

# NO-GAP SECOND-ORDER OPTIMALITY CONDITIONS FOR OPTIMAL CONTROL OF A NON-SMOOTH QUASILINEAR ELLIPTIC EQUATION

Christian Clason\*

Vu Huu Nhu<sup>†</sup>

Arnd Rösch<sup>‡</sup>

**Abstract** This paper deals with second-order optimality conditions for a quasilinear elliptic control problem with a nonlinear coefficient in the principal part that is finitely  $PC^2$  (continuous and  $C^2$  apart from finitely many points). We prove that the control-to-state operator is continuously differentiable even though the nonlinear coefficient is non-smooth. This enables us to establish “no-gap” second-order necessary and sufficient optimality conditions in terms of an abstract curvature functional, i.e., for which the sufficient condition only differs from the necessary one in the fact that the inequality is strict. A condition that is equivalent to the second-order sufficient optimality condition and could be useful for error estimates in, e.g., finite element discretizations is also provided.

## 1 INTRODUCTION

This work is concerned with the quasilinear elliptic optimal control problem

$$(P) \quad \begin{cases} \min_{u \in L^\infty(\Omega), y \in H_0^1(\Omega)} J(y, u) := G(y) + \frac{\nu}{2} \|u\|_{L^2(\Omega)}^2 \\ \text{s.t.} \quad -\operatorname{div}[(b + a(y))\nabla y] = u \text{ in } \Omega, \quad y = 0 \text{ on } \partial\Omega, \\ \alpha(x) \leq u(x) \leq \beta(x) \quad \text{a.e. } x \in \Omega, \end{cases}$$

with a  $C^2$ -functional  $G : H_0^1(\Omega) \rightarrow \mathbb{R}$  for a bounded convex domain  $\Omega \subset \mathbb{R}^N$ ,  $N \in \{2, 3\}$ , a Lipschitz continuous function  $b : \bar{\Omega} \rightarrow \mathbb{R}$ , a continuous and piecewise twice differentiable function  $a : \mathbb{R} \rightarrow \mathbb{R}$ , functions  $\alpha, \beta \in L^\infty(\Omega)$  satisfying  $\beta(x) - \alpha(x) \geq \gamma$  for some  $\gamma > 0$  and almost every  $x \in \Omega$ , and a positive constant  $\nu$ . For the precise assumptions on the data of (P), we refer to [Section 2](#).

The state equation in the optimal control problem (P) arises, for instance, in models of heat conduction with a nonlinear dependence on the temperature  $y$  that allows for different behavior in different temperature regimes with sharp phase transitions. Such situations occur, e.g., in the context of steel production, where the thermal conductivity does not change spatially but rather depends on the

---

This work was supported by the DFG under the grants CL 487/2-1 and RO 2462/6-1, both within the priority programme SPP 1962 “Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization”.

\*Faculty of Mathematics, University of Duisburg-Essen, Thea-Leymann-Strasse 9, 45127 Essen, Germany ([christian.clason@uni-due.de](mailto:christian.clason@uni-due.de), ORCID: [0000-0002-9948-842](https://orcid.org/0000-0002-9948-842))

<sup>†</sup>Faculty of Mathematics, University of Duisburg-Essen, Thea-Leymann-Strasse 9, 45127 Essen, Germany; *currently* Department of Scientific Fundamentals, Posts and Telecommunications Institute of Technology, Hanoi, Vietnam ([hvu.vu@uni-due.de](mailto:hvu.vu@uni-due.de), ORCID: [0000-0003-4279-3937](https://orcid.org/0000-0003-4279-3937))

<sup>‡</sup>Faculty of Mathematics, University of Duisburg-Essen, Thea-Leymann-Strasse 9, 45127 Essen, Germany ([arnd.roesch@uni-due.de](mailto:arnd.roesch@uni-due.de))

temperature; cf. [6, 22, 29]. When the conductivity coefficient is of class  $C^2$  in the state variable  $y$ , second-order necessary and sufficient conditions for such optimal control problems were already obtained in [14, 12]. However, if the coefficient is non-smooth, the standard tools for smooth problems are typically inapplicable, making the analytical and numerical treatment challenging. The goal of this paper is therefore to derive second-order necessary and sufficient optimality conditions for (P). Specifically, we introduce a curvature functional in terms of the jumps of the first-order derivatives of  $a$  in critical points (see Section 5.2). Using a second-order Taylor-type expansion and estimates of the functional in terms of the jumps of  $a'$  in the optimal state, we derive second-order necessary as well as sufficient conditions that are “no-gap” conditions in the sense that the only difference between necessary and sufficient conditions are in the fact that the inequality in the latter are strict (see Theorems 5.9 and 5.10). In addition, we derive an equivalent formulation of the second-order sufficient condition useful for proving error estimates for finite element discretizations of (P), which will be the focus of a follow-up work.

Let us comment on related work. As far as second-order sufficient optimality conditions (SSC) are concerned, there is a rich literature on SSC for smooth PDE constrained optimal control problems; see, e.g., the articles [13, 12, 10, 27, 32, 31], the seminal book [34], the survey [15], as well as the references therein.

Regarding second-order necessary optimality conditions (SNC) for optimization problems, it is well-known that the second-order derivative of a Lagrangian function (or of the reduced cost functional) is, in general, not less than a so-called “sigma-term” [9, Chap. 3]. This term is defined as the value of a support functional of a second-order tangent set and contributes prominently in the gap between SNC and SSC. If, in addition, the optimization problem satisfies the *polyhedricity* condition, then the “sigma-term” and hence the gap vanishes; see, e.g., [9, Prop. 3.53]. For related works for smooth semilinear PDE-constrained problems, we refer to [2, 26, 8, 3, 7] as well as the references therein for  $C^2$  coefficients, while [35] treats the case where the nonlinearities are of class  $C^1$ , but not  $C^2$ , and second-order sequentially directionally differentiable. In particular, [7] is to the best of our knowledge the first work deriving no-gap second order conditions for optimal control of PDEs with polyhedral constraint sets. Another approach to deal with SNC for problems governed by smooth quasilinear elliptic equations was followed in [14, 12]. There, the non-negativity of the second-order derivatives of auxiliary real functions at minimum points (see [14, pp. 710]) was employed to derive SNC that have a minimal gap in comparison with the corresponding SSC [14, Rem. 5.3.1]. Interestingly, an inspection shows that the problems considered in [14, 12] fulfill the polyhedricity condition.

However, there are comparatively few contributions on SSC for optimal control problems governed by non-smooth PDEs, and even less on SNC for such problems. In the literature, a common approach pursued in, e.g., [28, 4, 18] is to exploit a strong stationarity condition to obtain a second-order Taylor-type expansion of the mapping  $u \mapsto J(S(u), u)$ , where  $S$  is the control-to-state operator. In these papers, SSC are derived using an additional sign assumption on the Lagrange multipliers in the vicinity of the contact set that ensures a so-called “safety distance” [5, Rem. 4.13]. In contrast, here we can use that the gradient term  $\nabla \bar{y}$  occurring in the Taylor-type expansion (5.8) vanishes on the “active set” where the coefficient is non-differentiable (due to the finiteness assumption on the set of non-differentiability points of  $a$ ). The benefit of following this approach is that we do not need any sign assumption on the multipliers. A related approach of deriving no-gap second order-conditions for non-smooth problems in terms of a generalized curvature functional was introduced in [17, 18].

Finally, we mention that a generalized version of problem (P) was studied in [20] to derive the Clarke, Bouligand, and strong stationarity conditions for the case where the coefficient  $a$  is merely directionally differentiable and locally Lipschitz continuous. In [20], we proved the equivalence of Clarke and strong stationarity conditions when the function  $a$  is countably  $PC^1$  (continuous and  $C^1$  apart from countably many points). More interestingly, as we will see later, under the assumption that

$a$  is finitely  $PC^1$  the control-to-state operator is, indeed, first-order continuously differentiable because of the finiteness assumption and the occurrence of the divergence term in the linearized equation, cf. [Theorem 3.5](#) below.

The paper is organized as follows. [Section 2](#) is devoted to notation and the main assumptions of [\(P\)](#). In [Section 3](#), we provide some required properties of the state equation, where we present some results on the existence, uniqueness, and regularity of solutions and prove first-order differentiability of the control-to-state operator. [Section 4](#) is concerned with the existence of minimizers as well as first-order necessary optimality conditions of [\(P\)](#). The main results of the paper, the no-gap second-order necessary and sufficient conditions, are derived in [Section 5](#). Finally, the paper ends with appendices showing an a priori estimate for [\(P\)](#) on convex domains and verifying a central assumption on a jump functional for a one-dimensional example.

## 2 NOTATION AND MAIN ASSUMPTIONS

**Notation.** For a given point  $u \in X$  and  $\rho > 0$ , we denote by  $B_X(u, \rho)$  and  $\bar{B}_X(u, \rho)$  the open and closed balls, respectively, of radius  $\rho$  centered at  $u$ . For Banach spaces  $X$  and  $Y$ , the notation  $X \hookrightarrow Y$  means that  $X$  is continuously embedded in  $Y$ , and  $X \Subset Y$  means that  $X$  is compactly embedded in  $Y$ . For a Banach space  $X$  with dual  $X^*$ , the symbol  $\langle \cdot, \cdot \rangle_{X^*, X}$  denotes the duality pairing between  $X$  and  $X^*$ . For a function  $f : \Omega \rightarrow \mathbb{R}$  defined on a domain  $\Omega \subset \mathbb{R}^N$  and a subset  $A \subset \mathbb{R}$ , we denote by  $\{f \in A\}$  the set of all points  $x \in \Omega$  for which  $f(x) \in A$ . Similarly, for functions  $f_1, f_2$  and subsets  $A_1, A_2 \subset \mathbb{R}$ , the symbol  $\{f_1 \in A_1, f_2 \in A_2\}$  denotes the set of all points at which the values of  $f_1$  and  $f_2$  belonging to  $A_1$  and  $A_2$ , respectively. For any set  $\omega \subset \Omega$ , the symbol  $\mathbb{1}_\omega$  stands for the indicator function of  $\omega$ , i.e.,  $\mathbb{1}_\omega(x) = 1$  if  $x \in \omega$  and  $\mathbb{1}_\omega(x) = 0$  otherwise. Finally,  $C$  denotes a generic positive constant, which may be different at different places of occurrence. We also write, e.g.,  $C_\xi$  for a constant depending only on the parameter  $\xi$ .

We recall the following definition from, e.g., [\[33, Chap. 4\]](#) or [\[36, Def. 2.19\]](#). For an open subset  $O$  in  $\mathbb{R}$ , we say that a continuous function  $f : O \rightarrow \mathbb{R}$  is a  $PC^k$ -function,  $1 \leq k \leq \infty$ , if for each point  $t_0 \in O$  there exist a neighborhood  $O_{t_0} \subset O$  and a finite set of  $C^k$ -functions  $f_i : O_{t_0} \rightarrow \mathbb{R}$ ,  $i = 1, 2, \dots, m$ , such that

$$f(t) \in \{f_1(t), f_2(t), \dots, f_m(t)\} \quad \text{for all } t \in O_{t_0}.$$

This implies in particular that  $f$  is locally Lipschitz continuous; see, e.g., [\[33, Cor. 4.1.1\]](#). For a  $PC^k$ -function  $f : O \rightarrow \mathbb{R}$ ,  $1 \leq k \leq \infty$ , we can thus define the exceptional set

$$E_f := \{t \in O \mid f \text{ is not differentiable at } t\},$$

which by Rademacher's Theorem has Lebesgue measure zero. We shall say that a  $PC^k$ -function  $f$  is *finitely (countably)  $PC^k$*  if the set  $E_f$  is finite (countable); see e.g., [\[20\]](#).

**Example 2.1.** The functions  $\mathbb{R} \ni t \mapsto |t| \in \mathbb{R}$ ,  $\mathbb{R} \ni t \mapsto \max\{0, t\} \in \mathbb{R}$ , and  $\mathbb{R} \ni t \mapsto \min\{0, t\} \in \mathbb{R}$  are finitely  $PC^\infty$ .

Let  $f$  be a finitely  $PC^1$ -function on  $\mathbb{R}$  such that the set  $E_f$  is given as

$$E_f = \{t_1, t_2, \dots, t_K\} \quad \text{with } -\infty < t_1 < t_2 < \dots < t_K < \infty \text{ and } K \in \mathbb{N}.$$

For convenience, set  $t_0 := -\infty$  and  $t_{K+1} := \infty$ . By the decomposition theorem for piecewise smooth functions [\[21, Prop. 2D.7\]](#),  $f$  can then be expressed as

$$(2.1) \quad f(t) = \sum_{i=0}^K \mathbb{1}_{(t_i, t_{i+1}]}(t) f_i(t) \quad \text{for all } t \in \mathbb{R},$$

where  $f_i$ ,  $0 \leq i \leq K$ , are  $C^1$ -functions on  $\mathbb{R}$  such that

$$(2.2) \quad f_{i-1}(t_i) = f_i(t_i) \quad \text{for all } 1 \leq i \leq K.$$

Here and in what follows, we use the convention  $(t, +\infty] := (t, \infty)$ . For any  $i \in \{1, 2, \dots, K\}$ , we set

$$(2.3) \quad \sigma_i := |f'_{i-1}(t_i) - f'_i(t_i)|.$$

Minding (2.1), this term measures the jump of the derivative of  $f$  in the singular point  $t_i$  and will play an important part in the second-order optimality conditions for (P).

It is easy to see that if  $f$  is a finitely  $PC^1$ -function defined by (2.1), then it is directionally differentiable and its directional derivative is given by

$$f'(t; h) = \sum_{i=0}^K \left\{ \mathbb{1}_{(t_i, t_{i+1})}(t) f'_i(t) h + \mathbb{1}_{\{t_{i+1}\}}(t) \left[ \mathbb{1}_{(0, \infty)}(h) f'_{i+1}(t_{i+1}) h + \mathbb{1}_{(-\infty, 0)}(h) f'_i(t_{i+1}) h \right] \right\},$$

where we use the convention  $\mathbb{1}_{\{t_{K+1}\}} = \mathbb{1}_{\{\infty\}} = 0$ .

Throughout the paper, we need the following assumptions.

(A1)  $\Omega \subset \mathbb{R}^N$ ,  $N \in \{2, 3\}$ , is an open, convex, and bounded domain.

(A2) The function  $b : \bar{\Omega} \rightarrow \mathbb{R}$  is Lipschitz continuous with Lipschitz constant  $L_b > 0$  and satisfies

$$b(x) \geq \underline{b} > 0$$

for all  $x \in \bar{\Omega}$  and for some constant  $\underline{b}$ .

(A3)  $a : \mathbb{R} \rightarrow \mathbb{R}$  is a non-negative finitely  $PC^2$  and is defined by (2.1) with  $C^2$  non-negative functions  $a_i$  satisfying (2.2) and numbers  $t_i \in \mathbb{R}$ ,  $i \in \{1, 2, \dots, K\}$ ,  $t_0 := -\infty$ ,  $t_{K+1} := +\infty$ .

(A4) The functional  $G : H_0^1(\Omega) \rightarrow \mathbb{R}$  is of class  $C^2$ .

For any  $y \in C(\bar{\Omega})$ , we then define

$$(2.4) \quad I_y := \left\{ i \in \mathbb{N} \mid \exists x \in \bar{\Omega} \text{ such that } y(x) \in (t_i, t_{i+1}] \right\}.$$

Obviously, we then have

$$a(y(x)) = \sum_{i \in I_y} \mathbb{1}_{(t_i, t_{i+1}]}(y(x)) a_i(y(x)) \quad \text{for all } x \in \bar{\Omega}.$$

Furthermore, since the  $a_i$  are  $C^2$  and therefore Lipschitz continuous on bounded sets,  $a$  is also Lipschitz continuous on bounded sets (where by the assumption only finitely many selections  $a_i$  can be attained).

### 3 CONTROL-TO-STATE OPERATOR

In this section, we shall derive the required results for the state equation

$$(3.1) \quad \begin{cases} -\operatorname{div}[(b + a(y))\nabla y] = u & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega. \end{cases}$$

### 3.1 EXISTENCE, UNIQUENESS, AND REGULARITY OF SOLUTIONS TO THE STATE EQUATION

We first address the existence, uniqueness, and regularity of solutions to (3.1).

**Theorem 3.1** (cf. [14, Thm. 2.4 and Thm. 2.5]). *Let  $p, q > N$  be arbitrary. Let Assumptions (A1) to (A3) hold. Then, for any  $u \in W^{-1,p}(\Omega)$ , there exists a unique solution  $y_u \in W_0^{1,p}(\Omega)$  to (3.1). Moreover, for any bounded set  $U \in W^{-1,p}(\Omega)$ , a constant  $C_U$  exists such that*

$$(3.2) \quad \|y_u\|_{W_0^{1,p}(\Omega)} \leq C_U \quad \text{for all } u \in U.$$

Moreover, if  $U$  is a bounded set in  $L^q(\Omega)$ , then, for any  $u \in U$ , there holds that  $y_u \in H^2(\Omega) \cap W^{1,\infty}(\Omega)$  and

$$(3.3) \quad \|y_u\|_{H^2(\Omega)} + \|y_u\|_{W^{1,\infty}(\Omega)} \leq C_U.$$

*Proof.* Let  $p > N$ ,  $U$  be a bounded set in  $W^{-1,p}(\Omega)$ , and  $u \in U$  be arbitrary. By [14, Thm. 2.2], (3.1) has a unique solution  $y_u \in H_0^1(\Omega) \cap C(\bar{\Omega})$ , and there exists a constant  $C_{1,U}$  such that

$$(3.4) \quad \|y_u\|_{H_0^1(\Omega)} + \|y_u\|_{C(\bar{\Omega})} \leq C_{1,U} \quad \text{for all } u \in U.$$

We now use the Kirchhoff transformation (see [37, Chap. V])

$$(3.5) \quad K(x, t) := b(x)t + \int_0^t a(s) ds.$$

By setting  $\theta_u(x) := K(x, y_u(x))$  for  $x \in \bar{\Omega}$ , (3.1) can be rewritten as follows

$$(3.6) \quad \begin{cases} -\Delta \theta_u = u - \operatorname{div}(\nabla b y_u) & \text{in } \Omega, \\ \theta_u = 0 & \text{on } \partial\Omega. \end{cases}$$

Applying the maximal elliptic regularity for Poisson's equation on the convex domain (see, e.g., [23, Cor. 1]) to (3.6) yields that

$$\|\theta_u\|_{W_0^{1,p}(\Omega)} \leq C_{\Omega,p,N} \|u - \operatorname{div}(\nabla b y_u)\|_{W^{-1,p}(\Omega)}.$$

This, together with (3.4) and the global Lipschitz continuity of  $b$ , gives

$$(3.7) \quad \|\theta_u\|_{W_0^{1,p}(\Omega)} \leq C_{2,U} \quad \text{for all } u \in U$$

and for some constant  $C_{2,U}$ . Moreover, for any fixed  $x \in \Omega$ ,  $K(x, \cdot)$  is monotonically increasing due to Assumptions (A2) and (A3). It then has an inverse denoted by  $T(x, \cdot)$ . By a simple computation, we have for all  $1 \leq i \leq N$  that

$$\left| \frac{\partial y_u}{\partial x_i} \right| = \left| \frac{\partial T}{\partial x_i} + \frac{\partial T}{\partial s} \frac{\partial \theta_u}{\partial x_i} \right| \leq \frac{1}{\underline{b}} \left[ \left| \frac{\partial b}{\partial x_i} \right| |y_u| + \left| \frac{\partial \theta_u}{\partial x_i} \right| \right].$$

From this, (3.7), and (3.4), we derive (3.2).

It remains to prove (3.3). To this end, let  $U$  be a bounded set in  $L^q(\Omega)$  with  $q > N$  and take  $u \in U$  arbitrary but fixed. Since  $q > N$ , we have the continuous embedding  $L^q(\Omega) \hookrightarrow W^{-1,2q}(\Omega)$ . This gives  $y_u \in W^{1,2q}(\Omega)$ . We thus have the  $H^2$ - and  $W^{1,\infty}$ -regularity of  $y_u$  as well as (3.3) according to (3.4) and Lemma A.1.  $\square$

From now on, for each  $u \in W^{-1,p}(\Omega)$ ,  $p > N$ , we denote by  $y_u$  the unique solution to (3.1). The control-to-state operator  $W^{-1,p}(\Omega) \ni u \mapsto y_u \in W_0^{1,p}(\Omega)$  is denoted by  $S$ , which is uniformly bounded by Theorem 3.1.

### 3.2 DIFFERENTIABILITY OF THE CONTROL-TO-STATE OPERATOR

We now prove the first-order differentiability of the control-to-state operator even for the non-differentiable coefficient  $a$ . To this end, we will employ the differentiability of the implicit mapping [38, Thm. 2.1], which is a generalized version of the classical implicit function theorem [40, Chap. 4] and applies to a class of quasilinear PDEs.

We first derive the locally Lipschitz continuity of the control-to-state mapping  $S$ .

**Lemma 3.2.** *Let  $p > N$  and  $u \in W^{-1,p}(\Omega)$  be arbitrary. Let Assumptions (A1) to (A3) hold. Then the operator  $S$  is locally Lipschitz continuous at  $u$  as a function from  $W^{-1,p}(\Omega)$  to  $W_0^{1,p}(\Omega)$ . Moreover, for any bounded set  $U$  in  $W^{-1,p}(\Omega)$ , there exists a constant  $L_U$  such that*

$$(3.8) \quad \|S(u_1) - S(u_2)\|_{W_0^{1,p}(\Omega)} \leq L_U \|u_1 - u_2\|_{W^{-1,p}(\Omega)} \quad \text{for all } u_1, u_2 \in U.$$

*Proof.* It is enough to prove (3.8). Let  $u_1, u_2 \in U$  be arbitrary and set  $y_i := S(u_i)$  and  $\theta_i(x) := K(x, y_i(x))$ ,  $i = 1, 2$ , with  $K$  defined in (3.5). Similar to (3.6), we have

$$(3.9) \quad \begin{cases} -\Delta(\theta_1 - \theta_2) = u_1 - u_2 - \operatorname{div}[\nabla b(y_1 - y_2)] & \text{in } \Omega, \\ \theta_1 - \theta_2 = 0 & \text{on } \partial\Omega. \end{cases}$$

Applying [23, Cor. 1] to (3.9) and using the fact that  $\|\nabla b\|_{L^\infty(\Omega)} \leq L_b$  yields

$$(3.10) \quad \|\theta_1 - \theta_2\|_{W_0^{1,p}(\Omega)} \leq C_{\Omega,p,N} [\|u_1 - u_2\|_{W^{-1,p}(\Omega)} + L_b \|y_1 - y_2\|_{L^p(\Omega)}].$$

By the definition of  $\theta_i$ ,  $i = 1, 2$ , it follows for all  $x \in \bar{\Omega}$  that

$$\theta_1(x) - \theta_2(x) = b(x)(y_1(x) - y_2(x)) + \int_{y_2(x)}^{y_1(x)} a(s) \, ds.$$

From this and a straightforward computation, we derive for all  $1 \leq m \leq N$  that

$$(3.11) \quad \frac{\partial}{\partial x_m}(y_1 - y_2) = \frac{1}{b + a(y_1)} \left[ \frac{\partial}{\partial x_m}(\theta_1 - \theta_2) - \frac{\partial b}{\partial x_m}(y_1 - y_2) - (a(y_1) - a(y_2)) \frac{\partial y_2}{\partial x_m} \right].$$

For almost every  $x \in \Omega$ , since  $K(x, \cdot)$  is monotonically increasing, so is its inverse  $T(x, \cdot)$ . This implies for almost every  $x \in \Omega$  that  $\theta_1(x) \geq \theta_2(x)$  if and only if  $y_1(x) \geq y_2(x)$ . Consequently, we obtain

$$|\theta_1(x) - \theta_2(x)| = b(x)|y_1(x) - y_2(x)| + \left| \int_{y_2(x)}^{y_1(x)} a(s) \, ds \right| \geq \underline{b}|y_1(x) - y_2(x)|.$$

From this, the continuous embedding  $W_0^{1,p}(\Omega) \hookrightarrow C(\bar{\Omega})$ , and (3.10), there holds

$$(3.12) \quad \|y_1 - y_2\|_{L^\infty(\Omega)} \leq C_{\Omega,p,N,L_b,\underline{b}} [\|u_1 - u_2\|_{W^{-1,p}(\Omega)} + \|y_1 - y_2\|_{L^p(\Omega)}].$$

Furthermore, as a result of Theorem 3.1 and the continuous embedding  $W_0^{1,p}(\Omega) \hookrightarrow C(\bar{\Omega})$ , there exists a constant  $C_U > 0$  such that

$$(3.13) \quad \|y_i\|_{C(\bar{\Omega})}, \|y_i\|_{W_0^{1,p}(\Omega)} \leq C_{1,U} \quad \text{for } i = 1, 2.$$

The combination of (3.11) with (3.10), (3.12), (3.13), and Assumptions (A2) and (A3) implies that

$$\|y_1 - y_2\|_{W_0^{1,p}(\Omega)} \leq C_{2,U} [\|u_1 - u_2\|_{W^{-1,p}(\Omega)} + \|y_1 - y_2\|_{L^p(\Omega)}].$$

Combining this with Young's inequality and the continuous embedding  $W_0^{1,p}(\Omega) \hookrightarrow L^\infty(\Omega)$ , there holds for all  $\varepsilon > 0$  that

$$\begin{aligned} \|y_1 - y_2\|_{W_0^{1,p}(\Omega)} &\leq C_{2,U} \left[ \|u_1 - u_2\|_{W^{-1,p}(\Omega)} + \|y_1 - y_2\|_{L^\infty(\Omega)}^{(p-1)/p} \|y_1 - y_2\|_{L^1(\Omega)}^{1/p} \right] \\ &\leq C_{2,U} \left[ \|u_1 - u_2\|_{W^{-1,p}(\Omega)} + \varepsilon^{p/(p-1)} \|y_1 - y_2\|_{L^\infty(\Omega)} + \frac{1}{\varepsilon^p} \|y_1 - y_2\|_{L^1(\Omega)} \right] \\ &\leq C_{2,U} \left[ \|u_1 - u_2\|_{W^{-1,p}(\Omega)} + \varepsilon^{p/(p-1)} \|y_1 - y_2\|_{W_0^{1,p}(\Omega)} + \frac{1}{\varepsilon^p} \|y_1 - y_2\|_{L^1(\Omega)} \right]. \end{aligned}$$

By choosing  $\varepsilon = \varepsilon(p, C_{2,U}) > 0$  small enough, we arrive at

$$(3.14) \quad \|y_1 - y_2\|_{W_0^{1,p}(\Omega)} \leq C_U \left[ \|u_1 - u_2\|_{W^{-1,p}(\Omega)} + \|y_1 - y_2\|_{L^1(\Omega)} \right].$$

We now show that there is a constant  $L_U$  satisfying

$$(3.15) \quad \|y_1 - y_2\|_{L^{2p/(p-2)}(\Omega)} \leq L_U \|u_1 - u_2\|_{W^{-1,p}(\Omega)} \quad \text{for all } u_1, u_2 \in U,$$

which, together with (3.14), gives the desired conclusion. Assume to the contrary that (3.15) does not hold. Then we can find  $u_1^{(n)}, u_2^{(n)} \in U$  such that

$$\frac{1}{\eta_n} \|y_1^{(n)} - y_2^{(n)}\|_{L^{2p/(p-2)}(\Omega)} \rightarrow +\infty$$

with  $\eta_n := \|u_1^{(n)} - u_2^{(n)}\|_{W^{-1,p}(\Omega)}$  and  $y_i^{(n)} := S(u_i^{(n)})$ ,  $i = 1, 2$ . Obviously,  $\eta_n \rightarrow 0$  as  $n \rightarrow \infty$ . We now define a scalar function  $a_n$  on  $\Omega$  and a vector-valued function  $\mathbb{b}_n$  on  $\Omega$  by

$$a_n(x) := b(x) + a(y_1^{(n)}(x)), \quad \mathbb{b}_n(x) := \nabla y_2^{(n)}(x) \mathbb{1}_{\{y_1^{(n)} \neq y_2^{(n)}\}}(x) \frac{a(y_1^{(n)}(x)) - a(y_2^{(n)}(x))}{y_1^{(n)}(x) - y_2^{(n)}(x)}.$$

As a result of (3.2), a constant  $C_U$  exists such that

$$(3.16) \quad \|y_i^{(n)}\|_{W_0^{1,p}(\Omega)}, \|\mathbb{b}_n\|_{L^p(\Omega)} \leq C_U \quad \text{for all } n \in \mathbb{N}, i = 1, 2.$$

Setting  $w_n := y_1^{(n)} - y_2^{(n)}$  yields

$$\begin{cases} -\operatorname{div}[a_n \nabla w_n + \mathbb{b}_n w_n] = u_1^{(n)} - u_2^{(n)} & \text{in } \Omega, \\ w_n = 0 & \text{on } \partial\Omega. \end{cases}$$

Setting

$$\rho_n := \frac{\eta_n}{\|y_1^{(n)} - y_2^{(n)}\|_{L^{2p/(p-2)}(\Omega)}}, \quad \xi_n := \frac{\rho_n}{\eta_n} w_n, \quad h_n := \frac{\rho_n}{\eta_n} \left[ u_1^{(n)} - u_2^{(n)} \right],$$

we see that  $\rho_n \rightarrow 0$  and that  $\xi_n$  solves

$$(3.17) \quad \begin{cases} -\operatorname{div}[a_n \nabla \xi_n + \mathbb{b}_n \xi_n] = h_n & \text{in } \Omega, \\ \xi_n = 0 & \text{on } \partial\Omega. \end{cases}$$

Testing the above equation by  $\xi_n$  and employing the Hölder inequality yield

$$\underline{b} \|\nabla \xi_n\|_{L^2(\Omega)}^2 \leq \|h_n\|_{H^{-1}(\Omega)} \|\xi_n\|_{H_0^1(\Omega)} + \|\mathbb{b}_n\|_{L^p(\Omega)} \|\xi_n\|_{L^{2p/(p-2)}(\Omega)} \|\nabla \xi_n\|_{L^2(\Omega)},$$

which, together with (3.16) and the fact that  $\|\xi_n\|_{H_0^1(\Omega)} \leq C\|\nabla\xi_n\|_{L^2(\Omega)}$ , gives

$$(3.18) \quad \begin{aligned} \underline{b}\|\nabla\xi_n\|_{L^2(\Omega)} &\leq \|h_n\|_{H^{-1}(\Omega)} + C\|\mathbb{b}_n\|_{L^p(\Omega)}\|\xi_n\|_{L^{2p/(p-2)}} \\ &\leq C\|h_n\|_{W^{-1,p}(\Omega)} + C_U \\ &= C\rho_n + C_U \\ &\leq C + C_U \end{aligned}$$

for  $n$  large enough. From this and the compact embedding  $H_0^1(\Omega) \Subset L^{2p/(p-2)}(\Omega)$ , we can assume that  $\xi_n \rightharpoonup \xi$  in  $H_0^1(\Omega)$  and  $\xi_n \rightarrow \xi$  in  $L^{2p/(p-2)}(\Omega)$  for some  $\xi \in H_0^1(\Omega)$ . Moreover, there exist subsequences of  $\{y_1^{(n)}\}$  and  $\{\mathbb{b}_n\}$ , denoted in the same way, such that  $y_1^{(n)} \rightarrow y_*$  in  $C(\overline{\Omega})$  and  $\mathbb{b}_n \rightarrow \mathbb{b}$  in  $L^p(\Omega)^N$  for some  $y_* \in C(\overline{\Omega})$  and  $\mathbb{b} \in L^p(\Omega)^N$ . Therefore, we have  $a_n \rightarrow a_*$  in  $C(\overline{\Omega})$  with  $a_*(x) := b(x) + a(y_*(x))$ . Passing to the limit in (3.17), we deduce from the fact that  $h_n \rightarrow 0$  in  $H^{-1}(\Omega)$  that  $\xi$  fulfills

$$\begin{cases} -\operatorname{div}[a_*\nabla\xi + \mathbb{b}\xi] = 0 & \text{in } \Omega, \\ \xi = 0 & \text{on } \partial\Omega. \end{cases}$$

The uniqueness of solutions, see, e.g., [11, Thm. 2.6], implies that  $\xi = 0$ , which contradicts the fact that  $\|\xi\|_{L^{2p/(p-2)}(\Omega)} = \lim_{n \rightarrow \infty} \|\xi_n\|_{L^{2p/(p-2)}(\Omega)} = 1$ .  $\square$

For any  $y, \hat{y} \in C(\overline{\Omega})$  and any  $\tau_1, \tau_2 \in \mathbb{R}$  we define the set

$$(3.19) \quad \Omega_{\hat{y},i,j}^{[\tau_1,\tau_2]} := \{\hat{y} \in [t_i + \tau_1, t_j + \tau_2]\},$$

where  $t_i, i \in \{0, 1, \dots, K+1\}$ , are given in Assumption (A3). Similar sets such as  $\Omega_{\hat{y},i,j}^{[\tau_1,\tau_2]}$  are defined in the same way. We also define the function  $T_{y,\hat{y}} : \Omega \rightarrow \mathbb{R}$  via

$$(3.20) \quad T_{y,\hat{y}} := \mathbb{1}_{\{\hat{y} \notin E_a\}} [a(y) - a(\hat{y}) - a'(\hat{y})(y - \hat{y})].$$

From now on, let us fix a number  $\delta \in R$  such that

$$0 < \delta \leq \frac{t_{i+1} - t_i}{2} \quad \text{for all } 1 \leq i \leq K-1.$$

We need the following lemmas.

**Lemma 3.3.** *Let Assumption (A3) be satisfied. Then, for any  $y, \hat{y} \in C(\overline{\Omega})$  with  $\|y - \hat{y}\|_{C(\overline{\Omega})} < \delta$ , there holds*

$$(3.21) \quad T_{y,\hat{y}} = \sum_{i \in I_{\hat{y}}} (T_{y,\hat{y}}^{i,1} + T_{y,\hat{y}}^{i,2} + T_{y,\hat{y}}^{i,3})$$

with  $I_{\hat{y}}$  defined via (2.4) and

$$\begin{aligned} T_{y,\hat{y}}^{i,1} &:= \mathbb{1}_{\Omega_{y,\hat{y}}^{i,1}} [a_i(y) - a_i(\hat{y}) - a'_i(\hat{y})(y - \hat{y})], \\ T_{y,\hat{y}}^{i,2} &:= \mathbb{1}_{\Omega_{y,\hat{y}}^{i,2}} [a_{i-1}(y) - a_i(\hat{y}) - a'_i(\hat{y})(y - \hat{y})], \\ T_{y,\hat{y}}^{i,3} &:= \mathbb{1}_{\Omega_{y,\hat{y}}^{i,3}} [a_{i+1}(y) - a_i(\hat{y}) - a'_i(\hat{y})(y - \hat{y})], \end{aligned}$$

where

$$\begin{aligned} \Omega_{y,\hat{y}}^{i,1} &:= \Omega_{\hat{y},i,i+1}^{[\delta,-\delta]} \cup \left( \Omega_{\hat{y},i,i}^{(0,\delta)} \cap \Omega_{y,i,i}^{(0,2\delta)} \right) \cup \left( \Omega_{\hat{y},i+1,i+1}^{(-\delta,0)} \cap \Omega_{y,i+1,i+1}^{(-2\delta,0)} \right), \\ \Omega_{y,\hat{y}}^{i,2} &:= \Omega_{\hat{y},i,i}^{(0,\delta)} \cap \Omega_{y,i,i}^{(-\delta,0]}, \quad \Omega_{y,\hat{y}}^{i,3} := \Omega_{\hat{y},i+1,i+1}^{(-\delta,0)} \cap \Omega_{y,i+1,i+1}^{[0,\delta)}. \end{aligned}$$

Moreover, if  $y_n \rightarrow \hat{y}$  in  $W_0^{1,p}(\Omega)$  with  $p > N$ , then

$$(3.22) \quad \frac{1}{\|y_n - \hat{y}\|_{W_0^{1,p}(\Omega)}} \|T_{y_n,\hat{y}} \nabla \hat{y}\|_{L^p(\Omega)} \rightarrow 0.$$

*Proof.* We have

$$(3.23) \quad \begin{aligned} T_{y,\hat{y}} &= \sum_{i=0}^K \mathbb{1}_{\Omega_{\hat{y},i,i+1}^{(0,0)}} \left[ a(y) - a_i(\hat{y}) - a'_i(\hat{y})(y - \hat{y}) \right] \\ &= \sum_{i \in I_{\hat{y}}} \left[ \mathbb{1}_{\Omega_{\hat{y},i,i}^{(0,\delta)}} + \mathbb{1}_{\Omega_{\hat{y},i,i+1}^{[\delta,-\delta]}} + \mathbb{1}_{\Omega_{\hat{y},i+1,i+1}^{(-\delta,0)}} \right] \left[ a(y) - a_i(\hat{y}) - a'_i(\hat{y})(y - \hat{y}) \right], \end{aligned}$$

using the fact that  $\mathbb{1}_{\Omega_{\hat{y},i,i+1}^{(0,0)}} \equiv 0$  for all  $i \notin I_{\hat{y}}$ . Since  $\|y - \hat{y}\|_{C(\bar{\Omega})} < \delta$ , we have

$$\begin{aligned} \Omega_{\hat{y},i,i+1}^{[\delta,-\delta]} &= \Omega_{\hat{y},i,i+1}^{[\delta,-\delta]} \cap \Omega_{y,i,i+1}^{(0,0)}, \\ \Omega_{\hat{y},i,i}^{(0,\delta)} &= \Omega_{\hat{y},i,i}^{(0,\delta)} \cap \left( \Omega_{y,i,i}^{(-\delta,0]} \cup \Omega_{y,i,i}^{(0,2\delta)} \right), \\ \Omega_{\hat{y},i+1,i+1}^{(-\delta,0)} &= \Omega_{\hat{y},i+1,i+1}^{(-\delta,0)} \cap \left( \Omega_{y,i+1,i+1}^{(-2\delta,0)} \cup \Omega_{y,i+1,i+1}^{[0,\delta]} \right). \end{aligned}$$

Together with (3.23), these yield the claimed expression.

It remains to prove the limit (3.22). Let  $y_n \rightarrow \hat{y}$  in  $W_0^{1,p}(\Omega)$  and set  $\varepsilon_n := \|y_n - \hat{y}\|_{C(\bar{\Omega})}$ . Then  $\varepsilon_n \rightarrow 0$  since  $W_0^{1,p}(\Omega) \hookrightarrow C(\bar{\Omega})$ . We therefore can assume that  $\varepsilon_n < \delta$  and  $|y_n(x)| \leq M$  for some  $M > 0$  and all  $x \in \bar{\Omega}$  for  $n \in \mathbb{N}$  sufficiently large. Using (3.21), we obtain

$$(3.24) \quad \|T_{y_n,\hat{y}} \nabla \hat{y}\|_{L^p(\Omega)} \leq \sum_{i \in I_{\hat{y}}} \left( \|T_{y_n,\hat{y}}^{i,1} \nabla \hat{y}\|_{L^p(\Omega)} + \|T_{y_n,\hat{y}}^{i,2} \nabla \hat{y}\|_{L^p(\Omega)} + \|T_{y_n,\hat{y}}^{i,3} \nabla \hat{y}\|_{L^p(\Omega)} \right).$$

From the definition of  $T_{y_n,\hat{y}}^{i,1}$ , the dominated convergence theorem yields

$$(3.25) \quad \frac{1}{\varepsilon_n} \|T_{y_n,\hat{y}}^{i,1} \nabla \hat{y}\|_{L^p(\Omega)} \rightarrow 0 \quad \text{for all } i \in I_{\hat{y}}.$$

For  $\frac{1}{\varepsilon_n} \|T_{y_n,\hat{y}}^{i,2} \nabla \hat{y}\|_{L^p(\Omega)}$ , we have

$$\begin{aligned} |a_{i-1}(y_n) - a_i(\hat{y}) - a'_i(\hat{y})(y_n - \hat{y})| &= |a_{i-1}(y_n) - a_{i-1}(t_i) + a_i(t_i) - a_i(\hat{y}) - a'_i(\hat{y})(y_n - \hat{y})| \\ &\leq L_i (|y_n - t_i| + |t_i - \hat{y}| + |y_n - \hat{y}|) \\ &= 2L_i |y_n - \hat{y}| \\ &\leq 2L_i \varepsilon_n \end{aligned}$$

for almost every  $x \in \Omega_{y_n,\hat{y}}^{i,2}$  and for some constant  $L_i$  depending on  $M$ . Moreover,  $\mathbb{1}_{\Omega_{y_n,\hat{y}}^{i,2}} \rightarrow 0$  almost everywhere in  $\Omega$ . Combining this with the dominated convergence theorem yields

$$(3.26) \quad \frac{1}{\varepsilon_n} \|T_{y_n,\hat{y}}^{i,2} \nabla \hat{y}\|_{L^p(\Omega)} \rightarrow 0 \quad \text{for all } i \in I_{\hat{y}}.$$

Similarly,

$$(3.27) \quad \frac{1}{\varepsilon_n} \|T_{y_n,\hat{y}}^{i,3} \nabla \hat{y}\|_{L^p(\Omega)} \rightarrow 0 \quad \text{for all } i \in I_{\hat{y}}.$$

From (3.24)–(3.27) and the fact that  $\varepsilon_n \leq C \|y_n - \hat{y}\|_{W_0^{1,p}(\Omega)}$  and that  $I_{\hat{y}}$  is finite, we obtain the limit (3.22).  $\square$

**Lemma 3.4.** *Let  $y_n \rightarrow y$  in  $W_0^{1,p}(\Omega)$  with  $p > N$ . Under Assumption (A3), there holds*

$$\mathbb{1}_{\{y_n \notin E_a\}} a'(y_n) \nabla y_n - \mathbb{1}_{\{y \notin E_a\}} a'(y) \nabla y \rightarrow 0 \quad \text{in } L^p(\Omega).$$

*Proof.* Since  $y_n \rightarrow y$  in  $W_0^{1,p}(\Omega)$ , we have  $y_n \rightarrow y$  in  $C(\overline{\Omega})$ . Also, we can assume that  $\|y_n - y\|_{C(\overline{\Omega})} < \delta$  for all  $n \in \mathbb{N}$  large enough. Setting  $A_n := \mathbb{1}_{\{y_n \notin E_a\}} a'(y_n) \nabla y_n - \mathbb{1}_{\{y \notin E_a\}} a'(y) \nabla y$ , we have

$$A_n = \sum_{i=0}^K A_n^{(i)} \quad \text{with} \quad A_n^{(i)} := \mathbb{1}_{\Omega_{y_n, i, i+1}^{(0,0)}} a'_i(y_n) \nabla y_n - \mathbb{1}_{\Omega_{y, i, i+1}^{(0,0)}} a'_i(y) \nabla y.$$

It is sufficient to prove that  $A_n^{(i)} \rightarrow 0$  in  $L^p(\Omega)$  for all  $0 \leq i \leq K$ . To this end, we write  $A_n^{(i)}$  as

$$A_n^{(i)} = \left[ \mathbb{1}_{\Omega_{y_n, i, i+1}^{(0,0)}} - \mathbb{1}_{\Omega_{y, i, i+1}^{(0,0)}} \right] a'_i(y) \nabla y + \mathbb{1}_{\Omega_{y_n, i, i+1}^{(0,0)}} [a'_i(y_n) - a'_i(y)] \nabla y + \mathbb{1}_{\Omega_{y_n, i, i+1}^{(0,0)}} a'_i(y_n) [\nabla y_n - \nabla y].$$

The last two terms in the right-hand side tend to 0 in  $L^p(\Omega)$ . For the first term in the right-hand side, we have

$$\begin{aligned} \left[ \mathbb{1}_{\Omega_{y_n, i, i+1}^{(0,0)}} - \mathbb{1}_{\Omega_{y, i, i+1}^{(0,0)}} \right] a'_i(y) \nabla y &= \mathbb{1}_{\Omega_{y_n, i, i+1}^{(0,0)} \cap [\{y=t_i\} \cup \{y=t_{i+1}\}]} a'_i(y) \nabla y \\ &\quad + \mathbb{1}_{\Omega_{y_n, i, i}^{(0,\delta)} \cap \Omega_{y, i, i}^{(-\delta,0)}} a'_i(y) \nabla y + \mathbb{1}_{\Omega_{y_n, i+1, i+1}^{(-\delta,0)} \cap \Omega_{y, i+1, i+1}^{(0,\delta)}} a'_i(y) \nabla y \\ &\quad - \mathbb{1}_{\Omega_{y_n, i, i}^{(-\delta,0)} \cap \Omega_{y, i, i}^{(0,\delta)}} a'_i(y) \nabla y - \mathbb{1}_{\Omega_{y_n, i+1, i+1}^{(0,\delta)} \cap \Omega_{y, i+1, i+1}^{(-\delta,0)}} a'_i(y) \nabla y. \end{aligned}$$

In the above expression, the first term in the right-hand side vanishes almost everywhere since  $\nabla y(x) = 0$  for almost every  $x \in \{y = t_i\} \cup \{y = t_{i+1}\}$  (see, e.g., [16, Rem. 2.6] and [1, Prop. 5.8.2]), and the other terms tend to zero in  $L^p(\Omega)$  according to the dominated convergence theorem. Thus,

$$\left[ \mathbb{1}_{\Omega_{y_n, i, i+1}^{(0,0)}} - \mathbb{1}_{\Omega_{y, i, i+1}^{(0,0)}} \right] a'_i(y) \nabla y \rightarrow 0 \quad \text{in } L^p(\Omega)$$

and hence  $A_n^{(i)} \rightarrow 0$  in  $L^p(\Omega)$ .  $\square$

As seen in the proof of Lemma 3.4, the fact that for  $y \in W^{1,p}(\Omega)$ ,  $p \geq 1$ , the gradient  $\nabla y$  vanishes almost everywhere on  $\{y \in E_a\}$ , plays an important role. Furthermore, this fact also guarantees that

$$[a(y+z) - a(y) - a'(y; z)] \nabla y = T_{y+z, y} \nabla y$$

for  $y, z \in W_0^{1,p}(\Omega)$ , which will be crucial for proving continuous differentiability of the control-to-state mapping  $S$ . We will also use this to show a key limit in Proposition 5.8 (iii) below. We point out that do *not* need to assume that the set  $\{y \in E_a\}$  has small or even zero Lebesgue measure.

**Theorem 3.5.** *Under Assumptions (A1) to (A3), the control-to-state operator  $S : W^{-1,p}(\Omega) \rightarrow W_0^{1,p}(\Omega)$ ,  $p > N$ , is Fréchet differentiable. Moreover, for any  $u, v \in W^{-1,p}(\Omega)$ ,  $z_v := S'(u)v$  satisfies*

$$(3.28) \quad \begin{cases} -\operatorname{div}[(b + a(y_u)) \nabla z_v + \mathbb{1}_{\{y_u \notin E_a\}} a'(y_u) z_v \nabla y_u] = v & \text{in } \Omega, \\ z_v = 0 & \text{on } \partial\Omega, \end{cases}$$

with  $y_u := S(u)$ . Furthermore, the mapping  $W^{-1,p}(\Omega) \ni u \mapsto S'(u) \in W_0^{1,p}(\Omega)$  is continuous.

*Proof. Step 1: existence of  $S'$ .* We shall apply the differentiability of the implicit mapping [38, Thm. 2.1]. Consider the mapping  $F : W_0^{1,p}(\Omega) \times W^{-1,p}(\Omega) \rightarrow W^{-1,p}(\Omega)$  defined by

$$F(y, u) := -\operatorname{div}[(b + a(y)) \nabla y] - u.$$

We first show that  $F$  has a partial derivative in  $y$  given by

$$\frac{\partial F}{\partial y}(y, u)z = -\operatorname{div}[(b + a(y)) \nabla z + \mathbb{1}_{\{y \notin E_a\}} a'(y) z \nabla y]$$

for all  $y, z \in W_0^{1,p}(\Omega)$  and  $u \in W^{-1,p}(\Omega)$ . To this end, let  $z_n$  be an arbitrary sequence converging to zero in  $W_0^{1,p}(\Omega)$  and let  $\varphi$  be arbitrary in  $W_0^{1,p'}(\Omega)$ ,  $p' := p/(p-1)$ , with  $\|\varphi\|_{W_0^{1,p'}(\Omega)} \leq 1$ . Setting  $y_n := y + z_n$ , we then have

$$\begin{aligned}
& \left\langle F(y_n, u) - F(y, u) - \frac{\partial F}{\partial y}(y, u)z_n, \varphi \right\rangle_{W^{-1,p}(\Omega), W_0^{1,p'}(\Omega)} \\
&= \int_{\Omega} [(b + a(y_n))\nabla y_n - (b + a(y))\nabla y \\
&\quad - (b + a(y))\nabla z_n - a'(y; z_n)\nabla y] \cdot \nabla \varphi \, dx \\
&= \int_{\Omega} [(a(y_n) - a(y))\nabla z_n + (a(y_n) - a(y) - a'(y; z_n))\nabla y] \cdot \nabla \varphi \, dx \\
&\leq \|(a(y_n) - a(y))\nabla z_n + (a(y_n) - a(y) - a'(y; z_n))\nabla y\|_{L^p(\Omega)} \\
&\leq \|a(y_n) - a(y)\|_{L^\infty(\Omega)} \|\nabla z_n\|_{L^p(\Omega)} + \|(a(y_n) - a(y) - a'(y; z_n))\nabla y\|_{L^p(\Omega)} \\
&= \|a(y_n) - a(y)\|_{L^\infty(\Omega)} \|\nabla z_n\|_{L^p(\Omega)} + \|T_{y_n, y}\nabla y\|_{L^p(\Omega)}.
\end{aligned}$$

Setting  $\varepsilon_n := \|z_n\|_{W_0^{1,p}(\Omega)}$ , we obtain

$$\frac{1}{\|z_n\|_{W_0^{1,p}(\Omega)}} \|F(y_n, u) - F(y, u) - \frac{\partial F}{\partial y}(y, u)z_n\|_{W^{-1,p}(\Omega)} \leq \|a(y_n) - a(y)\|_{L^\infty(\Omega)} + \frac{1}{\varepsilon_n} \|T_{y_n, y}\nabla y\|_{L^p(\Omega)}.$$

Obviously,  $\|a(y_n) - a(y)\|_{L^\infty(\Omega)} \rightarrow 0$ . Combining this with (3.22) yields the partial differentiability of  $F$  in  $y$ .

We now prove that  $\frac{\partial F}{\partial y}(y_u, u)$ ,  $y_u := S(u)$ , is an isomorphism as a mapping from  $W_0^{1,p}(\Omega)$  to  $W^{-1,p}(\Omega)$ . From this, the Lipschitz continuity of  $S$  (see Lemma 3.2), and [38, Thm. 2.1], we then deduce the existence of Fréchet derivative at  $u$  of  $S$  as well as (3.28). It is enough to prove for any  $v \in W^{-1,p}(\Omega)$  that there exists a unique  $z_v \in W_0^{1,p}(\Omega)$  such that

$$(3.29) \quad \frac{\partial F}{\partial y}(y_u, u)z_v = v$$

and

$$(3.30) \quad \|z_v\|_{W_0^{1,p}(\Omega)} \leq C_u \|v\|_{W^{-1,p}(\Omega)}.$$

Setting  $a_u(x) := b(x) + a(y_u(x))$  and  $c_u(x) := \mathbb{1}_{\{y_u \notin E_a\}}(x) a'(y_u(x)) \nabla y_u(x)$  for almost every  $x \in \Omega$ , we have that  $a_u$  is continuous and there hold

$$(3.31) \quad \underline{b} \leq a_u(x) \leq C_u, \quad \|c_u\|_{L^p(\Omega)} \leq C_u \quad \text{for a.e. } x \in \Omega.$$

The equation (3.29) can thus be written as

$$(3.32) \quad \begin{cases} -\operatorname{div}[a_u \nabla z_v + c_u z_v] = v & \text{in } \Omega, \\ z_v = 0 & \text{on } \partial\Omega. \end{cases}$$

Due to [11, Thm. 2.6], (3.32) (and thus (3.29)) has a unique solution  $z_v \in H_0^1(\Omega) \hookrightarrow L^{2p/(p-2)}(\Omega)$ . Similar to (3.15) and (3.18), there hold that

$$(3.33) \quad \|z_v\|_{L^{2p/(p-2)}(\Omega)} \leq C_{1,u} \|v\|_{W^{-1,p}(\Omega)}$$

and

$$\underline{b} \|\nabla z_v\|_{L^2(\Omega)} \leq C_{1,u} \left[ \|v\|_{W^{-1,p}(\Omega)} + \|z_v\|_{L^{2p/(p-2)}(\Omega)} \right]$$

for some constant  $C_{1,u}$  independent of  $v$ . It then follows that

$$(3.34) \quad \|\nabla z_v\|_{L^2(\Omega)} \leq C_{2,u} \|v\|_{W^{-1,p}(\Omega)}.$$

Moreover, by using the chain rule and the product formula [24, Chap. 7] as well as Assumption (A3), we can rewrite (3.32) as

$$(3.35) \quad \begin{cases} -\Delta \hat{z}_v = v - \operatorname{div}[z_v \nabla b] & \text{in } \Omega, \\ z_v = 0 & \text{on } \partial\Omega, \end{cases}$$

with  $\hat{z}_v := (b + a(y_u))z_v$ . We now show that

$$(3.36) \quad \|\hat{z}_v\|_{L^\infty(\Omega)} \leq C_{\Omega,p} \|v\|_{W^{-1,p}(\Omega)}.$$

To this end, we first consider the case  $p \geq 6$ . We then have  $v \in W^{-1,p}(\Omega) \hookrightarrow W^{-1,6}(\Omega)$ . Besides, we have  $z_v \in H_0^1(\Omega) \hookrightarrow L^6(\Omega)$  and thus  $z_v \nabla b \in L^6(\Omega)^N$ . It follows that the right-hand side of (3.35) belongs to  $W^{-1,6}(\Omega)$ . Applying the Stampacchia Theorem [16, Thm. 12.4] and using the continuous embedding  $W^{-1,p}(\Omega) \hookrightarrow W^{-1,6}(\Omega)$ , we can conclude that

$$\begin{aligned} \|\hat{z}_v\|_{L^\infty(\Omega)} &\leq C_\Omega \left[ \|v\|_{W^{-1,6}(\Omega)} + \|z_v \nabla b\|_{L^6(\Omega)} \right] \\ &\leq C_{\Omega,p} \left[ \|v\|_{W^{-1,p}(\Omega)} + \|\nabla b\|_{L^\infty(\Omega)} \|z_v\|_{H_0^1(\Omega)} \right] \\ &\leq C_{\Omega,p} \|v\|_{W^{-1,p}(\Omega)}, \end{aligned}$$

where we have used (3.34) to obtain the last estimate. For the case  $N < p < 6$ , we see that the right-hand side of (3.35) belongs to  $W^{-1,p}(\Omega)$ . Using the Stampacchia Theorem again and exploiting the embedding  $H_0^1(\Omega) \hookrightarrow L^6(\Omega) \hookrightarrow L^p(\Omega)$  and (3.34) yield

$$\begin{aligned} \|\hat{z}_v\|_{L^\infty(\Omega)} &\leq C_{\Omega,p} \left[ \|v\|_{W^{-1,p}(\Omega)} + \|z_v \nabla b\|_{L^p(\Omega)} \right] \\ &\leq C_{\Omega,p} \left[ \|v\|_{W^{-1,p}(\Omega)} + \|z_v \nabla b\|_{L^6(\Omega)} \right] \\ &\leq C_{\Omega,p} \|v\|_{W^{-1,p}(\Omega)}. \end{aligned}$$

We now consider two cases:

Case 1:  $N = 2$ . Since the embedding  $H_0^1(\Omega) \hookrightarrow L^p(\Omega)$  is continuous, we have  $v - \operatorname{div}[z_v \nabla b] \in W^{-1,p}(\Omega)$ . Applying [23, Cor. 1] to (3.35) and exploiting (3.34) yields that  $\hat{z}_v \in W_0^{1,p}(\Omega)$  and that

$$\|\hat{z}_v\|_{W_0^{1,p}(\Omega)} \leq C_{\Omega,p,N,u} \|v\|_{W^{-1,p}(\Omega)}.$$

From this, the definition of  $\hat{z}_v$ , and (3.36), we derive  $z_v \in W_0^{1,p}(\Omega)$  and (3.30).

Case 2:  $N = 3$ . In this case, we have  $H_0^1(\Omega) \hookrightarrow L^6(\Omega)$ . If  $p \leq 6$ , we then have the desired conclusion using a similar argument as in the first case. If  $p > 6$ , then  $v - \operatorname{div}[z_v \nabla b] \in W^{-1,6}(\Omega)$ . Similar to the first case,  $z_v \in W_0^{1,6}(\Omega)$  and  $\|z_v\|_{W_0^{1,6}(\Omega)} \leq C_u \|v\|_{W^{-1,p}(\Omega)}$ . Finally, by using the continuous embedding  $W_0^{1,6}(\Omega) \hookrightarrow L^p(\Omega)$  and a bootstrapping argument, we arrive at the desired conclusion.

*Step 2. continuity of  $S'$ .* Taking any  $v, u_n, u_0 \in W^{-1,p}(\Omega)$  such that  $\|v\|_{W^{-1,p}(\Omega)} \leq 1$  as well as  $s_n := \|u_n - u_0\|_{W^{-1,p}(\Omega)} \rightarrow 0$  as  $n \rightarrow \infty$  and setting  $z_{vn} := S'(u_n)v, z_v := S'(u_0)v$ , we see that  $z_{vn}$  and  $z_v$  satisfy

$$\begin{cases} -\operatorname{div}[(b + a(y_0))\nabla(z_{vn} - z_v) + \mathbb{1}_{\{y_0 \notin E_a\}} a'(y_0)\nabla y_0(z_{vn} - z_v)] = \operatorname{div} g_{vn} & \text{in } \Omega, \\ z_{vn} - z_v = 0 & \text{on } \partial\Omega, \end{cases}$$

with  $y_0 := S(u_0), y_n := S(u_n)$ , and

$$g_{vn} := [\nabla z_{vn}(a(y_n) - a(y_0)) + (\mathbb{1}_{\{y_n \notin E_a\}} a'(y_n)\nabla y_n - \mathbb{1}_{\{y_0 \notin E_a\}} a'(y_0)\nabla y_0) z_{vn}].$$

Similar to (3.30), a constant  $C_1 := C_{1u_0}$  exists such that

$$(3.37) \quad \|z_{vn} - z_v\|_{W_0^{1,p}(\Omega)} \leq C_1 \|g_{vn}\|_{L^p(\Omega)}.$$

Furthermore, the local Lipschitz continuity of  $S$  (Lemma 3.2) implies that there is a constant  $C_2 = C_2(u_0)$  such that

$$\|z_{vn}\|_{W_0^{1,p}(\Omega)}, \|z_v\|_{W_0^{1,p}(\Omega)} \leq C_2 \|v\|_{W^{-1,p}(\Omega)} \leq C_2 \quad \text{for all } n \in \mathbb{N}.$$

This implies that

$$(3.38) \quad \begin{aligned} \|g_{vn}\|_{L^p(\Omega)} &\leq C_2 \|a(y_n) - a(y_0)\|_{L^\infty(\Omega)} + \|z_{vn}\|_{L^\infty(\Omega)} \|A_n\|_{L^p(\Omega)} \\ &\leq C_2 (\|a(y_n) - a(y_0)\|_{L^\infty(\Omega)} + C \|A_n\|_{L^p(\Omega)}), \end{aligned}$$

where we have used the continuous embedding  $W_0^{1,p}(\Omega) \hookrightarrow L^\infty(\Omega)$  to obtain the last inequality with a constant  $C$  independent of  $v$  and  $n$ . Here

$$A_n := \mathbb{1}_{\{y_n \notin E_a\}} a'(y_n)\nabla y_n - \mathbb{1}_{\{y_0 \notin E_a\}} a'(y_0)\nabla y_0$$

does not depend on  $v$ . Combining (3.37) with (3.38) yields

$$\|S'(u_n) - S'(u_0)\|_{\mathbb{L}(W^{-1,p}(\Omega), W_0^{1,p}(\Omega))} \leq C_3 [\|a(y_n) - a(y_0)\|_{L^\infty(\Omega)} + \|A_n\|_{L^p(\Omega)}].$$

Obviously, the first term in the right-hand side converges to zero since  $y_n \rightarrow y_0$  in  $C(\overline{\Omega})$ . In addition, as a result of Lemma 3.4, the second term tends to zero.  $\square$

We end this section with a direct consequence of the differentiability of  $S$ .

**Corollary 3.6.** *Let  $u$  and  $h$  be arbitrary in  $L^2(\Omega)$  and let  $\{s_n\} \subset (0, \infty)$  and  $\{h_n\} \subset L^2(\Omega)$  be such that  $s_n \rightarrow 0^+$  and  $h_n \rightarrow h$  in  $L^2(\Omega)$ . Then, for any  $p \in (N, 6)$ , there holds*

$$\frac{S(u + s_n h_n) - S(u)}{s_n} \rightarrow S'(u)h \quad \text{in } W_0^{1,p}(\Omega) \hookrightarrow C(\overline{\Omega}).$$

*Proof.* We write

$$\frac{S(u + s_n h_n) - S(u)}{s_n} - S'(u)h = \frac{S(u + s_n h_n) - S(u) - s_n S'(u)h_n}{s_n} + S'(u)(h_n - h).$$

From this, Theorem 3.5, and the compact embedding  $L^2(\Omega) \Subset W^{-1,p}(\Omega)$  and the continuous embedding  $W_0^{1,p}(\Omega) \hookrightarrow C(\overline{\Omega})$ , we derive the desired conclusion.  $\square$

**Remark 3.7.** All results in this and the following sections can be extended to the case where an additional semilinear term  $f(x, y(x))$  is present if  $f$  is a Carathéodory function and for almost every  $x \in \Omega$ , the mapping  $z \mapsto f(x, z)$  is monotone and of class  $C^2$  as in, e.g., [14]. However, in order to avoid additional technicalities, we have restricted the discussion to the case (3.1). Similarly, in assumption (A1), the convexity of  $\Omega$  was only used to establish the  $H^2$ - and  $W^{1,\infty}$ -regularity of solutions of elliptic equations. The results therefore remain valid for other domains guaranteeing that regularity, e.g., if  $\Omega$  is of class  $C^{1,1}$ .

## 4 EXISTENCE AND FIRST-ORDER OPTIMALITY CONDITIONS

The optimal control problem (P) can be rewritten in the form

$$(P) \quad \begin{cases} \min_{u \in L^\infty(\Omega)} j(u) := J(S(u), u) = G(S(u)) + \frac{\nu}{2} \|u\|_{L^2(\Omega)}^2 \\ \text{s.t. } u \in \mathcal{U}_{ad}, \end{cases}$$

where the admissible set is defined by

$$(4.1) \quad \mathcal{U}_{ad} := \{u \in L^\infty(\Omega) \mid \alpha(x) \leq u(x) \leq \beta(x) \text{ for a.e. } x \in \Omega\}.$$

The existence of a minimizer  $\bar{u} \in \mathcal{U}_{ad}$  of (P) follows as in [14, Thm. 3.1].

To derive first-order necessary optimality conditions, we first address the existence, uniqueness, and regularity of solutions to the adjoint state equation

$$(4.2) \quad \begin{cases} -\operatorname{div}[(b + a(y_u))\nabla\varphi] + \mathbb{1}_{\{y_u \notin E_a\}} a'(y_u) \nabla y_u \cdot \nabla\varphi = v & \text{in } \Omega, \\ \varphi = 0 & \text{on } \partial\Omega, \end{cases}$$

for  $u \in W^{-1,p}(\Omega)$ ,  $p > N$ ,  $v \in H^{-1}(\Omega)$ , and  $y_u := S(u)$ .

**Lemma 4.1.** *Let Assumptions (A1) to (A3) hold and  $p > N$ . Then, for any  $u \in W^{-1,p}(\Omega)$  and  $v \in H^{-1}(\Omega)$ , there exists a unique  $\varphi \in H_0^1(\Omega)$  solving (4.2). If, in addition,  $U$  is a bounded subset in  $L^p(\Omega)$ , then for any  $u \in U$  and  $v \in L^q(\Omega)$  with  $q > N$ , there exists a constant  $C_U$  such that  $\varphi \in H^2(\Omega) \cap W^{1,\infty}(\Omega)$  satisfies*

$$(4.3) \quad \|\varphi\|_{H^2(\Omega)} + \|\varphi\|_{W^{1,\infty}(\Omega)} \leq C_U \|v\|_{L^q(\Omega)}.$$

*Proof.* For any  $u \in W^{-1,p}(\Omega)$ , we define a function  $a_u$  and a vector-valued function  $\mathfrak{c}_u$  given by

$$a_u := b + a(y_u), \quad \mathfrak{c}_u := \mathbb{1}_{\{y_u \notin E_a\}} a'(y_u) \nabla y_u.$$

By Theorem 3.1, we have  $y_u \in W_0^{1,p}(\Omega)$  and thus  $a_u \in W^{1,p}(\Omega)$  and  $\mathfrak{c}_u \in L^p(\Omega)^N$ . Consider the operator

$$T_u : H_0^1(\Omega) \rightarrow H^{-1}(\Omega) \quad \varphi \mapsto T_u \varphi := -\operatorname{div}[a_u \nabla \varphi] + \mathfrak{c}_u \cdot \nabla \varphi.$$

Then  $T_u$  is an isomorphism (cf. [11, Thm. 2.6]). Therefore, for any  $v \in H^{-1}(\Omega)$ , there exists a unique solution  $\varphi \in H_0^1(\Omega)$  to (4.2) such that  $T_u \varphi = v$ .

It remains to show the  $W^{1,\infty}$ -regularity of  $\varphi$  and the estimate (4.3). Let  $U$  be a bounded subset in  $L^p(\Omega)$  and  $u \in U$ ,  $v \in L^q(\Omega)$  be arbitrary. Then  $y_u \in W^{1,\infty}(\Omega)$  by Theorem 3.1. It follows that  $\mathfrak{c}_u \in L^\infty(\Omega)^N$  and  $a_u$  is uniformly Lipschitz continuous. In addition,  $\varphi \in H_0^1(\Omega)$  satisfies

$$-\operatorname{div}[a_u \nabla \varphi] = v - \mathfrak{c}_u \cdot \nabla \varphi \in L^2(\Omega),$$

which together with the  $H^2$ -regularity of solutions (see, e.g., [25, Thm. 3.2.1.2]) gives  $\varphi \in H^2(\Omega)$ . Therefore,  $\varphi \in H_0^1(\Omega) \cap H^2(\Omega)$  is the strong solution to

$$-\Delta \varphi = \frac{1}{a_u} [v - \mathfrak{c}_u \cdot \nabla \varphi + \nabla a_u \cdot \nabla \varphi] \quad \text{in } \Omega.$$

By virtue of the chain rule (see, e.g., [24, Thm. 7.8]), one has  $\nabla a_u = \nabla b + \mathfrak{c}_u$ . We then have that  $\varphi \in H_0^1(\Omega) \cap H^2(\Omega)$  satisfies

$$(4.4) \quad -\Delta \varphi = f := \frac{1}{a_u} [v + \nabla b \cdot \nabla \varphi] \quad \text{in } \Omega$$

and there holds

$$\|\Delta\varphi\|_{L^2(\Omega)} \leq \frac{1}{b} \left[ \|v\|_{L^2(\Omega)} + L_b \|\varphi\|_{H_0^1(\Omega)} \right].$$

On the other hand, using a similar argument as in (3.34) and employing the continuous embedding  $L^q(\Omega) \hookrightarrow W^{-1,q}(\Omega)$ , we can show that

$$\|\varphi\|_{H_0^1(\Omega)} \leq C_{1,U} \|v\|_{W^{-1,q}(\Omega)} \leq C_{2,U} \|v\|_{L^q(\Omega)}.$$

We thus have

$$(4.5) \quad \|\varphi\|_{H^2(\Omega)} \leq C_{3,U} \|v\|_{L^q(\Omega)}.$$

Setting  $\tilde{q} := \min\{q, 6\} > N$  and using the continuous embedding  $H^2(\Omega) \hookrightarrow W^{1,6}(\Omega) \hookrightarrow W^{1,\tilde{q}}(\Omega)$  thus gives

$$\|f\|_{L^{\tilde{q}}(\Omega)} \leq C_{4,U} \|v\|_{L^q(\Omega)}.$$

From this, [Assumption \(A1\)](#), and the global boundedness of the gradient of solutions to Poisson's equation (see, e.g. [19, Thm. 3.1, Rem. 3], [30, 39]), we can conclude that  $\varphi \in W^{1,\infty}(\Omega)$  and that

$$\|\nabla\varphi\|_{L^\infty(\Omega)} \leq C_{5,U} \|f\|_{L^{\tilde{q}}(\Omega)} \leq C_{5,U} C_{4,U} \|v\|_{L^q(\Omega)},$$

which, along with (4.5), gives (4.3).  $\square$

From the chain rule, [Lemma 4.1](#), and an elementary calculus, we derive the following result.

**Theorem 4.2.** *Let Assumptions (A1) to (A4) hold. Then the reduced cost functional  $j : L^2(\Omega) \rightarrow \mathbb{R}$  is of class  $C^1$ . Moreover, for any  $u, h \in L^2(\Omega)$ , there holds*

$$j'(u)h = \int_{\Omega} (\varphi_u + vu) h \, dx,$$

where  $\varphi_u \in H_0^1(\Omega)$  solves (4.2) corresponding to the right-hand side term  $v = G'(y_u)$  and  $y_u$  solves (3.1).

We now arrive at first-order necessary optimality condition for the problem (P). Since the reduced functional is Fréchet differentiable by [Theorem 4.2](#), the proof of the following result, based on the variational inequality  $j'(\bar{u})(u - \bar{u}) \geq 0$  for all  $u \in \mathcal{U}_{ad}$ , is standard and therefore omitted.

**Theorem 4.3.** *Let Assumptions (A1) to (A4) hold. If  $\bar{u}$  is a local minimizer of (P), then there exists an adjoint state  $\bar{\varphi} \in H_0^1(\Omega)$  such that*

$$(4.6a) \quad \begin{cases} -\operatorname{div} [(b + a(\bar{y})) \nabla \bar{\varphi}] + \mathbb{1}_{\{\bar{y} \notin E_a\}} a'(\bar{y}) \nabla \bar{y} \cdot \nabla \bar{\varphi} = G'(\bar{y}) & \text{in } \Omega, \\ \bar{\varphi} = 0 & \text{on } \partial\Omega, \end{cases}$$

$$(4.6b) \quad \int_{\Omega} (\bar{\varphi} + v\bar{u})(u - \bar{u}) \, dx \geq 0 \quad \text{for all } u \in \mathcal{U}_{ad}$$

with  $\bar{y} := S(\bar{u})$ .

**Remark 4.4.** If the discrepancy term  $G$  is an integral functional of the form

$$G(y) = \int_{\Omega} g(x, y(x)) \, dx$$

where  $g : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$  is a Carathéodory function of class  $C^1$  with respect to the second variable such that [Assumption \(A4\)](#) holds, then we obtain from [Theorem 4.3](#) and [34, Lem. 2.26] the Pontryagin maximum principle

$$g(x, \bar{y}(x)) + \frac{\nu}{2} \bar{u}(x)^2 + \bar{\varphi}(x) \bar{u}(x) = \min \left\{ g(x, \bar{y}(x)) + \frac{\nu}{2} s^2 + \bar{\varphi}(x) s \mid s \in [\alpha(x), \beta(x)] \right\} \quad \text{for a.e. } x \in \Omega$$

due to the convexity of the mapping  $u \mapsto J(y, u)$ ; cf. [14, Thm. 4.1].

## 5 SECOND-ORDER OPTIMALITY CONDITIONS

Our goal is now to derive second-order necessary and sufficient conditions in terms of a non-smooth *curvature functional* characterizing the (generalized) curvature of the reduced functional  $j$  in critical directions. A similar approach was followed in [17, 18]. We will introduce the necessary technical notation in Section 5.1, prove preliminary estimates in Section 5.2, and finally derive the desired second-order conditions in Section 5.3.

### 5.1 NON-SMOOTH CURVATURE FUNCTIONAL

Intuitively, differentiating (P) (formally) and applying the sum and product rules, we see that the total curvature of  $j$  can be separated into three contributions: a *smooth* part involving only  $a$  and its derivatives in smooth points; a *first-order non-smooth* part involving only first (directional) derivatives of  $a$ ; and a *second-order non-smooth* part relating to second generalized derivatives of  $a$ .

Correspondingly, we define the following three partial curvature functionals at a point  $(u, y, \varphi) \in L^2(\Omega) \times H^1(\Omega) \times W^{1,\infty}(\Omega)$ . First, the smooth part of the curvature in directions  $(h_1, h_2) \in L^2(\Omega)^2$  is given by

$$(5.1) \quad Q_s(u, y, \varphi; h_1, h_2) := \frac{1}{2} G''(y)(S'(u)h_1 S'(u)h_2) + \frac{\nu}{2} \int_{\Omega} h_1 h_2 \, dx \\ - \frac{1}{2} \int_{\Omega} \mathbb{1}_{\{y \notin E_a\}} a''(y)(S'(u)h_1)(S'(u)h_2) \nabla y \cdot \nabla \varphi \, dx,$$

which is a bilinear form in  $(h_1, h_2)$ .

The first-order non-smooth part of the curvature is given by

$$Q_1(u, y, \varphi; h_1, h_2) := -\frac{1}{2} \int_{\Omega} [a'(y; S'(u)h_1) \nabla(S'(u)h_2) + a'(y; S'(u)h_2) \nabla(S'(u)h_1)] \cdot \nabla \varphi \, dx.$$

Although  $Q_1(u, y, \varphi; \cdot, \cdot)$  is not bilinear, it is positively homogeneous in each variable due to the positive homogeneity of the function  $a'(y(x); \cdot)$  for all  $x \in \bar{\Omega}$ , i.e.,

$$Q_1(u, y, \varphi; \tau_1 h_1, \tau_2 h_2) = \tau_1 \tau_2 Q_1(u, y, \varphi; h_1, h_2) \quad \text{for all } h_1, h_2 \in L^2(\Omega), \tau_1, \tau_2 \geq 0.$$

If  $a$  is a  $C^2$  function,  $Q_s + Q_1$  corresponds to the second derivative  $j''(u)(h_1, h_2)$  of the reduced functional; in this case our second-order conditions reduce to the results obtained in [14].

The critical part for our analysis is of course the second-order non-smooth part, which requires some additional notation. For any  $0 \leq i \leq K$ ,  $s \in \mathbb{R}$ , and  $h \in L^2(\Omega)$ , we define

$$\zeta_i(u, y; s, h) := [a'_{i-1}(t_i) - a'_i(t_i)] \mathbb{1}_{\Omega_{S(u+sh), y}^{i,2}}(t_i - S(u+sh)) \\ + [a'_{i+1}(t_{i+1}) - a'_i(t_{i+1})] \mathbb{1}_{\Omega_{S(u+sh), y}^{i,3}}(t_{i+1} - S(u+sh)),$$

with  $t_i$  as defined in Assumption (A3) and the sets  $\Omega_{y, \hat{y}}^{i,2}, \Omega_{y, \hat{y}}^{i,3}$  as defined in Lemma 3.3. Here we use the convention  $a_{-1} \equiv a_{K+1} \equiv 0$ ,  $a'_0(t_0) = a'_0(-\infty) = 0$ , and  $a'_K(t_{K+1}) = a'_K(\infty) = 0$ . We then define for positive null sequences  $\{s_n\} \in c_0^+ := \{\{s_n\} \subset (0, \infty) \mid s_n \rightarrow 0\}$ ,  $h \in L^2(\Omega)$ , and  $\{h_n\} \subset L^2(\Omega)$ ,

$$(5.2) \quad A_n(u, y; \{s_n\}, \{h_n\}) := \sum_{i=0}^K \zeta_i(u, y; s_n, h_n)$$

as well as

$$(5.3) \quad \tilde{Q}(u, y, \varphi; \{s_n\}, h) := \liminf_{n \rightarrow \infty} \frac{1}{s_n^2} \int_{\Omega} A_n(u, y; \{s_n\}, \{h\}) \nabla y \cdot \nabla \varphi \, dx,$$

where  $\{h\}$  denotes the constant sequence  $h_n \equiv h$ . (These terms will all be of independent use in the following.) The second-order non-smooth part of the curvature in direction  $h \in L^2(\Omega)$  is then given by

$$(5.4) \quad Q_2(u, y, \varphi; h) := \inf \{ \tilde{Q}(u, y, \varphi; \{s_n\}, h) \mid \{s_n\} \in c_0^+ \}.$$

Crucially,  $Q_2$  is positively homogeneous of degree 2 in  $h$ .

**Lemma 5.1.** *Let  $u, h \in L^2(\Omega)$ ,  $y \in H^1(\Omega)$ , and  $\varphi \in W^{1,\infty}(\Omega)$ . If  $Q_2(u, y, \varphi; h) < \infty$ , then*

$$Q_2(u, y, \varphi; th) = t^2 Q_2(u, y, \varphi; h) \quad \text{for all } t > 0.$$

*Proof.* We first observe that

$$\zeta_i(u, y; s_n, th) = \zeta_i(u, y; ts_n, h) \quad \text{and} \quad A_n(u, y; \{s_n\}, \{th\}) = A_n(u, y; \{ts_n\}, \{h\})$$

hold for all  $0 \leq i \leq K$ ,  $n \geq 1$  and for all  $u, h \in L^2(\Omega)$ ,  $y \in H^1(\Omega)$ ,  $\{s_n\} \in c_0^+$ , and  $t > 0$ . By the definition of  $\tilde{Q}$ , we have

$$\tilde{Q}(u, y, \varphi; \{s_n\}, th) = t^2 \tilde{Q}(u, y, \varphi; \{ts_n\}, h)$$

for all  $u, h \in L^2(\Omega)$ ,  $y \in H^1(\Omega)$ ,  $\varphi \in W^{1,\infty}(\Omega)$ ,  $\{s_n\} \in c_0^+$ , and  $t > 0$ . It follows that

$$\begin{aligned} Q_2(u, y, \varphi; th) &= \inf \{ \tilde{Q}(u, y, \varphi; \{s_n\}, th) \mid \{s_n\} \in c_0^+ \} \\ &= \inf \{ t^2 \tilde{Q}(u, y, \varphi; \{ts_n\}, h) \mid \{s_n\} \in c_0^+ \} \\ &= t^2 \inf \{ \tilde{Q}(u, y, \varphi; \{r_n\}, h) \mid \{r_n\} \in c_0^+ \} \quad (\text{by setting } r_n := ts_n) \\ &= t^2 Q_2(u, y, \varphi; h). \end{aligned} \quad \square$$

Note also that  $\zeta_i(u, y; s, h) = 0$  and therefore  $Q_2(u, y, \varphi; h) = 0$  if the functions  $a_i$  have the property that  $a'_{i-1}(t_i) = a'_i(t_i)$  for all  $1 \leq i \leq K$ , i.e., if the derivative  $a'$  is finitely  $PC^1$ . We also remark that  $Q_2$  is related to the term  $\sigma(h)$  used to bound the second derivative of the Lagrangian in [9, Chap. 3] and which there characterized the gap between necessary and sufficient second-order conditions.

Finally, to account for the control constraints, we recall the following basic notation standard in the study of second-order conditions; see, e.g., [9, 2]. Let  $K$  be a closed subset in  $L^2(\Omega)$  and let  $z \in K$  be arbitrary. The *radial*, *contingent tangent*, and *normal* cones to  $K$  at  $z$  are defined, respectively, as

$$\begin{aligned} \mathcal{R}(K; z) &:= \{ v \in L^2(\Omega) \mid \exists \bar{s} > 0 \text{ s.t. } z + sv \in K \text{ for all } s \in [0, \bar{s}] \}, \\ \mathcal{T}(K; z) &:= \{ v \in L^2(\Omega) \mid \exists s_n \rightarrow 0^+, v_n \rightarrow v \text{ in } L^2(\Omega) \text{ s.t. } z + s_n v_n \in K \text{ for all } n \in \mathbb{N} \}, \\ \mathcal{N}(K; z) &:= \left\{ w \in L^2(\Omega) \mid \int_{\Omega} w(x) (v(x) - z(x)) \, dx \leq 0 \text{ for all } v \in \mathcal{T}(K; z) \right\}. \end{aligned}$$

It is well-known that when  $K$  is convex, we have

$$\mathcal{T}(K; z) = \text{cl}_2 [\mathcal{R}(K; z)],$$

where  $\text{cl}_2(U)$  stands for the closure of a set  $U$  in  $L^2(\Omega)$ . For any  $w \in L^2(\Omega)$ , we denote the annihilator of  $w$  by

$$w^\perp := \left\{ h \in L^2(\Omega) \mid \int_{\Omega} w(x) h(x) \, dx = 0 \right\}.$$

Furthermore, we say that the set  $K$  is *polyhedral at*  $z \in K$  if for any  $w \in \mathcal{N}(K; z)$ , there holds

$$\text{cl}_2 [\mathcal{R}(K; z) \cap (w^\perp)] = \mathcal{T}(K; z) \cap (w^\perp).$$

We say that  $K$  is *polyhedral* if it is polyhedral at each point  $z \in K$ .

In the following, we will consider the admissible set  $\mathcal{U}_{ad}$ , defined in (4.1), as a subset in  $L^2(\Omega)$  rather than a subset in  $L^\infty(\Omega)$ . In this case,  $\mathcal{U}_{ad}$  is polyhedral, see [2, Lem. 4.13].

Furthermore, for a given triple  $(\bar{u}, \bar{y}, \bar{\varphi})$  with  $\bar{u} \in \mathcal{U}_{ad}$  satisfying the system (4.6), set

$$(5.5) \quad \bar{d} := \bar{\varphi} + v\bar{u}.$$

Obviously,  $-\bar{d} \in \mathcal{N}(\mathcal{U}_{ad}; \bar{u})$  by (4.6b). The *critical cone* of the problem (P) at  $\bar{u}$  is then defined via

$$(5.6) \quad C(\mathcal{U}_{ad}; \bar{u}) := \{h \in L^2(\Omega) \mid h \in \mathcal{T}(\mathcal{U}_{ad}; \bar{u}) \cap (\bar{d}^\perp)\}.$$

By [2, Lem. 4.11], the tangent cone  $\mathcal{T}(\mathcal{U}_{ad}; \bar{u})$  and the critical cone  $C(\mathcal{U}_{ad}; \bar{u})$  can, respectively, be characterized pointwise as

$$\mathcal{T}(\mathcal{U}_{ad}; \bar{u}) = \left\{ v \in L^2(\Omega) \mid v(x) \begin{cases} \geq 0 & \text{if } \bar{u}(x) = \alpha(x) \\ \leq 0 & \text{if } \bar{u}(x) = \beta(x) \end{cases} \text{ a.e. } x \in \Omega \right\}$$

and

$$(5.7) \quad C(\mathcal{U}_{ad}; \bar{u}) := \left\{ h \in L^2(\Omega) \mid h(x) \begin{cases} \geq 0 & \text{if } \bar{u}(x) = \alpha(x) \\ \leq 0 & \text{if } \bar{u}(x) = \beta(x) \\ = 0 & \text{if } \bar{d}(x) \neq 0 \end{cases} \text{ a.e. } x \in \Omega \right\}.$$

## 5.2 PRELIMINARY ESTIMATES

Throughout this section, let  $(\bar{u}, \bar{y}, \bar{\varphi})$  be a point that satisfies the system (4.6) and  $\bar{d}$  be given by (5.5). We start this section with a second-order Taylor-type expansion.

**Lemma 5.2.** *For any  $u \in L^2(\Omega)$  and  $y_u := S(u)$ , there holds*

$$(5.8) \quad \begin{aligned} j(u) - j(\bar{u}) &= \int_0^1 (1-s)G''(\bar{y} + s(y_u - \bar{y}))(y_u - \bar{y})^2 ds + \frac{v}{2}\|u - \bar{u}\|_{L^2(\Omega)}^2 \\ &\quad + \int_\Omega \bar{d}(u - \bar{u}) dx - \int_\Omega [a(y_u) - a(\bar{y})] \nabla \bar{\varphi} \cdot (\nabla y_u - \nabla \bar{y}) dx \\ &\quad - \int_\Omega [a(y_u) - a(\bar{y}) - \mathbb{1}_{\{\bar{y} \notin E_a\}} a'(\bar{y})(y_u - \bar{y})] \nabla \bar{\varphi} \cdot \nabla \bar{y} dx. \end{aligned}$$

*Proof.* Since  $j$  is Fréchet differentiable by Theorem 4.2 and  $G$  is  $C^2$  by Assumption (A4), we can use a Taylor expansion to write

$$(5.9) \quad \begin{aligned} j(u) - j(\bar{u}) &= G(y_u) - G(\bar{y}) + \frac{v}{2} \left( \|u\|_{L^2(\Omega)}^2 - \|\bar{u}\|_{L^2(\Omega)}^2 \right) \\ &= G'(\bar{y})(y_u - \bar{y}) + v \int_\Omega (u - \bar{u}) \bar{u} dx \\ &\quad + \int_0^1 (1-s)G''(\bar{y} + s(y_u - \bar{y}))(y_u - \bar{y})^2 ds + \frac{v}{2}\|u - \bar{u}\|_{L^2(\Omega)}^2 \\ &= \int_\Omega (b + a(\bar{y})) \nabla \bar{\varphi} \cdot (\nabla y_u - \nabla \bar{y}) + \mathbb{1}_{\{\bar{y} \notin E_a\}} a'(\bar{y}) \nabla \bar{y} \cdot \nabla \bar{\varphi} (y_u - \bar{y}) dx \\ &\quad - \int_\Omega \bar{\varphi}(u - \bar{u}) dx + \int_\Omega \bar{d}(u - \bar{u}) dx \\ &\quad + \int_0^1 (1-s)G''(\bar{y} + s(y_u - \bar{y}))(y_u - \bar{y})^2 ds + \frac{v}{2}\|u - \bar{u}\|_{L^2(\Omega)}^2, \end{aligned}$$

where we have employed (4.6a) and the definition of  $\bar{d}$  to obtain the last equality. Testing the state equations corresponding to  $y_u$  and  $\bar{y}$  by  $\bar{\varphi}$  and then subtracting yields

$$\int_{\Omega} \bar{\varphi}(u - \bar{u}) \, dx = \int_{\Omega} [(b + a(y_u))(\nabla y_u - \nabla \bar{y}) + (a(y_u) - a(\bar{y}))\nabla \bar{y}] \cdot \nabla \bar{\varphi} \, dx.$$

Inserting this equality into (5.9), we arrive at the desired conclusion.  $\square$

A crucial step of our analysis will be to bound  $Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h)$  purely in terms of the jumps of the derivatives of  $a$  in the optimal state  $\bar{y}$ . To do this, we define the jump functional  $\Sigma : W^{1,1}(\Omega) \rightarrow [0, \infty]$  via

$$(5.10) \quad \begin{aligned} \Sigma(y) &:= \limsup_{r \rightarrow 0^+} \frac{1}{r} \sum_{m=1}^N \sum_{i=1}^K \sigma_i \int_{\Omega} [\mathbb{1}_{\{0 < |y - t_i| \leq r\}} |\partial_{x_m} y|] \, dx \\ &= \limsup_{r \rightarrow 0^+} \frac{1}{r} \sum_{m=1}^N \sum_{i \in I_y^+} \sigma_i \int_{\Omega} [\mathbb{1}_{\{|y - t_i| \leq r\}} |\partial_{x_m} y|] \, dx, \end{aligned}$$

where the (non-negative)  $\{\sigma_i\}_{1 \leq i \leq K}$  are as defined in (2.3) and

$$I_y^+ := \{i \in \{1, \dots, K\} \mid \min_{x \in \bar{\Omega}} y(x) \leq t_i \leq \max_{x \in \bar{\Omega}} y(x)\}.$$

(Note that in contrast to  $I_y$  defined in (2.4), we exclude the largest value of  $i$  for which  $t_i < \min_{x \in \bar{\Omega}} y(x)$  but include the largest value of  $i$  for which  $t_i = \max_{x \in \bar{\Omega}} y(x)$ .)

Clearly, if  $a'$  is finitely  $PC^1$ , then  $\Sigma(y) = 0$ . On the other hand, if  $a'$  has points of discontinuity, then, intuitively,  $\Sigma(y)$  measures the oscillation of  $y \in W^{1,1}(\Omega)$  around these discontinuities (which maybe unbounded even if there are only finitely many such points). This is illustrated by the following one-dimensional example, whose derivation by straight-forward calculation is given in [Appendix B](#).

**Example 5.3.** Let  $\Omega := (\alpha_0, \beta_0) \subset \mathbb{R}$  be bounded and let  $y : \bar{\Omega} \rightarrow \mathbb{R}$  be a Lipschitz continuous function that is piecewise monotone (increasing or decreasing). Assume that  $a$  satisfies [Assumption \(A3\)](#) such that  $\min_{1 \leq i \leq K} \sigma_i > 0$  and that

$$y^{-1}(E_a) := \{x \in [\alpha_0, \beta_0] \mid y(x) \in E_a\} = \bigcup_{j \in J} [\underline{x}_j, \bar{x}_j]$$

for some  $\underline{x}_j, \bar{x}_j \in [\alpha_0, \beta_0]$ ,  $j \in J$ , with  $[\underline{x}_j, \bar{x}_j] \cap [\underline{x}_k, \bar{x}_k] = \emptyset$  for all  $j, k \in J$ ,  $j \neq k$ . Then, the following assertions hold:

- (i) if  $\text{card}(J) < \infty$ , then  $(\text{card}(J) - 1) \min_{1 \leq i \leq K} \sigma_i \leq \Sigma(y) \leq 2 \text{card}(J) \max_{1 \leq i \leq K} \sigma_i$ ;
- (ii) if  $\text{card}(J) = \infty$ , then  $\Sigma(y) = \infty$ ;

where  $\text{card}(J)$  stands for the cardinality of set  $J$ . In particular,  $\Sigma(y)$  is finite if and only if the index set  $J$  is finite.

Note that if  $\underline{x}_j < \bar{x}_j$ , then  $[\underline{x}_j, \bar{x}_j] \subset \{y \in E_a\}$ , which therefore has positive Lebesgue measure. This demonstrates that the assumption  $\Sigma(y) < \infty$ , which will be crucial throughout the following, does not imply that  $\{y \in E_a\}$  has measure zero.

We now prove some technical results on  $\Sigma$  and  $Q_2$  that will be needed in the following [Section 5.3](#). To keep the notation concise, from now on we will simply write  $\zeta_i(s, h) := \zeta_i(\bar{u}, \bar{y}; s, h)$  and  $A_n(\{s_n\}, \{h_n\}) := A_n(\bar{u}, \bar{y}; \{s_n\}, \{h_n\})$ .

**Lemma 5.4.** *Let  $\{s_n\} \in c_0^+$  be arbitrary and let  $\{h_n\}, \{v_n\} \subset L^2(\Omega)$  be such that  $h_n \rightarrow h$  and  $v_n \rightarrow h$  in  $L^2(\Omega)$  for some  $h \in L^2(\Omega)$ . If  $\Sigma(\bar{y}) < \infty$  and  $\bar{\varphi} \in W^{1,\infty}(\Omega)$ , then*

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^2} \int_{\Omega} [A_n(\{s_n\}, \{h_n\}) - A_n(\{s_n\}, \{v_n\})] \nabla \bar{y} \cdot \nabla \bar{\varphi} \, dx = 0.$$

*Proof.* Setting  $y_n := S(\bar{u} + s_n h_n)$ ,  $w_n := S(\bar{u} + s_n v_n)$  and exploiting the definition (5.2) yields

$$(5.11) \quad A_n(\{s_n\}, \{h_n\}) - A_n(\{s_n\}, \{v_n\}) = \sum_{i=0}^K \{ [a'_{i-1}(t_i) - a'_i(t_i)] M_{n,i} + [a'_{i+1}(t_{i+1}) - a'_i(t_{i+1})] L_{n,i} \}$$

with

$$M_{n,i} := \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,2}}(t_i - y_n) - \mathbb{1}_{\Omega_{w_n, \bar{y}}^{i,2}}(t_i - w_n), \quad L_{n,i} := \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,3}}(t_{i+1} - y_n) - \mathbb{1}_{\Omega_{w_n, \bar{y}}^{i,3}}(t_{i+1} - w_n).$$

Setting  $r_n := \|y_n - w_n\|_{L^\infty(\Omega)}$  and  $\kappa_n := \max\{\|y_n - \bar{y}\|_{L^\infty(\Omega)}, \|w_n - \bar{y}\|_{L^\infty(\Omega)}\}$  gives

$$(5.12) \quad \kappa_n \leq C s_n \quad \text{and} \quad \frac{r_n}{s_n} \rightarrow 0$$

for some positive constant  $C$  due to the differentiability of  $S$  and [Corollary 3.6](#), respectively. We can thus assume that  $0 \leq \kappa_n, r_n < \delta$  for all  $n \in \mathbb{N}$  large enough. Writing

$$M_{n,i} := -\mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,2}}(y_n - w_n) - \left( \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,2}} - \mathbb{1}_{\Omega_{w_n, \bar{y}}^{i,2}} \right) (w_n - t_i)$$

and using the fact that

$$\Omega_{y_n, \bar{y}}^{i,2} := \{ \bar{y} \in (t_i, t_i + \delta), y_n \in (t_i - \delta, t_i] \} \subset \{ 0 < \bar{y} - t_i \leq \kappa_n \},$$

we derive

$$|M_{n,i}| \leq \mathbb{1}_{\{0 < \bar{y} - t_i \leq \kappa_n\}} r_n + \left| \left( \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,2}} - \mathbb{1}_{\Omega_{w_n, \bar{y}}^{i,2}} \right) (w_n - t_i) \right|.$$

On the other hand, by a simple calculation, it holds that

$$\begin{aligned} & \left| \left( \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,2}} - \mathbb{1}_{\Omega_{w_n, \bar{y}}^{i,2}} \right) (w_n - t_i) \right| \\ &= \mathbb{1}_{\{\bar{y} \in (t_i, t_i + \delta)\}} \left| \mathbb{1}_{\{y_n \in (t_i - \delta, t_i]\}} - \mathbb{1}_{\{w_n \in (t_i - \delta, t_i]\}} \right| |w_n - t_i| \\ &= \mathbb{1}_{\{\bar{y} \in (t_i, t_i + \delta)\}} \left| \mathbb{1}_{\{y_n \in (t_i - \delta, t_i), w_n \in (t_i, t_i + r_n]\}} - \mathbb{1}_{\{w_n \in (t_i - \delta, t_i), y_n \in (t_i, t_i + r_n]\}} \right| |w_n - t_i| \\ &\leq \mathbb{1}_{\{0 < \bar{y} - t_i \leq \kappa_n\}} r_n. \end{aligned}$$

Here we have exploited the facts that

$$\{ \bar{y} \in (t_i, t_i + \delta), y_n \in (t_i - \delta, t_i] \} \cup \{ \bar{y} \in (t_i, t_i + \delta), w_n \in (t_i - \delta, t_i] \} \subset \{ 0 < \bar{y} - t_i \leq \kappa_n \}$$

and  $|w_n - t_i| \leq |w_n - y_n| \leq r_n$  almost everywhere on  $\{w_n \in (t_i - \delta, t_i], y_n \in (t_i, t_i + r_n]\}$ . We therefore have

$$(5.13) \quad |M_{n,i}| \leq 2r_n \mathbb{1}_{\{0 < \bar{y} - t_i \leq \kappa_n\}}.$$

Similarly, there holds that

$$(5.14) \quad |L_{n,i}| \leq 2r_n \mathbb{1}_{\{0 < t_{i+1} - \bar{y} \leq \kappa_n\}}.$$

Inserting (5.13) and (5.14) into (5.11) and exploiting the obtained result as well as (2.3) yields

$$\begin{aligned} & \frac{1}{s_n^2} \int_{\Omega} |[A_n(\{s_n\}, \{h_n\}) - A_n(\{s_n\}, \{v_n\})] \nabla \bar{y} \cdot \nabla \bar{\varphi}| \, dx \\ & \leq \frac{2r_n}{s_n^2} \int_{\Omega} \sum_{i=0}^K [\sigma_i \mathbb{1}_{\{0 < \bar{y} - t_i \leq \kappa_n\}} + \sigma_{i+1} \mathbb{1}_{\{0 < t_{i+1} - \bar{y} \leq \kappa_n\}}] |\nabla \bar{y} \cdot \nabla \bar{\varphi}| \, dx \\ & = \frac{2r_n \kappa_n}{s_n^2} \times \frac{1}{\kappa_n} \int_{\Omega} \sum_{i=1}^K \sigma_i \mathbb{1}_{\{0 < |\bar{y} - t_i| \leq \kappa_n\}} |\nabla \bar{y} \cdot \nabla \bar{\varphi}| \, dx. \end{aligned}$$

Here we have used the fact that  $\mathbb{1}_{\{0 < \bar{y} - t_0 \leq \kappa_n\}} \equiv \mathbb{1}_{\{0 < t_{K+1} - \bar{y} \leq \kappa_n\}} \equiv 0$ . Letting  $n \rightarrow \infty$  and using the definition (5.10), we arrive at

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^2} \int_{\Omega} |[A_n(\{s_n\}, \{h_n\}) - A_n(\{s_n\}, \{v_n\})] \nabla \bar{y} \cdot \nabla \bar{\varphi}| \, dx \leq \|\nabla \bar{\varphi}\|_{L^\infty(\Omega)} \Sigma(\bar{y}) \lim_{n \rightarrow \infty} \frac{2r_n \kappa_n}{s_n^2},$$

which, together with (5.12), completes the proof.  $\square$

Combining this with (5.3) and (5.4), we obtain the following estimate.

**Corollary 5.5.** *If  $\Sigma(\bar{y}) < \infty$  and  $\bar{\varphi} \in W^{1,\infty}(\Omega)$ , then for any  $\{s_n\} \in c_0^+$  and any  $h_n \rightharpoonup h$  in  $L^2(\Omega)$ ,*

$$\liminf_{n \rightarrow \infty} \frac{1}{s_n^2} \int_{\Omega} A_n(\{s_n\}, \{h_n\}) \nabla \bar{y} \cdot \nabla \bar{\varphi} \, dx = \tilde{Q}(\bar{u}, \bar{y}, \bar{\varphi}; \{s_n\}, h) \geq Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h).$$

We can use this result to show weak lower semi-continuity of  $Q_2$ .

**Proposition 5.6.** *If  $\Sigma(\bar{y}) < \infty$  and  $\bar{\varphi} \in W^{1,\infty}(\Omega)$ , then for any  $h_n \rightharpoonup h$  in  $L^2(\Omega)$ ,*

$$Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h) \leq \liminf_{n \rightarrow \infty} Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h_n).$$

*Proof.* Let  $\{h_n\} \subset L^2(\Omega)$  be arbitrary such that  $h_n \rightharpoonup h$  in  $L^2(\Omega)$ . Fixing  $n \in \mathbb{N}$  and using the definition (5.4) shows that there exists a sequence  $\{s_j^k(h_n)\}_{j,k \in \mathbb{N}} \in c_0^+$  such that

$$(5.15) \quad s_j^k(h_n) \rightarrow 0^+ \text{ as } j \rightarrow \infty \quad \text{for all } k \in \mathbb{N}$$

and

$$Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h_n) = \lim_{k \rightarrow \infty} \tilde{Q}(\bar{u}, \bar{y}, \bar{\varphi}; \{s_j^k(h_n)\}_{j \in \mathbb{N}}, h_n).$$

There thus exists a  $k_n \geq n$  satisfying

$$(5.16) \quad Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h_n) - \tilde{Q}(\bar{u}, \bar{y}, \bar{\varphi}; \{s_j^{k_n}(h_n)\}_{j \in \mathbb{N}}, h_n) > -\frac{1}{n}.$$

The limit in (5.15) leads to the existence of a  $j_n \in \mathbb{N}$  such that

$$(5.17) \quad 0 < s_j^{k_n}(h_n) < \frac{1}{n} \quad \text{for all } j \geq j_n.$$

Furthermore, from (5.3) and (5.2), a subsequence  $\{s_{j_q}^{k_n}(h_n)\}_{q \in \mathbb{N}}$  of  $\{s_j^{k_n}(h_n)\}_{j \in \mathbb{N}}$  exists satisfying

$$\tilde{Q}(\bar{u}, \bar{y}, \bar{\varphi}; \{s_{j_q}^{k_n}(h_n)\}_{j \in \mathbb{N}}, h_n) = \lim_{q \rightarrow \infty} \frac{1}{(s_{j_q}^{k_n}(h_n))^2} \int_{\Omega} \sum_{i=0}^K \zeta_i(s_{j_q}^{k_n}(h_n), h_n) \nabla \bar{y} \cdot \nabla \bar{\varphi} \, dx.$$

Then there is a  $q_n \in \mathbb{N}$  satisfying  $j_{q_n} \geq j_n$  and

$$(5.18) \quad \tilde{Q}(\bar{u}, \bar{y}, \bar{\varphi}; \{s_j^{k_n}(h_n)\}_{j \in \mathbb{N}}, h_n) - \frac{1}{r_n^2} \int_{\Omega} \sum_{i=0}^K \zeta_i(r_n, h_n) \nabla \bar{y} \cdot \nabla \bar{\varphi} \, dx > -\frac{1}{n}$$

with  $r_n := s_{j_{q_n}}^{k_n}(h_n)$ . By (5.17), we have  $r_n \rightarrow 0^+$  as  $n \rightarrow \infty$  and so  $\{r_n\}_{n \in \mathbb{N}} \in c_0^+$ . On the other hand, adding (5.16) and (5.18) yields that

$$Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h_n) - \frac{1}{r_n^2} \int_{\Omega} \sum_{i=0}^K \zeta_i(r_n, h_n) \nabla \bar{y} \cdot \nabla \bar{\varphi} \, dx > -\frac{2}{n}.$$

Taking the limit inferior then shows that

$$\begin{aligned} \liminf_{n \rightarrow \infty} Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h_n) &\geq \liminf_{n \rightarrow \infty} \frac{1}{r_n^2} \int_{\Omega} \sum_{i=0}^K \zeta_i(r_n, h_n) \nabla \bar{y} \cdot \nabla \bar{\varphi} \, dx \\ &= \liminf_{n \rightarrow \infty} \frac{1}{r_n^2} \int_{\Omega} A_n(\{r_n\}, \{h_n\}) \nabla \bar{y} \cdot \nabla \bar{\varphi} \, dx, \end{aligned}$$

where we have used the definition (5.2) to obtain the last identity. Together with Corollary 5.5, this yields the claim.  $\square$

**Lemma 5.7.** *Assume that  $\Sigma(\bar{y}) < \infty$  and  $\bar{\varphi} \in W^{1,\infty}(\Omega)$ , then for any  $h \in L^2(\Omega)$ ,*

$$|Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h)| \leq \Sigma(\bar{y}) \|\nabla \bar{\varphi}\|_{L^\infty(\Omega)} \|S'(\bar{u})h\|_{L^\infty(\Omega)}^2.$$

*Proof.* It suffices to show for any  $\{s_n\} \in c_0^+$  that

$$(5.19) \quad |\tilde{Q}(\bar{u}, \bar{y}, \bar{\varphi}; \{s_n\}, h)| \leq \Sigma(\bar{y}) \|\nabla \bar{\varphi}\|_{L^\infty(\Omega)} \|S'(\bar{u})h\|_{L^\infty(\Omega)}^2.$$

To this end, we first set  $y_n := S(\bar{u} + s_n h)$  and take  $p \in (N, 6)$  arbitrary. By the fact that  $s_n \rightarrow 0^+$  and  $L^2(\Omega) \Subset W^{-1,p}(\Omega)$ , we have  $\bar{u} + s_n h \rightarrow \bar{u}$  in  $W^{-1,p}(\Omega)$  and thus  $y_n \rightarrow \bar{y}$  in  $W_0^{1,p}(\Omega)$ . From this and the embedding  $W_0^{1,p}(\Omega) \hookrightarrow C(\bar{\Omega})$ , it holds that

$$\kappa_n := \|y_n - \bar{y}\|_{C(\bar{\Omega})} \rightarrow 0.$$

We can thus assume that  $\kappa_n < \delta$  for all  $n \in \mathbb{N}$  large enough. Moreover, from (5.2), there holds

$$(5.20) \quad A_n(\{s_n\}, \{h\}) := \sum_{i=0}^K \zeta_i(s_n, h) = \sum_{i \in I_{\bar{y}}} \zeta_i(s_n, h).$$

We see from (2.3) and the definition of  $\zeta_i(s_n, h)$ ,  $\Omega_{y_n, \bar{y}}^{i,2}$  and  $\Omega_{y_n, \bar{y}}^{i,3}$  that

$$\begin{aligned} \sum_{i \in I_{\bar{y}}} |\zeta_i(s_n, h)| &\leq \sum_{i \in I_{\bar{y}}} \left[ \mathbb{1}_{\{\bar{y} \in (t_i, t_i + \delta), y_n \in (t_i - \delta, t_i)\}} \sigma_i + \mathbb{1}_{\{\bar{y} \in (t_{i+1} - \delta, t_{i+1}), y_n \in [t_{i+1}, t_{i+1} + \delta)\}} \sigma_{i+1} \right] \|y_n - \bar{y}\| \\ &\leq \sum_{i \in I_{\bar{y}}} \left[ \mathbb{1}_{\{0 < \bar{y} - t_i \leq \|y_n - \bar{y}\|_{C(\bar{\Omega})}\}} \sigma_i + \mathbb{1}_{\{0 < t_{i+1} - \bar{y} \leq \|y_n - \bar{y}\|_{C(\bar{\Omega})}\}} \sigma_{i+1} \right] \|y_n - \bar{y}\|_{C(\bar{\Omega})} \\ &= \sum_{i=1}^K \mathbb{1}_{\{0 < |\bar{y} - t_i| \leq \kappa_n\}} \sigma_i \|y_n - \bar{y}\|_{C(\bar{\Omega})}. \end{aligned}$$

Consequently, it holds that

$$\frac{1}{s_n^2} \int_{\Omega} \sum_{i \in I_{\bar{y}}} |\zeta_i(s_n, h) \nabla \bar{\varphi} \cdot \nabla \bar{y}| \, dx \leq \|\nabla \bar{\varphi}\|_{L^\infty(\Omega)} \frac{\|y_n - \bar{y}\|_{C(\bar{\Omega})}^2}{s_n^2} \sum_{m=1}^N \sum_{i=1}^K \sigma_i \frac{1}{\kappa_n} \int_{\Omega} \mathbb{1}_{\{0 < |\bar{y} - t_i| \leq \kappa_n\}} |\partial_{x_m} \bar{y}| \, dx.$$

Passing to the limit, employing [Corollary 3.6](#), and using [\(5.10\)](#) then yields that

$$\limsup_{n \rightarrow \infty} \frac{1}{s_n^2} \int_{\Omega} \sum_{i \in I_{\bar{y}}} |\zeta_i(s_n, h) \nabla \bar{\varphi} \cdot \nabla \bar{y}| \, dx \leq \|\nabla \bar{\varphi}\|_{L^\infty(\Omega)} \|S'(\bar{u})h\|_{L^\infty(\Omega)}^2 \Sigma(\bar{y}).$$

The combination of this with [\(5.20\)](#) gives [\(5.19\)](#) and thus the claim.  $\square$

The following is the main result of this subsection.

**Proposition 5.8.** *Let Assumptions (A1) to (A4) hold. Assume further that  $G'(\bar{y}) \in L^{\bar{p}}(\Omega)$  for some  $\bar{p} > N$ . Let  $p \in (N, 6)$  be arbitrary and let  $\{s_n\} \in c_0^+$  and  $\{h_n\} \subset L^2(\Omega)$  be arbitrary such that  $h_n \rightarrow h$  in  $L^2(\Omega)$  for some  $h \in L^2(\Omega)$ . Then the following limits hold:*

- (i)  $\frac{1}{s_n^2} \int_0^1 (1-s) G''(\bar{y} + s(y_n - \bar{y}))(y_n - \bar{y})^2 \, ds \rightarrow \frac{1}{2} G''(\bar{y})(S'(\bar{u})h)^2$  with  $y_n := S(\bar{u} + s_n h_n)$ ;
- (ii)  $\frac{1}{s_n^2} \int_{\Omega} [a(y_n) - a(\bar{y})] \nabla \bar{\varphi} \cdot (\nabla y_n - \nabla \bar{y}) \, dx \rightarrow \int_{\Omega} a'(\bar{y}; S'(\bar{u})h) \nabla \bar{\varphi} \cdot \nabla (S'(\bar{u})h) \, dx$ ;
- (iii) if, in addition,  $\Sigma(\bar{y}) < \infty$ , then

$$H_n := \int_{\Omega} [a(y_n) - a(\bar{y}) - \mathbb{1}_{\{\bar{y} \notin E_a\}} a'(\bar{y})(y_n - \bar{y})] \nabla \bar{\varphi} \cdot \nabla \bar{y} \, dx, \quad n \in \mathbb{N},$$

satisfy

$$\limsup_{n \rightarrow \infty} \frac{1}{s_n^2} H_n = \frac{1}{2} \int_{\Omega} \mathbb{1}_{\{\bar{y} \notin E_a\}} a''(\bar{y})(S'(\bar{u})h)^2 \nabla \bar{\varphi} \cdot \nabla \bar{y} \, dx - \tilde{Q}(\bar{u}, \bar{y}, \bar{\varphi}; \{s_n\}, h).$$

*Proof.* (i): By [Corollary 3.6](#), we have

$$(5.21) \quad \frac{y_n - \bar{y}}{s_n} \rightarrow S'(\bar{u})h \quad \text{in } W_0^{1,p}(\Omega).$$

This and the dominated convergence theorem give assertion (i).

(ii): According to [Lemma 4.1](#) and the fact that  $G'(\bar{y}) \in L^{\bar{p}}(\Omega)$  and  $\bar{u} \in L^\infty(\Omega)$ , the adjoint state  $\bar{\varphi}$  belongs to  $W^{1,\infty}(\Omega)$ . Moreover, from [\(5.21\)](#) we have that

$$\frac{1}{s_n} \nabla(y_n - \bar{y}) \rightarrow \nabla S'(\bar{u})h \quad \text{in } L^p(\Omega)^N \quad \text{and} \quad \frac{1}{s_n} (y_n - \bar{y}) \rightarrow S'(\bar{u})h \quad \text{in } C(\bar{\Omega}).$$

Finally, for all  $x \in \bar{\Omega}$ , it holds that  $\frac{1}{s_n} [a(y_n(x)) - a(\bar{y}(x))] \rightarrow a'(\bar{y}(x); S'(\bar{u})h(x))$ ; see, e.g., [\[20, Lem. 3.5\]](#). Therefore, we obtain (ii) from the dominated convergence theorem.

(iii): According to [Lemma 3.2](#) and the continuous embedding  $W_0^{1,p}(\Omega) \hookrightarrow C(\bar{\Omega})$ , we obtain  $y_n \rightarrow \bar{y}$  in  $C(\bar{\Omega})$ , and we can thus assume that  $\|y_n - \bar{y}\|_{C(\bar{\Omega})} < \delta$  for all  $n \in \mathbb{N}$  large enough. Since  $\nabla \bar{y}$  vanishes almost everywhere on  $\{\bar{y} \in E_a\}$  [\[16, Rem. 2.6\]](#) and there exists a constant  $M > 0$  such that  $\max\{|y_n(x)|, |\bar{y}(x)|\} \leq M$  for all  $x \in \bar{\Omega}$ , we can write

$$(5.22) \quad \begin{aligned} H_n &= \int_{\Omega} \mathbb{1}_{\{\bar{y} \notin E_a\}} [a(y_n) - a(\bar{y}) - a'(\bar{y})(y_n - \bar{y})] \nabla \bar{\varphi} \cdot \nabla \bar{y} \, dx \\ &= \int_{\Omega} T_{y_n, \bar{y}} \nabla \bar{\varphi} \cdot \nabla \bar{y} \, dx = \sum_{i \in I_{\bar{y}}} \int_{\Omega} \left[ T_{y_n, \bar{y}}^{i,1} + T_{y_n, \bar{y}}^{i,2} + T_{y_n, \bar{y}}^{i,3} \right] \nabla \bar{\varphi} \cdot \nabla \bar{y} \, dx, \end{aligned}$$

where  $T_{y_n, \bar{y}}$  and  $T_{y_n, \bar{y}}^{i,j}$ ,  $j = 1, 2, 3$ , are defined in (3.20) and Lemma 3.3.

We now estimate  $T_{y_n, \bar{y}}^{i,j}$  for  $i \in I_{\bar{y}}$  and  $j = 1, 2, 3$ . Let us fix  $i \in I_{\bar{y}}$  and consider  $T_{y_n, \bar{y}}^{i,1}$ . We have

$$(5.23) \quad T_{y_n, \bar{y}}^{i,1} = g_n^{i,1} + g_n^{i,2} + g_n^{i,3}$$

with

$$\begin{aligned} g_n^{i,1} &:= \mathbb{1}_{\Omega_{\bar{y}, i, i+1}^{[\delta, -\delta]}} [a_i(y_n) - a_i(\bar{y}) - a'_i(\bar{y})(y_n - \bar{y})], \\ g_n^{i,2} &:= \mathbb{1}_{\Omega_{\bar{y}, i, i}^{(0, \delta)} \cap \Omega_{y_n, i, i}^{(0, 2\delta)}} [a_i(y_n) - a_i(\bar{y}) - a'_i(\bar{y})(y_n - \bar{y})], \\ g_n^{i,3} &:= \mathbb{1}_{\Omega_{\bar{y}, i+1, i+1}^{(-\delta, 0)} \cap \Omega_{y_n, i+1, i+1}^{(-2\delta, 0)}} [a_i(y_n) - a_i(\bar{y}) - a'_i(\bar{y})(y_n - \bar{y})]. \end{aligned}$$

A standard argument shows that

$$(5.24) \quad \frac{1}{s_n^2} \int_{\Omega} g_n^{i,1} \nabla \bar{\varphi} \cdot \nabla \bar{y} \, dx \rightarrow \frac{1}{2} \int_{\Omega} \mathbb{1}_{\Omega_{\bar{y}, i, i+1}^{[\delta, -\delta]}} a_i''(\bar{y}) (S'(\bar{u})h)^2 \nabla \bar{\varphi} \cdot \nabla \bar{y} \, dx.$$

Since  $\|y_n - \bar{y}\|_{C(\bar{\Omega})} < \delta$  and  $\|y_n - \bar{y}\|_{C(\bar{\Omega})} \rightarrow 0$  as  $n \rightarrow \infty$ , there holds

$$\mathbb{1}_{\Omega_{\bar{y}, i, i}^{(0, \delta)} \cap \Omega_{y_n, i, i}^{(0, 2\delta)}} - \mathbb{1}_{\Omega_{\bar{y}, i, i}^{(0, \delta)}} = -\mathbb{1}_{\Omega_{\bar{y}, i, i}^{(0, \delta)} \cap \Omega_{y_n, i, i}^{(-\delta, 0)}} \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

almost everywhere in  $\Omega$ , which together with the dominated convergence theorem yields

$$(5.25) \quad \frac{1}{s_n^2} \int_{\Omega} g_n^{i,2} \nabla \bar{\varphi} \cdot \nabla \bar{y} \, dx \rightarrow \frac{1}{2} \int_{\Omega} \mathbb{1}_{\Omega_{\bar{y}, i, i}^{(0, \delta)}} a_i''(\bar{y}) (S'(\bar{u})h)^2 \nabla \bar{\varphi} \cdot \nabla \bar{y} \, dx.$$

Similarly, it holds that

$$(5.26) \quad \frac{1}{s_n^2} \int_{\Omega} g_n^{i,3} \nabla \bar{\varphi} \cdot \nabla \bar{y} \, dx \rightarrow \frac{1}{2} \int_{\Omega} \mathbb{1}_{\Omega_{\bar{y}, i+1, i+1}^{(-\delta, 0)}} a_i''(\bar{y}) (S'(\bar{u})h)^2 \nabla \bar{\varphi} \cdot \nabla \bar{y} \, dx.$$

From (5.23)–(5.26) and the fact that  $\Omega_{\bar{y}, i, i+1}^{[\delta, -\delta]} \cup \Omega_{\bar{y}, i, i}^{(0, \delta)} \cup \Omega_{\bar{y}, i+1, i+1}^{(-\delta, 0)} = \Omega_{\bar{y}, i, i+1}^{(0, 0)} =: \{\bar{y} \in (t_i, t_{i+1})\}$ , we deduce that

$$\frac{1}{s_n^2} \int_{\Omega} T_{y_n, \bar{y}}^{i,1} \nabla \bar{\varphi} \cdot \nabla \bar{y} \, dx \rightarrow \frac{1}{2} \int_{\Omega} \mathbb{1}_{\{\bar{y} \in (t_i, t_{i+1})\}} a_i''(\bar{y}) (S'(\bar{u})h)^2 \nabla \bar{\varphi} \cdot \nabla \bar{y} \, dx,$$

which, together with the fact that  $I_{\bar{y}}$  is finite, implies that

$$(5.27) \quad \frac{1}{s_n^2} \sum_{i \in I_{\bar{y}}} \int_{\Omega} T_{y_n, \bar{y}}^{i,1} \nabla \bar{\varphi} \cdot \nabla \bar{y} \, dx \rightarrow \frac{1}{2} \int_{\Omega} \mathbb{1}_{\{\bar{y} \notin E_a\}} a_i''(\bar{y}) (S'(\bar{u})h)^2 \nabla \bar{\varphi} \cdot \nabla \bar{y} \, dx.$$

We now estimate  $T_{y_n, \bar{y}}^{i,2}$ . To this end, we write

$$\begin{aligned} (5.28) \quad T_{y_n, \bar{y}}^{i,2} &= \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,2}} [a_{i-1}(y_n) - a_i(\bar{y}) - a'_i(\bar{y})(y_n - \bar{y})] \\ &= \left\{ \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,2}} [a_{i-1}(y_n) - a_{i-1}(t_i) - a'_{i-1}(t_i)(y_n - t_i)] \right. \\ &\quad \left. + \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,2}} [a_i(t_i) - a_i(\bar{y}) - a'_i(\bar{y})(t_i - \bar{y})] + \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,2}} [a'_i(t_i) - a'_i(\bar{y})] (y_n - t_i) \right\} \\ &\quad + \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,2}} [a'_{i-1}(t_i) - a'_i(t_i)] (y_n - t_i) \\ &=: \hat{T}_{y_n, \bar{y}}^{i,2} - \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,2}} [a'_{i-1}(t_i) - a'_i(t_i)] (t_i - y_n) \end{aligned}$$

with  $\Omega_{y_n, \bar{y}}^{i,2} := \Omega_{\bar{y}, i, i}^{(0, \delta)} \cap \Omega_{y_n, i, i}^{(-\delta, 0]} = \{\bar{y} \in (t_i, t_i + \delta), y_n \in (t_i - \delta, t_i)\}$ . For almost every  $x \in \Omega_{y_n, \bar{y}}^{i,2}$ , we have

$$(5.29) \quad 0 \leq \bar{y}(x) - t_i, t_i - y_n(x) \leq |y_n(x) - \bar{y}(x)| \leq \|y_n - \bar{y}\|_{C(\bar{\Omega})} < \delta.$$

Combined with the fact that  $|y_n(x)| \leq M$  for all  $x \in \bar{\Omega}$  and that  $a_{i-1}$  is of class  $C^2$ , we obtain

$$\begin{aligned} \frac{1}{s_n^2} \left| \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,2}}(x) [a_{i-1}(y_n(x)) - a_{i-1}(t_i) - a'_{i-1}(t_i)(y_n(x) - t_i)] \right| &\leq \frac{1}{2s_n^2} C \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,2}}(x) \|y_n - \bar{y}\|_{C(\bar{\Omega})}^2 \\ &\leq \frac{1}{2} C \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,2}}(x) \rightarrow 0 \end{aligned}$$

for almost every  $x \in \Omega$  and for some constant  $C$ . Here we employed the Lipschitz continuity of  $S$  (see Lemma 3.2) and the embedding  $W_0^{1,p}(\Omega) \hookrightarrow C(\bar{\Omega})$  to deduce the last inequality. It therefore holds that

$$(5.30) \quad \frac{1}{s_n^2} \int_{\Omega} \left| \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,2}} [a_{i-1}(y_n) - a_{i-1}(t_i) - a'_{i-1}(t_i)(y_n - t_i)] \nabla \bar{\varphi} \cdot \nabla \bar{y} \right| dx \rightarrow 0.$$

Similarly, we obtain

$$(5.31) \quad \frac{1}{s_n^2} \int_{\Omega} \left| \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,2}} [a_i(t_i) - a_i(\bar{y}) - a'_i(\bar{y})(t_i - \bar{y})] \nabla \bar{\varphi} \cdot \nabla \bar{y} \right| dx \rightarrow 0$$

and

$$(5.32) \quad \frac{1}{s_n^2} \int_{\Omega} \left| \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,2}} [a'_i(t_i) - a'_i(\bar{y})] (y_n - t_i) \nabla \bar{\varphi} \cdot \nabla \bar{y} \right| dx \rightarrow 0.$$

Combining (5.30)–(5.32) thus gives

$$(5.33) \quad \frac{1}{s_n^2} \int_{\Omega} \left| \hat{T}_{y_n, \bar{y}}^{i,2} \nabla \bar{\varphi} \cdot \nabla \bar{y} \right| dx \rightarrow 0.$$

By the same argument as for (5.28) and (5.33),  $T_{y_n, \bar{y}}^{i,3}$  can be written in the form

$$(5.34) \quad T_{y_n, \bar{y}}^{i,3} = \hat{T}_{y_n, \bar{y}}^{i,3} - \mathbb{1}_{\Omega_{y_n, \bar{y}}^{i,3}} [a'_{i+1}(t_{i+1}) - a'_i(t_{i+1})] (t_{i+1} - y_n)$$

with  $\Omega_{y_n, \bar{y}}^{i,3} = \{\bar{y} \in (t_{i+1} - \delta, t_{i+1}), y_n \in [t_{i+1}, t_{i+1} + \delta)\}$ , and  $\hat{T}_{y_n, \bar{y}}^{i,3}$  satisfying

$$(5.35) \quad \frac{1}{s_n^2} \int_{\Omega} \left| \hat{T}_{y_n, \bar{y}}^{i,3} \nabla \bar{\varphi} \cdot \nabla \bar{y} \right| dx \rightarrow 0.$$

By combining (5.22) with the limits (5.27), (5.28), (5.33), (5.34), (5.35) and the definition (5.2), we can conclude that

$$\limsup_{n \rightarrow \infty} \frac{1}{s_n^2} H_n = \frac{1}{2} \int_{\Omega} \mathbb{1}_{\{\bar{y} \notin E_a\}} a''(\bar{y}) (S'(\bar{u})h)^2 \nabla \bar{\varphi} \cdot \nabla \bar{y} dx - \liminf_{n \rightarrow \infty} \frac{1}{s_n^2} \int_{\Omega} A_n(\{s_n\}, \{h_n\}) \nabla \bar{y} \cdot \nabla \bar{\varphi} dx.$$

Together with Corollary 5.5, this gives (iii).  $\square$

### 5.3 SECOND-ORDER CONDITIONS

We now have everything at hand to prove the following two theorems that are the main results of the paper, providing no-gap second-order necessary and sufficient conditions in terms of the curvature functionals  $Q_s$ ,  $Q_1$ , and  $Q_2$  defined in Section 5.1.

**Theorem 5.9 (second-order necessary optimality condition).** *Let Assumptions (A1) to (A4) hold. Assume that  $\bar{u}$  is a local optimal solution of (P) such that  $G'(\bar{y}) \in L^{\bar{p}}(\Omega)$  and  $\Sigma(\bar{y}) < \infty$  for some  $\bar{p} > N$  and  $\bar{y} := S(\bar{u})$ . Then there exists an adjoint state  $\bar{\varphi} \in H_0^1(\Omega) \cap W^{1,\infty}(\Omega)$  that together with  $\bar{u}, \bar{y}$  satisfies (4.6) as well as*

$$(5.36) \quad Q_s(\bar{u}, \bar{y}, \bar{\varphi}; h, h) + Q_1(\bar{u}, \bar{y}, \bar{\varphi}; h, h) + Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h) \geq 0 \quad \text{for all } h \in C(\mathcal{U}_{ad}; \bar{u}).$$

*Proof.* The existence of a  $\bar{\varphi} \in H_0^1(\Omega)$  satisfying (4.6) follows from Theorem 4.3, while the claimed regularity of  $\bar{\varphi}$  follows from Lemma 4.1. It remains to prove (5.36). To this end, let  $h \in C(\mathcal{U}_{ad}; \bar{u})$  and  $\{s_n\} \in c_0^+$  be arbitrary but fixed. We only need to show that

$$(5.37) \quad Q_s(\bar{u}, \bar{y}, \bar{\varphi}; h, h) + Q_1(\bar{u}, \bar{y}, \bar{\varphi}; h, h) + \tilde{Q}(\bar{u}, \bar{y}, \bar{\varphi}; \{s_n\}, h) \geq 0.$$

In order to verify (5.37), we first see from the definition of  $\tilde{Q}(\bar{u}, \bar{y}, \bar{\varphi}; \{s_n\}, h)$  that there exists a subsequence  $\{s_{n_k}\}$  satisfying

$$(5.38) \quad \tilde{Q}(\bar{u}, \bar{y}, \bar{\varphi}; \{s_n\}, h) = \lim_{k \rightarrow \infty} \frac{1}{s_{n_k}^2} \int_{\Omega} A_{n_k}(\{s_{n_k}\}, \{h\}) \nabla \bar{y} \cdot \nabla \bar{\varphi} \, dx.$$

Since  $\mathcal{U}_{ad}$  is polyhedric, there are sequences  $\{h_m\} \subset L^2(\Omega)$ ,  $\{q_m\} \subset \mathcal{U}_{ad}$ , and  $\{\lambda_m\} \subset (0, \infty)$  such that

$$h_m \rightarrow h \quad \text{in } L^2(\Omega), \quad h_m = \frac{q_m - \bar{u}}{\lambda_m}, \quad \text{and} \quad h_m \in \bar{d}^\perp \quad \text{for all } m \in \mathbb{N}.$$

Since  $s_{n_k} \rightarrow 0^+$  as  $k \rightarrow \infty$ , a subsequence, denoted by  $\{r_m\}$ , of  $\{s_{n_k}\}$  exists such that  $0 < r_m \leq \lambda_m$  for all  $m \in \mathbb{N}$ . This and the convexity of  $\mathcal{U}_{ad}$  yield that

$$\bar{u} + r_m h_m = \left(1 - \frac{r_m}{\lambda_m}\right) \bar{u} + \frac{r_m}{\lambda_m} q_m \in \mathcal{U}_{ad} \quad \text{for all } m \in \mathbb{N}.$$

Since  $\bar{u}$  is a local minimizer of (P), it holds that

$$\frac{1}{r_m^2} (j(\bar{u} + r_m h_m) - j(\bar{u})) \geq 0$$

for all  $m \in \mathbb{N}$  large enough. Taking the limit inferior and employing the fact that  $h_m \in \bar{d}^\perp$  and  $h_m \rightarrow h$  in  $L^2(\Omega)$ , we obtain from Lemma 5.2 and Proposition 5.8 that

$$Q_s(\bar{u}, \bar{y}, \bar{\varphi}; h, h) + Q_1(\bar{u}, \bar{y}, \bar{\varphi}; h, h) + \tilde{Q}(\bar{u}, \bar{y}, \bar{\varphi}; \{r_m\}, h) \geq 0.$$

Moreover, we have  $\tilde{Q}(\bar{u}, \bar{y}, \bar{\varphi}; \{r_m\}, h) = \tilde{Q}(\bar{u}, \bar{y}, \bar{\varphi}; \{s_n\}, h)$  as a result of (5.38) and the fact that  $\{r_m\}$  is a subsequence of  $\{s_{n_k}\}$ . This finally gives (5.37).  $\square$

**Theorem 5.10 (second-order sufficient optimality conditions).** *Let Assumptions (A1) to (A4) hold. Assume that  $\bar{u}$  is a feasible point of (P) such that  $G'(\bar{y}) \in L^{\bar{p}}(\Omega)$  and  $\Sigma(\bar{y}) < \infty$  for some  $\bar{p} > N$  and  $\bar{y} := S(\bar{u})$ . Assume further that there exists an adjoint state  $\bar{\varphi} \in H_0^1(\Omega) \cap W^{1,\infty}(\Omega)$  that together with  $\bar{u}, \bar{y}$  satisfies the first-order optimality conditions (4.6) as well as*

$$(5.39) \quad Q_s(\bar{u}, \bar{y}, \bar{\varphi}; h, h) + Q_1(\bar{u}, \bar{y}, \bar{\varphi}; h, h) + Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h) > 0 \quad \text{for all } h \in C(\mathcal{U}_{ad}; \bar{u}) \setminus \{0\}.$$

*Then there exist constants  $c, \rho > 0$  such that*

$$(5.40) \quad j(\bar{u}) + c \|u - \bar{u}\|_{L^2(\Omega)}^2 \leq j(u) \quad \text{for all } u \in \mathcal{U}_{ad} \cap \bar{B}_{L^2(\Omega)}(\bar{u}, \rho).$$

*In particular,  $\bar{u}$  is a strict local minimizer of (P).*

*Proof.* We argue by contradiction. Assume that there exists a sequence  $\{u_n\} \subset \mathcal{U}_{ad}$  such that

$$(5.41) \quad \|u_n - \bar{u}\|_{L^2(\Omega)} < \frac{1}{n} \quad \text{and} \quad j(\bar{u}) + \frac{1}{n} \|u_n - \bar{u}\|_{L^2(\Omega)}^2 > j(u_n), \quad n \in \mathbb{N}.$$

Setting  $s_n := \|u_n - \bar{u}\|_{L^2(\Omega)}$  and  $h_n := \frac{u_n - \bar{u}}{s_n}$  yields that  $\|h_n\|_{L^2(\Omega)} = 1$ . Then there exists a subsequence of  $\{h_n\}$ , also denoted in the same way, such that  $h_n \rightharpoonup h$  in  $L^2(\Omega)$  for some  $h \in L^2(\Omega)$ .

We first verify that  $h \in C(\mathcal{U}_{ad}; \bar{u})$ . First, we have that  $h_n \in \mathcal{R}(\mathcal{U}_{ad}; \bar{u})$  and thus  $h_n(x) \geq 0$  if  $\bar{u}(x) = \alpha(x)$  and  $h_n(x) \leq 0$  if  $\bar{u}(x) = \beta(x)$  for almost every  $x \in \Omega$ . From this and  $h_n \rightharpoonup h$ , we deduce that  $h(x) \geq 0$  if  $\bar{u}(x) = \alpha(x)$  and  $h(x) \leq 0$  if  $\bar{u}(x) = \beta(x)$  for almost every  $x \in \Omega$ . Consequently, it holds that  $h \in \mathcal{T}(\mathcal{U}_{ad}; \bar{u})$ . Since  $j$  is continuously differentiable as a function from  $L^2(\Omega)$  to  $\mathbb{R}$  according to [Theorem 4.2](#), a Taylor expansion thus gives

$$j(u_n) = j(\bar{u}) + j'(\bar{u})(u_n - \bar{u}) + o(\|u_n - \bar{u}\|_{L^2(\Omega)}).$$

This, together with the last inequality in (5.41), implies, for  $n$  large enough, that

$$j'(\bar{u})(u_n - \bar{u}) + o(\|u_n - \bar{u}\|_{L^2(\Omega)}) < \frac{1}{n} s_n^2.$$

Dividing the above inequality by  $s_n$  and then passing to the limit, we have  $j'(\bar{u})h \leq 0$ . Furthermore, it follows from (4.6b) that  $j'(\bar{u})v \geq 0$  for all  $v \in \mathcal{U}_{ad} - \bar{u}$  and thus for all  $v \in \mathcal{T}(\mathcal{U}_{ad}; \bar{u})$ . In particular, we have  $j'(\bar{u})h \geq 0$  and thus  $j'(\bar{u})h = 0$ . Hence, it holds that  $h \in \bar{d}^\perp$  and so  $h \in C(\mathcal{U}_{ad}; \bar{u})$ .

We now obtain a contradiction and thus complete the proof. Indeed, from the last inequality in (5.41), we obtain

$$\frac{1}{s_n^2} [j(u_n) - j(\bar{u})] < \frac{1}{n}.$$

Combining this with (4.6b), (5.2), and [Lemma 5.2](#) yields that

$$\begin{aligned} \frac{1}{n} &> \frac{1}{s_n^2} \left\{ \int_0^1 (1-s)G''(\bar{y} + s(y_n - \bar{y}))(y_n - \bar{y})^2 ds + \frac{\nu}{2} s_n^2 \|h_n\|_{L^2(\Omega)}^2 \right. \\ &\quad - \int_{\Omega} [a(y_n) - a(\bar{y})] \nabla \bar{\varphi} \cdot (\nabla y_n - \nabla \bar{y}) dx \\ &\quad \left. - \int_{\Omega} [a(y_n) - a(\bar{y}) - \mathbb{1}_{\{\bar{y} \notin E_a\}} a'(\bar{y})(y_n - \bar{y})] \nabla \bar{\varphi} \cdot \nabla \bar{y} dx \right\} \end{aligned}$$

with  $y_n := S(u_n)$ . Taking the limit inferior, employing [Proposition 5.8](#), and using that  $h_n \rightharpoonup h$  in  $L^2(\Omega)$ , we arrive at

$$0 \geq Q_s(\bar{u}, \bar{y}, \bar{\varphi}; h, h) + Q_1(\bar{u}, \bar{y}, \bar{\varphi}; h, h) + \tilde{Q}(\bar{u}, \bar{y}, \bar{\varphi}; \{s_n\}, h) + \frac{\nu}{2} (1 - \|h\|_{L^2(\Omega)}^2).$$

Consequently,

$$(5.42) \quad 0 \geq Q_s(\bar{u}, \bar{y}, \bar{\varphi}; h, h) + Q_1(\bar{u}, \bar{y}, \bar{\varphi}; h, h) + Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h) + \frac{\nu}{2} (1 - \|h\|_{L^2(\Omega)}^2).$$

Since the norm in  $L^2(\Omega)$  is weakly lower semicontinuous, there holds  $\|h\|_{L^2(\Omega)} \leq 1$  by definition of  $h_n$ . From this, (5.42), and (5.39), we have  $h = 0$ . Inserting  $h = 0$  into (5.42) and exploiting the fact that  $Q_2(\bar{u}, \bar{y}, \bar{\varphi}; 0) = 0$  yields that  $0 \geq \frac{\nu}{2} > 0$ , which is impossible.  $\square$

We finish this section by providing another version of the sufficient second-order optimality conditions that are equivalent to (5.39) and could be useful for estimating discretization errors in finite element approximations. The proof of the next result is partly based on [13, Thm. 4.4] with some modifications.

**Proposition 5.11.** *Let Assumptions (A1) to (A4) hold. Assume that  $\bar{u}$  is a feasible point of (P) such that  $G'(\bar{y}) \in L^{\bar{p}}(\Omega)$  and  $\Sigma(\bar{y}) < \infty$  for some  $\bar{p} > N$  and  $\bar{y} := S(\bar{u})$ . Assume further that there exists an adjoint state  $\bar{\varphi} \in H_0^1(\Omega) \cap W^{1,\infty}(\Omega)$  that together with  $\bar{u}, \bar{y}$  satisfies (4.6). Then, (5.39) holds if and only if there exist constants  $c_0, \tau > 0$  such that*

$$(5.43) \quad Q_s(\bar{u}, \bar{y}, \bar{\varphi}; h, h) + Q_1(\bar{u}, \bar{y}, \bar{\varphi}; h, h) + Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h) \geq c_0 \|h\|_{L^2(\Omega)}^2 \quad \text{for all } h \in C^\tau(\mathcal{U}_{ad}; \bar{u})$$

for

$$C^\tau(\mathcal{U}_{ad}; \bar{u}) := \left\{ h \in L^2(\Omega) \left| h(x) \begin{cases} \geq 0 & \text{if } \bar{u}(x) = \alpha(x) \\ \leq 0 & \text{if } \bar{u}(x) = \beta(x) \text{ a.e. } x \in \Omega \\ = 0 & \text{if } |\bar{d}(x)| > \tau \end{cases} \right. \right\}.$$

*Proof.* Since the inclusion  $C(\mathcal{U}_{ad}; \bar{u}) \subset C^\tau(\mathcal{U}_{ad}; \bar{u})$  holds, the inequality (5.39) is a direct consequence of (5.43). We thus only need to prove the implication “(5.39)  $\Rightarrow$  (5.43)”. To this end, assume that (5.43) is not fulfilled. Then, for any  $n \in \mathbb{N}$ , there exists  $v_n \in C^{1/n}(\mathcal{U}_{ad}; \bar{u})$  such that

$$Q_s(\bar{u}, \bar{y}, \bar{\varphi}; v_n, v_n) + Q_1(\bar{u}, \bar{y}, \bar{\varphi}; v_n, v_n) + Q_2(\bar{u}, \bar{y}, \bar{\varphi}; v_n) < \frac{1}{n} \|v_n\|_{L^2(\Omega)}^2.$$

Dividing this inequality by  $\|v_n\|_{L^2(\Omega)}^2$ , using the positive homogeneity of  $(Q_s + Q_1)(\bar{u}, \bar{y}, \bar{\varphi}; \cdot, \cdot)$  as well as Lemma 5.1, and setting  $h_n := \frac{v_n}{\|v_n\|_{L^2(\Omega)}}$ , we have that  $h_n \in C^{1/n}(\mathcal{U}_{ad}; \bar{u})$ ,  $\|h_n\|_{L^2(\Omega)} = 1$  and

$$(5.44) \quad Q_s(\bar{u}, \bar{y}, \bar{\varphi}; h_n, h_n) + Q_1(\bar{u}, \bar{y}, \bar{\varphi}; h_n, h_n) + Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h_n) < \frac{1}{n}.$$

Since  $\{h_n\}$  is bounded in  $L^2(\Omega)$ , there exists a subsequence, denoted in the same way, such that  $h_n \rightharpoonup h$  in  $L^2(\Omega)$  for some  $h \in L^2(\Omega)$ . Obviously, it holds that  $h \in C(\mathcal{U}_{ad}; \bar{u})$ . The compact embedding  $L^2(\Omega) \Subset W^{-1,p}(\Omega)$  for any  $p \in (N, 6)$  implies that  $h_n \rightarrow h$  in  $W^{-1,p}(\Omega)$ . Thus, we have  $S'(\bar{u})h_n \rightarrow S'(\bar{u})h$  in  $W_0^{1,p}(\Omega)$  and so in  $H_0^1(\Omega) \cap C(\bar{\Omega})$ . From this, the weak lower semicontinuity of the norm in  $L^2(\Omega)$ , the estimate (5.44), the definition of  $Q_s$  and  $Q_1$  and the weak lower semicontinuity of  $Q_2(\bar{u}, \bar{y}, \bar{\varphi}; \cdot)$  from Proposition 5.6 we deduce that

$$(5.45) \quad \begin{aligned} (Q_s + Q_1)(\bar{u}, \bar{y}, \bar{\varphi}; h, h) + Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h) &\leq \liminf_{n \rightarrow \infty} (Q_s + Q_1)(\bar{u}, \bar{y}, \bar{\varphi}; h_n, h_n) + \liminf_{n \rightarrow \infty} Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h_n) \\ &\leq \liminf_{n \rightarrow \infty} [(Q_s + Q_1)(\bar{u}, \bar{y}, \bar{\varphi}; h_n, h_n) + Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h_n)] \\ &\leq 0. \end{aligned}$$

From this and (5.39), it holds that  $h = 0$ . Again, we see from the definition of  $Q_s$  and  $Q_1$  and the fact that  $\|h_n\|_{L^2(\Omega)} = 1$  that

$$\begin{aligned} (Q_s + Q_1)(\bar{u}, \bar{y}, \bar{\varphi}; h_n, h_n) &= \frac{1}{2} G''(\bar{y})(S'(\bar{u})h_n)^2 + \frac{\nu}{2} - \int_{\Omega} a'(\bar{y}; S'(\bar{u})h_n) \nabla(S'(\bar{u})h_n) \cdot \nabla \bar{\varphi} \, dx \\ &\quad - \frac{1}{2} \int_{\Omega} \mathbb{1}_{\{\bar{y} \notin E_a\}} a''(\bar{y})(S'(\bar{u})h_n)^2 \nabla \bar{\varphi} \cdot \nabla \bar{y} \, dx. \end{aligned}$$

Combining this with the fact that  $\nabla(S'(\bar{u})h_n) \rightarrow 0$  in  $L^p(\Omega)^N$  and  $S'(\bar{u})h_n \rightarrow 0$  in  $C(\bar{\Omega}) \cap H_0^1(\Omega)$ , we can conclude from the dominated convergence theorem and Assumption (A4) that

$$\begin{aligned} \frac{\nu}{2} &= \lim_{n \rightarrow \infty} (Q_s + Q_1)(\bar{u}, \bar{y}, \bar{\varphi}; h_n, h_n) \leq \limsup_{n \rightarrow \infty} [-Q_2(\bar{u}, \bar{y}, \bar{\varphi}; h_n)] \\ &\leq \Sigma(\bar{y}) \|\nabla \bar{\varphi}\|_{L^\infty(\Omega)} \lim_{n \rightarrow \infty} \|S'(\bar{u})h_n\|_{L^\infty(\Omega)}^2 = 0, \end{aligned}$$

where we have used the limit (5.45) and Lemma 5.7 to obtain the last two estimates. This gives the desired contradiction and completes the proof.  $\square$

## 6 CONCLUSIONS

We have derived second-order optimality conditions for an optimal control problem governed by a quasilinear elliptic differential equation with a finitely  $PC^2$  coefficient. Showing that the control-to-state operator is in fact Fréchet differentiable (but in general not twice differentiable) allows using a second-order Taylor-type expansion to formulate necessary and sufficient conditions in terms of a new curvature functional related to the jump of the first derivatives of the non-smooth coefficients in critical points. These are no-gap conditions in the sense that the only difference between necessary and sufficient conditions lies in the fact that the inequality in the latter is strict. Furthermore, an equivalent formulation of the second-order sufficient optimality condition that could be used for discretization error estimates is also derived. Such estimates will be studied in a follow-up work.

## APPENDIX A REGULARITY OF A QUASILINEAR EQUATION ON CONVEX DOMAINS

**Lemma A.1.** *Let Assumptions (A1) to (A3) hold and  $q > N$ . If  $y_u \in W_0^{1,2q}(\Omega)$  is the unique solution to*

$$(A.1) \quad \begin{cases} -\operatorname{div}[(b + a(y_u)) \nabla y_u] = u & \text{in } \Omega, \\ y_u = 0 & \text{on } \partial\Omega, \end{cases}$$

with  $u \in L^q(\Omega)$  and satisfies  $\|y_u\|_{C(\bar{\Omega})} \leq M$  for some  $M > 0$ , then  $y_u \in H^2(\Omega) \cap W^{1,\infty}(\Omega)$  and

$$(A.2) \quad \|y_u\|_{H^2(\Omega)} + \|y_u\|_{W^{1,\infty}(\Omega)} \leq C \left( q, M, \|u\|_{L^q(\Omega)}, \|y_u\|_{W_0^{1,2q}(\Omega)} \right).$$

*Proof.* Define the function

$$a^M : \mathbb{R} \rightarrow \mathbb{R}, \quad a^M(t) := \begin{cases} a(t) & \text{if } |t| \leq 2M, \\ a(2M) & \text{if } t > 2M, \\ a(-2M) & \text{if } t < -2M. \end{cases}$$

Obviously,  $a^M$  is a  $PC^1$ -function and  $\nabla a^M \in L^\infty(\mathbb{R})^N$ . By the chain rule (see, e.g., [24, Thm. 7.8]), it holds that

$$\frac{\partial a(y_u)}{\partial x_i} = \frac{\partial a^M(y_u)}{\partial x_i} = \mathbb{1}_{\{y_u \notin E_a\}} a'(y_u) \frac{\partial y_u}{\partial x_i}.$$

Moreover, by employing Assumption (A3), we deduce that  $|\mathbb{1}_{\{y_u \notin E_a\}} a'(y_u)| \leq C_M$  for almost every  $x \in \Omega$  and for some constant  $C_M > 0$ . Consider now the equation

$$(A.3) \quad \begin{cases} -\Delta \tilde{y} = \frac{1}{b + a(y_u)} \left[ u + \nabla b \cdot \nabla y_u + \mathbb{1}_{\{y_u \notin E_a\}} a'(y_u) |\nabla y_u|^2 \right] & \text{in } \Omega, \\ \tilde{y} = 0 & \text{on } \partial\Omega. \end{cases}$$

Since the right-hand side of (A.3) belongs to  $L^q(\Omega)$  for  $q > N \geq 2$ , it holds that  $\tilde{y} \in H^2(\Omega)$  according to [25, Thm. 3.2.1.2]. Furthermore, we have from the fact that  $b(x) \geq \underline{b} > 0$  for all  $x \in \bar{\Omega}$  and the non-negativity of  $a$  that

$$\begin{aligned} \|\Delta \tilde{y}\|_{L^q(\Omega)} &\leq \frac{1}{\underline{b}} \left[ \|u\|_{L^q(\Omega)} + \|\nabla b \cdot \nabla y_u\|_{L^q(\Omega)} + \|\mathbb{1}_{\{y_u \notin E_a\}} a'(y_u) |\nabla y_u|^2\|_{L^q(\Omega)} \right] \\ &\leq \frac{1}{\underline{b}} \left[ \|u\|_{L^q(\Omega)} + L_b \|\nabla y_u\|_{L^q(\Omega)} + C_M \|\nabla y_u\|_{L^{2q}(\Omega)}^2 \right]. \end{aligned}$$

From this, [Assumption \(A1\)](#), and the global boundedness of the gradient of solutions to Poisson's equation (see, e.g. [19, Thm. 3.1, Rem. 3], [30, 39]), it follows that  $\tilde{y} \in W^{1,\infty}(\Omega)$  and

$$\|\nabla \tilde{y}\|_{L^\infty(\Omega)} \leq C \|\Delta \tilde{y}\|_{L^q(\Omega)}.$$

It is therefore sufficient to prove that  $\tilde{y} = y_u$ . To this end, taking any  $\phi \in H_0^1(\Omega)$  yields that  $(b+a(y_u))\phi \in H_0^1(\Omega)$ . Testing [\(A.3\)](#) by  $(b+a(y_u))\phi$ , a straightforward computation shows that

$$\int_{\Omega} (\nabla \tilde{y} - \nabla y_u) \cdot \nabla (b+a(y_u))\phi + (b+a(y_u))\nabla \phi \cdot \nabla \tilde{y} \, dx = \int_{\Omega} u\phi \, dx.$$

Now testing [\(A.1\)](#) by  $\phi$  and then subtracting the obtained identity from the above equality, we obtain

$$\int_{\Omega} (\nabla \tilde{y} - \nabla y_u) \cdot \nabla [(b+a(y_u))\phi] \, dx = 0.$$

Choosing  $\phi := \frac{1}{b+a(y_u)}(\tilde{y} - y_u) \in H_0^1(\Omega)$  then yields  $\tilde{y} = y_u$ .  $\square$

## APPENDIX B ESTIMATE OF JUMP FUNCTIONAL

In this appendix, we verify the estimates for  $\Sigma(y)$  in [Example 5.3](#).

We can assume without loss of generality that  $E_a = \{t\}$ , i.e.,  $K = 1$  and  $t_1 = t$ . We first consider the right end point  $\bar{x}_j$  of  $[\underline{x}_j, \bar{x}_j]$ . If  $\bar{x}_j = \beta_0$ , then the set  $\{0 < |y-t| < r\} \cap (\bar{x}_j, \infty)$  is empty for any  $r > 0$ , and thus

$$\lim_{r \rightarrow 0^+} \frac{1}{r} \int_{\Omega} \mathbb{1}_{\{0 < |y-t| < r\} \cap (\bar{x}_j, \infty)} \left| \frac{\partial y}{\partial x} \right| \, dx = 0.$$

If  $\bar{x}_j < \beta_0$ , then there exists an  $\bar{\varepsilon}_j > 0$  such that  $y$  is monotone on  $(\bar{x}_j, \bar{x}_j + \bar{\varepsilon}_j) \subset \Omega$  due the monotonicity assumption on  $y$ . Since  $\bar{x}_j$  is the right end point of  $[\underline{x}_j, \bar{x}_j]$ , the sets  $[\underline{x}_k, \bar{x}_k]$ ,  $k \in J$ , are mutually disjoint, and  $y$  is continuous, we have  $y(\bar{x}_j + \bar{\varepsilon}_j) \neq y(\bar{x}_j) = t$ . Setting  $\bar{r}_j := |y(\bar{x}_j + \bar{\varepsilon}_j) - y(\bar{x}_j)| > 0$ , we derive from the monotonicity of  $y$  on  $(\bar{x}_j, \bar{x}_j + \bar{\varepsilon}_j)$  that for all  $r \in (0, \bar{r}_j)$ ,

$$\{0 < |y-t| < r\} \cap (\bar{x}_j, \bar{x}_j + \bar{\varepsilon}_j) = \begin{cases} (\bar{x}_j, y^{-1}(t+r)) & \text{if } y \text{ is monotone increasing on } (\bar{x}_j, \bar{x}_j + \bar{\varepsilon}_j), \\ (\bar{x}_j, y^{-1}(t-r)) & \text{if } y \text{ is monotone decreasing on } (\bar{x}_j, \bar{x}_j + \bar{\varepsilon}_j). \end{cases}$$

This implies that

$$\lim_{r \rightarrow 0^+} \frac{1}{r} \int_{\Omega} \mathbb{1}_{\{0 < |y-t| < r\} \cap (\bar{x}_j, \bar{x}_j + \bar{\varepsilon}_j)} \left| \frac{\partial y}{\partial x} \right| \, dx = 1.$$

Similarly, we have for the left end point  $\underline{x}_j$  that there exists an  $\underline{\varepsilon}_j > 0$  such that

$$\lim_{r \rightarrow 0^+} \frac{1}{r} \int_{\Omega} \mathbb{1}_{\{0 < |y-t| < r\} \cap (\underline{x}_j - \underline{\varepsilon}_j, \underline{x}_j)} \left| \frac{\partial y}{\partial x} \right| \, dx = \begin{cases} 0 & \text{if } \underline{x}_j = \alpha_0, \\ 1 & \text{if } \underline{x}_j > \alpha_0. \end{cases}$$

Furthermore, if  $\underline{x}_j > \alpha_0$ , then  $y$  is monotone on  $(\underline{x}_j - \underline{\varepsilon}_j, \underline{x}_j) \subset \Omega$  and

$$\{0 < |y-t| < r\} \cap (\underline{x}_j - \underline{\varepsilon}_j, \underline{x}_j) = \begin{cases} (y^{-1}(t-r), \underline{x}_j) & \text{if } y \text{ is monotone increasing on } (\underline{x}_j - \underline{\varepsilon}_j, \underline{x}_j), \\ (y^{-1}(t+r), \underline{x}_j) & \text{if } y \text{ is monotone decreasing on } (\underline{x}_j - \underline{\varepsilon}_j, \underline{x}_j), \end{cases}$$

for all  $r \in (0, \underline{r}_j)$  with  $\underline{r}_j := |y(\underline{x}_j - \underline{\varepsilon}_j) - y(\underline{x}_j)| > 0$ . Let us set  $\varepsilon_j := \min\{\underline{\varepsilon}_j, \bar{\varepsilon}_j\}$  and  $r_j := \min\{\underline{r}_j, \bar{r}_j\}$ .

If  $y$  is monotone increasing on  $(\bar{x}_j, \bar{x}_j + \bar{\varepsilon}_j)$  and on  $(\underline{x}_j - \underline{\varepsilon}_j, \underline{x}_j)$ , then

$$\{0 < |y-t| < r\} \cap (\underline{x}_j - \underline{\varepsilon}_j, \bar{x}_j + \varepsilon_j) = (\bar{x}_j, y^{-1}(t+r)) \cup (y^{-1}(t-r), \underline{x}_j)$$

for all  $r \in (0, r_j)$ . Similar expressions hold for the other cases of monotonicity on each interval. We hence obtain

$$\lim_{r \rightarrow 0^+} \frac{1}{r} \int_{\Omega} \mathbb{1}_{\{0 < |y-t| < r\} \cap (\underline{x}_j - \varepsilon_j, \bar{x}_j + \varepsilon_j)} \left| \frac{\partial y}{\partial x} \right| dx = \begin{cases} 0 & \text{if } \underline{x}_j, \bar{x}_j \in \{\alpha_0, \beta_0\}, \\ 1 & \text{if } \underline{x}_j \in \{\alpha_0, \beta_0\}, \bar{x}_j \notin \{\alpha_0, \beta_0\} \text{ or vice versa,} \\ 2 & \text{if } \underline{x}_j, \bar{x}_j \notin \{\alpha_0, \beta_0\}. \end{cases}$$

A standard argument then yields the desired conclusions.

## REFERENCES

- [1] H. Attouch, G. Buttazzo, and G. Michaille, *Variational Analysis in Sobolev and BV Spaces*, MOS-SIAM Series on Optimization, SIAM, 2006, [doi:10.1137/1.9781611973488](https://doi.org/10.1137/1.9781611973488).
- [2] T. Bayen, J. F. Bonnans, and F. J. Silva, Characterization of local quadratic growth for strong minima in the optimal control of semi-linear elliptic equations, *Trans. Amer. Math. Soc.* 366 (2014), 2063–2087, [doi:10.1090/S0002-9947-2013-05961-2](https://doi.org/10.1090/S0002-9947-2013-05961-2).
- [3] T. Bayen and F. J. Silva, Second order analysis for strong solutions in the optimal control of parabolic equations, *SIAM J. Control Optim.* 54 (2016), 819–844, [doi:10.1137/141000415](https://doi.org/10.1137/141000415).
- [4] L. M. Betz, Second-order sufficient optimality conditions for optimal control of nonsmooth, semi-linear parabolic equations, *SIAM J. Control Optim.* 57 (2019), 4033–4062, [doi:10.1137/19m1239106](https://doi.org/10.1137/19m1239106).
- [5] T. Betz and C. Meyer, Second-order sufficient optimality conditions for optimal control of static elastoplasticity with hardening, *ESAIM: COCV* 21 (2015), 271–300, [doi:10.1051/cocv/2014024](https://doi.org/10.1051/cocv/2014024).
- [6] A. Binder, On an inverse problem arising in continuous casting of steel billets, *Appl. Anal.* 57 (1995), 341–366, [doi:10.1080/00036819508840356](https://doi.org/10.1080/00036819508840356).
- [7] J. F. Bonnans, Second-order analysis for control constrained optimal control problems of semilinear elliptic systems, *Appl. Math. Optim.* 38 (1998), 305–325, [doi:10.1007/s002459900093](https://doi.org/10.1007/s002459900093).
- [8] J. F. Bonnans and Housnaa, Optimal control problems with partially polyhedral constraints, *SIAM J. Control Optim.* 37 (1999), 1726–1741, [doi:10.1137/s0363012998333724](https://doi.org/10.1137/s0363012998333724).
- [9] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer-Verlag, Berlin, Heidelberg, 2000, [doi:10.1007/978-1-4612-1394-9](https://doi.org/10.1007/978-1-4612-1394-9).
- [10] E. Casas, Necessary and sufficient optimality conditions for elliptic control problems with finitely many pointwise state constraints, *ESAIM: COCV* 14 (2008), 575–589, [doi:10.1051/cocv:2007063](https://doi.org/10.1051/cocv:2007063).
- [11] E. Casas and V. Dharmo, Error estimates for the numerical approximation of a quasilinear Neumann problem under minimal regularity of the data, *Numer. Math.* 117 (2011), 115–145, [doi:10.1007/s00211-010-0344-1](https://doi.org/10.1007/s00211-010-0344-1).
- [12] E. Casas and V. Dharmo, Optimality conditions for a class of optimal boundary control problems with quasilinear elliptic equations, *Control Cybernet.* 40 (2011), 457–490.
- [13] E. Casas and M. Mateos, Second order optimality conditions for semilinear elliptic control problems with finitely many state constraints, *SIAM J. Control Optim.* 40 (2002), 1431–1454, [doi:10.1137/s0363012900382011](https://doi.org/10.1137/s0363012900382011).

- [14] E. Casas and F. Tröltzsch, First- and second-order optimality conditions for a class of optimal control problems with quasilinear elliptic equations, *SIAM J. Control Optim.* 48 (2009), 688–718, doi:10.1137/080720048.
- [15] E. Casas and F. Tröltzsch, Second order optimality conditions and their role in PDE control, *Jahresber. Dtsch. Math.-Ver.* 117 (2015), 3–44, doi:10.1365/s13291-014-0109-3.
- [16] M. Chipot, *Elliptic Equations: An Introductory Course*, Birkhäuser Verlag, Basel, 2009, doi:10.1007/978-3-7643-9982-5.
- [17] C. Christof and G. Wachsmuth, No-gap second-order conditions via a directional curvature functional, *SIAM Journal on Optimization* 28 (2018), 2097–2130, doi:10.1137/17m1140418.
- [18] C. Christof and G. Wachsmuth, On second-order optimality conditions for optimal control problems governed by the obstacle problems, *Optimization* (2020), 1–41, doi:10.1080/02331934.2020.1778686.
- [19] A. Cianchi and V. Maz’ya, Global gradient estimates in elliptic problems under minimal data and domain regularity, *Commun. Pure Appl. Anal.* 14 (2015), 285–311, doi:10.3934/cpaa.2015.14.285.
- [20] C. Clason, V.H. Nhu, and A. Rösch, Optimal control of a non-smooth quasilinear elliptic equation, *Mathematical Control and Related Fields* to appear, (2018), arXiv:1810.08007.
- [21] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mappings*, Springer Monographs in Mathematics, Springer, 2014, doi:10.1007/978-1-4939-1037-3.
- [22] H. W. Engl and T. Langthaler, Control of the solidification front by secondary cooling in continuous casting of steel, in *Case studies in industrial mathematics*, volume 2 of European Consort. Math. Indust., Teubner, Stuttgart, 1988, 51–77, doi:10.1007/978-3-663-12063-6\_3.
- [23] S. J. Fromm, Potential space estimates for Green potentials in convex domains, *Proc. Amer. Math. Soc.* 119 (1993), 225–233, doi:10.1090/s0002-9939-1993-1156467-3.
- [24] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin, Heidelberg, 2001, doi:10.1007/978-3-642-61798-0.
- [25] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman Advanced Pub. Program, 1985, doi:10.1137/1.9781611972030.
- [26] B. T. Kien, V.H. Nhu, and N. H. Son, Second-order optimality conditions for a semilinear elliptic optimal control problem with mixed pointwise constraints, *Set-Valued Var. Anal* 25 (2017), 177–210, doi:10.1007/s11228-016-0373-8.
- [27] K. Krumbiegel, I. Neitzel, and A. Rösch, Sufficient optimality conditions for the Moreau-Yosida-type regularization concept applied to semilinear elliptic optimal control problems with pointwise state constraints, *Ann. Acad. Rom. Sci. Ser. Math. Appl.* 2 (2010), 222–246.
- [28] 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 (2012), 520–547, doi:10.1051/cocv/2011105.
- [29] P. Kügler, Identification of a temperature dependent heat conductivity from single boundary measurements, *SIAM J. Numer. Anal.* 41 (2003), 1543–1563, doi:10.1137/s0036142902415900.

- [30] V. Maz'ya, Boundedness of the gradient of a solution to the Neumann–Laplace problem in a convex domain, *C. R. Math. Acad. Sci. Paris* 347 (2009), 517–520, [doi:10.1016/j.crma.2009.03.001](https://doi.org/10.1016/j.crma.2009.03.001).
- [31] 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.* 42 (2003), 138–154, [doi:10.1137/S0363012902403262](https://doi.org/10.1137/S0363012902403262).
- [32] A. Rösch and F. Tröltzsch, Sufficient second-order optimality conditions for an elliptic optimal control problem with pointwise control-state constraints, *SIAM J. Optim.* 17 (2006), 776–794, [doi:10.1137/050625850](https://doi.org/10.1137/050625850).
- [33] S. Scholtes, *Introduction to Piecewise Differentiable Equations*, Springer Science & Business Media, 2012, [doi:10.1007/978-1-4614-4340-7](https://doi.org/10.1007/978-1-4614-4340-7).
- [34] F. Tröltzsch, *Optimal Control of Partial Differential Equations*, volume 112 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2010, [doi:10.1090/gsm/112](https://doi.org/10.1090/gsm/112). Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.
- [35] N.D. Tuan, Sequence-based necessary second-order optimality conditions for semilinear elliptic optimal control problems with nonsmooth data, *Positivity* 23 (2019), 195–217, [doi:10.1007/s11117-018-0602-5](https://doi.org/10.1007/s11117-018-0602-5).
- [36] M. Ulbrich, *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*, MOS-SIAM Series on Optimization, SIAM, 2011, [doi:10.1137/1.9781611970692](https://doi.org/10.1137/1.9781611970692).
- [37] A. Visintin, *Models of Phase Transitions*, Birkhäuser, Boston, 1996, [doi:10.1007/978-1-4612-4078-5](https://doi.org/10.1007/978-1-4612-4078-5).
- [38] G. Wachsmuth, Differentiability of implicit functions: Beyond the implicit function theorem, *J. Math. Anal. Appl.* 414 (2014), 259–272, [doi:10.1016/j.jmaa.2014.01.007](https://doi.org/10.1016/j.jmaa.2014.01.007).
- [39] S. Yang, D. C. Chang, D. Yang, and Z. Fu, Gradient estimates via rearrangements for solutions of some Schrödinger equations, *Anal. Appl.* 16 (2018), 339–361, [doi:10.1142/S0219530517500142](https://doi.org/10.1142/S0219530517500142).
- [40] E. Zeidler, *Nonlinear Functional Analysis and its Applications. I: Fixed-Point Theorems*, Springer-Verlag, New York, 1986.