{D}\boldsymbol{\Sigma} - \mathcal{D}\bar{\boldsymbol{\Sigma}}) - \mathbb{H}^{-1}(\boldsymbol{\chi} - \bar{\boldsymbol{\chi}} - \boldsymbol{\chi}') : \mathcal{D}\bar{\boldsymbol{\Sigma}}$$
$$\quad - \mathbb{H}^{-1}\bar{\boldsymbol{\chi}} : (\mathcal{D}\boldsymbol{\Sigma} - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}').$$

Let β be defined by $\beta := 1/(\frac{1}{s} + \frac{1}{2} - \frac{1}{p})$ so that $\frac{1}{p} + \frac{1}{\beta} = \frac{1}{s} + \frac{1}{2} = \frac{1}{\alpha}$. Obviously $\alpha < 2$ such that $\left\| \mathcal{D}\bar{\boldsymbol{\Sigma}}(x) \right\|_{\mathbb{S}} \leq \sigma_0$ a.e. in Ω implies

$$\left\| \lambda - \bar{\lambda} - \lambda' \right\|_{L^\alpha(\Omega)} \leq c \left(\left\| \boldsymbol{\chi} - \bar{\boldsymbol{\chi}} \right\|_{L^p(\Omega;\mathbb{S})} \left\| \mathcal{D}\boldsymbol{\Sigma} - \mathcal{D}\bar{\boldsymbol{\Sigma}} \right\|_{L^\beta(\Omega;\mathbb{S})} + \left\| \boldsymbol{\chi} - \bar{\boldsymbol{\chi}} - \boldsymbol{\chi}' \right\|_S \right)$$
$$\quad + c \left\| \bar{\boldsymbol{\chi}} \right\|_{L^s(\Omega;\mathbb{S})} \left\| \boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}' \right\|_{S^2}.$$

Thus Theorem 3.7 and Lemma 3.4 give the assertion. □

4   Second-order sufficient optimality conditions

With the differentiability results of the previous section at hand we can now establish second-order sufficient condtions which guarantee local optimality for (**P**). As already indicated in the introduction, the analysis is based on a Taylor expansion of a tailored Lagrangian. For this purpose we require the following

Assumption 4.1. *Let* $J\colon V \times U \ni (\boldsymbol{u},\boldsymbol{g}) \mapsto J(\boldsymbol{u},\boldsymbol{g}) \in \mathbb{R}$ *be twice continuously Fréchet-differentiable.*

In preparation of our main result, we will next provide some auxiliary results. First observe that, in view of Theorem 2.3, Problem (**P**) is equivalent to

$$\text{Minimize} \quad J(\boldsymbol{u},\boldsymbol{g})$$
$$\text{s.t.} \quad \begin{cases} A\boldsymbol{\Sigma} + B^*\boldsymbol{u} + \lambda\mathcal{D}^*\mathcal{D}\boldsymbol{\Sigma} = 0 & \text{in } S^2, \\ \qquad\qquad\quad B\boldsymbol{\Sigma} = -\tau_N^*(\boldsymbol{g}) & \text{in } V', \\ 0 \leq \lambda(x) \perp \phi(\boldsymbol{\Sigma}(x)) \leq 0 & \text{a.e. in } \Omega. \end{cases}$$

Here the operator τ_N^* denotes the adjoint of the trace $\tau_N\colon V \to U$ on Γ_N, i.e.,

$$\tau_N^*\colon U \to V', \quad \langle \tau_N^*(\boldsymbol{g}), \boldsymbol{v} \rangle := \int_{\Gamma_N} \boldsymbol{g} \cdot \tau_N(\boldsymbol{v}) \, \mathrm{d}s, \quad \boldsymbol{v} \in V.$$

In all what follows we set

$$p = \min(\hat{p}, 3), \qquad (4.1)$$

where $\hat{p}$ is as defined in Theorem 2.4. Then τ_N^* continuously maps U into $W_D^{-1,p}(\Omega; \mathbb{R}^d)$ according to Sobolev's embedding theorem. For simplicity we denote τ_N^* with range in $W_D^{-1,p}(\Omega; \mathbb{R}^d)$ by the same symbol.

Remark 4.2. *Due to the continuity of τ_N^* the results of Section 3 imply*

(i) $\left\| \mathbf{\Sigma} - \bar{\mathbf{\Sigma}} \right\|_{L^p(\Omega; \mathbb{S}^2)} \leq L \left\| \tau_N^* \right\| \left\| \boldsymbol{g} - \bar{\boldsymbol{g}} \right\|_U$

(ii) $\left\| \lambda - \bar{\lambda} \right\|_{L^q(\Omega)} \leq c \left\| \tau_N^* \right\| \left\| \boldsymbol{g} - \bar{\boldsymbol{g}} \right\|_U,$

(iii) $\left\| \mathbf{\Sigma}' \right\|_{L^p(\Omega; \mathbb{S}^2)} \leq L \left\| \tau_N^* \right\| \left\| \boldsymbol{g} - \bar{\boldsymbol{g}} \right\|_U,$

(iv) $\left\| \mathbf{\Sigma} - \bar{\mathbf{\Sigma}} - \mathbf{\Sigma}' \right\|_{S^2} + \left\| \boldsymbol{u} - \bar{\boldsymbol{u}} - \boldsymbol{u}' \right\|_V = o\left(\left\| \boldsymbol{g} - \bar{\boldsymbol{g}} \right\|_U \right),$

(v) $\left\| \lambda - \bar{\lambda} - \lambda' \right\|_{L^\alpha(\Omega)} = o\left(\left\| \boldsymbol{g} - \bar{\boldsymbol{g}} \right\|_U \right),$

(vi) $\left\| \boldsymbol{u}' \right\|_V \leq c \left\| \tau_N^* \right\| \left\| \boldsymbol{g} - \bar{\boldsymbol{g}} \right\|_U$

with $\left\| \tau_N^* \right\| := \left\| \tau_N^* \right\|_{\mathcal{L}(U, W_D^{-1,p}(\Omega; \mathbb{R}^d))}.$

The next lemma covers the Lipschitz continuity of the function ϕ in the yield condition (1.1) on the admissble set $\mathcal{K}$:

Lemma 4.3. *Let $\nu \geq 1$ and $\mathbf{\Sigma}_1, \mathbf{\Sigma}_2 \in \mathcal{K}$. Then*

$$\left\| \phi(\mathbf{\Sigma}_1) - \phi(\mathbf{\Sigma}_2) \right\|_{L^\nu(\Omega)} \leq \sigma_0 \left\| \mathbf{\Sigma}_1 - \mathbf{\Sigma}_2 \right\|_{L^\nu(\Omega; \mathbb{S}^2)}.$$

holds true.

Proof. By definition of $\mathcal{K}$ we find

$$\left\| \phi(\mathbf{\Sigma}_1) - \phi(\mathbf{\Sigma}_2) \right\|_{L^\nu(\Omega)}^\nu = \int_\Omega \left| \frac{1}{2} (\mathcal{D}\mathbf{\Sigma}_2 + \mathcal{D}\mathbf{\Sigma}_1) : (\mathcal{D}\mathbf{\Sigma}_1 - \mathcal{D}\mathbf{\Sigma}_2) \right|^\nu \, \mathrm{d}x$$
$$\leq \sigma_0^\nu \left\| \mathbf{\Sigma}_1 - \mathbf{\Sigma}_2 \right\|_{L^\nu(\Omega; \mathbb{S}^2)}^\nu.$$

$\square$

Next we define a Lagrange functional which is especially suited for the discussion of (**P**). For this purpose let us introduce the space

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

Endowed with the norm $\left\| \boldsymbol{T} \right\|_{S^2} + \left\| \mathcal{D}\boldsymbol{T} \right\|_{L^\infty(\Omega; \mathbb{S})}$, S_∞^2 is a Banach space. Obviously every solution of (3.1) and (3.7), respectively, belongs to S_∞^2. With this space at hand the Lagrangian $\mathcal{L} \colon V \times S_\infty^2 \times L^2(\Omega) \times U \times S^2 \times V \times L^2(\Omega) \times L^\infty(\Omega)' \to \mathbb{R}$ is then defined by

$$\mathcal{L}(\boldsymbol{u}, \mathbf{\Sigma}, \lambda, \boldsymbol{g}, \mathbf{\Upsilon}, \boldsymbol{w}, \mu, \theta) := J(\boldsymbol{u}, \boldsymbol{g}) + (A\mathbf{\Sigma} + B^*\boldsymbol{u} + \lambda \mathcal{D}^* \mathcal{D}\mathbf{\Sigma}, \mathbf{\Upsilon})$$
$$+ \langle B\mathbf{\Sigma} + \tau_N^*(\boldsymbol{g}), \boldsymbol{w} \rangle - (\lambda, \mu) + \langle \phi(\mathbf{\Sigma}), \theta \rangle_{L^\infty(\Omega), L^\infty(\Omega)'}.$$

Obviously, $\mathcal{L}$ is twice continuously Fréchet-differentiable by Assumption 4.1. Note that we do not introduce a Lagrange multiplier associated with the complementarity relation $\lambda \phi(\mathbf{\Sigma}) = 0$ a.e. in Ω, which is typical for MPECs, cf. for instance the Lagrangian defined in [23]. Our main result then reads as follows:

Theorem 4.4. *Let Assumption 4.1 hold. Suppose further that $\bar{\boldsymbol{g}} \in U$ together with its state $(\bar{\mathbf{\Sigma}}, \bar{\boldsymbol{u}}, \bar{\lambda}) \in S^2 \times V \times L^2(\Omega)$, adjoint state $(\bar{\mathbf{\Upsilon}}, \bar{\boldsymbol{w}}) \in S^2 \times V$, and multipliers $(\bar{\mu}, \bar{\theta}) \in L^2(\Omega) \times L^2(\Omega)$ satisfy*

(1) *the following regularity conditions:*

$$\bar{\boldsymbol{\chi}} \in L^s(\Omega; \mathbb{S}), \quad \bar{\boldsymbol{\Upsilon}} \in L^{\eta}(\Omega; \mathbb{S}^2), \quad \bar{\mu} \in L^{\zeta}(\Omega), \quad \bar{\theta} \in L^{\vartheta}(\Omega)$$

with

$$s > \frac{2p}{p-2}, \quad \eta \geq \frac{2sp}{sp - 2p - 2s}, \quad \zeta > \frac{sp}{sp - p - 2s},$$

$$\vartheta > \max\left(\frac{sp}{sp - p - 2s}, \frac{2p}{p-2}\right), \tag{4.2}$$

where p is as defined in (4.1),

(2) *the following optimality system:*

$$A\bar{\boldsymbol{\Upsilon}} + B^*\bar{\boldsymbol{w}} + \bar{\lambda}\mathcal{D}^*\mathcal{D}\bar{\boldsymbol{\Upsilon}} + \bar{\theta}\mathcal{D}^*\mathcal{D}\bar{\boldsymbol{\Sigma}} = 0 \tag{4.3a}$$

$$B\bar{\boldsymbol{\Upsilon}} + \partial_{\boldsymbol{u}}J(\bar{\boldsymbol{u}}, \bar{\boldsymbol{g}}) = 0 \tag{4.3b}$$

$$\tau_N\bar{\boldsymbol{w}} + \partial_{\boldsymbol{g}}J(\bar{\boldsymbol{u}}, \bar{\boldsymbol{g}}) = 0 \tag{4.3c}$$

$$\mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\bar{\boldsymbol{\Upsilon}} = \bar{\mu} \tag{4.3d}$$

$$\bar{\mu}\bar{\lambda} = 0 \qquad a.e.\ in\ \Omega \tag{4.3e}$$

$$\bar{\theta}\phi(\bar{\boldsymbol{\Sigma}}) = 0 \qquad a.e.\ in\ \Omega \tag{4.3f}$$

$$\bar{\mu} \geq 0 \qquad a.e.\ in\ \mathcal{A}_1 \tag{4.3g}$$

$$\bar{\theta} \geq 0 \qquad a.e.\ in\ \mathcal{A}_2 \tag{4.3h}$$

with

$$\mathcal{A}_1 := \{x \in \Omega : -\tau_1 \leq \phi(\bar{\boldsymbol{\Sigma}}) \leq 0\}, \quad \mathcal{A}_2 := \{x \in \Omega : 0 \leq \bar{\lambda} \leq \tau_2\}$$

for some $\tau_1, \tau_2 > 0$,

(3) *the second-order condition:*
there exists $\bar{\alpha} > 0$ such that

$$\left. \begin{aligned} \partial^2_{(\boldsymbol{u}, \boldsymbol{\Sigma}, \lambda, \boldsymbol{g})} \mathcal{L}(\bar{\boldsymbol{u}}, \bar{\boldsymbol{\Sigma}}, \bar{\lambda}, \bar{\boldsymbol{g}}, \bar{\boldsymbol{\Upsilon}}, \bar{\boldsymbol{w}}, \bar{\mu}, \bar{\theta})(\boldsymbol{\Sigma}', \boldsymbol{u}', \lambda', \boldsymbol{h})^2 \\ \geq \bar{\alpha} \|\boldsymbol{h}\|_U^2 \quad \forall \boldsymbol{h} \in U, \end{aligned} \right\} \tag{SSC}$$

where $(\boldsymbol{\Sigma}', \boldsymbol{u}', \lambda')$ solves (3.4) with $\delta\ell = -\tau_N^*(\boldsymbol{h})$.

Then there exists an $\epsilon > 0$ such that the following quadratic growth condition

$$J(\boldsymbol{u}, \boldsymbol{g}) \geq J(\bar{\boldsymbol{u}}, \bar{\boldsymbol{g}}) + \frac{\bar{\alpha}}{2} \|\boldsymbol{g} - \bar{\boldsymbol{g}}\|_U^2 \tag{4.4}$$

is satisfied for all $\boldsymbol{g} \in U$ with $\|\boldsymbol{g} - \bar{\boldsymbol{g}}\|_U \leq \epsilon$. Thus $\bar{\boldsymbol{g}}$ is a strict local optimum of **(P)**.

Remark 4.5. *Let us compare these sufficient optimality conditions with the necessary ones. In [13, Section 3.3] so-called C-stationarity conditions are proven to be satisfied by every local optimum. These conditions coincide with (4.3), except that they only provide a sign condition for the product of the multipliers $\bar{\mu}$ and $\bar{\theta}$, whereas (4.3g) and (4.3h) contain sign conditions for each multiplier individually. More restrictive optimality conditions for MPECs are given by strong stationarity. To the best of our knowledge, it has not been proven so far that these conditions are necessary for local optimality for problems of type* **(P)**. *Only in the rather academic case of an additional control appearing as inhomogeneity in (2.1a) Herzog et al. proved strong stationarity conditions to be necessary in [14]. These conditions also contain sign conditions for each multiplier individually, but only on the strongly active set $\bar{\mathcal{A}}_s$ and the inactive set $\bar{\mathcal{I}}$, respectively. Hence the sign conditions in (4.3g) and (4.3h) are even more restrictive than strong stationarity. Moreover, higher regularity of the optimal hardening variable $\bar{\boldsymbol{\chi}}$ and the multipliers $\bar{\boldsymbol{\Upsilon}}$, $\bar{\mu}$, and*

$\bar{\theta}$ is required in (4.2). In summary, there is thus a significant gap between the necessary and sufficient optimality conditions for (**P**).

Remark 4.6. *A comparison to the second-order sufficient conditions for finite dimensional MPECs in [23] shows that the assertion of Theorem 4.4 represents a natural generalization of the finite dimensional result to a function space setting. This shows that the conditions in (4.3) and* (SSC) *are not as restrictive as indicated by Remark 4.5.*

Proof of Theorem 4.4. At first we note that Assumption 3.3 is satisfied because of $s > \frac{2p}{p-2}$ and (4.1). Let $\boldsymbol{g} \in U$ be arbitrary and $(\boldsymbol{u}, \boldsymbol{\Sigma}, \lambda)$ be the state and multiplier associated with $\boldsymbol{g}$.

We aim to deduce the quadratic growth condition (4.4) from a Taylor expansion of the Lagrangian. To this end we introduce the abbreviations $\boldsymbol{z} := (\boldsymbol{u}, \boldsymbol{\Sigma}, \lambda, \boldsymbol{g})$, $\bar{\boldsymbol{z}} := (\bar{\boldsymbol{u}}, \bar{\boldsymbol{\Sigma}}, \bar{\lambda}, \bar{\boldsymbol{g}})$, and $\bar{\boldsymbol{z}}' := (\bar{\boldsymbol{\Upsilon}}, \bar{\boldsymbol{w}}, \bar{\mu}, \bar{\theta})$ and consider

$$\mathcal{L}(\boldsymbol{z}, \bar{\boldsymbol{z}}') = \mathcal{L}(\bar{\boldsymbol{z}}, \bar{\boldsymbol{z}}') + \nabla_{\boldsymbol{z}} \mathcal{L}(\bar{\boldsymbol{z}}, \bar{\boldsymbol{z}}')(\boldsymbol{z} - \bar{\boldsymbol{z}}) + \frac{1}{2} \nabla_{\boldsymbol{z}}^2 \mathcal{L}(\bar{\boldsymbol{z}} + t(\boldsymbol{z} - \bar{\boldsymbol{z}}), \bar{\boldsymbol{z}}')(\boldsymbol{z} - \bar{\boldsymbol{z}})^2 \quad (4.5)$$

with $t \in [0, 1]$. In the following we discuss each expression of (4.5) separately.

The zero-order terms:
Due to (3.1a), (3.1b), (4.3e) and (4.3f) it holds

$$\mathcal{L}(\bar{\boldsymbol{z}}, \bar{\boldsymbol{z}}') = J(\bar{\boldsymbol{u}}, \bar{\boldsymbol{g}}). \quad (4.6)$$

and

$$\mathcal{L}(\boldsymbol{z}, \bar{\boldsymbol{z}}') = J(\boldsymbol{u}, \boldsymbol{g}) - (\lambda,\, \bar{\mu}) + \left(\phi(\boldsymbol{\Sigma}),\, \bar{\theta}\right). \quad (4.7)$$

Defining $\Omega_1 := \{x \in \Omega : \lambda(x) > 0\}$ and $E_1 := \Omega_1 \cap \mathcal{A}_1$, we see

$$-\int_\Omega \lambda \bar{\mu}\, \mathrm{d}x = -\int_{E_1} \underbrace{\lambda}_{>0} \underbrace{\bar{\mu}}_{\geq 0}\, \mathrm{d}x - \int_{\Omega_1 \setminus E_1} \lambda \bar{\mu}\, \mathrm{d}x \leq -\int_{\Omega_1 \setminus E_1} (\lambda - \bar{\lambda}) \bar{\mu}\, \mathrm{d}x$$

$$\leq \left\| \lambda - \bar{\lambda} \right\|_{L^q(\Omega)} \left\| \bar{\mu} \right\|_{L^\varsigma(\Omega)} |\Omega_1 \setminus E_1|^{\frac{1}{\beta}}.$$

Here we used that $\bar{\lambda} = 0$ a.e. in $\Omega_1 \setminus E_1$ due to (3.1c) and that there exists $\beta < p$ with $\frac{1}{q} + \frac{1}{\varsigma} + \frac{1}{\beta} = 1$ in consequence of (4.2), where q is as defined in (3.6). Furthermore (3.7c) and the definition of Ω_1 and $\mathcal{A}_1$ imply

$$\begin{aligned} |\Omega_1 \setminus E_1| &= \frac{1}{\tau_1^p} \int_{\Omega_1 \setminus E_1} \tau_1^p\, \mathrm{d}x \\ &< \frac{1}{\tau_1^p} \int_{\Omega_1 \setminus E_1} \left| \phi(\bar{\boldsymbol{\Sigma}}) \right|^p\, \mathrm{d}x = \frac{1}{\tau_1^p} \int_{\Omega_1 \setminus E_1} \left| \phi(\boldsymbol{\Sigma}) - \phi(\bar{\boldsymbol{\Sigma}}) \right|^p\, \mathrm{d}x. \end{aligned} \quad (4.8)$$

Thus, thanks to Lemma 4.3 and Remark 4.2, we arrive at

$$\begin{aligned} -\int_\Omega \lambda \bar{\mu}\, \mathrm{d}x &< c \left\| \lambda - \bar{\lambda} \right\|_{L^q(\Omega)} \left\| \bar{\mu} \right\|_{L^\varsigma(\Omega)} \left\| \phi(\boldsymbol{\Sigma}) - \phi(\bar{\boldsymbol{\Sigma}}) \right\|_{L^p(\Omega)}^{\frac{p}{\beta}} \\ &\leq c \left\| \lambda - \bar{\lambda} \right\|_{L^q(\Omega)} \left\| \bar{\mu} \right\|_{L^\varsigma(\Omega)} \left\| \boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}} \right\|_{L^p(\Omega;\mathbb{S}^2)}^{\frac{p}{\beta}} \leq c \left\| \boldsymbol{g} - \bar{\boldsymbol{g}} \right\|_U^{1+\frac{p}{\beta}}. \end{aligned} \quad (4.9)$$

Similarly we define $\Omega_2 := \{x \in \Omega : \phi(\boldsymbol{\Sigma}(x)) < 0\}$ and $E_2 := \Omega_2 \cap \mathcal{A}_2$. Analogously to above, one finds

$$\begin{aligned} \int_\Omega \phi(\boldsymbol{\Sigma}) \bar{\theta}\, \mathrm{d}x &= \int_{E_2} \underbrace{\phi(\boldsymbol{\Sigma})}_{<0} \underbrace{\bar{\theta}}_{\geq 0}\, \mathrm{d}x + \int_{\Omega_2 \setminus E_2} \phi(\boldsymbol{\Sigma}) \bar{\theta}\, \mathrm{d}x \\ &\leq \int_{\Omega_2 \setminus E_2} (\phi(\boldsymbol{\Sigma}) - \phi(\bar{\boldsymbol{\Sigma}})) \bar{\theta}\, \mathrm{d}x \\ &\leq \left\| \phi(\boldsymbol{\Sigma}) - \phi(\bar{\boldsymbol{\Sigma}}) \right\|_{L^p(\Omega)} \left\| \bar{\theta} \right\|_{L^\vartheta(\Omega)} |\Omega_2 \setminus E_2|^{\frac{1}{\gamma}}, \end{aligned}$$

where we used that $\phi(\bar{\boldsymbol{\Sigma}}) = 0$ in $\Omega_2 \setminus E_2$ because of (3.1c). Moreover, the condition on ϑ in (4.2) implies the existence of $\gamma < q$ with q as defined in (3.6) so that $\frac{1}{p} + \frac{1}{\vartheta} + \frac{1}{\gamma} = 1$. As above we deduce from (3.7c) and the definition of Ω_2 and $\mathcal{A}_2$

$$
\begin{aligned}
|\Omega_2 \setminus E_2| &= \frac{1}{\tau_2^q} \int_{\Omega_2 \setminus E_2} \tau_2^q \, \mathrm{d}x \\
&< \frac{1}{\tau_2^q} \int_{\Omega_2 \setminus E_2} \left|\bar{\lambda}\right|^q \, \mathrm{d}x = \frac{1}{\tau_2^q} \int_{\Omega_{0,2} \setminus \bar{\mathcal{A}}_0^{\tau_2}} \left|\lambda - \bar{\lambda}\right|^q \, \mathrm{d}x.
\end{aligned} \tag{4.10}
$$

Because of Lemma 4.3 and Remark 4.2 we thus obtain

$$
\begin{aligned}
\int_\Omega \phi(\boldsymbol{\Sigma}) \bar{\theta} \, \mathrm{d}x &< c \left\|\phi(\boldsymbol{\Sigma}) - \phi(\bar{\boldsymbol{\Sigma}})\right\|_{L^p(\Omega)} \left\|\bar{\theta}\right\|_{L^\vartheta(\Omega)} \left\|\lambda - \bar{\lambda}\right\|_{L^q(\Omega)}^{\frac{q}{\gamma}} \\
&\leq c \left\|\boldsymbol{g} - \bar{\boldsymbol{g}}\right\|_U^{1 + \frac{q}{\gamma}}.
\end{aligned} \tag{4.11}
$$

Due to $1 + p/\beta > 2$ and $1 + q/\gamma > 2$, cf. (4.2) and (3.6), (4.7)–(4.11) show

$$
\mathcal{L}(\boldsymbol{z}, \bar{\boldsymbol{z}}') \leq J(\boldsymbol{u}, \boldsymbol{g}) + o\left(\left\|\boldsymbol{g} - \bar{\boldsymbol{g}}\right\|_U^2\right). \tag{4.12}
$$

The first-order term:
From (4.3a)–(4.3d) it follows

$$
\begin{aligned}
&\nabla_{\boldsymbol{z}} \mathcal{L}(\bar{\boldsymbol{z}}, \bar{\boldsymbol{z}}')(\boldsymbol{z} - \bar{\boldsymbol{z}}) = \\
&= \underbrace{\left(A\bar{\boldsymbol{\Upsilon}} + B^*\bar{\boldsymbol{w}} + \bar{\lambda}\mathcal{D}^*\mathcal{D}\bar{\boldsymbol{\Upsilon}} + \bar{\theta}\mathcal{D}^*\mathcal{D}\bar{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}}\right)}_{=0} + \underbrace{\left\langle B\bar{\boldsymbol{\Upsilon}} + \partial_{\boldsymbol{u}}J(\bar{\boldsymbol{u}}, \bar{\boldsymbol{g}}), \boldsymbol{u} - \bar{\boldsymbol{u}}\right\rangle}_{=0} \\
&\quad + \underbrace{\left(\partial_{\boldsymbol{g}}J(\bar{\boldsymbol{u}}, \bar{\boldsymbol{g}}) + \tau_N\bar{\boldsymbol{w}}, \boldsymbol{g} - \bar{\boldsymbol{g}}\right)}_{=0} - \left(\lambda - \bar{\lambda}, \bar{\mu}\right) + \underbrace{\left(\lambda - \bar{\lambda}, \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\bar{\boldsymbol{\Upsilon}}\right)}_{=\left(\lambda - \bar{\lambda}, \bar{\mu}\right)} = 0.
\end{aligned} \tag{4.13}
$$

The second-order term:
The second derivative of $\mathcal{L}$ at $\boldsymbol{z}_t := \bar{\boldsymbol{z}} + t(\boldsymbol{z} - \bar{\boldsymbol{z}})$ in direction $(\boldsymbol{z} - \bar{\boldsymbol{z}})^2$ is given by

$$
\begin{aligned}
&\nabla_{\boldsymbol{z}}^2 \mathcal{L}(\boldsymbol{z}_t, \bar{\boldsymbol{z}}')(\boldsymbol{z} - \bar{\boldsymbol{z}})^2 = \\
&= \nabla_{(\boldsymbol{u}, \boldsymbol{g})}^2 J(\boldsymbol{u}_t, \boldsymbol{g}_t)(\boldsymbol{u} - \bar{\boldsymbol{u}}, \boldsymbol{g} - \bar{\boldsymbol{g}})^2 + 2\left((\lambda - \bar{\lambda})\mathcal{D}^*\mathcal{D}(\boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}}), \bar{\boldsymbol{\Upsilon}}\right) \\
&\quad + \left(\left\|\mathcal{D}\boldsymbol{\Sigma} - \mathcal{D}\bar{\boldsymbol{\Sigma}}\right\|_{\mathbb{S}}^2, \bar{\theta}\right) \\
&= \underbrace{\nabla_{(\boldsymbol{u}, \boldsymbol{g})}^2 J(\bar{\boldsymbol{u}}, \bar{\boldsymbol{g}})(\boldsymbol{u}', \boldsymbol{g} - \bar{\boldsymbol{g}})^2}_{=:D_1} + \left(\nabla_{(\boldsymbol{u}, \boldsymbol{g})}^2 J(\boldsymbol{u}_t, \boldsymbol{g}_t) - \nabla_{(\boldsymbol{u}, \boldsymbol{g})}^2 J(\bar{\boldsymbol{u}}, \bar{\boldsymbol{g}})\right)(\boldsymbol{u} - \bar{\boldsymbol{u}}, \boldsymbol{g} - \bar{\boldsymbol{g}})^2 \\
&\quad + 2\partial_{\boldsymbol{g}}\partial_{\boldsymbol{u}}J(\bar{\boldsymbol{u}}, \bar{\boldsymbol{g}})[\boldsymbol{u} - \bar{\boldsymbol{u}} - \boldsymbol{u}', \boldsymbol{g} - \bar{\boldsymbol{g}}] + \partial_{\boldsymbol{u}}^2 J(\bar{\boldsymbol{u}}, \bar{\boldsymbol{g}})(\boldsymbol{u} - \bar{\boldsymbol{u}} - \boldsymbol{u}')^2 \\
&\quad + 2\partial_{\boldsymbol{u}}^2 J(\bar{\boldsymbol{u}}, \bar{\boldsymbol{g}})[\boldsymbol{u} - \bar{\boldsymbol{u}} - \boldsymbol{u}', \boldsymbol{u}'] + 2\left((\lambda - \bar{\lambda})\mathcal{D}^*\mathcal{D}(\boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}}), \bar{\boldsymbol{\Upsilon}}\right) \\
&\quad + \left(\left\|\mathcal{D}\boldsymbol{\Sigma} - \mathcal{D}\bar{\boldsymbol{\Sigma}}\right\|_{\mathbb{S}}^2, \bar{\theta}\right).
\end{aligned}
$$

The regularity condition (4.2) implies $\frac{1}{q} + \frac{1}{2} + \frac{1}{\eta} = \frac{1}{\alpha} + \frac{1}{p} + \frac{1}{\eta} = \frac{1}{s} + \frac{1}{2} + \frac{1}{p} + \frac{1}{\eta} \leq 1$, cf. also (3.6) and Corollary 3.11. Hence due to Remark 4.2 (i)–(iii) it holds

$$
\begin{aligned}
&\left((\lambda - \bar{\lambda})(\mathcal{D}\boldsymbol{\Sigma} - \mathcal{D}\bar{\boldsymbol{\Sigma}}), \mathcal{D}\bar{\boldsymbol{\Upsilon}}\right) = \\
&= \underbrace{\left(\lambda'\mathcal{D}\boldsymbol{\Sigma}', \mathcal{D}\bar{\boldsymbol{\Upsilon}}\right)}_{=:D_2} + \left((\lambda - \bar{\lambda})(\mathcal{D}\boldsymbol{\Sigma} - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}'), \mathcal{D}\bar{\boldsymbol{\Upsilon}}\right) + \left((\lambda - \bar{\lambda} - \lambda')\mathcal{D}\boldsymbol{\Sigma}', \mathcal{D}\bar{\boldsymbol{\Upsilon}}\right) \\
&\geq \left(\lambda'\mathcal{D}\boldsymbol{\Sigma}', \mathcal{D}\bar{\boldsymbol{\Upsilon}}\right) - c\left\|\lambda - \bar{\lambda}\right\|_{L^q(\Omega)} \left\|\boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'\right\|_{S^2} \left\|\bar{\boldsymbol{\Upsilon}}\right\|_{L^\eta(\Omega;\mathbb{S}^2)} \\
&\quad - c\left\|\lambda - \bar{\lambda} - \lambda'\right\|_{L^\alpha(\Omega)} \left\|\boldsymbol{\Sigma}'\right\|_{L^p(\Omega;\mathbb{S}^2)} \left\|\bar{\boldsymbol{\Upsilon}}\right\|_{L^\eta(\Omega;\mathbb{S}^2)} \\
&\geq \left(\lambda'\mathcal{D}\boldsymbol{\Sigma}', \mathcal{D}\bar{\boldsymbol{\Upsilon}}\right) - c\left\|\boldsymbol{g} - \bar{\boldsymbol{g}}\right\|_U \left(\left\|\boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'\right\|_{S^2} + \left\|\lambda - \bar{\lambda} - \lambda'\right\|_{L^\alpha(\Omega)}\right).
\end{aligned}
$$

Furthermore, (4.2) yields $\frac{1}{p} + \frac{1}{2} + \frac{1}{\vartheta} < 1$ and consequently

$$
\begin{aligned}
&\left(\left\| \mathcal{D}\boldsymbol{\Sigma} - \mathcal{D}\bar{\boldsymbol{\Sigma}} \right\|_{\mathbb{S}}^2, \bar{\theta} \right) = \\
&= \underbrace{\left(\left\| \mathcal{D}\boldsymbol{\Sigma}' \right\|_{\mathbb{S}}^2, \bar{\theta} \right)}_{=:D_3} + \left(\left\| \mathcal{D}\boldsymbol{\Sigma} - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}' \right\|_{\mathbb{S}}^2, \bar{\theta} \right) + 2 \left((\mathcal{D}\boldsymbol{\Sigma} - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}') : \mathcal{D}\boldsymbol{\Sigma}', \bar{\theta} \right) \\
&\geq \left(\left\| \mathcal{D}\boldsymbol{\Sigma}' \right\|_{\mathbb{S}}^2, \bar{\theta} \right) - c \left\| \boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}' \right\|_{L^p(\Omega;\mathbb{S})} \left\| \boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}' \right\|_{S^2} \left\| \bar{\theta} \right\|_{L^\vartheta(\Omega)} \\
&\quad - c \left\| \boldsymbol{\Sigma}' \right\|_{L^p(\Omega;\mathbb{S})} \left\| \boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}' \right\|_{S^2} \left\| \bar{\theta} \right\|_{L^\vartheta(\Omega)} \\
&\geq \left(\left\| \mathcal{D}\boldsymbol{\Sigma}' \right\|_{\mathbb{S}}^2, \bar{\theta} \right) - c \left\| \boldsymbol{g} - \bar{\boldsymbol{g}} \right\|_U \left\| \boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}' \right\|_{S^2}.
\end{aligned}
$$

Note that $D_1 + 2D_2 + D_3 = \nabla_{\boldsymbol{z}}^2 \mathcal{L}(\bar{\boldsymbol{z}}, \bar{\boldsymbol{z}}')(\boldsymbol{\Sigma}', \boldsymbol{u}', \lambda', \boldsymbol{g} - \bar{\boldsymbol{g}})^2$. Thus, because of (SSC), Remark 4.2, Theorem 2.4 and Theorem 3.7, one obtains

$$
\begin{aligned}
&\nabla_{\boldsymbol{z}}^2 \mathcal{L}(\boldsymbol{z}_t, \bar{\boldsymbol{z}}')(\boldsymbol{z} - \bar{\boldsymbol{z}})^2 = \\
&\geq \nabla_{\boldsymbol{z}}^2 \mathcal{L}(\bar{\boldsymbol{z}}, \bar{\boldsymbol{z}}')(\boldsymbol{\Sigma}', \boldsymbol{u}', \lambda', \boldsymbol{g} - \bar{\boldsymbol{g}})^2 \\
&\quad - c \left\| \boldsymbol{g} - \bar{\boldsymbol{g}} \right\|_U \left(\left\| \boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}' \right\|_{S^2} + \left\| \lambda - \bar{\lambda} - \lambda' \right\|_{L^\alpha(\Omega)} \right) \\
&\quad - c \left\| \nabla_{(\boldsymbol{u},\boldsymbol{g})}^2 J(\boldsymbol{u}_t, \boldsymbol{g}_t) - \nabla_{(\boldsymbol{u},\boldsymbol{g})}^2 J(\bar{\boldsymbol{u}}, \bar{\boldsymbol{g}}) \right\|_{\mathcal{L}(V \times U, (V \times U)')} \left\| \boldsymbol{g} - \bar{\boldsymbol{g}} \right\|_U^2 \\
&\quad - c \left\| \partial_{\boldsymbol{g}} \partial_{\boldsymbol{u}} J(\bar{\boldsymbol{u}}, \bar{\boldsymbol{g}}) \right\|_{\mathcal{L}(U, V')} \left\| \boldsymbol{u} - \bar{\boldsymbol{u}} - \boldsymbol{u}' \right\|_V \left\| \boldsymbol{g} - \bar{\boldsymbol{g}} \right\|_U \\
&\quad - c \left\| \partial_{\boldsymbol{u}}^2 J(\bar{\boldsymbol{u}}, \bar{\boldsymbol{g}}) \right\|_{\mathcal{L}(V, V')} \left\| \boldsymbol{u} - \bar{\boldsymbol{u}} - \boldsymbol{u}' \right\|_V^2 \\
&\quad - c \left\| \partial_{\boldsymbol{u}}^2 J(\bar{\boldsymbol{u}}, \bar{\boldsymbol{g}}) \right\|_{\mathcal{L}(V, V')} \left\| \boldsymbol{u} - \bar{\boldsymbol{u}} - \boldsymbol{u}' \right\|_V \left\| \boldsymbol{u}' \right\|_V \\
&\geq \bar{\alpha} \left\| \boldsymbol{g} - \bar{\boldsymbol{g}} \right\|_U^2 - o \left(\left\| \boldsymbol{g} - \bar{\boldsymbol{g}} \right\|_U^2 \right),
\end{aligned}
\tag{4.14}
$$

where we used that J is twice continuously differentiable by assumption.

Altogether (4.5), (4.6) and (4.12)–(4.14) yield the existence of an $\epsilon > 0$ so that $J(\boldsymbol{u}, \boldsymbol{g}) \geq J(\bar{\boldsymbol{u}}, \bar{\boldsymbol{g}}) + \frac{\bar{\alpha}}{4} \left\| \boldsymbol{g} - \bar{\boldsymbol{g}} \right\|_U^2$ for all $\boldsymbol{g} \in U$ with $\left\| \boldsymbol{g} - \bar{\boldsymbol{g}} \right\|_U \leq \epsilon$, i.e., the desired quadratic growth condition. $\qquad\square$

Remark 4.7. *It is easily seen that, for every $p > 2$, there are numbers $s, \eta, \zeta, \vartheta \in [2, \infty[$ satisfying the conditions in (4.2). However, if p tends to 2, then the bounds for s, η, ζ, and ϑ tend to ∞.*

Remark 4.8. *By using the technique introduced in [6], it should be possible to include box constraints on the control into the above analysis and establish second-order sufficient conditions accounting for strongly active sets. However, in this case, a two-norm discrepancy will arise. As we wish to keep the discussion concise and thus focus on the difficulties induced by the VI, we omitted additional control constraints here.*

A Auxiliary results

Lemma A.1. *Let X be a Hilbert space and $x_1, x_2 \in X$. If $\|x_1\|_X = \|x_2\|_X$, then it holds*

$$
(x_1, x_1 - x_2)_X = \frac{1}{2} \|x_1 - x_2\|_X^2.
$$

Proof. The assertion directly follows from straight forward computation:

$$(x_1,\, x_1 - x_2)_X = \|x_1\|_X^2 - \frac{1}{2}\left(\|x_1\|_X^2 + \|x_2\|_X^2 - \|x_1 - x_2\|_X^2\right)$$
$$= \frac{1}{2}\,\|x_1 - x_2\|_X^2,$$

where we used $\|x_1\|_X = \|x_2\|_X$ for the last equation. □

Lemma A.2. *Let $E \subset \mathbb{R}^d$ be measurable and bounded, $\nu \in (1, \infty)$ and $f, f_n \in L^\nu(E)$, $n \in \mathbb{N}$. If $\sup_{n \in \mathbb{N}} \|f_n\|_{L^\nu(E)} \leq c$ and $f_n \to f$ a.e. in E, then $f_n \to f$ in $L^\mu(E)$ for $1 \leq \mu < \nu$.*

Proof. We define $g_n := |f_n - f|^\mu$. Obviously, g_n converges pointwise to zero a.e. in E and $g_n \in L^{\frac{\nu}{\mu}}(E)$. Moreover g_n is bounded and thus there exists a subsequence converging weakly in $L^{\frac{\nu}{\mu}}(E)$. Due to Egorov's Theorem the weak limit equals the pointwise limit. Thus the weak limit is unique, which implies weak convergence of the whole sequence g_n to zero. Consequently, $\int_\Omega g_n \, dx \to 0$, which implies the assertion. □

References

[1] V. Barbu. *Optimal Control of Variational Inequalities*, volume 100 of *Research Notes in Mathematics*. Pitman, Boston, 1984.

[2] 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.

[3] F. Bonnans and D. Tiba. Pontryagin's principle in the control of semilinear elliptic variational inequalities. *Applied Mathematics and Optimization*, 23:299–312, 1991.

[4] E. Casas, J. C. de los Reyes, and F. Tröltzsch. Sufficient second-order optimality conditions for semilinear control problems with pointwise state constraints. *SIAM Journal on Optimization*, 19(2):616–643, 2008.

[5] E. Casas and F. Tröltzsch. First- and second order optimality conditions for a class of optimal control problems with quasilinear ellitpic equations. *SIAM Journal on Control and Optimization*, 48(2):688–718, 2009.

[6] E. Casas, F. Tröltzsch, and A. Unger. Second order sufficient optimality conditions for a nonlinear elliptic control problem. *Zeitschrift für Analysis und ihre Anwendungen*, 15:687–707, 1996.

[7] P. Grisvard. *Elliptic Problems in Nonsmooth Domains*. Pitman, Boston, 1985.

[8] K. Gröger. A $W^{1,p}$-estimate for solutions to mixed boundary value problems for second order elliptic differential equations. *Mathematische Annalen*, 283:679–687, 1989.

[9] R. Haller-Dintelmann, C. Meyer, J. Rehberg, and A. Schiela. Hölder continuity and optimal control for nonsmooth elliptic problems. *Applied Mathematics and Optimization*, 60(3):397–428, 2009.

[10] W. Han and B. D. Reddy. *Plasticity*. Springer, New York, 1999.

[11] 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.

[12] R. Herzog, C. Meyer, and G. Wachsmuth. Integrability of displacement and stresses in linear and nonlinear elasticity with mixed boundary conditions. *Journal of Mathematical Analysis and Applications*, 382(2):802–813, 2011.

[13] R. Herzog, C. Meyer, and G. Wachsmuth. C-stationarity for optimal control of static plasticity with linear kinematic hardening. *SIAM Journal on Control and Optimization*, 50(5):3052–3082, 2012.

[14] R. Herzog, C. Meyer, and G. Wachsmuth. B- and strong stationarity for optimal control of static plasticity with hardening. *SIAM Journal on Optimization*, to appear.

[15] 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.

[16] Michael Hintermüller and Thomas Surowiec. First order optimality conditions for elliptic mathematical programs with equilibrium constraints via variational analysis. *SIAM Journal on Optimization*, 21(4):1561–1593, 2012.

[17] K. Ito and K. Kunisch. Optimal control of elliptic variational inequalities. *Applied Mathematics and Optimization*, 41:343–364, 2000.

[18] 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.

[19] K. Kunisch and D. Wachsmuth. Sufficient optimality conditions and semi-smooth newton methods for optimal control of stationary variational inequalities. *ESAIM: Control, Optimisation and Calculus of Variations*, 2011.

[20] K. Kunisch and D. Wachsmuth. Path-following for optimal control of stationary variational inequalities. *Computational Optimization and Applications*, pages 1–29, 2012.

[21] F. Mignot and J.-P. Puel. Optimal control in some variational inequalities. *SIAM Journal on Control and Optimization*, 22(3):466–476, 1984.

[22] 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.

[23] H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. *Mathematics of Operations Research*, 25(1):1–22, 2000.

[24] G. Wachsmuth. *Optimal control of quasistatic plasticity – An MPCC in function space*. PhD thesis, Chemnitz University of Technology, Germany, 2011.

[25] G. Wachsmuth. Differentiability of implicit functions: Beyond the implicit function theorem. Technical Report SPP1253-137, Priority Program 1253, German Research Foundation, 2012.

TU Dortmund, Faculty of Mathematics, Vogelpothsweg 87, D–44227 Dortmund, Germany

E-mail address: `tbetz@math.tu-dortmund.de`

TU Dortmund, Faculty of Mathematics, Vogelpothsweg 87, D–44227 Dortmund, Germany

E-mail address: `cmeyer@math.tu-dortmund.de`