ERROR ESTIMATES FOR THE NUMERICAL APPROXIMATION OF A NON-SMOOTH QUASILINEAR ELLIPTIC CONTROL PROBLEM

Christian Clason* Vu Huu Nhu† Arnd Rösch‡

Abstract In this paper, we carry out the numerical analysis of a non-smooth quasilinear elliptic optimal control problem, where the coefficient in the divergence term of the corresponding state equation is a finitely PC^2 (continuous and C^2 apart from finitely many points) function in the state variable. Although the nonlinearity of the quasilinear elliptic equation is non-smooth, the corresponding control-to-state operator is of class C^1 but not of class C^2. Analogously, the discrete control-to-state operators associated with the approximated control problems are proven to be of class C^1 only. An explicit formula of a second-order generalized derivative of the cost functional is also established. We then exploit a second-order sufficient optimality condition to prove a priori error estimates for a variational and a piecewise constant approximation of the continuous optimal control problem.

Key words Optimal control, non-smooth optimization, quasilinear elliptic equation, piecewise differentiable function, sufficient optimality condition, error estimate, finite element approximation.

1 INTRODUCTION

We investigate the non-smooth quasilinear elliptic optimal control problem

$$
\text{(P)} \qquad \begin{cases} \min\limits_{u \in L^\infty(\Omega)} j(u) := \displaystyle\int_\Omega L(x, y_u(x))\,\mathrm{d}x + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} \\ \text{s.t.} \quad -\operatorname{div}[(b + a(y_u))\nabla y_u] = u \quad \text{in}\,\Omega, \quad y_u = 0 \,\text{on}\,\partial\Omega, \\ \qquad \alpha \le u(x) \le \beta \quad \text{a.e.}\, x \in \Omega, \end{cases}
$$

where Ω is a bounded, convex and polygonal domain $\Omega \subset \mathbb{R}^2$; $L : \Omega \times \mathbb{R} \to \mathbb{R}$ is a Carathéodory function of class C^2 with respect to (w.r.t.) the second variable; $b : \overline{\Omega} \to \mathbb{R}$ is a Lipschitz continuous function; $a : \mathbb{R} \to \mathbb{R}$ is a non-smooth function; and $\alpha, \beta, \nu \in \mathbb{R}$ satisfy $\beta > \alpha$ and $\nu > 0$. For the precise hypotheses on the data of (P), we refer to Section 2.

The control problem (P) is interesting since the corresponding state equation arises, for instance, in models of heat conduction where the coefficient in the divergence term of the state equation is the heat conductivity and depends on the temperature y and on the spatial coordinate x; see, e.g. [4, 31]. When the data are of class C^2, the numerical analysis of the discrete approximation of such optimal control problems was investigated by Casas et al. in [11, 12] for Dirichlet boundary type and in [8] for Neumann boundary type.

*Faculty of Mathematics, Universität Duisburg-Essen, 45117 Essen, Germany; *current address:* Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, 8010 Graz, Austria (c.clason@uni-graz.at, ORCID: 0000-0002-9948-8426)

†Faculty of Fundamental Sciences, PHENIKAA University, Yen Nghia, Ha Dong, Hanoi 12116, Vietnam (nhu.vuhuu@phenikaa-uni.edu.vn, ORCID: 0000-0003-4279-3937)

‡Faculty of Mathematics, University of Duisburg-Essen, Thea-Leymann-Strasse 9, 45127 Essen, Germany (arnd.roesch@uni-due.de)

Let us briefly comment on other works concerning the error analysis of optimal control problems governed by partial differential equations (PDEs), in particular by elliptic PDEs. For control-constrained elliptic problems, we refer to the early papers [21, 22] for linear elliptic control problems; to [2, 10] for semilinear elliptic problems. For state-constrained control problems, we mention only the recent contributions [19, 29] and refer to the survey paper [26] for further references. Although the error analysis for smooth PDE-constrained problems has been intensively investigated, there are very few contributions on this topic for non-smooth PDE-constrained optimal control. Here we want to mention the work [27] concerning error analysis for optimal control of a coupled PDE–ODE system, where the non-smooth nonlinearity acts on a semilinear elliptic ODE. For works related to optimal control of obstacle problems, we refer to [28, 14] and the references therein. Based on a quadratic growth condition, a priori error estimates were established in [27, 28]. To the authors' best knowledge, this is the first work that exploits a second-order sufficient optimality condition to show a priori error estimates for the discretization of optimal control problems governed by non-smooth PDEs.

In this paper, our main aim is to derive the convergence analysis and error estimates of the discretization of (P) under the second-order sufficient optimality conditions [16]. Here, the coefficient a in the state equation is assumed to be a PC^2-function that has a finite set of non-differentiability points. As a consequence, there are two major difficulties in deriving the error analysis. The first issue arises in studying error estimates of the discretization of the adjoint state equation; the other is the lack of the second-order differentiability of the cost functional. We therefore cannot apply an abstract theorem on error estimates shown in [12]. To deal with the first issue, we introduce the function $Z_{y,\hat{y}}$ defined in (4.5) that pointwise measures the difference between the gradients of the superposition mappings of a associated with two distinct states y and $\hat{y}$. This allows us to derive thereafter both an L^2- and an H^1_0- error estimates for the approximation of the adjoint state equation. For handling the second issue, we will exploit an assumption on the optimal state in order to establish an explicit formula of a second-order generalized derivative of the objective functional. Based on the second-order sufficient optimality conditions for (P) from [16], we then prove a general error estimate for variational and piecewise constant approximations of the optimal control, which generalizes the one in [12, Thm. 2.14] and can be applied for the case where the cost functional j is of class C^1 but not necessarily C^2; see Theorem 6.8.

The plan of the paper is as follows. This section ends with our notation. In the next section, we make the assumptions for (P) and provide some preliminary results from [16]. An explicit formula of the curvature functional (a second-order generalized derivative) of the objective functional is stated in Section 3. Section 4 is devoted to the numerical approximation of the state equation by finite elements and the local well-posedness and differentiability of the discrete counterpart of the control-to-state operator. In Section 5, the error analysis of the adjoint state equation is investigated. Finally, the main results of the paper is presented in Section 6. There, the convergence and error estimates of local minima of discrete optimal control problems are, respectively, presented in Section 6.1 and Section 6.2.

Notation. We denote by $B_X(u,\rho)$ and $\overline{B}_X(u,\rho)$ the open and closed balls in a Banach space X of radius $\rho > 0$ centered at $u \in X$, respectively. For Banach spaces X and Y, the notation $X \hookrightarrow (\Subset) Y$ is understood that X is continuously (compactly) embedded in Y. Let X be a Banach space with its dual X^*, the symbol $\langle \cdot, \cdot \rangle_{X^*,X}$ stands for the dual product of X and X^*. For a given function $g : \overline{\Omega} \to \mathbb{R}$ and a subset $A \subset \mathbb{R}$, $\{g \in A\}$ denotes the set of all points $x \in \overline{\Omega}$ for which $g(x) \in A$. Similarly, for functions g_1, g_2 and subsets $A_1, A_2 \subset \mathbb{R}$, the symbol $\{g_1 \in A_1, g_2 \in A_2\}$ determines the set of all points at which the values of g_1 and g_2 belonging to A_1 and A_2, respectively. For any set $\omega \subset \overline{\Omega}$, we denote by $\mathbb{1}_\omega$ the indicator function of ω, i.e., $\mathbb{1}_\omega(x) = 1$ if $x \in \omega$ and $\mathbb{1}_\omega(x) = 0$ otherwise. For any constants $c > 0$ and $t \in \mathbb{R}$, we use the convention that $\frac{c}{0} = \infty$ and $(t, +\infty] := (t, \infty)$. Finally, we write the symbol C for a generic positive constant, which may be different at different places of occurrence and the notation, e.g. C_ξ for a constant depending only on the parameter ξ.

2 MAIN ASSUMPTIONS AND PRELIMINARY RESULTS

Let a be a finitely PC^k-function with $1 \le k \le \infty$, that is, there exist an integer $K \in \mathbb{N}$, numbers $t_1, t_2, \dots, t_K$ such that $t_1 < t_2 < \dots < t_K$, and C^k-functions a_i, $0 \le i \le K$, satisfying

$$(2.1)\qquad a(t) := \sum_{i=0}^{K} \mathbb{1}_{(t_i,t_{i+1}]}(t) a_i(t) \quad \text{for all} t \in \mathbb{R}, \quad \text{where} a_{i-1}(t_i) = a_i(t_i) \text{for all} 1 \le i \le K.$$

Here we use the the conventions $t_0 := -\infty$ and $t_{K+1} := \infty$ (see, e.g. [16] for the definition). Obviously, the function a is of class C^2 over each interval (t_i, t_{i+1}) for $0 \le i \le K$, but not even of class C^1 in general. By E_a, we denote the exceptional set of all non-differentiability points of a, that is, $E_a := \{t_i \mid 1 \le i \le K\}$. For any $t_i \in E_a$, we denote by $\{a'\}_{t_i+0}^{t_i-0}$ the difference between the one-sided derivatives of a at t_i from left and right, i.e.,

$$\{a'\}_{t_i+0}^{t_i-0} := \lim_{t\to t_i^-} a'(t) - \lim_{t\to t_i^+} a'(t) = a'_{i-1}(t_i) - a'_i(t_i).$$

By setting

$$(2.2)\qquad \sigma_i := |\{a'\}_{t_i+0}^{t_i-0}| = |a'_{i-1}(t_i) - a'_i(t_i)| \quad (1 \le i \le K),$$

we see that these terms determine the differentiability of a and play a crucial part in the second-order optimality conditions for (P); see [16]. Moreover, when a is a finitely PC^1-function defined via (2.1), then it is directionally differentiable and its directional derivative is given by

$$(2.3)\qquad a'(t;s) = \sum_{i=0}^{K} \left\{ \mathbb{1}_{(t_i,t_{i+1})}(t) a'_i(t)s + \mathbb{1}_{\{t_{i+1}\}}(t) \left[\mathbb{1}_{(0,\infty)}(s) a'_{i+1}(t_{i+1})s + \mathbb{1}_{(-\infty,0)}(s) a'_i(t_{i+1})s \right] \right\}$$

with the convention $\mathbb{1}_{\{t_{K+1}\}} = \mathbb{1}_{\{\infty\}} = 0$. The following assumptions shall be assumed in the whole paper except in Section 3, where we will only require the Lipschitz regularity of Ω instead of Assumption (A1) below.

(A1) $\Omega \subset \mathbb{R}^2$ is an open bounded convex polygonal.

(A2) The Lipschitz continuous function $b : \overline{\Omega} \to \mathbb{R}$ satisfies $b(x) \ge \underline{b} > 0$ for all $x \in \overline{\Omega}$.

(A3) $a : \mathbb{R} \to \mathbb{R}$ is a non-negative finitely PC^2 function and determined via (2.1).

(A4) $L : \Omega \times \mathbb{R} \to \mathbb{R}$ is a Carathéodory function that is of class C^2 w.r.t. the second variable with $L(\cdot, 0) \in L^1(\Omega)$. Besides, for any $M > 0$, there exist $C_M > 0$ and $\psi_M \in L^{\bar{p}}(\Omega)$ ($\bar{p} > 2$) such that $|\frac{\partial L}{\partial y}(x,y)| \le \psi_M(x)$ and $|\frac{\partial^2 L}{\partial y^2}(x,y)| \le C_M$ for all $y \in \mathbb{R}$ with $|y| \le M$, and a.e. $x \in \Omega$.

In the remainder of this subsection, we state some known results for the state equation, the adjoint state equation, and the optimality conditions for (P); see, e.g. [16]. Let us first consider the state equation

$$(2.4)\qquad -\operatorname{div}[(b + a(y))\nabla y] = u \,\text{in}\Omega, \quad y = 0 \,\text{on}\partial\Omega.$$

Theorem 2.1 (cf. [16, Thms. 3.1,3.5]). *Let Assumptions (A1) to (A3) hold. Then, the control-to-state operator $S : W^{-1,p}(\Omega) \ni u \mapsto y_u \in W_0^{1,p}(\Omega)$ with y_u being the unique solution to (2.4) is of class C^1. Moreover, for any $u, v \in W^{-1,p}(\Omega)$ with $p > 2$ and $y_u := S(u)$, $z_v := S'(u)v$ is the unique solution to*

$$(2.5)\qquad -\operatorname{div}[(b + a(y_u))\nabla z_v + \mathbb{1}_{\{y_u \notin E_a\}} a'(y_u) z_v \nabla y_u] = v \,\text{in}\Omega, \quad z_v = 0 \,\text{on}\partial\Omega.$$

Moreover, there exists a number $p_* > 2$ such that for any $p \in [2, p_*)$ and for any bounded set $U \subset L^p(\Omega)$, there hold $S(u) \in W_0^{1,p}(\Omega) \cap W^{2,p}(\Omega)$ and $\|S(u)\|_{W^{2,p}(\Omega)} \le C_U$.

Proof. The well-posedness and the continuous differentiability of S follows from [16, Thms. 3.1,3.5]. On the other hand, the H^2- and $W^{1,\infty}(\Omega)$-regularity of solutions to (2.4) was also shown in [16, Thm. 3.1] when the right-hand side u belongs to $L^q(\Omega)$ with $q > 2$. Moreover, if U is a bounded subset of $L^q(\Omega)$ with $q > 2$, then there holds $\|S(u)\|_{H^2(\Omega)} + \|S(u)\|_{W^{1,\infty}(\Omega)} \leq C_U$ for all $u \in U$. To show the higher $W^{2,p}$-regularity as well as the corresponding a priori estimate, we observe that (2.4) can be rewritten as

$$-\Delta y = \frac{1}{b + a(y)}[u + \nabla b \cdot \nabla y + \mathbb{1}_{\{y \notin E_a\}} a'(y)|\nabla y|^2] \text{ in}\Omega, \quad y = 0 \text{ on}\partial\Omega \tag{2.6}$$

(see, e.g. equations (A.1) and (A.3) in [16, Lem. A.1]). Since Ω is assumed to be a convex polygon in $\mathbb{R}^2$, [24, Thm. 4.4.3.7] shows that there exists a constant $p_* := 2/(2 - \min\{\pi\omega_{\max}^{-1}, 2\}) > 2$ depending on the maximal interior angle $\omega_{\max} < \pi$ of the domain Ω such that any solution y to (2.6) belongs to $W^{2,p}(\Omega)$ provided that $u \in L^p(\Omega)$ for all $p \in (2, p_*)$. Of course, we have $y \in H^2(\Omega)$ when $u \in L^2(\Omega)$ due to the convexity of Ω. Finally, the $W^{2,p}$-estimate of solutions y is derived by applying [24, Thm. 4.3.2.4] to (2.6) and using the a priori $W^{1,\infty}(\Omega)$-estimates of y, Assumption (A2) and Assumption (A3). □

We now consider the adjoint state equation

$$-\operatorname{div}[(b + a(y_u))\nabla\varphi] + \mathbb{1}_{\{y_u \notin E_a\}} a'(y_u)\nabla y_u \cdot \nabla\varphi = v \text{ in}\Omega, \quad \varphi = 0 \text{ on}\partial\Omega \tag{2.7}$$

for $u \in W^{-1,p}(\Omega)$, $p > 2$, $v \in H^{-1}(\Omega)$, and $y_u := S(u)$.

Theorem 2.2 ([16, Lem. 4.1]). *Let Assumptions (A1) to (A3) be satisfied and let $p, q > 2$ be arbitrary. Then, for any $u \in W^{-1,p}(\Omega)$, $v \in H^{-1}(\Omega)$, a unique $\varphi \in H_0^1(\Omega)$ exists and uniquely solves (2.7). Furthermore, if U is a bounded subset in $L^p(\Omega)$, then for any $u \in U$ and any $v \in L^q(\Omega)$, the solution φ of (2.7) belongs to $H^2(\Omega) \cap W^{1,\infty}(\Omega)$ and there holds $\|\varphi\|_{H^2(\Omega)} + \|\varphi\|_{W^{1,\infty}(\Omega)} \leq C_U\|v\|_{L^q(\Omega)}$.*

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

$$\min_{u \in \mathcal{U}_{ad}} j(u) = \int_\Omega L(x, S(u)(x))\,\mathrm{d}x + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} \tag{P}$$

with

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

Under Assumptions (A1) to (A4), the cost functional $j : L^2(\Omega) \to \mathbb{R}$ is of class C^1. Moreover, there holds

$$j'(u)v = \int_\Omega (\varphi_u + \nu u)v\,\mathrm{d}x \quad \text{for} u, v \in L^2(\Omega) \tag{2.8}$$

with $\varphi_u \in H_0^1(\Omega)$ solving (2.7) corresponding to the right-hand side term v substituted by $\frac{\partial L}{\partial y}(\cdot, S(u))$; see [16, Thm. 4.2]. We have the following first-order necessary optimality conditions.

Theorem 2.3 ([16, Thm. 4.3]). *Assume that Assumptions (A1) to (A4) are satisfied. Then there exists at least one minimizer $\bar{u}$ of (P). Moreover, there exists an adjoint state $\bar{\varphi} \in H_0^1(\Omega)$ such that for $\bar{y} := S(\bar{u})$,*

$$-\operatorname{div}[(b + a(\bar{y}))\nabla\bar{\varphi}] + \mathbb{1}_{\{\bar{y} \notin E_a\}} a'(\bar{y})\nabla\bar{y} \cdot \nabla\bar{\varphi} = \frac{\partial L}{\partial y}(x, \bar{y}) \text{ in}\Omega, \quad \bar{\varphi} = 0 \text{ on}\partial\Omega, \tag{2.9a}$$

$$\int_\Omega (\bar{\varphi} + \nu\bar{u})(u - \bar{u})\,\mathrm{d}x \geq 0 \quad \text{for all} u \in \mathcal{U}_{ad}. \tag{2.9b}$$

Assume that $\bar{\varphi} \in H_0^1(\Omega)$ satisfies (2.9). The *critical cone* of the problem (P) at $\bar{u}$ is defined as

$$C(\mathcal{U}_{ad}; \bar{u}) := \{v \in L^2(\Omega) \mid v \geq 0 \text{ if} \bar{u} = \alpha, v \leq 0 \text{ if} \bar{u} = \beta, v = 0 \text{ if} \bar{\varphi} + \nu\bar{u} \neq 0 \text{ a.e. in}\Omega\}. \tag{2.10}$$

In the rest of this section, we shall provide second-order sufficient optimality conditions for (P). To this end, the curvature functional of j is first introduced and can be separated into three contributions.

For any $(u, y, \varphi) \in L^2(\Omega) \times H^1(\Omega) \times W^{1,\infty}(\Omega)$, the smooth part and the first-order non-smooth part of the curvature in direction $(v_1, v_2) \in L^2(\Omega)^2$ are given by

$$Q_s(u, y, \varphi; v_1, v_2) := \frac{1}{2}\int_\Omega \frac{\partial^2 L}{\partial y^2}(\cdot, y) z_{v_1} z_{v_2}\,dx + \frac{\nu}{2}\int_\Omega v_1 v_2\,dx - \frac{1}{2}\int_\Omega \mathbb{1}_{\{y \notin E_a\}} a''(y) z_{v_1} z_{v_2} \nabla y \cdot \nabla\varphi\,dx,$$
$$Q_1(u, y, \varphi; v_1, v_2) := -\frac{1}{2}\int_\Omega [a'(y; z_{v_1})\nabla z_{v_2} + a'(y; z_{v_2})\nabla z_{v_1}] \cdot \nabla\varphi\,dx,$$

for $z_{v_i} := S'(u)v_i$, $i = 1, 2$, respectively. The critical part for our analysis is of course the second-order non-smooth part, which requires some additional notation. For ease of exposition, we use the following notation in the remainder. For any $y, \hat{y} \in C(\overline{\Omega}) \cap W^{1,1}(\Omega)$ and any $\tau_1, \tau_2 \in \mathbb{R}$ we define the set $\Omega_{\hat{y},i,j}^{[\tau_1,\tau_2]} := \{\hat{y} \in [t_i + \tau_1, t_j + \tau_2]\}$; similar sets such as $\Omega_{\hat{y},i,j}^{[\tau_1,\tau_2)}$ are defined in the same way and set

$$\begin{cases} \Omega_{y,\hat{y}}^{i,1} := \Omega_{\hat{y},i,i+1}^{[\delta,-\delta]} \cup (\Omega_{\hat{y},i,i}^{(0,\delta)} \cap \Omega_{y,i,i}^{(0,2\delta)}) \cup (\Omega_{\hat{y},i+1,i+1}^{(-\delta,0)} \cap \Omega_{y,i+1,i+1}^{(-2\delta,0)}), \\ \Omega_{y,\hat{y}}^{i,2} := \Omega_{\hat{y},i,i}^{(0,\delta)} \cap \Omega_{y,i,i}^{(-\delta,0]}, \qquad \Omega_{y,\hat{y}}^{i,3} := \Omega_{\hat{y},i+1,i+1}^{(-\delta,0)} \cap \Omega_{y,i+1,i+1}^{[0,\delta)}, \end{cases} \tag{2.11}$$

with some fixed number δ such that

$$0 < \delta \leq \frac{t_{i+1} - t_i}{2} \quad \text{for all} 1 \leq i \leq K - 1. \tag{2.12}$$

For any $0 \leq i \leq K$, $s \in \mathbb{R}$, $u, v \in L^2(\Omega)$, $y \in C(\overline{\Omega}) \cap H^1(\Omega)$, and $\varphi \in W^{1,\infty}(\Omega)$, we set

$$\zeta_i(u, y; s, v) := \{a'\}_{t_i+0}^{t_i-0} \mathbb{1}_{\Omega_{S(u+sv),y}^{i,2}}(t_i - S(u + sv)) - \{a'\}_{t_{i+1}+0}^{t_{i+1}-0} \mathbb{1}_{\Omega_{S(u+sv),y}^{i,3}}(t_{i+1} - S(u + sv)) \tag{2.13}$$

with $\{a'\}_{t_i+0}^{t_i-0}$ defined as in (2.2) and the conventions

$$a_{-1} \equiv a_{K+1} \equiv 0, \quad a_0'(t_0) = a_0'(-\infty) = 0, \quad a_K'(t_{K+1}) = a_K'(\infty) = 0, \quad \Omega_{\hat{y},y}^{0,2} = \Omega_{\hat{y},y}^{K,3} = \emptyset. \tag{2.14}$$

We then define for any $\{s_n\} \in c_0^+ := \{\{s_n\} \subset (0, \infty) \mid s_n \to 0\}$ and $v \in L^2(\Omega)$

$$\begin{aligned} \tilde{Q}(u, y, \varphi; \{s_n\}, v) &:= \liminf_{n\to\infty} \frac{1}{s_n^2}\int_\Omega \sum_{i=0}^{K} \zeta_i(u, y; s_n, v)\nabla y \cdot \nabla\varphi\,dx \\ &= \liminf_{n\to\infty} \frac{1}{s_n^2}\sum_{i=1}^{K}\{a'\}_{t_i+0}^{t_i-0}\int_\Omega (t_i - S(u + s_n v))\left[\mathbb{1}_{\Omega_{S(u+s_nv),y}^{i,2}} - \mathbb{1}_{\Omega_{S(u+s_nv),y}^{i-1,3}}\right]\nabla y \cdot \nabla\varphi\,dx. \end{aligned} \tag{2.15}$$

The second-order non-smooth part of the curvature in direction $v \in L^2(\Omega)$ is then given by

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

and finally the total curvature in direction v is

$$Q(u, y, \varphi; v) := Q_s(u, y, \varphi; v, v) + Q_1(u, y, \varphi; v, v) + Q_2(u, y, \varphi; v). \tag{2.16}$$

Moreover, according to [16, Prop. 5.6 & Lem. 5.7], Q_2 is weakly lower semicontinuous in the last variable and satisfies

$$|Q_2(u, S(u), \varphi; v)| \leq \Sigma(S(u))\|\nabla\varphi\|_{L^\infty(\Omega)}\|S'(u)v\|_{L^\infty(\Omega)}^2 \quad \text{for all} u, v \in L^2(\Omega) \text{and} \varphi \in W^{1,\infty}(\Omega),$$

with the *jump functional*

$$\Sigma(y) := \limsup_{r\to 0^+} \frac{1}{r}\sum_{m=1}^{2}\sum_{i=1}^{K}\sigma_i\int_\Omega [\mathbb{1}_{\{0<|y-t_i|\leq r\}}|\partial_{x_m} y|]\,dx, \quad y \in W^{1,1}(\Omega) \cap C(\overline{\Omega}) \tag{2.17}$$

for σ_i defined in (2.2). Also, from [16, Cor. 5.5], it holds for any $u \in L^2(\Omega)$, $\{s_n\} \in c_0^+$ and $v_n \rightharpoonup v$ in $L^2(\Omega)$ that

$$\liminf_{n\to\infty} \frac{1}{s_n^2} \int_\Omega \sum_{i=0}^{K} \zeta_i(u, S(u); s_n, v_n) \nabla S(u) \cdot \nabla\varphi \, dx = \tilde{Q}(u, S(u), \varphi; \{s_n\}, v) \geq Q_2(u, S(u), \varphi; v), \tag{2.18}$$

provided that $\Sigma(S(u)) < \infty$.

Theorem 2.4 (second-order sufficient optimality conditions, [16, Thm. 5.10]). *Let Assumptions (A1) to (A4) be valid. Assume that $\bar{u}$ is a feasible point of (P) such that $\Sigma(\bar{y}) < \infty$ for $\bar{y} := S(\bar{u})$. Assume further that there is a $\bar{\varphi} \in W_0^{1,\bar{p}}(\Omega) \cap W^{1,\infty}(\Omega)$, with $\bar{p}$ defined in Assumption (A4), that together with $\bar{u}$, $\bar{y}$ satisfies (2.9) and*

$$Q(\bar{u}, \bar{y}, \bar{\varphi}; v) > 0 \quad \text{for all } v \in C(\mathcal{U}_{ad}; \bar{u}) \setminus \{0\} \tag{2.19}$$

with Q defined in (2.16). Then there exist constants $c_0, \rho_0 > 0$ satisfying

$$j(\bar{u}) + c_0\|u - \bar{u}\|^2_{L^2(\Omega)} \leq j(u) \qquad \text{for all } u \in \mathcal{U}_{ad} \cap \overline{B}_{L^2(\Omega)}(\bar{u}, \rho_0).$$

It is noted that the term $2Q(\bar{u}, \bar{y}, \bar{\varphi}; v)$ can be seen as a second-order generalized derivative of j at $\bar{u}$ in the direction v; see, e.g. the proof of Theorem 5.9 in [16] and [30, Rem. 5.1]).

3 AN EXPLICIT FORMULA FOR THE CURVATURE FUNCTIONAL

In this section, Ω is a convex and bounded domain in $\mathbb{R}^2$ only and the symbol $\mathcal{H}^1$ stands for the one-dimensional Hausdorff measure on $\mathbb{R}^2$ that is scaled as in [20, Def. 2.1]. We need the following notion.

Definition 3.1. A function $y : \overline{\Omega} \to \mathbb{R}$ is called *uniformly locally convex-concave* on a set $V \subset \overline{\Omega}$ if an $\varepsilon > 0$ exists such that for any $x \in V$, y is either convex or concave on $B_{\mathbb{R}^2}(x, \varepsilon) \cap \overline{\Omega}$.

Furthermore, we recall from [1] that a connected component of a set V is any element of the class of connected subsets of V that is maximal with respect to inclusion and that a closed Lipschitz curve is a curve that admits a Lipschitz parametrization $\gamma : [a, b] \to \mathbb{R}^2$ which is injective on $[a, b)$ and satisfies $\gamma(a) = \gamma(b)$.

Proposition 3.2. *Let $t \in \mathbb{R}$ be arbitrary and let $y \in C^1(\overline{\Omega})$ be uniformly locally convex-concave on the level set $\{y = t\}$. Assume that C is a connected component of $\{y = t\}$. If ∇y vanishes at some point $x_0 \in C$, then $\nabla y(x) = 0$ for all $x \in C$.*

Proof. There is an $\varepsilon > 0$ such that, for any $x \in \{y = t\}$, the restriction $y\,|_{B_{\mathbb{R}^2}(x,\varepsilon)\cap\overline{\Omega}}$ is either convex or concave. Since $\nabla y(x_0) = 0$, then x_0 is a local extremal point of $y(x)$ and so is every point in $B_{\mathbb{R}^2}(x, \varepsilon) \cap C$. From this and the connection property of C, we have $\nabla y = 0$ on C. □

The following result is a direct consequence of Proposition 3.2.

Corollary 3.3. *Let $t \in \mathbb{R}$ be arbitrary and let $y \in C^1(\overline{\Omega})$ be uniformly locally convex-concave on the level set $\{y = t\}$. Assume that C is a connected component of $\{y = t\}$. If $\nabla y(x_0) \neq 0$ for some point $x_0 \in C$, then $\nabla y(x) \neq 0$ for all $x \in C$ and C is a closed Lipschitz simple curve in $\mathbb{R}^2$.*

Proof. Obviously, the gradient of y does not vanish at any point of C according to Proposition 3.2. From this and the Implicit Function Theorem, we deduce that C is a closed simple curve in $\mathbb{R}^2$, which is of class C^1 if $C \cap \partial\Omega = \emptyset$ and of class $C^{0,1}$ if otherwise. □

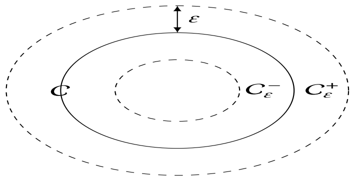


Figure 1: sets C_ε^+ and C_ε^-

From now on, for any $\varepsilon > 0$ and any set $V \subset \overline{\Omega}$, we denote by V^ε the open ε-neighborhood in Ω of V, that is, $V^\varepsilon := \{x \in \overline{\Omega} \mid \operatorname{dist}(x, V) < \varepsilon\}$, where $\operatorname{dist}(x, V)$ is the distance from x to V. If C is a closed simple curve in $\overline{\Omega}$, we define the open sets C_ε^+ and C_ε^- (illustrated in Figure 1) as follows

$$C_\varepsilon^- := \{x \in \overline{\Omega} \mid 0 < \operatorname{dist}(x, C) < \varepsilon \,\text{and}\, x \text{ is surrounded by} C\} \quad \text{and} \quad C_\varepsilon^+ := C^\varepsilon \backslash (C_\varepsilon^- \cup C).$$

The following proposition will play an important role in establishing an explicit formula of the curvature functional (2.16). Its proof rests on the following two lemmas.

Proposition 3.4. *Let $\{s_n\} \in c_0^+$, $\bar\varphi \in W^{1,\infty}(\Omega) \cap W^{2,1}(\Omega)$ and $\bar y, y_n \in C^1(\overline{\Omega})$ such that $y_n = \bar y = 0$ on $\partial\Omega$, $y_n \to \bar y$ in $C^1(\overline{\Omega})$ and $(y_n - \bar y)/s_n \to w$ in $W_0^{1,p}(\Omega)$ for some $p > 2$ and $w \in W_0^{1,p}(\Omega)$. Let C be a closed connected component of $\{\bar y = t_i\}$. Assume that either*

$$\min\{|\nabla \bar y(x)| : x \in C\} > 0 \text{ or } \begin{cases} \bar y \text{ is uniformly locally convex-concave on } C, \\ \nabla \bar y = 0 \text{ on} C, \\ \mathcal{H}^{N-1}(\{\bar y = t\} \cap C^{\tilde\varepsilon}) \le C_0 \text{ a.e.} t \in (t_i - r_0, t_i + r_0), N = 2, \end{cases} \tag{3.1}$$

for some constants $\tilde\varepsilon, r_0, C_0 > 0$. Then there exists an $\varepsilon_0 = \varepsilon_0(C) \in (0, \tilde\varepsilon)$ such that for any $\varepsilon \in (0, \varepsilon_0)$,

$$\frac{A_n(C,\varepsilon)}{s_n^2} \to \frac{1}{2} \int_C \mathbb{1}_{\{|\nabla \bar y| > 0\}} w^2 \frac{\nabla \bar y \cdot \nabla \bar\varphi}{|\nabla \bar y|} \,\mathrm{d}\mathcal{H}^1(x) \quad \text{as} \quad n \to \infty, \tag{3.2}$$

where $A_n(C,\varepsilon) := \int_{C^\varepsilon} (t_i - y_n)[\mathbb{1}_{\Omega^{i,2}_{y_n,\bar y}} - \mathbb{1}_{\Omega^{i-1,3}_{y_n,\bar y}}] \nabla \bar y \cdot \nabla \bar\varphi \,\mathrm{d}x$ with $\Omega^{i,2}_{y_n,\bar y}$ and $\Omega^{i-1,3}_{y_n,\bar y}$ defined in (2.11).

Remark 3.5. Note that (3.1) does not require that the level sets $\{\bar y = t_i\}$ have measure zero. When $N = 1$, the second condition in (3.1) means that for a.e. t in a neighborhood of t_i, the level sets $\{\bar y = t\}$ consist of finitely many points; in other words, the function $\bar y$ oscillates around the values t_i only finitely many times. In [16, Exam. 5.3], this condition was shown to be equivalent to the finiteness of the jump functional $\Sigma(\bar y)$ introduced in (2.17) for the case $N = 1$.

Lemma 3.6. *If $|\nabla \bar y(x)| > 0$ for all $x \in C$, then there exists an $\varepsilon_0 = \varepsilon_0(C) > 0$ such that*

$$\frac{A_n(C,\varepsilon)}{s_n^2} \to \frac{1}{2} \int_C w^2 \frac{\nabla \bar y \cdot \nabla \bar\varphi}{|\nabla \bar y|} \,\mathrm{d}\mathcal{H}^1(x) \quad \text{as} \quad n \to \infty \quad \text{for all} \varepsilon \in (0, \varepsilon_0). \tag{3.3}$$

Proof. The Implicit Function Theorem implies that C is a closed Lipschitz simple curve in $\overline{\Omega}$. By virtue of the limit $y_n \to \bar y$ in $C^1(\overline{\Omega})$, there is $\varepsilon_1 > 0$ such that, for n large enough,

$$|\nabla y_n(x)|, |\nabla \bar y(x)| > 0 \quad \text{for all} \quad x \in C^{\varepsilon_1}. \tag{3.4}$$

Since C is a closed component in $\{\bar y = t_i\}$, there holds $C^{\varepsilon_2} \cap (\{\bar y = t_i\} \backslash C)^{\varepsilon_2} = \emptyset$ for some constant $\varepsilon_2 > 0$. Moreover, the sign of $(\bar y - t_i)$ in $C_{\varepsilon_2}^-$ is opposite to the one in $C_{\varepsilon_2}^+$. Without loss of generality, we can thus assume that

$$\bar y < t_i \quad \text{on} C_{\varepsilon_2}^- \quad \text{and} \quad \bar y > t_i \quad \text{on} C_{\varepsilon_2}^+. \tag{3.5}$$

Set $\varepsilon_0 := \min\{\varepsilon_1, \varepsilon_2\}$, $\tau_n := \|y_n - \bar y\|_{C(\overline{\Omega})}$, fix any $\varepsilon \in (0, \varepsilon_0)$, and define the sets (depending on ε) $\Omega_n^2 := \Omega^{i,2}_{y_n,\bar y} \cap C^\varepsilon$ and $\Omega_n^3 := \Omega^{i-1,3}_{y_n,\bar y} \cap C^\varepsilon$. Obviously, for n large enough, we have

$$\Omega_n^2 = \{\bar y \in (t_i, t_i+\delta), y_n \in (t_i - \delta, t_i]\} \cap C^\varepsilon = \{0 < \bar y - t_i \le \bar y - y_n\} \cap C^\varepsilon \subset C_\varepsilon^+, \tag{3.6a}$$

$$\Omega_n^3 = \{\bar y \in (t_i - \delta, t_i), y_n \in [t_i, t_i+\delta)\} \cap C^\varepsilon = \{0 > \bar y - t_i \ge \bar y - y_n\} \cap C^\varepsilon \subset C_\varepsilon^-. \tag{3.6b}$$

Moreover, by (3.5) and the limit $y_n \to \bar y$ in $C(\overline{\Omega})$, one has $\{y_n = t_i\} \cap C^\varepsilon \neq \emptyset$ for all n large enough. From this and (3.4), we deduce from the Implicit Function Theorem that, for n sufficiently large, the set $C_n := \{y_n = t_i\} \cap C^\varepsilon$ is a closed Lipschitz simple curve in $\mathbb{R}^2$. We now consider two cases.
Case 1: $C = \partial\Omega$. In this case, we have $t_i = 0$ and thus $C_n = \partial\Omega$ for sufficiently large n. Moreover, one has $C_\varepsilon^+ = \emptyset$, $C^\varepsilon = C \cup C_\varepsilon^-$ and then $\Omega_n^2 = \emptyset$ for n large enough. On the other hand, $C_n = \{y_n = t_i\} \cap C^\varepsilon$ is a closed simple curve in $\mathbb{R}^2$ for n large enough and $y_n \to \bar y$ in $C(\overline{\Omega})$, we deduce from (3.5) that $y_n \le t_i$ on C^ε and therefore $\Omega_n^3 = \emptyset$ for sufficiently large n. From this and the definition of $A_n(C, \varepsilon)$, we have (3.3) because of the vanishing on $\partial\Omega$ of w.
Case 2: $C \neq \partial\Omega$. In this case, both C_ε^+ and C_ε^- are non-empty. To estimate $A_n(C,\varepsilon)$, we split it into two terms as follows:

$$A_n(C,\varepsilon) := \int_{C^\varepsilon} (t_i - y_n)[\mathbb{1}_{\Omega_n^2} - \mathbb{1}_{\Omega_n^3}]\nabla(\bar y - y_n)\cdot\nabla\bar\varphi \,\mathrm{d}x + \int_{C^\varepsilon} (t_i - y_n)[\mathbb{1}_{\Omega_n^2} - \mathbb{1}_{\Omega_n^3}]\nabla y_n \cdot \nabla\bar\varphi\,\mathrm{d}x =: B_n + C_n.$$

From the facts $|t_i - y_n| \le |\bar y - y_n|$ on $\Omega_n^2 \cup \Omega_n^3$ and $\Omega_n^2 \cup \Omega_n^3 \subset \{0 < |\bar y - t_i| \le \tau_n\}$, we deduce from Hölder's inequality for $p' := \frac{p}{p-1}$ that

$$|B_n| \le \|y_n - \bar y\|_{C(\overline{\Omega})} \|\nabla(y_n - \bar y)\|_{L^p(\Omega)} \|\nabla\bar\varphi \mathbb{1}_{\{0<|\bar y - t_i| \le \tau_n\}}\|_{L^{p'}(\Omega)} = o(s_n^2) \tag{3.7}$$

and thus

$$A_n(C,\varepsilon) = o(s_n^2) + C_n. \tag{3.8}$$

C_n can then be rewritten as

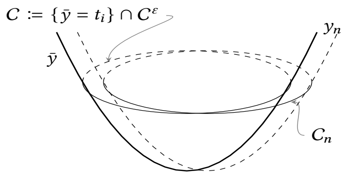
$$C_n = -\frac{1}{2}\int_{C^\varepsilon} [\mathbb{1}_{\Omega_n^2} - \mathbb{1}_{\Omega_n^3}]\nabla(y_n - t_i)^2 \cdot \nabla\bar\varphi\,\mathrm{d}x =: -\frac{1}{2}(C_n^2 - C_n^3) \tag{3.9}$$

with $C_n^j := \int_{\Omega_n^j} \nabla(y_n - t_i)^2 \cdot \nabla\bar\varphi dx$ for $j = 2, 3$. We now split the sequence $\{n\}$ into subsequences, still denoted by the same symbol, that satisfy one of the following conditions:

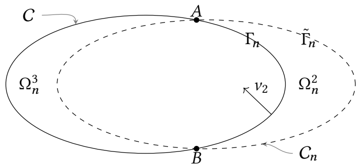
(a) $\Omega_n^2 \neq \emptyset$ and $\Omega_n^3 \neq \emptyset$ for all $n \ge 1$;
(b) $\Omega_n^2 \neq \emptyset$ and $\Omega_n^3 = \emptyset$ for all $n \ge 1$;
(c) $\Omega_n^2 = \emptyset$ and $\Omega_n^3 \neq \emptyset$ for all $n \ge 1$;
(d) $\Omega_n^2 = \Omega_n^3 = \emptyset$ for all $n \ge 1$.

We first consider (a); see Figure 2. Set $\Gamma_n := \partial\Omega_n^2 \cap \{\bar y = t_i\}$ and $\tilde\Gamma_n := \partial\Omega_n^2 \cap \{y_n = t_i\}$. Obviously, the outward unit normal vector ν_2 of $\partial\Omega_n^2$ is defined by $\nu_2 := -\frac{\nabla\bar y}{|\nabla\bar y|}$ on Γ_n and $\nu_2 := \frac{\nabla y_n}{|\nabla y_n|}$ on $\tilde\Gamma_n$. By the Lipschiptz continuity of the closed curves C and C_n, we can apply integration by parts to derive

$$\begin{aligned}
C_n^2 &= \int_{\Omega_n^2} \nabla(y_n - t_i)^2\cdot\nabla\bar\varphi\,\mathrm{d}x = -\int_{\Omega_n^2}(y_n - t_i)^2\Delta\bar\varphi dx + \int_{\Gamma_n\cup\tilde\Gamma_n}(y_n - t_i)^2\nabla\bar\varphi\cdot\nu_2\,\mathrm{d}\mathcal{H}^1(x)\\
&= o(s_n^2) - \int_{\Gamma_n}(y_n - t_i)^2\nabla\bar\varphi\cdot\frac{\nabla\bar y}{|\nabla\bar y|}\,\mathrm{d}\mathcal{H}^1(x) + \int_{\tilde\Gamma_n}(y_n - t_i)^2\nabla\bar\varphi\cdot\frac{\nabla y_n}{|\nabla y_n|}\,\mathrm{d}\mathcal{H}^1(x)\\
&= o(s_n^2) - \int_{\Gamma_n}(y_n - \bar y)^2\nabla\bar\varphi\cdot\frac{\nabla\bar y}{|\nabla\bar y|}\,\mathrm{d}\mathcal{H}^1(x).
\end{aligned}$$

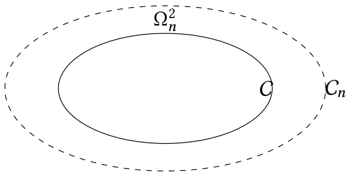


(a) closed simple curves C and C_n

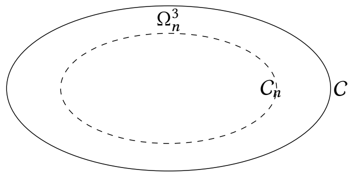


(b) subsets Ω_n^2 and Ω_n^3, boundary $\Gamma_n \cup \tilde\Gamma_n$ of Ω_n^2, and the outward normal vector $\nu_2 = -\frac{\nabla \bar y}{|\nabla \bar y|}$ on Γ_n

Figure 2: case where both Ω_n^2 and Ω_n^3 are non-empty



(a) $\Omega_n^2 \neq \emptyset$ and $\Omega_n^3 = \emptyset$



(b) $\Omega_n^2 = \emptyset$ and $\Omega_n^3 \neq \emptyset$

Figure 3: case where Ω_n^2 or Ω_n^3 is empty

Here we have used the fact

$$\left|\int_{\Omega_n^2} (y_n - t_i)^2 \Delta\bar\varphi \,\mathrm{d}x\right| \leq \int_{\{0<|\bar y - t_i| \leq \tau_n\}} |y_n - \bar y|^2 |\Delta \bar\varphi| \,\mathrm{d}x = o(s_n^2).$$

Analogously, by noting that $\frac{\nabla \bar y}{|\nabla \bar y|}$ is the outward unit normal vector to $\partial\Omega_n^3 \cap \{\bar y = t_i\}$, one has

$$C_n^3 = o(s_n^2) + \int_{C\setminus\Gamma_n} (y_n - \bar y)^2 \nabla\bar\varphi \cdot \frac{\nabla \bar y}{|\nabla \bar y|} \,\mathrm{d}\mathcal{H}^1(x).$$

We then derive from (3.9) that

$$C_n = o(s_n^2) + \frac{1}{2}\int_C (y_n - \bar y)^2 \nabla\bar\varphi \cdot \frac{\nabla \bar y}{|\nabla \bar y|} \,\mathrm{d}\mathcal{H}^1(x). \tag{3.10}$$

The combination of (3.10) with (3.8) and the limits $(y_n - \bar y)/s_n \to w$ in $W_0^{1,p}(\Omega)$ yields (3.3).

For (b) or (c) (see Figure 3), the argument is similar as above. Finally, for (d), we have $C = C_n$ and $A_n(C, \varepsilon) = 0$ for all $n \geq 1$. Moreover, there holds $w = 0$ on C as a result of the fact that $C = C_n$ and the limit $\frac{y_n - \bar y}{s_n} \to w$ in $W_0^{1,p}(\Omega) \hookrightarrow C(\overline{\Omega})$ with $p > 2$. Then (3.3) follows. □

Lemma 3.7. *Under the second condition in (3.1), there is an $\varepsilon_0 = \varepsilon_0(C) > 0$ such that $\frac{A_n(C,\varepsilon)}{s_n^2} \to 0$ as $n \to \infty$ for all $\varepsilon \in (0, \varepsilon_0)$.*

Proof. Without loss of generality, assume that there thus exists an $\varepsilon_3 > 0$ such that $\bar y$ is convex on $B(x, \varepsilon_3) \cap \overline{\Omega}$ for all $x \in C$; see Figure 4. We first set $\varepsilon_0 := \min\{\tilde\varepsilon, \varepsilon_2, \varepsilon_3\}$ with constants $\tilde\varepsilon$ given in (3.1) and ε_2 defined as in the proof of Lemma 3.6. Let us take $\varepsilon \in (0, \varepsilon_0)$ arbitrarily but fixed and reuse all symbols defined in the proof of Lemma 3.6. Moreover, the relations (3.6)–(3.9) are still valid. The convexity of $\bar y$ and the fact that $\nabla \bar y = 0$ on C imply that $\bar y > t_i$ on $C^\varepsilon \setminus C$. This and (3.6) guarantee that

$$\Omega_n^2 = \{y_n \leq t_i\} \cap C^\varepsilon \quad \text{and} \quad \Omega_n^3 = \emptyset \quad \text{for all } n \text{ large enough.} \tag{3.11}$$

As in the proof of Lemma 3.6, we now split the sequence $\{n\}$ into subsequences, also denoted by $\{n\}$, that satisfy one of the following conditions:

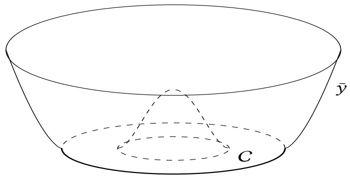


Figure 4: a closed component C of the level set $\{\bar y = t_i\}$ with positive measure $\text{meas}_{\mathbb{R}^2}(C)$

(a) $\Omega_n^2 \neq \emptyset$ for all $n \geq 1$; (b) $\Omega_n^2 = \emptyset$ for all $n \geq 1$.

For (b), the desired conclusion of the lemma is shown using the argument similar to that in the end part of the proof of Lemma 3.6. It remains to consider (a). To this end, by the Sard Theorem for Lipschitz maps [1, Thm. 2.5 (iv)], for a.e. $t \in \mathbb{R}$, the level set $\{y_n = t\}$ is either a point or a closed Lipschitz simple curve in $\mathbb{R}^2$. Therefore, for any $n \geq 1$, there exists a $t_n^i \in \mathbb{R}$ such that

$$\begin{cases} 0 \leq t_i - t_n^i = o(s_n) \quad \text{and} \quad \text{any connected component of} \{y_n = t_n^i\} \text{ is either a point} \\ \text{or a closed Lipschitz simple curve.} \end{cases} \tag{3.12}$$

Combining this with the definition of C_n^2 in (3.9) and the first identity in (3.11) yields

$$\begin{aligned} C_n^2 = -2\int_{\Omega_n^2} (t_i - y_n)\nabla y_n \cdot \nabla\bar\varphi \, \mathrm{d}x &= -2\int_{\{t_n^i \leq y_n \leq t_i\}\cap C^\varepsilon} (t_i - y_n)\nabla y_n \cdot \nabla\bar\varphi \, \mathrm{d}x \\ &\quad - 2\int_{\{y_n < t_n^i\}\cap C^\varepsilon} (t_i - y_n)\nabla y_n \cdot \nabla\bar\varphi \, \mathrm{d}x =: D_n^1 + D_n^2. \end{aligned}$$

Obviously, one has from the choice of t_n^i in (3.12) and the fact $\|y_n - \bar y\|_{W_0^{1,p}(\Omega)} = o(s_n)$ that

$$\begin{aligned} 2^{-1}|D_n^1| &\leq \int_{\mathbb{R}^2} \mathbb{1}_{\{t_n^i \leq y_n \leq t_i\}\cap C^\varepsilon} |t_i - y_n||\nabla y_n||\nabla\bar\varphi| \, \mathrm{d}x \\ &\leq \int_{\mathbb{R}^2} \mathbb{1}_{\{t_n^i \leq y_n \leq t_i\}\cap C^\varepsilon} |t_i - y_n||\nabla(y_n - \bar y)||\nabla\bar\varphi| \, \mathrm{d}x + \int_{\mathbb{R}^2} \mathbb{1}_{\{t_n^i \leq y_n \leq t_i\}\cap C^\varepsilon} |t_i - y_n||\nabla\bar y||\nabla\bar\varphi| \, \mathrm{d}x \\ &= o(s_n^2) + \int_{\mathbb{R}^2} \mathbb{1}_{\{t_n^i \leq y_n \leq t_i\}\cap C^\varepsilon} |t_i - y_n||\nabla\bar y||\nabla\bar\varphi| \, \mathrm{d}x. \end{aligned}$$

From the inclusion $\{t_n^i \leq y_n \leq t_i\} \subset \{t_n^i - \tau_n \leq \bar y \leq t_i + \tau_n\}$ and the coarea formula for Lipschitz mappings; see, e.g. [20, Thm. 2, p. 117] and [1, Sec. 2.7], we have

$$\begin{aligned} 2^{-1}|D_n^1| &\leq o(s_n^2) + \int_{\mathbb{R}} \left[\int_{\{\bar y = t\}} \mathbb{1}_{\{t_n^i \leq y_n \leq t_i\}\cap C^\varepsilon} |t_i - y_n||\nabla\bar\varphi| \, \mathrm{d}\mathcal{H}^1(x) \right] dt \\ &= o(s_n^2) + \int_{t_n^i - \tau_n}^{t_i + \tau_n} \left[\int_{\{\bar y = t\}} \mathbb{1}_{\{t_n^i \leq y_n \leq t_i\}\cap C^\varepsilon} |t_i - y_n||\nabla\bar\varphi| \, \mathrm{d}\mathcal{H}^1(x) \right] dt \\ &\leq o(s_n^2) + C_0(t_i - t_n^i + 2\tau_n)(t_i - t_n^i)\|\nabla\bar\varphi\|_{L^\infty(\Omega)} = o(s_n^2), \end{aligned}$$

due to the choice of t_n^i in (3.12) and the second condition in (3.1). For D_n^2, we see from (3.12) that $\{y_n = t_n^i\} \cap C^\varepsilon \neq \emptyset$ for n large enough. If $\{y_n < t_n^i\} \cap C^\varepsilon = \emptyset$, then $D_n^2 = 0$. Otherwise, let Γ_n^i be the boundary of $\{y_n < t_n^i\} \cap C^\varepsilon$. There are two possibilities in principle: either an infinite subsequence $\{k\}$ of $\{n\}$ exists and satisfies $\Gamma_k^i \cap \partial C^\varepsilon \neq \emptyset$, or there is no such an subsequence. Let us see that the first possibility is not actually a correct assumption. Indeed, if $\Gamma_k^i \cap \partial C^\varepsilon \neq \emptyset$, then $\|y_k - \bar y\|_{C(\overline\Omega)} \geq \bar y(x) - t_i > 0$ for all $x \in \partial C^\varepsilon$. This contradicts the limit $\|y_k - \bar y\|_{C(\overline\Omega)} \to 0$ as $k \to \infty$. Therefore, the second possibility always holds. It then must be true that $\Gamma_n^i \cap \partial C^\varepsilon = \emptyset$ and so $\Gamma_n^i = \{y_n = t_n^i\} \cap C^\varepsilon$ for n large enough.

Combining this with (3.12) yields that the boundary Γ_n^i belongs to class $C^{0,1}$ (here, we can remove all isolated points of Γ_n^i). By integration by parts, we have

$$
\begin{aligned}
D_n^2 &= -2\int_{\{y_n<t_n^i\}\cap C^\varepsilon}(t_i-y_n)\nabla y_n\cdot\nabla\bar\varphi\,\mathrm{d}x = \int_{\{y_n<t_n^i\}\cap C^\varepsilon}\nabla(y_n-t_i)^2\cdot\nabla\bar\varphi\,\mathrm{d}x\\
&= -\int_{\{y_n<t_n^i\}\cap C^\varepsilon}(y_n-t_i)^2\Delta\bar\varphi\,\mathrm{d}x + \int_{\Gamma_n^i}(y_n-t_i)^2\nabla\bar\varphi\cdot\nu_n^i\,\mathrm{d}\mathcal{H}^1(x)\\
&= -\int_{\{y_n<t_n^i\}\cap C^\varepsilon}(y_n-t_i)^2\Delta\bar\varphi\,\mathrm{d}x + (t_n^i-t_i)^2\int_{\Gamma_n^i}\nabla\bar\varphi\cdot\nu_n^i\,\mathrm{d}\mathcal{H}^1(x),
\end{aligned}
$$

where ν_n^i stands for the outward unit normal vector to Γ_n^i. From this and the fact that $\bar y(x)\geq t_i$ for all $x\in C^\varepsilon$, we have

$$
|D_n^2|\leq\int_{\{y_n<t_n^i\}\cap C^\varepsilon}(y_n-\bar y)^2|\Delta\bar\varphi|\,\mathrm{d}x + (t_n^i-t_i)^2\int_{\Gamma_n^i}|\nabla\bar\varphi|\,\mathrm{d}\mathcal{H}^1(x) = o(s_n^2),
$$

due to the choice of t_n^i in (3.12). In conclusion, we derive $D_n^1=o(s_n^2)$, $D_n^2=o(s_n^2)$ and thus $C_n^2=D_n^1+D_n^2=o(s_n^2)$. Besides, from the second identity in (3.11) and the definition of C_n^3 in (3.9), we have $C_n^3=0$. Combing this with (3.9) yields $C_n=o(s_n^2)$. We then deduce from (3.8) that $A_n(C,\varepsilon)=o(s_n^2)$. Consequently, the desired conclusion of the lemma follows. □

Similar to Proposition 3.4, we have the following result.

Proposition 3.8. *Under all assumptions of Proposition 3.4, there exists an $\varepsilon_0=\varepsilon_0(C)>0$ such that, for any $\varepsilon\in(0,\varepsilon_0)$,*

$$
\frac{\tilde A_n(C,\varepsilon)}{s_n^2}\to-\frac{1}{2}\int_C\mathbb{1}_{\{|\nabla\bar y|>0\}}w^2\frac{\nabla\bar y\cdot\nabla\bar\varphi}{|\nabla\bar y|}\,\mathrm{d}\mathcal{H}^1(x)\quad as\quad n\to\infty, \tag{3.13}
$$

where $\tilde A_n(C,\varepsilon):=\int_{C^\varepsilon}(t_i-\bar y)[\mathbb{1}_{\Omega^{i,2}_{y_n,\bar y}}-\mathbb{1}_{\Omega^{i-1,3}_{y_n,\bar y}}]\nabla\bar y\cdot\nabla\bar\varphi\,\mathrm{d}x$.

Proof. We first note from the definition (2.11) of the sets $\Omega^{i,2}_{y_n,y}$ and $\Omega^{i-1,3}_{y_n,y}$ that

$$
\Omega^{i,2}_{y_n,\bar y}=\{\bar y\in(t_i,t_i+\delta),\,y_n\in(t_i-\delta,t_i]\},\qquad \Omega^{i-1,3}_{y_n,\bar y}=\{\bar y\in(t_i-\delta,t_i),\,y_n\in[t_i,t_i+\delta)\}
$$

and thus

$$
\begin{aligned}
\Omega^{i,2}_{y_n,y}&=[\Omega^{i-1,3}_{y,y_n}\cup\{y\in(t_i,t_i+\delta),\,y_n=t_i\}]\setminus\{y=t_i,\,y_n\in(t_i-\delta,t_i)\},\\
\Omega^{i-1,3}_{y_n,y}&=[\Omega^{i,2}_{y,y_n}\cup\{y\in(t_i-\delta,t_i),\,y_n=t_i\}]\setminus\{y=t_i,\,y_n\in(t_i,t_i+\delta)\}.
\end{aligned}
$$

Combining this with the definition of $\tilde A_n(C,\varepsilon)$ and the fact $\nabla\bar y$ vanishes a.e. on $\{\bar y=t_i\}$ yields

$$
\begin{aligned}
\tilde A_n(C,\varepsilon)=&-\int_{C^\varepsilon}(t_i-\bar y)[\mathbb{1}_{\Omega^{i,2}_{\bar y,y_n}}-\mathbb{1}_{\Omega^{i-1,3}_{\bar y,y_n}}]\nabla\bar y\cdot\nabla\bar\varphi\,\mathrm{d}x\\
&+\int_{C^\varepsilon}(t_i-\bar y)[1_{\{\bar y\in(t_i,t_i+\delta),y_n=t_i\}}-\mathbb{1}_{\{\bar y\in(t_i-\delta,t_i),y_n=t_i\}}]\nabla\bar y\cdot\nabla\bar\varphi\,\mathrm{d}x=:-\tilde C_n+\tilde B_n.
\end{aligned}
$$

Since $\nabla y_n=0$ a.e. on $\{y_n=t_i\}$, $\tilde B_n$ can be rewritten as follows

$$
\tilde B_n=\int_{C^\varepsilon}(y_n-\bar y)[1_{\{\bar y\in(t_i,t_i+\delta),y_n=t_i\}}-\mathbb{1}_{\{\bar y\in(t_i-\delta,t_i),y_n=t_i\}}]\nabla(\bar y-y_n)\cdot\nabla\bar\varphi\,\mathrm{d}x.
$$

Analogous to (3.7), one has $\tilde{B}_n = o(s_n^2)$. This implies that

$$\tilde{A}_n(C, \varepsilon) = o(s_n^2) - \tilde{C}_n. \tag{3.14}$$

From the definitions of C_n in (3.9) (see also (3.6)) and of $\tilde{C}_n$, we can get $\tilde{C}_n$ by interchanging $\bar{y}$ and y_n in the integrand of C_n. Similar to (3.10), we have for the case $|\nabla \bar{y}| > 0$ on C that

$$\tilde{C}_n = o(s_n^2) + \frac{1}{2} \int_{\{y_n = t_i\} \cap C^\varepsilon} (\bar{y} - y_n)^2 \nabla \bar{\varphi} \cdot \frac{\nabla y_n}{|\nabla y_n|} \,\mathrm{d}\mathcal{H}^1(x)$$

and for the case $|\nabla \bar{y}| = 0$ on C that $\tilde{C}_n = o(s_n^2)$. From this and (3.14), we derive (3.13). □

As a consequence of Propositions 3.4 and 3.8, we have an explicit formula for the crucial terms $\tilde{Q}$ and Q_2 in (2.16), and an important limit that will play a significant role in establishing the error estimates for the numerical approximations of (P).

Corollary 3.9. *Assume that $S(B_{L^2(\Omega)}(\bar{u}, \rho))$ is bounded in $W^{2,p}(\Omega)$ for some $p > 2$, $\rho > 0$. Assume further that the level sets $\{\bar{y} = t_i\}, 1 \leq i \leq K$ and $\bar{y} := S(\bar{u})$, decompose into finitely many connected components and $\bar{y}$ fulfills (3.1) for all connected component C of $\{\bar{y} = t_i\}$. Let $v \in L^2(\Omega)$ and $\bar{\varphi} \in W^{1,\infty}(\Omega) \cap W^{2,1}(\Omega)$. Then, for any $\{s_n\} \in c_0^+$, $y_n := S(\bar{u} + s_n v)$, and $w := S'(\bar{u})v$, there hold*

$$Q_2(\bar{u}, \bar{y}, \bar{\varphi}; v) = \tilde{Q}(\bar{u}, \bar{y}, \bar{\varphi}; \{s_n\}, v) = \frac{1}{2} \sum_{i=1}^{K} \{a'\}_{t_i+0}^{t_i-0} \int_{\{\bar{y}=t_i\}} \mathbb{1}_{\{|\nabla \bar{y}|>0\}} w^2 \frac{\nabla \bar{y} \cdot \nabla \bar{\varphi}}{|\nabla \bar{y}|} \,\mathrm{d}\mathcal{H}^1(x), \tag{3.15}$$

$$\lim_{n \to \infty} \frac{1}{s_n^2} \int_\Omega (2t_i - \bar{y} - y_n)[\mathbb{1}_{\Omega^{i,2}_{y_n,\bar{y}}} - \mathbb{1}_{\Omega^{i-1,3}_{y_n,\bar{y}}}] \nabla \bar{y} \cdot \nabla \bar{\varphi} \,\mathrm{d}x = 0. \tag{3.16}$$

Proof. Assume that $\{\bar{y} = t_i\} = \bigcup_{k=1}^{n_i} C_k^i$, $1 \leq i \leq K$, with C_k^i being a connected component of $\{\bar{y} = t_i\}$. Let $\varepsilon_0 > 0$ be such that $(C_{k_1}^i)^{\varepsilon_0} \cap (C_{k_2}^j)^{\varepsilon_0} = \emptyset$ for all $(k_1, i) \neq (k_2, j)$. By Proposition 3.4, for any $1 \leq i \leq K$ and $1 \leq k \leq n_i$, there exists $\varepsilon_k^i > 0$ satisfying the claims of Proposition 3.4 in place of ε_0. We now set $\varepsilon_* := \min\{\varepsilon_0, \varepsilon_k^i \mid 1 \leq i \leq K, 1 \leq k \leq n_i\} > 0$ and fix $\varepsilon \in (0, \varepsilon_*)$ arbitrarily. We first prove (3.15). To this end, we deduce that $y_n \to \bar{y}$ in $W_0^{1,p}(\Omega)$ due to Theorem 2.1 and $y_n \in W^{2,p}(\Omega)$ for n large enough. Since $W^{2,p}(\Omega) \Subset C^1(\overline{\Omega})$ is compact, $y_n \to \bar{y}$ in $C^1(\overline{\Omega})$. Therefore, all assumptions required in Proposition 3.4 are fulfilled. From (2.11), we have

$$\Omega^{i,2}_{y_n,\bar{y}} \cup \Omega^{i-1,3}_{y_n,\bar{y}} \subset \{0 < |\bar{y} - t_i| \leq \tau_n\} \subset \bigcup_{k=1}^{n_i} (C_k^i)^\varepsilon$$

for n large enough, where $\tau_n := \|y_n - \bar{y}\|_{C(\overline{\Omega})}$. From the definition of $\tilde{Q} := \tilde{Q}(\bar{u}, \bar{y}, \bar{\varphi}; \{s_n\}, v)$ in (2.15) and of $A_n(C, \varepsilon)$ in Proposition 3.4, we then arrive at

$$\tilde{Q} = \liminf_{n \to \infty} \frac{1}{s_n^2} \sum_{i=1}^{K} \{a'\}_{t_i+0}^{t_i-0} \sum_{k=1}^{n_i} A_n(C_k^i, \varepsilon) = \sum_{i=1}^{K} \{a'\}_{t_i+0}^{t_i-0} \sum_{k=1}^{n_i} \frac{1}{2} \int_{C_k^i} \mathbb{1}_{\{|\nabla \bar{y}|>0\}} w^2 \frac{\nabla \bar{y} \cdot \nabla \bar{\varphi}}{|\nabla \bar{y}|} \,\mathrm{d}\mathcal{H}^1(x),$$

which, together with the disjoint decomposition $\{\bar{y} = t_i\} = \bigcup_{k=1}^{n_i} C_k^i$ yields the second identity in (3.15). Moreover, from the definition of Q_2, we therefore derive the first one in (3.15).

Analogously, by using the definitions of $A_n(C, \varepsilon)$ and of $\tilde{A}_n(C, \varepsilon)$, respectively, in Propositions 3.4 and 3.8, there holds

$$\int_\Omega (2t_i - \bar{y} - y_n)[\mathbb{1}_{\Omega^{i,2}_{y_n,\bar{y}}} - \mathbb{1}_{\Omega^{i-1,3}_{y_n,\bar{y}}}] \nabla \bar{y} \cdot \nabla \bar{\varphi} \,\mathrm{d}x = \sum_{k=1}^{n_i} [A_n(C_k^i, \varepsilon) + \tilde{A}_n(C_k^i, \varepsilon)].$$

We thus derive (3.16) by Propositions 3.4 and 3.8. □

The following result shows the finiteness of the jump functional $\Sigma(\bar y)$ determined in (2.17) under a bound of the $\mathcal{H}^1$-measure of the level sets of $\bar y$ (in particular, the assumption (3.1)).

Proposition 3.10. *Let $\bar u \in L^2(\Omega)$ be such that $\bar y := S(\bar u)$ is Lipschitz continuous on $\overline{\Omega}$. If there exist $r_0 > 0$ and $C > 0$ satisfying*

$$\mathcal{H}^1(\{\bar y = t\}) \le C \quad \text{for a.e.} t \in \bigcup_{i=1}^{K} (t_i - r_0, t_i + r_0),$$

then $\Sigma(\bar y) < \infty$.

Proof. Applying the coarea formula for Lipschitz mappings (see, e.g. [20, Thm. 2, p. 117] and [1, Sec. 2.7]), yields for any $r \in (0, r_0)$ that

$$\int_\Omega \mathbb{1}_{\{0<|\bar y - t_i|<r\}} |\nabla \bar y| dx = \int_{\mathbb{R}} \int_{\{\bar y = t\}} \mathbb{1}_{\{0<|\bar y - t_i|<r\}} \, d\mathcal{H}^1(x) dt = \int_{t_i - r}^{t_i + r} \int_{\{\bar y = t\}} d\mathcal{H}^1(x) dt \le 2rC.$$

From this and the definition of $\Sigma(\bar y)$ in (2.17), we conclude that $\Sigma(\bar y) < \infty$. □

4 ANALYSIS OF THE DISCRETE STATE EQUATION

In this section, we study the discrete version of the state equation (2.4) and show error estimates of solutions to the discrete state equation (4.1), local uniqueness of these solutions, and local differentiability of the solution operators of (4.1). To this end, we introduce a family of regular triangulations $\{\mathcal{T}_h\}_{h>0}$: $\overline{\Omega} = \bigcup_{T \in \mathcal{T}_h} T$ for all $h > 0$. For each element $T \in \mathcal{T}_h$, we denote by $\varrho(T)$ and $\delta(T)$ the diameter of T and the diameter of the largest ball contained in T, respectively. The mesh size of $\mathcal{T}_h$ will be denoted by $h := \max_{T \in \mathcal{T}_h} \varrho(T)$. This triangulation is assumed to be *regular* in the sense that there exist $\bar\varrho, \bar\delta > 0$ such that $\frac{\varrho(T)}{\delta(T)} \le \bar\delta$ and $\frac{h}{\varrho(T)} \le \bar\varrho$ for all $T \in \mathcal{T}_h$ and $h > 0$; see, e.g. [15].

We will employ the standard continuous piecewise linear finite elements for the state y and set

$$V_h := \left\{ v_h \in C(\overline{\Omega}) \mid v_{h|_T} \in \mathcal{P}_1 \text{ for all} T \in \mathcal{T}_h, v_h = 0 \text{ on} \partial\Omega \right\},$$

where $\mathcal{P}_1$ stands for the space of polynomials of degree equal at most 1. The discrete approximation of the state equation (2.4) for $y_h \in V_h$ is then

$$\int_\Omega (b + a(y_h)) \nabla y_h \cdot \nabla v_h \, dx = \int_\Omega u v_h \, dx \qquad \text{for all} v_h \in V_h. \tag{4.1}$$

While the existence of solutions to (4.1) follows from [7, Thm. 3.1], the uniqueness of solutions is still an open problem. However, if $a : \mathbb{R} \to R$ is assumed to be bounded, then we have uniqueness provided that h is small enough; see [7, Thm. 4.1]. Below, we provide some error estimates for solutions to (4.1) that are sufficiently close to the solutions of (2.4).

In what follows, we fix $\bar u \in L^2(\Omega)$ and set $\bar y := S(\bar u)$. From Theorem 2.1 and the continuous embedding $L^2(\Omega) \hookrightarrow W^{-1,p}(\Omega)$ for any $p > 1$, we then have $\bar y \in W_0^{1,p}(\Omega) \cap H^2(\Omega)$.

Theorem 4.1 ([11, Thm. 3.1]). *Let $U := \overline{B}_{L^2(\Omega)}(\bar u, \rho_0)$ for some $\rho_0 > 0$ and let $p \ge 2$. Assume that Assumptions (A1) to (A3) are fulfilled. Then there exists a constant $h_0 > 0$ such that for any $u \in U$ and $h < h_0$, there exists at least one solution $y_h(u)$ to (4.1) satisfying for $y_u := S(u)$*

$$\|y_u - y_h(u)\|_{L^2(\Omega)} + h\|y_u - y_h(u)\|_{H_0^1(\Omega)} + h\|y_u - y_h(u)\|_{L^\infty(\Omega)} \le C_U h^2, \tag{4.2}$$

$$\|y_u - y_h(u)\|_{W_0^{1,p}(\Omega)} \le C_{U,p} h^{2/p}. \tag{4.3}$$

Proof. The estimates for the norms in L^2, H_0^1, and $W^{1,p}$ are shown in [11, Thm. 3.1], while the estimate for the L^∞ norm can be obtained similar to estimate (3.11) in [11, Thm. 3.1]. □

The following theorem guarantees the local uniqueness of solutions to (4.1). Its proof is similar to that of [8, Thm. 4.2] with slight modifications and is thus omitted here.

Theorem 4.2. *Let $p > 2$ be arbitrary and let h_0 be defined in Theorem 4.1. Under Assumptions (A1) to (A3), there exist $h_1 \in (0, h_0)$, $\rho > 0$, and $\kappa_\rho > 0$ such that for any $h < h_1$ and any $u \in \overline{B}_{L^2(\Omega)}(\bar{u}, \rho)$, (4.1) admits a unique solution in $\overline{B}_{W_0^{1,p}(\Omega)}(\bar{y}, \kappa_\rho) \cap V_h$.*

From now on, let us fix $\tilde{p} \geq 4$ and let h_0, h_1, ρ, and κ_ρ be the constants defined in Theorems 4.1 and 4.2 for $p = \tilde{p}$. In the rest of this section, we shall investigate the differentiability of the discrete solution operator

$$S_h : B_{L^2(\Omega)}(\bar{u}, \rho) \ni u \mapsto y_h(u) \in \overline{B}_{W_0^{1,\tilde{p}}(\Omega)}(\bar{y}, \kappa_\rho) \cap V_h, \tag{4.4}$$

where $y_h(u)$ is the unique solution to (4.1) in $\overline{B}_{W_0^{1,\tilde{p}}(\Omega)}(\bar{y}, \kappa_\rho)$ from Theorem 4.2.

For any $y, \hat{y} \in C(\overline{\Omega}) \cap W^{1,1}(\Omega)$, we define functions $T_{y,\hat{y}}$ and $Z_{y,\hat{y}}$ on Ω via

$$T_{y,\hat{y}} := \mathbb{1}_{\{\hat{y} \notin E_a\}}[a(y) - a(\hat{y}) - a'(\hat{y})(y - \hat{y})], \quad Z_{y,\hat{y}} := \mathbb{1}_{\{y \notin E_a\}} a'(y) \nabla y - \mathbb{1}_{\{\hat{y} \notin E_a\}} a'(\hat{y}) \nabla \hat{y}. \tag{4.5}$$

In order to prove the differentiability of S_h, we need the following lemmas.

Lemma 4.3 ([16, Lem. 3.3]). *Let Assumption (A3) be fulfilled. Assume that $y_n \to y$ in $W_0^{1,p}(\Omega)$ as $n \to \infty$ with $p > 2$. Then*

$$\frac{1}{\|y_n - y\|_{W_0^{1,p}(\Omega)}} \|T_{y_n,y} \nabla y\|_{L^p(\Omega)} \to 0 \quad \text{as} \, n \to \infty.$$

Lemma 4.4. *Let Assumption (A3) be fulfilled and let $y, \hat{y} \in C(\overline{\Omega}) \cap W^{1,1}(\Omega)$ and $M > 0$ be arbitrary such that $\|y - \hat{y}\|_{C(\overline{\Omega})} < \delta$ with δ defined in (2.12) and $\|y\|_{C(\overline{\Omega})}, \|\hat{y}\|_{C(\overline{\Omega})} \leq M$. Then*

$$Z_{y,\hat{y}} = Z^{(1)}_{y,\hat{y}} + Z^{(2)}_{y,\hat{y}} + Z^{(3)}_{y,\hat{y}} + Z^{(4)}_{y,\hat{y}}, \tag{4.6}$$

for

$$\begin{aligned}
Z^{(1)}_{y,\hat{y}} &:= \sum_{i=0}^{K} \mathbb{1}_{\{\hat{y} \in (t_i, t_{i+1}), y \in (t_i, t_{i+1})\}} [a_i'(y) \nabla y - a_i'(\hat{y}) \nabla \hat{y}], \\
Z^{(2)}_{y,\hat{y}} &:= \sum_{i=1}^{K} \mathbb{1}_{\{\hat{y} = t_i\}} \mathbb{1}_{\{y \notin E_a\}} a'(y) \nabla y, \qquad Z^{(3)}_{y,\hat{y}} := \sum_{i=0}^{K} [\{a'\}_{t_i+0}^{t_i-0} \mathbb{1}_{\Omega^{i,2}_{y,\hat{y}}} - \{a'\}_{t_{i+1}+0}^{t_{i+1}-0} \mathbb{1}_{\Omega^{i,3}_{y,\hat{y}}}] \nabla \hat{y}, \\
Z^{(4)}_{y,\hat{y}} &:= \sum_{i=0}^{K} \Big\{ \mathbb{1}_{\Omega^{i,2}_{y,\hat{y}}} [a'_{i-1}(y) \nabla (y - \hat{y}) + (a'_{i-1}(y) - a'_{i-1}(t_i)) \nabla \hat{y} + (a_i'(t_i) - a_i'(\hat{y})) \nabla \hat{y}] \\
&\qquad + \mathbb{1}_{\Omega^{i,3}_{y,\hat{y}}} [a'_{i+1}(y) \nabla (y - \hat{y}) + (a'_{i+1}(y) - a'_{i+1}(t_{i+1})) \nabla \hat{y} + (a_i'(t_{i+1}) - a_i'(\hat{y})) \nabla \hat{y}] \Big\},
\end{aligned}$$

with the conventions (2.14) and the sets $\Omega^{i,j}_{y,\hat{y}}$ defined as in (2.11). Moreover, there exists a constant $C_M > 0$ such that a.e. in Ω,

$$\begin{cases} |Z^{(1)}_{y,\hat{y}}| + |Z^{(2)}_{y,\hat{y}}| \leq C_M [|y - \hat{y}| |\nabla \hat{y}| + |\nabla (y - \hat{y})|], \\ \qquad |Z^{(4)}_{y,\hat{y}}| \leq C_M [|y - \hat{y}| |\nabla \hat{y}| + |\nabla (y - \hat{y})|] \displaystyle\sum_{i=0}^{K} (\mathbb{1}_{\Omega^{i,2}_{y,\hat{y}}} + \mathbb{1}_{\Omega^{i,3}_{y,\hat{y}}}). \end{cases} \tag{4.7}$$

Consequently, $Z_{y,\hat{y}} \to 0$ in $L^p(\Omega)$ as $y \to \hat{y}$ in $W^{1,p}(\Omega) \cap C(\overline{\Omega})$ for any $p \geq 1$.

Proof. Clearly,

$$Z_{y,\hat{y}} = \sum_{i=0}^{K} (Z_i^1 + Z_i^2) \tag{4.8}$$

with $Z_i^1 := \mathbb{1}_{\{\hat{y}\in(t_i,t_{i+1})\}}[\mathbb{1}_{\{y\notin E_a\}}a'(y)\nabla y - a_i'(\hat{y})\nabla\hat{y}]$ and $Z_i^2 := \mathbb{1}_{\{\hat{y}=t_{i+1}\}}\mathbb{1}_{\{y\notin E_a\}}a'(y)\nabla y$. From the fact that $\mathbb{1}_{\{\hat{y}=t_{K+1}\}} = \mathbb{1}_{\{\hat{y}=\infty\}} = 0$ and the definition of $Z^{(2)}_{y,\hat{y}}$, we obtain

$$Z^{(2)}_{y,\hat{y}} = \sum_{i=0}^{K-1} Z_i^2 = \sum_{i=0}^{K} Z_i^2. \tag{4.9}$$

Since $\|y-\hat{y}\|_{C(\overline{\Omega})} < \delta$ and $\nabla y = 0$ a.e. in $\{y = t_i\} \cup \{y = t_{i+1}\}$ (see [13]), we can write

$$\begin{aligned} Z_i^1 = \mathbb{1}_{\{\hat{y}\in(t_i,t_i+\delta),y\in(t_i-\delta,t_i]\}}&[a_{i-1}'(y)\nabla y - a_i'(\hat{y})\nabla\hat{y}] \\ &+ \mathbb{1}_{\{\hat{y}\in(t_i,t_{i+1}),y\in(t_i,t_{i+1})\}}[a_i'(y)\nabla y - a_i'(\hat{y})\nabla\hat{y}] \\ &+ \mathbb{1}_{\{\hat{y}\in(t_{i+1}-\delta,t_{i+1}),y\in[t_{i+1},t_{i+1}+\delta)\}}[a_{i+1}'(y)\nabla y - a_i'(\hat{y})\nabla\hat{y}] =: Z_i^{1,1} + Z_i^{1,2} + Z_i^{1,3}. \end{aligned}$$

Thus, we have from the definition of $\Omega^{i,j}_{y,\hat{y}}$ in (2.11) and of $Z^{(1)}_{y,\hat{y}}$ that

$$Z^{(1)}_{y,\hat{y}} = \sum_{i=0}^{K} Z_i^{1,2}. \tag{4.10}$$

We now write

$$Z_i^{1,1} = \mathbb{1}_{\Omega^{i,2}_{y,\hat{y}}}[a_{i-1}'(t_i) - a_i'(t_i)]\nabla\hat{y} + \tilde{Z}_i^{1,1} \quad \text{and} \quad Z_i^{1,3} = \mathbb{1}_{\Omega^{i,3}_{y,\hat{y}}}[a_{i+1}'(t_{i+1}) - a_i'(t_{i+1})]\nabla\hat{y} + \tilde{Z}_i^{1,3}$$

with

$$\begin{aligned} \tilde{Z}_i^{1,1} &:= \mathbb{1}_{\Omega^{i,2}_{y,\hat{y}}}[a_{i-1}'(y)\nabla(y-\hat{y}) + (a_{i-1}'(y) - a_{i-1}'(t_i))\nabla\hat{y} + (a_i'(t_i) - a_i'(\hat{y}))\nabla\hat{y}], \\ \tilde{Z}_i^{1,3} &:= \mathbb{1}_{\Omega^{i,3}_{y,\hat{y}}}[a_{i+1}'(y)\nabla(y-\hat{y}) + (a_{i+1}'(y) - a_{i+1}'(t_{i+1}))\nabla\hat{y} + (a_i'(t_{i+1}) - a_i'(\hat{y}))\nabla\hat{y}]. \end{aligned}$$

Obviously, we have $Z^{(4)}_{y,\hat{y}} = \sum_{i=0}^{K}(\tilde{Z}_i^{1,1} + \tilde{Z}_i^{1,3})$ and there then holds $Z^{(3)}_{y,\hat{y}} + Z^{(4)}_{y,\hat{y}} = \sum_{i=0}^{K}(Z_i^{1,1} + Z_i^{1,3})$. From this and (4.8)–(4.10), we derive (4.6). Moreover, (4.7) is derived by combining the definition of $Z^{(k)}_{y,\hat{y}}$, Assumption (A3), the definition of $\Omega^{i,2}_{y,\hat{y}}$ and $\Omega^{i,3}_{y,\hat{y}}$, and the fact that $\nabla\hat{y} = 0$ a.e. in $\{\hat{y} = t_i\}$ (see; e.g. [13, Rem. 2.6]). Finally, the claimed convergence follows from (4.6), (4.7), the fact that $\mathbb{1}_{\Omega^{i,2}_{y,\hat{y}}}, \mathbb{1}_{\Omega^{i,3}_{y,\hat{y}}} \to 0$ a.e. in Ω as $y \to \hat{y}$ in $C(\overline{\Omega})$, and Lebesgue's dominated convergence theorem. □

For any $h \in (0, h_1)$ and $y_h \in V_h$, we now define the operator $D_{h,y_h} : V_h \to V_h^*$ via

$$\langle D_{h,y_h} w_h, z_h\rangle := \int_\Omega [(b + a(y_h))\nabla w_h + \mathbb{1}_{\{y_h\notin E_a\}}a'(y_h)\nabla y_h w_h] \cdot \nabla z_h \,\mathrm{d}x, \; w_h, z_h \in V_h. \tag{4.11}$$

Lemma 4.5. *Let all assumptions of Theorem 4.1 hold. Then for any $h \in (0, h_1)$ and any $\{y_h^k\} \subset V_h$ converging to $y_h \in V_h$ in $H_0^1(\Omega)$ as $k \to \infty$, there holds $\|D_{h,y_h^k} - D_{h,y_h}\|_{\mathbb{L}(V_h,V_h^*)} \to 0$.*

Proof. Let $w_h, v_h \in V_h$ be arbitrary such that $\|w_h\|_{H_0^1(\Omega)}, \|v_h\|_{H_0^1(\Omega)} \le 1$ and $h \in (0, h_1)$ be arbitrary but fixed. Assume that $\{y_h^k\} \subset V_h$ converges to $y_h \in V_h$ in $H_0^1(\Omega)$ as $k \to \infty$. By virtue of the inverse inequality [15, Thm. 3.2.6], we deduce that $y_h^k \to y_h$ in $W_0^{1,\tilde{p}}(\Omega)$ and hence in $C(\overline{\Omega})$ as $k \to \infty$. We can therefore assume that $\|y_h^k - y_h\|_{C(\overline{\Omega})} < \delta$ for all $k \in \mathbb{N}$ large enough. On the other hand, we have

$$\langle (D_{h,y_h^k} - D_{h,y_h})w_h, v_h\rangle = \int_\Omega [(a(y_h^k) - a(y_h))\nabla w_h + Z_{y_h^k,y_h} w_h] \cdot \nabla v_h \,\mathrm{d}x.$$

Together with the Hölder inequality, this yields that

$$\begin{aligned}\|D_{h,y_h^k} - D_{h,y_h}\|_{\mathbb{L}(V_h,V_h^*)} &\le \|a(y_n^k) - a(y_h)\|_{L^\infty(\Omega)} + \|Z_{y_h^k,y_h}\|_{L^{\tilde{p}}(\Omega)}\|w_h\|_{L^{2\tilde{p}/(\tilde{p}-2)}(\Omega)} \\ &\le \|a(y_h^k) - a(y_h)\|_{L^\infty(\Omega)} + C\|Z_{y_h^k,y_h}\|_{L^{\tilde{p}}(\Omega)},\end{aligned}$$

where we have employed the continuous embedding $H_0^1(\Omega) \hookrightarrow L^{2\tilde{p}/(\tilde{p}-2)}(\Omega)$ and the fact that $\|w_h\|_{H_0^1(\Omega)} \le 1$ to obtain the last inequality. The first term on the right-hand side of the last estimate tends to zero as $k \to \infty$ since $y_h^k \to y_h$ in $C(\overline{\Omega})$ as $k \to \infty$. Moreover, the second term tends to zero as a result of Lemma 4.4. □

Lemma 4.6. *Let all assumptions of Theorem 4.1 hold. Then there exists a constant $h_2 \in (0, h_1)$ such that for any $h \in (0, h_2)$ and any $y_h \in \overline{B}_{W_0^{1,\tilde{p}}(\Omega)}(\bar{y}, \kappa_\rho) \cap V_h$, the operator $D_{h,y_h} : V_h \to V_h^*$ is an isomorphism.*

Proof. Since V_h is finite-dimensional and D_{h,y_h} is linear, it suffices to prove that there exists an $h_2 \in (0, h_1)$ such that for any $h \in (0, h_2)$ and $y_h \in \overline{B}_{W_0^{1,\tilde{p}}(\Omega)}(\bar{y}, \kappa_\rho) \cap V_h$, the equation

$$D_{h,y_h} w_h = 0 \tag{4.12}$$

admits the unique solution $w_h = 0$. We argue by contradiction. Assume for any $k \ge 1$ that there exist $h_k \in (0, h_1)$, $y_{h_k} \in \overline{B}_{W_0^{1,\tilde{p}}(\Omega)}(\bar{y}, \kappa_\rho) \cap V_{h_k}$, and $w_{h_k} \in V_h\backslash\{0\}$ such that $h_k \to 0^+$ and w_{h_k} solves (4.12) for $h = h_k$ and $y_h = y_{h_k}$. By setting $\hat{w}_{h_k} := \frac{w_{h_k}}{\|w_{h_k}\|_{L^{2\tilde{p}/(\tilde{p}-2)}(\Omega)}}$, we deduce that

$$\|\hat{w}_{h_k}\|_{L^{2\tilde{p}/(\tilde{p}-2)}(\Omega)} = 1 \quad \text{and} \quad D_{h_k,y_{h_k}}\hat{w}_{h_k} = 0. \tag{4.13}$$

Furthermore, as a result of the embedding $W_0^{1,\tilde{p}}(\Omega) \Subset C(\overline{\Omega})$, there hold that $\|y_{h_k}\|_{C(\overline{\Omega})} \le M$ for all $k \ge 1$ and some constant $M > 0$ independent of k and that

$$y_{h_k} \to y \text{ in} C(\overline{\Omega}) \quad \text{for some} y \in W_0^{1,\tilde{p}}(\Omega). \tag{4.14}$$

Testing the second equation in (4.13) by $\hat{w}_{h_k}$, Hölder's inequality thus gives

$$\underline{b}\|\nabla\hat{w}_{h_k}\|_{L^2(\Omega)} \le \|\nabla y_{h_k}\|_{L^{\tilde{p}}(\Omega)}\|\mathbb{1}_{\{y_{h_k}\notin E_a\}}a'(y_{h_k})\|_{L^\infty(\Omega)}\|\hat{w}_{h_k}\|_{L^{2\tilde{p}/(\tilde{p}-2)}(\Omega)} \le C_{M,\rho}$$

for some constant $C_{M,\rho} > 0$. From this and the compact embedding $H_0^1(\Omega) \Subset L^{2\tilde{p}/(\tilde{p}-2)}(\Omega)$, a subsequence argument shows that we can assume that

$$\hat{w}_{h_k} \rightharpoonup \hat{w} \text{ in} H_0^1(\Omega) \quad \text{and} \quad \hat{w}_{h_k} \to \hat{w} \text{ in} L^{2\tilde{p}/(\tilde{p}-2)}(\Omega) \tag{4.15}$$

for some $\hat{w} \in H_0^1(\Omega)$. Moreover, there exist an element $\mathbb{b} \in L^{\tilde{p}}(\Omega)^2$ and a subsequence of $\{\mathbb{b}_k\}$ with $\mathbb{b}_k := \mathbb{1}_{\{y_{h_k}\notin E_a\}}a'(y_{h_k})\nabla y_{h_k}$, denoted in the same way, such that $\mathbb{b}_k \rightharpoonup \mathbb{b}$ weakly in $L^{\tilde{p}}(\Omega)^2$. By fixing any $v \in H^2(\Omega) \cap H_0^1(\Omega)$ and testing the last equation in (4.13) with $v_{h_k} := \Pi_{h_k}v \in V_{h_k}$, where Π_{h_k} is the interpolation operator, we have

$$\int_\Omega [(b + a(y_{h_k}))\nabla\hat{w}_{h_k} + \hat{w}_{h_k}\mathbb{b}_k] \cdot \nabla v_{h_k} \,\mathrm{d}x = 0 \quad \text{for all} k \ge 1.$$

Letting $k \to \infty$ and exploiting the limits (4.14), (4.15), $\mathbb{b}_k \rightharpoonup \mathbb{b}$ in $L^{\tilde{p}}(\Omega)^N$, and $v_{h_k} \to v$ in $H_0^1(\Omega)$, we can conclude that $\int_\Omega [(b + a(y))\nabla \hat{w} + \hat{w}\mathbb{b}] \cdot \nabla v \,dx = 0$. From this, the density of $H^2(\Omega) \cap H_0^1(\Omega)$ in $H_0^1(\Omega)$, and [7, Thm. 2.6], we conclude that $\hat{w} = 0$, contradicting the fact that $\|\hat{w}\|_{L^{2\tilde{p}/(\tilde{p}-2)}(\Omega)} = \lim_{k\to\infty} \|\hat{w}_{h_k}\|_{L^{2\tilde{p}/(\tilde{p}-2)}(\Omega)} = 1$. □

As a consequence of Lemmas 4.5 and 4.6 and the implicit function theorem, we obtain the differentiability of S_h.

Theorem 4.7. *Let all assumptions of Theorem 4.1 hold. Then, for any $h \in (0, h_2)$, the operator S_h defined in (4.4) is of class C^1. Moreover, for any $u \in B_{L^2(\Omega)}(\bar{u}, \rho)$, let $y_h(u) := S_h(u)$. Then for any $v \in L^2(\Omega)$, the Fréchet derivative $S_h'(u)v =: z_h$ is the unique solution to*

$$\int_\Omega [(b+a(y_h(u)))\nabla z_h + \mathbb{1}_{\{y_h(u)\notin E_a\}} a'(y_h(u)) z_h \nabla y_h(u)] \cdot \nabla w_h \,dx = \int_\Omega v w_h \,dx \quad \text{for all} w_h \in V_h. \tag{4.16}$$

Proof. We first consider for any $h \in (0, h_2)$ the mapping $F_h : B_{L^2(\Omega)}(\bar{u}, \rho) \times V_h \to V_h^*$ defined via

$$\langle F_h(u, y_h), v_h\rangle = \int_\Omega (b + a(y_h))\nabla y_h \cdot \nabla v_h - u v_h \,dx, \quad u \in B_{L^2(\Omega)}(\bar{u}, \rho),\, y_h, v_h \in V_h. \tag{4.17}$$

Clearly, $F_h(u, y_h(u)) = 0$ and F_h is continuously partially differentiable in u. We now prove that F_h is partially differentiable in y_h with $\frac{\partial F_h}{\partial y_h}(u, y_h) = D_{h,y_h}$, where D_{h,y_h} is defined in (4.11). We thus derive the differentiability of S_h according to Lemmas 4.5 and 4.6 as well as a simple computation. To this end, by taking any $v_h \in V_h$ and $\{w_h^k\} \subset V_h$ with $\|w_h^k\|_{H_0^1(\Omega)} \to 0$ as $k \to \infty$ and $\|v_h\|_{H_0^1(\Omega)} \le 1$, we deduce from a straightforward computation that

$$\langle F_h(u, y_h + w_h^k) - F_h(u, y_h) - D_{h,y_h} w_h^k, v_h\rangle = \int_\Omega [T_{y_h^k, y_h} \nabla y_h + (a(y_h^k) - a(y_h))\nabla w_h^k] \cdot \nabla v_h \,dx,$$

where $y_h^k := y_h + w_h^k$ and $T_{y,\hat{y}}$ is defined in (4.5). This gives

$$\|F_h(u, y_h + w_h^k) - F_h(u, y_h) - D_{h,y_h} w_h^k\|_{\mathbb{L}(V_h,V_h^*)} \le \|T_{y_h^k, y_h} \nabla y_h\|_{L^2(\Omega)} + \|a(y_h^k) - a(y_h)\|_{L^\infty(\Omega)} \|w_h^k\|_{H_0^1(\Omega)}.$$

Moreover, in view of inverse estimates [15, Thm. 3.2.6], we have $y_h^k \to y_h$ in $W_0^{1,\tilde{p}}(\Omega)$ and hence in $C(\overline{\Omega})$ as $k \to \infty$. Then Lemma 4.3 and the embedding $W_0^{1,\tilde{p}}(\Omega) \hookrightarrow H_0^1(\Omega)$ imply that

$$\frac{1}{\|w_h^k\|_{H_0^1(\Omega)}} \|F_h(u, y_h + w_h^k) - F_h(u, y_h) - D_{h,y_h} w_h^k\|_{\mathbb{L}(V_h,V_h^*)} \to 0,$$

which gives that $\frac{\partial F_h}{\partial y_h}(u, y_h) = D_{h,y_h}$. □

5 NUMERICAL ANALYSIS OF THE ADJOINT STATE EQUATION

In this section, we will carry out the numerical analysis of the adjoint equation (2.7). For any $h \in (0, h_2)$, $u \in B_{L^2(\Omega)}(\bar{u}, \rho)$, $v \in L^2(\Omega)$ and $y_h := S_h(u)$, we approximate (2.7) using the triangulation $\mathcal{T}_h$ by

$$\int_\Omega (b + a(y_h))\nabla \varphi_h \cdot \nabla w_h + \mathbb{1}_{\{y_h\notin E_a\}} a'(y_h) w_h \nabla y_h \cdot \nabla \varphi_h \,dx = \int_\Omega v w_h \,dx \quad \text{for all} w_h \in V_h. \tag{5.1}$$

From the bijectivity of D_{h,y_h} shown in Lemma 4.6, we deduce the existence and uniqueness of solutions to (5.1).

Theorem 5.1. *Let all assumptions of Theorem 4.7 hold. Then for all $h \in (0, h_2)$, $u \in B_{L^2(\Omega)}(\bar{u}, \rho)$, and $v \in L^2(\Omega)$, there exists a unique solution $\varphi_h \in V_h$ to (5.1).*

In order to derive error estimates for the full approximation (5.1) of (2.7), we first consider the continuous problem (2.7) with $y_h(u)$ in place of y_u.

Lemma 5.2. *Let all assumptions of Theorem 4.7 hold. Then for any $h \in (0, h_2)$, $u \in B_{L^2(\Omega)}(\bar{u}, \rho)$, $y_h := S_h(u)$, and $v \in L^2(\Omega)$, the equation*

$$-\operatorname{div}[(b + a(y_h))\nabla\tilde{\varphi}] + \mathbb{1}_{\{y_h \notin E_a\}} a'(y_h)\nabla y_h \cdot \nabla\tilde{\varphi} = v \quad \text{in}\Omega, \quad \tilde{\varphi} = 0 \quad \text{on}\partial\Omega, \tag{5.2}$$

has a unique solution $\tilde{\varphi}$ in $H^2(\Omega) \cap H_0^1(\Omega)$. Moreover,

$$\|\varphi - \tilde{\varphi}\|_{H_0^1(\Omega)} \le C_\rho h \|v\|_{L^2(\Omega)} \quad \text{and} \quad \|\varphi - \tilde{\varphi}\|_{L^2(\Omega)} \le C_\rho h^2 \|v\|_{L^2(\Omega)} \tag{5.3}$$

for some constant C_ρ independent of u, v, and h, where φ is the unique solution to (2.7).

Proof. From Theorem 2.1, the continuous embedding $W_0^{1,\tilde{p}}(\Omega) \hookrightarrow C(\overline{\Omega})$, and (4.3) for $p := \tilde{p} \ge 4$, there holds

$$\|\mathbb{1}_{\{y_h \notin E_a\}} a'(y_h)\nabla y_h\|_{L^{\tilde{p}}(\Omega)} + \|b + a(y_h(u))\|_{W^{1,\tilde{p}}(\Omega)} \le C_\rho \quad \text{for all} h \in (0, h_2), u \in \overline{B}_{L^2(\Omega)}(\bar{u}, \rho).$$

A standard argument then proves the existence of solutions $\tilde{\varphi}$ to (5.2) in $H^2(\Omega) \cap H_0^1(\Omega)$; see. e.g. [7, Thm. 2.6] and the proof of Lemma 4.1 in [16]. Moreover, we have

$$\|\tilde{\varphi}\|_{H^2(\Omega)} \le C_\rho \|v\|_{L^2(\Omega)}. \tag{5.4}$$

Setting $\psi := \varphi - \tilde{\varphi}$ and subtracting the equations corresponding to φ and $\tilde{\varphi}$ yields

$$-\operatorname{div}[(b + a(y_u))\nabla\psi] + \mathbb{1}_{\{y_u \notin E_a\}} a'(y_u)\nabla y_u \cdot \nabla\psi = g_{u,h} \quad \text{in}\Omega, \quad \psi = 0 \quad \text{on}\partial\Omega, \tag{5.5}$$

with

$$g_{u,h} := \operatorname{div}[(a(y_u) - a(y_h))\nabla\tilde{\varphi}] + Z_{y_h,y_u} \cdot \nabla\tilde{\varphi}.$$

By the chain rule [23, Thm. 7.8] and the fact that $y_u, y_h \in W_0^{1,\tilde{p}}(\Omega)$ and that $\tilde{\varphi} \in H^2(\Omega)$, we can write

$$g_{u,h} = (a(y_u) - a(y_h))\Delta\tilde{\varphi}. \tag{5.6}$$

Similar to (5.4), there holds

$$\|\psi\|_{H^2(\Omega)} \le C_\rho \|g_{u,h}\|_{L^2(\Omega)} \le C_\rho \|a(y_u) - a(y_h)\|_{L^\infty(\Omega)} \|\Delta\tilde{\varphi}\|_{L^2(\Omega)}.$$

Combing this with the Lipschitz continuity of a on bounded sets, the L^∞-estimate in (4.2), and (5.4) yields the first estimate in (5.3). To show the second estimate, set $z_{u,\psi} := S'(u)\psi$ and note that $\psi = S'(u)^* g_{u,h}$. We then deduce from (5.6) that

$$\|\psi\|_{L^2(\Omega)}^2 = \int_\Omega g_{u,h} z_{u,\psi} \,\mathrm{d}x \le C_\rho \|z_{u,\psi}\|_{L^\infty(\Omega)} \|\Delta\tilde{\varphi}\|_{L^2(\Omega)} \|y_u - y_h\|_{L^2(\Omega)}. \tag{5.7}$$

By Theorem 2.1 and the compact embedding $L^2(\Omega) \Subset W^{-1,\tilde{p}}(\Omega)$,

$$\sup\{\|S'(u)\|_{\mathbb{L}(W^{-1,\tilde{p}}(\Omega), W_0^{1,\tilde{p}}(\Omega))} \mid u \in \overline{B}_{L^2(\Omega)}(\bar{u}, \rho)\} \le C_\rho. \tag{5.8}$$

The continuous embeddings $W_0^{1,\tilde{p}}(\Omega) \hookrightarrow L^\infty(\Omega)$ and $L^2(\Omega) \hookrightarrow W^{-1,\tilde{p}}(\Omega)$ therefore yield

$$\|z_{u,\psi}\|_{L^\infty(\Omega)} \le C\|z_{u,\psi}\|_{W_0^{1,\tilde{p}}(\Omega)} \le C_\rho \|\psi\|_{W^{-1,\tilde{p}}(\Omega)} \le C_\rho \|\psi\|_{L^2(\Omega)}.$$

The inequality (5.7) thus yields

$$\|\psi\|_{L^2(\Omega)} \le C_\rho \|\Delta\tilde{\varphi}\|_{L^2(\Omega)} \|y_u - y_h\|_{L^2(\Omega)}.$$

This, (4.2) and (5.4) yield the last estimate in (5.3). □

From now on, we assume further that $\bar{u} \in L^\infty(\Omega)$ and set $\rho_* := \text{meas}_{\mathbb{R}^2}(\Omega)^{-1/2}\rho$. Clearly, $B_{L^\infty(\Omega)}(\bar{u}, \rho_*) \subset B_{L^2(\Omega)}(\bar{u}, \rho)$. Moreover, the L^∞-error estimate in (4.2) yields a constant C_∞ such that

$$\|S(u) - S_h(u)\|_{L^\infty(\Omega)} \le C_\infty h \quad \text{for all} u \in \overline{B}_{L^\infty(\Omega)}(\bar{u}, \rho_*), h \in (0, h_2). \tag{5.9}$$

For any $y \in W^{1,1}(\Omega) \cap C(\overline{\Omega})$ and $r > 0$, let

$$V(y, r) := \sum_{m=1}^{2}\sum_{i=1}^{K} \sigma_i \mathbb{1}_{\{0<|y-t_i|\le r\}} |\partial_{x_m} y| \quad \text{and} \quad \Sigma_r(y) := \frac{1}{r}V(y, r). \tag{5.10}$$

Proposition 5.3. *Let $r > 0$, $y \in W^{1,1}(\Omega) \cap C(\overline{\Omega})$, and $\hat{y} \in W^{1,\infty}(\Omega)$ be arbitrary and let $\kappa := r + \|y - \hat{y}\|_{C(\overline{\Omega})}$ and $\sigma_{\max} := \max\{\sigma_i \mid 1 \le i \le K\}$ with σ_i defined in (2.2). Then*

(i) $V(y, r) \le V(\hat{y}, \kappa) + \sigma_{\max} \sum_{m=1}^{2} |\partial_{x_m} y - \partial_{x_m} \hat{y}|$ for a.e. in Ω;

(ii) $\|V(\hat{y}, r)\|^2_{L^2(\Omega)} \le 2rK\sigma_{\max}\|\nabla \hat{y}\|_{L^\infty(\Omega)}\|\Sigma_r(\hat{y})\|_{L^1(\Omega)}$.

Proof. Since $\{0 < |y - t_i| \le r\} \subset \{|\hat{y} - t_i| \le \kappa\}$ and $\hat{y}$ vanishes a.e. in $\{\hat{y} = t_i\}$, the first claim holds. The proof of the second claim is straightforward. □

Lemma 5.4. *There exist an $h_3 \in (0, h_2]$ and a constant $L_{\rho_*} > 0$ such that for all $u \in B_{L^\infty(\Omega)}(\bar{u}, \rho_*)$ and $h \in (0, h_3)$,*

$$\|Z_{y_h,y}\|_{L^2(\Omega)} \le L_{\rho_*}h + \|V(y, \|y_h - y\|_{L^\infty(\Omega)})\|_{L^2(\Omega)} \tag{5.11}$$

with $y := S(u)$ and $y_h := S_h(u)$.

Proof. By Theorem 2.1, there exists a constant M_{1,ρ_*} such that

$$\|S(u)\|_{W^{1,\infty}(\Omega)} \le M_{1,\rho_*} \quad \text{for all} u \in \overline{B}_{L^\infty(\Omega)}(\bar{u}, \rho_*). \tag{5.12}$$

Setting $h_3 := \min\{h_2, \delta 2^{-1} C_\infty^{-1}\}$ and exploiting (5.9) shows that $\|y - y_h\|_{C(\overline{\Omega})} < \delta$ for any $u \in B_{L^\infty(\Omega)}(\bar{u}, \rho_*)$ and $h \in (0, h_3)$. From the definition of $Z_{y_h,y}$ in (4.5) and Lemma 4.4, we arrive at

$$Z_{y_h,y} = Z^{(1)}_{y_h,y} + Z^{(2)}_{y_h,y} + Z^{(3)}_{y_h,y} + Z^{(4)}_{y_h,y}. \tag{5.13}$$

By (4.7), (5.12), and Theorem 4.1, we have

$$\|Z^{(k)}_{y_h,y}\|_{L^2(\Omega)} \le L_{\rho_*}h \quad \text{for} k = 1, 2, 4. \tag{5.14}$$

On the other hand, we have

$$\begin{cases} \Omega^{i,2}_{y_h,y} := \{y \in (t_i, t_i + \delta), y_h \in (t_i - \delta, t_i]\} \subset \{0 < y - t_i \le \|y_h - y\|_{L^\infty(\Omega)}\}, \\ \Omega^{i,3}_{y_h,y} := \{y \in (t_{i+1} - \delta, t_{i+1}), y_h \in [t_{i+1}, t_{i+1} + \delta)\} \subset \{0 < t_{i+1} - y \le \|y_h - y\|_{L^\infty(\Omega)}\}, \end{cases} \tag{5.15}$$

which together with the definitions of $Z^{(3)}_{y_h,y}$ in Lemma 4.4 and of V in (5.10), and (2.2), show that $|Z^{(3)}_{y_h,y}| \le V(y, \|y_h - y\|_{L^\infty(\Omega)})$ a.e. in Ω. Combining this with (5.13) and (5.14), we obtain (5.11). □

Theorem 5.5. *If $\Sigma(\bar{y}) < \infty$, then there exist constants $\bar{h} := \bar{h}(\bar{u}) \in (0, h_3)$, $\bar{\rho} := \bar{\rho}(\bar{u}) \le \rho_*$, and $C_{\bar{u}} > 0$ such that*

$$\|\varphi - \varphi_h\|_{L^2(\Omega)} \le C_{\bar{u}} \varepsilon^u_h \|v\|_{L^2(\Omega)} \quad \text{and} \quad \|\varphi - \varphi_h\|_{H^1_0(\Omega)} \le C_{\bar{u}} h \|v\|_{L^2(\Omega)}$$

for all $h \in (0, \bar{h})$, $u \in B_{L^\infty(\Omega)}(\bar{u}, \bar{\rho})$, and $v \in L^2(\Omega)$, where

$$\varepsilon^u_h := h^{1+\bar{q}} + h\|V(S(u), \|S_h(u) - S(u)\|_{L^\infty(\Omega)})\|_{L^2(\Omega)} \tag{5.16}$$

with $\bar{q} := \frac{\tilde{p}-2}{\tilde{p}}$ for $\tilde{p} \ge 4$, and φ and φ_h are the unique solutions to (2.7) and (5.1), respectively.

Proof. Let $\tilde{\varphi}$ be the solution of (5.2). To simplify the notation, set $y_u := S(u)$ and $y_h := S_h(u)$ for any $h \in (0, h_3)$ and $u \in B_{L^\infty(\Omega)}(\bar{u}, \rho_*)$. We divide the proof into three steps.

Step 1: Existence of a constant C_{1,h_3,ρ_*} such that

$$\|\tilde{\varphi} - \varphi_h\|_{L^2(\Omega)} \leq C_{1,h_3,\rho_*}(h^{\bar{q}} + \|V(y_u, \|y_h - y_u\|_{L^\infty(\Omega)})\|_{L^2(\Omega)})\|\tilde{\varphi} - \varphi_h\|_{H_0^1(\Omega)} \tag{5.17}$$

for all $h \in (0, h_3)$, $u \in B_{L^\infty(\Omega)}(\bar{u}, \rho_*)$, and $v \in L^2(\Omega)$.

To prove (5.17), first set $\hat{z} := S'(u)(\tilde{\varphi} - \varphi_h)$. Then (5.8) and the embedding $L^2(\Omega) \hookrightarrow W^{-1,\tilde{p}}(\Omega)$ imply that

$$\|\hat{z}\|_{W_0^{1,\tilde{p}}(\Omega)} \leq C_{\rho_*}\|\tilde{\varphi} - \varphi_h\|_{L^2(\Omega)} \quad \text{for all} h \in (0, h_3), u \in B_{L^\infty(\Omega)}(\bar{u}, \rho_*), v \in L^2(\Omega). \tag{5.18}$$

Consider for any $u \in B_{L^\infty(\Omega)}(\bar{u}, \rho_*)$ and $h \in (0, h_3)$ the bilinear operators $B_u : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ and $B_{u,h} : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ defined via

$$\begin{aligned}
B_u(z, w) &:= \int_\Omega [(b + a(y_u))\nabla z + \mathbb{1}_{\{y_u \notin E_a\}} a'(y_u) z \nabla y_u] \cdot \nabla w \, \mathrm{d}x, \\
B_{u,h}(z, w) &:= \int_\Omega [(b + a(y_h))\nabla z + \mathbb{1}_{\{y_h \notin E_a\}} a'(y_h) z \nabla y_h] \cdot \nabla w \, \mathrm{d}x.
\end{aligned}$$

From this, the definition of $\hat{z}$, and (2.5), we obtain

$$\begin{aligned}
\|\tilde{\varphi} - \varphi_h\|^2_{L^2(\Omega)} &= B_u(\hat{z}, \tilde{\varphi} - \varphi_h) = B_u(\hat{z} - \Pi_h\hat{z}, \tilde{\varphi} - \varphi_h) + B_u(\Pi_h\hat{z}, \tilde{\varphi} - \varphi_h) \\
&= B_u(\hat{z} - \Pi_h\hat{z}, \tilde{\varphi} - \varphi_h) + [B_u(\Pi_h\hat{z}, \tilde{\varphi} - \varphi_h) - B_{u,h}(\Pi_h\hat{z}, \tilde{\varphi} - \varphi_h)],
\end{aligned}$$

where we have used the fact that $B_{u,h}(\Pi_h\hat{z}, \tilde{\varphi} - \varphi_h) = 0$ which follows from combining (5.1) with (5.2). From this and Hölder's inequality, there holds

$$\begin{aligned}
\|\tilde{\varphi} - \varphi_h\|^2_{L^2(\Omega)} &\leq \|b + a(y_u)\|_{L^\infty(\Omega)}\|\hat{z} - \Pi_h\hat{z}\|_{H_0^1(\Omega)}\|\tilde{\varphi} - \varphi_h\|_{H_0^1(\Omega)} \\
&\quad + \|\mathbb{1}_{\{y_u \notin E_a\}} a'(y_u)\nabla y_u\|_{L^\infty(\Omega)}\|\hat{z} - \Pi_h\hat{z}\|_{L^2(\Omega)}\|\tilde{\varphi} - \varphi_h\|_{H_0^1(\Omega)} \\
&\quad + \|a(y_u) - a(y_h)\|_{L^4(\Omega)}\|\Pi_h\hat{z}\|_{W_0^{1,4}(\Omega)}\|\tilde{\varphi} - \varphi_h\|_{H_0^1(\Omega)} \\
&\quad + \|Z_{y_h,y_u}\|_{L^2(\Omega)}\|\Pi_h\hat{z}\|_{L^\infty(\Omega)}\|\tilde{\varphi} - \varphi_h\|_{H_0^1(\Omega)}.
\end{aligned} \tag{5.19}$$

Moreover, we have from Assumption (A3), the continuous embedding $H^{N/4}(\Omega) \hookrightarrow L^4(\Omega)$ with $N = 2$ (see, e.g. [24, Thm. 1.4.4.1]), interpolation theory [6, Thm. 14.2.7], and (4.2) that

$$\begin{aligned}
\|a(y_u) - a(y_h)\|_{L^4(\Omega)} &\leq C\|y_u - y_h\|_{L^4(\Omega)} \leq C\|y_u - y_h\|_{H^{N/4}(\Omega)} \\
&\leq C\|y_u - y_h\|^{1-\frac{2}{4}}_{L^2(\Omega)}\|y_u - y_h\|^{\frac{2}{4}}_{H_0^1(\Omega)} \leq Ch^{\frac{3}{2}}.
\end{aligned} \tag{5.20}$$

We then deduce from this, (5.19), (5.12), the assumptions on b and a, and (5.11) that

$$\begin{aligned}
\|\tilde{\varphi} - \varphi_h\|^2_{L^2(\Omega)} \leq C_{\rho_*}\|\tilde{\varphi} - \varphi_h\|_{H_0^1(\Omega)}\big[&\|\hat{z} - \Pi_h\hat{z}\|_{H_0^1(\Omega)} + h^{\frac{3}{2}}\|\Pi_h\hat{z}\|_{W_0^{1,4}(\Omega)} \\
&+ (h + \|V(y_u, \|y_h - y_u\|_{L^\infty(\Omega)})\|_{L^2(\Omega)})\|\Pi_h\hat{z}\|_{L^\infty(\Omega)}\big].
\end{aligned} \tag{5.21}$$

Moreover, from standard interpolation error estimates [6, 15], we obtain

$$\begin{cases}
\|\hat{z} - \Pi_h\hat{z}\|_{H_0^1(\Omega)} \leq C_1 h^{\frac{\tilde{p}-2}{\tilde{p}}}\|\hat{z}\|_{W_0^{1,\tilde{p}}(\Omega)}, \quad \|\hat{z} - \Pi_h\hat{z}\|_{W_0^{1,4}(\Omega)} \leq C_1 h^{\frac{\tilde{p}-4}{2\tilde{p}}}\|\hat{z}\|_{W_0^{1,\tilde{p}}(\Omega)}, \\
\|\hat{z} - \Pi_h\hat{z}\|_{L^\infty(\Omega)} \leq C_2 h^{1-\frac{2}{\tilde{p}}}\|\hat{z}\|_{W_0^{1,\tilde{p}}(\Omega)}
\end{cases} \tag{5.22}$$

for some constants C_1, C_2 independent of h and $\hat{z}$. This, the triangle inequality, and the embedding $W_0^{1,\tilde{p}}(\Omega) \hookrightarrow L^\infty(\Omega) \cap W_0^{1,4}(\Omega)$ (due to $\tilde{p} \geq 4$) as well as (5.18) give

$$\|\hat{z} - \Pi_h \hat{z}\|_{H_0^1(\Omega)} \leq C_3 h^{\bar{q}} \|\tilde{\varphi} - \varphi_h\|_{L^2(\Omega)}, \quad \|\Pi_h \hat{z}\|_{L^\infty(\Omega)} + \|\Pi_h \hat{z}\|_{W_0^{1,4}(\Omega)} \leq C_4 \|\tilde{\varphi} - \varphi_h\|_{L^2(\Omega)}.$$

Inserting these estimates into (5.21) yields (5.17).

Step 2: Existence of a constant C_{2,h_3,ρ_*} such that

$$\|\tilde{\varphi} - \varphi_h\|_{H_0^1(\Omega)} \leq C_{2,h_3,\rho_*}(h\|v\|_{L^2(\Omega)} + \|\tilde{\varphi} - \varphi_h\|_{L^2(\Omega)}) \tag{5.23}$$

for all $h \in (0, h_3)$, $u \in B_{L^\infty(\Omega)}(\bar{u}, \rho_*)$, and $v \in L^2(\Omega)$.

To show this, we first consider for any $u \in B_{L^\infty(\Omega)}(\bar{u}, \rho_*)$ and $h \in (0, h_3)$, the bilinear mapping $S_{u,h} : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ defined by

$$S_{u,h}(\omega, \psi) := \int_\Omega (b + a(y_h)) \nabla\omega \cdot \nabla\psi \, dx.$$

From Assumptions (A2) and (A3), we obtain

$$\underline{b}\|\tilde{\varphi} - \varphi_h\|^2_{H_0^1(\Omega)} \leq S_{u,h}(\tilde{\varphi} - \varphi_h, \tilde{\varphi} - \varphi_h) = S_{u,h}(\tilde{\varphi} - \varphi_h, \tilde{\varphi} - \Pi_h\tilde{\varphi}) + S_{u,h}(\tilde{\varphi} - \varphi_h, \Pi_h\tilde{\varphi} - \varphi_h). \tag{5.24}$$

Moreover, the Cauchy–Schwarz inequality, the uniform boundedness of $\{y_h\}$ on $C(\overline{\Omega})$, and Assumptions (A2) and (A3) yield that

$$\begin{aligned} S_{u,h}(\tilde{\varphi} - \varphi_h, \tilde{\varphi} - \Pi_h\tilde{\varphi}) &\leq C_{\rho_*}\|\tilde{\varphi} - \varphi_h\|_{H_0^1(\Omega)}\|\tilde{\varphi} - \Pi_h\tilde{\varphi}\|_{H_0^1(\Omega)} \\ &\leq C_{\rho_*} h\|\tilde{\varphi}\|_{H^2(\Omega)}\|\tilde{\varphi} - \varphi_h\|_{H_0^1(\Omega)} \leq C_{\rho_*} h\|v\|_{L^2(\Omega)}\|\tilde{\varphi} - \varphi_h\|_{H_0^1(\Omega)}, \end{aligned} \tag{5.25}$$

where we have exploited the interpolation error [15] and (5.4) in order to obtain the last two estimates. Now using (5.1) and (5.2), we deduce from Hölder's inequality that

$$\begin{aligned} S_{u,h}(\tilde{\varphi} - \varphi_h, \Pi_h\tilde{\varphi} - \varphi_h) &= -\int_\Omega \mathbb{1}_{\{y_h \notin E_a\}} a'(y_h)\nabla y_h \cdot \nabla(\tilde{\varphi} - \varphi_h)(\Pi_h\tilde{\varphi} - \varphi_h)\, dx \\ &\leq \|\mathbb{1}_{\{y_h \notin E_a\}} a'(y_h)\|_{L^\infty(\Omega)}\|\nabla y_h\|_{L^4(\Omega)}\|\tilde{\varphi} - \varphi_h\|_{H_0^1(\Omega)}\|\Pi_h\tilde{\varphi} - \varphi_h\|_{L^4(\Omega)}. \end{aligned}$$

Combing this with the uniform boundedness in $C(\overline{\Omega})$ of $\{y_h\}$ and the embedding $W_0^{1,\tilde{p}}(\Omega) \hookrightarrow W_0^{1,4}(\Omega)$, we obtain that

$$S_{u,h}(\tilde{\varphi} - \varphi_h, \Pi_h\tilde{\varphi} - \varphi_h) \leq C_{\rho_*}(\|\bar{y}\|_{W_0^{1,\tilde{p}}(\Omega)} + \kappa_\rho)\|\tilde{\varphi} - \varphi_h\|_{H_0^1(\Omega)}\|\Pi_h\tilde{\varphi} - \varphi_h\|_{L^4(\Omega)}.$$

The combination of a triangle inequality and [15, Thm. 3.1.6] further implies that

$$\begin{aligned} \|\Pi_h\tilde{\varphi} - \varphi_h\|_{L^4(\Omega)} \leq \|\Pi_h\tilde{\varphi} - \tilde{\varphi}\|_{L^4(\Omega)} + \|\tilde{\varphi} - \varphi_h\|_{L^4(\Omega)} &\leq Ch^{2-\frac{2}{4}}\|\tilde{\varphi}\|_{H^2(\Omega)} + \|\tilde{\varphi} - \varphi_h\|_{L^4(\Omega)} \\ &\leq Ch\|v\|_{L^2(\Omega)} + \|\tilde{\varphi} - \varphi_h\|_{L^4(\Omega)}, \end{aligned}$$

where we have used (5.4) to obtain the last inequality. Similar to (5.20), we find that

$$\|\tilde{\varphi} - \varphi_h\|_{L^4(\Omega)} \leq C\|\tilde{\varphi} - \varphi_h\|^{1-\frac{2}{4}}_{L^2(\Omega)}\|\tilde{\varphi} - \varphi_h\|^{\frac{2}{4}}_{H_0^1(\Omega)}.$$

We then have

$$S_{u,h}(\tilde{\varphi} - \varphi_h, \Pi_h\tilde{\varphi} - \varphi_h) \leq C_{\rho_*}\|\tilde{\varphi} - \varphi_h\|_{H_0^1(\Omega)}\left(h\|v\|_{L^2(\Omega)} + \|\tilde{\varphi} - \varphi_h\|^{\frac{1}{2}}_{L^2(\Omega)}\|\tilde{\varphi} - \varphi_h\|^{\frac{1}{2}}_{H_0^1(\Omega)}\right),$$

which, together with (5.24) and (5.25), yields

$$\|\tilde{\varphi} - \varphi_h\|_{H_0^1(\Omega)} \le C_{\rho_*} \left(h\|v\|_{L^2(\Omega)} + \|\tilde{\varphi} - \varphi_h\|^{\frac{1}{2}}_{L^2(\Omega)} \|\tilde{\varphi} - \varphi_h\|^{\frac{1}{2}}_{H_0^1(\Omega)} \right).$$

Applying the Cauchy–Schwarz inequality then gives (5.23).

Step 3: Existence of constants $\bar{h} \in (0, h_3)$ and $\bar{\rho} \in (0, \rho_*]$.

To show this, we first obtain from the definition of $\Sigma(\bar{y})$ in (2.17) and of Σ_r in (5.10) the existence of a $r_* > 0$ such that $\|\Sigma_r(\bar{y})\|_{L^1(\Omega)} \le \Sigma(\bar{y}) + 1$ for all $r \in (0, r_*)$. This together with (ii) in Proposition 5.3 yields

$$\|V(\bar{y}, r)\|_{L^2(\Omega)} \le C_{\bar{y}} r^{1/2} (\Sigma(\bar{y}) + 1)^{1/2} \quad \text{for all} r \in (0, r_*). \tag{5.26}$$

Now Proposition 5.3 (i), (5.9), Theorem 2.1, and the monotonic growth of $V(y, \cdot)$ imply that

$$\begin{aligned}
&\|V(y_u, \|y_h - y_u\|_{L^\infty(\Omega)})\|_{L^2(\Omega)} \\
&\quad \le \|V(\bar{y}, \|y_h - y_u\|_{L^\infty(\Omega)} + \|y_u - \bar{y}\|_{L^\infty(\Omega)})\|_{L^2(\Omega)} + C_{\bar{y},\rho_*}\|y_u - \bar{y}\|_{H_0^1(\Omega)} \\
&\quad \le \|V(\bar{y}, C_\infty h + C_{\rho_*}\|u - \bar{u}\|_{L^\infty(\Omega)})\|_{L^2(\Omega)} + C_{\bar{y},\rho_*} C_{\rho_*}\|u - \bar{u}\|_{L^\infty(\Omega)}
\end{aligned} \tag{5.27}$$

for some constants C_{ρ_*} and $C_{\bar{y},\rho_*}$ and for all $h \in (0, h_3)$ and $u \in B_{L^\infty(\Omega)}(\bar{u}, \rho_*)$. Next, fix $\bar{h} \in (0, h_3)$ and $\bar{\rho} \le \rho_*$ such that

$$\bar{h}^{\bar{q}} + C_{\bar{y}} [C_\infty \bar{h} + C_{\rho_*}\bar{\rho}]^{\frac{1}{2}} (\Sigma(\bar{y}) + 1)^{\frac{1}{2}} + C_{\bar{y},\rho_*} C_{\rho_*} \bar{\rho} \le \frac{1}{2C_{1,h_3,\rho_*} C_{2,h_3,\rho_*}} \quad \text{and} \quad C_\infty \bar{h} + C_{\rho_*}\bar{\rho} < r_*,$$

where C_{1,h_3,ρ_*} and C_{2,h_3,ρ_*} are defined in (5.17) and (5.23), respectively. From this, (5.26), and (5.27), we conclude that

$$\bar{h}^{\bar{q}} + \|V(y_u, \|y_h - y_u\|_{L^\infty(\Omega)})\|_{L^2(\Omega)} \le \frac{1}{2C_{1,h_3,\rho_*} C_{2,h_3,\rho_*}} \text{ for all} h \in (0, \bar{h}), u \in B_{L^\infty(\Omega)}(\bar{u}, \bar{\rho}).$$

The combination of this with (5.17) and (5.23) yields $\|\tilde{\varphi} - \varphi_h\|_{H_0^1(\Omega)} \le 2C_{2,h_3,\rho_*} h\|v\|_{L^2(\Omega)}$, and together with (5.17), we obtain $\|\tilde{\varphi} - \varphi_h\|_{L^2(\Omega)} \le C\varepsilon_h^u \|v\|_{L^2(\Omega)}$. Combining the last two estimates with Lemma 5.2 and the triangle inequality, we arrive at the desired conclusion. □

Remark 5.6. The convergence rate in Theorem 5.5 is limited by the regularity of the solutions to (2.5), which can only be guaranteed to belong to $W_0^{1,\tilde{p}}(\Omega)$ due to the nondifferentiability of a.

Let us briefly contrast this with the situation in the smooth case: For solutions to (2.5) in $H^2(\Omega)$, the $W_0^{1,\tilde{p}}(\Omega)$ norm in the a priori estimate (5.18) could be replaced by the $H^2(\Omega)$-norm, and the first interpolation error estimate in (5.22) could be replaced by

$$\|\hat{z} - \Pi_h \hat{z}\|_{H_0^1(\Omega)} \le C_1 h \|\hat{z}\|_{H^2(\Omega)}.$$

As a result, the term ε_h^u in (5.16) could be chosen as

$$\varepsilon_h^u := h^2 + h\|V(S(u), \|S_h(u) - S(u)\|_{L^\infty(\Omega)})\|_{L^2(\Omega)}.$$

Furthermore, if a is of class C^2, then the terms σ_i in (2.2) vanish and thus the function $V(\cdot, \cdot)$ defined in (5.10) is identical to zero. In this case, we obtain $\varepsilon_h^u = h^2$, which is the optimal choice for ε_h^u when considering the corresponding optimal control problem; see estimates (3.27) and (4.21) in [12] as well as [11, Thm. 4.5].

Remark 5.7. Let us comment on the discrete counterpart of the linearized state equation (2.5). When the right-hand side of (2.5) belongs to $L^2(\Omega)$, the solutions to (2.5) only belong to $W_0^{1,\tilde{p}}(\Omega)$ instead of $H^2(\Omega)$. Therefore, it is possible to derive an error estimate in $H_0^1(\Omega)$ for the discretized linearized state equation in the same manner as for Theorem 5.5, i.e., that there exist constants $h^* > 0$ and $\rho^* > 0$ such that

$$\|z_v - z_h\|_{H_0^1(\Omega)} \le C_{\bar{u}} \|v\|_{L^2(\Omega)} \left(h^{\frac{\tilde{p}-2}{\tilde{p}}} + \|V(S(u), \|S_h(u) - S(u)\|_{L^\infty(\Omega)})\|_{L^2(\Omega)} \right)$$

for all $h \in (0, h^*)$, $u \in B_{L^\infty(\Omega)}(\bar{u}, \rho^*)$, and $v \in L^2(\Omega)$, where z_v and z_h stand for the solutions to (2.5) and (4.16), respectively. On the other hand, error estimates in $L^2(\Omega)$ for $z_v - z_h$ of order $O(h^q)$ for some $q > \frac{\tilde{p}-2}{\tilde{p}}$ cannot be obtained in this way, since the term $Z_{y,\hat{y}}$ in (4.5) is in general not differentiable in the weak sense even if $y, \hat{y} \in H^2(\Omega)$.

6 DISCRETIZATION OF THE CONTROL PROBLEM

In this section, we show convergence of the discrete optimal solutions to a strict local minimizer $\bar{u}$ of (P) as $h \to 0^+$ as well as error estimates. Recall that $\bar{u}$ is called a *strict local minimizer* of (P) if there exists a constant $\bar{\varepsilon} > 0$ such that

$$j(\bar{u}) < j(u) \quad \text{for all} u \in \overline{B}_{L^\infty(\Omega)}(\bar{u}, \bar{\varepsilon}) \cap \mathcal{U}_{ad}.$$

We can obviously assume that $\bar{\varepsilon} < \rho_*$. For any $h \in (0, h_2)$, the discretized cost functional is then given by

$$j_h : B_{L^2(\Omega)}(\bar{u}, \rho) \to \mathbb{R}, \qquad j_h(u) := \int_\Omega L(x, (S_h(u))(x))\,\mathrm{d}x + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)}.$$

Using Theorem 4.7 and Assumption (A4), we can show differentiability of j_h. The proof of the following result is straightforward and therefore omitted.

Theorem 6.1. *For any $h \in (0, h_2)$, the discrete cost functional $j_h : B_{L^2(\Omega)}(\bar{u}, \rho) \to \mathbb{R}$ is of class C^1, and its derivative at $u \in B_{L^2(\Omega)}(\bar{u}, \rho)$ is given by*

$$j_h'(u)w = \int_\Omega (\varphi_h(u) + \nu u) w \,\mathrm{d}x \qquad \textit{for all} w \in L^2(\Omega),$$

where $\varphi_h(u) \in V_h$ is the unique solution to (5.1) with $v = \frac{\partial L}{\partial y}(\cdot, S_h(u))$.

In the following, we will consider two different discretizations of the control:

(i) *variational discretization*: $\mathcal{U}_h = L^\infty(\Omega)$ (see, e.g., [25]);

(ii) *piecewise constant discretization*:

$$\mathcal{U}_h = \mathcal{U}_h^0 := \left\{ u \in L^\infty(\Omega) \mid u_{|T} \in \mathbb{R} \text{for all} T \in \mathcal{T}_h \right\}.$$

Unless specified, any claim for $\mathcal{U}_h$ should be understood to hold for both cases. For any $h \in (0, h_2)$, set now $\mathcal{U}_{ad,h} := \mathcal{U}_{ad} \cap \mathcal{U}_h$ and consider the discretized optimal control problem defined via

$$(P_h^{\bar{\varepsilon}}) \qquad \min_{u_h \in \mathcal{U}_{ad,h} \cap \overline{B}_{L^\infty(\Omega)}(\bar{u}, \bar{\varepsilon})} j_h(u_h).$$

We now provide a result on the existence of global minimizers and the associated optimality conditions of $(P_h^{\bar{\varepsilon}})$. Its proof is elementary and is thus omitted.

Theorem 6.2. *Let Assumptions (A1) to (A4) hold. Then there exists a constant $h_* \in (0, h_2)$ such that for any $h \in (0, h_*)$, $(P_h^{\bar{\varepsilon}})$ admits at least one global minimizer $\bar{u}_h \in \mathcal{U}_{ad,h} \cap \overline{B}_{L^\infty(\Omega)}(\bar{u}, \bar{\varepsilon})$. Moreover, there exists a function $\bar{\varphi}_h \in V_h$ that together with $\bar{u}_h$ and $\bar{y}_h := S_h(\bar{u}_h)$ satisfies*

$$\int_\Omega (b + a(\bar{y}_h))\nabla\bar{\varphi}_h \cdot \nabla w_h + \mathbb{1}_{\{\bar{y}_h \notin E_a\}} a'(\bar{y}_h) w_h \nabla\bar{y}_h \cdot \nabla\bar{\varphi}_h \,\mathrm{d}x = \int_\Omega \frac{\partial L}{\partial y}(x, \bar{y}_h) w_h \,\mathrm{d}x, \tag{6.1a}$$

$$\int_\Omega (\bar{\varphi}_h + \nu\bar{u}_h)(u_h - \bar{u}_h)\,\mathrm{d}x \ge 0 \quad \textit{for all} w_h \in V_h, u_h \in \mathcal{U}_{ad,h} \cap \overline{B}_{L^\infty(\Omega)}(\bar{u}, \bar{\varepsilon}). \tag{6.1b}$$

6.1 Convergence of discrete minimizers

From now on, let $\bar{u} \in \mathcal{U}_{ad}$ be a strict local minimizer of (P), and for any $h \in (0, h_*)$ let $\bar{u}_h \in \mathcal{U}_{ad,h} \cap \overline{B}_{L^\infty(\Omega)}(\bar{u}, \bar{\varepsilon})$ be any global minimizer of $(P_h^{\bar{\varepsilon}})$ from Theorem 6.2. The goal of this subsection is to show the convergence $\bar{u}_h \to \bar{u}$ as $h \to 0^+$.

We first state a convergence result in $L^2(\Omega)$, whose proof is similar to that of [18, Thm. 4.2] and is thus omitted here.

Theorem 6.3. *Let all assumptions of Theorem 6.2 hold. Then $\|\bar{u}_h - \bar{u}\|_{L^2(\Omega)} \to 0$ as $h \to 0^+$.*

In order to show convergence in $L^\infty(\Omega)$, we first need the following lemma.

Lemma 6.4. *Let $q > 2$ be given and let $\bar{h}$ and $\bar{\rho}$ be defined in Theorem 5.5. If $\Sigma(\bar{y}) < \infty$, then for any $h \in (0, \bar{h})$ and $u, v \in \mathcal{U}_{ad}$ such that $v \in B_{L^\infty(\Omega)}(\bar{u}, \bar{\rho})$, there hold*

$$\|y_u - y_h(v)\|_{L^2(\Omega)} + h\|y_u - y_h(v)\|_{H_0^1(\Omega)} \le C(h^2 + \|u - v\|_{L^2(\Omega)}), \tag{6.2}$$

$$\|\varphi_u - \varphi_v\|_{H^2(\Omega)} \le C\|u - v\|_{L^2(\Omega)}, \tag{6.3}$$

$$\|\varphi_u - \varphi_h(v)\|_{L^2(\Omega)} \le C(\varepsilon_h^v + \|u - v\|_{L^2(\Omega)}), \tag{6.4}$$

$$\|\varphi_u - \varphi_h(v)\|_{H_0^1(\Omega)} \le C(h + \|u - v\|_{L^2(\Omega)}), \tag{6.5}$$

$$\|\varphi_u - \varphi_h(v)\|_{L^\infty(\Omega)} \le C(h^{1-\frac{2}{q}} + \|u - v\|_{L^2(\Omega)}), \tag{6.6}$$

for some constant C independent of u, v, and h. Here $y_u := S(u)$ and $y_h(v) := S_h(v)$, while φ_u is the unique solution to (2.7) for $v = \frac{\partial L}{\partial y}(\cdot, y_u)$ and $\varphi_h(v)$ is the unique solution to (5.1) for $y_h = S_h(v)$ and $v = \frac{\partial L}{\partial y}(\cdot, y_h(v))$.

Proof. First, a standard argument yields (6.2). For the other estimates, let $\varphi_{v,h}$ be the solution to (2.7) for $v = \frac{\partial L}{\partial y}(\cdot, y_h(v))$ and y_u replaced by $y_v := S(v)$. We need to show that

$$\|\varphi_{v,h} - \varphi_u\|_{H^2(\Omega)} \le C(\|u - v\|_{L^2(\Omega)} + h^2) \tag{6.7}$$

for some constant $C > 0$ independent of u, v, and h. To this end, we subtract the equations for φ_u and $\varphi_{v,h}$ to obtain that $\varphi_{v,h} - \varphi_u \in H_0^1(\Omega)$ and

$$-\operatorname{div}[(b + a(y_v))\nabla(\varphi_{v,h} - \varphi_u)] + \mathbb{1}_{\{y_v \notin E_a\}} a'(y_v)\nabla y_v \cdot \nabla(\varphi_{v,h} - \varphi_u) = g_{u,v,h} \tag{6.8}$$

with

$$g_{u,v,h} := -\operatorname{div}[(a(y_u) - a(y_v))\nabla\varphi_u] + Z_{y_u,y_v} \cdot \nabla\varphi_u + \frac{\partial L}{\partial y}(\cdot, y_h(v)) - \frac{\partial L}{\partial y}(\cdot, y_u).$$

Theorem 2.2 and Assumption (A4) imply that $\varphi_u \in H^2(\Omega)$. From this, the product formula, the chain rule [23], and the finiteness of the set E_a, we deduce that

$$\operatorname{div}[(a(y_u) - a(y_v))\nabla\varphi_u] = Z_{y_u,y_v} \cdot \nabla\varphi_u + (a(y_u) - a(y_v))\Delta\varphi_u.$$

This shows that

$$g_{u,v,h} = (a(y_v) - a(y_u))\Delta\varphi_u + \frac{\partial L}{\partial y}(\cdot, y_h(v)) - \frac{\partial L}{\partial y}(\cdot, y_u) \in L^2(\Omega).$$

The standard stability estimate for the solution $\varphi_{v,h} - \varphi_u$ to (6.8) thus gives

$$\|\varphi_{v,h} - \varphi_u\|_{H^2(\Omega)} \le C\|g_{u,v,h}\|_{L^2(\Omega)} \le C[\|y_u - y_v\|_{L^\infty(\Omega)} + \|y_h(v) - y_u\|_{L^2(\Omega)}]$$

for some constant $C > 0$ not depending on u, v, and h, where we have employed the boundedness of $\{S(w) \mid w \in \mathcal{U}_{ad}\}$ in $C(\overline{\Omega})$, the fact that $\|\Delta\varphi_u\|_{L^2(\Omega)} \le C$ due to Theorem 2.2, and Assumptions (A3) and (A4) to derive the last estimate. From this, (6.2), and the fact that $\|y_u - y_v\|_{L^\infty(\Omega)} \le C\|u - v\|_{L^2(\Omega)}$, we obtain (6.7). The estimate (6.3) is shown by a similar argument.

We now prove (6.4)–(6.6). According to the triangle inequality, (6.7), Theorem 5.5, and the boundedness in $L^2(\Omega)$ of $\{\frac{\partial L}{\partial y}(\cdot, y_h(v)) \mid v \in B_{L^\infty(\Omega)}(\bar u, \bar\rho), h \le h_2\}$, we obtain (6.4) and (6.5). Finally, for (6.6), we first see from the continuous embedding $H_0^1(\Omega) \hookrightarrow L^q(\Omega)$, the interpolation error and inverse estimates [6] for $N = 2$ that

$$\begin{aligned}
\|\varphi_{v,h} - \varphi_h(v)\|_{L^\infty(\Omega)} &\le \|\varphi_{v,h} - \Pi_h\varphi_{v,h}\|_{L^\infty(\Omega)} + \|\Pi_h\varphi_{v,h} - \varphi_h(v)\|_{L^\infty(\Omega)} \\
&\le Ch^1\|\varphi_{v,h}\|_{H^2(\Omega)} + Ch^{-\frac{2}{q}}\|\Pi_h\varphi_{v,h} - \varphi_h(v)\|_{L^q(\Omega)} \\
&\le Ch^1\|\varphi_{v,h}\|_{H^2(\Omega)} + Ch^{-\frac{2}{q}}\|\Pi_h\varphi_{v,h} - \varphi_h(v)\|_{H_0^1(\Omega)} \\
&\le C\left[h^1\|\varphi_{v,h}\|_{H^2(\Omega)} + h^{-\frac{2}{q}}\|\Pi_h\varphi_{v,h} - \varphi_{v,h}\|_{H_0^1(\Omega)} + h^{-\frac{2}{q}}\|\varphi_{v,h} - \varphi_h(v)\|_{H_0^1(\Omega)}\right],
\end{aligned}$$

which together with Theorem 5.5 and the interpolation error estimate from [6, Thm. 4.4.20] yields

$$\|\varphi_{v,h} - \varphi_h(v)\|_{L^\infty(\Omega)} \le C\left[h^{1-\frac{2}{q}}\|\varphi_{v,h}\|_{H^2(\Omega)} + h^{1-\frac{2}{q}}\|\frac{\partial L}{\partial y}(\cdot, y_h(v))\|_{L^2(\Omega)}\right].$$

Combining this with the uniform boundedness of $\varphi_{v,h}$ in $H^2(\Omega)$ and of $\frac{\partial L}{\partial y}(\cdot, y_h(v))$ in $L^2(\Omega)$ for all $v \in B_{L^\infty(\Omega)}(\bar u, \bar\rho)$ and $h \le h_2$, we conclude that $\|\varphi_{v,h} - \varphi_h(v)\|_{L^\infty(\Omega)} \le Ch^{1-\frac{2}{q}}$. From this, (6.7), and the embedding $H^2(\Omega) \hookrightarrow L^\infty(\Omega)$, the triangle inequality thus leads to (6.6). □

From Theorem 6.3 and the estimate (6.6) in Lemma 6.4, we obtain the desired convergence result in $L^\infty(\Omega)$. Its proof is similar to that of [11, Thm. 5.3] with some modifications and it is thus omitted.

Theorem 6.5. *Let all assumptions of Theorem 6.3 hold. If $\Sigma(\bar y) < \infty$, then $\|\bar u_h - \bar u\|_{L^\infty(\Omega)} \to 0$ as $h \to 0^+$.*

Remark 6.6. By Theorem 6.5, it holds that $\|\bar u_h - \bar u\|_{L^\infty(\Omega)} \le \frac{\bar\varepsilon}{2}$ for all $h \in (0, \tilde h)$ and for some $\tilde h > 0$. Now for any $h \in (0, \tilde h)$ and $u_h \in \mathcal{U}_{ad,h}$, we have that $w_h := t(u_h - \bar u_h) + \bar u_h \in \overline{B}_{L^\infty(\Omega)}(\bar u_h, \frac{\bar\varepsilon}{2})$ for $t > 0$ small enough and hence that $w_h \in \mathcal{U}_{ad,h} \cap \overline{B}_{L^\infty(\Omega)}(\bar u, \bar\varepsilon)$. The variational inequality (6.1b) then implies that

$$\int_\Omega (\bar\varphi_h + \nu\bar u_h)(u_h - \bar u_h)\,dx \ge 0 \quad \text{for all} u_h \in \mathcal{U}_{ad,h}.$$

6.2 ERROR ESTIMATES FOR DISCRETE MINIMIZERS

We finally turn to error estimates for discrete local minimizers $\bar u_h$ under the second-order sufficient optimality condition (2.19). We need the following technical lemma.

Lemma 6.7. *Let Assumptions (A1) to (A4) hold. Assume further that the level sets $\{\bar y = t_i\}$, $1 \le i \le K$, decompose into finitely many connected components and that $\bar y$ fulfills (3.1) for each connected component of $\{\bar y = t_i\}$. Let $\{v_n\} \subset L^2(\Omega)$, $v \in L^2(\Omega)$, $\bar\varphi \in W^{1,\infty}(\Omega) \cap W^{2,1}(\Omega)$, and $\{s_n\} \in c_0^+$ be arbitrary such that $\|v_n\|_{L^2(\Omega)} = 1$ and $v_n \rightharpoonup v$ in $L^2(\Omega)$. Setting $u_n := \bar u + s_n v_n$, $y_n := S(u_n)$ for $n \ge 1$ and $w := S'(\bar u)v$, then the following assertions hold:*

(i) If $w_n \to w$ in $C(\overline\Omega)$, then

$$\limsup_{n\to\infty} \frac{1}{s_n}\int_\Omega Z_{y_n,\bar y} \cdot \nabla\bar\varphi w_n\,dx = -2Q_2(\bar u, \bar y, \bar\varphi; v) + \int_\Omega [\mathbb{1}_{\{\bar y \notin E_a\}} a''(\bar y)w^2\nabla\bar y + a'(\bar y; w)\nabla w] \cdot \nabla\bar\varphi\,dx.$$

(ii) If in addition (2.9a) is fulfilled, then

$$\liminf_{n\to\infty} \frac{j'(u_n)v_n - j'(\bar u)v_n}{s_n} = 2Q(\bar u, \bar y, \bar\varphi; v) + \nu\left(1 - \|v\|^2_{L^2(\Omega)}\right).$$

Proof. We first observe from [16, Cor. 3.6] that

$$\frac{y_n - \bar y}{s_n} \to w = S'(\bar u)v \text{ in} W^{1,\tilde p}_0(\Omega) \hookrightarrow C(\overline{\Omega}). \tag{6.9}$$

Ad (i): We first deduce for $n \in \mathbb{N}$ large enough that

$$\tau_n := \|y_n - \bar y\|_{C(\overline{\Omega})} \le Cs_n < \delta. \tag{6.10}$$

Moreover, there exists a constant $M > 0$ such that $\|y_n\|_{C(\overline{\Omega})}, \|\bar y\|_{C(\overline{\Omega})} \le M$ for all $n \ge 1$. According to Lemma 4.4, we have

$$Z_{y_n,\bar y} = Z^{(1)}_{y_n,\bar y} + Z^{(2)}_{y_n,\bar y} + Z^{(3)}_{y_n,\bar y} + Z^{(4)}_{y_n,\bar y} \tag{6.11}$$

with $Z^{(k)}_{y_n,\bar y}$, $k = 1, 2, 3, 4$, defined in Lemma 4.4. For $Z^{(1)}_{y_n,\bar y}$, using (6.10) yields

$$Z^{(1)}_{y_n,\bar y} = \sum_{i=0}^{K} [\mathbb{1}_{\{\bar y\in(t_i,t_{i+1})\}} - \mathbb{1}_{\{\bar y\in(t_i,t_{i+1}), y_n\in(t_i-\delta,t_i]\cup[t_{i+1},t_{i+1}+\delta)\}}][(a_i'(y_n) - a_i'(\bar y))\nabla y_n + a_i'(\bar y)\nabla(y_n - \bar y)].$$

Since $\mathbb{1}_{\{\bar y\in(t_i,t_{i+1}), y_n\in(t_i-\delta,t_i]\cup[t_{i+1},t_{i+1}+\delta)\}} \to 0$ a.e. in Ω, we have from (6.9) and the Lebesgue dominated convergence theorem that

$$\begin{aligned}\frac{1}{s_n}\int_\Omega Z^{(1)}_{y_n,\bar y}\cdot\nabla\bar\varphi w_n \,dx &\to \int_\Omega \sum_{i=0}^{K} \mathbb{1}_{\bar y\in(t_i,t_{i+1})}[a_i''(\bar y)w\nabla\bar y + a_i'(\bar y)\nabla w]\cdot\nabla\bar\varphi w\,dx\\ &= \int_\Omega \mathbb{1}_{\{\bar y\notin E_a\}}[a''(\bar y)w^2\nabla\bar y + a'(\bar y)w\nabla w]\cdot\nabla\bar\varphi\,dx.\end{aligned}$$

For $Z^{(2)}_{y_n,\bar y}$, we see from (6.10) and the fact $\nabla\bar y = 0$ a.e. on $\{\bar y = t_i\}$ that

$$Z^{(2)}_{y_n,\bar y} = \sum_{i=1}^{K} \mathbb{1}_{\{\bar y=t_i, y_n\in(t_i-\delta,t_i)\}} a_{i-1}'(y_n)\nabla(y_n - \bar y) + \mathbb{1}_{\{\bar y=t_i, y_n\in(t_i,t_i+\delta)\}} a_i'(y_n)\nabla(y_n - \bar y).$$

Setting $\hat w_n := \frac{y_n-\bar y}{s_n}$ and exploiting (6.9) yields $w_n - \hat w_n \to w - w = 0$ in $W^{1,\tilde p}_0(\Omega) \hookrightarrow C(\overline{\Omega})$. From this, the limit $y_n \to \bar y$ in $C(\overline{\Omega})$, and the continuity of a_{i-1}' and a_i', the dominated convergence theorem implies that

$$\begin{aligned}&\lim_{n\to\infty}\frac{1}{s_n}\int_\Omega Z^{(2)}_{y_n,\bar y}\cdot\nabla\bar\varphi w_n\,dx\\ &\qquad= \lim_{n\to\infty}\sum_{i=1}^{K}\int_\Omega [\mathbb{1}_{\{\bar y=t_i, y_n\in(t_i-\delta,t_i)\}} a_{i-1}'(t_i)\hat w_n + \mathbb{1}_{\{\bar y=t_i, y_n\in(t_i,t_i+\delta)\}} a_i'(t_i)\hat w_n]\nabla\hat w_n\cdot\nabla\bar\varphi\,dx.\end{aligned}$$

As a result of (2.3) and the fact that $\hat w_n < 0$ on $\{\bar y = t_i, y_n \in (t_i - \delta, t_i)\}$, there holds

$$\mathbb{1}_{\{\bar y=t_i, y_n\in(t_i-\delta,t_i)\}} a_{i-1}'(t_i)\hat w_n = \mathbb{1}_{\{\bar y=t_i, y_n\in(t_i-\delta,t_i)\}} a'(t_i; \hat w_n).$$

Similarly, one has $\mathbb{1}_{\{\bar y=t_i, y_n\in(t_i,t_i+\delta)\}} a_i'(t_i)\hat w_n = \mathbb{1}_{\{\bar y=t_i, y_n\in(t_i,t_i+\delta)\}} a'(t_i;\hat w_n)$. We thus have

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{s_n}\int_\Omega Z^{(2)}_{y_n,\bar y}\cdot\nabla\bar\varphi w_n\,\mathrm{d}x &= \lim_{n\to\infty}\sum_{i=1}^K\int_\Omega \mathbb{1}_{\{\bar y=t_i, y_n\in(t_i-\delta,t_i)\cup(t_i,t_i+\delta)\}} a'(t_i;\hat w_n)\nabla\hat w_n\cdot\nabla\bar\varphi\,\mathrm{d}x\\ &= \lim_{n\to\infty}\sum_{i=1}^K\int_\Omega \mathbb{1}_{\{\bar y=t_i\}} a'(t_i;\hat w_n)\nabla\hat w_n\cdot\nabla\bar\varphi\,\mathrm{d}x,\end{aligned}$$

where we have used (6.10) and the fact that $\nabla y_n = \nabla\bar y = 0$ a.e. on $\{y_n = \bar y = t_i\}$ to obtain the last identity. We thus conclude from the continuity of $a'(t_i;\cdot)$ due to [5, Prop. 2.49], (6.9), and the dominated convergence theorem that

$$\lim_{n\to\infty}\frac{1}{s_n}\int_\Omega Z^{(2)}_{y_n,\bar y}\cdot\nabla\bar\varphi w_n\,\mathrm{d}x = \int_\Omega \mathbb{1}_{\{\bar y\in E_a\}} a'(\bar y;w)\nabla w\cdot\nabla\bar\varphi\,\mathrm{d}x. \tag{6.12}$$

For $Z^{(4)}_{y_n,\bar y}$, we have from (4.7) that

$$|Z^{(4)}_{y_n,\bar y}| \leq C_M\left[|y_n-\bar y||\nabla\bar y| + |\nabla(y_n-\bar y)|\right]\sum_{i=0}^K\left(\mathbb{1}_{\Omega^{i,2}_{y_n,\bar y}} + \mathbb{1}_{\Omega^{i,3}_{y_n,\bar y}}\right)\quad \text{a.e. in}\Omega.$$

This, together with the fact that $\mathbb{1}_{\Omega^{i,2}_{y_n,\bar y}} + \mathbb{1}_{\Omega^{i,3}_{y_n,\bar y}} \to 0$ a.e. in Ω as well as (6.9), yields

$$\lim_{n\to\infty}\frac{1}{s_n}\int_\Omega Z^{(4)}_{y_n,\bar y}\cdot\nabla\bar\varphi w_n\,\mathrm{d}x = 0. \tag{6.13}$$

It remains to estimate $Z^{(3)}_{y_n,\bar y}$. To this end, we first deduce from (5.15) and the coarea formula for Lipschitz mappings (see, e.g. [20, Thm. 2, p. 117]) or [1, Sec. 2.7]) that

$$\begin{aligned} s_n^{-1}\left|\int_\Omega \mathbb{1}_{\Omega^{i,2}_{y_n,\bar y}}\nabla\bar y\cdot\nabla\bar\varphi(w_n-\hat w_n)dx\right| &\leq s_n^{-1}\|w_n-\hat w_n\|_{L^\infty(\Omega)}\|\nabla\bar\varphi\|_{L^\infty(\Omega)}\int_\Omega \mathbb{1}_{\{0<\bar y-t_i\leq\tau_n\}}|\nabla\bar y|dx\\ &= s_n^{-1}\|w_n-\hat w_n\|_{L^\infty(\Omega)}\|\nabla\bar\varphi\|_{L^\infty(\Omega)}\int_{t_i}^{t_i+\tau_n}\int_{\{\bar y=t_i\}} d\mathcal{H}^1(x)dt\\ &\leq Cs_n^{-1}\tau_n\|w_n-\hat w_n\|_{L^\infty(\Omega)}\|\nabla\bar\varphi\|_{L^\infty(\Omega)}\to 0,\end{aligned}$$

where we have used (6.10) and the fact that $w_n - \hat w_n \to 0$ in $C(\overline\Omega)$ to obtain the last limit. Similarly, $s_n^{-1}|\int_\Omega \mathbb{1}_{\Omega^{i,3}_{y_n,\bar y}}\nabla\bar y\cdot\nabla\bar\varphi(w_n-\hat w_n)dx| \to 0$. From these limits and the definition of $Z^{(3)}_{y_n,\bar y}$, we deduce that $s_n^{-1}\int_\Omega |Z^{(3)}_{y_n,\bar y}\cdot\nabla\bar\varphi(w_n-\hat w_n)|\,\mathrm{d}x \to 0$ and thus

$$\lim_{n\to\infty}\frac{1}{s_n}\int_\Omega Z^{(3)}_{y_n,\bar y}\cdot\nabla\bar\varphi w_n\,\mathrm{d}x = \lim_{n\to\infty}\frac{1}{s_n}\int_\Omega Z^{(3)}_{y_n,\bar y}\cdot\nabla\bar\varphi\hat w_n\,\mathrm{d}x = \lim_{n\to\infty}\frac{1}{s_n^2}\int_\Omega Z^{(3)}_{y_n,\bar y}\cdot\nabla\bar\varphi(y_n-\bar y)\,\mathrm{d}x \tag{6.14}$$

provided that one of these three limits exists. For $\zeta_i(\bar u,\bar y;s_n,v_n)$ defined in (2.13), this yields

$$\begin{aligned} P_n &:= Z^{(3)}_{y_n,\bar y}(y_n-\bar y) + 2\sum_{i=0}^K\zeta_i(\bar u,\bar y;s_n,v_n)\nabla\bar y\\ &= \sum_{i=1}^K\mathbb{1}_{\Omega^{i,2}_{y_n,\bar y}}\{a'\}^{t_i-0}_{t_i+0}(2t_i-\bar y-y_n)\nabla\bar y - \sum_{i=0}^{K-1}\mathbb{1}_{\Omega^{i,3}_{y_n,\bar y}}\{a'\}^{t_{i+1}-0}_{t_{i+1}+0}(2t_{i+1}-\bar y-y_n)\nabla\bar y\\ &= \sum_{i=1}^K\{a'\}^{t_i-0}_{t_i+0}(2t_i-\bar y-y_n)[\mathbb{1}_{\Omega^{i,2}_{y_n,\bar y}} - \mathbb{1}_{\Omega^{i-1,3}_{y_n,\bar y}}]\nabla\bar y.\end{aligned}$$

By (3.16), we deduce that $\lim_{n\to\infty} \frac{1}{s_n^2}\int_\Omega P_n \cdot \nabla\bar\varphi \,\mathrm{d}x = 0$. Combining this with (6.14) and (2.18), we can conclude that

$$\limsup_{n\to\infty} \frac{1}{s_n}\int_\Omega Z^{(3)}_{y_n,\bar y}\cdot\nabla\bar\varphi w_n \,\mathrm{d}x = \limsup_{n\to\infty}\left[\frac{1}{s_n^2}\int_\Omega P_n\cdot\nabla\bar\varphi\,\mathrm{d}x - \frac{2}{s_n^2}\int_\Omega\sum_{i=0}^K \zeta_i(\bar u,\bar y;s_n,v_n)\nabla\bar y\cdot\nabla\bar\varphi\,\mathrm{d}x\right]$$
$$= -2\tilde Q(\bar u,\bar y,\bar\varphi;\{s_n\},v) = -2Q_2(\bar u,\bar y,\bar\varphi;v),$$

where the last equality has been derived from Corollary 3.9. Together with (6.11)–(6.13), we obtain (i).

Ad (ii): Defining the functional $G : L^\infty(\Omega)\to\mathbb{R}$ via $G(y) := \int_\Omega L(x,y(x))\,\mathrm{d}x$ and employing Assumption (A4), we deduce that G is of class C^2 and that its derivatives are given by

$$G'(y)y_1 := \int_\Omega \frac{\partial L}{\partial y}(x,y(x))y_1(x)\,\mathrm{d}x \quad\text{and}\quad G''(y)y_1y_2 := \int_\Omega\frac{\partial^2 L}{\partial y^2}(x,y(x))y_1(x)y_2(x)\,\mathrm{d}x$$

for all $y, y_1, y_2\in L^\infty(\Omega)$. We see from the chain rule that for any $u,v\in L^2(\Omega)$,

$$j'(u)v = G'(S(u))S'(u)v + \nu\int_\Omega uv\,\mathrm{d}x.$$

This, together with a Taylor expansion and the fact that $\|v_n\|_{L^2(\Omega)} = 1$, yields

$$\begin{aligned}(6.15)\quad \frac{1}{s_n}[j'(u_n)v_n - j'(\bar u)v_n] &= \frac{1}{s_n}[G'(y_n)S'(u_n)v_n - G'(\bar y)S'(\bar u)v_n] + \nu\int_\Omega v_n^2\,\mathrm{d}x\\ &= \frac{1}{s_n}[G'(y_n)-G'(\bar y)]S'(\bar u)v_n + \frac{1}{s_n}G'(y_n)[S'(u_n)v_n - S'(\bar u)v_n] + \nu\\ &= \frac{1}{s_n}\int_0^1 G''(\bar y + s(y_n-\bar y))(y_n-\bar y)S'(\bar u)v_n\,ds + \frac{1}{s_n}G'(\bar y)[S'(u_n)v_n - S'(\bar u)v_n]\\ &\quad + \frac{1}{s_n}[G'(y_n)-G'(\bar y)][S'(u_n)v_n - S'(\bar u)v_n] + \nu.\end{aligned}$$

Obviously, the third term on the right-hand side of (6.15) tends to 0 since $v_n\to v$ in $W^{-1,\tilde p}(\Omega)$, S' is continuous, and G is of class C^2. Moreover, it follows from Assumption (A4), (6.9), and the dominated convergence theorem that the first term on the right-hand side of (6.15) tends to $G''(\bar y)(S'(\bar u)v)^2$. It remains to estimate the limes inferior of the second term on the right-hand side of (6.15). Subtracting the equations for $z_n^{(1)} := S'(u_n)v_n$ and $z_n^{(2)} := S'(\bar u)v_n$, we find that $z_n := z_n^{(1)} - z_n^{(2)}\in H^1_0(\Omega)$ satisfies

$$(6.16)\qquad -\operatorname{div}[(b+a(\bar y))\nabla z_n + \mathbb{1}_{\{\bar y\notin E_a\}}a'(\bar y)\nabla\bar y z_n] = \operatorname{div}[(a(y_n)-a(\bar y))\nabla z_n^{(1)} + Z_{y_n,\bar y}z_n^{(1)}] =: g_n.$$

We then have $z_n = S'(\bar u)g_n$, which together with (2.9a) yields

$$(6.17)\qquad G'(\bar y)[z_n^{(1)} - z_n^{(2)}] = \langle G'(\bar y), z_n\rangle_{H^{-1}(\Omega),H^1_0(\Omega)} = \langle g_n,\bar\varphi\rangle_{H^{-1}(\Omega),H^1_0(\Omega)} = -(B_n + C_n)$$

for $B_n := \int_\Omega (a(y_n)-a(\bar y))\nabla z_n^{(1)}\cdot\nabla\bar\varphi\,\mathrm{d}x$ and $C_n := \int_\Omega Z_{y_n,\bar y}\cdot\nabla\bar\varphi z_n^{(1)}\,\mathrm{d}x$. As a result of Theorem 2.1 and the fact that $v_n\to v$ in $W^{-1,\tilde p}(\Omega)$, there holds $z_n^{(1)}\to S'(\bar u)v$ in $W^{1,\tilde p}_0(\Omega)$. Besides, from (6.9) and [17, Lem. 3.5], we have

$$\frac{a(y_n(x)) - a(\bar y(x))}{s_n}\to a'(\bar y(x);(S'(\bar u)v)(x))\quad\text{for all} x\in\overline\Omega.$$

The dominated convergence theorem thus implies that

$$\frac{1}{s_n}B_n\to\int_\Omega a'(\bar y;S'(\bar u)v)\nabla(S'(\bar u)v)\cdot\nabla\bar\varphi\,\mathrm{d}x.$$

This, along with (6.17) and assertion (i), ensures that

$$\liminf_{n\to\infty} \frac{1}{s_n} G'(\bar y)[S'(u_n)v_n - S'(\bar u)v_n] = 2Q_2(\bar u, \bar y, \bar\varphi; v) - 2\int_\Omega a'(\bar y; S'(\bar u)v)\nabla(S'(\bar u)v)\cdot\nabla\bar\varphi\,dx \\ - \int_\Omega \mathbb{1}_{\{\bar y\notin E_a\}} a''(\bar y)(S'(\bar u)v)^2\nabla\bar y\cdot\nabla\bar\varphi\,dx.$$

Using these limits, (6.15), and the definition of Q in (2.16), we arrive at (ii). □

The following theorem is one of main results of the paper, which extends [12, Thm. 2.14] to the case where the cost functional j is of class C^1 but not necessarily C^2.

Theorem 6.8. *Let all assumptions of Theorem 6.5 hold. Assume that the second-order sufficient condition (2.19) is fulfilled. Assume further that the level sets $\{\bar y = t_i\}$, $1 \leq i \leq K$, decompose into finitely many connected components and $\bar y$ fulfills (3.1) for each connected component of $\{\bar y = t_i\}$. Then there exist constants $C > 0$ and $\hat h \in (0, \min\{\bar h, h_*\})$ such that*

$$\|\bar u_h - \bar u\|^2_{L^2(\Omega)} \leq C[(\varepsilon_h^{\bar u_h})^2 + \|\bar u - u_h\|^2_{L^2(\Omega)} + j'(\bar u)(u_h - \bar u)] \tag{6.18}$$

for all $h \in (0,\hat h)$ and $u_h \in \mathcal{U}_{ad,h} \cap \overline{B}_{L^\infty(\Omega)}(\bar u, \hat\varepsilon)$ with $\hat\varepsilon := \min\{\bar\varepsilon, \bar\rho\}$.

Proof. For simplicity of notation, we set $\varepsilon_h := \varepsilon_h^{\bar u_h}$. We first show that

$$[j'(\bar u_h) - j'(\bar u)](\bar u_h - \bar u) \leq j'(\bar u)(u_h - \bar u) \\ + C[\varepsilon_h\|\bar u_h - \bar u\|_{L^2(\Omega)} + \varepsilon_h\|u_h - \bar u\|_{L^2(\Omega)} + \|u_h - \bar u\|_{L^2(\Omega)}\|\bar u_h - \bar u\|_{L^2(\Omega)}] \tag{6.19}$$

for some constant $C > 0$, for all $u_h \in \mathcal{U}_{ad,h} \cap \overline{B}_{L^\infty(\Omega)}(\bar u, \hat\varepsilon)$ and $h \in (0, \min\{\bar h, h_*\})$. To this end, let us take any $h \in (0, \min\{\bar h, h_*\})$, $u \in \mathcal{U}_{ad} \cap \overline{B}_{L^\infty(\Omega)}(\bar u, \hat\varepsilon)$, and $u_h \in \mathcal{U}_{ad,h} \cap \overline{B}_{L^\infty(\Omega)}(\bar u, \hat\varepsilon)$. We deduce from (2.8), Theorem 6.1, Lemma 6.4, and the Cauchy–Schwarz inequality that

$$\left|[j_h'(u) - j'(u)](u_h - \bar u)\right| = \left|\int_\Omega (\varphi_h(u) - \varphi_u)(u_h - \bar u)\,dx\right| \leq C\varepsilon_h^u\|u_h - \bar u\|_{L^2(\Omega)}. \tag{6.20}$$

Moreover, we deduce from $j_h'(\bar u_h)(\bar u_h - u_h) \leq 0$ and $j'(\bar u)(\bar u - \bar u_h) \leq 0$ that

$$\begin{aligned}
&[j'(\bar u_h) - j'(\bar u)](\bar u_h - \bar u) \\
&= [j_h'(\bar u_h) - j'(\bar u_h)](\bar u - \bar u_h) + [j_h'(\bar u_h) - j'(\bar u)](u_h - \bar u) + j_h'(\bar u_h)(\bar u_h - u_h) + j'(\bar u)(u_h - \bar u_h) \\
&\leq [j_h'(\bar u_h) - j'(\bar u_h)](\bar u - \bar u_h) + [j_h'(\bar u_h) - j'(\bar u)](u_h - \bar u) + j'(\bar u)[(u_h - \bar u) + (\bar u - \bar u_h)] \\
&\leq [j_h'(\bar u_h) - j'(\bar u_h)](\bar u - \bar u_h) + [(j_h'(\bar u_h) - j'(\bar u_h)) + (j'(\bar u_h) - j'(\bar u))](u_h - \bar u) + j'(\bar u)(u_h - \bar u).
\end{aligned} \tag{6.21}$$

Applying (6.20) yields that

$$\begin{cases} |[j_h'(\bar u_h) - j'(\bar u_h)](\bar u - \bar u_h)| \leq C\varepsilon_h\|\bar u - \bar u_h\|_{L^2(\Omega)}, \\ |[j_h'(\bar u_h) - j'(\bar u_h)](u_h - \bar u)| \leq C\varepsilon_h\|\bar u - u_h\|_{L^2(\Omega)}. \end{cases} \tag{6.22}$$

Using (2.8), (6.3), $H^2(\Omega) \hookrightarrow L^2(\Omega)$, and the Cauchy–Schwarz inequality yields

$$|[j'(\bar u_h) - j'(\bar u)](u_h - \bar u)| = \left|\int_\Omega [(\varphi_{\bar u_h} - \bar\varphi) + \nu(\bar u_h - \bar u)](u_h - \bar u)\,dx\right| \leq C\|\bar u_h - \bar u\|_{L^2(\Omega)}\|u_h - \bar u\|_{L^2(\Omega)}. \tag{6.23}$$

From this, (6.21), and (6.22), we derive (6.19).

We now prove the conclusion of the theorem by contradiction. To that purpose, we suppose that there exist $h_n \to 0^+$ and $u_{h_n} \in \mathcal{U}_{ad,h_n} \cap \overline{B}_{L^\infty(\Omega)}(\bar{u}, \hat{\varepsilon})$ such that

$$\|\bar{u}_{h_n} - \bar{u}\|^2_{L^2(\Omega)} > n\left[(\varepsilon_{h_n})^2 + \|\bar{u} - u_{h_n}\|^2_{L^2(\Omega)} + j'(\bar{u})(u_{h_n} - \bar{u})\right] \quad \text{for all} n \geq 1$$

or, equivalently, with $s_n := \|\bar{u}_{h_n} - \bar{u}\|_{L^2(\Omega)}$ that

$$\frac{1}{n} > \frac{(\varepsilon_{h_n})^2}{s_n^2} + \frac{\|\bar{u} - u_{h_n}\|^2_{L^2(\Omega)}}{s_n^2} + \frac{j'(\bar{u})(u_{h_n} - \bar{u})}{s_n^2} \quad \text{for all} n \geq 1. \tag{6.24}$$

By setting $v_n := \frac{\bar{u}_{h_n} - \bar{u}}{s_n}$, we can suppose that

$$s_n \to 0^+, \quad \|v_n\|_{L^2(\Omega)} = 1, \quad \bar{u}_{h_n} = \bar{u} + s_n v_n, \quad v_n \rightharpoonup v \text{ in} L^2(\Omega) \quad \text{for some} v \in L^2(\Omega).$$

We first show that v is an element of the critical cone $C(\mathcal{U}_{ad}; \bar{u})$ defined in (2.10). To this end, we first deduce that $v \geq 0$ a.e. on $\{\bar{u} = \alpha\}$ and $v \leq 0$ a.e. on $\{\bar{u} = \beta\}$. Moreover, since $j'(\bar{u})v_n \geq 0$, there holds $j'(\bar{u})v \geq 0$. On the other hand, from (6.22) and (6.23) for $u_h := \bar{u}_{h_n}$, we obtain that

$$\limsup_{n\to\infty}\left\{[j'(\bar{u}_{h_n}) - j'_{h_n}(\bar{u}_{h_n})]v_n + [j'(\bar{u}) - j'(\bar{u}_{h_n})]v_n\right\} \leq \lim_{n\to\infty} C(\varepsilon_{h_n} + s_n) = 0,$$

which yields

$$\begin{aligned} j'(\bar{u})v = \lim_{n\to\infty} j'(\bar{u})v_n &= \lim_{n\to\infty} [j'_{h_n}(\bar{u}_{h_n})v_n + (j'(\bar{u}_{h_n}) - j'_{h_n}(\bar{u}_{h_n}))v_n + (j'(\bar{u}) - j'(\bar{u}_{h_n}))v_n] \\ &\leq \lim_{n\to\infty} j'_{h_n}(\bar{u}_{h_n})v_n = \lim_{n\to\infty}\frac{1}{s_n}[j'_{h_n}(\bar{u}_{h_n})(u_{h_n} - \bar{u}) + j'_{h_n}(\bar{u}_{h_n})(\bar{u}_{h_n} - u_{h_n})]. \end{aligned}$$

From this and the fact that $j'_{h_n}(\bar{u}_{h_n})(\bar{u}_{h_n} - u_{h_n}) \leq 0$, we obtain

$$j'(\bar{u})v \leq \lim_{n\to\infty}\frac{1}{s_n} j'_{h_n}(\bar{u}_{h_n})(u_{h_n} - \bar{u}) \leq \lim_{n\to\infty} \|\varphi_{h_n}(\bar{u}_{h_n}) + \nu\bar{u}_{h_n}\|_{L^2(\Omega)} \frac{\|u_{h_n} - \bar{u}\|_{L^2(\Omega)}}{s_n} \to 0,$$

where we have used Theorem 6.1 and the Cauchy–Schwarz inequality to derive the last estimate and the boundedness of $\{\|\varphi_{h_n}(\bar{u}_{h_n}) + \nu\bar{u}_{h_n}\|_{L^2(\Omega)}\}$ (due to Lemma 6.4) as well as (6.24) to pass to the limit. There therefore holds that $j'(\bar{u})v = 0$. This and [3, Lem. 4.11] lead to $v(x) = 0$ whenever $\bar{\varphi}(x) + \nu\bar{u}(x) \neq 0$. We thus have $v \in C(\mathcal{U}_{ad}; \bar{u})$.

We now derive a contradiction and thus complete the proof. To this end, we divide (6.19) (with $h := h_n$) by s_n^2 to obtain

$$\frac{1}{s_n}[j'(\bar{u}_{h_n}) - j'(\bar{u})]v_n \leq \frac{j'(\bar{u})(u_{h_n} - \bar{u})}{s_n^2} + C\left(\frac{\varepsilon_{h_n}}{s_n} + \frac{\varepsilon_{h_n}}{s_n}\frac{\|u_{h_n} - \bar{u}\|_{L^2(\Omega)}}{s_n} + \frac{\|u_{h_n} - \bar{u}\|_{L^2(\Omega)}}{s_n}\right).$$

Taking the limes inferior as $n \to \infty$, employing (6.24), and using Lemma 6.7 (ii), we conclude that

$$2Q(\bar{u}, \bar{y}, \bar{\varphi}; v) + \nu\left(1 - \|v\|^2_{L^2(\Omega)}\right) \leq 0. \tag{6.25}$$

Combining this with (2.19) and the fact that $\|v\|_{L^2(\Omega)} \leq \liminf_{n\to\infty} \|v_n\|_{L^2(\Omega)} = 1$, we have $v = 0$. Inserting this into (6.25) leads to $0 < \nu \leq 0$, which is the desired contradiction. □

Theorem 6.9 (variational discretization). *Assume that $\mathcal{U}_h = L^\infty(\Omega)$. Under all assumptions of Theorem 6.8, there exists a constant $C > 0$ such that for any $h \in (0, \hat{h})$,*

$$\|\bar{u}_h - \bar{u}\|_{L^2(\Omega)} \leq C\left(h^{1+\bar{q}} + \|\Sigma_{\kappa_h}(\bar{y})\|^{\frac{1}{2}}_{L^1(\Omega)} h^{\frac{3}{2}}\right) \tag{6.26}$$

with $\kappa_h := C_\infty h + \|S(\bar{u}_h) - \bar{y}\|_{L^\infty(\Omega)}$, and $\bar{q}$ and Σ_{κ_h} defined in (5.16) and (5.10), respectively.

Proof. Choosing $u_h := \bar u$ in (6.18) yields

$$\|\bar u_h - \bar u\|_{L^2(\Omega)} \le C\varepsilon_h^{\bar u_h} \quad \text{for all} h \in (0, \hat h). \tag{6.27}$$

Setting $r_h := \|S_h(\bar u_h) - S(\bar u_h)\|_{L^\infty(\Omega)}$ and using (5.9) yield $r_h \le C_\infty h$. Exploiting (5.16), the Cauchy–Schwarz inequality, Proposition 5.3, and the monotonic growth of $V(\bar y, \cdot)$, there holds

$$\begin{aligned}
(\varepsilon_h^{\bar u_h})^2 &\le 2h^{2+2\bar q} + 2h^2 \|V(S(\bar u_h), r_h)\|^2_{L^2(\Omega)} \\
&\le 2h^{2+2\bar q} + Ch^2 \left[\|V(\bar y, r_h + \|S(\bar u_h) - \bar y\|_{L^\infty(\Omega)})\|^2_{L^2(\Omega)} + \|S(\bar u_h) - \bar y\|^2_{H_0^1(\Omega)} \right] \\
&\le 2h^{2+2\bar q} + Ch^2 \left[\|V(\bar y, C_\infty h + \|S(\bar u_h) - \bar y\|_{L^\infty(\Omega)})\|^2_{L^2(\Omega)} + \|S(\bar u_h) - \bar y\|^2_{H_0^1(\Omega)} \right] \\
&\le C \left[h^{2+2\bar q} + h^2 \|\Sigma_{\kappa_h}(\bar y)\|_{L^1(\Omega)} (C_\infty h + \|S(\bar u_h) - \bar y\|_{L^\infty(\Omega)}) + h^2 \|S(\bar u_h) - \bar y\|^2_{H_0^1(\Omega)} \right] \\
&\le C \left[h^{2+2\bar q} + h^2 \|\Sigma_{\kappa_h}(\bar y)\|_{L^1(\Omega)} (h + \|\bar u_h - \bar u\|_{L^2(\Omega)}) + h^2 \|\bar u_h - \bar u\|^2_{L^2(\Omega)} \right],
\end{aligned}$$

where we have used Theorem 2.1 to derive the last inequality. From this and (6.27), a simple computation gives (6.26). ◻

From the fact that $1 + \bar q = 1 + \frac{\tilde p - 2}{\tilde p} \to 2$ as $\tilde p \to \infty$ and that $\|\Sigma_{\kappa_h}(\bar y)\|_{L^1(\Omega)} \to \Sigma(\bar y)$ as $h \to 0^+$, we have the following corollary of Theorem 6.9.

Corollary 6.10. *Assume that all assumptions of Theorem 6.9 are fulfilled. Then the following assertions hold for $h \to 0^+$:*

(i) if $\sigma_i = 0$ for all $1 \le i \le K$, then $\|\bar u_h - \bar u\|_{L^2(\Omega)} = O(h^{2-\varepsilon})$ for any $\varepsilon > 0$;

(ii) if $\Sigma(\bar y) = 0$, then $\|\bar u_h - \bar u\|_{L^2(\Omega)} = o(h^{\frac{3}{2}})$;

(iii) if $\Sigma(\bar y) > 0$, then $\|\bar u_h - \bar u\|_{L^2(\Omega)} = O(h^{\frac{3}{2}})$.

Similarly, we obtain from Theorem 6.8 error estimates for piecewise constant controls.

Theorem 6.11 (piecewise constant controls). *Assume that $\mathcal{U}_h = \mathcal{U}_h^0$. Under all assumptions of Theorem 6.8, there exist constants $C > 0$ and $\hat h_* \in (0, \hat h)$ such that*

$$\|\bar u_h - \bar u\|_{L^2(\Omega)} \le Ch \quad \textit{for all} h \in (0, \hat h_*). \tag{6.28}$$

Proof. According to Theorem 2.2 and (2.9b), $\bar\varphi$ and $\bar u$ are Lipschitz continuous on $\overline{\Omega}$. Hence constants $C_1 > 0$ and $\hat h_* \in (0, \hat h)$ exist such that for any $h \in (0, \hat h_*)$, there exists a $u_h \in \mathcal{U}_{ad,h}$ satisfying $\|\bar u - u_h\|_{L^\infty(\Omega)} \le C_1 h$ and $j'(\bar u)(\bar u - u_h) = 0$; see, e.g. [9, Lem. 4.17]. Combining this with (6.18) and the fact that $\varepsilon_h^{\bar u_h} \le C_2 h$ for all $h \in (0, \hat h)$ yields (6.28). ◻

7 CONCLUSIONS

We have studied the numerical approximation of an optimal control problem governed by a class of quasilinear elliptic equations with non-smooth coefficients in the divergence part. The convergence of a sequence of local minimizers of some discrete control problems to a strict local minimizer of the original problem is shown. A priori error estimates for variational and piecewise constant discretizations are derived under a second-order sufficient condition for the optimal control problem and an assumption on the optimal state which is used to establish an explicit formula of the curvature of the cost functional.

ACKNOWLEDGMENTS

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". Part of this work was completed during a visit of the second author to the Vietnam Institute for Advanced Study in Mathematics (VIASM). The second author would like to thank the VIASM for their financial support and hospitality.

REFERENCES

[1] G. Alberti, S. Bianchini, and G. Crippa, Structure of level sets and Sard-type properties of Lipschitz maps, *Ann. Sc. Norm. Super. Pisa Cl. Sci.(5)* XII (2013), 863–902, doi:10.2422/2036-2145.201107_006.

[2] N. Arada, E. Casas, and F. Tröltzsch, Error estimates for the numerical approximation of a semilinear elliptic control problem, *Comput. Optim. Appl.* 23 (2002), 201–229, doi:10.1023/a:1020576801966.

[3] 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.

[4] A. Bejan, *Convection Heat Transfer*, J. Wiley & Sons, 4 edition, 2013, doi:10.1002/9781118671627.

[5] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer-Verlag, Berlin, Heidelberg, 2000, doi:10.1007/978-1-4612-1394-9.

[6] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Texts in Applied Mathematics, Springer-Verlag, New York, 2008, doi:10.1007/978-0-387-75934-0.

[7] E. Casas and V. Dhamo, 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.

[8] E. Casas and V. Dhamo, Error estimates for the numerical approximation of Neumann control problems governed by a class of quasilinear elliptic equations, *Comput. Optim. Appl.* 52 (2012), 719–756, doi:10.1007/s10589-011-9440-0.

[9] E. Casas, M. Mateos, and J. P. Raymond, Error estimates for the numerical approximation of a distributed control problem for the steady-state Navier-Stokes equations, *SIAM J. Control Optim.* 46 (2007), 952–982, doi:10.1137/060649999.

[10] E. Casas, M. Mateos, and F. Tröltzsch, Error estimates for the numerical approximation of boundary semilinear elliptic control problems, *Comput. Optim. Appl.* 31 (2005), 193–219, doi:10.1007/s10589-005-2180-2.

[11] E. Casas and F. Tröltzsch, Numerical analysis of some optimal control problems governed by a class of quasilinear elliptic equations, *ESAIM:COCV* 17 (2011), 771–800, doi:10.1051/cocv/2010025.

[12] E. Casas and F. Tröltzsch, A general theorem on error estimates with application to a quasilinear elliptic optimal control problem, *Comput. Optim. Appl.* 53 (2012), 173–206, doi:10.1007/s10589-011-9453-8.

[13] M. Chipot, *Elliptic Equations: An Introductory Course*, Birkhäuser Verlag, Basel, 2009, doi:10.1007/978-3-7643-9982-5.

[14] C. Christof and C. Meyer, A note on a priori L^p-error estimates for the obstacle problem, *Numerische Mathematik* 139 (2018), 27–45, doi:10.1007/s00211-017-0931-5.

[15] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Classics in Applied Mathematics, SIAM, 2002, doi:10.1137/1.9780898719208.

[16] C. Clason, V. H. Nhu, and A. Rösch, No-gap second-order optimality conditions for optimal control of a non-smooth quasilinear elliptic equation, *ESAIM: COCV* 27 (2021), 62, doi:10.1051/cocv/2020092.

[17] C. Clason, V. H. Nhu, and A. Rösch, Optimal control of a non-smooth quasilinear elliptic equation, *Mathematical Control and Related Fields* 11 (2021), 521–554, doi:10.3934/mcrf.2020052.

[18] J. C. De Los Reyes and V. Dhamo, Error estimates for optimal control problems of a class of quasilinear equations arising in variable viscosity fluid flow, *Numer. Math.* 132 (2016), 691–720, doi:10.1007/s00211-015-0737-2.

[19] K. Deckelnick and M. Hinze, Convergence of a finite element approximation to a state-constrained elliptic control problem, *SIAM J. Numer. Anal.* 45 (2007), 1937–1953, doi:10.1137/060652361.

[20] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Function*, CRC Press, New York, 4th edition, 1992, doi:10.1201/b18333.

[21] R. S. Falk, Approximation of a class of optimal control problems with order of convergence estimates, *J. Math. Anal. Appl.* 44 (1973), 28–47, doi:10.1016/0022-247x(73)90022-x.

[22] T. Geveci, On the approximation of the solution of an optimal control problem governed by an elliptic equation, *RAIRO. Anal. Numér.* 13 (1979), 313–328, doi:10.1051/m2an/1979130403131.

[23] 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.

[24] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman Advanced Pub. Program, 1985, doi:10.1137/1.9781611972030.

[25] M. Hinze, A variational discretization concept in control constrained optimization: The linear-quadratic case, *Comput. Optim. Appl.* 30 (2005), 45–61, doi:10.1007/s10589-005-4559-5.

[26] M. Hinze and F. Tröltzsch, Discrete concepts versus error analysis in PDE-constrained optimization, *GAMM-Mitteilungen* 33 (2010), 148–162, doi:10.1002/gamm.201010012.

[27] M. Holtmannspötter and A. Rösch, A priori error estimates for the finite element approximation of a nonsmooth optimal control problem governed by a coupled semilinear PDE-ODE system, *SIAM J. Control Optim.* 59 (2021), 3329–3358, doi:10.1137/20m1331263.

[28] C. Meyer and O. Thoma, A priori finite element error analysis for optimal control of the obstacle problem, *SIAM J. Numer. Anal.* 51 (2013), 605–628, doi:10.1137/110836092.

[29] I. Neitzel, J. Pfefferer, and A. Rösch, Finite element discretization of state-constrained elliptic optimal control problems with semilinear state equation, *SIAM J. Control Optim.* 53 (2015), 874–904, doi:10.1137/140960645.

[30] V. H. Nhu, On the no-gap second-order optimality conditions for a non-smooth semilinear elliptic optimal control, *Optimization* (2021), 1–31, doi:10.1080/02331934.2021.1939698.

[31] Y. B. Zel'dovich and Y. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena*, Academic Press, 1966, doi:10.1115/1.3607836.