

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

THOMAS BETZ¹ AND CHRISTIAN MEYER¹

Abstract. The paper is concerned with the optimal control of static elastoplasticity 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.

Mathematics Subject Classification. 49K20, 74C05, 74P10, 35R45.

Received March 22, 2013. Revised April 25, 2014.

Published online December 9, 2014.

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.* [19, 26] 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, 16–18, 23, 25]. In [14, 15] C- and B-stationarity conditions for optimal control of static

Keywords and phrases. Second-order sufficient conditions, optimal control of variational inequalities, bouligand differentiability.

¹ TU Dortmund, Faculty of Mathematics, Vogelpothsweg 87, 44227 Dortmund, Germany.
tbetz@math.tu-dortmund.de; cmeyer@math.tu-dortmund.de

elastoplasticity are derived. Moreover, the necessity of strong stationarity for local optimality is established in [15] 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, 7] 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 contributions in this field are papers by Mignot [22] and by Kunisch and Wachsmuth [20]. Mignot proved that the obstacle control problem is convex if the desired state is behind the obstacle and thus not reachable. Kunisch and Wachsmuth presented sufficient conditions for the optimal control of the obstacle problem in the general case. In a follow-up paper [21] 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.

Furthermore, there are only few differentiability results concerning VIs which are not of obstacle type. For the static elastoplasticity system under consideration Herzog *et al.* showed in [15] that the solution mapping associated to the VI is weakly directionally differentiable. This is however not sufficient for a substantial second-order analysis. Therefore the main step towards second-order sufficient conditions is to improve these differentiability results. To be more precise, we prove that the solution operator is Bouligand differentiable in appropriate spaces under mild assumptions on the data. Once this result is proven, the derivation of second-order conditions is then rather straight forward. In particular the analysis of [20] can be adapted to optimal control of static elastoplasticity.

The outline of the paper is as follows: After fixing the notation and stating the precise problem and the standing assumptions 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 $\mathbb{S} := \mathbb{R}_{\text{sym}}^{d \times d}$ 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 $\boldsymbol{\sigma} : \boldsymbol{\tau} = \sum_{ij} \sigma_{ij} \tau_{ij}$. 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

$$\begin{aligned} W_D^{1,p}(\Omega; \mathbb{R}^d) &:= \{\mathbf{u} \in W^{1,p}(\Omega; \mathbb{R}^d) : \mathbf{u} = \mathbf{0} \text{ on } \Gamma_D\}, \\ V &:= W_D^{1,2}(\Omega; \mathbb{R}^d), \\ W_D^{-1,p}(\Omega; \mathbb{R}^d) &:= (W_D^{1,p'}(\Omega; \mathbb{R}^d))', \\ U &:= L^2(\Gamma_N; \mathbb{R}^d), \\ S &:= L^2(\Omega; \mathbb{S}). \end{aligned}$$

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 $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi}) \in S^2$ and displacement $\mathbf{u} \in V$ so that $\boldsymbol{\Sigma} \in \mathcal{K}$ and

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

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

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

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 $\varepsilon(v) = 1/2(\nabla v + (\nabla v)^T)$ 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.*,

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

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

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

Note that

$$\mathcal{D}^* \mathcal{D}\Sigma = \begin{pmatrix} \mathcal{D}\Sigma \\ \mathcal{D}\Sigma \end{pmatrix}$$

holds. For a detailed introduction into this and other common plasticity models we refer to [11].

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

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

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

Remark 1.1. Due to Sobolev’s embedding theorem the source term ℓ as defined in (P) is not just an element of V' but satisfies $\ell \in W_D^{-1,4}$ for $d = 2$ and $\ell \in W_D^{-1,3}$ for $d = 3$, (*cf.* also Sect. 4).

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

Assumption 1.2.

- (1) The domain $\Omega \subset \mathbb{R}^d, d \in \{2, 3\}$, is a bounded Lipschitz domain in the sense of Chapter 1.2 from [9]. 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], Def. 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.* $\tau : \mathbb{C}^{-1}(x)\sigma = \sigma : \mathbb{C}^{-1}(x)\tau$ and an analogous relation holds for \mathbb{H}^{-1} for all $\sigma, \tau \in \mathbb{S}$.

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 $\partial\Omega \setminus \Gamma_D \cap \Gamma_D$ is finite and no connected component of Γ_D consists of a single point, *cf.* [10]. 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 [10]. We point out that a broad class of non-smooth domains is regular in the sense of Gröger.

The conditions in Assumption 1.2, (3) are for instance fulfilled by isotropic and homogeneous materials, where the compliance tensor is given by

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

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 [11], Sect. 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 ([12], Props. 3.1, 3.2 and Lem. 3.3). *For every $\ell \in V'$, problem (VI) possesses a unique solution $(\boldsymbol{\Sigma}, \mathbf{u}) \in S^2 \times V$.*

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

Definition 2.2. The control-to-state map $V' \ni \ell \mapsto (\boldsymbol{\Sigma}, \mathbf{u}) \in S^2 \times V$ 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 ([14], Thm. 2.2). *Let $\ell \in V'$ be given. The pair $(\boldsymbol{\Sigma}, \mathbf{u}) \in S^2 \times V$ is the unique solution of (VI) if and only if there exists a plastic multiplier $\lambda \in L^2(\Omega)$ such that*

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

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

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

holds. Moreover, λ is unique.

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 $(\boldsymbol{\Sigma}, \mathbf{u})$ 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*

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

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 [27], Theorem 4.4.4 and Proposition 4.4.5. We aim to apply Theorem 1.1 of [13]. To this end, we will reformulate (VI) and (2.1), respectively, as a single nonlinear PDE in \mathbf{u} , cf. (2.7) below, and verify Assumption 1.5(2) of [13] for the underlying nonlinearity.

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

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

Combining both equations yields

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

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 $\mathbb{C}^{-1}, \boldsymbol{\sigma}, \mathbb{H}^{-1}, \boldsymbol{\chi}, \boldsymbol{\varepsilon}(\mathbf{u})$ 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

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

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

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

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

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

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

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

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. $x \in \Omega$

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

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

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

We define $M_x : \mathbb{S} \rightarrow \mathbb{S}$ by $\boldsymbol{\sigma} \mapsto \mathbb{C}^{-1}(x)\boldsymbol{\sigma} + \mathbb{H}^{-1}(x)(\boldsymbol{\sigma} - \text{Proj}_{\bar{K}}^{\mathbb{H}^{-1}(x)}(\boldsymbol{\sigma}))$. In view of the monotonicity of $\mathbb{H}^{-1}(x)(\boldsymbol{\sigma} - \text{Proj}_{\bar{K}}^{\mathbb{H}^{-1}(x)}(\boldsymbol{\sigma}))$ (see, e.g., [12], Lem. 4.1), we have for every $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathbb{S}$

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

where \underline{m} is the coercivity constant of \mathbb{C}^{-1} . Thus $M_x(\cdot)$ is strongly monotone and coercive because of $M_x(\mathbf{0}) = \mathbf{0}$. Due to the boundedness of \mathbb{C}^{-1} and \mathbb{H}^{-1} and the non-expansiveness of $\text{Proj}_K^{\mathbb{H}^{-1}(x)}$ 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 $\boldsymbol{\sigma} = M^{-1}(\cdot, \boldsymbol{\varepsilon}(\mathbf{u}))$ and hence, due to (2.1b), \mathbf{u} is a solution of

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

which is the desired nonlinear PDE in the displacement field only. In order to apply Theorem 1.1 of [13], we have to verify that M^{-1} satisfies ([13], 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

$$(M_x^{-1}(\boldsymbol{\sigma}) - M_x^{-1}(\boldsymbol{\tau}), \boldsymbol{\sigma} - \boldsymbol{\tau})_{\mathbb{S}} \geq \underline{m}/\overline{m}^2 \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{\mathbb{S}}^2,$$

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 $\|\mathbb{C}^{-1}(x) - \mathbb{C}_n^{-1}(x)\|_{\mathcal{L}(\mathbb{S})} \rightarrow 0$ and $\|\mathbb{H}^{-1}(x) - \mathbb{H}_n^{-1}(x)\|_{\mathcal{L}(\mathbb{S})} \rightarrow 0$ 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 $\mathbb{I}_{\mathbb{S}} : \mathbb{S} \rightarrow \mathbb{S}$ denotes the identity. Obviously $\tilde{\mathbb{C}}_n^{-1} : \Omega \rightarrow \mathbb{S}$, $\tilde{\mathbb{H}}_n^{-1} : \Omega \rightarrow \mathbb{S}$ are simple functions. Moreover $\tilde{\mathbb{C}}_n^{-1}(x), \tilde{\mathbb{H}}_n^{-1}(x) \in \mathbb{S}$ are coercive and $\tilde{\mathbb{C}}_n^{-1}(x) \rightarrow \mathbb{C}^{-1}(x)$ and $\tilde{\mathbb{H}}_n^{-1}(x) \rightarrow \mathbb{H}^{-1}(x)$ f.a.a. $x \in \Omega$. Analogous arguments as for M_x show that $M_x^n : \mathbb{S} \rightarrow \mathbb{S}$, defined by $\boldsymbol{\sigma} \mapsto \tilde{\mathbb{C}}_n^{-1}(x)\boldsymbol{\sigma} + \tilde{\mathbb{H}}_n^{-1}(x)(\boldsymbol{\sigma} - \text{Proj}_K^{\tilde{\mathbb{H}}_n^{-1}(x)}(\boldsymbol{\sigma})) \in \mathbb{S}$, 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 $\boldsymbol{\chi} := \text{Proj}_K^{\mathbb{H}^{-1}(x)}(\boldsymbol{\sigma})$ and $\boldsymbol{\chi}_n := \text{Proj}_K^{\tilde{\mathbb{H}}_n^{-1}(x)}(\boldsymbol{\sigma})$ we find

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

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 $\|\boldsymbol{\chi}_n - \boldsymbol{\chi}\|_{\mathbb{S}} \rightarrow 0$. Hence $\{\boldsymbol{\chi}_n(x)\}_{n \in \mathbb{N}}$ is bounded in \mathbb{S} . Consequently it holds

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

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

$$\begin{aligned} \|(M^n)^{-1}(x, \sigma) - M^{-1}(x, \sigma)\|_{\mathbb{S}} &= \|(M_x^n)^{-1}(M_x^n((M_x^n)^{-1}(\sigma))) - (M_x^n)^{-1}(M_x^n(M_x^{-1}(\sigma)))\|_{\mathbb{S}} \\ &\leq L_x^n \|M_x^n((M_x^n)^{-1}(\sigma)) - M_x^n(M_x^{-1}(\sigma))\|_{\mathbb{S}} \\ &= \max(1, 2/\underline{m}) \|M_x(M_x^{-1}(\sigma)) - M_x^n(M_x^{-1}(\sigma))\|_{\mathbb{S}} \xrightarrow{n \rightarrow \infty} 0, \end{aligned}$$

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

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

f.a.a. $x \in \Omega$ and all $\sigma, \tau \in \mathbb{S}$. Therefore M^{-1} satisfies Assumption 1.5(2) in [13], and Theorem 1.1 (cf. [13]), 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 $\mathbf{u} \in W_D^{1,p}(\Omega; \mathbb{R}^d)$. Moreover, the associated solution map $W_D^{-1,p}(\Omega; \mathbb{R}^d) \ni \ell \mapsto \mathbf{u} \in W_D^{1,p}(\Omega; \mathbb{R}^d)$ 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 $\sigma = M^{-1}(\cdot, \varepsilon(\mathbf{u})) \in L^p(\Omega; \mathbb{S})$. Finally, since \mathbb{H}^{-1} is uniformly coercive by Assumption 1.2, (2.3) implies $\chi \in L^p(\Omega; \mathbb{S})$, and both σ and χ depend Lipschitz continuously on \mathbf{u} and thus on ℓ . \square

Remark 2.5. We underline that the proof of Theorem 1.1 from [13], adapts a technique introduced by Gröger in [8] for scalar second-order differential operators to the case of nonlinear elasticity. For this purpose Assumption 1.2(1) is required.

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

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

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

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 $G: \ell \rightarrow (\Sigma, \mathbf{u})$ 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 [15] 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 $(\bar{\Sigma}, \bar{\mathbf{u}}, \bar{\lambda})$, i.e., $(\bar{\Sigma}, \bar{\mathbf{u}}, \bar{\lambda}) \in S \times V \times L^2(\Omega)$ solves

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

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

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

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

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

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

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

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

Theorem 3.1 ([15], Thm. 3.2). *For every $\bar{\ell} \in V'$ and every $\delta\ell \in V'$, the control-to-state map $G: V' \rightarrow S^2 \times V$ 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 $(\Sigma', \mathbf{u}') \in \mathcal{S}_\ell \times V$ of the following variational inequality:

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

$$B\Sigma' = \delta\ell, \quad (3.3b)$$

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

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

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

Theorem 3.2 ([15], Prop. 3.13). *Let $\bar{\ell} \in V'$ and $\delta\ell \in V'$ be given and let $(\bar{\Sigma}, \bar{\mathbf{u}}, \bar{\lambda})$ be the state and plastic multiplier associated with $\bar{\ell}$. A pair $(\Sigma', \mathbf{u}') \in S^2 \times V$ is the unique solution of (3.3) if and only if there exists a multiplier $\lambda' \in L^2(\Omega)$ such that*

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

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

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

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

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

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 $\bar{\chi} \in L^s(\Omega; \mathbb{S})$ with

$$s > \frac{2p}{p-2}. \quad (3.5)$$

Moreover we set

$$q = \frac{sp}{s+p}. \quad (3.6)$$

Note that $q > 2$ and $p' < q < p$ due to (3.5), where p' is the integrability exponent conjugated to p , i.e., $1/p + 1/p' = 1$.

Remark 3.4.

- (i) If the right hand sides ℓ and $\delta\ell$ in (3.1) and (3.4) are elements of $U = L^2(\Gamma_N; \mathbb{R}^d)$, and thus feasible controls for (\mathbf{P}) , assumption (3.3) (i) is automatically fulfilled (cf. Rem. 1.1).

- (ii) Note that we assume the higher integrability (3.5) only for the particular solution $\bar{\chi}$ of (3.1), where the control-to-state map is differentiated. Later on, in Section 4, this will be a fixed stationary point of (P). Let us point out that we do not assume the control-to-state operator G to map from $W_D^{-1,p}(\Omega; \mathbb{R}^d)$ into $L^p(\Omega; \mathbb{S}) \times L^s(\Omega; \mathbb{S}) \times W_D^{1,p}(\Omega; \mathbb{R}^d)$ (cf. Thm. 2.4). Nevertheless such higher integrability of the solution could probably be proven, provided that the data of the problem are sufficiently smooth. This is for instance the case, if volume loads are applied without mixed boundary conditions, i.e., $\Gamma_D = \partial\Omega$, and $\partial\Omega$, \mathbb{C} , and \mathbb{H} are sufficiently smooth (cf. e.g. [24], Thm. 5.2).

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

Next let us consider the following perturbed problem:

$$A\Sigma + B^*u + \lambda D^*D\Sigma = 0 \quad \text{in } S^2, \quad (3.7a)$$

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

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

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

Lemma 3.5. *Let $(\bar{\Sigma}, \bar{\lambda})$ and (Σ, λ) be given by the solution of (3.1) and (3.7), respectively. Then it holds*

- (i) $\|\Sigma - \bar{\Sigma}\|_{L^p(\Omega; \mathbb{S}^2)} \leq L \|\delta\ell\|_{W^{-1,p}(\Omega; \mathbb{R}^d)}$,
- (ii) $D\Sigma - D\bar{\Sigma} \rightarrow 0$ in $L^m(\Omega; \mathbb{S})$ for all $1 \leq m < \infty$, if $\delta\ell \rightarrow 0$ in $W_D^{-1,p}(\Omega; \mathbb{R}^d)$,
- (iii) $\|\lambda - \bar{\lambda}\|_{L^q(\Omega)} \leq c \|\delta\ell\|_{W_D^{-1,p}(\Omega; \mathbb{R}^d)}$.

Proof. Let $(\delta\ell_n)_{n \in \mathbb{N}} \subset W_D^{-1,p}(\Omega; \mathbb{R}^d)$ be an arbitrary sequence with $\delta\ell_n \rightarrow 0$ for $n \rightarrow \infty$ and let (Σ_n, λ_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 $\|D\Sigma_n(x) - D\bar{\Sigma}(x)\|_{\mathbb{S}} \leq 2\sigma_0$ a.e. in Ω , i.e., $D\Sigma_n - D\bar{\Sigma}$ is bounded in $L^m(\Omega; \mathbb{S})$. In addition we know $D\Sigma_n(x) - D\bar{\Sigma}(x) \rightarrow 0$ 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 $\bar{\lambda}$

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

Completely analogously one shows $\sigma_0^2 \lambda_n = \lambda_n D\Sigma_n : D\Sigma_n = -\mathbb{H}^{-1} \chi_n : D\Sigma_n$ a.e. in Ω . Hence it holds

$$\lambda_n - \bar{\lambda} = \frac{1}{\sigma_0^2} (-\mathbb{H}^{-1} \bar{\chi} : (D\Sigma_n - D\bar{\Sigma}) - \mathbb{H}^{-1} (\chi_n - \bar{\chi}) : D\Sigma_n). \quad (3.9)$$

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

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

which establishes (iii). \square

To establish a regularity result for the directional derivative (Σ', u', λ') let us now consider another perturbed problem:

$$A\Sigma_t + B^*u_t + \lambda_t D^*D\Sigma_t = 0 \quad \text{in } S^2, \quad (3.10a)$$

$$B\Sigma_t = \bar{\ell} + t\delta\ell \quad \text{in } V', \quad (3.10b)$$

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

with $t > 0$ given.

Lemma 3.6. *Let $(\bar{\Sigma}, \bar{\lambda})$ be given by the solution of (3.1), (Σ_t, λ_t) by the solution of (3.10) and (Σ', λ') by the solution of (3.4). Then it holds*

- (i) $\frac{\Sigma_t - \bar{\Sigma}}{t} \rightharpoonup \Sigma'$ 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 $\|\Sigma'\|_{L^p(\Omega; \mathbb{S}^2)} \leq L \|\delta\ell\|_{W_D^{-1,p}(\Omega; \mathbb{R}^d)}$.

Proof. Let $(t_n)_{n \in \mathbb{N}} \subset \mathbb{R}^+$ be an arbitrary sequence of positive real numbers converging to zero and let $(\Sigma_{t_n}, \lambda_{t_n})$ 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 $(\Sigma_{t_n} - \bar{\Sigma})/t_n \rightharpoonup \Sigma'$ in S^2 . Moreover it holds

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

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

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

Due to (i) and (3.6) it holds $\mathbb{H}^{-1}\bar{\chi} : (\mathcal{D}\Sigma_{t_n} - \mathcal{D}\bar{\Sigma})/t_n \rightharpoonup \mathbb{H}^{-1}\bar{\chi} : \mathcal{D}\Sigma'$ in $L^q(\Omega)$. If we choose $m = s$, Lemma 3.5 (ii) implies $\mathbb{H}^{-1}(\chi_{t_n} - \bar{\chi})/t_n : \mathcal{D}\Sigma_{t_n} \rightharpoonup \mathbb{H}^{-1}\chi' : \mathcal{D}\bar{\Sigma}$ in $L^q(\Omega)$. Consequently, $(\lambda_{t_n} - \bar{\lambda})/t_n \rightharpoonup 1/\sigma_0^2 (-\mathbb{H}^{-1}\bar{\chi} : \mathcal{D}\Sigma' - \mathbb{H}^{-1}\chi' : \mathcal{D}\bar{\Sigma})$ in $L^q(\Omega)$. Since $\mathcal{D}\Sigma' : \mathcal{D}\bar{\Sigma} = 0$ 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 $\lambda \mathcal{D}\Sigma' : \mathcal{D}\bar{\Sigma} = 0$ a.e. in Ω . In view of (2.2) we thus know

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

Hence

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

Testing (3.4a) with $(\mathbf{0}, \mu)$, $\mu \in S$, yields

$$\mathbb{H}^{-1}\chi' + \bar{\lambda} \mathcal{D}\Sigma' + \lambda' \mathcal{D}\bar{\Sigma} = 0 \quad \text{a.e. in } \Omega. \quad (3.15)$$

Because of (3.2), (3.1c), and (3.4c)–(3.4e) one deduces $\sigma_0^2 \lambda' = \lambda' \mathcal{D}\bar{\Sigma} : \mathcal{D}\bar{\Sigma}$ a.e. in Ω and therefore

$$\sigma_0^2 \lambda' = -\mathbb{H}^{-1}\chi' : \mathcal{D}\bar{\Sigma} - \bar{\lambda} \mathcal{D}\Sigma' : \mathcal{D}\bar{\Sigma}. \quad (3.16)$$

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

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

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

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

Corollary 3.7. For \mathbf{u}' given by the solution of (3.3) there exists a constant $c > 0$ such that

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

Proof. Let $\boldsymbol{\tau} \in S$ be arbitrary and define $\hat{\mathbf{T}} \in S^2$ by $\hat{\mathbf{T}} := (\boldsymbol{\tau}, -\boldsymbol{\tau})$. Due to $\mathcal{D}\hat{\mathbf{T}} = 0$ it then follows $\hat{\mathbf{T}} + \boldsymbol{\Sigma}' \in \mathcal{S}_\ell$. Thus we are allowed to test (3.3a) with $\hat{\mathbf{T}} + \boldsymbol{\Sigma}'$, which together with Lemma 3.6 (iii) yields

$$\begin{aligned} \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{u}') : \boldsymbol{\tau} \, dx &\leq (A\boldsymbol{\Sigma}', \hat{\mathbf{T}}) \leq c \|\boldsymbol{\Sigma}'\|_{L^p(\Omega;\mathbb{S}^2)} \|\boldsymbol{\tau}\|_{L^{p'}(\Omega;\mathbb{S})} \\ &\leq c \|\delta\ell\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)} \|\boldsymbol{\tau}\|_{L^{p'}(\Omega;\mathbb{S})} \quad \forall \boldsymbol{\tau} \in S. \end{aligned}$$

By the Hahn–Banach theorem $\boldsymbol{\varepsilon}(\mathbf{u}')$ 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 3.8. Let $(\bar{\boldsymbol{\Sigma}}, \bar{\mathbf{u}}, \bar{\lambda})$ be the solution of (3.1), $(\boldsymbol{\Sigma}, \mathbf{u}, \lambda)$ the solution of (3.7) and $(\boldsymbol{\Sigma}', \mathbf{u}', \lambda')$ the solution of (3.4). Then it holds

$$\|\boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'\|_{S^2} + \|\mathbf{u} - \bar{\mathbf{u}} - \mathbf{u}'\|_V = o\left(\|\delta\ell\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)}\right), \quad (3.17)$$

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

Remark 3.9. We point out that a norm gap is needed for the proof of Theorem 3.8, 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. [28]).

Proof of Theorem 3.8. Let $(\delta\ell_n)_{n \in \mathbb{N}} \subset W_D^{-1,p}(\Omega;\mathbb{R}^d)$ be an arbitrary sequence with $\delta\ell_n \rightarrow 0$ for $n \rightarrow \infty$. Furthermore by $(\boldsymbol{\Sigma}_n, \mathbf{u}_n, \lambda_n)$ we denote the solution of (3.7) with right hand side $\bar{\ell} + \delta\ell_n$ and by $(\boldsymbol{\Sigma}'_n, \mathbf{u}'_n, \lambda'_n)$ 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

$$\begin{aligned} &(A(\boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n), \boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n) + \underbrace{(B^*(\mathbf{u}_n - \bar{\mathbf{u}} - \mathbf{u}'_n), \boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n)}_{=: I_1} \\ &+ (\lambda_n \mathcal{D}\boldsymbol{\Sigma}_n - \bar{\lambda} \mathcal{D}\bar{\boldsymbol{\Sigma}} - \lambda'_n \mathcal{D}\boldsymbol{\Sigma}'_n - \lambda'_n \mathcal{D}\bar{\boldsymbol{\Sigma}}, \mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}'_n) = 0. \end{aligned} \quad (3.18)$$

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 (3.18) implies

$$\begin{aligned} c \|\boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n\|_{S^2}^2 &\leq (A(\boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n), \boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n) \\ &= - \int_{\Omega} \underbrace{\bar{\lambda}}_{\geq 0} \underbrace{\|\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}'_n\|_{\mathbb{S}}^2}_{\geq 0} \, dx \\ &\quad - \int_{\Omega} (\lambda_n - \bar{\lambda} - \lambda'_n) \mathcal{D}\boldsymbol{\Sigma}_n : (\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}'_n) \, dx \\ &\quad - \int_{\Omega} \lambda'_n (\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}}) : (\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}} - \mathcal{D}\boldsymbol{\Sigma}'_n) \, dx \leq I_s + I_b + I_i + I_{\Omega}, \end{aligned} \quad (3.19)$$

where

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

Moreover, I_s is given by

$$I_s := - \int_{\bar{\mathcal{A}}_s} (\lambda_n - \bar{\lambda} - \lambda'_n) \mathcal{D}\Sigma_n : (\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma'_n) \, dx,$$

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 3.5 there exist subsequences λ_{n_k} and Σ_{n_k} with $\lambda_{n_k}(x) \rightarrow \bar{\lambda}(x)$ and $\Sigma_{n_k}(x) \rightarrow \bar{\Sigma}(x)$ a.e. on Ω . For the sake of convenience we denote these subsequences again by λ_n and Σ_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 3.6(iii) implies

$$\begin{aligned} I_\Omega &\leq \|\lambda'_n\|_{L^q(\Omega)} \|\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma}\|_{L^m(\Omega;\mathbb{S})} \|\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma'_n\|_S \\ &\leq c \|\delta\ell_n\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)} \|\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma}\|_{L^m(\Omega;\mathbb{S})} \|\Sigma_n - \bar{\Sigma} - \Sigma'_n\|_{S^2}. \end{aligned} \tag{3.20}$$

Estimation of I_s :

If $x \in \bar{\mathcal{A}}_s$, then it holds $\bar{\lambda}(x) > 0$ and (3.1c) gives $\phi(\bar{\Sigma}(x)) = 0$. 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 $\phi(\Sigma_n(x)) = 0$, i.e., $\|\mathcal{D}\Sigma_n(x)\|_{\mathbb{S}} = \sigma_0$, for all $n \geq N_x$ and f.a.a. $x \in \bar{\mathcal{A}}_s$. From Lemma A.1 we conclude

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

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

$$z_n \rightarrow 0 \quad \text{in } L^\xi(\bar{\mathcal{A}}_s) \quad \text{for all } 1 \leq \xi < \infty. \tag{3.21}$$

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

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

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 Lemmas 3.5 and 3.6(iii) we thus arrive at

$$\begin{aligned} I_s &\leq \|\lambda_n - \bar{\lambda} - \lambda'_n\|_{L^q(\Omega)} \|\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma}\|_{L^m(\Omega;\mathbb{S})} \|\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma'_n\|_{S^2} \\ &\quad + \sigma_0 \|\lambda_n - \bar{\lambda} - \lambda'_n\|_{L^q(\Omega)} \|z_n\|_{L^\xi(\bar{\mathcal{A}}_s)} \|\Sigma_n - \bar{\Sigma}\|_{L^p(\Omega;S^2)} \\ &\leq c \|\delta\ell_n\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)} \|\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma}\|_{L^m(\Omega;\mathbb{S})} \|\Sigma_n - \bar{\Sigma} - \Sigma'_n\|_{S^2} \\ &\quad + c \|\delta\ell_n\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)}^2 \|z_n\|_{L^\xi(\bar{\mathcal{A}}_s)}. \end{aligned} \tag{3.22}$$

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

$$(\lambda_n - \bar{\lambda}) \mathcal{D}\bar{\Sigma} : (\mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma_n) = (\lambda_n - \bar{\lambda}) \frac{1}{2} \|\mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma_n\|_{\mathbb{S}}^2 \tag{3.23}$$

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{B} , cf. (3.2) and (3.4c)–(3.4e). Therefore we obtain

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

where we used (3.23) for the last estimate. Thus, as in (3.22) 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

$$\begin{aligned}
 I_b &\leq \|\lambda_n - \bar{\lambda} - \lambda'_n\|_{L^q(\Omega)} \|\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma}\|_{L^m(\Omega; \mathbb{S})} \|\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma'_n\|_{S^2} \\
 &\quad + \frac{1}{2} \|\lambda_n - \bar{\lambda}\|_{L^q(\Omega)} \|\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma}\|_{L^\xi(\Omega; \mathbb{S})} \|\Sigma_n - \bar{\Sigma}\|_{L^p(\Omega; \mathbb{S}^2)} \\
 &\leq c \|\delta\ell_n\|_{W_D^{-1,p}(\Omega; \mathbb{R}^d)} \|\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma}\|_{L^m(\Omega; \mathbb{S})} \|\Sigma_n - \bar{\Sigma} - \Sigma'_n\|_{S^2} \\
 &\quad + c \|\delta\ell_n\|_{W_D^{-1,p}(\Omega; \mathbb{R}^d)}^2 \|\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma}\|_{L^\xi(\Omega; \mathbb{S})}.
 \end{aligned} \tag{3.24}$$

Estimation of I_i :

For $x \in \bar{I}$ there holds $\phi(\bar{\Sigma}(x)) < 0$ 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 $\phi(\Sigma_n(x)) < 0$ and $\lambda_n(x) = 0$ for all $n \geq \tilde{N}_x$ and for a.a. $x \in \bar{I}$. In particular this implies $(\lambda_n - \bar{\lambda}) / \|\delta\ell_n\|_{W^{-1,p}(\Omega; \mathbb{R}^d)} \rightarrow 0$ a.e. in \bar{I} . Furthermore we know $\|\lambda_n - \bar{\lambda}\|_{L^q(\mathcal{I}(\ell))} / \|\delta\ell_n\|_{W^{-1,p}(\Omega; \mathbb{R}^d)} \leq c \forall n \in \mathbb{N}$ with $2 < q$, cf. Lemma 3.5(iii) and (3.6). Thus, Lemma A.2 implies

$$\omega_n := \frac{\lambda_n - \bar{\lambda}}{\|\delta\ell_n\|_{W_D^{-1,p}(\Omega; \mathbb{R}^d)}} \rightarrow 0 \quad \text{in } L^2(\bar{I}). \tag{3.25}$$

Due to (3.4e), it follows

$$\begin{aligned}
 I_i &= - \int_{\bar{I}} (\lambda_n - \bar{\lambda} - \lambda'_n) \mathcal{D}\Sigma_n : (\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma'_n) \, dx \\
 &= \int_{\bar{I}} \omega_n \|\delta\ell_n\|_{W^{-1,p}(\Omega; \mathbb{R}^d)} \mathcal{D}\Sigma_n : (\mathcal{D}\Sigma_n - \mathcal{D}\bar{\Sigma} - \mathcal{D}\Sigma'_n) \, dx \\
 &\leq \sigma_0 \|\omega_n\|_{L^2(\bar{I})} \|\delta\ell_n\|_{W_D^{-1,p}(\Omega; \mathbb{R}^d)} \|\Sigma_n - \bar{\Sigma} - \Sigma'_n\|_{S^2}.
 \end{aligned} \tag{3.26}$$

In summary, (3.20), (3.22), (3.24), and (3.26) together with (3.19), Lemma 3.6(iii), and Young’s inequality yield

$$\begin{aligned} \frac{\|\boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n\|_{S^2}^2}{\|\delta\ell_n\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)}^2} &\leq c \left(\|\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}}\|_{L^m(\Omega;\mathbb{S})} + \|\omega_n\|_{L^2(\bar{\Omega})} \right)^2 \\ &+ c \left(\|z_n\|_{L^\xi(\bar{\mathcal{A}}_s)} + \|\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\bar{\boldsymbol{\Sigma}}\|_{L^\xi(\Omega;\mathbb{S})} \right) \xrightarrow{n \rightarrow \infty} 0, \end{aligned} \tag{3.27}$$

where we used Lemma 3.5(ii), (3.21) and (3.25) to pass to the limit.

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

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

so that

$$\begin{aligned} (A(\boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n), \tilde{\mathbf{T}}) + (B^*(\mathbf{u}_n - \bar{\mathbf{u}} - \mathbf{u}'_n), \tilde{\mathbf{T}}) \\ + \underbrace{(\lambda_n \mathcal{D}\boldsymbol{\Sigma}_n - \bar{\lambda} \mathcal{D}\bar{\boldsymbol{\Sigma}} - \bar{\lambda} \mathcal{D}\boldsymbol{\Sigma}'_n - \lambda'_n \mathcal{D}\bar{\boldsymbol{\Sigma}}, \mathcal{D}\tilde{\mathbf{T}})}_{=: I_2} = 0. \end{aligned} \tag{3.28}$$

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

$$\begin{aligned} \|\mathbf{u}_n - \bar{\mathbf{u}} - \mathbf{u}'_n\|_V^2 &\leq c (B^*(\mathbf{u}_n - \bar{\mathbf{u}} - \mathbf{u}'_n), \tilde{\mathbf{T}}) \leq c (A(\boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n), \tilde{\mathbf{T}}) \\ &\leq c \|\mathbf{u}_n - \bar{\mathbf{u}} - \mathbf{u}'_n\|_V \|\boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n\|_{S^2}. \end{aligned} \tag{3.29}$$

Hence (3.27) gives $\|\mathbf{u}_n - \bar{\mathbf{u}} - \mathbf{u}'_n\|_V / \|\delta\ell_n\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)} \rightarrow 0$.

Since the above arguments hold for every subsequence of $(\boldsymbol{\Sigma}_n, \mathbf{u}_n, \lambda_n)$, a standard argument implies (3.17) for the whole sequence. \square

In view of Lemma A.2 we can enhance the result of Theorem 3.8.

Corollary 3.10. *Let $(\bar{\boldsymbol{\Sigma}}, \bar{\mathbf{u}}, \bar{\lambda})$ be the solution of (3.1), $(\boldsymbol{\Sigma}, \mathbf{u}, \lambda)$ the solution of (3.7) and $(\boldsymbol{\Sigma}', \mathbf{u}', \lambda')$ the solution of (3.4). Then it holds*

$$\|\boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'\|_{L^\beta(\Omega;\mathbb{S}^2)} + \|\mathbf{u} - \bar{\mathbf{u}} - \mathbf{u}'\|_{W_D^{1,\beta}(\Omega;\mathbb{R}^d)} = o\left(\|\delta\ell\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)}\right) \tag{3.30}$$

for all $1 \leq \beta < p$.

Proof. As in the proof of Theorem 3.8 let $(\delta\ell_n)_{n \in \mathbb{N}} \subset W_D^{-1,p}(\Omega;\mathbb{R}^d)$ be an arbitrary sequence with $\delta\ell_n \rightarrow 0$ for $n \rightarrow \infty$. Furthermore by $(\boldsymbol{\Sigma}_n, \mathbf{u}_n, \lambda_n)$ we denote the solution of (3.7) with right hand side $\bar{\ell} + \delta\ell_n$ and by $(\boldsymbol{\Sigma}'_n, \mathbf{u}'_n, \lambda'_n)$ the solution of (3.4) with right hand side $\delta\ell_n$. Thanks to Theorem 3.8 there exist subsequences, denoted in the same way, such that

$$\frac{\boldsymbol{\Sigma}_n - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'_n}{\|\delta\ell\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)}} \rightarrow 0 \quad \text{a.e. in } \Omega, \quad \frac{\mathbf{u}_n - \bar{\mathbf{u}} - \mathbf{u}'_n}{\|\delta\ell\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)}} \rightarrow 0 \quad \text{a.e. in } \Omega.$$

Moreover these subsequences are bounded in $L^p(\Omega;\mathbb{S}^2)$ and $W_D^{1,p}(\Omega;\mathbb{R}^d)$, respectively, cf. Theorem 2.4, Lemma 3.6 and Corollary 3.7. By Lemma A.2 we get (3.30) for the subsequences and a standard argument gives (3.30) for the whole sequence. \square

Theorem 3.8 entails two consequences stated in the following corollaries:

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

$$\|\lambda - \bar{\lambda} - \lambda'\|_{L^\gamma(\Omega)} = o\left(\|\delta\ell\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)}\right) \tag{3.31}$$

for all $1 \leq \gamma < q$.

Proof. In view of (3.8), (3.16) and (3.13) it holds

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

Since $q = 1/(\frac{1}{s} + \frac{1}{p})$, cf. (3.6), every $\gamma < q$ can be written as $\gamma = 1/(\frac{1}{s} + \frac{1}{\beta})$ for some $\beta < p$. The boundedness $\|\mathcal{D}\bar{\Sigma}(x)\|_{\mathbb{S}} \leq \sigma_0$ a.e. in Ω implies

$$\begin{aligned} \|\lambda - \bar{\lambda} - \lambda'\|_{L^\gamma(\Omega)} &\leq c \|\chi - \bar{\chi}\|_{L^\beta(\Omega; \mathbb{S})} \|\mathcal{D}\Sigma - \mathcal{D}\bar{\Sigma}\|_{L^s(\Omega; \mathbb{S})} \\ &\quad + c \|\chi - \bar{\chi} - \chi'\|_{L^\beta(\Omega; \mathbb{S}^2)} \\ &\quad + c \|\bar{\chi}\|_{L^s(\Omega; \mathbb{S})} \|\Sigma - \bar{\Sigma} - \Sigma'\|_{L^\beta(\Omega; \mathbb{S}^2)}. \end{aligned}$$

Thus Corollary 3.10 and Lemma 3.5 yield the claim. □

Corollary 3.12. *For all $1 \leq \beta < p$ the control-to-state map $G : W_D^{-1,p}(\Omega; \mathbb{R}^d) \rightarrow L^\beta(\Omega; \mathbb{S}^2) \times W_D^{1,\beta}(\Omega; \mathbb{R}^d)$ is directionally differentiable at $\bar{\ell}$ in direction $\delta\ell$.*

Furthermore the mapping $W_D^{-1,p}(\Omega; \mathbb{R}^d) \ni \ell \mapsto \lambda \in L^\gamma(\Omega)$ is directionally differentiable at $\bar{\ell}$ in direction $\delta\ell$ for all $1 \leq \gamma < q$.

Proof. Since \mathcal{S}_ℓ is a cone, the mapping $\delta\ell \mapsto \delta G(\bar{\ell}, \delta\ell)$ is positively homogeneous so that $(t\Sigma', tu', t\lambda')$ is the solution of (3.4) with right hand side $t\delta\ell$. Consequently, (3.30) yields $(\|\Sigma_t - \bar{\Sigma} - t\Sigma'\|_{L^\beta(\Omega; \mathbb{S}^2)} + \|u_t - \bar{u} - tu'\|_{W_D^{1,\beta}(\Omega; \mathbb{R}^d)})/t \xrightarrow{t \searrow 0} 0$. The second assertion follows analogously from (3.31). □

Remark 3.13. 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 3.8 that the control-to-state map is only directionally differentiable without the remainder term property (3.17).

4. SECOND-ORDER SUFFICIENT OPTIMALITY CONDITIONS

With the differentiability results of the previous section at hand we can now establish second-order sufficient conditions which guarantee local optimality for (P). We will present two different second-order sufficient conditions:

1. The first set of conditions given in Theorem 4.6 allows for a general (smooth) objective. It involves a particularly chosen Lagrange functional, see (4.2) below. However, the coercivity of the Hessian of this Lagrangian is required on a cone which is larger than the classical cone of critical directions. The proof is based on a Taylor expansion of the Lagrangian and employs the remainder term properties of the directional derivative in Theorem 3.8.
2. Theorem 4.12 states another set of second-order sufficient conditions. These apply if the objective provides a particular structure, namely $J(\mathbf{u}, \mathbf{g}) = i(\mathbf{u}) + j(\mathbf{g})$, where i and j are smooth and j is in addition uniformly convex. In this case it is sufficient to assume the coercivity of the Hessian of the Lagrangian only on the cone of critical directions. The proof follows the lines of [20] and argues by contradiction.

Both proofs require the objective to be sufficiently smooth. To be more precise, we need the following

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

In preparation of our second-order analysis, we will next provide some auxiliary results. First observe that, in view of Theorem 2.3, Problem (P) is equivalent to

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

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

$$\tau_N^*: U \rightarrow V', \quad \langle \tau_N^*(\mathbf{g}), \mathbf{v} \rangle := \int_{\Gamma_N} \mathbf{g} \cdot \tau_N(\mathbf{v}) \, ds, \quad \mathbf{v} \in V.$$

In all what follows we set

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

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. In dimension $d = 2$ we can choose $p = \min(\hat{p}, 4)$ by Sobolev’s embedding theorem. In order to keep the discussion concise we set $p = \min(\hat{p}, 3)$ for both $d = 3$ and $d = 2$.

Remark 4.3. Due to the continuity of τ_N^* the results of Section 3 imply

- (i) $\|\boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}}\|_{L^p(\Omega; \mathbb{S}^2)} \leq L \|\tau_N^*\| \|\mathbf{g} - \bar{\mathbf{g}}\|_U$
- (ii) $\|\lambda - \bar{\lambda}\|_{L^q(\Omega)} \leq c \|\tau_N^*\| \|\mathbf{g} - \bar{\mathbf{g}}\|_U$,
- (iii) $\|\boldsymbol{\Sigma}'\|_{L^p(\Omega; \mathbb{S}^2)} \leq L \|\tau_N^*\| \|\mathbf{g} - \bar{\mathbf{g}}\|_U$,
- (iv) $\|\boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}'\|_{L^\beta(\Omega; \mathbb{S}^2)} + \|\mathbf{u} - \bar{\mathbf{u}} - \mathbf{u}'\|_{W_D^{1,\beta}(\Omega; \mathbb{R}^d)} = o(\|\mathbf{g} - \bar{\mathbf{g}}\|_U) \quad \forall \beta \in [1, p)$,
- (v) $\|\lambda - \bar{\lambda} - \lambda'\|_{L^\gamma(\Omega)} = o(\|\mathbf{g} - \bar{\mathbf{g}}\|_U) \quad \forall \gamma \in [1, q)$,
- (vi) $\|\mathbf{u}'\|_{W_D^{1,p}(\Omega; \mathbb{R}^d)} \leq c \|\tau_N^*\| \|\mathbf{g} - \bar{\mathbf{g}}\|_U$

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

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

Lemma 4.4. *Let $\nu \geq 1$ and $\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2 \in \mathcal{K}$. Then*

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

holds true.

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

$$\begin{aligned} \|\phi(\boldsymbol{\Sigma}_1) - \phi(\boldsymbol{\Sigma}_2)\|_{L^\nu(\Omega)}^\nu &= \int_\Omega \left| \frac{1}{2}(\mathcal{D}\boldsymbol{\Sigma}_2 + \mathcal{D}\boldsymbol{\Sigma}_1) : (\mathcal{D}\boldsymbol{\Sigma}_1 - \mathcal{D}\boldsymbol{\Sigma}_2) \right|^\nu \, dx \\ &\leq \sigma_0^\nu \|\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2\|_{L^\nu(\Omega; \mathbb{S}^2)}^\nu. \end{aligned}$$

□

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

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

Endowed with the norm $\|\mathbf{T}\|_{S^2} + \|\mathcal{D}\mathbf{T}\|_{L^\infty(\Omega; \mathbb{S})}$, S_∞^2 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 is defined by

$$\begin{aligned} \mathcal{L}: V \times S_\infty^2 \times L^2(\Omega) \times U \times S^2 \times V \times L^2(\Omega) \times L^\infty(\Omega)' &\rightarrow \mathbb{R}, \\ \mathcal{L}(\mathbf{u}, \boldsymbol{\Sigma}, \lambda, \mathbf{g}, \boldsymbol{\Upsilon}, \mathbf{w}, \mu, \theta) &:= J(\mathbf{u}, \mathbf{g}) + (A\boldsymbol{\Sigma} + B^*\mathbf{u} + \lambda\mathcal{D}^*\mathcal{D}\boldsymbol{\Sigma}, \boldsymbol{\Upsilon}) \\ &+ \langle B\boldsymbol{\Sigma} + \tau_N^*(\mathbf{g}), \mathbf{w} \rangle - (\lambda, \mu) + \langle \phi(\boldsymbol{\Sigma}), \theta \rangle_{L^\infty(\Omega), L^\infty(\Omega)'}. \end{aligned} \tag{4.2}$$

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 $\lambda\phi(\boldsymbol{\Sigma}) = 0$ a.e. in Ω , which is typical for MPECs, cf. for instance the Lagrangian defined in [26].

Before we are in the position to establish our first set of second-order conditions, we need to prove a last auxiliary result:

Lemma 4.5. *Let $(\boldsymbol{\Sigma}, \mathbf{u}, \lambda)$ be the state and multiplier associated to $\mathbf{g} \in U$ and let $(\bar{\boldsymbol{\Sigma}}, \bar{\mathbf{u}}, \bar{\lambda})$ be the state and multiplier associated to $\bar{\mathbf{g}} \in U$. Furthermore let $\bar{\mu} \in L^\zeta(\Omega)$ and $\bar{\theta} \in L^r(\Omega)$ with $\zeta, r > \frac{sp}{sp-p-2s}$, where $s > 2p/(p-2)$ and p is as defined in (4.1). If $\bar{\mu} \geq 0$ on $\mathcal{A}_1 := \{x \in \Omega : -\tau_1 \leq \phi(\bar{\boldsymbol{\Sigma}}) \leq 0\}$ for some $\tau_1 > 0$ and $\bar{\theta} \geq 0$ on $\mathcal{A}_2 := \{x \in \Omega : 0 \leq \bar{\lambda} \leq \tau_2\}$ for some $\tau_2 > 0$, then there are $\nu_1, \nu_2 > 2$ so that*

$$-\int_\Omega \lambda \bar{\mu} \, dx \leq c \|\mathbf{g} - \bar{\mathbf{g}}\|_U^{\nu_1}, \quad \int_\Omega \phi(\boldsymbol{\Sigma}) \bar{\theta} \, dx \leq c \|\mathbf{g} - \bar{\mathbf{g}}\|_U^{\nu_2}.$$

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

$$\begin{aligned} -\int_\Omega \lambda \bar{\mu} \, dx &= -\int_{E_1} \underbrace{\lambda}_{>0} \underbrace{\bar{\mu}}_{\geq 0} \, dx - \int_{\Omega_1 \setminus E_1} \lambda \bar{\mu} \, dx \\ &\leq -\int_{\Omega_1 \setminus E_1} (\lambda - \bar{\lambda}) \bar{\mu} \, dx \\ &\leq \|\lambda - \bar{\lambda}\|_{L^q(\Omega)} \|\bar{\mu}\|_{L^\zeta(\Omega)} |\Omega_1 \setminus E_1|^{\frac{1}{\beta}}. \end{aligned} \tag{4.3}$$

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 $\frac{1}{q} + \frac{1}{\zeta} + \frac{1}{\beta} = 1$ in consequence of $\zeta > \frac{sp}{sp-p-2s}$, where q is as defined in (3.6). Furthermore (3.7c) and the definition of Ω_1 and \mathcal{A}_1 imply

$$\begin{aligned} |\Omega_1 \setminus E_1| &= \frac{1}{\tau_1^p} \int_{\Omega_1 \setminus E_1} \tau_1^p \, dx \\ &< \frac{1}{\tau_1^p} \int_{\Omega_1 \setminus E_1} |\phi(\bar{\boldsymbol{\Sigma}})|^p \, dx = \frac{1}{\tau_1^p} \int_{\Omega_1 \setminus E_1} |\phi(\boldsymbol{\Sigma}) - \phi(\bar{\boldsymbol{\Sigma}})|^p \, dx. \end{aligned} \tag{4.4}$$

Thus, thanks to Lemma 4.4 and Remark 4.3, we arrive at

$$\begin{aligned} -\int_\Omega \lambda \bar{\mu} \, dx &< c \|\lambda - \bar{\lambda}\|_{L^q(\Omega)} \|\bar{\mu}\|_{L^\zeta(\Omega)} \|\phi(\boldsymbol{\Sigma}) - \phi(\bar{\boldsymbol{\Sigma}})\|_{L^p(\Omega)}^{\frac{p}{\beta}} \\ &\leq c \|\lambda - \bar{\lambda}\|_{L^q(\Omega)} \|\bar{\mu}\|_{L^\zeta(\Omega)} \|\boldsymbol{\Sigma} - \bar{\boldsymbol{\Sigma}}\|_{L^p(\Omega; \mathbb{S}^2)}^{\frac{p}{\beta}} \leq c \|\mathbf{g} - \bar{\mathbf{g}}\|_U^{1+\frac{p}{\beta}} \end{aligned} \tag{4.5}$$

with $\nu_1 := 1 + p/\beta > 2$. Similarly we define $\Omega_2 := \{x \in \Omega : \phi(\Sigma(x)) < 0\}$ and $E_2 := \Omega_2 \cap \mathcal{A}_2$. Analogously to (4.3), one finds

$$\int_{\Omega} \phi(\Sigma)\bar{\theta} \, dx \leq \|\phi(\Sigma) - \phi(\bar{\Sigma})\|_{L^p(\Omega)} \|\bar{\theta}\|_{L^r(\Omega)} |\Omega_2 \setminus E_2|^{\frac{1}{r}},$$

where we used that $r > \frac{sp}{sp-p-2s}$ implies the existence of $\gamma < q$ with q as defined in (3.6) so that $\frac{1}{p} + \frac{1}{r} + \frac{1}{\gamma} = 1$. As in (4.4) we deduce from (3.7c) and the definition of Ω_2 and \mathcal{A}_2 that

$$|\Omega_2 \setminus E_2| < \frac{1}{\tau_2^q} \int_{\Omega_2 \setminus E_2} |\lambda - \bar{\lambda}|^q \, dx. \quad (4.6)$$

Because of Lemma 4.4 and Remark 4.3 we thus obtain similarly to (4.5)

$$\int_{\Omega} \phi(\Sigma)\bar{\theta} \, dx \leq c \|\mathbf{g} - \bar{\mathbf{g}}\|_U^{1+\frac{q}{\gamma}}$$

with $\nu_2 := 1 + q/\gamma > 2$. □

Our first result on sufficient second-order conditions reads as follows:

Theorem 4.6. *Let Assumption 4.1 hold. Suppose further that $\bar{\mathbf{g}} \in U$ together with its state $(\bar{\Sigma}, \bar{\mathbf{u}}, \bar{\lambda}) \in S^2 \times V \times L^2(\Omega)$, adjoint state $(\bar{\mathbf{Y}}, \bar{\mathbf{w}}) \in S^2 \times V$, and multipliers $(\bar{\mu}, \bar{\theta}) \in L^2(\Omega) \times L^2(\Omega)$ satisfy*

(1) *the following regularity conditions:*

$$\bar{\chi} \in L^s(\Omega; \mathbb{S}), \quad \bar{\mathbf{Y}} \in L^\eta(\Omega; \mathbb{S}^2), \quad \bar{\mu} \in L^\zeta(\Omega), \quad \bar{\theta} \in L^r(\Omega)$$

with

$$s > \frac{2p}{p-2}, \quad \eta, \zeta, r > \frac{sp}{sp-p-2s}, \quad (4.7)$$

where p is as defined in (4.1),

(2) *the following optimality system:*

$$A\bar{\mathbf{Y}} + B^*\bar{\mathbf{w}} + \bar{\lambda}\mathcal{D}^*\mathcal{D}\bar{\mathbf{Y}} + \bar{\theta}\mathcal{D}^*\mathcal{D}\bar{\Sigma} = 0 \quad (4.8a)$$

$$B\bar{\mathbf{Y}} + \partial_{\mathbf{u}}J(\bar{\mathbf{u}}, \bar{\mathbf{g}}) = 0 \quad (4.8b)$$

$$\tau_N\bar{\mathbf{w}} + \partial_{\mathbf{g}}J(\bar{\mathbf{u}}, \bar{\mathbf{g}}) = 0 \quad (4.8c)$$

$$\mathcal{D}\bar{\Sigma} : \mathcal{D}\bar{\mathbf{Y}} = \bar{\mu} \quad (4.8d)$$

$$\bar{\mu}\bar{\lambda} = 0 \quad \text{a.e. in } \Omega \quad (4.8e)$$

$$\bar{\theta}\phi(\bar{\Sigma}) = 0 \quad \text{a.e. in } \Omega \quad (4.8f)$$

$$\bar{\mu} \geq 0 \quad \text{a.e. in } \mathcal{A}_1 \quad (4.8g)$$

$$\bar{\theta} \geq 0 \quad \text{a.e. in } \mathcal{A}_2 \quad (4.8h)$$

with

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

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

(3) *the second-order condition:*

there exists $\alpha > 0$ such that

$$\partial_{(\mathbf{u}, \Sigma, \lambda, \mathbf{g})}^2 \mathcal{L}(\bar{\mathbf{u}}, \bar{\Sigma}, \bar{\lambda}, \bar{\mathbf{g}}, \bar{\mathbf{Y}}, \bar{\mathbf{w}}, \bar{\mu}, \bar{\theta})(\mathbf{u}', \Sigma', \lambda', \mathbf{h})^2 \geq \alpha \|\mathbf{h}\|_U^2 \quad (\text{SSC})$$

for all $\mathbf{h} \in U$ and $(\Sigma', \mathbf{u}', \lambda')$ solving (3.4) with $\delta\ell = -\tau_N^*(\mathbf{h})$.

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

$$J(\mathbf{u}, \mathbf{g}) \geq J(\bar{\mathbf{u}}, \bar{\mathbf{g}}) + \frac{\alpha}{4} \|\mathbf{g} - \bar{\mathbf{g}}\|_U^2 \tag{4.9}$$

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

Remark 4.7. The Lagrangian \mathcal{L} is twice continuously differentiable in its space of definition in (4.2). If the dual variables $(\bar{\mathbf{Y}}, \bar{\theta})$ satisfy the integrability conditions in (4.7), its second derivative with respect to the primal variables $(\mathbf{u}, \boldsymbol{\Sigma}, \lambda, \mathbf{g})$ is given by

$$\begin{aligned} & \partial_{(\mathbf{u}, \boldsymbol{\Sigma}, \lambda, \mathbf{g})}^2 \mathcal{L}(\bar{\mathbf{u}}, \bar{\boldsymbol{\Sigma}}, \bar{\lambda}, \bar{\mathbf{g}}, \bar{\mathbf{Y}}, \bar{\mathbf{w}}, \bar{\mu}, \bar{\theta})(\delta \mathbf{u}, \delta \boldsymbol{\Sigma}, \delta \lambda, \delta \mathbf{g})^2 \\ &= \nabla_{(\mathbf{u}, \mathbf{g})}^2 J(\bar{\mathbf{u}}, \bar{\mathbf{g}})(\delta \mathbf{u}, \delta \mathbf{g})^2 + 2 \int_{\Omega} \delta \lambda \mathcal{D} \delta \boldsymbol{\Sigma} : \mathcal{D} \bar{\mathbf{Y}} \, dx + \int_{\Omega} \|\mathcal{D} \delta \boldsymbol{\Sigma}\|_{\mathbb{S}}^2 \bar{\theta} \, dx \end{aligned} \tag{4.10}$$

and defines a continuous bilinear form on the space $V \times L^p(\Omega; \mathbb{S}^2) \times L^q(\Omega) \times U$. In the following we denote the integrals in (4.10) as L^2 -scalarproducts by (\cdot, \cdot) .

Remark 4.8. Let us compare the sufficient optimality conditions in Theorem 4.6 with the necessary ones. In Section 3.3 from [14], so-called C-stationarity conditions are proven to be satisfied by every local optimum. These conditions coincide with (4.8), except that they only provide a sign condition for the product of the multipliers $\bar{\mu}$ and $\bar{\theta}$, whereas (4.8g) and (4.8h) 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 [15]. These conditions also contain sign conditions for each multiplier individually, but only on the strongly active set \mathcal{A}_s and the inactive set $\bar{\mathcal{I}}$, respectively. Hence the sign conditions in (4.8g) and (4.8h) are even more restrictive than strong stationarity. Moreover, higher regularity of the optimal hardening variable $\bar{\chi}$ and the multipliers $\bar{\mathbf{Y}}$, $\bar{\mu}$, and $\bar{\theta}$ is required in (4.7) in contrast to C-stationarity, where these variables are only element of L^2 -type spaces. Notice however that we do not impose any differentiability properties on neither $\bar{\chi}$ nor $\bar{\mathbf{Y}}$, $\bar{\mu}$, or $\bar{\theta}$. Nevertheless, in summary, there is thus a significant gap between the necessary and sufficient optimality conditions for **(P)**.

Proof of Theorem 4.6. At first we note that Assumption 3.3 is satisfied because of $s > \frac{2p}{p-2}$ and (4.1). Let $\mathbf{g} \in U$ be arbitrary and $(\mathbf{u}, \boldsymbol{\Sigma}, \lambda) \in V \times S_{\infty}^2 \times L^2(\Omega)$ be the state and multiplier associated with \mathbf{g} .

We aim to deduce the quadratic growth condition (4.9) from a Taylor expansion of the Lagrangian. To this end we introduce the abbreviations $\mathbf{z} := (\mathbf{u}, \boldsymbol{\Sigma}, \lambda, \mathbf{g})$, $\bar{\mathbf{z}} := (\bar{\mathbf{u}}, \bar{\boldsymbol{\Sigma}}, \bar{\lambda}, \bar{\mathbf{g}})$, and $\bar{\boldsymbol{\omega}} := (\bar{\mathbf{Y}}, \bar{\mathbf{w}}, \bar{\mu}, \bar{\theta})$. Since \mathcal{L} is twice continuously differentiable in its space of definition, there is a $t \in [0, 1]$ such that

$$\mathcal{L}(\mathbf{z}, \bar{\boldsymbol{\omega}}) = \mathcal{L}(\bar{\mathbf{z}}, \bar{\boldsymbol{\omega}}) + \nabla_{\mathbf{z}} \mathcal{L}(\bar{\mathbf{z}}, \bar{\boldsymbol{\omega}})(\mathbf{z} - \bar{\mathbf{z}}) + \frac{1}{2} \nabla_{\mathbf{z}}^2 \mathcal{L}(\bar{\mathbf{z}} + t(\mathbf{z} - \bar{\mathbf{z}}), \bar{\boldsymbol{\omega}})(\mathbf{z} - \bar{\mathbf{z}})^2. \tag{4.11}$$

In the following we discuss each expression of (4.11) separately.

The zero-order terms:

Due to (3.1a), (3.1b), (4.8e) and (4.8f) it holds

$$\mathcal{L}(\bar{\mathbf{z}}, \bar{\boldsymbol{\omega}}) = J(\bar{\mathbf{u}}, \bar{\mathbf{g}}), \quad \mathcal{L}(\mathbf{z}, \bar{\boldsymbol{\omega}}) = J(\mathbf{u}, \mathbf{g}) - (\lambda, \bar{\mu}) + (\phi(\boldsymbol{\Sigma}), \bar{\theta}) \tag{4.12}$$

and Lemma 4.5 shows

$$\mathcal{L}(\mathbf{z}, \bar{\boldsymbol{\omega}}) \leq J(\mathbf{u}, \mathbf{g}) + o\left(\|\mathbf{g} - \bar{\mathbf{g}}\|_U^2\right). \tag{4.13}$$

The first-order term:

From (4.8a)–(4.8d) it follows

$$\begin{aligned} \nabla_z \mathcal{L}(\bar{z}, \bar{\omega})(z - \bar{z}) &= \underbrace{(A\bar{\mathbf{Y}} + B^*\bar{\omega} + \bar{\lambda}D^*D\bar{\mathbf{Y}} + \bar{\theta}D^*D\bar{\Sigma}, \Sigma - \bar{\Sigma})}_{=0} + \underbrace{\langle B\bar{\mathbf{Y}} + \partial_u J(\bar{u}, \bar{g}), u - \bar{u} \rangle}_{=0} \\ &+ \underbrace{(\partial_g J(\bar{u}, \bar{g}) + \tau_N \bar{\omega}, g - \bar{g})}_{=0} - (\lambda - \bar{\lambda}, \bar{\mu}) + \underbrace{(\lambda - \bar{\lambda}, D\bar{\Sigma} : D\bar{\mathbf{Y}})}_{=(\lambda - \bar{\lambda}, \bar{\mu})} = 0. \end{aligned} \quad (4.14)$$

The second-order term:

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

$$\begin{aligned} \nabla_z^2 \mathcal{L}(z_t, \bar{\omega})(z - \bar{z})^2 &= \nabla_{(u,g)}^2 J(u_t, g_t)(u - \bar{u}, g - \bar{g})^2 + 2((\lambda - \bar{\lambda})D^*D(\Sigma - \bar{\Sigma}), \bar{\mathbf{Y}}) \\ &+ \left(\|D\Sigma - D\bar{\Sigma}\|_{\mathbb{S}}^2, \bar{\theta} \right) \\ &= \underbrace{\nabla_{(u,g)}^2 J(\bar{u}, \bar{g})(u', g - \bar{g})^2}_{=:D_1} + \left(\nabla_{(u,g)}^2 J(u_t, g_t) - \nabla_{(u,g)}^2 J(\bar{u}, \bar{g}) \right) (u - \bar{u}, g - \bar{g})^2 \\ &+ 2\partial_g \partial_u J(\bar{u}, \bar{g})[u - \bar{u} - u', g - \bar{g}] + \partial_u^2 J(\bar{u}, \bar{g})(u - \bar{u} - u')^2 \\ &+ 2\partial_u^2 J(\bar{u}, \bar{g})[u - \bar{u} - u', u'] + 2((\lambda - \bar{\lambda})D^*D(\Sigma - \bar{\Sigma}), \bar{\mathbf{Y}}) \\ &+ \left(\|D\Sigma - D\bar{\Sigma}\|_{\mathbb{S}}^2, \bar{\theta} \right). \end{aligned}$$

The regularity condition (4.7) implies the existence of $\beta < p$ and $\gamma < q$ such that $\frac{1}{q} + \frac{1}{\beta} + \frac{1}{\eta} = \frac{1}{\gamma} + \frac{1}{p} + \frac{1}{\eta} = 1$, cf. the proof of Lemma 4.5. Hence due to Remark 4.3 (i)–(iii) it holds

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

Furthermore, (4.7) yields $\frac{1}{p} + \frac{1}{\beta} + \frac{1}{r} < 1$ and consequently

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

Recall that $D_1 + 2D_2 + D_3 = \nabla_z^2 \mathcal{L}(\bar{z}, \bar{\omega})(\Sigma', \mathbf{u}', \lambda', \mathbf{g} - \bar{\mathbf{g}})^2$. Thus, because of (SSC), Remark 4.3, Theorems 2.4 and 3.8, one obtains

$$\begin{aligned}
\nabla_z^2 \mathcal{L}(z_t, \bar{\omega})(z - \bar{z})^2 &\geq \nabla_z^2 \mathcal{L}(\bar{z}, \bar{\omega})(\Sigma', \mathbf{u}', \lambda', \mathbf{g} - \bar{\mathbf{g}})^2 \\
&\quad - c \|\mathbf{g} - \bar{\mathbf{g}}\|_U \left(\|\Sigma - \bar{\Sigma} - \Sigma'\|_{L^\beta(\Omega; \mathbb{S}^2)} + \|\lambda - \bar{\lambda} - \lambda'\|_{L^\gamma(\Omega)} \right) \\
&\quad - c \left\| \nabla_{(\mathbf{u}, \mathbf{g})}^2 J(\mathbf{u}_t, \mathbf{g}_t) - \nabla_{(\mathbf{u}, \mathbf{g})}^2 J(\bar{\mathbf{u}}, \bar{\mathbf{g}}) \right\|_{\mathcal{L}(V \times U, (V \times U)')} \|\mathbf{g} - \bar{\mathbf{g}}\|_U^2 \\
&\quad - c \|\partial_{\mathbf{g}} \partial_{\mathbf{u}} J(\bar{\mathbf{u}}, \bar{\mathbf{g}})\|_{\mathcal{L}(U, V')} \|\mathbf{u} - \bar{\mathbf{u}} - \mathbf{u}'\|_V \|\mathbf{g} - \bar{\mathbf{g}}\|_U \\
&\quad - c \|\partial_{\mathbf{u}}^2 J(\bar{\mathbf{u}}, \bar{\mathbf{g}})\|_{\mathcal{L}(V, V')} \|\mathbf{u} - \bar{\mathbf{u}} - \mathbf{u}'\|_V^2 \\
&\quad - c \|\partial_{\mathbf{u}}^2 J(\bar{\mathbf{u}}, \bar{\mathbf{g}})\|_{\mathcal{L}(V, V')} \|\mathbf{u} - \bar{\mathbf{u}} - \mathbf{u}'\|_V \|\mathbf{u}'\|_V \\
&\geq \alpha \|\mathbf{g} - \bar{\mathbf{g}}\|_U^2 - o\left(\|\mathbf{g} - \bar{\mathbf{g}}\|_U^2\right),
\end{aligned} \tag{4.15}$$

where we used that J is twice continuously differentiable by assumption. Altogether (4.11), (4.12) and (4.13)–(4.15) yield the existence of an $\epsilon > 0$ so that $J(\mathbf{u}, \mathbf{g}) \geq J(\bar{\mathbf{u}}, \bar{\mathbf{g}}) + \frac{\alpha}{4} \|\mathbf{g} - \bar{\mathbf{g}}\|_U^2$ for all $\mathbf{g} \in U$ with $\|\mathbf{g} - \bar{\mathbf{g}}\|_U \leq \epsilon$, i.e., the desired quadratic growth condition. \square

Remark 4.9. It is easily seen that, for every $p > 2$, there are numbers $s, \eta, \zeta, r \in [2, \infty[$ satisfying the conditions in (4.7). However, if p tends to 2, then the bounds for s, η, ζ , and r tend to ∞ . If \hat{p} as defined in (4.1) satisfies $\hat{p} > 4$, then the assumption on s in (4.7) is automatically fulfilled. In view of Sobolev's embedding theorem such high integrability can not be expected for controls in $U = L^2(\Gamma_N; \mathbb{R}^d)$ but for controls in $L^\nu(\Gamma_N; \mathbb{R}^d)$ with ν sufficiently high.

Remark 4.10. By using the technique introduced in [7], 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.

We now come to the second version of sufficient optimality conditions. If the objective functional satisfies additional properties, then we are able to tighten condition (SSC). To be more precise, we can restrict (SSC) to the cone of critical directions so that (SSC) becomes a classical condition (cf. [26], Thm. 7). Moreover we can weaken the integrability condition for the hardening variable in (4.7). In the following, as before, $(\bar{\mathbf{u}}, \bar{\Sigma}, \bar{\lambda}, \bar{\mathbf{g}}) \in V \times S^2 \times L^2(\Omega) \times U$ again denotes a stationary point. The precise assumptions on the objective then look as follows:

Assumption 4.11. The objective functional $J: V \times U \rightarrow \mathbb{R}$ is given by $J(\mathbf{u}, \mathbf{g}) = i(\mathbf{u}) + j(\mathbf{g})$, where the functions i, j satisfy:

- i) $i: V \rightarrow \mathbb{R}$ is twice continuously Fréchet-differentiable,
- ii) $j: U \rightarrow \mathbb{R}$ is twice continuously Fréchet-differentiable and there is $\nu > 0$ such that

$$\partial^2 j(\bar{\mathbf{g}})z^2 \geq \nu \|z\|_U^2 \quad \forall \mathbf{h} \in U.$$

Theorem 4.12. Let Assumption 4.11 hold. Suppose further that $\bar{\mathbf{g}} \in U$ together with its state $(\bar{\Sigma}, \bar{\mathbf{u}}, \bar{\lambda}) \in S^2 \times V \times L^2(\Omega)$, adjoint state $(\bar{\mathbf{Y}}, \bar{\mathbf{w}}) \in S^2 \times V$, and multipliers $(\bar{\mu}, \bar{\theta}) \in L^2(\Omega) \times L^2(\Omega)$ satisfy

- (1) the following regularity conditions:

$$\bar{\chi} \in L^s(\Omega; \mathbb{S}), \quad \bar{\mathbf{Y}} \in L^\eta(\Omega; \mathbb{S}^2), \quad \bar{\mu} \in L^\zeta(\Omega), \quad \bar{\theta} \in L^r(\Omega)$$

with

$$s > \frac{p}{p-2}, \quad \eta, \zeta, r > \frac{sp}{sp-p-2s}, \tag{4.16}$$

where p is as defined in (4.1),

(2) the optimality system (4.8) with

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

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

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

$$\partial_{(\mathbf{u}, \Sigma, \lambda, \mathbf{g})}^2 \mathcal{L}(\bar{\mathbf{u}}, \bar{\Sigma}, \bar{\lambda}, \bar{\mathbf{g}}, \bar{\mathbf{Y}}, \bar{\mathbf{w}}, \bar{\mu}, \bar{\theta})(\mathbf{u}', \Sigma', \lambda', \mathbf{h})^2 \geq \alpha \|\mathbf{h}\|_U^2 \tag{SSC}$$

for all $\mathbf{h} \in U$ and $(\Sigma', \mathbf{u}', \lambda')$ solving (3.4) with $i'(\bar{\mathbf{u}})\mathbf{u}' + j'(\bar{\mathbf{g}})\mathbf{h} = 0$.

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

$$J(\mathbf{u}, \mathbf{g}) \geq J(\bar{\mathbf{u}}, \bar{\mathbf{g}}) + \alpha \|\mathbf{g} - \bar{\mathbf{g}}\|_U^2 \tag{4.17}$$

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

Remark 4.13. A comparison to the second-order sufficient conditions for finite dimensional MPECs in [26] shows that the assertion of Theorem 4.6 can be seen as a natural generalization of the finite dimensional result in the following sense:

The second-order sufficient conditions in Theorem 7 from [26], Scheel and Scholtes involve the strong stationarity conditions in combination with the coercivity of the Hessian of the Lagrangian on the cone of critical directions. In our case this cone consists of all directions $(\mathbf{u}', \Sigma', \lambda', \mathbf{h})$ which solve (3.4) and satisfy $J'(\bar{\mathbf{u}}, \bar{\mathbf{g}})(\mathbf{u}', \mathbf{h}) = 0$. Thus the second-order sufficient conditions from [26] exactly coincide with these of Theorem 4.12 except that we had to slightly tighten the sign conditions of strong stationarity. As already mentioned in Remark 4.8, strong stationarity conditions are equivalent to (4.8) except that they require the sign conditions in (4.8g) and (4.8h) only on the smaller sets $\bar{\mathcal{A}}_s$ and $\bar{\mathcal{I}}$. However, when turning from component-wise inequality constraints in finite dimensions to pointwise constraints in a function space setting, one needs an additional 'safety distance', induced by the larger sets \mathcal{A}_1 and \mathcal{A}_1 . In this way the sufficient conditions in Theorem 4.12 can be regarded as a natural extension of their finite dimensional counterpart.

Proof of Theorem 4.12. The proof is similar to Theorem 2.12 from [20]. Nevertheless, for convenience of the reader, we recall the arguments. We argue by contradiction and assume that the quadratic growth condition (4.17) is violated. Then there exist a sequence $(\mathbf{g}_n)_{n \in \mathbb{N}} \subset U$ and a sequence of associated states $(\Sigma_n, \mathbf{u}_n, \lambda_n)_{n \in \mathbb{N}}$ with $(\Sigma_n, \mathbf{u}_n, \lambda_n, \mathbf{g}_n) \rightarrow (\bar{\Sigma}, \bar{\mathbf{u}}, \bar{\lambda}, \bar{\mathbf{g}})$ in $L^p(\Omega; \mathbb{S}^2) \times W_D^{1,p}(\Omega; \mathbb{R}^d) \times L^q(\Omega) \times U$, cf. Lemma 3.5, such that

$$J(\bar{\mathbf{u}}, \bar{\mathbf{g}}) + \frac{1}{n} \|\mathbf{g}_n - \bar{\mathbf{g}}\|_U^2 > J(\mathbf{u}_n, \mathbf{g}_n). \tag{4.18}$$

For the sake of convenience we introduce the abbreviations

$$\rho_n := \|\mathbf{g}_n - \bar{\mathbf{g}}\|_U, \quad \mathbf{h}_n := \frac{\mathbf{g}_n - \bar{\mathbf{g}}}{\rho_n}, \quad \mathbf{Z}_n := \frac{\Sigma_n - \bar{\Sigma}}{\rho_n}, \quad \mathbf{v}_n := \frac{\mathbf{u}_n - \bar{\mathbf{u}}}{\rho_n}, \quad \iota_n := \frac{\lambda_n - \bar{\lambda}}{\rho_n}.$$

Due to the boundedness of \mathbf{h}_n and the compactness of the trace operator there exists a subsequence, denoted in the same way, with

$$\mathbf{h}_n \rightharpoonup \mathbf{h} \quad \text{in } U, \quad \tau_N^*(\mathbf{h}_n) \rightarrow \tau_N^*(\mathbf{h}) \quad \text{in } W_D^{-1,\beta}(\Omega; \mathbb{R}^d) \quad \forall 1 \leq \beta < p. \tag{4.19}$$

Let $(\Sigma_{\rho_n}, \mathbf{u}_{\rho_n}, \lambda_{\rho_n})$ denote the state associated to $\bar{\mathbf{g}} + \rho_n \mathbf{h}$ and let $(\Sigma'_{\mathbf{h}}, \mathbf{u}'_{\mathbf{h}}, \lambda'_{\mathbf{h}})$ denote the solution of (3.4) with right-hand side $-\tau_N^*(\mathbf{h})$. Theorem 2.4, Corollary 3.12, Remark 3.13 and (4.19) yield

$$\begin{aligned} \|\mathbf{Z}_n - \Sigma'_{\mathbf{h}}\|_{L^\beta(\Omega; \mathbb{S}^2)} &\leq \left\| \frac{\Sigma_{\rho_n} - \bar{\Sigma}}{\rho_n} - \Sigma'_{\mathbf{h}} \right\|_{L^\beta(\Omega; \mathbb{S}^2)} + \left\| \frac{\Sigma_n - \Sigma_{\rho_n}}{\rho_n} \right\|_{L^\beta(\Omega; \mathbb{S}^2)} \\ &\leq \left\| \frac{\Sigma_{\rho_n} - \bar{\Sigma}}{\rho_n} - \Sigma'_{\mathbf{h}} \right\|_{L^\beta(\Omega; \mathbb{S}^2)} + L \left\| \frac{\tau_N^*(\mathbf{g}_n - \bar{\mathbf{g}} - \rho_n \mathbf{h})}{\rho_n} \right\|_{W_D^{-1, \beta}(\Omega; \mathbb{R}^d)} \\ &\leq \left\| \frac{\Sigma_{\rho_n} - \bar{\Sigma}}{\rho_n} - \Sigma'_{\mathbf{h}} \right\|_{L^\beta(\Omega; \mathbb{S}^2)} + L \|\tau_N^*(\mathbf{h}_n) - \tau_N^*(\mathbf{h})\|_{W_D^{-1, \beta}(\Omega; \mathbb{R}^d)} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Similar arguments hold true for \mathbf{v}_n and ι_n so that we get

$$\mathbf{Z}_n \rightarrow \Sigma'_{\mathbf{h}} \quad \text{in } L^\beta(\Omega; \mathbb{S}^2) \quad \forall 1 \leq \beta < p, \quad (4.20a)$$

$$\mathbf{v}_n \rightarrow \mathbf{u}'_{\mathbf{h}} \quad \text{in } W_D^{1, \beta}(\Omega; \mathbb{R}^d) \quad \forall 1 \leq \beta < p, \quad (4.20b)$$

$$\iota_n \rightarrow \lambda'_{\mathbf{h}} \quad \text{in } L^\gamma(\Omega) \quad \forall 1 \leq \gamma < q. \quad (4.20c)$$

with q as defined in (3.6). Thanks to (3.1a), (3.1b) and (4.8a)–(4.8c) we obtain

$$\begin{aligned} J(\mathbf{u}_n, \mathbf{g}_n) - J(\bar{\mathbf{u}}, \bar{\mathbf{g}}) &= i'(\bar{\mathbf{u}})(\mathbf{u}_n - \bar{\mathbf{u}}) + j'(\bar{\mathbf{g}})(\mathbf{g}_n - \bar{\mathbf{g}}) + \frac{1}{2}i''(\tilde{\mathbf{u}}_n)(\mathbf{u}_n - \bar{\mathbf{u}})^2 + \frac{1}{2}j''(\tilde{\mathbf{g}}_n)(\mathbf{g}_n - \bar{\mathbf{g}})^2 \\ &\quad + \underbrace{(A(\Sigma_n - \bar{\Sigma}) + B^*(\mathbf{u}_n - \bar{\mathbf{u}}) + \lambda_n \mathcal{D}^* \mathcal{D} \Sigma_n - \bar{\lambda} \mathcal{D}^* \mathcal{D} \bar{\Sigma}, \bar{\mathbf{Y}})}_{=0} \\ &\quad + \underbrace{\langle B(\Sigma_n - \bar{\Sigma}) + \tau_N^*(\mathbf{g}_n) - \tau_N^*(\bar{\mathbf{g}}), \bar{\mathbf{w}} \rangle}_{=0} + (\bar{\lambda} \mathcal{D}^* \mathcal{D} \Sigma_n, \bar{\mathbf{Y}}) - (\bar{\lambda} \mathcal{D}^* \mathcal{D} \Sigma_n, \bar{\mathbf{Y}}) \\ &= i'(\bar{\mathbf{u}})(\mathbf{u}_n - \bar{\mathbf{u}}) + j'(\bar{\mathbf{g}})(\mathbf{g}_n - \bar{\mathbf{g}}) + \frac{1}{2}i''(\tilde{\mathbf{u}}_n)(\mathbf{u}_n - \bar{\mathbf{u}})^2 + \frac{1}{2}j''(\tilde{\mathbf{g}}_n)(\mathbf{g}_n - \bar{\mathbf{g}})^2 \\ &\quad + (A\bar{\mathbf{Y}} + B^*\bar{\mathbf{w}} + \bar{\lambda} \mathcal{D}^* \mathcal{D} \bar{\mathbf{Y}}, \Sigma_n - \bar{\Sigma}) + \langle B\bar{\mathbf{Y}}, \mathbf{u}_n - \bar{\mathbf{u}} \rangle + (\tau_N \bar{\mathbf{w}}, \mathbf{g}_n - \bar{\mathbf{g}}) \\ &\quad + ((\lambda_n - \bar{\lambda}) \mathcal{D}^* \mathcal{D} \bar{\mathbf{Y}}, \Sigma_n) \\ &= \frac{1}{2}i''(\tilde{\mathbf{u}}_n)(\mathbf{u}_n - \bar{\mathbf{u}})^2 + \frac{1}{2}j''(\tilde{\mathbf{g}}_n)(\mathbf{g}_n - \bar{\mathbf{g}})^2 - (\bar{\theta} \mathcal{D}^* \mathcal{D} \bar{\Sigma}, \Sigma_n - \bar{\Sigma}) \\ &\quad + ((\lambda_n - \bar{\lambda}) \mathcal{D}^* \mathcal{D} \bar{\mathbf{Y}}, \Sigma_n), \end{aligned}$$

where $(\tilde{\mathbf{u}}_n, \tilde{\mathbf{g}}_n)$ denotes an element between $(\bar{\mathbf{u}}, \bar{\mathbf{g}})$ and $(\mathbf{u}_n, \mathbf{g}_n)$. In view of (4.18) it holds

$$\begin{aligned} &((\lambda_n - \bar{\lambda}) \mathcal{D}^* \mathcal{D} \bar{\mathbf{Y}}, \Sigma_n) - (\bar{\theta} \mathcal{D}^* \mathcal{D} \bar{\Sigma}, \Sigma_n - \bar{\Sigma}) \\ &< \frac{\rho_n^2}{n} - \frac{1}{2} (i''(\tilde{\mathbf{u}}_n)(\mathbf{u}_n - \bar{\mathbf{u}})^2 + j''(\tilde{\mathbf{g}}_n)(\mathbf{g}_n - \bar{\mathbf{g}})^2), \end{aligned}$$

which is equivalent to

$$\begin{aligned} &(\iota_n \mathcal{D}^* \mathcal{D} \bar{\mathbf{Y}}, \Sigma_n) - (\bar{\theta} \mathcal{D}^* \mathcal{D} \bar{\Sigma}, \mathbf{Z}_n) \\ &< \frac{\rho_n}{n} - \frac{1}{2} (i''(\tilde{\mathbf{u}}_n)[\mathbf{u}_n - \bar{\mathbf{u}}, \mathbf{v}_n] + j''(\tilde{\mathbf{g}}_n)[\mathbf{g}_n - \bar{\mathbf{g}}, \mathbf{h}_n]). \end{aligned} \quad (4.21)$$

The integrability conditions give $1/\gamma + 1/\eta + 1/p < 1$ and $1/r + 1/\beta < 1$ for some $1 \leq \gamma < q$ and $1 \leq \beta < p$ so that Lemma 3.5, (4.20a) and (4.20c) imply

$$(\iota_n \mathcal{D}^* \mathcal{D} \bar{\mathbf{Y}}, \Sigma_n) \rightarrow (\lambda'_{\mathbf{h}} \mathcal{D}^* \mathcal{D} \bar{\mathbf{Y}}, \bar{\Sigma}), \quad (\bar{\theta} \mathcal{D}^* \mathcal{D} \bar{\Sigma}, \mathbf{Z}_n) \rightarrow (\bar{\theta} \mathcal{D}^* \mathcal{D} \bar{\Sigma}, \Sigma'_{\mathbf{h}}), \quad (4.22)$$

where we used $\|\mathcal{D}\bar{\Sigma}\|_{\mathfrak{S}} \leq \sigma_0$ for the last convergence. Since the second derivatives of i and j are continuous bilinear forms by Assumption 4.11 we infer

$$|i''(\tilde{\mathbf{u}}_n)[\mathbf{u}_n - \bar{\mathbf{u}}, \mathbf{v}_n]| \leq \|i''(\tilde{\mathbf{u}}_n)\|_{\mathcal{L}(V, V')} \|\mathbf{u}_n - \bar{\mathbf{u}}\|_V \|\mathbf{v}_n\|_V \leq c \|\mathbf{u}_n - \bar{\mathbf{u}}\|_V,$$

since $(\tilde{\mathbf{u}}_n)_{n \in \mathbb{N}}$ and $(\mathbf{v}_n)_{n \in \mathbb{N}}$ are bounded due to (weak) convergence. Together with an analogous estimate for $j''(\tilde{\mathbf{g}}_n)[\mathbf{g}_n - \bar{\mathbf{g}}, \mathbf{h}_n]$, (4.19) and (4.20b) consequently imply

$$i''(\tilde{\mathbf{u}}_n)[\mathbf{u}_n - \bar{\mathbf{u}}, \mathbf{v}_n] \rightarrow 0, \quad j''(\tilde{\mathbf{g}}_n)[\mathbf{g}_n - \bar{\mathbf{g}}, \mathbf{h}_n] \rightarrow 0. \quad (4.23)$$

Note that \mathbf{v}_n and \mathbf{h}_n are bounded due to (weak) convergence. From (4.21)–(4.23) we deduce

$$(\lambda'_h, \mathcal{D}\bar{\Upsilon} : \mathcal{D}\bar{\Sigma}) - (\bar{\theta}, \mathcal{D}\bar{\Sigma} : \mathcal{D}\Sigma'_h) \leq 0. \quad (4.24)$$

Moreover (3.4c)–(3.4e) and (4.8d)–(4.8h) yield

$$\begin{aligned} \lambda'_h = \bar{\theta} = 0 & \quad \text{in } \bar{\mathcal{L}}, & \mathcal{D}\bar{\Upsilon} : \mathcal{D}\bar{\Sigma} = \mathcal{D}\bar{\Sigma} : \mathcal{D}\Sigma'_h = 0 & \quad \text{in } \bar{\mathcal{A}}_s, \\ \lambda'_h \geq 0 & \quad \text{in } \bar{\mathcal{B}}, & \mathcal{D}\bar{\Upsilon} : \mathcal{D}\bar{\Sigma} \geq 0 & \quad \text{in } \bar{\mathcal{B}}, \\ \bar{\theta} \geq 0 & \quad \text{in } \bar{\mathcal{B}}, & \mathcal{D}\bar{\Sigma} : \mathcal{D}\Sigma'_h \leq 0 & \quad \text{in } \bar{\mathcal{B}}, \end{aligned}$$

so that

$$(\lambda'_h, \mathcal{D}\bar{\Upsilon} : \mathcal{D}\bar{\Sigma}) - (\bar{\theta}, \mathcal{D}\bar{\Sigma} : \mathcal{D}\Sigma'_h) \geq 0.$$

Together with (4.24) this gives

$$(\lambda'_h, \mathcal{D}\bar{\Upsilon} : \mathcal{D}\bar{\Sigma}) - (\bar{\theta}, \mathcal{D}\bar{\Sigma} : \mathcal{D}\Sigma'_h) = 0. \quad (4.25)$$

Because of (4.8a)–(4.8c), (3.4a) and (3.4b) we get the identity

$$\begin{aligned} (\lambda'_h, \mathcal{D}\bar{\Upsilon} : \mathcal{D}\bar{\Sigma}) - (\bar{\theta}, \mathcal{D}\bar{\Sigma} : \mathcal{D}\Sigma'_h) &= (\lambda'_h \mathcal{D}^* \mathcal{D}\bar{\Sigma}, \bar{\Upsilon}) - (\bar{\theta} \mathcal{D}^* \mathcal{D}\bar{\Sigma}, \Sigma'_h) \\ &= -(A\Sigma'_h + B^* \mathbf{u}'_h + \bar{\lambda} \mathcal{D}^* \mathcal{D}\Sigma'_h, \bar{\Upsilon}) + (A\bar{\Upsilon} + B^* \bar{\mathbf{w}} + \bar{\lambda} \mathcal{D}^* \mathcal{D}\bar{\Upsilon}, \Sigma'_h) \\ &= \langle B\Sigma'_h, \bar{\mathbf{w}} \rangle - \langle B\bar{\Upsilon}, \mathbf{u}'_h \rangle = -(\tau_N \bar{\mathbf{w}}, \mathbf{h}) - \langle B\bar{\Upsilon}, \mathbf{u}'_h \rangle = i'(\bar{\mathbf{u}}) \mathbf{u}'_h + j'(\bar{\mathbf{g}}) \mathbf{h} \end{aligned}$$

and in view of (4.25) we thus arrive at

$$i'(\bar{\mathbf{u}}) \mathbf{u}'_h + j'(\bar{\mathbf{g}}) \mathbf{h} = 0. \quad (4.26)$$

As in the proof of Theorem 4.6 we introduce the abbreviations $\mathbf{z}_n := (\mathbf{u}_n, \Sigma_n, \lambda_n, \mathbf{g}_n)$, $\bar{\mathbf{z}} := (\bar{\mathbf{u}}, \bar{\Sigma}, \bar{\lambda}, \bar{\mathbf{g}})$, and $\bar{\omega} := (\bar{\Upsilon}, \bar{\mathbf{w}}, \bar{\mu}, \bar{\theta})$. Analogously to (4.12) and (4.14) we obtain $\mathcal{L}(\bar{\mathbf{z}}, \bar{\omega}) = J(\bar{\mathbf{u}}, \bar{\mathbf{g}})$, $\mathcal{L}(\mathbf{z}_n, \bar{\omega}) = J(\mathbf{u}_n, \mathbf{g}_n) - (\lambda_n, \bar{\mu}) + (\phi(\Sigma_n), \bar{\theta})$, and $\nabla_{\mathbf{z}} \mathcal{L}(\bar{\mathbf{z}}, \bar{\omega})(\mathbf{z}_n - \bar{\mathbf{z}}) = 0$. Hence we have

$$\begin{aligned} J(\mathbf{u}_n, \mathbf{g}_n) - J(\bar{\mathbf{u}}, \bar{\mathbf{g}}) &= \mathcal{L}(\mathbf{z}_n, \bar{\omega}) + (\lambda_n, \bar{\mu}) - (\phi(\Sigma_n), \bar{\theta}) - \mathcal{L}(\bar{\mathbf{z}}, \bar{\omega}) \\ &= \frac{1}{2} \nabla_{\mathbf{z}}^2 \mathcal{L}(\bar{\mathbf{z}}, \bar{\omega})(\mathbf{z}_n - \bar{\mathbf{z}})^2 + (\lambda_n, \bar{\mu}) - (\phi(\Sigma_n), \bar{\theta}), \end{aligned} \quad (4.27)$$

where $\tilde{\mathbf{z}}_n$ denotes an element between $\bar{\mathbf{z}}$ and \mathbf{z}_n , so that (4.18) and (4.27) give

$$\frac{1}{2} \nabla_{\mathbf{z}}^2 \mathcal{L}(\tilde{\mathbf{z}}_n, \bar{\omega})(\mathbf{z}_n - \bar{\mathbf{z}})^2 < \frac{\rho_n^2}{n} - (\lambda_n, \bar{\mu}) + (\phi(\Sigma_n), \bar{\theta}). \quad (4.28)$$

From Lemma 4.5 we infer

$$\frac{1}{2} \nabla_{\mathbf{z}}^2 \mathcal{L}(\tilde{\mathbf{z}}_n, \bar{\omega}) \left(\frac{\mathbf{z}_n - \bar{\mathbf{z}}}{\rho_n} \right)^2 < \frac{1}{n} + \frac{o(\rho_n^2)}{\rho_n^2}$$

or equivalently

$$\frac{1}{2} \nabla_z^2 \mathcal{L}(\bar{z}, \bar{\omega}) \left(\frac{z_n - \bar{z}}{\rho_n} \right)^2 < \frac{1}{n} + \frac{o(\rho_n^2)}{\rho_n^2} - \frac{1}{2} \left(\nabla_z^2 \mathcal{L}(\tilde{z}_n, \bar{\omega}) \left(\frac{z_n - \bar{z}}{\rho_n} \right)^2 - \nabla_z^2 \mathcal{L}(\bar{z}, \bar{\omega}) \left(\frac{z_n - \bar{z}}{\rho_n} \right)^2 \right). \quad (4.29)$$

Recalling (4.10), we notice

$$\nabla_z^2 \mathcal{L}(\tilde{z}_n, \bar{\omega}) \left(\frac{z_n - \bar{z}}{\rho_n} \right)^2 - \nabla_z^2 \mathcal{L}(\bar{z}, \bar{\omega}) \left(\frac{z_n - \bar{z}}{\rho_n} \right)^2 = [\nabla_{(\mathbf{u}, \mathbf{g})}^2 J(\tilde{\mathbf{u}}_n, \tilde{\mathbf{g}}_n) - \nabla_{(\mathbf{u}, \mathbf{g})}^2 J(\bar{\mathbf{u}}, \bar{\mathbf{g}})] \left(\frac{\mathbf{u}_n - \bar{\mathbf{u}}}{\rho_n}, \frac{\mathbf{g}_n - \bar{\mathbf{g}}}{\rho_n} \right)^2. \quad (4.30)$$

Since $\nabla_{(\mathbf{u}, \mathbf{g})}^2 J$ is continuous by assumption and the sequence $(z_n - \bar{z})/\rho_n$ is bounded because of (weak) convergence, cf. (4.19) and (4.20), the right-hand side of (4.29) converges to zero. For the discussion of the left-hand side we use

$$\nabla_z^2 \mathcal{L}(\bar{z}, \bar{\omega}) \left(\frac{z_n - \bar{z}}{\rho_n} \right)^2 = i''(\bar{\mathbf{u}})(\mathbf{v}_n)^2 + j''(\bar{\mathbf{g}})(\mathbf{h}_n)^2 + 2(\iota_n \mathcal{D}^* \mathcal{D} \mathbf{Z}_n, \bar{\mathbf{Y}}) + (\|\mathbf{Z}_n\|_{\mathbb{S}}^2, \bar{\theta}), \quad (4.31)$$

cf. (4.10). By Assumption 4.11 we know

$$j''(\bar{\mathbf{g}})(\mathbf{h})^2 \leq \liminf_{k \rightarrow \infty} j''(\bar{\mathbf{g}})(\mathbf{h}_k)^2. \quad (4.32)$$

Moreover we can chose $1 \leq \beta < p$ and $1 \leq \gamma < q$ such that $1/\beta + 1/\gamma + 1/\eta \leq 1$, $2/\beta + 1/r \leq 1$ due to (4.16) and get

$$\begin{aligned} (\iota_n \mathcal{D}^* \mathcal{D} \mathbf{Z}_n, \bar{\mathbf{Y}}) &\rightarrow (\lambda'_h \mathcal{D}^* \mathcal{D} \Sigma'_h, \bar{\mathbf{Y}}), \\ (\|\mathbf{Z}_n\|_{\mathbb{S}}^2, \bar{\theta}) &\rightarrow (\|\Sigma'_h\|_{\mathbb{S}}^2, \bar{\theta}). \end{aligned} \quad (4.33)$$

From (4.20b) and (4.29)–(4.33) we deduce

$$\nabla_z^2 \mathcal{L}(\bar{z}, \bar{\omega}) (\Sigma'_h, \mathbf{u}'_h, \lambda'_h, \mathbf{h})^2 = \lim_{n \rightarrow \infty} \nabla_z^2 \mathcal{L}(\bar{z}, \bar{\omega}) \left(\frac{z_n - \bar{z}}{\rho_n} \right)^2 \leq 0,$$

so that $(\widetilde{\text{SSC}})$ and (4.26) imply $\mathbf{h} = 0$. As (3.4) is uniquely solvable we conclude that $\Sigma'_h = \mathbf{0}$, $\mathbf{u}'_h = \mathbf{0}$ and $\lambda'_h = 0$. By definition it holds $\|\mathbf{h}_n\|_U = 1$. Therefore (4.29), (4.31) and Assumption 4.11 yield

$$\begin{aligned} 0 &< \frac{\nu}{2} \leq \frac{1}{2} j''(\bar{\mathbf{g}})(\mathbf{h}_n)^2 \\ &< -\frac{1}{2} [\nabla_{(\mathbf{u}, \mathbf{g})}^2 J(\tilde{\mathbf{u}}_n, \tilde{\mathbf{g}}_n) - \nabla_{(\mathbf{u}, \mathbf{g})}^2 J(\bar{\mathbf{u}}, \bar{\mathbf{g}})] \left(\frac{\mathbf{u}_n - \bar{\mathbf{u}}}{\rho_n}, \frac{\mathbf{g}_n - \bar{\mathbf{g}}}{\rho_n} \right)^2 \\ &\quad + \frac{1}{n} + \frac{o(\rho_n^2)}{\rho_n^2} - \frac{1}{2} i''(\bar{\mathbf{u}})(\mathbf{v}_n)^2 - (\iota_n \mathcal{D}^* \mathcal{D} \mathbf{Z}_n, \bar{\mathbf{Y}}) - \frac{1}{2} (\|\mathbf{Z}_n\|_{\mathbb{S}}^2, \bar{\theta}). \end{aligned}$$

In consequence of the continuity of $\nabla_{(\mathbf{u}, \mathbf{g})}^2 J$ and (4.20) the right-hand side converges to zero, which is a contradiction. \square

We end the paper with an equivalent formulation of $(\widetilde{\text{SSC}})$ in Theorem 4.17 below. For this purpose we first state some auxiliary results:

Lemma 4.14. *Let $\mathbf{h}, \mathbf{h}_n \in U$ be given and let $(\Sigma'_n, \mathbf{u}'_n, \lambda'_n)$ be the solution of (3.4) with right-hand side $-\tau_N^*(\mathbf{h}_n)$. If $(\Sigma'_n, \mathbf{u}'_n, \lambda'_n, \mathbf{h}_n) \rightarrow (\Sigma', \mathbf{u}', \lambda', \mathbf{h})$ in $S^2 \times V \times L^2(\Omega) \times U$, then $(\Sigma', \mathbf{u}', \lambda', \mathbf{h})$ solves (3.4).*

Proof. Since $(\boldsymbol{\Sigma}'_n, \mathbf{u}'_n, \mathbf{h}_n)$ solves (3.3), it holds

$$\begin{aligned} & (A\boldsymbol{\Sigma}'_n, \mathbf{T}) - (A\boldsymbol{\Sigma}'_n, \boldsymbol{\Sigma}'_n) + (B^*\mathbf{u}'_n, \mathbf{T}) - (B^*\mathbf{u}'_n, \boldsymbol{\Sigma}'_n) \\ & + (\bar{\lambda}, \mathcal{D}\boldsymbol{\Sigma}'_n : \mathcal{D}\mathbf{T}) - \left(\bar{\lambda}, \|\mathcal{D}\boldsymbol{\Sigma}'_n\|_{\mathbb{S}}^2 \right) \geq 0 \quad \forall \mathbf{T} \in \mathcal{S}_\ell, \end{aligned} \quad (4.34a)$$

$$B\boldsymbol{\Sigma}'_n = -\tau_N^*(\mathbf{h}_n). \quad (4.34b)$$

For $\mathbf{T} \in \mathcal{S}_\ell$ fixed but arbitrary the weak continuity of A and B implies

$$(A\boldsymbol{\Sigma}'_n, \mathbf{T}) \rightarrow (A\boldsymbol{\Sigma}', \mathbf{T}), \quad (B^*\mathbf{u}'_n, \mathbf{T}) \rightarrow (B^*\mathbf{u}', \mathbf{T}). \quad (4.35)$$

Due to (4.34b) and the compactness of the trace operator we infer

$$(B^*\mathbf{u}'_n, \boldsymbol{\Sigma}'_n) = \langle \mathbf{u}'_n, -\tau_N^*(\mathbf{h}_n) \rangle \rightarrow \langle \mathbf{u}', -\tau_N(\mathbf{h}) \rangle = (B^*\mathbf{u}', \boldsymbol{\Sigma}'). \quad (4.36)$$

As in the proof of Lemma 3.6, cf. (3.15), we derive

$$\bar{\lambda}\mathcal{D}\boldsymbol{\Sigma}'_n = -\mathbb{H}^{-1}\boldsymbol{\chi}'_n - \lambda'_n\mathcal{D}\bar{\boldsymbol{\Sigma}}.$$

The right-hand side converges weakly in S because of the boundedness of $\mathcal{D}\bar{\boldsymbol{\Sigma}}$ and accordingly

$$\bar{\lambda}\mathcal{D}\boldsymbol{\Sigma}'_n \rightharpoonup \bar{\lambda}\mathcal{D}\boldsymbol{\Sigma}' = -\mathbb{H}^{-1}\boldsymbol{\chi}' - \lambda'\mathcal{D}\bar{\boldsymbol{\Sigma}} \quad \text{in } S. \quad (4.37)$$

Therefore we deduce

$$(\bar{\lambda}, \mathcal{D}\boldsymbol{\Sigma}'_n : \mathcal{D}\mathbf{T}) \rightarrow (\bar{\lambda}, \mathcal{D}\boldsymbol{\Sigma}' : \mathcal{D}\mathbf{T}). \quad (4.38)$$

The operator A is continuous and coercive. Therefore $S^2 \ni \mathbf{T} \mapsto -(A\mathbf{T}, \mathbf{T}) \in \mathbb{R}$ is continuous and concave so that

$$\limsup_{n \rightarrow \infty} (-(A\boldsymbol{\Sigma}'_n, \boldsymbol{\Sigma}'_n)) \leq -(A\boldsymbol{\Sigma}', \boldsymbol{\Sigma}'). \quad (4.39)$$

Next, we introduce the space

$$S_\lambda^2 := \left\{ \mathbf{T} \in S^2 : \sqrt{\bar{\lambda}}\mathcal{D}\mathbf{T} \in S \right\}$$

endowed with the scalarproduct

$$(\boldsymbol{\Sigma}, \mathbf{T})_{S_\lambda^2} = (\boldsymbol{\Sigma}, \mathbf{T}) + (\bar{\lambda}, \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\mathbf{T}).$$

In the following we show that S_λ^2 is a Hilbert space. Let $(\mathbf{T}_n)_{n \in \mathbb{N}}$ be a Cauchy sequence in S_λ^2 . Then $(\mathbf{T}_n)_{n \in \mathbb{N}}$ and $(\sqrt{\bar{\lambda}}\mathcal{D}\mathbf{T}_n)_{n \in \mathbb{N}}$ are Cauchy sequences in the Hilbert spaces S^2 and S . Consequently, there exist $\tilde{\mathbf{T}} \in S^2$ and $\tilde{\boldsymbol{\mu}} \in S$ such that $\mathbf{T}_n \rightarrow \tilde{\mathbf{T}}$ in S^2 and $\sqrt{\bar{\lambda}}\mathcal{D}\mathbf{T}_n \rightarrow \tilde{\boldsymbol{\mu}}$ in S . Moreover there are subsequences converging pointwisely a.e. in Ω so that $\tilde{\boldsymbol{\mu}} = \sqrt{\bar{\lambda}}\mathcal{D}\tilde{\mathbf{T}}$ a.e. in Ω . Thus, S_λ^2 is complete and therefore a Hilbert space. Finally, we define the continuous function

$$f: S_\lambda^2 \rightarrow \mathbb{R}, \quad \mathbf{T} \mapsto (\bar{\lambda}, \mathcal{D}\mathbf{T} : \mathcal{D}\mathbf{T}).$$

Thanks to (3.1c) $-f$ is concave and hence weakly upper semicontinuous. Since S_λ^2 is a Hilbert space, all linear functionals on S_λ^2 can be represented by scalarproducts. From (4.37) we thus infer $\boldsymbol{\Sigma}'_n \rightharpoonup \boldsymbol{\Sigma}'$ in S_λ^2 so that

$$\limsup_{n \rightarrow \infty} (-f(\boldsymbol{\Sigma}'_n)) \leq -\left(\bar{\lambda}, \|\mathcal{D}\boldsymbol{\Sigma}'\|_{\mathbb{S}}^2 \right). \quad (4.40)$$

Altogether (4.34)–(4.36), (4.38)–(4.40) yield that $(\boldsymbol{\Sigma}', \mathbf{u}')$ satisfies (3.3a). The weak continuity of B and the compactness of the trace operator give (3.3b). It remains to show that $\boldsymbol{\Sigma}' \in \mathcal{S}_\ell$. Testing with $\mathcal{D}\boldsymbol{\Sigma}'$ in (4.37) we get $\sqrt{\bar{\lambda}}\mathcal{D}\boldsymbol{\Sigma}' \in S$ because of the boundedness of $\mathcal{D}\bar{\boldsymbol{\Sigma}}$. Since equality and inequality conditions remain valid for the weak limit, we have $\boldsymbol{\Sigma}' \in \mathcal{S}_\ell$. Thus $(\boldsymbol{\Sigma}', \mathbf{u}')$ solves (3.3), i.e. the VI associated to the directional derivative. Furthermore, since λ' fulfills (4.37), the same argument that led from (3.15) to (3.16) shows that λ' coincides with the multiplier in (3.4). Hence $(\boldsymbol{\Sigma}', \mathbf{u}', \lambda')$ indeed solves (3.4). \square

Corollary 4.15. *Let $\mathbf{h}, \mathbf{h}_n \in U$ be given and let $(\boldsymbol{\Sigma}'_n, \mathbf{u}'_n, \lambda'_n)$ be the solution of (3.4) with right-hand side $-\tau_N^*(\mathbf{h}_n)$. If $(\boldsymbol{\Sigma}'_n, \mathbf{u}'_n, \lambda'_n, \mathbf{h}_n) \rightarrow (\boldsymbol{\Sigma}', \mathbf{u}', \lambda', \mathbf{h})$ in $S^2 \times V \times L^2(\Omega) \times U$, then $(\boldsymbol{\Sigma}'_n, \mathbf{u}'_n) \rightarrow (\boldsymbol{\Sigma}', \mathbf{u}')$ in $S^2 \times V$.*

Proof. By Lemma 4.14 we know that $(\boldsymbol{\Sigma}', \mathbf{u}', \lambda', \mathbf{h})$ solves (3.4). Subtracting (3.4a) from the corresponding equation for $(\boldsymbol{\Sigma}'_n, \mathbf{u}'_n, \lambda'_n)$ and testing with $\boldsymbol{\Sigma}' - \boldsymbol{\Sigma}'_n$ yields

$$\begin{aligned} & (A(\boldsymbol{\Sigma}' - \boldsymbol{\Sigma}'_n), \boldsymbol{\Sigma}' - \boldsymbol{\Sigma}'_n) + (B^*(\mathbf{u}' - \mathbf{u}'_n), \boldsymbol{\Sigma}' - \boldsymbol{\Sigma}'_n) \\ & + (\bar{\lambda}\mathcal{D}(\boldsymbol{\Sigma}' - \boldsymbol{\Sigma}'_n), \mathcal{D}(\boldsymbol{\Sigma}' - \boldsymbol{\Sigma}'_n)) + ((\lambda' - \lambda'_n)\mathcal{D}\bar{\boldsymbol{\Sigma}}, \mathcal{D}(\boldsymbol{\Sigma}' - \boldsymbol{\Sigma}'_n)) = 0. \end{aligned}$$

In view of the coercivity of A it follows

$$\begin{aligned} c \|\boldsymbol{\Sigma}' - \boldsymbol{\Sigma}'_n\|_{S^2}^2 & \leq \underbrace{(B^*(\mathbf{u}' - \mathbf{u}'_n), \boldsymbol{\Sigma}' - \boldsymbol{\Sigma}'_n)}_{=: I_n} - \underbrace{(\bar{\lambda}, \|\mathcal{D}(\boldsymbol{\Sigma}' - \boldsymbol{\Sigma}'_n)\|_F^2)}_{=: II_n} \\ & \quad - \underbrace{((\lambda' - \lambda'_n)\mathcal{D}\bar{\boldsymbol{\Sigma}}, \mathcal{D}(\boldsymbol{\Sigma}' - \boldsymbol{\Sigma}'_n))}_{=: III_n}. \end{aligned} \quad (4.41)$$

By the same arguments as in the proof of Lemma 4.14, cf. (4.36) and (4.40), we obtain

$$I_n \rightarrow 0 \quad \text{and} \quad \limsup_{n \rightarrow \infty} (-II_n) \leq 0. \quad (4.42)$$

Due to (3.4c)–(3.4e) and the boundedness of $\mathcal{D}\bar{\boldsymbol{\Sigma}}$, the assumed weak convergence yields

$$\begin{aligned} III_n & = \underbrace{(\lambda'\mathcal{D}\bar{\boldsymbol{\Sigma}}, \mathcal{D}\boldsymbol{\Sigma}')}_{=0} - (\lambda'_n\mathcal{D}\bar{\boldsymbol{\Sigma}}, \mathcal{D}\boldsymbol{\Sigma}') - (\lambda'\mathcal{D}\bar{\boldsymbol{\Sigma}}, \mathcal{D}\boldsymbol{\Sigma}'_n) + \underbrace{(\lambda'_n\mathcal{D}\bar{\boldsymbol{\Sigma}}, \mathcal{D}\boldsymbol{\Sigma}'_n)}_{=0} \\ & = -(\lambda'_n\mathcal{D}\bar{\boldsymbol{\Sigma}}, \mathcal{D}\boldsymbol{\Sigma}') - (\lambda'\mathcal{D}\bar{\boldsymbol{\Sigma}}, \mathcal{D}\boldsymbol{\Sigma}'_n) \xrightarrow{n \rightarrow \infty} 0. \end{aligned} \quad (4.43)$$

Altogether (4.41)–(4.43) imply $\boldsymbol{\Sigma}'_n \rightarrow \boldsymbol{\Sigma}'$ in S^2 . To prove the strong convergence of \mathbf{u}'_n we subtract (3.4a) from the corresponding equation for $(\boldsymbol{\Sigma}'_n, \mathbf{u}'_n, \lambda'_n)$ and test with $\bar{\mathbf{T}} := (\varepsilon(\mathbf{u}') - \varepsilon(\mathbf{u}'_n), -\varepsilon(\mathbf{u}') + \varepsilon(\mathbf{u}'_n)) \in S^2$. Similarly to (3.28)–(3.29) we conclude $\mathbf{u}'_n \rightarrow \mathbf{u}'$ in V . \square

Again, from Lemma A.2 we infer

Corollary 4.16. *Let $\mathbf{h}, \mathbf{h}_n \in U$ be given and let $(\boldsymbol{\Sigma}'_n, \mathbf{u}'_n, \lambda'_n)$ be the solution of (3.4) with right-hand side $-\tau_N^*(\mathbf{h}_n)$. If $(\boldsymbol{\Sigma}'_n, \mathbf{u}'_n, \lambda'_n, \mathbf{h}_n) \rightarrow (\boldsymbol{\Sigma}', \mathbf{u}', \lambda', \mathbf{h})$ in $L^p(\Omega; \mathbb{S}^2) \times V \times L^2(\Omega) \times U$ with $p \geq 2$, then $\boldsymbol{\Sigma}'_n \rightarrow \boldsymbol{\Sigma}'$ in $L^\beta(\Omega; \mathbb{S}^2)$ for all $1 \leq \beta < p$.*

Proof. Thanks to Lemma 4.15 there exists a subsequence, denoted in the same way, with $\boldsymbol{\Sigma}'_n \rightarrow \boldsymbol{\Sigma}'$ a.e. in Ω . Moreover the subsequence is bounded in $L^p(\Omega; \mathbb{S}^2)$ because of the weak convergence so that Lemma A.2 implies the claim for the subsequence. By a standard argument we obtain the result for the whole sequence. \square

We are now in the position to state the equivalent reformulation of ($\widetilde{\text{SSC}}$). In case of optimal control of PDEs such a result is well known, see for instance [5] and, based on the above findings, especially Corollaries 4.15 and 4.16, the same arguments also apply in our case.

Theorem 4.17. *Let Assumption 4.11 hold. Furthermore let $(\bar{\boldsymbol{\Sigma}}, \bar{\mathbf{u}}, \bar{\lambda}) \in S^2 \times V \times L^2(\Omega)$ be the state associated to $\bar{\mathbf{g}} \in U$. If $\bar{\boldsymbol{\chi}} \in L^s(\Omega; \mathbb{S})$, $(\bar{\mathbf{Y}}, \bar{\mathbf{w}}) \in L^n(\Omega; \mathbb{S}^2) \times L^2(\Omega)$ and $(\bar{\mu}, \bar{\theta}) \in L^s(\Omega) \times L^r(\Omega)$ satisfy (4.16), then the following two statements are equivalent:*

a) *There exists $\alpha > 0$ such that*

$$\partial_{(\mathbf{u}, \boldsymbol{\Sigma}, \lambda, \mathbf{g})}^2 \mathcal{L}(\bar{\mathbf{u}}, \bar{\boldsymbol{\Sigma}}, \bar{\lambda}, \bar{\mathbf{g}}, \bar{\mathbf{Y}}, \bar{\mathbf{w}}, \bar{\mu}, \bar{\theta})(\mathbf{u}', \boldsymbol{\Sigma}', \lambda', \mathbf{h})^2 \geq \alpha \|\mathbf{h}\|_U^2$$

for all $\mathbf{h} \in U$ and $(\boldsymbol{\Sigma}', \mathbf{u}', \lambda')$ solving (3.4) with $i'(\bar{\mathbf{u}})\mathbf{u}' + j'(\bar{\mathbf{g}})\mathbf{h} = 0$.

b) For all $\mathbf{h} \in U \setminus \{0\}$ and $(\boldsymbol{\Sigma}', \mathbf{u}', \lambda')$ solving (3.4) with $i'(\bar{\mathbf{u}})\mathbf{u}' + j'(\bar{\mathbf{g}})\mathbf{h} = 0$ it holds

$$\partial_{(\mathbf{u}, \boldsymbol{\Sigma}, \lambda, \mathbf{g})}^2 \mathcal{L}(\bar{\mathbf{u}}, \bar{\boldsymbol{\Sigma}}, \bar{\lambda}, \bar{\mathbf{g}}, \bar{\boldsymbol{\Upsilon}}, \bar{\mathbf{w}}, \bar{\mu}, \bar{\theta})(\mathbf{u}', \boldsymbol{\Sigma}', \lambda', \mathbf{h})^2 > 0.$$

Proof. The arguments are similar to Theorem 4.4 from [5], It suffices to show b) \implies a): We argue by contradiction and suppose that a) is not satisfied. Then for every $\alpha > 0$ there exist $\mathbf{h}_\alpha \in U$ and $(\boldsymbol{\Sigma}'_\alpha, \mathbf{u}'_\alpha, \lambda'_\alpha)$ solving (3.4) with $i'(\bar{\mathbf{u}})\mathbf{u}'_\alpha + j'(\bar{\mathbf{g}})\mathbf{h}_\alpha = 0$ such that

$$\partial_{(\mathbf{u}, \boldsymbol{\Sigma}, \lambda, \mathbf{g})}^2 \mathcal{L}(\bar{\mathbf{u}}, \bar{\boldsymbol{\Sigma}}, \bar{\lambda}, \bar{\mathbf{g}}, \bar{\boldsymbol{\Upsilon}}, \bar{\mathbf{w}}, \bar{\mu}, \bar{\theta})(\boldsymbol{\Sigma}'_\alpha, \mathbf{u}'_\alpha, \lambda'_\alpha, \mathbf{h}_\alpha)^2 < \alpha \|\mathbf{h}_\alpha\|_U^2. \quad (4.44)$$

Since S_ℓ is a cone, the solution map of (3.4) is positively homogeneous and we can assume $\|\mathbf{h}_\alpha\|_U = 1$. Thanks to Lemma 3.6 and Corollary 3.7 $(\boldsymbol{\Sigma}'_\alpha, \mathbf{u}'_\alpha, \lambda'_\alpha)$ is bounded in $L^p(\Omega; \mathbb{S}^2) \times W_D^{1,p}(\Omega; \mathbb{R}) \times L^q(\Omega)$. Consequently, there exist subsequences, denoted in the same way, such that $(\boldsymbol{\Sigma}'_\alpha, \mathbf{u}'_\alpha, \lambda'_\alpha, \mathbf{h}_\alpha) \rightharpoonup (\boldsymbol{\Sigma}', \mathbf{u}', \lambda', \mathbf{h})$ in $L^p(\Omega; \mathbb{S}^2) \times W_D^{1,p}(\Omega; \mathbb{R}^d) \times L^q(\Omega) \times U$ and $i'(\bar{\mathbf{u}})\mathbf{u}' + j'(\bar{\mathbf{g}})\mathbf{h} = 0$. In view of (4.44) and (4.10) we get

$$\begin{aligned} 0 &\geq \limsup_{\alpha \rightarrow 0} \partial_{(\mathbf{u}, \boldsymbol{\Sigma}, \lambda, \mathbf{g})}^2 \mathcal{L}(\bar{\mathbf{u}}, \bar{\boldsymbol{\Sigma}}, \bar{\lambda}, \bar{\mathbf{g}}, \bar{\boldsymbol{\Upsilon}}, \bar{\mathbf{w}}, \bar{\mu}, \bar{\theta})(\boldsymbol{\Sigma}'_\alpha, \mathbf{u}'_\alpha, \lambda'_\alpha, \mathbf{h}_\alpha)^2 \\ &\geq \liminf_{\alpha \rightarrow 0} \partial_{(\mathbf{u}, \boldsymbol{\Sigma}, \lambda, \mathbf{g})}^2 \mathcal{L}(\bar{\mathbf{u}}, \bar{\boldsymbol{\Sigma}}, \bar{\lambda}, \bar{\mathbf{g}}, \bar{\boldsymbol{\Upsilon}}, \bar{\mathbf{w}}, \bar{\mu}, \bar{\theta})(\boldsymbol{\Sigma}'_\alpha, \mathbf{u}'_\alpha, \lambda'_\alpha, \mathbf{h}_\alpha)^2 \\ &= \liminf_{\alpha \rightarrow 0} \left\{ \partial^2 i(\bar{\mathbf{u}})(\mathbf{u}'_\alpha)^2 + \partial^2 j(\bar{\mathbf{g}})\mathbf{h}_\alpha^2 + 2(\lambda'_\alpha \mathcal{D}\boldsymbol{\Sigma}'_\alpha, \mathcal{D}\bar{\boldsymbol{\Upsilon}}) + \left(\|\mathcal{D}\boldsymbol{\Sigma}'_\alpha\|_{\mathbb{S}}^2, \bar{\theta} \right) \right\}. \end{aligned} \quad (4.45)$$

The integrability conditions imply the existence of $1 \leq \beta < p$ with $1/q + 1/\beta + 1/\eta \leq 1$ and $2/\beta + 1/\vartheta \leq 1$. By Corollary 4.15 and Corollary 4.16 there are subsubsequences, denoted in the same way, with $\mathbf{u}'_\alpha \rightarrow \mathbf{u}'$ in V and $\boldsymbol{\Sigma}'_\alpha \rightarrow \boldsymbol{\Sigma}'$ in $L^\beta(\Omega; \mathbb{S}^2)$. Due to Assumption 4.11 $U \ni \mathbf{h} \mapsto \partial^2 j(\bar{\mathbf{g}})\mathbf{h}^2 \in \mathbb{R}$ is weakly lower semicontinuous and we infer from (4.45) that

$$\begin{aligned} 0 &\geq \partial^2 i(\bar{\mathbf{u}})(\mathbf{u}')^2 + \partial^2 j(\bar{\mathbf{g}})\mathbf{h}^2 + 2(\lambda' \mathcal{D}\boldsymbol{\Sigma}', \mathcal{D}\bar{\boldsymbol{\Upsilon}}) + \left(\|\mathcal{D}\boldsymbol{\Sigma}'\|_{\mathbb{S}}^2, \bar{\theta} \right) \\ &= \partial_{(\mathbf{u}, \boldsymbol{\Sigma}, \lambda, \mathbf{g})}^2 \mathcal{L}(\bar{\mathbf{u}}, \bar{\boldsymbol{\Sigma}}, \bar{\lambda}, \bar{\mathbf{g}}, \bar{\boldsymbol{\Upsilon}}, \bar{\mathbf{w}}, \bar{\mu}, \bar{\theta})(\boldsymbol{\Sigma}', \mathbf{u}', \lambda', \mathbf{h})^2. \end{aligned} \quad (4.46)$$

Since $(\boldsymbol{\Sigma}', \mathbf{u}', \lambda', \mathbf{h})$ solves (3.4) by Lemma 4.14, we deduce from b) that $\mathbf{h} = 0$ and thus $(\boldsymbol{\Sigma}', \mathbf{u}', \lambda', \mathbf{h}) = 0$. Using Assumption 4.11, (4.45), and the strong convergence $\mathbf{u}'_\alpha \rightarrow \mathbf{0}$, $\boldsymbol{\Sigma}'_\alpha \rightarrow \mathbf{0}$, we conclude

$$\begin{aligned} \nu &= \nu \limsup_{\alpha \rightarrow 0} \|\mathbf{h}_\alpha\|_U^2 \leq \limsup_{\alpha \rightarrow 0} \partial^2 j(\bar{\mathbf{g}})\mathbf{h}_\alpha^2 \\ &\leq \limsup_{\alpha \rightarrow 0} \left\{ \partial_{(\mathbf{u}, \boldsymbol{\Sigma}, \lambda, \mathbf{g})}^2 \mathcal{L}(\bar{\mathbf{u}}, \bar{\boldsymbol{\Sigma}}, \bar{\lambda}, \bar{\mathbf{g}}, \bar{\boldsymbol{\Upsilon}}, \bar{\mathbf{w}}, \bar{\mu}, \bar{\theta})(\boldsymbol{\Sigma}'_\alpha, \mathbf{u}'_\alpha, \lambda'_\alpha, \mathbf{h}_\alpha)^2 - \partial^2 i(\bar{\mathbf{u}})(\mathbf{u}'_\alpha)^2 \right. \\ &\quad \left. - 2(\lambda'_\alpha \mathcal{D}\boldsymbol{\Sigma}'_\alpha, \mathcal{D}\bar{\boldsymbol{\Upsilon}}) - \left(\|\mathcal{D}\boldsymbol{\Sigma}'_\alpha\|_{\mathbb{S}}^2, \bar{\theta} \right) \right\} \\ &\leq \limsup_{\alpha \rightarrow 0} \partial_{(\mathbf{u}, \boldsymbol{\Sigma}, \lambda, \mathbf{g})}^2 \mathcal{L}(\bar{\mathbf{u}}, \bar{\boldsymbol{\Sigma}}, \bar{\lambda}, \bar{\mathbf{g}}, \bar{\boldsymbol{\Upsilon}}, \bar{\mathbf{w}}, \bar{\mu}, \bar{\theta})(\boldsymbol{\Sigma}'_\alpha, \mathbf{u}'_\alpha, \lambda'_\alpha, \mathbf{h}_\alpha)^2 \\ &\quad - \lim_{\alpha \rightarrow 0} \left\{ \partial^2 i(\bar{\mathbf{u}})(\mathbf{u}'_\alpha)^2 + 2(\lambda'_\alpha \mathcal{D}\boldsymbol{\Sigma}'_\alpha, \mathcal{D}\bar{\boldsymbol{\Upsilon}}) + \left(\|\mathcal{D}\boldsymbol{\Sigma}'_\alpha\|_{\mathbb{S}}^2, \bar{\theta} \right) \right\} \\ &\leq 0. \end{aligned}$$

This is the desired contradiction. \square

Remark 4.18. The proof above can analogously be applied to (SSC) provided that the objective J contains a Tikhonov type term like j .

APPENDIX 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:

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

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

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 $f_n \rightarrow f$ a.e. in E , then $f_n \rightarrow f$ in $L^\kappa(E)$ for $1 \leq \kappa < \nu$.*

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

Acknowledgements. The authors would like to thank Roland Herzog and Gerd Wachsmuth (TU Chemnitz) for their helpful advices to improve the manuscript. Moreover, we are thankful to Daniel Wachsmuth (University of Würzburg), who pointed out the estimates (4.4) and (4.6) to us. Finally, we are very grateful to the anonymous referees for their helpful advices and comments. This work was supported by two DFG grants within the Priority Program SPP 1253 (*Optimization with Partial Differential Equations*) and the Collaborative Research Center SFB 708 (*3D-Surface Engineering of Tools for Sheet Metal Forming – Manufacturing, Modeling, Machining*), which is gratefully acknowledged.

REFERENCES

- [1] V. Barbu, Optimal Control of Variational Inequalities, vol. 100 of *Res. Notes Math.* Pitman, Boston (1984).
- [2] M. Bergounioux, Optimal control problems governed by abstract elliptic variational inequalities with state constraints. *SIAM J. Control Optim.* **36** (1998) 273–289.
- [3] F. Bonnans and D. Tiba, Pontryagin's principle in the control of semilinear elliptic variational inequalities. *Appl. Math. Optim.* **23** (1991) 299–312.
- [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 J. Optim.* **19** (2008) 616–643.
- [5] E. Casas and M. Mateos, Second order optimality conditions for semilinear elliptic control problems with finitely many state constraints. *SIAM J. Control Optim.* **40** (2002) 1431–1454.
- [6] E. Casas and F. Tröltzsch, First- and second order optimality conditions for a class of optimal control problems with quasilinear elliptic equations. *SIAM J. Control Optim.* **48** (2009) 688–718.
- [7] 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** (1996) 687–707.
- [8] K. Gröger, A $W^{1,p}$ -estimate for solutions to mixed boundary value problems for second order elliptic differential equations. *Mathematische Annalen* **283** (1989) 679–687.
- [9] P. Grisvard, Elliptic Problems in Nonsmooth Domains. Pitman, Boston (1985).
- [10] R. Haller-Dintelmann, C. Meyer, J. Rehberg, and A. Schiela, Hölder continuity and optimal control for nonsmooth elliptic problems. *Appl. Math. Optim.* **60** (2009) 397–428.
- [11] W. Han and B.D. Reddy, Plasticity. Springer, New York (1999).
- [12] R. Herzog and C. Meyer, Optimal control of static plasticity with linear kinematic hardening. *J. Appl. Math. Mech.* **91** (2011) 777–794.
- [13] R. Herzog, C. Meyer, and G. Wachsmuth, Integrability of displacement and stresses in linear and nonlinear elasticity with mixed boundary conditions. *J. Math. Anal. Appl.* **382** (2011) 802–813.
- [14] R. Herzog, C. Meyer, and G. Wachsmuth, C-stationarity for optimal control of static plasticity with linear kinematic hardening. *SIAM J. Control Optim.* **50** (2012) 3052–3082.
- [15] R. Herzog, C. Meyer, and G. Wachsmuth, B- and strong stationarity for optimal control of static plasticity with hardening. *SIAM J. Optim.* **23** (2013) 321–352.
- [16] M. Hintermüller and I. Kopacka, Mathematical programs with complementarity constraints in function space: C- and strong stationarity and a path-following algorithm. *SIAM J. Optim.* **20** (2009) 868–902.

- [17] M. Hintermüller and Th. Surowiec, First order optimality conditions for elliptic mathematical programs with equilibrium constraints via variational analysis. *SIAM J. Optim.* **21** (2012) 1561–1593.
- [18] K. Ito and K. Kunisch, Optimal control of elliptic variational inequalities. *Appl. Math. Optim.* **41** (2000) 343–364.
- [19] C. Kanzow and A. Schwartz, Mathematical programs with equilibrium constraints: enhanced Fritz John-conditions, new constraint qualifications, and improved exact penalty results. *SIAM J. Optim.* **20** (2010) 2730–2753.
- [20] K. Kunisch and D. Wachsmuth, Sufficient optimality conditions and semi-smooth newton methods for optimal control of stationary variational inequalities. *ESAIM: COCV* **18** (2012) 520–547.
- [21] K. Kunisch and D. Wachsmuth, Path-following for optimal control of stationary variational inequalities. *Comput. Optim. Appl.* (2012) 1–29.
- [22] F. Mignot, Contrôle dans les inéquations variationelles elliptiques. *J. Funct. Anal.* **22** (1976) 130–185.
- [23] F. Mignot and J.-P. Puel, Optimal control in some variational inequalities. *SIAM J. Control Optim.* **22** (1984) 466–476.
- [24] P. Neff and D. Knees, Regularity up to the boundary for nonlinear elliptic systems arising in time-incremental infinitesimal elasto-plasticity. *SIAM J. Math. Anal.* **40** (2008) 21–43.
- [25] J. Outrata, J. Jarušek and J. Stará, On optimality conditions in control of elliptic variational inequalities. *Set-Valued Var. Anal.* **19** (2011) 23–42.
- [26] H. Scheel and S. Scholtes, Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. *Math. Oper. Res.* **25** (2000) 1–22.
- [27] G. Wachsmuth, *Optimal control of quasistatic plasticity – An MPCC in function space*. Ph.D. thesis, Chemnitz University of Technology, Germany (2011).
- [28] G. Wachsmuth, Differentiability of implicit functions: Beyond the implicit function theorem. Technical Report SPP1253-137, Priority Program 1253, German Research Foundation (2012).