ruhr.paD
UA Ruhr Zentrum für
partielle Differentialgleichungen

Optimal Control of a Viscous Two-Field Gradient Damage Model

L. M. Susu

Preprint 2018-10

Optimal Control of a Viscous Two-Field Gradient Damage Model

Livia M. Susu[1,*]

[1] Thea-Leymann Str. 9, D-45127 Essen

Received XXXX, revised XXXX, accepted XXXX
Published online XXXX

Key words optimal control of PDEs, non-smooth optimization, viscous damage evolution

Subject classification 49J20, 49K27, 74R05

The paper deals with a class of optimal control problems governed by a viscous damage model including two damage variables which are coupled through a penalty term. The existence and directional differentiability of the associated control-to-state operator is established. Moreover, necessary optimality conditions are derived. It is shown that these are equivalent to an optimality system, provided that a strict complementarity condition is satisfied.

1 Introduction

This paper is concerned with the optimal control of a viscous gradient damage model. The latter one involves two damage variables which are connected through a penalty term in the stored energy functional. It is inspired by the one presented in [4], which is a popular model that is widely used in computational mechanics. The reasoning behind considering two damage variables instead of one is of numerical nature, see [4] for more details. While one damage variable provides a local character and carries the non-smooth time evolution, the other one accounts for nonlocal effects. Because of theoretical reasons, see [18, Section 2.2], we deal with a slightly modified version of the model in [4]. The viscous gradient damage model considered in this paper reads as follows

$$\left.\begin{aligned} (\boldsymbol{u}(t), \varphi(t)) \in \operatorname*{arg\,min}_{(\boldsymbol{u},\varphi)\in V\times H^1(\Omega)} \mathcal{E}(t, \boldsymbol{u}, \varphi, d(t)), \\ -\partial_d \mathcal{E}(t, \boldsymbol{u}(t), \varphi(t), d(t)) \in \partial\mathcal{R}_\gamma(\dot{d}(t)), \quad d(0) = d_0 \text{ a.e. in } \Omega \end{aligned}\right\} \tag{P}$$

for almost all $t \in (0, T)$. Herein, d and φ denote the local and non-local damage variable, respectively, and $\boldsymbol{u}$ stands for the displacement of the body occupying the domain $\Omega \subset \mathbb{R}^N$, $N = 2, 3$. The problem (P) describes the effect of a force ℓ on an elastic body in terms

* Corresponding author E-mail: livia.susu@uni-due.de, Phone: +49 201 183-7506,

of the displacement $\boldsymbol{u}$, the non-local damage φ and the local damage d. The stored energy $\mathcal{E} : [0,T] \times V \times H^1(\Omega) \times L^2(\Omega) \to \mathbb{R}$ is given by

$$\mathcal{E}(t, \boldsymbol{u}, \varphi, d) := \frac{1}{2} \int_\Omega g(\varphi)\mathbb{C}\varepsilon(\boldsymbol{u}) : \varepsilon(\boldsymbol{u})\, dx - \langle \ell(t), \boldsymbol{u} \rangle_V + \frac{\alpha}{2}\|\nabla\varphi\|_2^2 + \frac{\beta}{2}\|\varphi - d\|_2^2, \tag{1.1}$$

where $\varepsilon = 1/2(\nabla + \nabla^\top)$ is the linearized strain, $\mathbb{C}$ the elasticity tensor, and ℓ the applied load. The function g describes the influence of the damage on the elastic behavior of the body. Furthermore, $\alpha > 0$ denotes the gradient regularization parameter and $\beta > 0$ stands for the penalization parameter. The viscous dissipation functional $\mathcal{R}_\gamma : L^2(\Omega) \to [0, \infty]$ appearing in (P) is defined as

$$\mathcal{R}_\gamma(\eta) := \begin{cases} r \int_\Omega \eta\, dx + \frac{\gamma}{2}\|\eta\|_2^2, & \text{if } \eta \geq 0 \text{ a.e. in } \Omega, \\ \infty, & \text{otherwise,} \end{cases}$$

where $r > 0$ stands for the threshold value which triggers the damage evolution and $\gamma > 0$ is the viscosity parameter. Moreover, the initial damage is supposed to satisfy $d_0 \in L^2(\Omega)$. For a more detailed description of the model, as well as its motivation, see [18, Section 2] and [4]. Throughout the paper, we will often refer to (P) as "penalized damage model".

In many practical applications it is of interest to gain information about those loads which (locally) minimize a given cost functional, e.g. one may be interested to minimize the damage or the distance to a desired damage or/and displacement. This motivates our goal in this paper: to derive necessary optimality conditions for an optimal control problem governed by the penalized damage model, where the load is used as control. The optimization problem studied throughout this paper is therefore given by

$$\begin{aligned} \min_{\ell \in \mathfrak{L}} \quad & \mathcal{J}(\boldsymbol{u}, \varphi, d, \ell) \\ \text{s.t.} \quad & (\boldsymbol{u}, \varphi, d) \text{ solves (P) with right-hand side } \ell. \end{aligned} \tag{P_{min}}$$

The assumptions on the objective $\mathcal{J}$ and the control set $\mathfrak{L}$ are to be introduced and motivated in Section 5 below. As we will see, the structure of (P) accounts for a non-smooth optimal control problem. To be more precise, we deal with an optimal control problem governed by a non-smooth parabolic PDE. Thus, the control-to-state operator associated to the penalized damage model is not necessarily Gâteaux-differentiable, and the standard adjoint calculus for the derivation of qualified optimality conditions is not applicable, at least not without further ado.

Let us put our work into perspective. While the optimal control of smooth parabolic equations was investigated by many authors, see e.g. [25] and the references therein, less papers are dealing with non-smooth equations. Most of the contributions in this field focus on variational inequalities of the first kind, such as the parabolic obstacle problem, see [1,2,5,6,11,16,17]. In all these contributions, the lack of differentiability of the control-to-state operator is overcome by employing regularization and relaxation techniques. The thereby derived optimality systems are in the best case of intermediate strength, such as C stationarity, see e.g. [13, 15, 16]. For the optimal control of the parabolic obstacle problem, a strong stationarity system can be found in [21], but no rigorous proof is given there. Recently, an optimality system of strong stationary type was derived in [19] for an optimal control problem governed by a non-smooth parabolic PDE. This was possible due to the presence of so-called ‘ample controls’, which are necessary for deriving strong stationarity in most existing contributions, see e.g. [3, 14, 22] (elliptic VIs). To the best of our knowledge, [26] is the only paper where a strong stationary optimality system is derived in the absence thereof, however by requiring that the (unknown) optimizer satisfies certain assumptions (constraint qualifications). Hence, it is not surprising, that without regularizing, additional assumptions are needed in order to derive an optimality system for (P_{min}). Therefor we make use of the special structure of the constraint in (P_{min}). To be more precise, we employ the fact that (P) can be reduced to an ordinary (non-smooth) differential equation in Banach space, cf. [18].

The paper is organized as follows. Section 2 collects the notations and standing assumptions. In Section 3 we introduce the control-to-state operator by carefully choosing its domain of definition, while in Section 4 we investigate its (directional) differentiability. This is the preparatory step for deriving first order necessary optimality conditions, which are established in Section 5. Therein one arrives at an assumption under which the control-to-state operator is Gâteaux-differentiable. By means thereof, we can derive the optimality system for (P_{min}). Section 6 is concerned with a short discussion regarding the existence of solutions for (P_{min}).

2 Notation and standing assumptions

In what follows, $T > 0$ is fixed and $N \in \{2, 3\}$ is the spatial dimension. Throughout the paper, c and C denote generic positive constants. By bold-face letters we denote vector-valued variables and vector-valued spaces. The Frobenius norm on $\mathbb{R}^{N\times N}$, as well as the euclidean norm on $\mathbb{R}^N$ are denoted by $|\cdot|$, whereas the inducing scalar product in $\mathbb{R}^{N\times N}$ is represented by $(\cdot : \cdot)$. Let X and Y be Banach spaces. The open ball in X around $x \in X$ with radius $R > 0$ is denoted by $B_X(x, R)$. The space of linear and bounded operators from X to Y is called $\mathcal{L}(X, Y)$ and if $X = Y$, it is called $\mathcal{L}(X)$. The dual of the space X will be denoted by X^* and for the dual pairing between X and X^* we write $\langle ., .\rangle_X$. If X is compactly embedded in Y, we write $X \hookrightarrow\hookrightarrow Y$ and $X \overset{d}{\hookrightarrow} Y$ means that X is dense in Y. Let $s \in [1, \infty]$. By $\|\cdot\|_s$ we abbreviate the notation for the $L^s(\Omega)$-norm, and by $(\cdot, \cdot)_2$, the $L^2(\Omega)$-scalar product. For frequently used function spaces we introduce the following abbreviations:

$$\boldsymbol{W}^{1,s}_D(\Omega) := \{v \in \boldsymbol{W}^{1,s}(\Omega) : v|_{\Gamma_D} = 0\}, \quad V := \boldsymbol{W}^{1,2}_D(\Omega).$$

The dual space of $\boldsymbol{W}^{1,s'}_D(\Omega)$ is denoted by $\boldsymbol{W}^{-1,s}(\Omega)$, where s' is the conjugate exponent of s. For the following often employed subspaces of the Bochner-Sobolev space $W^{1,s}(0, T; X)$ we use the notations:

$$\begin{aligned} W^{1,s}_0(0, T; X) &:= \{z \in W^{1,s}(0, T; X) : z(0) = 0\}, \\ W^{1,s}_T(0, T; X) &:= \{z \in W^{1,s}(0, T; X) : z(T) = 0\}, \\ H^1_0(0, T; X) &:= W^{1,2}_0(0, T; X). \end{aligned}$$

By $\operatorname{div} : L^s(\Omega; \mathbb{R}^{N\times N}_{sym}) \to \boldsymbol{W}^{1,s'}(\Omega)^*$ we denote the distributional vector-valued divergence and $\Delta : W^{1,s}(\Omega) \to W^{1,s'}(\Omega)^*$ is the distributional Laplace operator. The Nemytskii operator (considered with different domains and ranges) associated to $\max : \mathbb{R} \to \mathbb{R},\ x \mapsto \max\{x, 0\}$ is denoted by $\max$. By χ_M we denote the characteristic function associated to the set M. Time derivatives are frequently denoted by a dot.

Let us now state our standing assumptions.

Assumption 2.1 The domain $\Omega \subset \mathbb{R}^N$, $N \in \{2, 3\}$, is a bounded Lipschitz domain, see [8, Chap. 1.2]. Its boundary is denoted by Γ and consists of two disjoint measurable parts Γ_N and Γ_D such that $\Gamma = \Gamma_N \cup \Gamma_D$. While Γ_N is an open subset, Γ_D is a closed subset of Γ. Moreover, Γ_D is assumed to have positive measure.

In addition, the set $\Omega \cup \Gamma_N$ is regular in the sense of Gröger, cf. [9, Definition 2]. That is, for every point $\boldsymbol{x} \in \Gamma$, there exists an open neighborhood $\mathcal{U}_{\boldsymbol{x}} \subset \mathbb{R}^N$ of $\boldsymbol{x}$ and a bi-Lipschitz map (a Lipschitz continuous and bijective map with Lipschitz continuous inverse) $\Psi_{\boldsymbol{x}} : \mathcal{U}_{\boldsymbol{x}} \to \mathbb{R}^N$ such that $\Psi_{\boldsymbol{x}}(\boldsymbol{x}) = \boldsymbol{0} \in \mathbb{R}^N$ and $\Psi_{\boldsymbol{x}}\big(\mathcal{U}_{\boldsymbol{x}} \cap (\Omega \cup \Gamma_N)\big)$ equals one of the following sets:

$$
\begin{aligned}
E_1 &:= \big\{\boldsymbol{y} \in \mathbb{R}^N : |y| < 1,\ y_N < 0\big\}\,,\\
E_2 &:= \big\{\boldsymbol{y} \in \mathbb{R}^N : |y| < 1,\ y_N \le 0\big\}\,,\\
E_3 &:= \{\boldsymbol{y} \in E_2 : y_N < 0 \text{ or } y_1 > 0\}\,.
\end{aligned}
$$

A detailed characterization of Gröger regular sets in two and three spatial dimensions is given in [10, Section 5].

Assumption 2.2 The function $g : \mathbb{R} \to [\epsilon, 1]$, where $\epsilon \in (0, 1]$, satisfies

$$
g \in C^2(\mathbb{R}) \text{ and } g, g' \in C^{0,1}(\mathbb{R}). \tag{2.1}
$$

With a little abuse of notation, the Nemytskii operators associated to g, g' and g'', considered with different domains and ranges, will be denoted by the same symbol.

The coefficient function g measures how the elastic properties of the body are preserved depending on the value of the damage. Since with increasing damage the material becomes weaker, it would make sense to impose that g is monotonically decreasing. This property of g is needed e.g. if one aims to show that the non-local damage variable admits just non-negative values, as the local damage variable does. For this, it suffices in fact that g decreases only on $\mathbb{R}^-$. However, since we do not need this result in our analysis, we do not require that g has this property here.

We emphasize that, due to the condition $g \ge \epsilon$, our model constitutes a partial damage model. By contrast, $\lim_{\varphi\to\infty} g(\varphi) = 0$ is assumed in [4, (2)], which assures that complete material rigidity loss occurs in the case of complete damage. However, in order to guarantee coercivity of the bilinear form associated with the balance of momentum in (3.1a) below, we have to impose a positive lower bound on g.

Assumption 2.3 The fourth-order tensor $\mathbb{C} \in L^\infty(\Omega; \mathcal{L}(\mathbb{R}^{N\times N}_{\text{sym}}))$ is symmetric and uniformly coercive, i.e., there is a constant $\gamma_{\mathbb{C}} > 0$ such that

$$
\mathbb{C}(x)\boldsymbol{\sigma} : \boldsymbol{\sigma} \ge \gamma_{\mathbb{C}} |\boldsymbol{\sigma}|^2 \quad \forall\, \boldsymbol{\sigma} \in \mathbb{R}^{N\times N}_{\text{sym}} \text{ and f.a.a. } x \in \Omega. \tag{2.2}
$$

Our last assumption concerns the balance of momentum associated with the energy functional in (1.1). For its precise statement we need the following

Definition 2.4 For given $\varphi \in L^1(\Omega)$ and $\overline{p} \in (1,\infty)$, we define the linear form $A_\varphi : \boldsymbol{W}_D^{1,\overline{p}}(\Omega) \to \boldsymbol{W}^{-1,\overline{p}}(\Omega)$ as

$$\langle A_\varphi \boldsymbol{u}, \boldsymbol{v}\rangle := \int_\Omega g(\varphi)\mathbb{C}\varepsilon(\boldsymbol{u}) : \varepsilon(\boldsymbol{v})\, dx.$$

Assumption 2.5 For the rest of the paper we require the following:

1. There exists $p > N$ such that, for all $\overline{p} \in [2,p]$ and all $\varphi \in L^1(\Omega)$, the operator $A_\varphi : \boldsymbol{W}_D^{1,\overline{p}}(\Omega) \to \boldsymbol{W}^{-1,\overline{p}}(\Omega)$ is continuously invertible. Moreover, there exists a constant $c > 0$, independent of φ and $\overline{p}$, such that

   $$\|A_\varphi^{-1}\|_{\mathcal{L}(\boldsymbol{W}^{-1,\overline{p}}(\Omega),\boldsymbol{W}_D^{1,\overline{p}}(\Omega))} \le c$$

   holds for all $\overline{p} \in [2,p]$ and all $\varphi \in L^1(\Omega)$.

2. The penalization parameter β is sufficiently large, depending only on the data, see [18, (3.35)] and Remark 3.2 below.

Let us point out that p stands for the integrability exponent in Assumption 2.5.1 throughout the entire paper.

Remark 2.6 The critical assumption is Assumption 2.5.1. If N = 2, then this condition is automatically fulfilled, see [18, Lemma 3.2] and [12]. The situation changes however if one turns to $N = 3$. In this case, this assumption can be guaranteed by imposing additional and rather restrictive conditions on the data, in particular on the ellipticity and boundedness constants associated with $\mathbb{C}$ and g, see [18, Remark 3.20] and [12] for more details. However, as explained in [18, Remark 3.21], one could alternatively modify the energy functional in (1.1) by replacing $\|\nabla\varphi\|_2^2$ with its $H^{3/2}$-seminorm. This would allow to drop Assumption 2.5.1 in the three dimensional case, too. However, we chose not to work with the $H^{3/2}$-seminorm, as the associated bilinear form is difficult to realize in numerical computations.

3 Control-to-state operator

In this section we investigate the constraint in (P_{min}), i.e., the problem (P). The unique solvability thereof was already established in [18, Section 3.2], in the context where $\ell \in C^{0,1}([0,T]; \boldsymbol{W}^{-1,p}(\Omega))$ was fixed. Therein it was reasonable to focus on the time dependence of the solution operators, not on the dependence on the load, which was just one of the given data. In particular, one was interested in the differentiability with respect to time of the solution operators, which in combination with the smoothness in time of the load played an essential role for proving the viability of the penalization approach, see [20]. However, for the sole purpose of showing unique solvability of (P), it suffices that the load ℓ belongs to $L^\infty(0,T; \boldsymbol{W}^{-1,p}(\Omega))$. The situation changes when we want to control the load in the context of analyzing the problem (P_{min}). As we will see, this calls for working with (variable) loads in a bounded subset of $L^\infty(0,T; \boldsymbol{W}^{-1,p}(\Omega))$, see Definition 3.3 below.

In the following, we abbreviate the open ball

$$\mathcal{B}_M := B_{\boldsymbol{W}^{-1,p}(\Omega)}(0, M),$$

for a given $M > 0$. By employing the exact same arguments as in [18, Sections 3 and 5], one proves

Lemma 3.1 *Let $M > 0$ be given. Then, for any pair $(\ell, d) \in \boldsymbol{W}^{-1,p}(\Omega) \times L^2(\Omega)$, the following elliptic system*

$$-\operatorname{div} g(\bar{\varphi})\mathbb{C}\varepsilon(\bar{\boldsymbol{u}}) = \ell \qquad \text{in } \boldsymbol{W}^{-1,p}(\Omega) \tag{3.1a}$$

$$-\alpha\Delta\bar{\varphi} + \beta\,\bar{\varphi} + \frac{1}{2}\,g'(\bar{\varphi})\mathbb{C}\,\varepsilon(\bar{\boldsymbol{u}}) : \varepsilon(\bar{\boldsymbol{u}}) = \beta d \quad \text{in } H^1(\Omega)^*. \tag{3.1b}$$

admits a unique solution $(\bar{\boldsymbol{u}}, \bar{\varphi}) \in \boldsymbol{W}_D^{1,p}(\Omega) \times H^1(\Omega)$. The solution operator associated to (3.1a) *is denoted by $\mathcal{U} : \boldsymbol{W}^{-1,p}(\Omega) \times H^1(\Omega) \ni (\ell, \varphi) \mapsto \mathcal{U}(\ell, \varphi) \in \boldsymbol{W}_D^{1,p}(\Omega)$, whereas the solution operator of* (3.1b) *is called $\Phi : \mathcal{B}_M \times L^2(\Omega) \ni (\ell, d) \mapsto \Phi(\ell, d) \in H^1(\Omega)$. These are well-defined and have the following properties:*

- *The operator $\mathcal{U}$ satisfies*

$$\|\mathcal{U}(\ell, \varphi)\|_{\boldsymbol{W}_D^{1,p}(\Omega)} \leq c\|\ell\|_{\boldsymbol{W}^{-1,p}(\Omega)} \quad \forall\, (\ell, \varphi) \in \boldsymbol{W}^{-1,p}(\Omega) \times H^1(\Omega), \tag{3.2}$$

  where $c > 0$ is independent of ℓ and φ.

- *$\mathcal{U}$ is continuous from $\boldsymbol{W}^{-1,p}(\Omega) \times L^1(\Omega)$ to $\boldsymbol{W}_D^{1,s}(\Omega)$ for every $s \in [2, p)$ and Lipschitz continuous from $\mathcal{B}_M \times H^1(\Omega)$ to V.*

- *Φ is Lipschitz continuous from $\mathcal{B}_M \times L^2(\Omega)$ to $H^1(\Omega)$.*

Moreover, $\mathcal{U}$ and Φ are continuously Fréchet-differentiable, in the sense that $\mathcal{U} \in C^1(\boldsymbol{W}^{-1,p}(\Omega) \times H^1(\Omega); V)$ and $\Phi \in C^1(\mathcal{B}_M \times L^2(\Omega); H^1(\Omega))$. Their derivatives satisfy the estimates

$$\|\mathcal{U}'(\ell, \varphi)\|_{\mathcal{L}(\boldsymbol{W}^{-1,p}(\Omega)\times H^1(\Omega);V)} \leq L(M), \quad \|\Phi'(\ell, d)\|_{\mathcal{L}(\boldsymbol{W}^{-1,p}(\Omega)\times L^2(\Omega);H^1(\Omega))} \leq L(M) \tag{3.3}$$

for all $(\ell, \varphi, d) \in \mathcal{B}_M \times H^1(\Omega) \times L^2(\Omega)$, where $L(M)$ is a positive constant dependent on M.

Remark 3.2 We point out that the choice of the domain for the operator Φ is due to the fact that ℓ varies in $\boldsymbol{W}^{-1,p}(\Omega)$. A closer inspection of the proof of [18, Lemma 3.16] shows that the unique solvability of (3.1) is guaranteed when β exceeds a threshold involving $\|\ell\|_{\boldsymbol{W}^{-1,p}(\Omega)}$, cf. [18, (3.35)], see also [18, Section 3]. Since ℓ is now a variable, we need to require that $\|\ell\|_{\boldsymbol{W}^{-1,p}(\Omega)} < M$, with some $M > 0$, in order to guarantee that β depends only on the given data (including M) and is thus independent of the (variable) load ℓ.

Let now $\ell \in L^\infty(0, T; \boldsymbol{W}^{-1,p}(\Omega))$ be fixed. Provided that the threshold for β exceeds a certain value, which depends on $\|\ell\|_{L^\infty(0,T;\boldsymbol{W}^{-1,p}(\Omega))}$ this time, we have due to Lemma 3.1 and [18, Theorem 3.17, Lemma 3.22] that (P) reduces to the operator differential equation

$$\dot{d}(t) = \frac{1}{\gamma} \max\big(-\beta(d(t) - \Phi(\ell(t), d(t))) - r\big) \quad \text{f.a.a. } t \in (0, T), \quad d(0) = d_0. \tag{3.4}$$

A contraction argument then yields that (3.4) admits a unique solution d which belongs to $W^{1,\infty}(0,T;L^2(\Omega))$. The problem (P) is then uniquely solvable by $\boldsymbol{u} = \mathcal{U}\big(\ell(\cdot), \Phi(\ell(\cdot), d(\cdot)\big)$, $\varphi = \Phi\big(\ell(\cdot), d(\cdot)\big)$ and d. Moreover, from the properties of $\mathcal{U}$ and Φ in Lemma 3.1 we obtain that $\boldsymbol{u} \in L^\infty(0,T;\boldsymbol{W}_D^{1,s}(\Omega))$, with $s \in [2,p)$, and $\varphi \in L^\infty(0,T;H^1(\Omega))$, respectively. For the existence of the control-to-state operator we need however that this result holds for variable ℓ as well. To ensure that β does not depend on the variable ℓ, we consider only controls belonging to a bounded set, namely

Definition 3.3 (Admissible loads) For a given $M > 0$ we define

$$\mathfrak{B}_M := B_{L^\infty(0,T;\boldsymbol{W}^{-1,p}(\Omega))}(0,M).$$

Note that this implies $\mathfrak{B}_M \subset \{\ell \in L^\infty(0,T;\boldsymbol{W}^{-1,p}(\Omega)) : \ell(t) \in \mathcal{B}_M \text{ f.a.a. } t \in (0,T)\}$.

We are now in the position to introduce

Definition 3.4 (Control-to-state operator associated to (P)) Let $M > 0$ be given. Then we define $\mathcal{S} : \mathfrak{B}_M \to L^\infty(0,T;\boldsymbol{W}_D^{1,s}(\Omega)) \times L^\infty(0,T;H^1(\Omega)) \times W^{1,\infty}(0,T;L^2(\Omega))$, with $s \in [2,p)$, as

$$\mathcal{S}(\ell) := (\boldsymbol{u}, \varphi, d),$$

where $(\boldsymbol{u}, \varphi, d)$ is the unique solution of (P). This satisfies the system of differential equations

$$-\operatorname{div} g(\varphi(t))\mathbb{C}\varepsilon(\boldsymbol{u}(t)) = \ell(t) \qquad \text{in } \boldsymbol{W}^{-1,p}(\Omega), \tag{3.5a}$$

$$-\alpha\Delta\varphi(t) + \beta\,\varphi(t) + \frac{1}{2}\,g'(\varphi(t))\mathbb{C}\,\varepsilon(\boldsymbol{u}(t)) : \varepsilon(\boldsymbol{u}(t)) = \beta d(t) \qquad \text{in } H^1(\Omega)^*, \tag{3.5b}$$

$$\dot{d}(t) - \frac{1}{\gamma}\max\big(-\beta(d(t) - \varphi(t)) - r\big) = 0, \qquad d(0) = d_0 \tag{3.5c}$$

f.a.a. $t \in (0,T)$. For $i \in \{1,2,3\}$ we denote by $\mathcal{S}_i$ the operator which associates to any $\ell \in \mathfrak{B}_M$ the i-th component of $\mathcal{S}(\ell)$.

Lemma 3.5 *Let $M > 0$ and $s \in [2,p)$ be given. Then the control-to-state operator $\mathcal{S} : \mathfrak{B}_M \to L^\infty(0,T;\boldsymbol{W}_D^{1,s}(\Omega)) \times L^\infty(0,T;H^1(\Omega)) \times W^{1,\infty}(0,T;L^2(\Omega))$ is Lipschitz continuous.*

P r o o f. It is easy to see that the second term in the operator differential equation in (3.5c) is Lipschitz continuous w.r.t. $(\ell(t), d(t)) \in \mathcal{B}_M \times L^2(\Omega)$ f.a.a. $t \in (0,T)$, where $M > 0$ is given. This follows from the Lipschitz continuity of $\max : L^2(\Omega) \to L^2(\Omega)$ and $\Phi : \mathcal{B}_M \times L^2(\Omega) \to H^1(\Omega)$, see Lemmata A.1.$(i)$ and 3.1, respectively. In light of Gronwall's inequality, it can then be shown that $\mathcal{S}_3 : \mathfrak{B}_M \to W^{1,\infty}(0,T;L^2(\Omega))$ is Lipschitz continuous, which as a result of the Lipschitz continuity of $\mathcal{U}$ and Φ, cf. Lemma 3.1, gives in turn the Lipschitz continuity of $\mathcal{S} : \mathfrak{B}_M \to L^\infty(0,T;V) \times L^\infty(0,T;H^1(\Omega)) \times W^{1,\infty}(0,T;L^2(\Omega))$. □

4 Directional differentiability

Since we aim to derive necessary optimality conditions in primal form for (P_{min}) in the next section, we address the differentiability of the control-to-state operator $\mathcal{S}$ in what follows. Notice that this is not expected to be Gâteaux-differentiable, since in (3.5c) the evolution of the local damage d is described via the $\max$-function, which is not Gâteaux-differentiable at 0.

For the sake of convenience we work with a fixed pair $(\ell, \delta\ell) \in \mathfrak{B}_M \times L^\infty(0,T; \boldsymbol{W}^{-1,p}(\Omega))$ in the rest of the section, where $M > 0$ is given. In order to have a better overview of the upcoming results, we abbreviate

$$
\begin{aligned}
\ell_\tau &:= \ell + \tau\delta\ell, \qquad (\boldsymbol{u}, \varphi, d) := \mathcal{S}(\ell), \quad (\boldsymbol{u_\tau}, \varphi_\tau, d_\tau) := \mathcal{S}(\ell_\tau),\\
\boldsymbol{\delta u} &:= \mathcal{U}'\big(\ell(\cdot), \varphi(\cdot)\big)\big(\delta\ell(\cdot), \delta\varphi(\cdot)\big), \quad \delta\varphi := \Phi'\big(\ell(\cdot), d(\cdot)\big)\big(\delta\ell(\cdot), \delta d(\cdot)\big),
\end{aligned}
\tag{4.1}
$$

where $\delta d \in L^\infty(0,T; L^2(\Omega))$ is arbitrary, but fixed and $\tau > 0$ is small enough such that $\ell_\tau \in \mathfrak{B}_M$. Notice that such a τ exists, since $\mathfrak{B}_M$ is open, and all the operators in (4.1) are well defined. Moreover, by the continuity of $\mathcal{U}'$ and Φ', cf. Lemma 3.1, we deduce the measurability of $\mathcal{U}'\big(\ell(\cdot), \varphi(\cdot)\big)$ and $\Phi'\big(\ell(\cdot), d(\cdot)\big)$, which in view of (3.3) yields that $\mathcal{U}'\big(\ell(\cdot), \varphi(\cdot)\big)$ belongs to $L^\infty\big(0,T; \mathcal{L}(\boldsymbol{W}^{-1,p}(\Omega) \times H^1(\Omega); V)\big)$ and $\Phi'\big(\ell(\cdot), d(\cdot)\big) \in L^\infty\big(0,T; \mathcal{L}(\boldsymbol{W}^{-1,p}(\Omega) \times L^2(\Omega); H^1(\Omega)\big)$. Since $(\delta\ell, \delta d) \in L^\infty(0,T; \boldsymbol{W}^{-1,p}(\Omega) \times L^2(\Omega))$, we have

$$
\delta\varphi \in L^\infty(0,T; H^1(\Omega)), \quad \boldsymbol{\delta u} \in L^\infty(0,T; V). \tag{4.2}
$$

Lemma 4.1 *There exist constants $C, c > 0$, so that for the quantities defined in* (4.1) *we have the estimates*

$$
\Big\|\frac{\boldsymbol{u_\tau}(t) - \boldsymbol{u}(t)}{\tau} - \boldsymbol{\delta u}(t)\Big\|_V \le C\Big\|\frac{d_\tau(t) - d(t)}{\tau} - \delta d(t)\Big\|_2 + cR_\Phi(t,\tau) + R_{\mathcal{U}}(t,\tau), \tag{4.3a}
$$

$$
\Big\|\frac{\varphi_\tau(t) - \varphi(t)}{\tau} - \delta\varphi(t)\Big\|_{H^1(\Omega)} \le C\Big\|\frac{d_\tau(t) - d(t)}{\tau} - \delta d(t)\Big\|_2 + R_\Phi(t,\tau) \tag{4.3b}
$$

f.a.a. $t \in (0,T)$, where $\tau > 0$ is small enough, independent of t, and $R_{\mathcal{U}}, R_\Phi : (0,T) \times (0,1) \to [0,\infty)$ are mappings which satisfy

$$
R_{\mathcal{U}}(\cdot,\tau), R_\Phi(\cdot,\tau) \to 0 \textit{ in } L^\varrho(0,T) \quad \textit{as } \tau \searrow 0, \tag{4.4}
$$

for any $\varrho \in [1,\infty)$.

Proof. Let us fix $\tau > 0$ such that $\ell + \tau\delta\ell \in \mathfrak{B}_M$. Then, by the Lipschitz continuity of $\mathcal{U}$, cf. Lemma 3.1, see also Definition 3.3, we arrive at

$$\begin{aligned}
&\Big\|\frac{\boldsymbol{u}_{\boldsymbol{\tau}}(t) - \boldsymbol{u}(t)}{\tau} - \boldsymbol{\delta u}(t)\Big\|_V \\
&= \Big\|\frac{\mathcal{U}\big(\ell_\tau(t), \varphi_\tau(t)\big) - \mathcal{U}\big(\ell(t), \varphi(t)\big)}{\tau} - \mathcal{U}'\big(\ell(t), \varphi(t)\big)\big(\delta\ell(t), \delta\varphi(t)\big)\Big\|_V \\
&\leq \Big\|\frac{\mathcal{U}\big(\ell_\tau(t), \varphi_\tau(t)\big) - \mathcal{U}\big(\ell(t) + \tau\delta\ell(t), \varphi(t) + \tau\delta\varphi(t)\big)}{\tau}\Big\|_V \\
&\quad + \Big\|\frac{\mathcal{U}\big(\ell(t) + \tau\delta\ell(t), \varphi(t) + \tau\delta\varphi(t)\big) - \mathcal{U}\big(\ell(t), \varphi(t)\big)}{\tau} - \mathcal{U}'\big(\ell(t), \varphi(t)\big)\big(\delta\ell(t), \delta\varphi(t)\big)\Big\|_V \\
&\leq L_{\mathcal{U}}\Big\|\frac{\big(\ell_\tau(t), \varphi_\tau(t)\big) - \big(\ell(t), \varphi(t)\big)}{\tau} - (\delta\ell(t), \delta\varphi(t))\Big\|_{\boldsymbol{W}^{-1,p}(\Omega)\times H^1(\Omega)} + R_{\mathcal{U}}(t, \tau) \\
&= L_{\mathcal{U}}\Big\|\frac{\varphi_\tau(t) - \varphi(t)}{\tau} - \delta\varphi(t)\Big\|_{H^1(\Omega)} + R_{\mathcal{U}}(t, \tau) \quad \text{f.a.a. } t \in (0, T),
\end{aligned} \tag{4.5}$$

where $L_{\mathcal{U}} > 0$ and $R_{\mathcal{U}}(t, \tau)$ is the abbreviation for the second addend in the first inequality above. Since $\mathcal{U}$ is continuously differentiable at $(\ell(t), \varphi(t))$, it holds

$$R_{\mathcal{U}}(t, \tau) \to 0 \ \text{ as } \tau \searrow 0 \quad \text{f.a.a. } t \in (0, T).$$

Employing the Lipschitz continuity of $\mathcal{U}$ together with (3.3) further gives

$$R_{\mathcal{U}}(t, \tau) \leq C\|(\delta\ell(t), \delta\varphi(t))\|_{\boldsymbol{W}^{-1,p}(\Omega)\times H^1(\Omega)} \quad \text{f.a.a. } t \in (0, T).$$

From (4.2) we know that $(\delta\ell(\cdot), \delta\varphi(\cdot)) \in L^\infty(0, T; \boldsymbol{W}^{-1,p}(\Omega) \times H^1(\Omega))$, and thus, by Lebesgue's dominated convergence theorem, we now have

$$R_{\mathcal{U}}(\cdot, \tau) \to 0 \text{ in } L^\varrho(0, T) \quad \text{as } \tau \searrow 0, \quad \forall\, \varrho \in [1, \infty).$$

The estimate (4.3b) follows by the exact same arguments. By inserting (4.3b) in (4.5), one immediately obtains (4.3a), which completes the proof. □

The next lemma provides the candidate for the derivative of $\mathcal{S}_3$ at ℓ in direction $\delta\ell$.

Lemma 4.2 *The equation*

$$\left.\begin{aligned}
\dot\eta(t) = \frac{1}{\gamma}\max{}'\big(-\beta(d(t) - \varphi(t)) - r\,; -\beta\big(\eta(t) - \Phi'\big(\ell(t), d(t)\big)\big(\delta\ell(t), \eta(t)\big)\big)\big)& \\
\text{f.a.a. } t \in (0, T),& \\
\eta(0) = 0&
\end{aligned}\right\} \tag{4.6}$$

admits a unique solution $\eta \in W_0^{1,\infty}(0, T; L^2(\Omega))$, where $d = \mathcal{S}_3(\ell)$ and $\varphi = \mathcal{S}_2(\ell)$.

Proof. We begin by defining $f : (0,T) \times L^2(\Omega) \to L^2(\Omega)$ as

$$f(t,\eta) = \frac{1}{\gamma} \max{}'\big(-\beta(d(t)-\varphi(t)) - r\,; -\beta\big(\eta - \Phi'\big(\ell(t),d(t)\big)\big(\delta\ell(t),\eta\big)\big)\big). \quad (4.7)$$

Definition 3.4 tells us that $d-\varphi \in L^\infty(0,T;L^2(\Omega))$ and for $\eta \in L^\infty(0,T;L^2(\Omega))$ we deduce from (4.1) and (4.2) that

$$\eta - \Phi'\big(\ell(\cdot),d(\cdot)\big)\big(\delta\ell(\cdot),\eta(\cdot)\big) \in L^\infty(0,T;L^2(\Omega)).$$

Therefore, $f(\cdot,\eta(\cdot)) \in L^\infty(0,T;L^2(\Omega))$ for $\eta \in L^\infty(0,T;L^2(\Omega))$, by Lemma A.1.$(ii)$. Let now $\eta_i \in L^2(\Omega)$ be arbitrary, but fixed, and let us abbreviate $\delta\varphi_i := \Phi'(\ell(\cdot),d(\cdot))(\delta\ell(\cdot),\eta_i)$ for $i = 1,2$. From Lemma A.1.(i) we know that the operator $\max{}'(y;\cdot) : L^2(\Omega) \to L^2(\Omega)$ is Lipschitz continuous for any $y \in L^2(\Omega)$, and as a consequence of (3.3), we have the estimate

$$\begin{aligned}\|f(t,\eta_1) - f(t,\eta_2)\|_2 &\le \frac{1}{\gamma}\| - \beta\big(\eta_1 - \delta\varphi_1(t)\big) + \beta\big(\eta_2 - \delta\varphi_2(t)\big)\|_2 \\ &\le \frac{\beta}{\gamma}\big(\|\eta_1-\eta_2\|_2 + \|\Phi'(\ell(t),d(t))\|_{\mathcal{L}(\boldsymbol{W}^{-1,p}(\Omega)\times L^2(\Omega);H^1(\Omega))}\|\eta_1-\eta_2\|_2\big) \\ &\le \frac{\beta}{\gamma}\,(L+1)\,\|\eta_1-\eta_2\|_2 \quad \text{f.a.a. } t \in (0,T).\end{aligned}$$

The properties of f shown above together with a fixed point argument allow us to deduce the existence of a unique (local) solution for (4.6) on some interval $[0,t]$, $t \le T$. The global solution of (4.6) is obtained by concatenating the local solutions. □

With the candidate for the derivative of $\mathcal{S}_3$ at hand, we can now prove its directional differentiability at ℓ in direction $\delta\ell$, which is covered by the following

Lemma 4.3 (Directional differentiability of $\mathcal{S}_3$) *Let $\varrho \in [1,\infty)$. Then, the following convergence holds true*

$$\frac{\mathcal{S}_3(\ell+\tau\delta\ell) - \mathcal{S}_3(\ell)}{\tau} \to \eta \ \text{ in } W^{1,\varrho}(0,T;L^2(\Omega)) \quad \text{as } \tau \searrow 0,$$

where $\eta \in W_0^{1,\infty}(0,T;L^2(\Omega))$ is the unique solution of (4.6).

Proof. For simplicity, we use the notations $d_\tau = \mathcal{S}_3(\ell+\tau\delta\ell)$ and $d = \mathcal{S}_3(\ell)$ again, where $\tau > 0$ is small enough such that $\ell + \tau\delta\ell \in \mathfrak{B}_M$. In view of (3.5c), (4.6) and the properties of $\max$, cf. Lemma A.1, one derives in the exact same way as (4.5) the following estimate f.a.a. $t \in (0,T)$:

$$\begin{aligned}&\gamma\Big\|\frac{\dot d_\tau(t) - \dot d(t)}{\tau} - \dot\eta(t)\Big\|_2 \\ &\le \beta\Big\|\frac{-\big(d_\tau(t)-\varphi_\tau(t)\big) + \big(d(t)-\varphi(t)\big)}{\tau} + \eta(t) - \Phi'\big(\ell(t),d(t)\big)\big(\delta\ell(t),\eta(t)\big)\Big\|_2 + R_{\max}(t,\tau),\end{aligned} \quad (4.8)$$

where $R_{\max} : (0,T) \times (0,1) \to [0,\infty)$ satisfies $R_{\max}(\cdot,\tau) \to 0$ in $L^\varrho(0,T)$ for $\tau \searrow 0$. On account of (4.3b), the estimate (4.8) can be continued as follows

$$\begin{aligned} &\gamma \Big\| \frac{\dot d_\tau(t) - \dot d(t)}{\tau} - \dot\eta(t) \Big\|_2 \\ &\leq \beta \Big\| \frac{d_\tau(t) - d(t)}{\tau} - \eta(t) \Big\|_2 + \beta \Big\| \frac{\varphi_\tau(t) - \varphi(t)}{\tau} - \Phi'\big(\ell(t), d(t)\big)\big(\delta\ell(t), \eta(t)\big) \Big\|_2 + R_{\max}(t,\tau) \\ &\leq C \Big\| \frac{d_\tau(t) - d(t)}{\tau} - \eta(t) \Big\|_2 + cR_\Phi(t,\tau) + R_{\max}(t,\tau) \quad \text{f.a.a. } t \in (0,T). \end{aligned} \tag{4.9}$$

Note that $R(\cdot,\tau) := cR_\Phi(\cdot,\tau) + R_{\max}(\cdot,\tau) \geq 0$ f.a.a. $t \in (0,T)$ and $R(\cdot,\tau) \to 0$ in $L^\varrho(0,T)$ as $\tau \searrow 0$. Integrating (4.9) further yields for all $t \in [0,T]$

$$\gamma \Big\| \frac{d_\tau(t) - d(t)}{\tau} - \eta(t) \Big\|_2 \leq C \int_0^t \Big\| \frac{d_\tau(s) - d(s)}{\tau} - \eta(s) \Big\|_2 \, ds + \int_0^t R(s,\tau)\, ds.$$

The positivity of $R(\cdot,\tau)$ together with Gronwall's inequality then implies

$$\Big\| \frac{d_\tau - d}{\tau} - \eta \Big\|_{L^\varrho(0,T;L^2(\Omega))} \leq C \|R(\cdot,\tau)\|_{L^\varrho(0,T)} \to 0 \quad \text{as } \tau \searrow 0,$$

which, combined with (4.9), gives the desired result. □

We now have all the necessary tools to conclude the main result of this section.

Proposition 4.4 (Directional differentiability of the control-to-state operator) *Let $M > 0$ and $\varrho \in [1,\infty)$ be given. Then, the operator $\mathcal{S} : \mathfrak{B}_M \to L^\varrho(0,T;V) \times L^\varrho(0,T;H^1(\Omega)) \times W^{1,\varrho}(0,T;L^2(\Omega))$ is directionally differentiable. The derivative at $\ell \in \mathfrak{B}_M$ in direction $\delta\ell \in L^\infty(0,T;\boldsymbol{W}^{-1,p}(\Omega))$, which we denote by $\mathcal{S}'(\ell;\delta\ell) := (\boldsymbol{\delta u}, \delta\varphi, \delta d)$, belongs to $L^\infty(0,T;V) \times L^\infty(0,T;H^1(\Omega)) \times W_0^{1,\infty}(0,T;L^2(\Omega))$. Moreover, this satisfies the following system*

$$\left\{ \begin{aligned} \boldsymbol{\delta u}(t) &= \mathcal{U}'(\ell(t),\varphi(t))(\delta\ell(t),\delta\varphi(t)), \\ \delta\varphi(t) &= \Phi'(\ell(t),d(t))(\delta\ell(t),\delta d(t)), \\ \dot{\delta d}(t) &= \frac{1}{\gamma} \max{}'\big(-\beta(d(t)-\varphi(t)) - r; -\beta(\delta d(t) - \delta\varphi(t))\big), \\ \delta d(0) &= 0 \end{aligned} \right\} \tag{4.10}$$

f.a.a. $t \in (0,T)$, where we abbreviate $\varphi := \mathcal{S}_2(\ell)$ and $d := \mathcal{S}_3(\ell)$.

Proof. Let $\ell \in \mathfrak{B}_M$ and $\delta\ell \in L^\infty(0,T;\boldsymbol{W}^{-1,p}(\Omega))$ be arbitrary, but fixed. From Lemma 4.3 we know that $\mathcal{S}_3 : \mathfrak{B}_M \to W^{1,\varrho}(0,T;L^2(\Omega))$ is directionally differentiable and $\mathcal{S}_3'(\ell;\delta\ell) \in W_0^{1,\infty}(0,T;L^2(\Omega))$ is the unique solution of the operator differential equation in (4.10). The directional differentiability of $\mathcal{S}_1$ and $\mathcal{S}_2$ is then an immediate consequence of Lemma 4.1. To see this, one constructs $L^\varrho(0,T)$-norms on both sides in (4.3). Moreover, according to Lemma 4.1, (4.1) and (4.2) we have

$$\begin{aligned} \mathcal{S}_1'(\ell;\delta\ell) &= \boldsymbol{\delta u} = \mathcal{U}'\big(\ell(\cdot),\varphi(\cdot)\big)\big(\delta\ell(\cdot),\delta\varphi(\cdot)\big) \in L^\infty(0,T;V), \\ \mathcal{S}_2'(\ell;\delta\ell) &= \delta\varphi = \Phi'\big(\ell(\cdot),d(\cdot)\big)\big(\delta\ell(\cdot),\delta d(\cdot)\big) \in L^\infty(0,T;H^1(\Omega)). \end{aligned}$$

This completes the proof. □

Remark 4.5 Note that due to the Lipschitz continuity of $\mathcal{S}$, cf. Lemma 3.5, the control-to-state operator is in fact Hadamard directional differentiable, see e.g. [24, Lemma 3.1.2].

5 Optimality system

This section is concerned with deriving an optimality system for the optimal control problem governed by the penalized damage model, i.e., (P_{min}). For convenience, let us recall that this reads as follows

$$\begin{aligned} \min_{\ell \in \mathfrak{L}} \quad & \mathcal{J}(\boldsymbol{u}, \varphi, d, \ell) \\ \text{s.t.} \quad & (\boldsymbol{u}, \varphi, d) \text{ solves (P)}, \end{aligned} \tag{P_{min}}$$

where $\mathfrak{L}$ denotes the set of admissible loads and $\mathcal{J}$ is the objective functional. In the last section we saw that the control-to-state operator is not necessarily Gâteaux-differentiable, and thus, the standard adjoint calculus for the derivation of optimality conditions is not applicable without further ado. However, by requiring that a strict complementarity condition is fulfilled, one is able to derive an optimality system equivalent to the classical purely primal optimality condition.

We begin by fixing the assumptions on the optimal control problem.

Assumption 5.1 The control set $\mathfrak{L}$ is a nonempty, convex and bounded subset of $L^\infty(0,T;\boldsymbol{W}^{-1,p}(\Omega))$.

Note that the boundedness assumption on the control set implies that there exists some $M > 0$ so that $\mathfrak{L} \subset \mathfrak{B}_M$. Therefore, the control-to-state operator is defined and directional differentiable on $\mathfrak{L}$.

Assumption 5.2 The objective functional $\mathcal{J} : L^2(0,T;\boldsymbol{W}_D^{1,\nu}(\Omega)) \times L^2(0,T;H^1(\Omega)) \times L^2(0,T;L^2(\Omega)) \times L^2(0,T;\boldsymbol{W}^{-1,p}(\Omega)) \to \mathbb{R}$ is continuously Fréchet-differentiable, where $\nu \in (1,2)$ is given.

For example, one can choose the functional $\mathcal{J}_{ex} : L^2(0,T;L^2(\Omega)) \times L^2(0,T;L^2(\Omega)) \times L^2(0,T;V^*)$ given by

$$\mathcal{J}_{ex}(\boldsymbol{u}, d, \ell) := \frac{1}{2} \|\boldsymbol{u} - \boldsymbol{u}_d\|^2_{L^2(0,T;L^2(\Omega))} + \frac{\alpha_1}{2} \|d\|^2_{L^2(0,T;L^2(\Omega))} + \frac{\alpha_2}{2} \|\ell\|^2_{L^2(0,T;V^*)}$$

as objective, where $\boldsymbol{u}_d \in L^2(0,T;L^2(\Omega))$ is a desired displacement and $\alpha_1, \alpha_2 > 0$. Note that $\mathcal{J}_{ex}$ fulfills Assumption 5.2 for $\nu \in (1,2)$ and $\nu \in [6/5,2)$ in the two-dimensional and three-dimensional case, respectively. This is due to $\boldsymbol{W}_D^{1,\nu}(\Omega) \hookrightarrow L^2(\Omega)$ (see e.g. [25, Theorem 7.1]) and $\boldsymbol{W}^{-1,p}(\Omega) \hookrightarrow V^*$. Moreover, note that $\mathcal{J}$ is continuously Fréchet-differentiable.

Remark 5.3 Although it might seem more reasonable to consider in Assumption 5.2 the Hilbert space $L^2(0,T;V)$ for the displacement, instead of $L^2(0,T;\boldsymbol{W}_D^{1,\nu}(\Omega))$, where $\nu \in (1,2)$, we choose to work with the latter one in order to obtain $L^2(0,T;V)$-regularity in both dimensions for the adjoint state $\boldsymbol{w}$ in Lemma 5.7 below. Otherwise, the adjoint state belongs to $L^2(0,T;\boldsymbol{W}_D^{1,\zeta}(\Omega))$, where $\zeta < 2$ in the three-dimensional case, while in the two-dimensional case it belongs to $L^2(0,T;V)$. For a better overview of the analysis in the proof of Lemma 5.7 below, it is convenient not to make this distinction.

With the properties of $\mathfrak{L}$ and $\mathcal{J}$ at hand, we can now derive necessary optimality conditions in primal form:

Lemma 5.4 *Let Assumptions 5.1 and 5.2 hold. Then, any local solution $\bar{\ell}$ of the problem (P_{min}) satisfies*

$$\partial_{(\boldsymbol{u},\varphi,d)}\mathcal{J}(\mathcal{S}(\bar{\ell}),\bar{\ell})\big(\mathcal{S}'(\bar{\ell};\delta\ell-\bar{\ell})\big)+\partial_{\ell}\mathcal{J}(\mathcal{S}(\bar{\ell}),\bar{\ell})(\delta\ell-\bar{\ell})\geq 0\quad\forall\,\delta\ell\in\mathfrak{L}. \tag{VI}$$

P r o o f. By Proposition 4.4, Assumption 5.2, and [14, Lemma 3.9], the composite mapping $\mathfrak{L}\ni\ell\mapsto\mathcal{J}(\mathcal{S}(\ell),\ell)\in\mathbb{R}$ is directionally differentiable at $\bar{\ell}\in\mathfrak{L}$ with

$$f'(\bar{\ell};\delta\ell-\bar{\ell})=\mathcal{J}'(\mathcal{S}(\bar{\ell}),\bar{\ell})(\mathcal{S}'(\bar{\ell};\delta\ell-\bar{\ell}),\delta\ell-\bar{\ell})\quad\forall\,\delta\ell\in L^{\infty}(0,T;\boldsymbol{W}^{-1,p}(\Omega)).$$

From the convexity of $\mathfrak{L}$ we then deduce that any local minimizer of (P_{min}) satisfies (VI), which completes the proof. □

Clearly, if $\mathcal{S}_3$ is Gâteaux-differentiable at $\bar{\ell}$, then all the terms in (VI) are linear in $\delta\ell-\bar{\ell}$. Since, within our scope of deriving an optimality system, we want to write the left-hand side in (VI) as a linear form in $\delta\ell-\bar{\ell}$, in the following we investigate under which conditions the nonlinearity of $\mathcal{S}_3'(\bar{\ell};\cdot)$ can be overcome. Therefor, a closer inspection of the operator differential equation in (4.10) is required. Prior to this, let us define for simplicity

Definition 5.5 For a given $\ell\in\mathfrak{L}$, we define for almost all $t\in(0,T)$ the following sets (up to sets of zero measure):

- $\Omega_t^+:=\{x\in\Omega:-\beta(d(t,x)-\varphi(t,x))-r>0\}$,
- $\Omega_t^0:=\{x\in\Omega:-\beta(d(t,x)-\varphi(t,x))-r=0\}$,
- $\Omega_t^-:=\{x\in\Omega:-\beta(d(t,x)-\varphi(t,x))-r<0\}$,

where we abbreviate $\varphi:=\mathcal{S}_2(\ell)$ and $d:=\mathcal{S}_3(\ell)$. We emphasize that the above defined sets ultimately depend on ℓ and the given data.

Let now $\ell\in\mathfrak{L}$ and $\delta\ell\in L^{\infty}(0,T;\boldsymbol{W}^{-1,p}(\Omega))$ be given. In view of Proposition 4.4 and (A.1) combined with Definition 5.5, $\delta d:=\mathcal{S}_3'(\ell;\delta\ell)$ is characterized as the unique solution of the operator differential equation

$$\dot{\delta d}(t)=\begin{cases}-\dfrac{\beta}{\gamma}(\delta d(t)-\delta\varphi(t)) & \text{a.e. in }\Omega_t^+\\ \max\big(-\dfrac{\beta}{\gamma}(\delta d(t)-\delta\varphi(t))\big) & \text{a.e. in }\Omega_t^0\\ 0 & \text{a.e. in }\Omega_t^-\end{cases}\quad\text{f.a.a. }t\in(0,T),\ \ \delta d(0)=0. \tag{5.1}$$

Here we use the notation $\delta\varphi:=\Phi'(\ell(\cdot),d(\cdot))(\delta\ell(\cdot),\delta d(\cdot))$ again, where $d:=\mathcal{S}_3(\ell)$.

From (5.1) we read, in view of the nonlinearity of $\max\{\cdot,0\}$, that as long as there exist $0\leq t_1<t_2\leq T$ such that $\mu(\Omega_t^0)>0$ f.a.a. $t\in(t_1,t_2)$, the operator $\mathcal{S}_3$ is not necessarily Gâteaux-differentiable at ℓ. This is also shown by straight forward computation. Indeed, the linearity of $\mathcal{S}_3'(\ell;\cdot)$ is ensured if the following assumption is satisfied, as we will next see.

Assumption 5.6 (Strict complementarity) The set Ω_t^0 associated to $\ell \in \mathfrak{L}$ has measure zero for almost all $t \in (0,T)$, i.e.,

$$\mu(\Omega_t^0) = 0 \quad \text{f.a.a. } t \in (0,T).$$

To prove that $\mathcal{S}_3'(\ell;\cdot)$ is linear under Assumption 5.6, note that (5.1) reads

$$\dot{\delta d}(t) = f_t(\delta\ell(t), \delta d(t)) \quad \text{f.a.a. } t \in (0,T), \quad \delta d(0) = 0, \tag{5.2}$$

where $f_t : \boldsymbol{W}^{-1,p}(\Omega) \times L^2(\Omega) \to L^2(\Omega)$ is given by

$$f_t(\delta\ell, \delta d) := -\frac{\beta}{\gamma}\chi_{\Omega_t^+}\big(\delta d - \Phi'(\ell(t), d(t))(\delta\ell, \delta d)\big).$$

Observe that, since Φ is Gâteaux-differentiable at $(\ell(t), d(t))$, f_t is linear for almost all $t \in (0,T)$. Now let $\delta\ell_1, \delta\ell_2 \in L^\infty(0,T;\boldsymbol{W}^{-1,p}(\Omega))$ and $a, b \in \mathbb{R}$. With the notation $\delta d_i := \mathcal{S}_3'(\ell;\delta\ell_i)$, where $i = 1,2$, (5.2) leads to

$$\begin{aligned} a\,\dot{\delta d}_1(t) + b\,\dot{\delta d}_2(t) &= a f_t(\delta\ell_1(t), \delta d_1(t)) + b f_t(\delta\ell_2(t), \delta d_2(t)) \\ &= f_t(a\,\delta\ell_1(t) + b\,\delta\ell_2(t), a\,\delta d_1(t) + b\,\delta d_2(t)) \quad \text{f.a.a. } t \in (0,T) \end{aligned}$$

and $a\delta d_1(0) + b\delta d_2(0) = 0$. Hence, $a\,\delta d_1 + b\,\delta d_2$ is the unique solution of (5.2) associated to $a\,\delta\ell_1 + b\,\delta\ell_2$, that is, $a\,\mathcal{S}_3'(\ell;\delta\ell_1) + b\,\mathcal{S}_3'(\ell;\delta\ell_2) = \mathcal{S}_3'(\ell; a\,\delta\ell_1 + b\,\delta\ell_2)$.

At the end of this section we make a few comments on the strict complementarity assumption, including possible alternative approaches, such as regularization, see Remark 5.11 below.

The following result is an essential tool for deriving the optimality system for (P_{min}), as it provides the candidates for the adjoint states associated to $\ell \in \mathfrak{L}$.

Lemma 5.7 (Adjoint equation) *Let Assumptions 5.1 and 5.2 hold. Moreover, let* $\ell \in \mathfrak{L}$ *be given and define* $(\boldsymbol{u}, \varphi, d) := \mathcal{S}(\ell)$. *Then, there exists a unique* $(\boldsymbol{w}, \upsilon, \xi) \in L^2(0,T;V) \times L^2(0,T;H^1(\Omega)) \times W_T^{1,2}(0,T;L^2(\Omega))$, *which satisfies the following system of equations*

$$-\operatorname{div}\big(g(\varphi(t))\mathbb{C}\varepsilon(\boldsymbol{w}(t)) + g'(\varphi(t))\upsilon(t)\mathbb{C}\varepsilon(\boldsymbol{u}(t))\big) = \partial_{\boldsymbol{u}}\mathcal{J}(\cdot)(t) \quad \text{in } V^*, \tag{5.3a}$$

$$\left.\begin{aligned} -\alpha\Delta\upsilon(t) + \beta\big(\upsilon(t) - \frac{1}{\gamma}\chi_{\Omega_t^+}\xi(t)\big) + g'(\varphi(t))\mathbb{C}\varepsilon(\boldsymbol{u}(t)) : \varepsilon(\boldsymbol{w}(t)) \\ + \frac{1}{2}g''(\varphi(t))\upsilon(t)\mathbb{C}\varepsilon(\boldsymbol{u}(t)) : \varepsilon(\boldsymbol{u}(t)) = \partial_\varphi\mathcal{J}(\cdot)(t) \quad \text{in } H^1(\Omega)^*, \end{aligned}\right\} \tag{5.3b}$$

$$\left.\begin{aligned} -\dot{\xi}(t) &= \beta\big(\upsilon(t) - \frac{1}{\gamma}\chi_{\Omega_t^+}\xi(t)\big) + \partial_d\mathcal{J}(\cdot)(t) \quad \text{in } L^2(\Omega), \\ \xi(T) &= 0 \end{aligned}\right\} \tag{5.3c}$$

f.a.a. $t \in (0,T)$, *where* $(\cdot)$ *stands for* $(\mathcal{S}(\ell), \ell)$.

P r o o f. Before we begin to discuss the solvability of (5.3), we observe that, since $\boldsymbol{W}_D^{1,\nu}(\Omega)$, $H^1(\Omega)$ and $L^2(\Omega)$ are reflexive Banach spaces, the partial derivatives $\partial_{\boldsymbol{u}}\mathcal{J}(\cdot)$, $\partial_\varphi\mathcal{J}(\cdot)$ and $\partial_d\mathcal{J}(\cdot)$ belong to $L^2(0,T;\boldsymbol{W}^{-1,\nu'}(\Omega))$, $L^2(0,T;H^1(\Omega)^*)$ and $L^2(0,T;L^2(\Omega))$, respectively. Note that, due to $\nu' > 2$, we have $\partial_{\boldsymbol{u}}\mathcal{J}(\cdot) \in L^2(0,T;V^*)$.

(i) Solvability of (5.3a). We search f.a.a. $t \in (0,T)$ and for any $\upsilon \in H^1(\Omega)$ for $\boldsymbol{w}$ such that

$$-\operatorname{div}\big(g(\varphi(t))\mathbb{C}\varepsilon(\boldsymbol{w})\big) = \partial_{\boldsymbol{u}}\mathcal{J}(\cdot)(t) + \operatorname{div}\big(g'(\varphi(t))\upsilon\mathbb{C}\varepsilon(\boldsymbol{u}(t))\big) \quad \text{in } V^*. \tag{5.4}$$

First we notice that $H^1(\Omega) \hookrightarrow L^{2p/(p-2)}(\Omega)$, as a result of $p > N$, cf. Assumption 2.5.1. Then, in view of Assumption 2.2 and the fact that $\boldsymbol{u}(t) \in \boldsymbol{W}^{1,p}_D(\Omega)$, we obtain with Hölder's inequality that, f.a.a. $t \in (0,T)$ and for any $\upsilon \in H^1(\Omega)$, it holds

$$\|\operatorname{div}\big(g'(\varphi(t))\upsilon\mathbb{C}\varepsilon(\boldsymbol{u}(t))\big)\|_{V^*} \le C\|\upsilon\|_{2p/(p-2)}\|\mathcal{U}(\ell(t),\varphi(t))\|_{\boldsymbol{W}^{1,p}_D(\Omega)} \le C\|\upsilon\|_{H^1(\Omega)}. \tag{5.5}$$

Note that we used (3.2) and the fact that ℓ belongs to a bounded subset of $L^\infty(0,T;\boldsymbol{W}^{-1,p}(\Omega))$ for the last inequality. Assumption 2.5.1 now ensures that (5.4) is uniquely solvable at almost all $t \in (0,T)$ and for any $\upsilon \in H^1(\Omega)$ with

$$\boldsymbol{w}(t,\upsilon) = A^{-1}_{\varphi(t)}\big(\partial_{\boldsymbol{u}}\mathcal{J}(\cdot)(t) + \operatorname{div}\big(g'(\varphi(t))\upsilon\mathbb{C}\varepsilon(\boldsymbol{u}(t))\big)\big) \in V. \tag{5.6}$$

Moreover, we have

$$\|\boldsymbol{w}(t,\upsilon)\|_V \le C\big(\|\partial_{\boldsymbol{u}}\mathcal{J}(\cdot)(t)\|_{V^*} + \|\upsilon\|_{H^1(\Omega)}\big) \ \text{ f.a.a. } t \in (0,T),\ \forall\, \upsilon \in H^1(\Omega). \tag{5.7}$$

We now show that the Nemytskii operator associated to $\boldsymbol{w}$ maps $L^2(0,T;H^1(\Omega))$ to $L^2(0,T;V)$. This result will be useful at the end of the proof, after the time regularity of the solution of the operator differential equation (5.3c) is established. First, let us observe that $H^1(\Omega) \ni \varphi \mapsto A^{-1}_\varphi \in \mathcal{L}(\boldsymbol{W}^{-1,q}(\Omega), V)$ is continuous, provided that $q > 2$, whence the Bochner measurability of $t \mapsto A^{-1}_{\varphi(t)} \in \mathcal{L}(\boldsymbol{W}^{-1,q}(\Omega), V)$. Let now $\upsilon \in L^2(0,T;H^1(\Omega))$ be arbitrary, but fixed. Similarly to (5.5), we obtain the following estimate f.a.a. $t \in (0,T)$

$$\|\operatorname{div}\big(g'(\varphi(t))\upsilon(t)\mathbb{C}\varepsilon(\boldsymbol{u}(t))\big)\|_{\boldsymbol{W}^{-1,\rho}(\Omega)} \le \|g'(\varphi(t))\|_\varrho\|\upsilon(t)\|_\kappa\|\mathbb{C}\varepsilon(\boldsymbol{u}(t))\|_s. \tag{5.8}$$

Here the index $\rho > 2$ is defined via $1/\rho := 1/\varrho + 1/\kappa + 1/s < 1/2$, where $\kappa > 2p/(p-2)$ is chosen so that the embedding $H^1(\Omega) \hookrightarrow L^\kappa(\Omega)$ holds true, whereas $\varrho < \infty$ and $s \in [2,p)$ are chosen large enough, such that $1/\varrho + 1/\kappa + 1/s < (p-2)/2p + 1/p = 1/2$. Notice that it is possible to choose κ as above, in view of $p > N$, see Assumption 2.5.1. Further, we infer from [7, Theorem 4] that $g' : H^1(\Omega) \to L^\varrho(\Omega)$ is continuous, and thus, the mapping $t \mapsto g'(\varphi(t)) \in L^\varrho(\Omega)$ is Bochner measurable, as $\varphi : [0,T] \to H^1(\Omega)$ is. Moreover, $t \mapsto \mathbb{C}\varepsilon(\boldsymbol{u}(t)) \in L^s(\Omega;\mathbb{R}^{N\times N})$ is Bochner measurable as well, as a result of the regularity of $\boldsymbol{u}$. On account of (5.8), this implies the Bochner measurability of $t \mapsto \partial_{\boldsymbol{u}}\mathcal{J}(\cdot)(t) + \operatorname{div}\big(g'(\varphi(t))\upsilon(t)\mathbb{C}\varepsilon(\boldsymbol{u}(t))\big) \in \boldsymbol{W}^{-1,\omega}(\Omega)$, where $\omega := \min\{\rho,\nu'\} > 2$. Therefor we also employed that $\upsilon : [0,T] \to H^1(\Omega) \hookrightarrow L^\kappa(\Omega)$ is Bochner measurable, by assumption. Altogether, we can now deduce by (5.6) the Bochner measurability of $t \mapsto \boldsymbol{w}(t,\upsilon(t)) \in V$. Assumption 5.2 together with (5.7) finally yields

$$\boldsymbol{w}(\cdot,\upsilon(\cdot)) \in L^2(0,T;V) \text{ for } \upsilon \in L^2(0,T;H^1(\Omega)). \tag{5.9}$$

(ii) Solvability of (5.3b). For almost all $t \in (0,T)$ and for any $\xi \in L^2(\Omega)$, we now search for $\upsilon \in H^1(\Omega)$ that solves

$$\left.\begin{aligned} -\alpha\triangle\upsilon + \beta\upsilon + g'(\varphi(t))\mathbb{C}\varepsilon(\boldsymbol{u}(t)) : \varepsilon(\boldsymbol{w}(t,\upsilon)) + \frac{1}{2}g''(\varphi(t))\upsilon\mathbb{C}\varepsilon(\boldsymbol{u}(t)) : \varepsilon(\boldsymbol{u}(t)) \\ = \partial_\varphi\mathcal{J}(\cdot)(t) + \frac{\beta}{\gamma}\chi_{\Omega_t^+}\xi \quad \text{in } H^1(\Omega)^*. \end{aligned}\right\} \tag{5.10}$$

To this end, we first define $B := -\alpha\triangle + \beta I$ and the mapping $F : \boldsymbol{W}^{-1,p}(\Omega) \times H^1(\Omega) \ni (\ell,\varphi) \mapsto \frac{1}{2}g'(\varphi)\mathbb{C}\varepsilon(\mathcal{U}(\ell,\varphi)) : \varepsilon(\mathcal{U}(\ell,\varphi)) \in H^1(\Omega)^*$. In view of (5.6) and [18, Lemma 5.3], it holds

$$\boldsymbol{w}(t,\upsilon) = A^{-1}_{\varphi(t)}\partial_{\boldsymbol{u}}\mathcal{J}(\cdot)(t) + \partial_\varphi\mathcal{U}(\ell(t),\varphi(t))(\upsilon) \quad \text{f.a.a. } t \in (0,T),\ \forall\, \upsilon \in H^1(\Omega).$$

By taking a look at [18, Lemma 5.9], one sees now that (5.10) is in fact equivalent to

$$B\upsilon + \partial_\varphi F(\ell(t),\varphi(t))\upsilon = \iota(t,\xi) \quad \text{in } H^1(\Omega)^*,$$

where the mapping $\iota : (0,T) \times L^2(\Omega) \to H^1(\Omega)^*$ is given by

$$\iota(t,\xi) := \partial_\varphi\mathcal{J}(\cdot)(t) + \frac{\beta}{\gamma}\chi_{\Omega_t^+}\xi - g'(\varphi(t))\mathbb{C}\varepsilon(\boldsymbol{u}(t)) : \varepsilon\big(A^{-1}_{\varphi(t)}\big(\partial_{\boldsymbol{u}}\mathcal{J}(\cdot)(t)\big)\big). \tag{5.11}$$

Assumption 2.2, (3.2), Assumption 2.5.1 combined with Hölder's inequality with $1/p + 1/2 + (p-2)/2p = 1$, and $H^1(\Omega) \hookrightarrow L^{2p/(p-2)}(\Omega)$ (see Assumption 2.5.1) yield

$$\big\|g'(\varphi(t))\mathbb{C}\varepsilon(\boldsymbol{u}(t)) : \varepsilon\big(A^{-1}_{\varphi(t)}\big(\partial_{\boldsymbol{u}}\mathcal{J}(\cdot)(t)\big)\big)\big\|_{H^1(\Omega)^*} \le C\|\partial_{\boldsymbol{u}}\mathcal{J}(\cdot)(t)\|_{V^*} \ \text{f.a.a. } t \in (0,T). \tag{5.12}$$

Thus, ι is well defined. From [18, Proof of Lemma 5.11], see also [18, Definition 5.7], we know that $B + \partial_\varphi F(\ell(t),\varphi(t)) \in \mathcal{L}(H^1(\Omega), H^1(\Omega)^*)$ is continuously invertible f.a.a. $t \in (0,T)$ with

$$\|\big(B + \partial_\varphi F(\ell(t),\varphi(t))\big)^{-1}\|_{\mathcal{L}(H^1(\Omega)^*,H^1(\Omega))} \le c \quad \text{f.a.a. } t \in (0,T). \tag{5.13}$$

Hence, (5.10) is uniquely solvable f.a.a. $t \in (0,T)$ and for any $\xi \in L^2(\Omega)$ with

$$\upsilon(t,\xi) = \big(B + \partial_\varphi F(\ell(t),\varphi(t))\big)^{-1}\iota(t,\xi) \in H^1(\Omega). \tag{5.14}$$

Moreover, due to (5.13) and (5.11), we have

$$\|\upsilon(t,\xi_1) - \upsilon(t,\xi_2)\|_{H^1(\Omega)} \le \frac{2}{\alpha}\|\iota(t,\xi_1) - \iota(t,\xi_2)\|_{H^1(\Omega)^*} \le \frac{2\beta}{\alpha\gamma}\|\xi_1 - \xi_2\|_2 \tag{5.15}$$

f.a.a. $t \in (0,T)$ and for all $\xi_1,\xi_2 \in L^2(\Omega)$. In preparation for the next part of the proof, we prove that υ belongs to $L^2(0,T;H^1(\Omega))$ if $\xi \in L^2(0,T;L^2(\Omega))$ in the following. Firstly, we observe that

$$\big(B + \partial_\varphi F(\ell(\cdot),\varphi(\cdot))\big)^{-1} \in L^\infty\big(0,T;\mathcal{L}(H^1(\Omega)^*,H^1(\Omega))\big). \tag{5.16}$$

This is a result of the continuity of $\partial_\varphi F : \boldsymbol{W}^{-1,p}(\Omega) \times H^1(\Omega) \to \mathcal{L}(H^1(\Omega), H^1(\Omega)^*)$, cf. [18, Lemma 5.9], which gives the Bochner measurability of the mapping in (5.16), and estimate (5.13). Secondly, it holds

$$\iota(\cdot, \xi(\cdot)) \in L^2(0,T; H^1(\Omega)^*) \quad \text{for } \xi \in L^2(0,T; L^2(\Omega)). \tag{5.17}$$

To see this, note that Definition 5.5 allows us to write $\chi_{\Omega_t^+}(x)\xi(t)(x) = (\chi_{Q_+}\xi)(t,x)$ f.a.a. $(t,x) \in (0,T) \times \Omega$, where $Q_+ := \{(t,x) \in (0,T) \times \Omega : -\beta(d(t,x) - \varphi(t,x)) - r > 0\}$ and $\xi \in L^2(0,T; L^2(\Omega))$. Since $\chi_{Q_+} \in L^\infty((0,T) \times \Omega)$, we have $\chi_{Q_+}\xi \in L^2((0,T) \times \Omega)$, whence

$$t \mapsto \chi_{\Omega_t^+}\xi(t) \in L^2(0,T; L^2(\Omega)) \text{ for } \xi \in L^2(0,T; L^2(\Omega)). \tag{5.18}$$

By arguing as at the end of part (i), we infer that $t \mapsto g'(\varphi(t))\mathbb{C}\varepsilon(\boldsymbol{u}(t)) : \varepsilon\big(A_{\varphi(t)}^{-1}\big(\partial_{\boldsymbol{u}}\mathcal{J}(\cdot)(t)\big)\big) \in H^1(\Omega)^*$ is Bochner measurable, on account of Hölder's inequality. From (5.12) we then obtain $t \mapsto g'(\varphi(t))\mathbb{C}\varepsilon(\boldsymbol{u}(t)) : \varepsilon\big(A_{\varphi(t)}^{-1}\big(\partial_{\boldsymbol{u}}\mathcal{J}(\cdot)(t)\big)\big) \in L^2(0,T; H^1(\Omega)^*)$, which together with (5.18) and $\partial_\varphi\mathcal{J}(\cdot) \in L^2(0,T; H^1(\Omega)^*)$ now gives (5.17). This in combination with (5.16) finally yields

$$\upsilon(\cdot, \xi(\cdot)) \in L^2(0,T; H^1(\Omega)) \text{ for } \xi \in L^2(0,T; L^2(\Omega)), \tag{5.19}$$

in view of (5.14).

(iii) Solvability of (5.3c). By means of the 'solution operator' $(t,\xi) \mapsto \upsilon(t,\xi)$ of the elliptic system, (5.3) reduces to

$$\left.\begin{aligned} &-\dot{\xi}(t) = \beta\big(\upsilon(t,\xi(t)) - \frac{1}{\gamma}\chi_{\Omega_t^+}\xi(t)\big) + \partial_d\mathcal{J}(\cdot)(t) \quad \text{f.a.a. } t \in (0,T), \\ &\xi(T) = 0. \end{aligned}\right\} \tag{5.20}$$

We see that, via the transformation $\widetilde{\xi}(\cdot) = \xi(T - \cdot)$, (5.20) is equivalent to

$$\dot{\widetilde{\xi}}(t) = f(t, \widetilde{\xi}(t)) \ \text{ f.a.a. } t \in (0,T), \quad \widetilde{\xi}(0) \ = 0, \tag{5.21}$$

where $f : (0,T) \times L^2(\Omega) \to L^2(\Omega)$ is given by

$$f(t, \widetilde{\xi}) = \beta\Big(\upsilon(T-t, \widetilde{\xi}) - \frac{1}{\gamma}\chi_{\Omega_{T-t}^+}\widetilde{\xi}\Big) + \partial_d\mathcal{J}(\cdot)(T-t). \tag{5.22}$$

To show that (5.21) is uniquely solvable, we proceed as in the proof of Lemma 4.2. Let now $\widetilde{\xi} \in L^2(0,T; L^2(\Omega))$ be arbitrary, but fixed. This implies $t \mapsto \widetilde{\xi}(T-t) \in L^2(0,T; L^2(\Omega))$ and according to (5.19), we then have $t \mapsto \upsilon(T-t, \widetilde{\xi}(t)) \in L^2(0,T; H^1(\Omega))$. With (5.18) we establish $t \mapsto \chi_{\Omega_{T-t}^+}\widetilde{\xi}(t) \in L^2(0,T; L^2(\Omega))$, which combined with $t \mapsto \partial_d\mathcal{J}(\cdot)(T-t) \in L^2(0,T; L^2(\Omega))$ gives in turn $f(\cdot, \widetilde{\xi}(\cdot)) \in L^2(0,T; L^2(\Omega))$, in view of (5.22). Moreover, (5.15) implies

$$\begin{aligned} \|f(t,\widetilde{\xi}_1) - f(t,\widetilde{\xi}_2)\|_2 &\le \beta\|\upsilon(T-t,\widetilde{\xi}_1) - \upsilon(T-t,\widetilde{\xi}_2)\|_2 + \beta/\gamma\|\widetilde{\xi}_1 - \widetilde{\xi}_2\|_2 \\ &\le L\|\widetilde{\xi}_1 - \widetilde{\xi}_2\|_2 \quad \forall \widetilde{\xi}_1, \widetilde{\xi}_2 \in L^2(\Omega), \ \text{ f.a.a. } t \in (0,T). \end{aligned}$$

With the above properties of f, the unique solvability of (5.21) follows by a fixed point argument. Therefrom we also obtain $\widetilde{\xi} \in W_0^{1,2}(0,T;L^2(\Omega))$. This means that (5.20) is uniquely solvable with $t \mapsto \xi(t) = \widetilde{\xi}(T-t) \in W_T^{1,2}(0,T;L^2(\Omega))$. Now (5.9) and (5.19) give the regularity of the adjoint states in the elliptic system (5.3a)-(5.3b), which completes the proof. □

Remark 5.8 Clearly, given $\ell \in \mathfrak{L}$, the system (5.3) admits a unique solution (with the exact same regularity) for an arbitrary right-hand side $h \in L^2(0,T;\boldsymbol{W}^{-1,\nu'}(\Omega)) \times L^2(0,T;H^1(\Omega)^*) \times L^2(0,T;L^2(\Omega))$ (instead of $\partial_{(\boldsymbol{u},\varphi,d)}\mathcal{J}(\mathcal{S}(\ell),\ell)$), where $\nu \in (1,2)$ is given. The thereby induced solution operator, which we call Υ here, can be interpreted as an 'artificial adjoint operator' of $\mathcal{S}'(\ell)$, provided that Assumption 5.6 holds true for ℓ. We call this adjoint operator 'artificial', since its range is a subset of a product space and not a subset of $L^\infty(0,T;\boldsymbol{W}^{-1,p}(\Omega))^*$, as one would expect. This holds however true for the range of the first component of Ψ. To see this, we linearize the state equation (5.23) below at ℓ in some arbitrary direction $\delta\ell \in L^\infty(0,T;\boldsymbol{W}^{-1,p}(\Omega))$, and argue as in the proof of Theorem 5.9 below, where we replace the partial derivatives of $\mathcal{J}$ in (5.24) by h as above. Then, instead of (5.32) below, one has the equality

$$\int_0^T \sum_{i=1}^3 \langle h_i(t), \mathcal{S}_i'(\ell)(\delta\ell)(t)\rangle \, dt = \int_0^T \langle \delta\ell(t), \Upsilon_1(h)(t)\rangle_V \, dt,$$

i.e., $\langle h, \mathcal{S}'(\ell)(\delta\ell)\rangle_{L^2(0,T;\boldsymbol{W}_D^{1,\nu}(\Omega)\times H^1(\Omega)\times L^2(\Omega))} = \langle \Upsilon_1 h, \delta\ell\rangle_{L^2(0,T;\boldsymbol{W}^{-1,p}(\Omega))}$ for any $h \in L^2(0,T;\boldsymbol{W}^{-1,\nu'}(\Omega) \times H^1(\Omega)^* \times L^2(\Omega))$ and any $\delta\ell \in L^\infty(0,T;\boldsymbol{W}^{-1,p}(\Omega))$. Here Υ_1 stands for the first component of the operator Υ. We point out that Υ_1 is the classical adjoint operator of $\mathcal{S}'(\ell) : L^\infty(0,T;\boldsymbol{W}^{-1,p}(\Omega)) \to L^2(0,T;\boldsymbol{W}_D^{1,\nu}(\Omega) \times H^1(\Omega) \times L^2(\Omega))$, as a result of the above identity and as it maps as follows

$$L^2(0,T;\boldsymbol{W}^{-1,\nu'}(\Omega) \times H^1(\Omega)^* \times L^2(\Omega)) \overset{\Upsilon_1}{\longmapsto} L^2(0,T;V) \subset L^\infty(0,T;\boldsymbol{W}^{-1,p}(\Omega))^*,$$

in view of Lemma 5.7 and $L^\infty(0,T;\boldsymbol{W}^{-1,p}(\Omega)) \hookrightarrow L^2(0,T;V^*)$.

We are now in the position to state the main result of this section.

Theorem 5.9 (Optimality system) *Let Assumptions 5.1 and 5.2 hold. Moreover, let $\bar{\ell}$ be a local solution of (P_{min}) with associated states*

$$(\bar{\boldsymbol{u}}, \bar{\varphi}, \bar{d}) = \mathcal{S}(\bar{\ell}) \in L^\infty(0,T;\boldsymbol{W}_D^{1,s}(\Omega)) \times L^\infty(0,T;H^1(\Omega)) \times W^{1,\infty}(0,T;L^2(\Omega)),$$

where $s \in [2,p)$, and suppose that Assumption 5.6 is fulfilled for $\bar{\ell}$. Then, there exist unique adjoint states

$$(\boldsymbol{w}, \upsilon, \xi) \in L^2(0,T;V) \times L^2(0,T;H^1(\Omega)) \times W_T^{1,2}(0,T;L^2(\Omega))$$

so that the following optimality system is satisfied f.a.a. $t \in (0,T)$*:*

$$-\operatorname{div} g(\bar\varphi(t))\mathbb{C}\varepsilon(\bar{\boldsymbol{u}}(t)) = \bar\ell(t) \quad \text{in } \boldsymbol{W}^{-1,p}(\Omega), \tag{5.23a}$$

$$-\alpha\Delta\bar\varphi(t) + \beta\,\bar\varphi(t) + \frac{1}{2}\,g'(\bar\varphi(t))\mathbb{C}\,\varepsilon(\bar{\boldsymbol{u}}(t)) : \varepsilon(\bar{\boldsymbol{u}}(t)) = \beta\bar d(t) \quad \text{in } H^1(\Omega)^*, \tag{5.23b}$$

$$\dot{\bar d}(t) = \frac{1}{\gamma}\max(-\beta(\bar d(t) - \bar\varphi(t)) - r), \quad \bar d(0) = d_0, \tag{5.23c}$$

$$-\operatorname{div}\big(g(\bar\varphi(t))\mathbb{C}\varepsilon(\boldsymbol{w}(t)) + g'(\bar\varphi(t))\upsilon(t)\mathbb{C}\varepsilon(\bar{\boldsymbol{u}}(t))\big) = \partial_{\boldsymbol{u}}\mathcal{J}(\cdot)(t) \ \text{ in } V^*, \tag{5.24a}$$

$$\left.\begin{aligned} &-\alpha\Delta\upsilon(t) + \beta\big(\upsilon(t) - \frac{1}{\gamma}\chi_{\Omega_t^+}\xi(t)\big) + g'(\bar\varphi(t))\mathbb{C}\varepsilon(\bar{\boldsymbol{u}}(t)) : \varepsilon(\boldsymbol{w}(t)) \\ &\qquad + \frac{1}{2}g''(\bar\varphi(t))\upsilon(t)\mathbb{C}\varepsilon(\bar{\boldsymbol{u}}(t)) : \varepsilon(\bar{\boldsymbol{u}}(t)) = \partial_\varphi\mathcal{J}(\cdot)(t) \ \text{ in } H^1(\Omega)^*, \end{aligned}\right\} \tag{5.24b}$$

$$\left.\begin{aligned} &-\dot\xi(t) = \beta\big(\upsilon(t) - \frac{1}{\gamma}\chi_{\Omega_t^+}\xi(t)\big) + \partial_d\mathcal{J}(\cdot)(t) \ \text{ in } L^2(\Omega), \\ &\xi(T) = 0, \end{aligned}\right\} \tag{5.24c}$$

$$\langle \boldsymbol{w} + \partial_\ell\mathcal{J}(\cdot), \delta\ell - \bar\ell\rangle_{L^2(0,T;\boldsymbol{W}^{-1,p}(\Omega))} \geq 0 \quad \forall\,\delta\ell \in \mathfrak{L}, \tag{5.25}$$

where $(\cdot)$ *denotes* $(\mathcal{S}(\bar\ell), \bar\ell)$.

P r o o f. The state equation is a direct result of Definition 3.4, while from Lemma 5.7 we know that (5.24) admits a unique solution $(\boldsymbol{w}, \upsilon, \xi)$ with the desired regularity. Hence, it remains to prove the gradient inequality (5.25). To this end, we test the linearized counterpart of (5.23) with the adjoint states and show that the (integrated over time) sum of the resulting equations is the (integrated over time) sum of the equations in (5.24) tested with the directional derivatives of $\mathcal{S}$. Using this in (VI) will ultimately give the claim. For a better overview, we split the rest of the proof in two parts: in the first one we derive a linearization for the state equation, while in the second one we test as depicted above and finalize the proof.

(i) Let $\delta\ell \in L^\infty(0,T;\boldsymbol{W}^{-1,p}(\Omega))$ be arbitrary, but fixed and define $(\boldsymbol{\delta u}, \delta\varphi, \delta d) := \mathcal{S}'(\bar\ell; \delta\ell - \bar\ell)$. Proposition 4.4 and Assumption 5.6 ensure that $(\boldsymbol{\delta u}, \delta\varphi, \delta d)$ satisfies the system

$$\boldsymbol{\delta u}(t) = \mathcal{U}'(\bar\ell(t), \bar\varphi(t))(\delta\ell(t) - \bar\ell(t), \delta\varphi(t)), \tag{5.26a}$$

$$\delta\varphi(t) = \Phi'(\bar\ell(t), \bar d(t))(\delta\ell(t) - \bar\ell(t), \delta d(t)), \tag{5.26b}$$

$$\dot{\delta d}(t) = -\frac{\beta}{\gamma}\chi_{\Omega_t^+}(\delta d(t) - \delta\varphi(t)), \quad \delta d(0) = 0 \tag{5.26c}$$

f.a.a. $t \in (0,T)$.

From the differentiability results in [18, Section 5] we can deduce the equations which characterize $\boldsymbol{\delta u}$ and $\delta\varphi$ as follows. Completely analogously to the result in [18, Lemma 5.2], we have $-\operatorname{div}\big(g(\bar\varphi(t))\mathbb{C}\varepsilon(\partial_\ell \mathcal{U}(\bar\ell(t),\bar\varphi(t))\eta)\big) = \eta \quad \forall \eta \in \boldsymbol{W}^{-1,p}(\Omega)$, which in combination with [18, Lemma 5.3] leads to

$$-\operatorname{div}\big(g(\bar\varphi(t))\mathbb{C}\varepsilon(\boldsymbol{\delta u}(t))\big) = \delta\ell(t) - \bar\ell(t) + \operatorname{div}\big(g'(\bar\varphi(t))\delta\varphi(t)\mathbb{C}\varepsilon(\bar{\boldsymbol{u}}(t))\big) \quad \text{in } V^* \tag{5.27}$$

f.a.a. $t \in (0,T)$. Further, as in [18, Lemma 5.9, Proposition 5.12], we arrive at

$$\left.\begin{aligned} -\alpha\Delta\delta\varphi(t) + \beta\delta\varphi(t) + \frac{1}{2}g''(\bar\varphi(t))\delta\varphi(t)\mathbb{C}\varepsilon(\bar{\boldsymbol{u}}(t)) : \varepsilon(\bar{\boldsymbol{u}}(t)) \\ = \beta\delta d(t) - g'(\bar\varphi(t))\mathbb{C}\varepsilon(\bar{\boldsymbol{u}}(t)) : \varepsilon(\boldsymbol{\delta u}(t)) \quad \text{in } H^1(\Omega)^* \end{aligned}\right\} \tag{5.28}$$

f.a.a. $t \in (0,T)$. Let us recall here that $(\boldsymbol{\delta u}, \delta\varphi, \delta d)$ belongs to $L^\infty(0,T;V)\times L^\infty(0,T;H^1(\Omega))\times W_0^{1,\infty}(0,T;L^2(\Omega))$, cf. Proposition 4.4.

(ii) We test (5.27), (5.28), and (5.26c) with $\boldsymbol{w}(t) \in V$, $\upsilon(t) \in H^1(\Omega)$, and $\xi(t) \in L^2(\Omega)$, respectively, at almost all $t \in (0,T)$. For the sake of convenience, we bring all the resulting terms containing $\boldsymbol{\delta u}$ and $\delta\varphi$ together, by abbreviating

$$\begin{aligned} \iota(t) := &-\langle \operatorname{div}\big(g(\bar\varphi(t))\mathbb{C}\varepsilon(\boldsymbol{\delta u}(t)) + g'(\bar\varphi(t))\delta\varphi(t)\mathbb{C}\varepsilon(\bar{\boldsymbol{u}}(t))\big), \boldsymbol{w}(t)\rangle_V \\ &+ \langle -\alpha\Delta\delta\varphi(t) + \beta\delta\varphi(t) + \frac{1}{2}g''(\bar\varphi(t))\delta\varphi(t)\mathbb{C}\varepsilon(\bar{\boldsymbol{u}}(t)) : \varepsilon(\bar{\boldsymbol{u}}(t)), \upsilon(t)\rangle_{H^1(\Omega)} \\ &+ \langle g'(\bar\varphi(t))\mathbb{C}\varepsilon(\bar{\boldsymbol{u}}(t)) : \varepsilon(\boldsymbol{\delta u}(t)), \upsilon(t)\rangle_{H^1(\Omega)} - \frac{\beta}{\gamma}(\chi_{\Omega_t^+}\delta\varphi(t), \xi(t))_2 \quad \text{f.a.a. } t \in (0,T). \end{aligned}$$

Thereby, the sum of the above tested equations reads as follows

$$\iota(t) + (\dot{\delta d}(t), \xi(t))_2 = \beta(\delta d(t), \upsilon(t))_2 - \frac{\beta}{\gamma}(\chi_{\Omega_t^+}\delta d(t), \xi(t))_2 + \langle \delta\ell(t) - \bar\ell(t), \boldsymbol{w}(t)\rangle_V \tag{5.29}$$

f.a.a. $t \in (0,T)$. Furthermore, as a result of (5.24a) and (5.24b), we observe that

$$\iota(t) = \langle \partial_{\boldsymbol{u}}\mathcal{J}(\cdot)(t), \boldsymbol{\delta u}(t)\rangle_V + \langle \partial_\varphi \mathcal{J}(\cdot)(t), \delta\varphi(t)\rangle_{H^1(\Omega)} \quad \text{f.a.a. } t \in (0,T). \tag{5.30}$$

Testing (5.24c) with $\delta d(t)$ and integration by parts results in

$$\begin{aligned} \int_0^T (\partial_d \mathcal{J}(\cdot)(t), \delta d(t))_2 \, dt &= -\int_0^T (\dot\xi(t), \delta d(t))_2 \, dt - \int_0^T \beta(\upsilon(t), \delta d(t))_2 - \frac{\beta}{\gamma}(\chi_{\Omega_t^+}\xi(t), \delta d(t))_2 \, dt \\ &= -(\xi(T), \delta d(T))_2 + (\xi(0), \delta d(0))_2 + \int_0^T (\dot{\delta d}(t), \xi(t))_2 \\ &\quad - \int_0^T \beta(\upsilon(t), \delta d(t))_2 - \frac{\beta}{\gamma}(\chi_{\Omega_t^+}\xi(t), \delta d(t))_2 \, dt \\ &= \int_0^T (\dot{\delta d}(t), \xi(t))_2 - \int_0^T \beta(\upsilon(t), \delta d(t))_2 - \frac{\beta}{\gamma}(\chi_{\Omega_t^+}\xi(t), \delta d(t))_2 \, dt, \end{aligned} \tag{5.31}$$

where we employed $\xi \in W_T^{1,2}(0,T;L^2(\Omega))$, $\delta d \in W_0^{1,\infty}(0,T;L^2(\Omega))$. Inserting (5.30) and (5.31) in (5.29) leads to

$$\int_0^T \langle \partial_{\boldsymbol{u}} \mathcal{J}(\cdot)(t), \boldsymbol{\delta u}(t)\rangle_V + \langle \partial_\varphi \mathcal{J}(\cdot)(t), \delta\varphi(t)\rangle_{H^1(\Omega)} + (\partial_d \mathcal{J}(\cdot)(t), \delta d(t))_2 \, dt$$
$$= \int_0^T \langle \delta\ell(t) - \bar{\ell}(t), \boldsymbol{w}(t)\rangle_V \, dt. \tag{5.32}$$

As $\bar{\ell}$ is a local minimizer of (P_{min}), it satisfies the variational inequality in Lemma 5.4, which in light of (5.32) now reads

$$\int_0^T \langle \delta\ell(t) - \bar{\ell}(t), \boldsymbol{w}(t)\rangle_V \, dt + \langle \partial_\ell \mathcal{J}(\mathcal{S}(\bar{\ell}), \bar{\ell}), \delta\ell - \bar{\ell}\rangle_{L^2(0,T;\boldsymbol{W}^{-1,p}(\Omega))}$$
$$= \langle \boldsymbol{w}, \delta\ell - \bar{\ell}\rangle_{L^2(0,T;\boldsymbol{W}^{-1,p}(\Omega))} + \langle \partial_\ell \mathcal{J}(\mathcal{S}(\bar{\ell}), \bar{\ell}), \delta\ell - \bar{\ell}\rangle_{L^2(0,T;\boldsymbol{W}^{-1,p}(\Omega))} \geq 0 \quad \forall \delta\ell \in \mathfrak{L}, \tag{5.33}$$

where we employed the reflexivity of V, $H^1(\Omega)$, and $L^2(\Omega)$, the embedding $V \hookrightarrow \boldsymbol{W}^{-1,p}(\Omega)^*$, as well as the regularity of $\boldsymbol{w}$. This completes the proof. □

From the above proof we deduce that the optimality system (5.23)-(5.25) is in fact equivalent to the first order necessary optimality condition in Lemma 5.4, provided that the strict complementarity assumption is fulfilled:

Proposition 5.10 *Suppose that Assumptions 5.1 and 5.2 hold true. Moreover, let Assumption 5.6 be fulfilled for some $\bar{\ell} \in \mathfrak{L}$. Then, if $\bar{\ell}$ together with its states*

$$(\bar{\boldsymbol{u}}, \bar{\varphi}, \bar{d}) \in L^\infty(0,T;\boldsymbol{W}_D^{1,s}(\Omega)) \times L^\infty(0,T;H^1(\Omega)) \times W^{1,\infty}(0,T;L^2(\Omega)),$$

where $s \in [2,p)$, and adjoint states

$$(\boldsymbol{w}, \upsilon, \xi) \in L^2(0,T;V) \times L^2(0,T;H^1(\Omega)) \times W_T^{1,2}(0,T;L^2(\Omega)),$$

satisfies the optimality system (5.23)-(5.25)*, it also satisfies the variational inequality* (VI).

We end this section with some remarks related with Assumption 5.6. First, let us point out that one can reformulate (P_{min}) as an MPCC, provided that the set $\mathfrak{L}$ is chosen accordingly. In view of (3.5) and since the $\max\{\cdot, 0\}$-function is a complementarity function, the unique solution $(\boldsymbol{u}, \varphi, d)$ of the problem (P) with right-hand side $\ell \in \mathfrak{L}$ is characterized by

$$\begin{aligned}
-\operatorname{div} g(\varphi(t))\mathbb{C}\varepsilon(\boldsymbol{u}(t)) &= \ell(t) && \text{in } \boldsymbol{W}^{-1,p}(\Omega),\\
-\alpha\Delta\varphi(t) + \beta\,\varphi(t) + \frac{1}{2}\, g'(\varphi(t))\mathbb{C}\,\varepsilon(\boldsymbol{u}(t)) : \varepsilon(\boldsymbol{u}(t)) &= \beta d(t) && \text{in } H^1(\Omega)^*,\\
0 \leq \dot{d}(t) \perp \beta(d(t) - \varphi(t)) + r + \gamma\dot{d}(t) &\geq 0 \ \text{ a.e. in } \Omega, && d(0) = d_0
\end{aligned}$$

f.a.a. $t \in (0,T)$. Thus, if the control set can be described e.g. only by inequalities, e.g. $\mathfrak{L} = \overline{\mathfrak{B}_M}$, then the problem (P_{min}) falls into the class of MPCCs, see [23] for the definition thereof in the finite-dimensional case.

Remark 5.11 (Strict complementarity) In contrast to our approach, most authors dealing with time-dependent MPCCs make use of regularization and penalization techniques, see e.g. [1, 17] (parabolic obstacle problem), [27] (quasistatic plasticity), and [5, 15] (Allen-Cahn and Cahn-Hilliard VIs). The optimality systems obtained thereby are in the best case of intermediate strength. This is not surprising at all, since one loses information when passing to the limit in the regularized/penalized problem. Roughly speaking, the strict complementarity assumption is the price one has to pay for not regularizing, and thus for obtaining a stronger optimality system as in the case of regularizing.

Remark 5.12 ('Ample controls') When it comes to the optimal control of non-smooth problems, strong stationary optimality systems have been mostly derived in the presence of 'ample controls', i.e., (distributed) controls that are not restricted by additional constraints. The literature here is rather scarce. We refer to [3, 14, 22] (elliptic VIs) and to [19] (non-smooth parabolic equations). Until recently, it was an open question whether such a system can be derived in the absence of 'ample controls', see also [22, Section 4]. It turns out that the necessity of strong stationarity can indeed be proven for the obstacle problem with pointwise constraints on the control. This was shown in [26], however by requiring that the (unknown) optimizer satisfies certain assumptions (constraint qualifications). There one obtains a strong stationary optimality system, which is a generalization of the optimality system derived by [22] in the more restrictive case of 'ample controls'.

6 Existence of solutions for (P_{min})

In this section we shortly address the existence of solutions for (P_{min}). To be more precise, we derive a setting so that the direct method of variational calculus can be applied for (P_{min}). To this end, let us recall that this can be written as

$$\min_{\ell \in \mathfrak{L}} f(\ell), \tag{6.1}$$

where $f : \ell \mapsto \mathcal{J}(\mathcal{S}(\ell), \ell)$ is the reduced objective functional. The first thing to observe is that the direct method of variational calculus cannot be applied to solve (6.1) without further ado. The control set $\mathfrak{L}$ is indeed a bounded subset of the reflexive Banach space $L^2(0, T; V^*)$, in light of $L^\infty(0, T; \boldsymbol{W}^{-1,p}(\Omega)) \hookrightarrow L^2(0, T; V^*)$. Since cf. Assumption 5.1, $\mathfrak{L}$ is convex, it would suffice to impose that it is also closed in $L^2(0, T; V^*)$, in order to obtain its weak compactness (in $L^2(0, T; V^*)$). However, $f : \mathfrak{L} \subset L^2(0, T; V^*) \to \mathbb{R}$ is not necessarily weakly lower semicontinuous. This would be the case if $\mathcal{S}$ were (at least) weakly continuous, which is not to be expected due to the structure of (3.5). Since by Lemma 3.5, $\mathcal{S}$ is Lipschitz continuous, it makes sense to require that the control set is a bounded subset of a reflexive Banach space which compactly embeds in $L^\infty(0, T; \boldsymbol{W}^{-1,p}(\Omega))$. An example for such a control set, which as we will see, satisfies all the conditions needed for showing existence of solutions, is

$$\mathfrak{L} := \{\ell \in H_0^1(0, T; L^p(\Omega; \mathbb{R}^N)) : \|\dot{\ell}\|_{L^2(0,T;L^p(\Omega;\mathbb{R}^N))} \le b\}, \tag{6.2}$$

where $b > 0$ is a given bound.

We now prove that the optimal control problem (P_{min}) admits solutions for $\mathfrak{L}$ given by (6.2) and for a continuous objective $\mathcal{J}$. We begin by noticing that $\mathfrak{L}$ is a nonempty, convex, closed and bounded subset of the reflexive Banach space $H^1(0,T;L^p(\Omega;\mathbb{R}^N))$, which in particular means that $\mathfrak{L}$ is weakly compact. Moreover, $f:\mathfrak{L}\to\mathbb{R}$ is weakly lower semicontinuous. To see this, consider a sequence $\{\ell_n\}\subset\mathfrak{L}$ with $\ell_n\rightharpoonup\ell$ in $H^1(0,T;L^p(\Omega;\mathbb{R}^N))$ as $n\to\infty$. Due to $\boldsymbol{W}_D^{1,p'}(\Omega)\hookrightarrow\hookrightarrow L^{p'}(\Omega;\mathbb{R}^N)$, we have the compact embedding $H^1(0,T;L^p(\Omega;\mathbb{R}^N))\hookrightarrow\hookrightarrow L^\infty(0,T;\boldsymbol{W}^{-1,p}(\Omega))$. With Lemma 3.5 we then infer

$$\mathcal{S}(\ell_n)\to\mathcal{S}(\ell)\quad\text{in } L^\infty(0,T;V)\times L^\infty(0,T;H^1(\Omega))\times W^{1,\infty}(0,T;L^2(\Omega))$$

as $n\to\infty$. Since $\mathcal{J}$ is continuous on $L^2(0,T;\boldsymbol{W}_D^{1,\nu}(\Omega))\times L^2(0,T;H^1(\Omega))\times L^2(0,T;L^2(\Omega))\times L^2(0,T;\boldsymbol{W}^{-1,p}(\Omega))$, the above convergence gives in turn the weak lower semicontinuity of f. Finally, a standard argument yields that (6.1), and thus (P_{min}), admits solutions. Clearly, the main result in Section 5 applies for this setting provided that $\mathcal{J}$ fulfills Assumption 5.2.

A Directional differentiability of max

Lemma A.1 *(i) The Nemytskii operator associated to* $\max:\mathbb{R}\to\mathbb{R},\ \max(\cdot):=\max\{\cdot,0\}$, *maps* $L^2(\Omega)$ *to* $L^2(\Omega)$. *Moreover,* $\max:L^2(\Omega)\to L^2(\Omega)$ *is Lipschitz continuous and directionally differentiable. For any* $y,h\in L^2(\Omega)$, *the derivative satisfies*

$$\max{}'(y;h)(x)=\begin{cases} h(x) & \text{if } y(x)>0\\ \max\{h(x),0\} & \text{if } y(x)=0\\ 0 & \text{if } y(x)<0\end{cases}\qquad \text{f.a.a. } x\in\Omega. \tag{A.1}$$

In addition, at any $y\in L^2(\Omega)$, *the operator* $\max{}'(y;\cdot):L^2(\Omega)\to L^2(\Omega)$ *is Lipschitz continuous (with Lipschitz constant* 1*).*
(ii) The Nemytskii operator associated to $\max:L^2(\Omega)\to L^2(\Omega)$ *maps* $L^\infty(0,T;L^2(\Omega))$ *to* $L^\infty(0,T;L^2(\Omega))$. *Moreover,* $\max:L^\infty(0,T;L^2(\Omega))\to L^\infty(0,T;L^2(\Omega))$ *is Lipschitz continuous. The operator* $\max:L^\infty(0,T;L^2(\Omega))\to L^\varrho(0,T;L^2(\Omega))$ *is directionally differentiable for any* $\varrho\in[1,\infty)$, *with* $\max{}'(\cdot;\cdot)\in L^\infty(0,T;L^2(\Omega))$. *For any* $y,h\in L^\infty(0,T;L^2(\Omega))$ *the derivative is given by*

$$\max{}'(y;h)(t,x)=\begin{cases} h(t,x) & \text{if } y(t,x)>0\\ \max\{h(t,x),0\} & \text{if } y(t,x)=0\\ 0 & \text{if } y(t,x)<0\end{cases}\qquad \text{f.a.a. } (t,x)\in(0,T)\times\Omega. \tag{A.2}$$

P r o o f. (i) The first thing to notice is that the Nemytskii operator $\max:L^2(\Omega)\to L^2(\Omega)$ is well defined and Lipschitz continuous, since $\max\{\cdot,0\}$ is Lipschitz continuous. Furthermore, straight forward computation shows that $\max:\mathbb{R}\to\mathbb{R}$ is directionally differentiable with

$$\max{}'(v;\delta v)=\begin{cases} \delta v & \text{if } v>0\\ \max\{\delta v,0\} & \text{if } v=0\\ 0 & \text{if } v<0\end{cases}\qquad \forall\, v,\delta v\in\mathbb{R}. \tag{A.3}$$

As a consequence thereof, we deduce that

$$|\max{}'(v;\delta v)| \leq |\delta v| \quad \forall\, v, \delta v \in \mathbb{R}. \tag{A.4}$$

Let now $y, h \in L^2(\Omega)$ be arbitrary, but fixed. The directional differentiability of $\max : \mathbb{R} \to \mathbb{R}$ yields

$$\Big|\frac{\max(y(x) + \tau h(x)) - \max(y(x))}{\tau} - \max{}'(y(x); h(x))\Big| \stackrel{\tau \searrow 0}{\longrightarrow} 0 \quad \text{f.a.a. } x \in \Omega.$$

On the other hand, the global Lipschitz continuity of $\max : \mathbb{R} \to \mathbb{R}$ with constant 1 together with (A.4) implies for $\tau \neq 0$ that

$$\Big|\frac{\max(y(x) + \tau h(x)) - \max(y(x))}{\tau} - \max{}'(y(x); h(x))\Big| \leq 2\,|h(x)| \;\; \text{f.a.a. } x \in \Omega.$$

Now, Lebesgue's dominated convergence theorem gives that $\max : L^2(\Omega) \to L^2(\Omega)$ is directionally differentiable with directional derivative given by (A.1), in view of (A.3). It now remains to show that the operator $\max{}'(y;\cdot) : L^2(\Omega) \to L^2(\Omega)$ is Lipschitz continuous (with Lipschitz constant 1) at any $y \in L^2(\Omega)$. This follows immediately from the definition of the directional derivative and the Lipschitz continuity of the operator $\max$, which imply that

$$\begin{aligned}\|\max{}'(y; h_1) - \max{}'(y; h_2)\|_2 &= \lim_{\tau \searrow 0} \Big\|\frac{\max(y + \tau h_1) - \max(y + \tau h_2)}{\tau}\Big\|_2 \\ &\leq \|h_1 - h_2\|_2 \quad \forall\, h_1, h_2 \in L^2(\Omega).\end{aligned}$$

(ii) The result follows by the exact same arguments. Notice that the norm gap is due to Lebesgue's dominated convergence theorem. □

References

[1] Viorel Barbu. Necessary conditions for distributed control problems governed by parabolic variational inequalities. *SIAM Journal on Control and Optimization*, 19(1):64–86, 1981.

[2] M. Bergounioux and H. Zidani. A fully discrete approximation for control problems governed by parabolic variational inequalities. *SIAM J. Numer. Anal.*, 39(6):2014–2033, 2002.

[3] J. C. De los Reyes and C. Meyer. Strong stationarity conditions for a class of optimization problems governed by variational inequalities of the second kind. *Journal of Optimization Theory and Applications*, 168(2):375–409, 2016.

[4] B.J. Dimitrijevic and K. Hackl. A method for gradient enhancement of continuum damage models. *Technische Mechanik*, 28(1):43–52, 2008.

[5] M.H. Farshbaf-Shaker. A penalty approach to optimal control of Allen-Cahn variational inequalities: MPEC-view. *Numerical Functional Analysis and Optimization*, 33(11):1321–1349, 2012.

[6] A. Friedman. Optimal control for parabolic variational inequalities. *SIAM Journal on Control and Optimization*, 25(2):482–497, 1987.

[7] H. Goldberg, W. Kampowsky, and F. Tröltzsch. On Nemytskij operators in L^p spaces of abstract functions. *Mathematische Nachrichten*, 155:127–140, 1992.

[8] P. Grisvard. *Elliptic Problems in non-smooth Domains*. Pitman, Boston, 1985.

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

[10] R. Haller-Dintelmann, C. Meyer, J. Rehberg, and A. Schiela. Hölder continuity and optimal control for non-smooth elliptic problems. *Applied Mathematics and Optimization*, 60:397–428, 2009.
[11] Z.X. He. State constrained control problems governed by variational inequalities. *SIAM Journal on Control and Optimization*, 25:1119–1144, 1987.
[12] R. Herzog and C. Meyer and G. Wachsmuth. Integrability of displacement and stresses in linear and nonlinear elasticity with mixed boundary conditions. *Journal of Mathematical Analysis and Applications*, 382(2):802-813, 2011.
[13] R. Herzog, C. Meyer, and G. Wachsmuth. C-stationarity for optimal control of static plasticity with linear kinematic hardening. *SIAM J. Optim.*, 50(5):3052–3082, 2012.
[14] R. Herzog, C. Meyer, and G. Wachsmuth. B- and strong stationarity for optimal control of static plasticity with hardening. *SIAM J. Optim.*, 23(1):321–352, 2013.
[15] M. Hintermüller and D. Wegner. Distributed optimal control of the Cahn-Hilliard system including the case of a double-obstacle homogeneous free energy density. *SIAM J. Control and Optimization*, 50(1):388–418, 2012.
[16] M. Hintermüller and D. Wegner. Optimal control of a semi-discrete Cahn-Hilliard-Navier-Stokes system. *SIAM J. Control and Optimization*, 52(1):747–772, 2014.
[17] K. Ito and K. Kunisch. Optimal control of parabolic variational inequalities. *Journal de Mathématiques Pures et Appliqués*, 93(4):329–360, 2010.
[18] C. Meyer and L.M. Susu. Analysis of a viscous two-field gradient damage model. Part I: Existence and uniqueness. Technical Report 526, Ergebnisberichte des Instituts für Angewandte Mathematik, TU Dortmund, 2015.
[19] C. Meyer and L.M. Susu. Optimal control of non-smooth, semilinear parabolic equations. *SIAM Journal on Control and Optimization*, 55(4):2206–2234, 2017.
[20] C. Meyer and L.M. Susu. Analysis of a viscous two-field gradient damage model. Part II: Penalization limit. Technical Report 538, Ergebnisberichte des Instituts für Angewandte Mathematik, TU Dortmund, 2016.
[21] F. Mignot and J.-P. Puel. Contrôle optimal d'un système gouverné par une inéquation variationelle parabolique. *Comptes Rendus de l'Académie des Sciences Paris*, 298(12):277–280, 1984.
[22] F. Mignot and J.-P. Puel. Optimal control in some variational inequalities. *SIAM Journal on Control and Optimization*, 22(3):466–476, 1984.
[23] H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. *Mathematics of Operations Research*, 25(1):1–22, 2000.
[24] W. Schirotzek. *Non-Smooth Analysis*. Springer, Berlin, 2007.
[25] F. Tröltzsch. *Optimal Control of Partial Differential Equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 2010. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.
[26] G. Wachsmuth. Strong stationarity for optimal control of the obstacle problem with control constraints. *SIAM Journal on Optimization*, 24(4):1914–1932, 2014.
[27] G. Wachsmuth. Optimal control of quasistatic plasticity with linear kinematic hardening II: Regularization and differentiability. *Zeitschrift für Analysis und ihre Anwendungen*, 34(4):391–418, 2015.