ruhr.paD
UA Ruhr Zentrum für
partielle Differentialgleichungen

Second-order sufficient optimality conditions for optimal control of non-smooth, semilinear parabolic equations

L. M. Betz (Susu)

Preprint 2018-11

SECOND-ORDER SUFFICIENT OPTIMALITY CONDITIONS FOR OPTIMAL CONTROL OF NON-SMOOTH, SEMILINEAR PARABOLIC EQUATIONS

LIVIA BETZ*

Abstract. This paper is concerned with an optimal control problem governed by a non-smooth, semilinear parabolic PDE. The nonlinearity in the state equation is only directionally differentiable, locally Lipschitz continuous, and is allowed to have infinitely many non-differentiable points. By employing its limited properties, Bouligand-differentiability of the control-to-state map is shown (in an extended sense). This enables us to establish second-order sufficient optimality conditions. We provide concrete settings where these reduce to the first-order necessary optimality condition.

Key words. Optimal control of PDEs, non-smooth optimization, second-order sufficient conditions, Bouligand differentiability, two-norm discrepancy

AMS subject classifications. 49J20, 35K58, 49K99

1. Introduction. In this paper we establish second-order sufficient conditions for the following optimal control problem:

$$\left.\begin{aligned} \min_{u\in L^r(0,T;L^2(\Omega))} \quad & J(y,u) \\ \text{s.t.} \quad & \dot y(t) + A\,y(t) + f(y(t)) = B\,u(t) \quad \text{a.e. in } (0,T) \\ & y(0) = 0, \end{aligned}\right\} \tag{P}$$

where $\Omega \subset \mathbb{R}^n$, $n = 2,3$, is a bounded Lipschitz domain, J is a smooth function, A is a linear unbounded operator and f is a non-smooth mapping. The precise statements will be given at the end of this section and in Assumption 2.1 below. The essential feature of (P) is that the nonlinearity f appearing in the state equation is only directionally differentiable. Thus, the second-order analysis cannot be performed by classical techniques for smooth optimization problems in Banach spaces. Moreover, we deal with a non-Hilbert space for the control, which gives rise to additional challenges (such as the two-norm discrepancy).

Optimal control problems subject to non-smooth constraints are challenging even in the finite dimensional case, see e.g. [30] and the references therein. Difficulties arise from the non-smoothness of the control-to-state mapping, which does not allow to apply the standard Karush-Kuhn-Tucker (KKT) theory. For this reason, various optimality conditions of different strength have been introduced, such as e.g. Clarke (C), Bouligand (B), and strong stationarity. In the spirit thereof, stationarity concepts for the infinite dimensional case are defined in [19]. The most rigorous stationarity concept is strong stationarity. In a previous work [24], necessary optimality conditions of this type were established for (P), from which we will benefit in the present paper.

While second-order sufficient optimality conditions (SSC) for the optimal control of smooth PDEs have been intensively investigated, see e.g. [4–8, 11, 15, 27, 29] and the references therein, the literature on SSC for the optimal control of *non-smooth problems* is rather rare. To the best of our knowledge, the only contributions in this

*Faculty of Mathematics, University of Duisburg-Essen, Germany (former last name: Susu).

field deal with elliptic VIs. These were addressed in [22] (obstacle problem) and [3] (static elastoplasticity). In [25] it was proven that the obstacle control problem is convex if the desired state is behind the obstacle. This result was extended in [22], where sufficient conditions for the optimal control of the obstacle problem in the general case were presented. To the best of the author's knowledge, the investigation of second-order sufficient conditions for optimization problems governed by non-smooth parabolic PDEs is an open research topic.

What distinguishes the problem (P) from the ones analyzed in [3] and [22] is not only the parabolic component but also the very general non-smooth mapping f. For example, we allow the set of non-differentiable points of f to be at most countable. In our second-order analysis, we require (in addition to strong stationarity, positive-definiteness/coercivity of the Hessian of the "Lagrangian") only a sign condition on the adjoint state. This is standard when it comes to deriving SSC for *infinite dimensional* control problems *which feature non-smooth solution operators*, as it ensures a so-called "safety distance", cf. [3, Rem. 4.13]. A corresponding assumption regarding the sign of the adjoint state is required in [22] and [3] as well. However, our SSC are comparatively sharper, since in [22] an additional sign condition on the multiplier is imposed, while in [3] additional regularity assumptions on the adjoint state and multipliers are made (besides strong stationarity, positive-definiteness/coercivity of the Hessian of the "Lagrangian").

Letting the non-smooth nonlinearity aside, a further challenge in this paper arises from the fact that we work with a non-Hilbert space for the control (which we motivate in Remark 2.6 below). For this reason we need to consider the two-norm discrepancy. In order to be able to carry out the second order analysis in this framework, it is crucial to show an improved Bouligand-differentiability result for the solution operator of the state equation (Theorem 3.4 below). In this context, we will exploit semi group theory arguments and we shall rely on the *global* Lipschitz continuity of $f'(;)$ w.r.t. its direction, cf. Remark 3.3 below.

The paper is organized as follows. In Section 2 we state the precise assumptions on the data and lay the foundations for our analysis, by recalling some useful results from [24]. The findings in Section 3 play a fundamental role, as they lead to an essential Bouligand-differentiability result for the control-to-state map. This in turn constitutes the basis for the investigation of SSC, which will be performed in Section 4. Here we present our main results, namely two sets of second-order sufficient conditions for the optimal control of (P) (Theorems 4.13 and 4.16 below). In Section 5 we introduce a setting where the necessary optimality condition alone is sufficient for optimality.

Notation. Throughout the paper, C and c denote generic positive constants. If X and Y are two linear normed spaces, the space of linear and bounded operators from X to Y is denoted by $\mathcal{L}(X,Y)$. For the open ball in X around $x \in X$ with radius R we write $B_X(x,R)$. The symbol X^* stands for the dual space of X, while $\langle .,. \rangle_X$ stands for the dual pairing between X and 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. If X and Y are Banach spaces, we use the notation $[X,Y]_\theta$ for the complex and $(X,Y)_{\theta,\omega}$ for the real interpolation space, respectively, where $\theta \in (0,1)$ and $\omega \in [1,\infty]$, see e.g. [31]. If a linear operator A is the infinitesimal generator of a semigroup, the latter will be denoted by $\{e^{tA}\}_{t\geq 0}$, see also [26, Chp. 2.5]. In all what follows, $T > 0$ is a fixed final time and $\Omega \subset \mathbb{R}^n$, $n = 2,3$, is a bounded Lipschitz domain in the sense of [23, Chp.

1.1.9]. For simplicity, we abbreviate $Q := (0,T) \times \Omega$. The boundary of Ω consists of two disjoint measurable parts Γ_D and Γ_N. By $W^{1,q}_D(\Omega)$ we denote the closure of the set $\{\psi|_\Omega : \psi \in C_0^\infty(\mathbb{R}^n),\ \operatorname{supp}(\psi) \cap \Gamma_D = \emptyset\}$ with respect to the $W^{1,q}$-norm, where $q \in (1,\infty)$. For the dual space associated with $W^{1,q'}_D(\Omega)$ we use the symbol $W^{-1,q}(\Omega)$, where q' stands for the conjugate exponent. Moreover, we abbreviate

$$\mathbb{W}_0^r(W^{1,q}_D(\Omega), W^{-1,q}(\Omega)) := \{v \in W^{1,r}(0,T; W^{-1,q}(\Omega)) \cap L^r(0,T; W^{1,q}_D(\Omega)) : v(0) = 0\},$$
$$\mathbb{W}_T^{r'}(W^{1,q'}_D(\Omega), W^{-1,q'}(\Omega)) := \{v \in W^{1,r'}(0,T; W^{-1,q'}(\Omega)) \cap L^{r'}(0,T; W^{1,q'}_D(\Omega)) : v(T) = 0\},$$

where $r \in (1,\infty)$.

2. Standing assumptions and known results. This section is devoted to collecting the assumptions on the data as well as crucial results from [24] concerning the state equation.

ASSUMPTION 2.1. *For the quantities in* (P) *we require the following:*

1. *Let*
$$n < q \le 2n/(n-2) \quad \text{and} \quad \frac{2q}{q-n} < r < \infty \tag{2.1}$$
be fixed.
2. *The operator $A : W^{-1,q}(\Omega) \to W^{-1,q}(\Omega)$ is linear, unbounded, and closed. Its domain of definition is given by $W^{1,q}_D(\Omega)$. In addition, $0 \notin \sigma(A)$, where $\sigma(A)$ denotes the spectrum of A.*
3. *Moreover, A satisfies maximal parabolic $L^r(0,T; W^{-1,q}(\Omega))$-regularity, i.e., for every $g \in L^r(0,T; W^{-1,q}(\Omega))$, the equation $\dot w + Aw = g$ admits a unique solution $w \in \mathbb{W}_0^r(W^{1,q}_D(\Omega), W^{-1,q}(\Omega))$.*
4. *The nonlinearity $f : \mathbb{R} \to \mathbb{R}$ is assumed to be monotone increasing. Moreover, it is Lipschitz continuous on bounded sets, i.e., for all $M > 0$, there exists a constant $L_M > 0$ such that*
$$|f(z_1) - f(z_2)| \le L_M |z_1 - z_2| \quad \forall\, z_1, z_2 \in [-M, M].$$
5. *The function f is directionally differentiable at every point, i.e.,*
$$\Big|\frac{f(x + \tau h) - f(x)}{\tau} - f'(x;h)\Big| \xrightarrow{\tau \searrow 0} 0 \quad \forall x, h \in \mathbb{R}.$$
Moreover, the set of non-differentiable points of f is at most countable.
6. *The operator $B : L^2(\Omega) \to (W^{-1,q}(\Omega), W^{1,q}_D(\Omega))_{\varsigma,\infty}$ is linear and bounded, where $\varsigma \in (0, 1/2)$ is fixed.*
7. *The objective $J : L^2(0,T; W^{1,q}_D(\Omega)) \times L^2(Q) \to \mathbb{R}$ is twice continuously Fréchet-differentiable.*

REMARK 2.2. *With a little abuse of notation, the Nemytskii-operators associated with f and B, considered with different ranges, will be denoted by the same symbol.*

Comments regarding the above assumptions are provided at the end of this section. From now on, Assumption 2.1 is tacitly assumed in the following without mentioning it every time. We point out that this fits in the general setting of [24, Assumptions 2.1 and 2.5], as explained in the following. With the notations from the preceding contribution [24] we have

$$X = W^{-1,q}(\Omega), \quad \mathcal{D} = W^{1,q}_D(\Omega), \quad U = L^2(\Omega), \quad Y = L^\infty(\Omega).$$

Note that the maximal parabolic regularity assumption on A implies that A generates an analytic semigroup, see [2, Section 3]. The Nemytskii operator $f : L^\infty(\Omega) \to L^\infty(\Omega)$ is well-defined and Lipschitz continuous on bounded sets, i.e., for every $M > 0$, there exists $L_M > 0$ so that

$$\|f(y_1) - f(y_2)\|_{L^\infty(\Omega)} \leq L_M \, \|y_1 - y_2\|_{L^\infty(\Omega)} \quad \forall \, y_1, y_2 \in \overline{B_{L^\infty(\Omega)}(0, M)}. \tag{2.2}$$

When considered with domain $L^\infty(\Omega)$ and range $L^\beta(\Omega)$, $\beta < \infty$, f is directionally differentiable, see proof of [24, Lemma 6.4] for details. Therein, Assumption 2.1.1 on r and q is justified as well. The condition $q > n$ guarantees that there exists $\theta \in (0, 1)$ such that $(W^{-1,q}(\Omega), W_D^{1,q}(\Omega))_{\theta,\infty} \hookrightarrow L^\infty(\Omega)$, see [24, (6.4)], while the relation between r and q in (2.1) ensures that $r(1-\theta) > 1$, i.e., [24, (2.4)]. These turn out to be essential not only for the existence of solutions of the state equation, see [24], but also for the upcoming second-order analysis, see Remark 2.6 below. Note that, in view of [31, Thm. 1.15.2 (d), p.101] and [26, Thm. 2.6.13, p.74], it holds

$$\|e^{-tA}\|_{\mathcal{L}(W^{-1,q}(\Omega), L^\infty(\Omega))} \leq c \, t^{-\theta} \quad \forall \, t \in (0, T], \tag{2.3}$$

which will be crucial in the proof of Theorem 3.4 below. Let us mention that we dropped the density assumption on B and the convexity assumption on J, as they were needed in [24] just for deriving (strong stationary) necessary optimality conditions and for proving the existence of global minimizers, respectively, which is not the case in this paper.

The embedding

$$\mathbb{W}_0^r(W_D^{1,q}(\Omega), W^{-1,q}(\Omega)) \hookrightarrow\hookrightarrow C([0,T]; C(\bar{\Omega})) = C(\bar{Q}) \tag{2.4}$$

will be crucial in the next sections and is a consequence of [1, Thm. 3] combined with $W_D^{1,q}(\Omega) \hookrightarrow\hookrightarrow W^{-1,q}(\Omega)$, $(W^{-1,q}(\Omega), W_D^{1,q}(\Omega))_{\hat{\theta},1} \hookrightarrow [W^{-1,q}(\Omega), W_D^{1,q}(\Omega)]_{\hat{\theta}} \hookrightarrow C(\bar{\Omega})$, see [1, Section 3], [16, Thm. 3.5], and [31, Thm. 4.6.1.e)]. Note that the last embedding is true if $\hat{\theta} \in (0, 1)$ is chosen such that $r(1-\hat{\theta}) > 1$ and $2\hat{\theta} - 1 > n/q$, which is possible in view of $1/r < (q-n)/2q$, cf. (2.1).

Since the setting in Assumption 2.1 is just a special case of [24, Assumption 2.1], we can apply the general results in [24, Sections 2-3] on our state equation. We begin by introducing the control-to-state mapping.

Definition 2.3. *The solution operator of*

$$\begin{aligned} \dot{y}(t) + A\, y(t) + f(y(t)) &= B\, u(t) \quad \text{a.e. in } (0,T), \\ y(0) &= 0 \end{aligned} \tag{2.5}$$

is denoted by $S : L^r(0,T; L^2(\Omega)) \ni u \mapsto y \in \mathbb{W}_0^r(W_D^{1,q}(\Omega), W^{-1,q}(\Omega))$. Note that, in view of [24, Proposition 2.11], this is well-defined.

Proposition 2.4. *[24, Proposition 2.11] The control-to-state mapping S is Lipschitz continuous on bounded sets, i.e., for every $R > 0$, there exists a constant $L_R > 0$ such that, for all $u_1, u_2 \in \overline{B_{L^r(0,T;L^2(\Omega))}(0, R)}$, it holds*

$$\|S(u_1) - S(u_2)\|_{\mathbb{W}_0^r(W_D^{1,q}(\Omega), W^{-1,q}(\Omega))} \leq L_R \|u_1 - u_2\|_{L^r(0,T;L^2(\Omega))}. \tag{2.6}$$

Theorem 2.5. *[24, Lemma 3.3, Theorem 3.4] The solution operator $S : L^r(0,T; L^2(\Omega)) \to \mathbb{W}_0^r(W_D^{1,q}(\Omega), W^{-1,q}(\Omega))$ is directionally differentiable and its directional derivative*

$\eta = S'(u;h)$ at $u \in L^r(0,T;L^2(\Omega))$ in direction $h \in L^r(0,T;L^2(\Omega))$ is given by the unique solution of

$$\begin{aligned} \dot\eta(t) + A\eta(t) + f'(y(t);\eta(t)) &= B\,h(t) \quad \text{a.e. in } (0,T) \\ \eta(0) &= 0, \end{aligned} \tag{2.7}$$

with $y = S(u)$. The solution operator of (2.7), namely $S'(u;\cdot) : L^r(0,T;L^2(\Omega)) \ni h \mapsto \eta \in \mathbb{W}_0^r(W_D^{1,q}(\Omega), W^{-1,q}(\Omega))$ is globally Lipschitz continuous.

Some remarks concerning Assumption 2.1 are in order:

Remark 2.6. *Note that Assumption 2.1.1 does not allow us to consider the $L^2(Q)$-Hilbert space for the control, since $r > 2$ even in two dimensions. We deliberately choose to work with such a setting, although additional assumptions on the nonlinearity f would enable us to set $r = 2$, cf. [24, Remark 6.5]. We proceed in this way due to the following reason. The condition (2.1) guarantees that the control-to-state mapping is (locally) Lipschitz continuous with range in $C(\bar Q)$ (see (2.6) and (2.4)). This will be indispensable for the derivation of SSC: given a point y in a neighborhood of the state $\bar y$, we have to be able to make assertions about the distance between $y(t,x)$ and $\bar y(t,x)$, see proofs of Lemmas 4.8 and 4.10 below. Let us point out that if $r = 2$, then one cannot expect $L^\infty(Q)$-regularity for the state, see [32, Chp. 5]. We also emphasize that if we weaken the assumption on r as suggested by [24, Remark 6.5], see also Lemma A.1, (i.e., if we require that $\frac{2q}{q-n+\varsigma 2q} < r < \infty$, where ς is given by Assumption 2.1.6), then the crucial embedding (2.4) is no longer true.*

In the elliptic case, the local Lipschitz continuity of the control-to-state operator with range in the space of essentially bounded functions is crucial too. Thanks to the Stampacchia method, see [32], one can choose $L^2(\Omega)$ as space for the control to guarantee this. Let us emphasize that the entire second-order analysis for the elliptic version of (P) can be performed in the same way as in the parabolic setting.

Remark 2.7. *In two dimensions, Assumption 2.1.2 is satisfied by the operator $A = -\operatorname{div}\kappa\nabla$ defined as*

$$A : W_D^{1,q}(\Omega) \ni y \mapsto \int_\Omega \kappa\nabla y \nabla\cdot\, dx \in W^{-1,q}(\Omega),$$

if $\Omega \cup \Gamma_N$ is regular in the sense of Gröger, cf. [17], and the coefficient function $\kappa \in L^\infty(\Omega;\mathbb{R}^{n\times n})$ is uniformly elliptic and symmetric. The papers [21, Appendix] and [12] provide many settings such that $-\operatorname{div}\kappa\nabla$ fulfills Assumption 2.1.2 in three dimensions too, e.g., if $\Gamma_N = \emptyset$, κ is uniformly continuous and may jump across a C^1-interface, and Ω is a strong Lipschitz domain in the sense of [23, Chp. 1.1.9]. For more details, see [24, Remark 6.3]. In the more recent contribution [13] very mild conditions on Ω are stated such that Assumption 2.1.3 is guaranteed in both dimensions for $A = -\operatorname{div}\kappa\nabla$ and κ as above. We refer to [13, Thm. 4.6(c)], where rough settings are allowed, e.g. domains which are not Lipschitz and where Ω is not required to lie on the same side of Γ_D. Note further that A satisfies maximal parabolic $L^s(0,T;W^{-1,q}(\Omega))$-regularity for every $s \in (1,\infty)$, cf. [14].

Remark 2.8. *The monotonicity property of f in Assumption 2.1.4 can be replaced by the more general [24, Assumption 2.5], see the proof of [24, Lemma 6.6]. Alternatively, one could require that f satisfies certain growth conditions, cf. [24, Remark 2.6].*

Remark 2.9. *Semilinear parabolic PDEs with non-smooth nonlinearities f of the type (2.5) arise for instance in the modeling of combustion processes, see e.g. [34]. In this* *case, the combustion nonlinearity f features a so-called ignition-temperature $\Theta > 0$. For example, f could be identically zero on the interval $(-\infty, \Theta]$, i.e., there is no reaction below the ignition-temperature. Once Θ is reached, ignition and combustion suddenly occur (so that Θ is a kink point of f). This leads to an abrupt change in the physical regime modeled by the non-smooth function f. For more details, we refer to [34, Chp. 1.2]. In Figure 5.1 below we depict an ignition-type nonlinearity with three ignition temperatures.*

3. Bouligand-differentiability of the control-to-state operator. This section addresses an essential property of the operator S that will be needed in the proof of the main result, namely its Bouligand-differentiability (in an extended sense). To show this, we need the following

PROPOSITION 3.1. *Let $\{\xi_n\}$ be a uniformly bounded sequence in $C(\bar{Q})$ with $\xi_n(t,x) \to 0$ a.e. in Q as $n \to \infty$. Then*

$$\frac{\|f(y+\xi_n) - f(y) - f'(y;\xi_n)\|_{L^\beta(0,T;L^\varrho(\Omega))}}{\|\xi_n\|_{L^{\hat{\beta}}(0,T;L^{\hat{\varrho}}(\Omega))}} \to 0 \quad \forall\, y \in C(\bar{Q})$$

for all $1 \leq \beta < \hat{\beta} \leq \infty$ and $1 \leq \varrho < \hat{\varrho} \leq \infty$. In particular, f is Bouligand- and thus, directionally differentiable from $C(\bar{Q})$ to $L^\beta(Q)$ for every $1 \leq \beta < \infty$. Furthermore, $f'(y;\xi) \in L^\infty(Q)\ \forall\, y, \xi \in C(\bar{Q})$.

Proof. From Assumptions 2.1.4 and 2.1.5 it follows that $f : \mathbb{R} \to \mathbb{R}$ is Bouligand-differentiable at any $z \in \mathbb{R}$, i.e.,

$$\frac{|f(z+v_n) - f(z) - f'(z;v_n)|}{|v_n|} \to 0 \quad \text{ as } v_n \to 0. \tag{3.1}$$

Note that, for all $M > 0$, it holds

$$|f'(z;v)| \leq L_{M+1}\,|v| \quad \forall\, v \in \mathbb{R},\, z \in [-M, M], \tag{3.2}$$

where $L_{M+1} > 0$ is given by Assumption 2.1.4. This is due to the definition of the directional derivative and its positive homogeneity w.r.t. direction, see also the proof of [24, Lemma 3.1].

Let now $\{\xi_n\} \subset C(\bar{Q})$ be a uniformly bounded sequence with $\xi_n(t,x) \to 0$ a.e. in Q. As a result of (3.1), we have the following convergence

$$g_n(t,x) := \frac{|f(y(t,x)+\xi_n(t,x)) - f(y(t,x)) - f'(y(t,x);\xi_n(t,x))|}{|\xi_n(t,x)|} \to 0 \quad \text{a.e. in } Q$$

as $n \to \infty$. Moreover, due to (3.2), it holds

$$|f'(y(t,x);\xi_n(t,x))| \leq L_{\|y\|_{C(\bar{Q})}+1}\,|\xi_n(t,x)| \quad \forall\,(t,x) \in Q.$$

By assumption, there exists $c > 0$ such that $\|\xi_n\|_{C(\bar{Q})} \leq c$ for all n and by employing the estimate in Assumption 2.1.4 we have

$$|f(y(t,x)+\xi_n(t,x)) - f(y(t,x))| \leq L_{\|y\|_{C(\bar{Q})}+c}\,|\xi_n(t,x)| \quad \forall\,(t,x) \in Q.$$

Thus, by Lebesgue's dominated convergence theorem, we obtain

$$g_n \to 0 \quad \text{in } L^{\beta'}(0,T;L^{\varrho'}(\Omega)) \text{ as } n \to \infty,$$

where we abbreviate $\beta' = \beta\hat{\beta}/(\hat{\beta} - \beta) \in [1, \infty)$ and $\varrho' = \varrho\hat{\varrho}/(\hat{\varrho} - \varrho) \in [1, \infty)$. In view of the definition of g_n, Hölder's inequality, and the above convergence we finally obtain

$$
\begin{aligned}
\frac{\|f(y+\xi_n) - f(y) - f'(y;\xi_n)\|_{L^\beta(0,T;L^\varrho(\Omega))}}{\|\xi_n\|_{L^{\hat{\beta}}(0,T;L^{\hat{\varrho}}(\Omega))}} &= \frac{\|g_n \xi_n\|_{L^\beta(0,T;L^\varrho(\Omega))}}{\|\xi_n\|_{L^{\hat{\beta}}(0,T;L^{\hat{\varrho}}(\Omega))}} \\
&\leq \frac{\|g_n\|_{L^{\beta'}(0,T;L^{\varrho'}(\Omega))} \|\xi_n\|_{L^{\hat{\beta}}(0,T;L^{\hat{\varrho}}(\Omega))}}{\|\xi_n\|_{L^{\hat{\beta}}(0,T;L^{\hat{\varrho}}(\Omega))}} \\
&\to 0 \quad \text{as } n \to \infty.
\end{aligned}
$$

Note that $f'(y;\xi) \in L^\infty(Q)$ for all $y, \xi \in C(\bar{Q})$, in light of (3.2). $\square$

Similarly to (3.2), we observe that, for all $M > 0$, it holds

$$
|f'(z;v_1) - f'(z;v_2)| \leq L_{M+1}\, |v_1 - v_2| \quad \forall\, v_1, v_2 \in \mathbb{R},\, z \in [-M, M], \tag{3.3}
$$

where $L_{M+1} > 0$ is given by Assumption 2.1.4. This is due to the estimate in Assumption 2.1.4 and the positive homogeneity of the directional derivative w.r.t. direction, see also the proof of [24, Lemma 3.1]. By means of (3.3), we next prove that the directional derivative $h \mapsto S'(u;h)$ can be extended on the Hilbert-space $L^2(Q)$ for any $u \in L^r(0,T;L^2(\Omega))$. We also obtain a crucial estimate, namely (3.5) below, which will be the key ingredient for showing the extended Bouligand-differentiability of S.

Lemma 3.2. *Let $u \in L^r(0,T;L^2(\Omega))$ be given and $y := S(u)$. For any right-hand side $h \in L^2(Q)$, there exists a unique solution $\eta \in \mathbb{W}_0^2(W_D^{1,q}(\Omega), W^{-1,q}(\Omega))$ of*

$$
\begin{aligned}
\dot{\eta}(t) + A\eta(t) + f'(y(t);\eta(t)) &= B\,h(t) \quad \text{a.e. in } (0,T) \\
\eta(0) &= 0.
\end{aligned}
\tag{3.4}
$$

The resulting solution operator of (2.7) is denoted by $S^u : L^2(Q) \ni h \mapsto \eta \in \mathbb{W}_0^2(W_D^{1,q}(\Omega), W^{-1,q}(\Omega))$. There exists a constant $c = c(u) > 0$ such that

$$
\|S^u h\|_{\mathbb{W}_0^2(W_D^{1,q}(\Omega), W^{-1,q}(\Omega))} \leq c\, \|h\|_{L^2(Q)} \quad \forall\, h \in L^2(Q). \tag{3.5}
$$

Moreover, the operator $S^u : L^2(Q) \ni h \mapsto \eta \in \mathbb{W}_0^2(W_D^{1,q}(\Omega), W^{-1,q}(\Omega))$ is weakly continuous.

Proof. Let $h \in L^2(Q)$ be arbitrary, but fixed. We first address the integral equation associated to (3.4):

$$
\eta(t) = \int_0^t e^{-(t-s)A}\big(Bh(s) - f'(y(s);\eta(s))\big)\, ds \quad \forall\, t \in [0,T]. \tag{3.6}
$$

As in [24, Lem. 3.3], the unique solvability of (3.6) follows by Banach's contraction principle. We underline only the key aspects which have to be considered here, since most of the arguments are rather standard. We first see that the term on the right-hand side in (3.6) maps $\eta \in C([0,T];L^2(\Omega))$ to $C([0,T];L^2(\Omega))$. This is a result of

(A.2), (3.2) together with $y \in C(\bar{Q})$, Assumption 2.1.6, and (A.3), which imply

$$
\begin{aligned}
&\int_0^t \|e^{-(t-s)A}\big(Bh(s) - f'(y(s);\eta(s))\big)\|_2 \, ds \\
&\leq \Big(\int_0^T \|e^{-sA}\|^2_{\mathcal{L}((W^{-1,q}(\Omega),W^{1,q}_D(\Omega))_{\varsigma,\infty},L^2(\Omega))} \, ds\Big)^{1/2} \|Bh\|_{L^2(0,T;(W^{-1,q}(\Omega),W^{1,q}_D(\Omega))_{\varsigma,\infty})} \\
&\quad + \int_0^t \|e^{-(t-s)A}\|_{\mathcal{L}(L^2(\Omega),L^2(\Omega))} \|\eta(s)\|_{L^2(\Omega)} \, ds \\
&\leq c\|h\|_{L^2(Q)} + \int_0^t (t-s)^{-\gamma} \|\eta(s)\|_{L^2(\Omega)} \, ds \quad \forall\, t \in [0,T],
\end{aligned}
\tag{3.7}
$$

where $\gamma \in (1/2, 1)$. The continuity in time follows by standard arguments of semigroup theory, cf. [26, Thm. 2.6.8(d) and Thm. 2.6.13(d)]. Further, the term on the right-hand side in (3.6) is Lipschitz continuous w.r.t. $\eta \in C([0,T];L^2(\Omega))$, since we can estimate as in (3.7), where this time we use (3.3). This yields the contractivity of the fixed point mapping on small time intervals. Then, a concatenation argument yields the global (in time) existence. Thus, (3.6) admits a unique solution $\eta \in C([0,T];L^2(\Omega))$. In view of (3.7), we can apply a generalized Gronwall's inequality, cf. [18, Lemma 7.1.1, p. 188], which gives in turn

$$
\|\eta\|_{C([0,T];L^2(\Omega))} \leq c\|h\|_{L^2(Q)}. \tag{3.8}
$$

We now recall that A satisfies maximal parabolic $L^2(0,T;W^{-1,q}(\Omega))$-regularity, according to Assumption 2.1.3 and [14], see also Remark 2.7. A boot strapping technique similar to [24, Prop. 2.11] finally yields that $\eta \in \mathbb{W}^2_0(W^{1,q}_D(\Omega), W^{-1,q}(\Omega))$. Note that $f'(y;\eta) \in L^2(Q) \hookrightarrow L^2(0,T;W^{-1,q}(\Omega))$, by (3.2) and the assumption on q, see (2.1).

In the light of $(\partial_t + A)^{-1} \in \mathcal{L}(L^2(0,T;W^{-1,q}(\Omega)); \mathbb{W}^2_0(W^{1,q}_D(\Omega), W^{-1,q}(\Omega)))$, (3.2), (3.8), and the boundedness of B, we conclude (3.5) with c depending on $\|y\|_{C(\bar{Q})}$, and thus on u.

The weak continuity of S^u then follows by the exact same arguments as in [24, Lem. 2.13], by relying on the compact embedding

$$
\mathbb{W}^2_0(W^{1,q}_D(\Omega), W^{-1,q}(\Omega)) \hookrightarrow\hookrightarrow L^\vartheta(0,T;L^q(\Omega)) \quad \forall\, \vartheta \in [1,\infty). \tag{3.9}
$$

This is a consequence of [1, Thm. 3] combined with the embeddings $W^{1,q}_D(\Omega) \hookrightarrow\hookrightarrow W^{-1,q}(\Omega)$, $(W^{-1,q}(\Omega), W^{1,q}_D(\Omega))_{\frac{1}{2},1} \hookrightarrow [W^{-1,q}(\Omega), W^{1,q}_D(\Omega)]_{\frac{1}{2}} = L^q(\Omega)$, see [1, Section 3] and [16, Thm. 3.5]. □

Remark 3.3. *Let us underline that the global Lipschitz continuity of the nonlinearity in* (3.4) *(along with the resulting global growth condition* (3.2)*) are the key tools in the proof of Lemma 3.2. These allow us to work with the global Lipschitz continuous Nemytskii operator $f'(y;\cdot)$ mapping from $L^2(\Omega)$ to $L^2(\Omega)$. If the nonlinearity does not satisfy some sort of global growth condition (e.g. if it is only Lipschitz continuous on bounded sets), then the associated Nemytskii operator is well-defined only on $L^\infty(\Omega)$ (such as in our state equation* (2.5)*). In this case the fixed point function should map between $C([0,T];L^\infty(\Omega))$ and $C([0,T];L^\infty(\Omega))$. An inspection of the proof of Lemma A.1 then shows that the $L^\infty(\Omega)$-norm of the integral in* (3.6) *is not finite for*

$h \in L^2(Q)$, *but only for* $h \in L^s(0,T;L^2(\Omega))$ *with some* $s > 2$, *see [24, Rem. 6.5].*

THEOREM 3.4. *Let* $u \in L^r(0,T;L^2(\Omega))$ *be fixed and* $h \to 0$ *in* $L^r(0,T;L^2(\Omega))$. *Then*

$$\frac{\|S(u+h) - S(u) - S'(u;h)\|_{\mathbb{W}_0^r(W_D^{1,q}(\Omega),W^{-1,q}(\Omega))}}{\|h\|_{L^2(Q)}} \to 0.$$

Thus, the control-to-state mapping $S : L^r(0,T;L^2(\Omega)) \to \mathbb{W}_0^r(W_D^{1,q}(\Omega),W^{-1,q}(\Omega))$ *is Bouligand-differentiable.*

Proof. Let $u, h \in L^r(0,T;L^2(\Omega))$ be arbitrary, but fixed and set $y := S(u)$, $y^h := S(u+h)$ and $\eta := S'(u;h)$. By subtracting (2.5) and (2.7) from (2.5) with right-hand side $u+h$ we have

$$\begin{gathered}\frac{d}{dt}(y^h - y - \eta) + A(y^h - y - \eta) = -f(y^h) + f(y) + f'(y;\eta),\\ (y^h - y - \eta)(0) = 0.\end{gathered} \tag{3.10}$$

The associated integral equation reads

$$(y^h - y - \eta)(t) = \int_0^t e^{-(t-s)A}\big(-f(y^h(s)) + f(y(s)) + f'(y(s);\eta(s))\big)\,ds,$$

see e.g. [18]. Consequently, one obtains

$$\begin{aligned}&\|(y^h - y - \eta)(t)\|_{L^\infty(\Omega)}\\ &\quad\le \int_0^t \|e^{-(t-s)A}\|_{\mathcal{L}(W^{-1,q}(\Omega),L^\infty(\Omega))}\Big(\overbrace{\|f\big(y^h(s)\big) - f(y(s)+\eta(s))\|_{W^{-1,q}(\Omega)}}^{=:\hat{A}_h(s)}\\ &\qquad\qquad + \underbrace{\|f(y(s)+\eta(s)) - f(y(s)) - f'(y(s);\eta(s))\|_{W^{-1,q}(\Omega)}}_{=:\hat{B}_h(s)}\Big)ds\end{aligned} \tag{3.11}$$

for all $t \in [0,T]$. We assume that $\|h\|_{L^r(0,T;L^2(\Omega))} \le 1$, since $h \to 0$ later anyway. Proceeding exactly as in the proof of [24, Lemma 6.6], we find

$$\begin{aligned}\|y^h\|_{C(\bar{Q})} &\le C(1 + \|u+h\|_{L^r(0,T;L^2(\Omega))}) \le C(2 + \|u\|_{L^r(0,T;L^2(\Omega))}) =: \rho_1,\\ \|y+\eta\|_{C(\bar{Q})} &\le C(1 + \|u\|_{L^r(0,T;L^2(\Omega))} + L_u) =: \rho_2,\end{aligned}$$

where for the last estimate we used Theorem 2.5 combined with (2.4). Now, applying (2.2) with $M := \max\{\rho_1, \rho_2\}$ yields

$$\hat{A}_h(t) \le L_M\,\|(y^h - y - \eta)(t)\|_{L^\infty(\Omega)} \quad \forall\, t \in [0,T], \tag{3.12}$$

in view of $L^\infty(\Omega) \hookrightarrow W^{-1,q}(\Omega)$. To estimate $\hat{B}_h$, let us first observe that $L^2(\Omega) \hookrightarrow W^{-1,q}(\Omega)$, as a result of $q \le 2n/(n-2)$, see (2.1). Now, by employing (3.5) and (3.9), we have

$$\begin{aligned}\|\hat{B}_h\|_{L^r(0,T)} &\le C\,\frac{\|f(y+\eta) - f(y) - f'(y;\eta)\|_{L^r(0,T;L^2(\Omega))}}{\|\eta\|_{L^{r+1}(0,T;L^q(\Omega))}}\|\eta\|_{L^{r+1}(0,T;L^q(\Omega))}\\ &\le C\,\frac{\|f(y+\eta) - f(y) - f'(y;\eta)\|_{L^r(0,T;L^2(\Omega))}}{\|\eta\|_{L^{r+1}(0,T;L^q(\Omega))}}\|h\|_{L^2(Q)},\end{aligned}$$

provided that $\eta \neq 0$. In case that $\eta = 0$, we deduce $\hat{B}_h = 0$, by the definition of $\hat{B}_h$. Further, we observe that $\|h\|_{L^r(0,T;L^2(\Omega))} \to 0$ implies $S'(u;h) \to 0$ in $C(\bar{Q})$, as a consequence of Theorem 2.5 combined with (2.4). Thanks to $q > 2$, Proposition 3.1 gives in turn

$$\frac{\|\hat{B}_h\|_{L^r(0,T)}}{\|h\|_{L^2(Q)}} \leq C \frac{\|f(y+\eta) - f(y) - f'(y;\eta)\|_{L^r(0,T;L^2(\Omega))}}{\|\eta\|_{L^{r+1}(0,T;L^q(\Omega))}} \to 0, \tag{3.13}$$

if $\|h\|_{L^r(0,T;L^2(\Omega))} \to 0$. Now we return to (3.11), where inserting (3.12) results in

$$\begin{aligned}
&\|(y^h - y - \eta)(t)\|_{L^\infty(\Omega)} \\
&\quad \leq L_M \int_0^t \|e^{-(t-s)A}\|_{\mathcal{L}(W^{-1,q}(\Omega), L^\infty(\Omega))} \|(y^h - y - \eta)(s)\|_{L^\infty(\Omega)}\, ds \\
&\qquad + \int_0^t \|e^{-(t-s)A}\|_{\mathcal{L}(W^{-1,q}(\Omega), L^\infty(\Omega))} \hat{B}_h(s)\, ds \\
&\quad \leq L_M \int_0^t c\,(t-s)^{-\theta} \|(y^h - y - \eta)(s)\|_{L^\infty(\Omega)}\, ds \\
&\qquad + \|e^{-\cdot A}\|_{L^{r'}(0,T;\mathcal{L}(W^{-1,q}(\Omega), L^\infty(\Omega)))} \|\hat{B}_h\|_{L^r(0,T)} \quad \forall\, t \in [0,T].
\end{aligned}$$

Here, $\theta \in (0,1)$ denotes the exponent in (2.3). In view of the latter, the mapping $t \mapsto e^{-tA}$ belongs indeed to $L^{r'}(0,T;\mathcal{L}(W^{-1,q}(\Omega), L^\infty(\Omega)))$, since $r'\theta < 1$. By means of a generalized Gronwall's inequality, cf. [18, Lemma 7.1.1, p. 188], we have

$$\|(y^h - y - \eta)(t)\|_{L^\infty(\Omega)} \leq C \|\hat{B}_h\|_{L^r(0,T)} \quad \forall\, t \in [0,T].$$

Then, by (3.13), the estimate

$$\frac{\|y^h - y - \eta\|_{C(\bar{Q})}}{\|h\|_{L^2(Q)}} \leq C \frac{\|\hat{B}_h\|_{L^r(0,T)}}{\|h\|_{L^2(Q)}} \to 0 \quad \text{ if } \|h\|_{L^r(0,T;L^2(\Omega))} \to 0 \tag{3.14}$$

follows. Since A satisfies maximal parabolic $L^r(0,T;W^{-1,q}(\Omega))$-regularity, see Assumption 2.1.3, we have

$$\begin{aligned}
&\|y^h - y - \eta\|_{\mathbb{W}_0^r(W_D^{1,q}(\Omega), W^{-1,q}(\Omega))} \\
&\quad \leq \|(\partial_t + A)^{-1}\|_{\mathcal{L}(L^r(0,T;W^{-1,q}(\Omega)), \mathbb{W}_0^r(W_D^{1,q}(\Omega), W^{-1,q}(\Omega)))} \| - f(y^h) + f(y) + f'(y;\eta)\|_{L^r(0,T;W^{-1,q}(\Omega))} \\
&\quad \leq c\,(\|\hat{A}_h\|_{L^r(0,T)} + \|\hat{B}_h\|_{L^r(0,T)}) \\
&\quad \leq c\,(\|y^h - y - \eta\|_{C(\bar{Q})} + \|\hat{B}_h\|_{L^r(0,T)}),
\end{aligned}$$

by (3.10) and (3.12). Now the desired assertion follows from (3.14). ☐

4. Second-order sufficient optimality conditions. This section is devoted to establishing second-order sufficient conditions (SSC) which guarantee local optimality for (P). Recall that this reads as follows:

$$\left.\begin{aligned}
\min_{u \in L^r(0,T;L^2(\Omega))} \quad & J(y,u) \\
\text{s.t.} \quad & \dot{y}(t) + A\,y(t) + f(y(t)) = B\,u(t) \quad \text{a.e. in } (0,T) \\
& y(0) = 0.
\end{aligned}\right\} \tag{P}$$

The upcoming second-order analysis relies on the extended Bouligand-differentiability from Theorem 3.4 and the (local) Lipschitz continuity of S with range in $C(\bar{Q})$, see (2.6) and (2.4). The main findings are stated in Theorems 4.13 and 4.16 below. Similarly to [3], we present two versions of second-order sufficient optimality conditions. The first set of conditions involves the positive-definiteness of the Hessian of a "Lagrangian" on the cone of critical directions and applies only to objectives with a particular structure, see Assumption 4.11 below. The second set of SSC allows for general (smooth) objectives. In this case, the Hessian of the "Lagrangian" is supposed to be coercive on a larger cone, cf. Assumption 4.14 below. We point out that our SSC are comparatively sharp, as explained at the end of this section. There we will also see that if the nonlinearity f is twice continuously differentiable, then our main results comply with the classical ones.

We begin by recalling the necessary optimality condition for (P) (in form of strong stationarity) established in [24].

Theorem 4.1. *[24, Thm. 5.3, Thm. 6.7] Suppose that the range of B is dense in $W^{-1,q}(\Omega)$. Let $\bar{u} \in L^r(0,T;L^2(\Omega))$ be locally optimal for (P) with associated state $\bar{y} = S(\bar{u}) \in \mathbb{W}_0^r(W_D^{1,q}(\Omega), W^{-1,q}(\Omega))$. Then there exists a unique adjoint state $p \in \mathbb{W}_T^{r'}(W_D^{1,q'}(\Omega), W^{-1,q'}(\Omega))$ and a unique multiplier $\lambda \in L^{r'}(0,T;L^s(\Omega))$ with $s = \frac{nq}{nq-n-q}$ such that*

$$\dot{\bar{y}} + A\,\bar{y} + f(\bar{y}) = B\,\bar{u}, \quad \bar{y}(0) = 0, \tag{4.1a}$$

$$-\dot{p} + A^* p + \lambda = \partial_y J(\bar{y},\bar{u}), \quad p(T) = 0, \tag{4.1b}$$

$$\lambda(t,x) \in [f'_+(\bar{y}(t,x))\,p(t,x), f'_-(\bar{y}(t,x))\,p(t,x)] \;\; \text{a.e. in } Q, \tag{4.1c}$$

$$B^* p + \partial_u J(\bar{y},\bar{u}) = 0, \tag{4.1d}$$

where, for an arbitrary $z \in \mathbb{R}$, the right- and left-sided derivative of $f : \mathbb{R} \to \mathbb{R}$ are defined through $f'_+(z) := f'(z;1)$ and $f'_-(z) := -f'(z;-1)$, respectively.

In the remaining of the section, let $(\bar{u}, \bar{y}, \lambda, p)$ be a fixed point which satisfies the system (4.1) and possesses the same regularity as in Theorem 4.1.

Lemma 4.2. *For any $u \in L^r(0,T;L^2(\Omega))$, there exists $\gamma \in [0,1]$ such that*

$$\begin{aligned} J(y,u) - J(\bar{y},\bar{u}) \geq &\int_Q p(t,x)\big(f(\bar{y}) - f(y) + f'(\bar{y}; y-\bar{y})\big)(t,x)\; d(t,x) \\ &+ \frac{1}{2} J''(y_\gamma, u_\gamma)(y-\bar{y}, u-\bar{u})^2, \end{aligned} \tag{4.2}$$

where we abbreviate $y := S(u)$ and $(y_\gamma, u_\gamma) := (\bar{y},\bar{u}) + \gamma((y,u) - (\bar{y},\bar{u}))$.

Proof. Let $u \in L^r(0,T;L^2(\Omega))$ be arbitrary, but fixed. Using the optimality system, we find

$$\begin{aligned} &\partial_y J(\bar{y},\bar{u})(y-\bar{y}) + \partial_u J(\bar{y},\bar{u})(u-\bar{u}) \\ &\qquad = \langle -\dot{p} + A^* p + \lambda, y - \bar{y}\rangle_{L^r(0,T;W_D^{1,q}(\Omega))} - \langle B^* p, u - \bar{u}\rangle_{L^r(0,T;L^2(\Omega))} \\ &\qquad = \langle p, \dot{y} - \dot{\bar{y}} + A(y-\bar{y}) - B(u-\bar{u})\rangle_{L^r(0,T;W^{-1,q}(\Omega))} + \langle \lambda, y-\bar{y}\rangle_{L^r(0,T;W_D^{1,q}(\Omega))} \\ &\qquad \geq \langle p, f(\bar{y}) - f(y)\rangle_{L^r(0,T;W^{-1,q}(\Omega))} + \int_Q p(t,x) f'(\bar{y}; y-\bar{y})(t,x)\; d(t,x). \end{aligned}$$

For the last equality, we applied the formula of integration by parts from [2, Proposition 5.1] in combination with the initial and final time conditions in (2.5) and (4.1b), respectively. The above inequality can be deduced from the state equation (2.5) and (4.1c) together with the positive homogeneity of the directional derivative. The desired assertion follows now from the continuous Fréchet-differentiability of J, cf. Assumption 2.1.7. □

The key idea in the proofs of Theorems 4.13 and 4.16 below is to write the integral in (4.2) as the sum of a nonnegative term, $-1/2\int_{\mathcal{M}} p(t,x) f''(\bar{y}(t,x)) S'(\bar{u}; u-\bar{u})(t,x)^2\, d(t,x)$, where $\mathcal{M}$ is a suitable subset of Q, and $o(\|u-\bar{u}\|^2_{L^2(Q)})$ for $u \to \bar{u}$ in $L^r(0,T;L^2(\Omega))$. In preparation therefor, we discuss the first term on the right-hand side in (4.2) on three different subsets of Q, by mainly distinguishing between those (t,x) for which f is differentiable or not at $\bar{y}(t,x)$, see (4.3), (4.7), and (4.10) below.

We identify the non-smooth and smooth points of the function f by means of the following sets:

$$\begin{aligned} \mathcal{N} &:= \{z \in \mathbb{R} \,|\, f \text{ is not differentiable at } z\}, \\ \mathcal{S} &:= \{z \in \mathbb{R} \,|\, f \text{ is differentiable at } z\}. \end{aligned}$$

Recall that $\mathcal{N}$ is at most countable, cf. Assumption 2.1.5. This ensures that the sets in (4.3), (4.7), and (4.10) below are measurable.

Next, we introduce the notion of local convexity/concavity for the nonlinearity f. This will play a crucial role in the next two lemmas.

Definition 4.3. *We say that the function f is convex around $y \in \mathbb{R}$ if there exists $\rho > 0$ so that f is convex on the interval $(y-\rho, y+\rho)$. Analogously, we say that f is concave around $y \in \mathbb{R}$ if there exists $\rho > 0$ so that f is concave on the interval $(y-\rho, y+\rho)$.*

We also define the following subset of Q:

$$Q_{\mathfrak{n}} := \{(t,x) \in Q \,|\, \exists\, z \in \mathcal{N} \text{ so that } \bar{y}(t,x) = z\}. \tag{4.3}$$

Lemma 4.4. *Suppose that at any $z \in \mathcal{N}$, the function f is either convex or concave around z with radius $\rho_z > 0$. If $\inf_{z\in\mathcal{N}} \rho_z > 0$, then there exists $\varepsilon > 0$ so that, for any $y \in C(\bar{Q})$ with $\|y-\bar{y}\|_{C(\bar{Q})} < \varepsilon$, there holds*

$$p(t,x)\big(f(\bar{y}) - f(y) + f'(\bar{y}; y-\bar{y})\big)(t,x) \geq 0 \qquad \text{a.e. in } Q_{\mathfrak{n}}.$$

Proof. Let $z \in \mathcal{N}$ be a non-smooth point of f for which the set $\mathfrak{m}_z := \{(t,x) \in Q \,|\, \bar{y}(t,x) = z\}$ has positive measure. If f is convex around z, then straightforward computation shows that $f'_+(z) > f'_-(z)$, since z is a non-differentiable point. Thus, $f'_+(\bar{y}(t,x)) > f'_-(\bar{y}(t,x))$ in $\mathfrak{m}_z$. On the other hand, from (4.1c) we deduce that the interval $[f'_+(\bar{y}(t,x))\, p(t,x), f'_-(\bar{y}(t,x))\, p(t,x)]$ is nonempty for a.a. $(t,x) \in Q$. Thus,

$$p(t,x) \leq 0 \qquad \text{a.e. in } \mathfrak{m}_z. \tag{4.4}$$

Since f is convex on $(z-\rho_z, z+\rho_z)$, it holds

$$f(v) - f(z) \geq f'(z; v-z) \quad \text{for all } v \in (z-\rho_z, z+\rho_z). \tag{4.5}$$

Now, we define $\varepsilon := \inf_{z\in\mathcal{N}} \rho_z > 0$. Let $y \in C(\bar{Q})$ with $\|y - \bar{y}\|_{C(\bar{Q})} < \varepsilon$ be arbitrary, but fixed. Then, $y(t,x) \in (z - \rho_z, z + \rho_z)$ in $\mathfrak{m}_z$, and from (4.5) combined with (4.4) we deduce

$$p(t,x)\big(f(\bar{y}) - f(y) + f'(\bar{y}; y - \bar{y})\big)(t,x) \geq 0 \quad \text{a.e. in } \mathfrak{m}_z. \tag{4.6}$$

Analogously, if f is concave around z, one has $p(t,x) \geq 0$ a.e. in $\mathfrak{m}_z$, and (4.6) follows in the same way as above. Since $Q_{\mathfrak{n}} = \cup_{z\in\mathcal{N}} \mathfrak{m}_z$ by definition, the proof is now complete. ☐

REMARK 4.5. *In view of* (4.4)*, one always knows the sign of the adjoint state p a.e. where $\bar{y}(t,x)$ is a non-differentiable point of f. Unfortunately, this is no longer the case in the smooth points. No matter how close $\bar{y}(t,x) \in \mathcal{S}$ is to some $z \in \mathcal{N}$, a sign for $p(t,x)$ could not be provided. This turns out to be a problem precisely for those (t,x) for which $\bar{y}(t,x)$ is "too close" to the non-smooth points, as explained in Remark 4.7.(iii) below. This is why we need to assume a certain sign condition for p on this critical subset of Q, below also known as $Q_{\mathfrak{s},\delta}$.*

ASSUMPTION 4.6. *Suppose that, at any $z \in \mathcal{N}$, the function f is either convex or concave around z with radius $\rho_z > 0$. Moreover, assume that, for any $z \in \mathcal{N}$, there exists $\delta_z \in (0, \rho_z)$ so that the following conditions are fulfilled:*

1. $\inf_{z\in\mathcal{N}} \rho_z - \delta_z > 0$,
2. *if f is convex around z: $p(t,x) \leq 0$ a.e. where $\bar{y}(t,x) \in (z - \delta_z, z + \delta_z) \setminus \{z\}$,*
3. *if f is concave around z: $p(t,x) \geq 0$ a.e. where $\bar{y}(t,x) \in (z - \delta_z, z + \delta_z) \setminus \{z\}$.*

From now on, $\{\delta_z\}_{z\in\mathcal{N}}$ is supposed to be a fixed set of strict positive, as small as possible values that satisfy Assumption 4.6. By means thereof, we define:

$$Q_{\mathfrak{s},\delta} := \{(t,x) \in Q \,|\, \bar{y}(t,x) \in \cup_{z\in\mathcal{N}} (z - \delta_z, z + \delta_z) \setminus \{z\}\}. \tag{4.7}$$

REMARK 4.7. *(i) For each $z \in \mathcal{N}$, it is desirable to choose $\delta_z > 0$ so (small) that $\{(t,x) \in Q \,|\, \bar{y}(t,x) \in (z - \delta_z, z + \delta_z) \setminus \{z\}\}$ has measure zero, if possible. In this case, Assumption 4.6.2-3 is automatically fulfilled at $z \in \mathcal{N}$. On the other hand, $\rho_z > 0$ should be chosen large, so that the condition in Assumption 4.6.1 is satisfied (and so that f is convex or concave on $(z - \rho_z, z + \rho_z)$). We point out that the neighborhoods of non-smooth points, where f is locally convex or concave are allowed to overlap. The same is true for the intervals $(z - \delta_z, z + \delta_z)$. This does not affect our analysis. However, keep in mind that in view of Assumption 4.6.2-3, $\delta_z > 0$ should be as small as possible.*

(ii) We emphasize that Assumption 4.6.2-3 is characteristic for the second-order analysis of problems with non-differentiable solution mappings. It complies with [22, Assumption 1.(iii)] (elliptic obstacle problem) and [3, (4.8g)-(4.8h)] (static elastoplasticity), where sign conditions for the adjoint state and the multiplier, respectively, are required on a set corresponding to $Q_{\mathfrak{s},\delta}$ (i.e., on the set where $\bar{y}$ is "too close" to the non-smooth points). As explained in [3, Rem. 4.13], this sort of assumption is due to the infinite-dimensional non-smooth framework and it ensures an additional so-called "safety distance". At the end of this section, we go into more detail by comparing our upcoming main result (Theorem 4.13 below) with the main results from [3, 22]. We will see that our SSC are comparatively sharp and can be interpreted as a natural generalization of the SSC for finite dimensional MPECs [30]. Moreover, in Section 5 below we provide settings for which Assumption 4.6 is guaranteed.

(iii) On the set $Q \setminus (Q_{\mathfrak{n}} \cup Q_{\mathfrak{s},\delta})$ we do not need any (sign) conditions. Here we can evaluate the term $(f(\bar{y}) - f(y) + f'(\bar{y}; y - \bar{y})(t,x)$ by means of a Taylor expansion for any $y \in B_{C(\bar{Q})}(\bar{y}, \varepsilon)$, where $\varepsilon > 0$ is chosen appropriately, see (4.11) in the proof of Lemma 4.10 below. Unfortunately, this cannot be done on the critical set $Q_{\mathfrak{s},\delta}$: it is not clear if a $C(\bar{Q})$-neighborhood of $\bar{y}$ exists so that, for all y in this neighborhood, it holds $[\bar{y}(t,x), y(t,x)] \subset \mathcal{S}$ for all $(t,x) \in Q_{\mathfrak{s},\delta}$.
(iv) Let us point out that if there exists $\delta > 0$ so that $|\bar{y}(t,x) - z| \geq \delta$ a.e. in $Q \setminus Q_{\mathfrak{n}}$ for all $z \in \mathcal{N}$, then Assumption 4.6 is no longer needed, and Lemma 4.8 below can be omitted. In this case, we can find a neighborhood as depicted above. Then we can argue as in the proof of Lemma 4.10 below and obtain the therein showed result for the entire set $Q \setminus Q_{\mathfrak{n}}$.

LEMMA 4.8. *Let Assumption 4.6 hold true. Then, there exists $\varepsilon > 0$ so that, for all $y \in C(\bar{Q})$ with $\|y - \bar{y}\|_{C(\bar{Q})} < \varepsilon$, it holds*

$$p(t,x)\big(f(\bar{y}) - f(y) + f'(\bar{y}; y - \bar{y})\big)(t,x) \geq 0 \quad \text{a.e. in } Q_{\mathfrak{s},\delta}.$$

Proof. We define $\varepsilon := \inf_{z \in \mathcal{N}} \rho_z - \delta_z > 0$ and consider $y \in C(\bar{Q})$ with $\|y - \bar{y}\|_{C(\bar{Q})} < \varepsilon$ arbitrary, but fixed. Let $z \in \mathcal{N}$ and denote by $\mathfrak{q}_z$ the set $\{(t,x) \in Q \,|\, \bar{y}(t,x) \in (z - \delta_z, z + \delta_z) \setminus \{z\}\}$. Then, due to $|y(t,x) - z| - |\bar{y}(t,x) - z| \leq \|y - \bar{y}\|_{C(\bar{Q})} < \rho_z - \delta_z$ for all $(t,x) \in Q$, we have

$$|y(t,x) - z| < \rho_z \quad \forall\, (t,x) \in \mathfrak{q}_z. \tag{4.8}$$

If f is convex around z, the inequality $f(v) - f(w) \geq f'(w; v - w)$ is true for all $v, w \in (z - \rho_z, z + \rho_z)$. Since $\rho_z > \delta_z$ by assumption, the definition of $\mathfrak{q}_z$ together with (4.8) and the sign assumption on p, cf. Assumption 4.6.2, now yield

$$p(t,x)\big(f(\bar{y}) - f(y) + f'(\bar{y}; y - \bar{y})\big)(t,x) \geq 0 \quad \text{a.e. in } \mathfrak{q}_z. \tag{4.9}$$

In case that f is concave around z, one arrives at (4.9) in the same way as in the convex case, by making use of Assumption 4.6.3. Since $z \in \mathcal{N}$ was arbitrary and $Q_{\mathfrak{s},\delta} = \cup_{z \in \mathcal{N}} \mathfrak{q}_z$, the desired assertion follows now from (4.9). □

Given the set of strict positive values $\{\delta_z\}_{z \in \mathcal{N}}$ from Assumption 4.6, we define:

$$Q_{\mathfrak{s}} := \{(t,x) \in Q \,|\, |\bar{y}(t,x) - z| \geq \delta_z \ \forall\, z \in \mathcal{N}\}. \tag{4.10}$$

ASSUMPTION 4.9. *From now on, we assume that, in addition to Assumption 4.6, $\hat{\varepsilon} := \inf_{z \in \mathcal{N}} \delta_z / 2 > 0$ holds, and that the nonlinearity f is twice continuously differentiable on $\{v \in \mathbb{R} |\, |v - z| \geq \delta_z/2 \ \ \forall\, z \in \mathcal{N}\}$.*

Notice that, as a direct consequence of Assumption 4.9, the mapping $(t,x) \in Q_{\mathfrak{s}} \mapsto f''(\bar{y}(t,x)) \in \mathbb{R}$ belongs to $C(Q_{\mathfrak{s}})$.

The next lemma is the last essential step before proving the main result.

LEMMA 4.10. *Let Assumption 4.9 be satisfied. Then it holds*

$$\begin{aligned}\int_{Q_{\mathfrak{s}}} p(t,x)\big(f(\bar{y}) - f(S(u)) + f'(\bar{y}; S(u) - \bar{y})\big)(t,x)\, d(t,x)\\ = -\frac{1}{2} \int_{Q_{\mathfrak{s}}} p(t,x) f''(\bar{y}(t,x)) S'(\bar{u}; u - \bar{u})(t,x)^2\, d(t,x) + \mathfrak{r}(u),\end{aligned}$$

where

$$\frac{\mathfrak{r}(u)}{\|u-\bar u\|^2_{L^2(Q)}} \to 0 \qquad \text{as } u \to \bar u \text{ in } L^r(0,T;L^2(\Omega)).$$

Proof. Let $y \in C(\bar Q)$ with $\|y-\bar y\|_{C(\bar Q)} < \hat\varepsilon$ be arbitrary, but fixed. Note that $\hat\varepsilon = \inf_{z\in\mathcal{N}} \delta_z/2 > 0$, cf. Assumption 4.9. From

$$|\bar y(t,x)-z| - |y(t,x)-z| \le \|y-\bar y\|_{C(\bar Q)} < \delta_z/2 \qquad \text{for all } (t,x)\in Q,\ \forall\, z\in\mathcal{N}$$

we deduce that $|y(t,x)-z| > \delta_z/2$ in $Q_{\mathfrak{s}}, \forall\, z \in \mathcal{N}$. Since for all $z\in\mathcal{N}$ it holds $|\bar y(t,x)-z| \ge \delta_z$ and $|y(t,x)-\bar y(t,x)| < \delta_z/2$ for all $(t,x)\in Q_{\mathfrak{s}}$, every point between $\bar y(t,x)$ and $y(t,x)$ belongs to $\{v\in\mathbb{R} \,|\, |v-z| > \delta_z/2\, \forall\, z\in\mathcal{N}\}$ for all $(t,x)\in Q_{\mathfrak{s}}$. Thus, f is twice continuously differentiable on $[\bar y(t,x), y(t,x)]$ for all $(t,x)\in Q_{\mathfrak{s}}$, in view of Assumption 4.9. This allows us to write the Taylor formula

$$\begin{aligned} f(y(t,x)) = f(\bar y(t,x)) &+ f'(\bar y(t,x))(y(t,x)-\bar y(t,x)) + 1/2 f''(\bar y(t,x))(y(t,x)-\bar y(t,x))^2 \\ &+ o((y(t,x)-\bar y(t,x))^2) \quad \forall (t,x)\in Q_{\mathfrak{s}}. \end{aligned} \tag{4.11}$$

Further, from (2.6) and (2.4) we know that there exists a constant $C = C(\bar u) > 0$ such that

$$\|S(v)-\bar y\|_{C(\bar Q)} \le C\|v-\bar u\|_{L^r(0,T;L^2(\Omega))} \tag{4.12}$$

for all $v \in \overline{B_{L^r(0,T;L^2(\Omega))}(\bar u,1)}$. Let now $\epsilon := \min\{\hat\varepsilon/2C, 1\} > 0$ and $u \in \overline{B_{L^r(0,T;L^2(\Omega))}(\bar u,\epsilon)}$, $u \ne \bar u$, be arbitrary, but fixed. Then, due to (4.12), we have $\|S(u)-\bar y\|_{C(\bar Q)} < \hat\varepsilon$ and as a result of (4.11), it holds

$$\begin{aligned} f(S(u)(t,x)) = f(\bar y(t,x)) &+ f'(\bar y(t,x))(S(u)(t,x)-\bar y(t,x)) + 1/2 f''(\bar y(t,x))(S(u)(t,x)-\bar y(t,x))^2 \\ &+ \underbrace{o((S(u)(t,x)-\bar y(t,x))^2)}_{=:h_u(t,x)} \quad \forall (t,x)\in Q_{\mathfrak{s}}. \end{aligned} \tag{4.13}$$

Hence,

$$\begin{aligned} &\int_{Q_{\mathfrak{s}}} p(t,x)\big(f(\bar y) - f(S(u)) + f'(\bar y; S(u)-\bar y)\big)(t,x)\, d(t,x) \\ &\quad = -\frac{1}{2} \underbrace{\int_{Q_{\mathfrak{s}}} p(t,x)\, f''(\bar y(t,x))\big(S(u)(t,x)-\bar y(t,x)\big)^2\, d(t,x)}_{\hat A_u} - \underbrace{\int_{Q_{\mathfrak{s}}} p(t,x) h_u(t,x)\, d(t,x)}_{\hat B_u}. \end{aligned} \tag{4.14}$$

In the following, we discuss the terms $\hat A_u$ and $\hat B_u$ separately. We begin with the first term on the right-hand side of (4.14). For convenience, we abbreviate

$$F(u) := S(u) - S(\bar u) - S'(\bar u; u-\bar u).$$

In view of Theorem 3.4, it holds

$$\frac{\|F(u)\|_{C(\bar Q)}}{\|u-\bar u\|_{L^2(Q)}} \to 0 \qquad \text{as } u\to\bar u \text{ in } L^r(0,T;L^2(\Omega)), \tag{4.15}$$

where we again employed (2.4). Now we write $\hat{A}_u$ as

$$\hat{A}_u = \int_{Q_{\mathfrak{s}}} p(t,x)\, f''(\bar{y}(t,x)) S'(\bar{u}; u-\bar{u})(t,x)^2 \, d(t,x) + \underbrace{\int_{Q_{\mathfrak{s}}} p(t,x)\, f''(\bar{y}(t,x)) F(u)(t,x) \Big(F(u)(t,x) + 2S'(\bar{u}; u-\bar{u})(t,x) \Big) \, d(t,x)}_{D_u} . \tag{4.16}$$

As a result of [1, Section 3], [16, Thm. 3.5], and [31, Chp. 4.6.1-2], there exist $\hat{\beta} > 2$ and $\hat{\varrho} > q'$ so that

$$p \in \mathbb{W}_T^{r'}(W_D^{1,q'}(\Omega), W^{-1,q'}(\Omega)) \hookrightarrow L^{\hat{\beta}}(0,T; L^{\hat{\varrho}}(\Omega)). \tag{4.17}$$

We obtain

$$\frac{|D_u|}{\|u-\bar{u}\|^2_{L^2(Q)}} \leq \|p\|_{L^2(0,T;L^{q'}(\Omega))} \|f''(\bar{y}(\cdot))\|_{L^\infty(Q_{\mathfrak{s}})} \frac{\|F(u)\|_{C(\bar{Q})}}{\|u-\bar{u}\|_{L^2(Q)}} \frac{(\|F(u)\|_{C(\bar{Q})} + 2\|S'(\bar{u}; u-\bar{u})\|_{L^2(0,T;L^q(\Omega))})}{\|u-\bar{u}\|_{L^2(Q)}} \longrightarrow 0 \quad \text{if } \|u-\bar{u}\|_{L^r(0,T;L^2(\Omega))} \to 0, \tag{4.18}$$

in the light of (4.15), (3.5) and (3.9). Next, we address the term $\hat{B}_u$. By relying again on (3.5), (3.9), and (4.15), we see that for all $\beta \in [1,2)$ it holds

$$\|(S(u)-\bar{y})^2\|_{L^\beta(0,T;L^q(\Omega))} \leq c\, \|u-\bar{u}\|^2_{L^2(Q)} \quad \forall\, u \in \overline{B_{L^r(0,T;L^2(\Omega))}(\bar{u},\epsilon)}, \tag{4.19}$$

where $\epsilon > 0$ is small enough. Note that here we also employed $\mathbb{W}_0^2(W_D^{1,q}(\Omega), W^{-1,q}(\Omega)) \hookrightarrow L^2(0,T; L^\infty(\Omega))$, since $q > n$. Further, from (4.13), we infer by applying the mean value theorem, that

$$\begin{aligned}
|h_u(t,x)| &\leq |f(S(u)(t,x)) - f(\bar{y}(t,x)) - f'(\bar{y}(t,x))(S(u)(t,x) - \bar{y}(t,x))| \\
&\quad + L/2(S(u)(t,x) - \bar{y}(t,x))^2 \\
&= |f''(\tilde{y}_u(t,x))|(S(u)(t,x) - \bar{y}(t,x))^2 + L/2(S(u)(t,x) - \bar{y}(t,x))^2 \quad \forall\, (t,x) \in Q_{\mathfrak{s}},
\end{aligned}$$

where $\tilde{y}_u(t,x) := \gamma_u(t,x)(S(u)(t,x) - \bar{y}(t,x)) + \bar{y}(t,x)$, with some $\gamma_u(t,x) \in (0,1)$, and $L > 0$ is a constant depending on $\bar{y}$. Here we used Assumption 4.9 and the fact that $\bar{y} \in C(\bar{Q})$. Due to $\|S(u) - \bar{y}\|_{C(\bar{Q})} < \hat{\varepsilon}$, we have $|\tilde{y}_u(t,x)| \leq \hat{\varepsilon} + \|\bar{y}\|_{C(\bar{Q})}$ for all $(t,x) \in Q_{\mathfrak{s}}$, whence $|f''(\tilde{y}_u(t,x))| \leq L$ for all $(t,x) \in Q_{\mathfrak{s}}$ follows, by Assumption 4.9. Hence, we deduce

$$g_u(t,x) := \frac{h_u(t,x)}{(S(u)(t,x) - \bar{y}(t,x))^2} \leq 3L/2 \quad \text{for all } (t,x) \in Q_{\mathfrak{s}}.$$

Notice that, from (4.13) we have $h_u(t,x) = 0$ if $S(u)(t,x) = \bar{y}(t,x)$ at $(t,x) \in Q_{\mathfrak{s}}$, in which case we define $g_u(t,x) := 0$. Since $\|u - \bar{u}\|_{L^r(0,T;L^2(\Omega))} \to 0$ implies $S(u)(t,x) \to \bar{y}(t,x)$ for all $(t,x) \in Q$, the definition of h_u (see (4.13)) gives $g_u(t,x) \to 0$ for all $(t,x) \in Q_{\mathfrak{s}}$ if $\|u - \bar{u}\|_{L^r(0,T;L^2(\Omega))} \to 0$. By Lebesgue's dominated convergence we then have

$$g_u \to 0 \text{ in } L^\vartheta(Q_{\mathfrak{s}}) \quad \forall\, \vartheta \in [1,\infty)$$

if $u \to \bar{u}$ in $L^r(0,T;L^2(\Omega))$. Together with (4.19) and (4.17), this yields

$$\frac{|\hat{B}_u|}{\|u-\bar{u}\|^2_{L^2(Q)}} \leq \|p\|_{L^{\hat{\beta}}(0,T;L^{\hat{\varrho}}(\Omega))} \frac{\|g_u\|_{L^\vartheta(Q_{\mathfrak{s}})} \|(S(u)-\bar{y})^2\|_{L^\beta(0,T;L^q(\Omega))}}{\|u-\bar{u}\|^2_{L^2(Q)}} \longrightarrow 0 \tag{4.20}$$

if $\|u-\bar{u}\|_{L^r(0,T;L^2(\Omega))} \to 0$, where we set $\vartheta := 1/(1-1/q-1/\hat{\varrho}) \in [1,\infty)$ and $\beta := 1/(1-1/\hat{\beta}-1/\vartheta) \in [1,2)$. By inserting (4.16), (4.18) and (4.20) in (4.14), we finally arrive at the desired result.
$\square$

We are now in the position to establish the first version of second-order sufficient optimality conditions for (P).

ASSUMPTION 4.11. *In addition to Assumptions 4.6 and 4.9, we require that*

1. *The objective $J : L^2(Q) \times L^2(Q) \to \mathbb{R}$ is given by $J(y,u) = g(y) + j(u)$, where $g : L^2(Q) \to \mathbb{R}$ and $j : L^2(Q) \to \mathbb{R}$ are both twice continuously Fréchet-differentiable. There exists $\nu > 0$ with*
$$j''(\bar{u})(h,h) \geq \nu \|h\|^2_{L^2(Q)} \quad \forall\, h \in L^r(0,T;L^2(\Omega)). \tag{4.21}$$
2. *For all $h \in L^2(Q) \setminus \{0\}$ and $\eta = S^{\bar{u}}h$ with $g'(\bar{y})\eta + j'(\bar{u})h = 0$, it holds*
$$g''(\bar{y})(\eta,\eta) + j''(\bar{u})(h,h) - \int_{Q_{\mathfrak{s}}} p(t,x) f''(\bar{y}(t,x))\eta(t,x)^2 \, d(t,x) > 0, \tag{4.22}$$
where $Q_{\mathfrak{s}}$ is the set associated with $\{\delta_z\}_{z\in\mathcal{N}}$ given by (4.10) and $S^{\bar{u}}$ is the solution operator of the (extended) 'linearized' equation (3.4), see Lemma 3.2.

REMARK 4.12. *(i) Assumption 4.11.1 is satisfied by the quadratic functional $\hat{J}(y,u) := \frac{1}{2}\|y-y_d\|^2_{L^2(Q)} + \frac{\nu}{2}\|u-u_d\|^2_{L^2(Q)}$, where $\nu > 0$ and $y_d, u_d \in L^2(Q)$.*
(ii) Similarly to [3, Thm. 4.17], it can be shown that (4.22) is equivalent to
$$g''(\bar{y})(\eta,\eta) + j''(\bar{u})(h,h) - \int_{Q_{\mathfrak{s}}} p(t,x) f''(\bar{y}(t,x))\eta(t,x)^2 \, d(t,x) \geq \kappa \|h\|^2_{L^2(Q)}, \tag{4.23}$$
where $\kappa > 0$, provided that Assumption 4.11.1 is true. Since we work with $L^r(0,T;L^2(\Omega))$ instead of $L^2(Q)$ as space for the control (see Remark 2.6), the coercivity condition (4.21), and thus (4.23), gives rise to the so called two-norm discrepancy, cf. [10, Sec. 3.2]. As a consequence, the quadratic growth in Theorem 4.13 below is expected to hold in the weaker $L^2(Q)$-norm, cf. [10, Sec. 4.3], in particular [10, Thm. 4.11 and 4.13]. We point out that the set $\{h \in L^2(Q) : g'(\bar{y})S^{\bar{u}}h + j'(\bar{u})h = 0\}$ corresponds to the so-called cone of critical directions, see [10, Def. 4.7]. Hence, Assumption 4.11.2 complies with [10, (4.16)]. Note that in [10, Thm. 4.13], it is possible to show local optimality in the sense of $L^2(Q)$ due to the presence of control constraints, see also [9].
(iii) We could replace (4.21) with the more restrictive coercivity condition
$$j''(\bar{u})(h,h) \geq \nu \|h\|^2_{L^r(0,T;L^2(\Omega))} \quad \forall\, h \in L^r(0,T;L^2(\Omega)), \tag{4.24}$$
in which case we obtain $L^r(0,T;L^2(\Omega))$-quadratic growth in Theorem 4.13 below, but (4.24) is not expected to be true in general, since it may lead to an equivalence between $L^r(0,T;L^2(\Omega))$- and $L^2(Q)$-norm. To see this, consider $j(u) = \int_Q F(u)\, d(t,x)$ with $F \in C^2(\mathbb{R})$ in (4.24), see also the examples presented in [10, Sec. 3.2].

THEOREM 4.13. *Let $(\bar{u}, \bar{y}, \lambda, p)$ satisfy the first-order optimality system (4.1) given by Theorem 4.1. If Assumptions 4.6, 4.9 and 4.11 are fulfilled, then there exist $\alpha > 0$ and $R > 0$ such that*
$$J(\bar{y},\bar{u}) + \alpha\|u-\bar{u}\|^2_{L^2(Q)} \leq J(S(u),u) \quad \forall\, u \in B_{L^r(0,T;L^2(\Omega))}(\bar{u},R). \tag{4.25}$$

In particular, $\bar{u}$ is locally optimal for (P).

Proof. The proof is inspired by the proof of [22, Thm. 2.12], see also the proof of [3, Thm. 4.12]. We assume that (4.25) is not satisfied. Thus, there exists a sequence $\{u_k\}_k \subset L^r(0,T;L^2(\Omega))$ with $u_k \to \bar{u}$ in $L^r(0,T;L^2(\Omega))$ and

$$J(\bar{y},\bar{u}) + \frac{1}{k}\|u_k - \bar{u}\|^2_{L^2(Q)} > J(y_k,u_k) \quad \forall\, k \in \mathbb{N},$$

where $y_k := S(u_k)$ for the rest of the proof. For simplicity, we define $\sigma_k := \|u_k - \bar{u}\|_{L^2(Q)}$ and $h_k := \frac{u_k-\bar{u}}{\sigma_k} \in L^r(0,T;L^2(\Omega))$. Then, the above inequality reads

$$J(\bar{y},\bar{u}) + \frac{1}{k}\sigma_k^2 > J(y_k,u_k) \quad \forall\, k \in \mathbb{N}. \tag{4.26}$$

Since $\|h_k\|_{L^2(Q)} = 1$ and $L^2(Q)$ is reflexive, we can extract a subsequence, denoted by the same symbol, so that

$$h_k \rightharpoonup h \quad \text{in } L^2(Q) \quad \text{as } k \to \infty. \tag{4.27}$$

By Lemma 3.2, the operator $S^{\bar{u}} : L^2(Q) \to \mathbb{W}^2_0(W^{1,q}_D(\Omega), W^{-1,q}(\Omega))$ is weakly continuous. Together with (3.9), this implies

$$S'(\bar{u};h_k) \to S^{\bar{u}}h \quad \text{in } L^{\vartheta}(0,T;L^q(\Omega)) \;\; \forall\, \vartheta \in [1,\infty) \quad \text{as } k \to \infty. \tag{4.28}$$

Moreover, according to Theorem 3.4 combined with (2.4), it holds

$$\frac{S(u_k) - S(\bar{u}) - S'(\bar{u};u_k - \bar{u})}{\sigma_k} \to 0 \quad \text{in } C(\bar{Q}) \;\; \text{as } k \to \infty.$$

Hence, by the positive homogeneity of the directional derivative, we have

$$\frac{S(u_k) - S(\bar{u})}{\sigma_k} \to S^{\bar{u}}h \quad \text{in } L^2(Q) \;\; \text{as } k \to \infty. \tag{4.29}$$

Next, we show that $g'(\bar{y})S^{\bar{u}}h + j'(\bar{u})h = 0$. As a result of (4.26), it holds

$$g'(\bar{y})(y_k - \bar{y}) + j'(\bar{u})(u_k - \bar{u}) < \frac{\sigma_k^2}{k} - \frac{1}{2}g''(\tilde{y}_k)(y_k - \bar{y})^2 - \frac{1}{2}j''(\tilde{u}_k)(u_k - \bar{u})^2 \quad \text{for all } k, \tag{4.30}$$

where we abbreviate $(\tilde{y}_k, \tilde{u}_k) = (\bar{y},\bar{u}) + \gamma_k((y_k,u_k) - (\bar{y},\bar{u}))$ with some $\gamma_k \in [0,1]$. Due to $u_k \to \bar{u}$ in $L^r(0,T;L^2(\Omega))$, we have $y_k \to \bar{y}$ in $C(\bar{Q})$, see Proposition 2.4 and (2.4). Since g and j are twice continuously Fréchet-differentiable, we obtain

$$(g''(\tilde{y}_k) - g''(\bar{y}))\Big(\frac{y_k - \bar{y}}{\sigma_k}\Big)^2 + g''(\bar{y})\Big(\frac{y_k - \bar{y}}{\sigma_k}\Big)^2 \to g''(\bar{y})\big(S^{\bar{u}}h, S^{\bar{u}}h\big) \quad \text{as } k \to \infty, \tag{4.31}$$

in view of (4.29), and

$$\liminf_{k\to\infty} (j''(\tilde{u}_k) - j''(\bar{u}))\Big(\frac{u_k - \bar{u}}{\sigma_k}\Big)^2 + j''(\bar{u})\Big(\frac{u_k - \bar{u}}{\sigma_k}\Big)^2 \geq j''(\bar{u})(h,h) \quad \text{as } k \to \infty, \tag{4.32}$$

in the light of (4.27) and (4.21) (which tells us that the mapping $h \mapsto j''(\bar{u})(h,h)$ is convex on $L^2(Q)$). Dividing by σ_k in (4.30) and passing to the limit therein then yields

$$\begin{aligned}\lim_{k\to\infty} g'(\bar{y})\Big(\frac{y_k - \bar{y}}{\sigma_k}\Big) + j'(\bar{u})\Big(\frac{u_k - \bar{u}}{\sigma_k}\Big)& \\ \leq \lim_{k\to\infty}\frac{\sigma_k}{k} - \lim_{k\to\infty}\frac{\sigma_k}{2}g''(\tilde{y}_k)\Big(\frac{y_k - \bar{y}}{\sigma_k}\Big)^2 - \lim_{k\to\infty}\frac{\sigma_k}{2}j''(\tilde{u}_k)\Big(\frac{u_k - \bar{u}}{\sigma_k}\Big)^2 &= 0.\end{aligned} \tag{4.33}$$

On the other hand, (4.29) and (4.27) lead to

$$\lim_{k\to\infty} g'(\bar y)\Big(\frac{y_k-\bar y}{\sigma_k}\Big)+j'(\bar u)\Big(\frac{u_k-\bar u}{\sigma_k}\Big)=g'(\bar y)S^{\bar u}h+j'(\bar u)h\geq 0, \tag{4.34}$$

where the last inequality is due to (4.1) and the proofs of [24, Thm. 5.7 and Thm. 6.10]. Note that the density of the range of B in $W^{-1,q}(\Omega)$ is not needed here. Thus, by (4.33) and (4.34), we have

$$g'(\bar y)S^{\bar u}h+j'(\bar u)h=0. \tag{4.35}$$

Now, we discuss the difference of the values of the objective J. From Lemma 4.2 we know that

$$\begin{aligned} J(y_k,u_k)-J(\bar y,\bar u)\geq &\int_Q p(t,x)\big(f(\bar y)-f(y_k)+f'(\bar y;y_k-\bar y)\big)(t,x)\,d(t,x)\\ &+\frac12 g''(\tilde y_k)(y_k-\bar y)^2+\frac12 j''(\tilde u_k)(u_k-\bar u)^2 \quad \text{for all } k. \end{aligned}\tag{4.36}$$

We begin by estimating the first term on the right-hand side in (4.36). To this end, we want to employ Lemmas 4.4, 4.8 and 4.10. Note that $\inf_{z\in\mathcal N}\rho_z\geq \inf_{z\in\mathcal N}(\rho_z-\delta_z)+\inf_{z\in\mathcal N}\delta_z>0$, by Assumptions 4.6 and 4.9, so that all these three lemmas are applicable. Let now $\varepsilon_1>0$ and $\varepsilon_2>0$ be given by Lemmas 4.4 and 4.8, respectively. From (2.6) and (2.4) we know that there exists a constant $C=C(\bar u)>0$ such that

$$\|S(v)-\bar y\|_{C(\bar Q)}\leq C\|v-\bar u\|_{L^r(0,T;L^2(\Omega))} \tag{4.37}$$

for all $v\in\overline{B_{L^r(0,T;L^2(\Omega))}(\bar u,1)}$. We set $\epsilon:=\min\{\varepsilon_1/2C,\varepsilon_2/2C,1\}>0$ and choose $\bar k$ large enough such that $u_k\in\overline{B_{L^r(0,T;L^2(\Omega))}(\bar u,\epsilon)}$ for all $k\geq\bar k$. Then, by (4.37), we have $\|S(u_k)-\bar y\|_{C(\bar Q)}<\min\{\varepsilon_1,\varepsilon_2\}$ for all $k\geq\bar k$. We are now in the position to apply Lemmas 4.4, 4.8 and 4.10, by means of which (4.36) can be continued as

$$\begin{aligned} J(y_k,u_k)-J(\bar y,\bar u)\geq &-\frac12\int_{Q_{\mathfrak s}} p(t,x)f''(\bar y(t,x))S'(\bar u;u_k-\bar u)(t,x)^2\,d(t,x)+\mathfrak r(u_k)\\ &+\frac12 g''(\tilde y_k)(y_k-\bar y)^2+\frac12 j''(\tilde u_k)(u_k-\bar u)^2 \quad \text{for all } k\geq\bar k, \end{aligned}\tag{4.38}$$

where

$$\frac{\mathfrak r(u_k)}{\sigma_k^2}\to 0\quad \text{as } k\to\infty. \tag{4.39}$$

Note that (4.38) is a result of $Q=Q_{\mathfrak n}\cup Q_{\mathfrak s,\delta}\cup Q_{\mathfrak s}$, see definitions (4.3), (4.7) and (4.10). In view of (4.26) and (4.38), it further holds

$$\begin{aligned} 1/k> &-\frac12\int_{Q_{\mathfrak s}} p(t,x)f''(\bar y(t,x))S'\Big(\bar u;\frac{u_k-\bar u}{\sigma_k}\Big)(t,x)^2\,d(t,x)+\frac{\mathfrak r(u_k)}{\sigma_k^2}\\ &+\frac12 g''(\tilde y_k)\Big(\frac{y_k-\bar y}{\sigma_k}\Big)^2+\frac12 j''(\tilde u_k)\Big(\frac{u_k-\bar u}{\sigma_k}\Big)^2 \quad \text{for all } k\geq\bar k, \end{aligned}\tag{4.40}$$

where we employed the positive homogeneity of the directional derivative. Since $p\in L^{\hat\beta}(0,T;L^{\hat\varrho}(\Omega))$ with $\hat\beta>2$, $\hat\varrho>q'$, see (4.17), (4.28) yields

$$p\big(S'(\bar u;h_k)-S^{\bar u}h\big)\to 0\ \text{ in } L^2(0,T;L^1(\Omega))\quad \text{as } k\to\infty.$$

Together with $f''(\bar{y}(\cdot)) \in C(Q_{\mathfrak{s}})$, see Assumption 4.9, the uniform boundedness of $\{\|S'(\bar{u};h_k) + S^{\bar{u}}h\|_{L^2(0,T;L^\infty(\Omega))}\}_k$, cf. (3.5) and (4.27), this implies

$$\int_{Q_{\mathfrak{s}}} p(t,x) f''(\bar{y}(t,x)) S'(\bar{u};h_k)(t,x)^2 \, d(t,x) \to \int_{Q_{\mathfrak{s}}} p(t,x) f''(\bar{y}(t,x)) (S^{\bar{u}}h)(t,x)^2 \, d(t,x) \tag{4.41}$$

for $k \to \infty$. We now build $\liminf_{k\to\infty}$ in (4.40), which by (4.41), (4.39), (4.31) and (4.32), gives in turn

$$0 \geq -\frac{1}{2} \int_{Q_{\mathfrak{s}}} p(t,x) f''(\bar{y}(t,x)) (S^{\bar{u}}h)(t,x)^2 \, d(t,x) + \frac{1}{2} g''(\bar{y})\big(S^{\bar{u}}h, S^{\bar{u}}h\big) + \frac{1}{2} j''(\bar{u})(h,h). \tag{4.42}$$

By Assumption 4.11.2, which can be applied in view of (4.35), we deduce from (4.42) that $h = 0$. This leads to $S^{\bar{u}}h = 0$. As a result thereof, (4.41) and (4.31) now read

$$\begin{aligned} \int_{Q_{\mathfrak{s}}} p(t,x) f''(\bar{y}(t,x)) S'(\bar{u};h_k)(t,x)^2 \, d(t,x) \to 0 \quad \text{as } k \to \infty, \\ g''(\tilde{y}_k)\Big(\frac{y_k - \bar{y}}{\sigma_k}\Big)^2 \to 0 \quad \text{as } k \to \infty. \end{aligned} \tag{4.43}$$

Thanks to $\|h_k\|_{L^2(Q)} = 1$, the coercivity condition (4.21), and (4.40), we obtain

$$\begin{aligned} 1/k > &-\frac{1}{2} \int_{Q_{\mathfrak{s}}} p(t,x) f''(\bar{y}(t,x)) S'\Big(\bar{u}; \frac{u_k - \bar{u}}{\sigma_k}\Big)(t,x)^2 \, d(t,x) + \frac{\mathfrak{r}(u_k)}{\sigma_k^2} \\ &+ \frac{1}{2} g''(\tilde{y}_k)\Big(\frac{y_k - \bar{y}}{\sigma_k}\Big)^2 + \frac{1}{2}(j''(\tilde{u}_k) - j''(\bar{u}))\Big(\frac{u_k - \bar{u}}{\sigma_k}\Big)^2 + \frac{\nu}{2} \underbrace{\Big\|\frac{u_k - \bar{u}}{\sigma_k}\Big\|_{L^2(Q)}^2}_{=1} \quad \forall\, k \geq \bar{k}. \end{aligned}$$

We pass again to the limit $k \to \infty$ on both sides, which in view of (4.43), (4.39), and the continuity of j'', results in $0 \geq \nu/2 > 0$. This finally gives the contradiction and completes the proof. $\square$

Next, we establish second-order sufficient optimality conditions for (P) which allow for an arbitrary (twice continuously Fréchet-differentiable) objective.

Assumption 4.14. *In addition to Assumptions 4.6 and 4.9, we assume that there exists $\kappa > 0$ such that*

$$J''(\bar{y},\bar{u})(\eta,h)^2 - \int_{Q_{\mathfrak{s}}} p(t,x) f''(\bar{y}(t,x)) \eta(t,x)^2 \, d(t,x) \geq \kappa \|h\|_{L^2(Q)}^2 \tag{4.44}$$

for all $h \in L^r(0,T;L^2(\Omega))$ and $\eta = S'(\bar{u};h)$. Here $Q_{\mathfrak{s}}$ denotes again the set associated with $\{\delta_z\}_{z\in\mathcal{N}}$ given by (4.10).

Remark 4.15. *(i) According to Assumption 4.14, the price for allowing an arbitrary objective J is the more restrictive condition (4.44). Unlike in Assumption 4.11.2, we now deal with a coercivity property which has to be satisfied for the set of all directions, instead of the cone of critical directions. If $J = \hat{J}$ from Remark 4.12.(i), (4.44) is satisfied when*

$$p(t,x) f''(\bar{y}(t,x)) \leq 1 \quad \text{a.e. in } Q_{\mathfrak{s}}. \tag{4.45}$$

In the next section, we provide conditions on the given data which ensure (4.45).
(ii) Since we ask that (4.44) holds in the $L^2(Q)$-norm, we deal again with the two-norm

discrepancy, see Remark 4.12.(ii). If we require (4.44) *to hold in the* $L^r(0,T;L^2(\Omega))$*-norm, then we can overcome the two-norm discrepancy and we have* $L^r(0,T;L^2(\Omega))$*-quadratic growth in Theorem 4.16 below. However, in this case the issues mentioned in Remark 4.12.(iii) apply.*

THEOREM 4.16. *Let* $(\bar u, \bar y, \lambda, p)$ *satisfy the first-order optimality system* (4.1) *given by Theorem 4.1. If Assumptions 4.6, 4.9 and 4.14 are fulfilled, then there exist* $\alpha > 0$ *and* $R > 0$ *such that*

$$J(\bar y, \bar u) + \alpha \|u - \bar u\|_{L^2(Q)}^2 \le J(S(u), u) \quad \forall\, u \in B_{L^r(0,T;L^2(\Omega))}(\bar u, R). \tag{4.46}$$

In particular, $\bar u$ *is locally optimal for* (P).

Proof. In the proof of Theorem 4.13 we already checked that, under Assumptions 4.6 and 4.9, Lemma 4.4 is applicable (in addition to Lemmas 4.8 and 4.10). We define again $\epsilon := \min\{\varepsilon_1/2C, \varepsilon_2/2C, 1\} > 0$ and fix $u \in \overline{B_{L^r(0,T;L^2(\Omega))}(\bar u, \epsilon)}$, $u \ne \bar u$, where $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ are given by Lemmas 4.4 and 4.8, respectively, and $C = C(\bar u) > 0$ is a constant so that (4.37) holds. Then, one has $\|y - \bar y\|_{C(\bar Q)} < \min\{\varepsilon_1, \varepsilon_2\}$, see the proof of Theorem 4.13, where $y := S(u)$ from now on. From Lemma 4.2 combined with Lemmas 4.4, 4.8, and 4.10 we have

$$\begin{aligned} J(y,u) - J(\bar y, \bar u) \ge &-\frac{1}{2}\int_{Q_s} p(t,x) f''(\bar y(t,x)) S'(\bar u; u-\bar u)(t,x)^2 \, d(t,x) + \mathfrak{r}(u) \\ &+ \frac{1}{2} J''(y_{\gamma_u}, u_{\gamma_u})(y - \bar y, u - \bar u)^2, \end{aligned} \tag{4.47}$$

where $\mathfrak{r}(u)/\|u - \bar u\|_{L^2(Q)}^2 \to 0$ if $u \to \bar u$ in $L^r(0,T;L^2(\Omega))$ and $(y_{\gamma_u}, u_{\gamma_u}) = (\bar y, \bar u) + \gamma_u((y,u) - (\bar y, \bar u))$ with some $\gamma_u \in [0,1]$. Further, by means of some algebraic manipulations, see the proof of [3, Thm. 4.6], we can write

$$\begin{aligned} \hat{\mathfrak{r}}(u) :&= J''(y_{\gamma_u}, u_{\gamma_u})(y - \bar y, u - \bar u)^2 - J''(\bar y, \bar u)(S'(\bar u; u - \bar u), u - \bar u)^2 \\ &= \big(J''(y_{\gamma_u}, u_{\gamma_u}) - J''(\bar y, \bar u)\big)(y - \bar y, u - \bar u)^2 + \partial_y^2 J(\bar y, \bar u)(y - \bar y - S'(\bar u; u - \bar u))^2 \\ &\quad + 2\partial_y^2 J(\bar y, \bar u)(y - \bar y - S'(\bar u; u - \bar u), S'(\bar u; u - \bar u)) \\ &\quad + 2\partial_y \partial_u J(\bar y, \bar u)(y - \bar y - S'(\bar u; u - \bar u), u - \bar u). \end{aligned} \tag{4.48}$$

Due to Assumption 2.1.7 and (3.5) combined with Theorem 3.4 (which imply $\|y - \bar y\|_{L^2(0,T;W^{1,q}_D(\Omega))} \le c\|u - \bar u\|_{L^2(Q)}$) we have

$$\hat{\mathfrak{r}}(u)/\|u - \bar u\|_{L^2(Q)}^2 \longrightarrow 0 \ \text{ if } u \to \bar u \text{ in } L^r(0,T;L^2(\Omega)).$$

Inserting (4.48) in (4.47) and employing Assumption 4.14 as well as the convergence properties of $\mathfrak{r}(u)$ and $\hat{\mathfrak{r}}(u)$ results in

$$\begin{aligned} J(S(u),u) - J(\bar y, \bar u) &\ge -\frac{1}{2}\int_{Q_s} p(t,x) f''(\bar y(t,x)) S'(\bar u; u - \bar u)(t,x)^2 \, d(t,x) \\ &\qquad + \frac{1}{2} J''(\bar y, \bar u)(S'(\bar u; u - \bar u), u - \bar u)^2 + \mathfrak{r}(u) + \frac{1}{2}\hat{\mathfrak{r}}(u) \\ &\ge \underbrace{\Big(\frac{\kappa}{2} - \frac{|\mathfrak{r}(u) + \hat{\mathfrak{r}}(u)/2|}{\|u - \bar u\|_{L^2(Q)}^2}\Big)}_{\ge \kappa/4} \|u - \bar u\|_{L^2(Q)}^2 \end{aligned}$$

for all $u \in \overline{B_{L^r(0,T;L^2(\Omega))}(\bar u, R)}$ with $R > 0$ small enough. The proof is now complete. $\square$

We conclude this section by comparing our SSC with the classical SSC in the smooth case (see e.g. [10]) and the SSC from the contributions [3, 22] (non-smooth case), respectively.

First, let us see what happens when $f : \mathbb{R} \to \mathbb{R}$ is twice continuously differentiable. Then, $\mathcal{N} = \emptyset$ and $Q_{\mathfrak{s}} = Q$, and thus, Assumptions 4.6 and 4.9 are automatically fulfilled. Moreover, (4.22) reads

$$\partial^2_{(y,u)}\mathfrak{L}(\bar y, \bar u, p)(S^{\bar u}h, h)^2 > 0 \quad \forall\, h \in L^2(Q) \setminus \{0\}, \tag{4.49}$$

where $\partial^2_{(y,u)}\mathfrak{L}(\bar y, \bar u, p)$ is the continuous extension of the second derivative on the space $\mathbb{W}^2_0(W^{1,q}_D(\Omega), W^{-1,q}(\Omega)) \times L^2(Q)$ of the Lagrangian $\mathfrak{L} : \mathbb{W}^r_0(W^{1,q}_D(\Omega), W^{-1,q}(\Omega)) \times L^r(0,T;L^2(\Omega)) \times L^{r'}(0,T;W^{1,q'}_D(\Omega)) \to \mathbb{R}$,

$$\mathfrak{L}(y,u,p) = g(y) + j(u) - \langle p, \dot y + A\, y + f(y) - Bu\rangle_{L^r(0,T;W^{-1,q}(\Omega))}.$$

Note that $\{h \in L^2(Q) : g'(\bar y)S^{\bar u}h + j'(\bar u)h = 0\} = L^2(Q)$, by the linearity of $S^{\bar u}$ and a density argument. Hence, the result in Theorem 4.13 is true if (4.1) (which in the smooth case coincides with the classical KKT system), Assumption 4.11.1 and (4.49) are satisfied. This complies with the assertion in [10, Thm. 4.13], where a concrete objective and control constraints are considered. Let us point out that the structural assumption on J is to be expected, since this is essential in infinite dimensions in order to obtain a contradiction at the end of the proof of Theorem 4.13, see also the proof of [10, Thm. 4.13]. Due to the reasons explained in Remark 4.12.(ii), we deal with the two-norm discrepancy, which is also the case in [10, Thm. 4.13]. For the second version of SSC, the discussion is similar, cf. also Remark 4.15. In conclusion, if the nonlinearity $f : \mathbb{R} \to \mathbb{R}$ is twice continuously differentiable, then both sets of SSC derived in this section coincide with the classical ones.

We now turn to the second-order sufficient conditions in the non-smooth case. For finite dimensional MPECs, these consist of strong stationarity (necessary conditions for local optimality) and the coercivity/positive-definiteness of the Hessian (w.r.t. the primal variables) of the Lagrangian on the cone of critical directions, see [30, Thm. 7]. Thus, the SSC from [30] coincide with our SSC in Theorem 4.13 except that we tighten the sign condition for the adjoint state (Assumption 4.6); recall that, in view of (4.4), one always knows the sign of p a.e. where $\bar y(t,x)$ is a non-differentiable point of f, see Remark 4.5. We underline that this sort of assumption is standard in the context of deriving SSC for *infinite dimensional* control problems *which feature non-smooth solution operators*. It ensures an additional so-called "safety distance", cf. [3, Rem. 4.13], see also Remark 4.7.(ii). We emphasize that a corresponding condition to our Assumption 4.6 is required in [3] (static elastoplasticity) and [22] (elliptic obstacle problem) as well (the only contributions known to the author which deal with SSC in the non-smooth case). To be more precise, we refer to [22, Assumption 1.(iii)] and [3, (4.8g)-(4.8h)], where sign conditions for the adjoint state and the multiplier, respectively, are required in those (t,x) for which $\bar y(t,x)$ is "too close" to the non-smooth points. Not only do the SSC in these contributions contain a corresponding condition to our Assumption 4.6, but they also consist of additional assumptions: in [22, Assumption 1.(iv)], a sign condition (in the variational sense) is imposed on the multiplier, while in [3, (4.16)], regularity assumptions on the multipliers and adjoint state are made. Therefore, our SSC are comparatively sharp, as they reduce only to the classical requirements (strong stationarity and the coercivity of the Hessian of the Lagrangian) and Assumption 4.6. Note that we can provide different settings where the latter is fulfilled (see Section 5 below). All in all, the sufficient conditions in Theorem 4.13 can be regarded as a natural extension of their finite dimensional counterpart. Note that, as already mentioned above, the structural assumption on the objective (Assumption 4.11.1) is due to the fact that we work in infinite dimensions and appears in the smooth case as well. It is contained in the SSC from [3,22] too and can be overcome by requiring the alternative version of SSC presented in Theorem 4.16.

5. Second-order sufficient conditions for a concrete setting. In this section, we derive conditions on the given data under which Assumption 4.6.2-3 and (4.45) are guaranteed. We consider the optimal control problem

$$\left.\begin{aligned} \min_{u\in L^r(0,T;L^2(\Omega))} \quad & \frac{1}{2}\|y-y_d\|_{L^2(Q)}^2+\frac{1}{2}\|u\|_{L^2(Q)}^2 \\ \text{s.t.} \quad & \dot y(t)-\triangle y(t)+f(y(t))=u(t) \quad \text{a.e. in } (0,T), \\ & y(0)=0. \end{aligned}\right\} \qquad (P_{ex})$$

In all what follows, $B: L^2(\Omega)\hookrightarrow (W^{-1,q}(\Omega), W^{1,q}_D(\Omega))_{\varsigma,\infty}$ is the embedding operator, in which case $\varsigma = n/2q - n/4 + 1/2$, see [1, Section 3], [16, Thm. 3.5], and [31, Thm. 4.6.1.c)]. Note that B satisfies Assumption 2.1.6, provided that $n<q<2n/(n-2)$. The exponent r is supposed to fulfill (2.1) where $q\in(n,2n/(n-2)]$ is fixed. In (P_{ex}), $\triangle: W^{1,q}_D(\Omega)\to W^{-1,q}(\Omega)$ denotes the Laplace operator in the distributional sense, i.e., $A:=-\triangle = -\operatorname{div}\nabla$. We assume that Ω is such that Assumptions 2.1.2-3 are satisfied by $-\triangle$, cf. Remark 2.7, and $\partial\Omega=\Gamma_D$. Recall that A satisfies maximal parabolic $L^s(0,T;W^{-1,q}(\Omega))$-regularity for every $s\in(1,\infty)$. By $C_\Omega>0$ we denote the Poincaré constant associated with the domain Ω. We suppose that the nonlinearity satisfies $f(0)=0$ and that f is convex around any $z\in\mathcal{N}$ with radius $\rho_z>0$ (in addition to Assumptions 2.1.4-5). The desired state y_d belongs to $L^r(0,T;L^2(\Omega))$.

For the above-described problem we show that if (5.9) below is satisfied, then $p(t,x)\le 0$ a.e. in Q, i.e., Assumption 4.6.2 holds true. Let us define $\delta_z:=\rho_z/2$ for all $z\in\mathcal{N}$ in Assumption 4.6. If, in addition to (5.9) below, f is twice continuously differentiable on $\{v\in\mathbb{R}|\,|v-z|\ge\rho_z/4 \;\;\forall\, z\in\mathcal{N}\}$, and (5.15) below holds, we can also prove (4.45). Thus, if $\mathcal{N}$ is finite, we can provide a setting where all assumptions in Theorems 4.13 and 4.16 are satisfied, see also Remarks 4.12 and 4.15. In this context, (5.1) is not only necessary but also sufficient for optimality.

Before we begin with the proof, we mention that a nonlinearity f which satisfies all the above conditions (i.e., Lipschitz continuous on bounded sets, directionally differentiable, convex around any $z\in\mathcal{N}$, monotone increasing with $f(0)=0$ and twice continuously differentiable on a subset of $\mathcal{S}$) is depicted in Figure 5.1. This could arise in combustion processes where different ignition temperatures are given, see Remark 2.9.

Throughout this section, $(\bar u,\bar y,p,\lambda)\in L^r(0,T;L^2(\Omega))\times \mathbb{W}^r_0(W^{1,q}_D(\Omega),W^{-1,q}(\Omega))\times \mathbb{W}^{r'}_T(W^{1,q'}_D(\Omega),W^{-1,q'}(\Omega))\times L^{r'}(0,T;L^s(\Omega))$(with $s=\frac{nq}{nq-n-q}$) is a fixed point that satisfies the first-order optimality system given by Theorem 4.1. In the setting con-

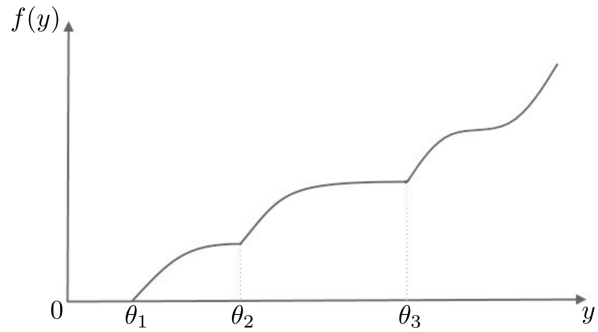

FIG. 5.1. *An ignition-type nonlinearity with ignition temperatures θ_1, θ_2, and θ_3*

sidered here, (4.1) reads

$$\dot{\bar{y}} - \triangle\,\bar{y} + f(\bar{y}) = \bar{u}, \quad \bar{y}(0) = 0, \tag{5.1a}$$

$$-\dot{p} - \triangle p + \lambda = \bar{y} - y_d, \quad p(T) = 0, \tag{5.1b}$$

$$\lambda(t,x) \in [f'_+(\bar{y}(t,x))\,p(t,x), f'_-(\bar{y}(t,x))\,p(t,x)] \text{ a.e. in } Q, \tag{5.1c}$$

$$p + \bar{u} = 0. \tag{5.1d}$$

Due to the monotonicity of f (see Assumption 2.1.4) and $f(0) = 0$, we have $f(y)y \geq 0$ and $f'(y;h)h \geq 0$ for all $y, h \in \mathbb{R}$. By (5.1c), we find

$$f(\bar{y}(t,x))\bar{y}(t,x) \geq 0, \quad \lambda(t,x)p(t,x) \geq 0, \ \lambda(t,x)p^+(t,x) \geq 0 \quad \text{a.e. in } Q. \tag{5.2}$$

where we abbreviate $p^+ := \max\{p, 0\}$.

(I) We first deal with Assumption 4.6.2. We start by showing that $\|p\|_{L^2(Q)} \leq K$, where $K > 0$ is some constant which depends only on the given parameters (step (i) below). This will enable us to derive conditions on the data such that $\bar{y} \leq y_d$ a.e. in Q holds (step (ii) below). By means of this inequality, we can then conclude that $p \leq 0$ a.e. in Q (step (iii) below).

(i) As a consequence of (5.1d), we have $p \in L^r(0,T;L^2(\Omega))$, which gives in turn $\lambda \in L^r(0,T;L^2(\Omega))$, by (5.1c) and (3.2). Thus, $t \mapsto (\bar{y}-y_d-\lambda)(T-t) \in L^r(0,T;L^2(\Omega)) \hookrightarrow L^r(0,T;W^{-1,q}(\Omega))$, since $q \leq 2n/(n-2)$. As $-\triangle$ satisfies maximal parabolic $L^r(0,T;W^{-1,q}(\Omega))$-regularity, we now deduce from (5.1b) that $p \in \mathbb{W}_0^r(W_D^{1,q}(\Omega), W^{-1,q}(\Omega))$, where we used the transformation $t \mapsto T-t$. Therefore, we can test the adjoint equation (5.1b) with $p(T-\cdot)$, which leads to

$$\begin{aligned} \frac{1}{2}\|p(T-t)\|^2_{L^2(\Omega)} - \frac{1}{2}\underbrace{\|p(T)\|^2_{L^2(\Omega)}}_{=0} + \frac{1}{C_\Omega^2}\int_0^t \|p(T-s)\|^2_{L^2(\Omega)}\,ds \\ \leq \int_0^t \int_\Omega (\bar{y}-y_d)(T-s,x)p(T-s,x)\,d(s,x) \quad \forall\, t \in [0,T], \end{aligned} \tag{5.3}$$

in view of the formula of integration by parts, Poincaré-Friedrichs's inequality and (5.2). We proceed in the same way regarding the state equation. We test (5.1a) with y and employ the first inequality in (5.2), as well as Poincaré-Friedrichs's inequality.

The resulting estimate combined with (5.1d) is then used on the right-hand side of (5.3) at $t := T$, which yields

$$\frac{1}{C_\Omega^2}\int_0^T \|p(s)\|_{L^2(\Omega)}^2\, ds \le \underbrace{\|\bar y(0)\|_{L^2(\Omega)}^2}_{=0} - \frac{1}{C_\Omega^2}\int_0^T \|\bar y(s)\|_{L^2(\Omega)}^2\, ds - \int_Q y_d(s,x)p(s,x)\, d(s,x). \tag{5.4}$$

From (5.4) we have

$$2\|p\|_{L^2(Q)}\|\bar y\|_{L^2(Q)} \le \|p\|_{L^2(Q)}^2 + \|\bar y\|_{L^2(Q)}^2 \le C_\Omega^2 \|y_d\|_{L^2(Q)}\|p\|_{L^2(Q)}. \tag{5.5}$$

Dividing by $\|p\|_{L^2(Q)}$ results in

$$\|p\|_{L^2(Q)} \le C_\Omega^2\|y_d\|_{L^2(Q)}, \quad \|\bar y\|_{L^2(Q)} \le \frac{C_\Omega^2\|y_d\|_{L^2(Q)}}{2}. \tag{5.6}$$

(ii) In order to prove that $p \le 0$, we derive conditions which guarantee that $\bar y \le y_d$ a.e. in Q. To this end, we insert (5.5) on the right-hand side in (5.3) and obtain

$$\|p(T-t)\|_{L^2(\Omega)}^2 \le (C_\Omega^2+2)\|y_d\|_{L^2(Q)}\|p\|_{L^2(Q)} \le (C_\Omega^2+2)C_\Omega^2\|y_d\|_{L^2(Q)}^2 \quad \forall\, t\in[0,T], \tag{5.7}$$

where for the last inequality we used (5.6). Via a comparison principle, cf. [28, Lem. A.1, Prop. 3.3 and 3.4], where one relies on the monotonicity of f, it can be shown that

$$|\bar y(t,x)| \le \tilde y(t,x) \le \underbrace{C_e\|(\partial_t - \triangle)^{-1}\|_{\mathcal{L}(L^r(0,T;W^{-1,q}(\Omega)),\mathbb{W}_0^r(W_D^{1,q}(\Omega),W^{-1,q}(\Omega)))}}_{=:K_1}\|p\|_{L^r(0,T;L^2(\Omega))} \quad \text{a.e. in } Q, \tag{5.8}$$

where $\tilde y \in \mathbb{W}_0^r(W_D^{1,q}(\Omega), W^{-1,q}(\Omega))$ is the unique solution of $\dot{\tilde y} - \triangle\,\tilde y = |\bar u|$ and $C_e > 0$ is the product of the embedding constants of $\mathbb{W}_0^r(W_D^{1,q}(\Omega), W^{-1,q}(\Omega)) \hookrightarrow C(\bar Q)$ and $L^2(\Omega) \hookrightarrow W^{-1,q}(\Omega)$, cf. (2.4). Note that for (5.8) we employed again (5.1d). Thus, in view of (5.7) and (5.8), we deduce that if

$$K_1\, T^{1/r}\sqrt{C_\Omega^2+2}\,C_\Omega\|y_d\|_{L^2(Q)} \le y_d \text{ a.e. in } Q, \tag{5.9}$$

then $\bar y \le y_d$ a.e. in Q.

(iii) Now, to see that (5.9) implies the desired result, we test the equation (5.1b) with $p^+ \in L^r(0,T;W_D^{1,q}(\Omega))$, see [20, Thm. A.1]. We arrive at

$$\int_0^t \langle -\dot p(T-s), p^+(T-s)\rangle_{W_D^{1,q}(\Omega)}\, ds + \int_0^t\int_\Omega \underbrace{\nabla p(T-s)(x)\nabla p^+(T-s)(x)}_{\ge 0}\, d(s,x)$$
$$+\int_0^t\int_\Omega \underbrace{\lambda(T-s,x)p^+(T-s,x)}_{\ge 0,\text{ see } (5.2)}\, d(s,x) = \int_0^t\int_\Omega \underbrace{(\bar y - y_d)(T-s,x)}_{\le 0}\, p^+(T-s,x)\, d(s,x) \quad \forall\, t\in[0,T]. \tag{5.10}$$

Thanks to an argument similar to [33, Lemma 3.2] combined with (5.10), we have

$$\int_0^t \langle -\dot p(T-s), p^+(T-s)\rangle_{W_D^{1,q}(\Omega)}\, ds = 1/2\|p^+(T-t)\|_{L^2(\Omega)}^2 - 1/2\underbrace{\|p^+(T)\|_{L^2(\Omega)}^2}_{=0} \le 0 \ \text{ for all } t\in[0,T],$$

which gives in turn $p \leq 0$ a.e. in Q.

(II) In order to derive conditions which guarantee (4.45), we first show that there exists a constant $c > 0$, dependent only on the given data, so that $-p \leq c$ a.e. in Q. To this end, we apply a comparison principle again and assume in the following that (5.9) holds. Consider the equation

$$-\dot{\tilde{p}} - \triangle\tilde{p} = \bar{y} - y_d, \quad \tilde{p}(T) = 0. \tag{5.11}$$

Since $-\triangle$ satisfies maximal parabolic $L^r(0,T;W^{-1,q}(\Omega))$-regularity, there exists a unique $\tilde{p} \in \mathbb{W}^r_T(W^{1,q}_D(\Omega), W^{-1,q}(\Omega)) \hookrightarrow C(\bar{Q})$ which solves (5.11). To see this, one uses the transformation $t \mapsto T - t$. Thus,

$$-\tilde{p}(t,x) \leq \|\tilde{p}\|_{C(\bar{Q})} \leq K_1\|\bar{y} - y_d\|_{L^r(0,T;L^2(\Omega))} \leq c \quad \text{for all } (t,x) \in Q, \tag{5.12}$$

where K_1 denotes the constant from (5.8). Note that a value for c (dependent only on the given data) can be obtained from (5.7) and (5.8). In view of (5.2) and $p \leq 0$, $\lambda \leq 0$ follows. We "compare" (5.1b) with (5.11) and we see that

$$-p(t,x) \leq -\tilde{p}(t,x) \leq c \quad \text{a.e. in } Q, \tag{5.13}$$

as a result of (5.12). Here we relied on [33, Lemma 3.3]. Before we proceed with the proof, let us recall that we defined $\delta_z := \rho_z/2$ for all $z \in \mathcal{N}$, which means that

$$Q_{\mathfrak{s}} = \{(t,x) \in Q \,|\, |\bar{y}(t,x) - z| \geq \rho_z/2 \;\forall\, z \in \mathcal{N}\}.$$

Note that (4.45) is automatically satisfied a.e. in $\{(t,x) \in Q_{\mathfrak{s}} | f''(\bar{y}(t,x)) \geq 0\}$, since $p \leq 0$. On the other hand, from (5.13) one has

$$\begin{aligned} f''(\bar{y}(t,x))p(t,x) &\leq -cf''(\bar{y}(t,x)) = c|f''(\bar{y}(t,x))| \\ &\leq c \sup_{v\in\mathcal{M}} |f''(v)| \\ &< \infty \quad \text{f.a.a. } (t,x) \in Q_{\mathfrak{s}} \text{ with } f''(\bar{y}(t,x)) < 0, \end{aligned} \tag{5.14}$$

where we abbreviate $\mathcal{M} := \{v \in \mathbb{R} \,|\, |v - z| \geq \rho_z/2 \,\forall\, z \in \mathcal{N},\; |v| \leq \|\bar{y}\|_{C(\bar{Q})},\; f''(v) < 0\}$. The last inequality in (5.14) is true, since f'' is continuous on $\{v \in \mathbb{R} |\, |v - z| \geq \rho_z/4 \;\; \forall\, z \in \mathcal{N}\}$, by assumption. Thus, if in addition to (5.9),

$$c \sup_{v\in\mathcal{M}} |f''(v)| \leq 1 \tag{5.15}$$

holds, then (4.45) is guaranteed. Note that (4.45) is automatically fulfilled if f is convex on $\{v \in \mathbb{R} \,|\, |v - z| \geq \rho_z/4 \,\forall\, z \in \mathcal{N},\; |v| \leq \|\bar{y}\|_{C(\bar{Q})} + 1\}$, since in this case $f''(\bar{y}(t,x)) \geq 0$ for all $(t,x) \in Q_{\mathfrak{s}}$.

Note that, if $\mathcal{N}$ is finite, then the inf-conditions in Assumptions 4.6 and 4.9 are true. In conclusion, for the setting considered here, (5.9) and (5.15) imply that every strong stationary point of (P) satisfies all assumptions in Theorems 4.13 and 4.16, and thus, the necessary optimality condition (5.1) is also sufficient for local optimality.

Appendix A.

Lemma A.1. *For every $0 < \hat{\varsigma} \leq \beta < 1$ it holds*

$$\|e^{-tA}\|_{\mathcal{L}((W^{-1,q}(\Omega),W^{1,q}_D(\Omega))_{\hat{\varsigma},1},(W^{-1,q}(\Omega),W^{1,q}_D(\Omega))_{\beta,\infty})} \leq c\, t^{\hat{\varsigma}-\beta} \quad \forall\, t \in (0,T]. \tag{A.1}$$

For all $\varsigma \in (0,1/2)$ we have

$$\Big(\int_0^T \|e^{-sA}\|^2_{\mathcal{L}((W^{-1,q}(\Omega),W^{1,q}_D(\Omega))_{\varsigma,\infty},L^2(\Omega))}\,ds\Big)^{1/2} < \infty. \tag{A.2}$$

Moreover, for any $\gamma \in (1/2,1)$ it holds

$$\|e^{-tA}\|_{\mathcal{L}(W^{-1,q}(\Omega),L^2(\Omega))} \le t^{-\gamma} \quad \forall\, t \in (0,T]. \tag{A.3}$$

Proof. Let $t \in (0,T]$ and $\psi \in (W^{-1,q}(\Omega),W^{1,q}_D(\Omega))_{\hat\varsigma,1}$ be arbitrary, but fixed. Then, by [31, Theorem 1.15.2], and [26, Theorems 2.6.8d), 2.6.13], we have

$$\begin{aligned}\|e^{-tA}\psi\|_{(W^{-1,q}(\Omega),W^{1,q}_D(\Omega))_{\beta,\infty}} &\le c\,\|A^{\beta}e^{-tA}\psi\|_{W^{-1,q}(\Omega)}\\ &\le c\,\|A^{\beta-\hat\varsigma}e^{-tA}\|_{\mathcal{L}(W^{-1,q}(\Omega),W^{-1,q}(\Omega))}\|A^{\hat\varsigma}\psi\|_{W^{-1,q}(\Omega)}\\ &= c\,\|A^{\beta-\hat\varsigma}e^{-tA}\|_{\mathcal{L}(W^{-1,q}(\Omega),W^{-1,q}(\Omega))}\|\psi\|_{\mathcal{D}(A^{\hat\varsigma})}\\ &\le c\,t^{\hat\varsigma-\beta}\|\psi\|_{(W^{-1,q}(\Omega),W^{1,q}_D(\Omega))_{\hat\varsigma,1}},\end{aligned} \tag{A.4}$$

where $\mathcal{D}(A^{\hat\varsigma})$ is the domain of $A^{\hat\varsigma}$. This proves (A.1). Next we show (A.2). To this end, consider $\varsigma \in (0,1/2)$ arbitrary, but fixed. We first observe that there exist $\hat\varsigma, \beta \in (0,1)$ such that $\varsigma > \hat\varsigma > \beta - 1/2 > 0$. This implies $(W^{-1,q}(\Omega),W^{1,q}_D(\Omega))_{\varsigma,\infty} \hookrightarrow (W^{-1,q}(\Omega),W^{1,q}_D(\Omega))_{\hat\varsigma,1}$, cf. [1, Section 3], and

$$(W^{-1,q}(\Omega),W^{1,q}_D(\Omega))_{\beta,\infty} \hookrightarrow [W^{-1,q}(\Omega),W^{1,q}_D(\Omega)]_{\frac12} = L^q(\Omega) \hookrightarrow L^2(\Omega), \tag{A.5}$$

see [16, Thm. 3.5]. From (A.1) we have

$$\|e^{-tA}\|_{\mathcal{L}((W^{-1,q}(\Omega),W^{1,q}_D(\Omega))_{\varsigma,\infty},L^2(\Omega))} \le c\,t^{\hat\varsigma-\beta} \quad \forall\, t \in (0,T], \tag{A.6}$$

with the above choice of $\hat\varsigma$ and β. Due to $2(\hat\varsigma-\beta) > -1$ and (A.6), we can now infer (A.2). Further, a computation similar to (A.4) shows that

$$\|e^{-tA}\|_{\mathcal{L}(W^{-1,q}(\Omega),(W^{-1,q}(\Omega),W^{1,q}_D(\Omega))_{\gamma,\infty})} \le c\,t^{-\gamma} \quad \forall\, t \in (0,T], \quad \forall\, \gamma \in (0,1).$$

Together with (A.5), this implies (A.3). The proof is now complete.
□

Acknowledgment. The author is very thankful to Christian Meyer (TU Dortmund) for pointing out the references [3, 22]. This work was supported by the DFG grant YO 159/2-1 within the Priority Program SPP 1962 (Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization), which is gratefully acknowledged.

REFERENCES

[1] H. Amann. Linear parabolic problems involving measures. *Rev. R. Acad. Cien. Serie A. Mat.*, 95(1):85–119, 2001.

[2] H. Amann. Nonautonomous parabolic equations involving measures. *J. Math. Sci.*, 130:4780–4802, 2005.

[3] T. Betz and C. Meyer. Second-order sufficient optimality conditions for optimal control of static elastoplasticity with hardening. *ESAIM: Control, Optimisation and Calculus of Variations*, 21:271–300, 2014.

[4] J. F. Bonnans. Second-order analysis for control constrained optimal control problems of semilinear elliptic systems. *Applied Mathematics and Optimization*, 38(3):303–325, 1998.

[5] J.F. Bonnans and A. Shapiro. *Pertubation Analysis of Optimization Problems.* Springer, New York, 2000.

[6] 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.*, 19:616–643, 2008.

[7] E. Casas and M. Mateos. Second order optimality conditions for semilinear elliptic control problems with finitely many state constraints. *SIAM J. Control Optim.*, 5(40):1431–1454, 2002.

[8] 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.*, 48:688–718, 2009.

[9] E. Casas and F. Tröltzsch. Second order analysis for optimal control problems: Improving results expected from abstract theory. *SIAM J. on Optimization*, 22:261–279, 2012.

[10] E. Casas and F. Tröltzsch. Second order optimality conditions and their role in PDE control. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 117(1):3–44, 2015.

[11] E. Casas, F. Tröltzsch, and A. Unger. Second order sufficient optimality conditions for a nonlinear elliptic control problem. *ZAA*, 15:687–707, 1996.

[12] K. Disser, H.-C. Kaiser, and J. Rehberg. Optimal Sobolev regularity for linear second-order divergence elliptic operators occuring in real-world problems. *SIAM J. Math. Anal.*, 47:1719–1746, 2015.

[13] K. Disser, A.F.M. ter Elst, and J. Rehberg. Hölder estimates for parabolic operators on domains with rough boundary. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, XVII(5):65–79, 2017.

[14] G. Dore. L^p-regularity for abstract differential equations. In *Functional Analysis and Related Topics*, volume 1540 of *Lect. Notes Math.*, pages 25–38. Springer, Berlin, 1993. Proc. of the international conference in memory of K. Yosida, Kyoto University, Japan, 1991.

[15] H. Goldberg and F. Tröltzsch. Second order sufficient optimality conditions for a class of non-linear parabolic boundary control problems. *SIAM J. Control Optim.*, 31:1007–1025, 1993.

[16] J. A. Griepentrog, Gröger, H.-C. K., Kaiser, and J. Rehberg. Interpolation for function spaces related to mixed boundary value problems. *Mathematische Nachrichten*, 241:110–120, 2002.

[17] 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.

[18] D. Henry. *Geometric Theory of Semilinear Parabolic Equations.* Springer Verlag, Berlin, 1981.

[19] M. Hintermüller and I. Kopacka. Mathematical programs with complementarity constraints in function space: C- and strong stationarity and a path-following algorithm. *SIAM Journal on Optimization*, 20(2):868–902, 2009.

[20] D. Kinderlehrer and G. Stampacchia. *An introduction to variational inequalities and their applications.* SIAM, Philadelphia, 2000.

[21] K. Krumbiegel and J. Rehberg. Second order sufficient optimality conditions for parabolic optimal control problems with pointwise state constraints. *SIAM J. Control Optim*, 51:304–331, 2013.

[22] K. Kunisch and D. Wachsmuth. Sufficient optimality conditions and semi-smooth Newton methods for optimal control of stationary variational inequalities. *ESAIM: Control, Optimisation and Calculus of Variations*, 18(2):520–547, 2012.

[23] V.G. Maz'ya. *Sobolev Spaces.* Springer, Berlin, 1985.

[24] C. Meyer and L.M. Susu. Optimal control of nonsmooth, semilinear parabolic equations. *SIAM Journal on Control and Optimization*, 55(4):2206–2234, 2017.

[25] F Mignot. Contrôle dans les inéquations variationelles elliptiques. *Journal of Functional Analysis*, 22(2):130–185, 1976.

[26] A. Pazy. *Semigroups of Linear Operators and Application to Partial Differential Equations.* Springer, New York, 1983.

[27] J.-P. Raymond and F. Tröltzsch. Second order sufficient optimality conditions for nonlinear parabolic control problems with state constraints. *Discrete and Continuous Dynamical Systems*, 6:431–450, 2000.

[28] J. P. Raymond and H. Zidani. Hamiltonian Pontryagin's principles for control problems governed by semilinear parabolic equations. *Appl. Math. Optim.*, 39:143–177, 1999.

[29] A. Rösch and F. Tröltzsch. Sufficient second-order optimality conditions for a parabolic optimal control problem with pointwise control-state constraints. *SIAM J. Control Optim.*, 1(42):138–154, 2003.

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

[31] H. Triebel. *Interpolation Theory, Function Spaces, Differential Operators*. North-Holland, Amsterdam, 1978.

[32] 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.

[33] D. Wachsmuth. The regularity of the positive part of functions in $L^2(I; H^1(\Omega)) \cap H^1(I; H^1(\Omega)^*)$ with applications to parabolic equations. 57:327–332, 2016.

[34] J. Xin. *An introduction to fronts in random media*, volume 5 of *Surveys and Tutorials in the Applied Mathematical Sciences*. Springer, New York, 2009.