Semismooth Newton Method for Boundary Bilinear Control

Eduardo Casas, Konstantinos Chrysafinos, and Mariano Mateos

***Abstract*—We study a control-constrained optimal control problem governed by a semilinear elliptic equation. The control acts in a bilinear way on the boundary, and can be interpreted as a heat transfer coefficient. A detailed study of the state equation is performed and differentiability properties of the control-to-state mapping are shown. First and second order optimality conditions are derived. Our main result is the proof of superlinear convergence of the semismooth Newton method to local solutions satisfying no-gap second order sufficient optimality conditions as well as a strict complementarity condition.**

***Index Terms*—Optimal control, bilinear control, semismooth Newton method, convergence analysis**

I. Introduction

IN this letter, we propose a semismooth Newton method to solve the following bilinear optimal control problem:

$$\text{(P)} \min_{u\in U_{\rm ad}} J(u) := \int_\Omega L(x, y_u(x))\,\mathrm{d}x + \frac{\nu}{2}\int_\Gamma u^2(x)\,\mathrm{d}x,$$

where y_u is the state associated with the control u solution of

$$\begin{cases} Ay + a(x,y) = 0 & \text{in } \Omega, \\ \partial_{n_A} y + uy = g & \text{on } \Gamma. \end{cases} \tag{1}$$

Here $\Omega \subset \mathbb{R}^d$, $d = 2$ or 3, is a bounded open connected set with a Lipschitz boundary Γ, $\nu > 0$ and

$$U_{\rm ad} = \{u \in L^2(\Gamma) : \alpha \le u(x) \le \beta \text{ a.e. in } \Gamma\},$$

with $0 \le \alpha < \beta < \infty$. The remaining assumptions regarding the data of the control problem will be given in Sections II and III. Typical examples would include the tracking type functional $L(x,y) = \frac{1}{2}(y - y_d(x))^2$ for some target state y_d and nonlinearities such as $a(x,y) = y^3$ or $a(x,y) = \exp(y)$.

Bilinear control plays an important role not only for the purposes of parameter identification, but also as ways of changing the intrinsic properties of the controlled system. Applications of bilinear control to very distinct fields such as nuclear and thermal control processes, ecologic and physiologic control or socioeconomic systems can be found in the early reference [10], where they are investigated in the framework of ordinary differential equations. In the recent paper [16], the author underlines the importance of bilinear boundary control of partial differential equations in several applications, providing references for them. The goal of that paper is not the analysis of an optimization algorithm, but the obtention of error estimates for the finite element approximation of (P), assuming that the state equation is linear.

Our main goal is to analyze the convergence of the semismooth Newton method applied to (P). The novelty of this paper is twofold. First, the convergence analysis is carried out under the assumptions of no-gap second order optimality conditions and a strict complementarity condition, which are the usual ones to study numerical optimization algorithms in finite dimensional constrained optimization problems; see e.g. [12]. This improves the previous results [1], [9], [13] for distributed controls and [7], [8] for boundary controls, where conditions leading to local convexity were assumed. Second, as far as we know, there are no results in this direction for boundary bilinear controls. In [3] we considered a problem with distributed control acting as a source in the equation; in [2] we turned our attention to a bilinear control problem where the control appears as a reaction coefficient in the partial differential equation. In the paper at hand, the control appears as the Robin coefficient on the boundary condition and a new difficulty appears: the control-to-state mapping is not differentiable $L^2(\Gamma)$ if $d = 3$. In this paper, we focus on the aspects of the proofs that are essentially different from those in [2] and [3], and refer to those papers when necessary.

II. State equation

Let us state the assumptions associated to the state equation.

Assumption 2.1: The operator A is defined in Ω by

$$Ay = -\sum_{i,j=1}^d \partial_{x_j}[a_{ij}(x)\partial_{x_i} y] + a_0 y.$$

We suppose that $a_0, a_{ij} \in L^\infty(\Omega)$ for $1 \le i,j \le d$ with $0 \le$

Submitted on September 8, 2023. The first and third authors were supported by MCIN/ AEI/10.13039/501100011033/ under research project PID2020-114837GB-I00. The second author was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the "First Call for H.F.R.I. Research Projects to support Faculty members and Researchers and the procurement of high-cost research equipment grant" (Project Number: 3270)

Eduardo Casas is with the Departamento de Matemática Aplicada y Ciencias de la Computación, E.T.S.I. Industriales y de Telecomunicación, Universidad de Cantabria, Av. Los Castros s/n, 39005 Santander, Spain (email eduardo.casas@unican.es).

Konstantinos Chrysafinos is with the Department of Mathematics, School of Applied Mathematics and Physical Sciences, National Technical University of Athens, Zografou Campus, Athens 15780, Greece and IACM, FORTH, 20013 Heraklion, Crete, Greece (email chrysafinos@math.ntua.gr).

Mariano Mateos is with the Departamento de Matemáticas, Campus de Gijón, Universidad de Oviedo, 33203, Gijón, Spain. (email mmateos@ uniovi.es).

$a_0 \not\equiv 0$, and there exist $0 < \tilde{\lambda}_A \le \tilde{\Lambda}_A < \infty$ satisfying

$$\tilde{\lambda}_A|\xi|^2 \le \sum_{i,j=1}^{d} a_{ij}(x)\xi_i\xi_j \le \tilde{\Lambda}_A|\xi|^2 \text{ for a.a. } x\in\Omega \text{ and } \forall\xi\in\mathbb{R}^d.$$

Notice that Assumption 2.1 implies the existence of $0 < \lambda_A < \Lambda_A$ such that the bilinear form

$$\mathfrak{a}(y,z) = \int_\Omega \left(\sum_{i,j=1}^{d} a_{ij}\partial_{x_i}y\partial_{x_j}z + a_0yz \right) \mathrm{d}x$$

satisfies

$$\mathfrak{a}(y,y) \ge \lambda_A\|y\|^2_{H^1(\Omega)} \qquad \forall y \in H^1(\Omega), \tag{2}$$

$$\mathfrak{a}(y,z) \le \Lambda_A\|y\|_{H^1(\Omega)}\|z\|_{H^1(\Omega)} \qquad \forall y,z \in H^1(\Omega). \tag{3}$$

Assumption 2.2: We assume that $a : \Omega \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the second variable satisfying for a.a. $x \in \Omega$:

- $a(\cdot,0) \in L^p(\Omega)$ for some $p > d/2$,
- $\dfrac{\partial a}{\partial y}(x,y) \ge 0 \ \forall y \in \mathbb{R}$,
- $\forall M{>}0 \ \exists C_{a,M}$ s.t. $\displaystyle\sum_{j=1}^{2}\left|\frac{\partial^j a}{\partial y^j}(x,y)\right| \le C_{a,M} \ \forall |y| \le M$,
- $\forall\varepsilon{>}0$ and $\forall M{>}0 \ \exists\rho{>}0$ s. t. $\left|\dfrac{\partial^2 a}{\partial y^2}(x,y_1) - \dfrac{\partial^2 a}{\partial y^2}(x,y_2)\right| \le \varepsilon$

for all $|y_1|,|y_2| \le M$ with $|y_1 - y_2| \le \rho$.

All the above constants are independent of x.

We suppose that $g \in L^q(\Gamma)$ with $q > d-1$ and, without loss of generality, that $q \le d$.

To deal with the nonlinearity of the state equation, we observe that $q = 2$ is not enough in dimension $d = 3$. The proof of the differentiability of the relation control-to-state requires $q > 2$. For linear state equations, $q = 2$ is enough; see [16].

For $d = 2$ or 3 it is known that $H^{1/2}(\Gamma) \subset L^4(\Gamma)$ and there exists C_Γ such that

$$\|y\|_{L^4(\Gamma)} \le C_\Gamma\|y\|_{H^1(\Omega)}, \quad \forall y \in H^1(\Omega). \tag{4}$$

Throughout this paper the following notation will be used: we fix $s = 2$ if $d = 2$ or $s = q$ if $d = 3$ and define the set

$$\mathcal{A}_0 := \{u \in L^s(\Gamma) : u \ge 0\}. \tag{5}$$

We denote $B_r(\bar{u}) = \{u \in L^s(\Gamma) : \ \|u - \bar{u}\|_{L^s(\Gamma)} < r\}$.

Theorem 2.3: There exists $\mu > 0$ such that for every $u \in \mathcal{A}_0$ equation (1) has a unique solution $y_u \in Y := H^1(\Omega) \cap C^{0,\mu}(\bar{\Omega})$. Furthermore, the following estimates hold:

$$\|y_u\|_{H^1(\Omega)} \le C\left(\|a(\cdot,0)\|_{L^p(\Omega)} + \|g\|_{L^q(\Gamma)}\right), \tag{6}$$

$$\|y_u\|_{L^\infty(\Omega)} \le M_\infty(\|a(\cdot,0)\|_{L^p(\Omega)} + \|g\|_{L^q(\Gamma)}), \tag{7}$$

$$\|y_u\|_{C^{0,\mu}(\bar{\Omega})} \le C_{\mu,\infty}(\|a(\cdot,0)\|_{L^p(\Omega)} + \|u\|_{L^s(\Gamma)} + \|g\|_{L^q(\Gamma)}), \tag{8}$$

where C, M_∞ and $C_{\mu,\infty}$ are independent of u.

Proof: We define the mapping

$$b : \Omega \times \mathbb{R} \longrightarrow \mathbb{R}, \ b(x,y) := a(x,y) - a(x,0).$$

Assumption 2.2 implies that $b(x,0) = 0$ and $\frac{\partial b}{\partial y}(x,y) \ge 0$. Equation (1) can be written in the variational form

$$\mathfrak{a}(y,z) + \int_\Omega b(x,y)z\,\mathrm{d}x + \int_\Gamma uyz\,\mathrm{d}x = \int_\Omega -a(x,0)z\,\mathrm{d}x + \int_\Gamma gz\,\mathrm{d}x \quad \forall z \in H^1(\Omega). \tag{9}$$

Using (3), Cauchy's inequality, (4), (2) and the nonnegativity of u imposed in (5), we infer that

$$\mathfrak{a}(y,z) + \int_\Gamma uyz\,\mathrm{d}x \le \Lambda_u\|y\|_{H^1(\Omega)}\|z\|_{H^1(\Omega)}, \tag{10}$$

$$\mathfrak{a}(y,y) + \int_\Gamma uy^2\,\mathrm{d}x \ge \lambda_A\|y\|^2_{H^1(\Omega)}, \tag{11}$$

where $\Lambda_u{=}\Lambda_A + \|u\|_{L^2(\Gamma)}C_\Gamma^2$. The proof of existence and uniqueness of a solution in $H^1(\Omega) \cap L^\infty(\Omega)$ of (9) as well as estimates (6) and (7) follow as in [4, Theorem 3.1]. The $L^\infty(\Omega)$ estimate is obtained following the approach of [14, Theorem 4.1] and using that $u \ge 0$ and $b(x,s)s \ge 0 \ \forall s \in \mathbb{R}$.

To prove (8) we write (1) in the form

$$\begin{cases} Ay = -a(x,y) & \text{in } \Omega, \\ \partial_{n_A}y = -uy + g & \text{on } \Gamma. \end{cases}$$

From Assumption 2.2 and the mean value theorem we infer

$$|a(x,y)| \le |a(x,0)| + C_{a,K}K,$$

where $K = \|y\|_{L^\infty(\Omega)}$. In addition, we have $\|-uy\|_{L^s(\Gamma)} \le K\|u\|_{L^s(\Gamma)}$. Then, from [11, Proposition 3.6] we infer that y belongs to $C^{0,\mu}(\bar{\Omega})$ and satisfies (8) for some $\mu \in (0,1]$. ∎

Next we consider the differentiability of the mapping $u \to y_u$.

Theorem 2.4: There exists an open set $\mathcal{A}$ in $L^s(\Gamma)$ such that $\mathcal{A}_0 \subset \mathcal{A}$ and equation (1) has a unique solution $y_u \in Y$ $\forall u \in \mathcal{A}$. Further, the mapping $G : \mathcal{A} \longrightarrow Y$ defined by $G(u) := y_u$ is of class C^2 and $\forall u \in \mathcal{A}$ and $\forall v, v_1, v_2 \in L^s(\Gamma)$ the functions $z = G'(u)v$ and $w = G''(u)(v_1,v_2)$ are the unique solutions of the equations:

$$\begin{cases} Az + \dfrac{\partial a}{\partial y}(x,y_u)z = 0 & \text{in } \Omega, \\ \partial_{n_A}z + uz = -vy_u & \text{on } \Gamma, \end{cases} \tag{12}$$

$$\begin{cases} Aw + \dfrac{\partial a}{\partial y}(x,y_u)w = 0 & \text{in } \Omega, \\ \partial_{n_A}w + uw = -v_1z_{u,v_2} - v_2z_{u,v_1} & \text{on } \Gamma, \end{cases} \tag{13}$$

where $z_{u,v_i} = G'(u)v_i$, $i = 1,2$.

Proof: We consider the space

$$Y_A := \{y \in Y : Ay \in L^p(\Omega), \ \partial_{n_A}y \in L^q(\Gamma)\}$$

endowed with the graph norm. We note that Y_A is a Banach space. We also define the mapping $\mathcal{F} : L^s(\Gamma) \times Y_A \longrightarrow L^p(\Omega) \times L^q(\Gamma)$ by

$$\mathcal{F}(u,y) := (Ay + a(\cdot,y), \partial_{n_A}y + uy - g).$$

Since $q \le s$, $\mathcal{F}$ is well defined and of class C^2 due to Assumption 2.2. For every $(\bar{u},\bar{y}) \in \mathcal{A}_0 \times Y_A$ the derivative $\frac{\partial \mathcal{F}}{\partial y}(\bar{u},\bar{y}) : Y_A \longrightarrow L^p(\Omega) \times L^q(\Gamma)$, given by

$$\frac{\partial \mathcal{F}}{\partial y}(\bar{u},\bar{y})z = \left(Az + \frac{\partial a}{\partial y}(\cdot,\bar{y})z, \partial_{n_A}z + \bar{u}z\right) \ \forall z \in Y_A,$$

is linear and continuous. The open mapping theorem implies that $\frac{\partial \mathcal{F}}{\partial y}(\bar{u},\bar{y})$ is an isomorphism if and only if the equation,

$$\begin{cases} Az + \dfrac{\partial a}{\partial y}(x,\bar{y})z = f & \text{in } \Omega, \\ \partial_{n_A} z + \bar{u}z = h & \text{on } \Gamma, \end{cases}$$

has unique solution $z \in Y_A$ for all $(f,h) \in L^p(\Omega) \times L^q(\Gamma)$. This fact follows from Theorem 2.3. Observing that $\mathcal{F}(u,y_u) = 0$ for all $u \in \mathcal{A}_0$ and taking $\bar{y} = y_{\bar{u}}$, the implicit function theorem implies the existence of $\varepsilon_{\bar{u}} > 0$ and $\varepsilon_{\bar{y}} > 0$ such that $\forall u \in B_{\varepsilon_{\bar{u}}}(\bar{u})$ the equation $\mathcal{F}(u,y) = 0$ has a unique solution y_u in the open ball $B_{\varepsilon_{\bar{y}}}(\bar{y}) \subset Y_A \subset Y$. Moreover, the mapping $u \in B_{\varepsilon_{\bar{u}}}(\bar{u}) \to y_u \in B_{\varepsilon_{\bar{y}}}(\bar{y})$ is of class C^2. Without loss of generality, we assume $\varepsilon_{\bar{u}} < \frac{1}{2}\lambda_A/(|\Gamma|^{\frac{s-2}{s}} C_\Gamma^2)$. Actually, for every $u \in B_{\varepsilon_{\bar{u}}}$ the equation $\mathcal{F}(u,y) = 0$ has unique solution $y \in Y$. Indeed, let y_1, y_2 denote two solutions of $\mathcal{F}(u,y) = 0$. We set $y = y_2 - y_1$, subtract the corresponding equations, and apply the mean value theorem to deduce that y satisfies

$$\begin{cases} Ay + \dfrac{\partial a}{\partial y}(x, y_1 + \theta_x y)y = 0 & \text{in } \Omega, \\ \partial_{n_A} y + uy = 0 & \text{on } \Gamma, \end{cases} \tag{14}$$

where $\theta_x : \Omega \to [0,1]$ is a measurable function. Adding and subtracting appropriate terms on the boundary, equation (14) can be written as

$$\begin{cases} Ay + \dfrac{\partial a}{\partial y}(x, y_1 + \theta_x y)y = 0 & \text{in } \Omega, \\ \partial_{n_A} y + \bar{u}y = -(u - \bar{u})y & \text{on } \Gamma. \end{cases} \tag{15}$$

Testing the variational form of (15) with y we get

$$\lambda_A \|y\|^2_{H^1(\Omega)} \le \varepsilon_{\bar{u}} |\Gamma|^{\frac{s-2}{s}} C_\Gamma^2 \|y\|^2_{H^1(\Omega)}.$$

Since $\varepsilon_{\bar{u}} < \frac{1}{2}\lambda_A/(|\Gamma|^{\frac{s-2}{s}} C_\Gamma^2)$, $y = 0$ holds. Defining in $L^s(\Gamma)$ the open set $\mathcal{A} = \cup_{\bar{u}\in\mathcal{A}_0} B_{\varepsilon_{\bar{u}}}(\bar{u})$ and $G : \mathcal{A} \longrightarrow Y$ such that $G(u) = y_u$, we have that G is of class C^2. Finally, equations (12) and (13) are obtained differentiating with respect to u the identity $\mathcal{F}(u, G(u)) = 0$. ■

Remark 2.5: Theorems 2.3 and 2.4 are valid if we use the operator A^* instead of A, where $A^*\varphi = -\sum_{i,j=1}^d \partial_{x_j}[a_{ji}(x)\partial_{x_i}\varphi] + a_0\varphi$. Therefore, for every $\bar{u} \in \mathcal{A}_0$ we obtain the existence of $\varepsilon^*_{\bar{u}} > 0$ such that, for every $(f,h) \in L^p(\Omega) \times L^q(\Gamma)$ and $u \in B_{\varepsilon^*_{\bar{u}}}(\bar{u})$, the equation

$$\begin{cases} A^*\varphi + \dfrac{\partial a}{\partial y}(x, y_u)\varphi = f & \text{in } \Omega, \\ \partial_{n_{A^*}} \varphi + u\varphi = h & \text{on } \Gamma, \end{cases}$$

has a unique solution $\varphi \in Y$. Without loss of generality, we can assume that $\varepsilon_{\bar{u}} \le \varepsilon^*_{\bar{u}}$, so the equation is uniquely solvable in Y for all $u \in \mathcal{A}$.

III. Analysis of the optimal control problem

In this section we proceed to the analysis of the optimal control problem. To this end we make the following hypotheses on J.

Assumption 3.1: The function $L : \Omega \times \mathbb{R} \longrightarrow \mathbb{R}$ is Carathéodory and of class C^2 with respect to the second variable. Further the following properties hold for a.a. $x \in \Omega$:

- $L(\cdot, 0) \in L^1(\Omega)$,
- $\forall M > 0,\ \exists L_M \in L^p(\Omega)$ such that $\Big|\dfrac{\partial L}{\partial y}(x,y)\Big| \le L_M(x)$,
- $\forall M > 0,\ \exists C_{L,M} \in \mathbb{R}$ such that $\Big|\dfrac{\partial^2 L}{\partial y^2}(x,y)\Big| \le C_{L,M}$,
- $\forall \varepsilon > 0$ and $\forall M > 0\ \exists \rho > 0$ such that
$$\Big|\frac{\partial^2 L}{\partial y^2}(x,y_1) - \frac{\partial^2 L}{\partial y^2}(x,y_2)\Big| \le \varepsilon$$

for all $|y|, |y_1|, |y_2| \le M$ with $|y_1 - y_2| \le \rho$. All the above constants are independent of x.

The following theorem states the differentiability properties of the minimizing functional.

Theorem 3.2: The functional $J : \mathcal{A} \longrightarrow \mathbb{R}$ is of class C^2 and its derivatives are given by the expressions:

$$J'(u)v = \int_\Gamma (\nu u - y_u\varphi_u)v \,\mathrm{d}x, \tag{16}$$

$$J''(u)(v_1,v_2) = \int_\Omega \Big[\frac{\partial^2 L}{\partial y^2}(x,y_u) - \varphi_u \frac{\partial^2 a}{\partial y^2}(x,y_u)\Big] z_{u,v_1} z_{u,v_2} \,\mathrm{d}x - \int_\Gamma \Big[v_1 z_{u,v_2} + v_2 z_{u,v_1}\Big]\varphi_u \,\mathrm{d}x + \nu \int_\Gamma v_1 v_2 \,\mathrm{d}x, \tag{17}$$

for all $u \in \mathcal{A}$ and all $v, v_1, v_2 \in L^s(\Gamma)$, where $z_{u,v_i} = G'(u)v_i$, $i = 1,2$ and $\varphi_u \in Y$ is the adjoint state, the unique solution of the equation

$$\begin{cases} A^*\varphi + \dfrac{\partial a}{\partial y}(x,y_u)\varphi = \dfrac{\partial L}{\partial y}(x,y_u) & \text{in } \Omega, \\ \partial_{n_{A^*}}\varphi + u\varphi = 0 & \text{on } \Gamma. \end{cases} \tag{18}$$

Proof: Existence, uniqueness and regularity of φ_u follow from Remark 2.5, Assumption 3.1, and Theorem 2.4. The proofs of (16) and (17) are standard and can be established working identically to [2, Theorem 3.4]. ■

According to Theorem 3.2 the mapping $\Phi : \mathcal{A} \longrightarrow Y$ given by $\Phi(u) := \varphi_u$ is well defined. Let us prove that it is C^1.

Theorem 3.3: The mapping Φ is of class C^1 and for all $u \in \mathcal{A}$ and $v \in L^s(\Gamma)$ the function $\eta_{u,v} = \Phi'(u)v$ is the unique solution of

$$\begin{cases} A^*\eta + \dfrac{\partial a}{\partial y}(x,y_u)\eta = \Big[\dfrac{\partial^2 L}{\partial y^2}(x,y_u) - \varphi_u \dfrac{\partial^2 a}{\partial y^2}(x,y_u)\Big] z_{u,v} & \text{in } \Omega, \\ \partial_{n_{A^*}}\eta + u\eta = -v\varphi_u & \text{on } \Gamma, \end{cases} \tag{19}$$

where $z_{u,v} = G'(u)v$.

Proof: Using Assumption 3.1 and the fact that $y_u, \varphi_u, z_{u,v} \in L^\infty(\Omega)$ we obtain that the right hand side of (19) belongs to $L^p(\Omega) \times L^s(\Gamma)$. Existence, uniqueness, and regularity of $\eta_{u,v}$ follow from Remark 2.5. To establish the differentiability of Φ we define

$$Y_{A^*} = \{\varphi \in Y : A^*\varphi \in L^p(\Omega) \text{ and } \partial_{n_{A^*}}\varphi \in L^q(\Gamma)\}$$

and $\mathcal{G} : \mathcal{A} \times Y_{A^*} \longrightarrow L^p(\Omega) \times L^q(\Gamma)$ by

$$\mathcal{G}(u,\varphi) := \Big(A^*\varphi + \frac{\partial a}{\partial y}(\cdot,y_u)\varphi - \frac{\partial L}{\partial y}(\cdot,y_u), \partial_{n_{A^*}}\varphi + u\varphi\Big).$$

From assumptions 2.2 and 3.1 we deduce that $\mathcal{G}$ is of class C^1. Moreover, $\frac{\partial \mathcal{G}}{\partial\varphi}(u,\varphi) : Y_{A^*} \longrightarrow L^p(\Omega) \times L^q(\Gamma)$ is a linear and continuous mapping, and $\forall \eta \in Y_{A^*}$ we have that

$$\frac{\partial \mathcal{G}}{\partial\varphi}(u,\varphi)\eta = \left(A^*\eta + \frac{\partial a}{\partial y}(\cdot,y_u)\eta, \partial_{n_{A^*}}\eta + u\eta\right).$$

Using again Remark 2.5 we get that

$$\begin{cases} A^*\eta + \dfrac{\partial a}{\partial y}(x,y_u)\eta = f & \text{in } \Omega, \\ \partial_{n_{A^*}}\eta + u\eta = h & \text{on } \Gamma, \end{cases}$$

has a unique solution in Y_{A^*} for all $(f,h) \in L^p(\Omega) \times L^q(\Gamma)$. Hence, $\frac{\partial \mathcal{G}}{\partial\varphi}(u,\varphi) : Y_{A^*} \longrightarrow L^p(\Omega)\times L^q(\Gamma)$ is an isomorphism. Then, applying the implicit function theorem and differentiating the identity $\mathcal{G}(u,\Phi(u)) = 0$ the result follows. ∎

Combining (19) with (17) we deduce the following alternative representation formula for $J''(u)$.

Corollary 3.4: For every $v_1, v_2 \in L^s(\Gamma)$ and all $u \in \mathcal{A}$, the following identities hold

$$\begin{aligned} J''(u)(v_1,v_2) &= \int_\Gamma \Big[\nu v_1 - (\varphi_u z_{u,v_1} + y_u\eta_{u,v_1})\Big] v_2 \,\mathrm{d}x \\ &= \int_\Gamma \Big[\nu v_2 - (\varphi_u z_{u,v_2} + y_u\eta_{u,v_2})\Big] v_1 \,\mathrm{d}x. \end{aligned} \tag{20}$$

Remark 3.5: In dimension $d = 3$, we can also extend $J'(u)$ and $J''(u)$ respectively to continuous linear and bilinear forms in $L^2(\Gamma)$ and $L^2(\Gamma)^2$ by the same expressions given above. Indeed, we notice that for all $v \in L^2(\Gamma)$, the Lax-Milgram Theorem implies that equations (12) and (19) have a unique solution in $H^1(\Omega) \subset L^2(\Omega)$.

Theorem 3.6: Problem (P) has at least one solution. Moreover, if $\bar u \in U_{\rm ad}$ is a local minimizer of (P) then there exist $\bar y, \bar\varphi \in Y$ such that

$$\begin{cases} A\bar y + a(x,\bar y) = 0 & \text{in } \Omega, \\ \partial_{n_A}\bar y + \bar u\bar y = g & \text{on } \Gamma, \end{cases} \tag{21}$$

$$\begin{cases} A^*\bar\varphi + \dfrac{\partial a}{\partial y}(x,\bar y)\bar\varphi = \dfrac{\partial L}{\partial y}(x,\bar y) & \text{in } \Omega, \\ \partial_{n_{A^*}}\bar\varphi + \bar u\bar\varphi = 0 & \text{on } \Gamma, \end{cases} \tag{22}$$

$$\bar u(x) = \mathrm{Proj}_{[\alpha,\beta]}\left(\frac{1}{\nu}\bar y(x)\bar\varphi(x)\right) \qquad \forall x \in \Gamma. \tag{23}$$

Moreover, the regularity $\bar u \in C^{0,\mu}(\Gamma)$ holds.

Existence of optimal solutions follows using standard techniques. First order optimality conditions are an immediate consequence of (16) and the convexity of $U_{\rm ad}$. The Hölder continuity of $\bar u$ is a consequence of (23), the same regularity for $\bar y$ and $\bar\varphi$, and the Lipschitz property of the projection $\mathrm{Proj}_{[\alpha,\beta]}(t) = \max\{\alpha,\min\{\beta,t\}\}$. In this paper a local minimizer is intended in the $L^2(\Gamma)$ sense.

From now on $(\bar u,\bar y,\bar\varphi) \in U_{\rm ad} \times Y^2$ will denote a triplet satisfying (21)-(23). Associated with this triplet we define the cone of critical directions

$$C_{\bar u} = \{v \in L^2(\Gamma) : v(x) = 0 \text{ if } \nu\bar u(x) - \bar y(x)\bar\varphi(x) \neq 0 \text{ a.e. in } \Gamma \text{ and (24) holds}\},$$

$$v(x)\begin{cases} \geq 0 & \text{if } \bar u(x) = \alpha, \\ \leq 0 & \text{if } \bar u(x) = \beta. \end{cases} \tag{24}$$

We proceed now to the second order optimality conditions. The proof of the following theorem is standard; see, e.g. [5, Theorem 2.3].

Theorem 3.7: If $\bar u$ is a local minimizer of (P), then $J''(\bar u)v^2 \geq 0$ $\forall v \in C_{\bar u}$ holds. Conversely, if $\bar u \in U_{\rm ad}$ satisfies the first order optimality conditions (21)–(23) and $J''(\bar u)v^2 > 0$ $\forall v \in C_{\bar u} \setminus \{0\}$, then there exist $\varepsilon > 0$ and $\delta > 0$ such that

$$J(\bar u) + \frac{\delta}{2}\|u - \bar u\|^2_{L^2(\Gamma)} \leq J(u)\, \forall u \in U_{\rm ad} \text{ with } \|u - \bar u\|_{L^2(\Gamma)} \leq \varepsilon.$$

Definition 3.8: Let us define

$$\Sigma_{\bar u} = \{x \in \Gamma : \bar u(x) \in \{\alpha,\beta\} \text{ and } \nu\bar u(x) - \bar y(x)\bar\varphi(x) = 0\}.$$

We say that the strict complementarity condition is satisfied at $\bar u$ if $|\Sigma_{\bar u}| = 0$, where $|\cdot|$ stands for the $(d-1)$ dimensional Lebesgue measure on Γ.

For every $\tau \geq 0$, we define the subspace

$$T^\tau_{\bar u} = \{v \in L^2(\Gamma) : v(x) = 0 \text{ if } |\nu\bar u(x) - \bar y(x)\bar\varphi(x)| > \tau\}.$$

Theorem 3.9: Assume that $\bar u$ satisfies the strict complementarity condition. Then, the following properties hold:

1- $T^0_{\bar u} = C_{\bar u}$.

2- If $\bar u$ satisfies the second order optimality condition $J''(\bar u)v^2 > 0$ $\forall v \in C_{\bar u} \setminus \{0\}$, then $\exists \tau > 0$ and $\kappa > 0$ such that

$$J''(\bar u)v^2 \geq \kappa\|v\|^2_{L^2(\Gamma)} \;\forall v \in T^\tau_{\bar u}. \tag{25}$$

For the proof the reader is referred to [2, Theorem 3.10].

IV. Convergence of the semismooth Newton method

We define $F : \mathcal{A} \longrightarrow L^s(\Gamma)$ by $F(u) = u - \mathrm{Proj}_{[\alpha,\beta]}\left(\frac{1}{\nu}y_u\varphi_u\right)$. From theorems 2.4 and 3.2 we deduce that F is well defined. Due to Theorem 3.6, any local minimizer of (P) is a solution of $F(u) = 0$. If a local minimizer $\bar u$ satisfies $J''(\bar u)v^2 > 0$ $\forall v \in C_{\bar u} \setminus \{0\}$, it is the unique stationary point in $B_\delta(\bar u) \cap U_{\rm ad}$; see [5, Corollary 2.6]. We are going to apply the semismooth Newton method sketched in Algorithm 1 to solve this equation. Here $\partial F(u)$ is a set valued mapping such that F is ∂F semismooth in the sense stated in [15, Chapter 3]. Local superlinear convergence follows from the semismoothness of F and the uniform boundedness of the norms of the inverses of the operators M_j. In order to define $\partial F(u)$ $\forall u \in \mathcal{A}$ we introduce some additional functions.

Algorithm 1: Semismooth Newton method.

```
1 Initialize Choose u_0 ∈ A. Set j = 0.
2 for j ≥ 0 do
3     Choose M_j ∈ ∂F(u_j) and solve M_j v_j = −F(u_j).
4     Set u_{j+1} = u_j + v_j and j = j + 1.
5 end
```

$$\begin{aligned} &S : \mathcal{A} \longrightarrow L^s(\Gamma), \quad S(u) = \frac{1}{\nu}G(u)\Phi(u), \\ &\psi : \mathbb{R} \longrightarrow \mathbb{R}, \quad \psi(t) = \mathrm{Proj}_{[\alpha,\beta]}(t), \\ &\Psi : \mathcal{A} \longrightarrow L^s(\Gamma), \quad \Psi(u)(x) = \psi(S(u)(x)). \end{aligned}$$

For every $u \in \mathcal{A}$ we define

$$\partial\Psi(u) = \big\{N \in \mathcal{L}(L^s(\Gamma), L^s(\Gamma)) : Nv = hS'(u)v \ \forall v \in L^s(\Gamma) \text{ and for some measurable function } h : \Omega \longrightarrow \mathbb{R} \text{ such that } h(x) \in \partial\psi(S(u)(x))\big\}.$$

We observe that ψ is a Lipschitz function and by $\partial\psi(t)$ we denote the subdifferential in Clarke's sense; see [6, Chapter 2]. Note that

$$\partial\psi(t) = \begin{cases} \{1\} & \text{if } t \in (\alpha, \beta), \\ \{0\} & \text{if } t \notin [\alpha, \beta], \\ [0,1] & \text{if } t \in \{\alpha, \beta\}. \end{cases}$$

According to [15, Prop. 2.26], ψ is 1-order $\partial\psi$-semismooth.

Theorem 4.1: Ψ is $\partial\Psi$-semismooth in $\mathcal{A}$.

The proof follows that of [2, Theorem 4.3]. Along that proof, the Lipschitz continuity of S is obtained, which we state as a lemma.

Lemma 4.2: For all $\bar{u} \in \mathcal{A}_0$, there exists $L_S > 0$ such that

$$\|S(u_1) - S(u_2)\|_{C(\Gamma)} \le L_S \|u_1 - u_2\|_{L^s(\Gamma)} \ \forall u_1, u_2 \in B_{\varepsilon_{\bar{u}}}(\bar{u})$$

where $\varepsilon_{\bar{u}}$ is the one introduced in Theorem 2.4.

Corollary 4.3: The function $F : \mathcal{A} \longrightarrow L^s(\Gamma)$ is ∂F-semismooth in $\mathcal{A}$, where

$$\partial F(u) = \{M = I - N : N \in \partial\Psi(u)\},$$

and I denotes the identity in $L^s(\Gamma)$.

We select the operators $M_u : L^s(\Gamma) \longrightarrow L^s(\Gamma)$ for every $u \in \mathcal{A}$ as follows. First, we define the function $\lambda : \mathbb{R} \longrightarrow \mathbb{R}$ by

$$\lambda(t) = \begin{cases} 1 & \text{if } t \in (\alpha, \beta), \\ 0 & \text{otherwise.} \end{cases}$$

It is obvious that $\lambda(t) \in \partial\psi(t)$ for every $t \in \mathbb{R}$. We define $M_u : L^s(\Gamma) \longrightarrow L^s(\Gamma)$ by $M_u v = v - h_u \cdot S'(u)v$, where $h_u(x) = \lambda(S(u)(x)) = \lambda\big(\frac{1}{\nu} y_u(x)\varphi_u(x)\big)$. It is immediate that $M_u \in \partial F(u)$. For this selection we have the following result.

Theorem 4.4: Let $(\bar{u}, \bar{y}, \bar{\varphi}) \in U_{\rm ad} \times Y^2$ satisfy the first order optimality conditions (21)–(23), the strict complementarity condition $|\Sigma_{\bar{u}}| = 0$, and the second order sufficient optimality condition $J''(\bar{u})v^2 > 0$ for every $v \in C_{\bar{u}} \setminus \{0\}$. Then, there exist $\delta > 0$ and $C > 0$ such that for every $u \in B_\delta(\bar{u}) \subset \mathcal{A}$ the linear operator $M_u : L^s(\Gamma) \longrightarrow L^s(\Gamma)$ is an isomorphism and $\|M_u^{-1}\| \le C$ holds.

Proof: For any $u \in \mathcal{A}$, we define

$$\mathbb{A}_u = \{x \in \Gamma : \frac{1}{\nu} y_u(x)\varphi_u(x) \notin (\alpha, \beta)\},$$

$$\mathbb{I}_u = \{x \in \Gamma : \frac{1}{\nu} y_u(x)\varphi_u(x) \in (\alpha, \beta)\},$$

so $M_u v = v - \frac{1}{\nu}[z_{u,v}\varphi_u + y_u \eta_{u,v}]\chi_{\mathbb{I}_u}$. Here $\chi_{\mathbb{S}}$ stands for the characteristic function of a set $\mathbb{S}$. M_u being obviously continuous, it is enough to prove that the equation $M_u v = w$ has a unique solution $v \in L^s(\Gamma)$ for every $w \in L^s(\Gamma)$. Clearly, $v = w$ in $\mathbb{A}_u$, and hence, denoting $b = w + \frac{1}{\nu}[z_{u,\chi_{\mathbb{A}_u} w}\varphi_u + y_u \eta_{u,\chi_{\mathbb{A}_u} w}] \in L^s(\Gamma)$, to compute v we have to solve

$$\chi_{\mathbb{I}_u} v - \frac{1}{\nu}[z_{u,\chi_{\mathbb{I}_u} v}\varphi_u + y_u \eta_{u,\chi_{\mathbb{I}_u} v}] = b \text{ in } \mathbb{I}_u. \tag{26}$$

Using (20), it is obvious that this equation is the optimality condition of the unconstrained quadratic optimization problem

$$(Q) \min_{v \in L^2(\mathbb{I}_u)} \mathbb{J}(v) = \frac{1}{2\nu} J''(u)(\chi_{\mathbb{I}_u} v)^2 - \int_{\mathbb{I}_u} bv \, \mathrm{d}x.$$

Here and elsewhere, for every measurable set $\Sigma \subset \Gamma$ and $v \in L^1(\Sigma)$, $\chi_\Sigma v$ denotes the extension by 0 to $\Gamma \setminus \Sigma$. The continuity of J'' established in Theorem 3.2 and (25) imply the existence of $\delta_0 > 0$ such that

$$J''(u)v^2 \ge \frac{\kappa}{2}\|v\|^2_{L^2(\Gamma)} \quad \forall v \in T^\tau_{\bar{u}} \text{ and } \forall u \in B_{\delta_0}(\bar{u}).$$

Setting $\delta = \min\{\delta_0, \varepsilon_{\bar{u}}, \frac{\tau}{\nu L_S}\}$, where $\varepsilon_{\bar{u}}$ and L_S are introduced in Theorem 2.4 and Lemma 4.2 respectively, we have that $L^2(\mathbb{I}_u) \subset T^\tau_{\bar{u}}$ for all $u \in B_\delta(\bar{u})$; see [2, Theorem 4.5]. Therefore (Q) has a unique solution $v \in L^2(\mathbb{I}_u)$. Since $z_{u,\chi_{\mathbb{I}_u} v}, \eta_{u,\chi_{\mathbb{I}_u} v} \in L^s(\Gamma)$, (26) implies that $v \in L^s(\mathbb{I}_u)$ and, consequently, v is the unique solution of the equation $M_u v = w$ in $L^s(\Gamma)$.

To prove the uniform boundness of M_u^{-1} we proceed in two steps. First, using the same technique as in [2, Theorem 4.5], we obtain the existence of a constant $C_2 > 0$ such that $\|v\|_{L^2(\Gamma)} \le C_2\|w\|_{L^s(\Gamma)}$. If $d = 2$ the proof is complete with $C = C_2$. For $d = 3$, we use Remark 3.5 and the previous estimate to obtain $\|z_{u,\chi_{\mathbb{I}_u} v}\|_{L^s(\Gamma)} + \|\eta_{u,\chi_{\mathbb{I}_u} v}\|_{L^s(\Gamma)} \le C_3\|v\|_{L^2(\Gamma)} \le C_3 C_2 \|w\|_{L^s(\Gamma)}$ for some $C_3 > 0$. Applying this and (26), we get $\|v\|_{L^s(\mathbb{I}_u)} \le \|b\|_{L^s(\Gamma)} + C_4(\|z_{u,\chi_{\mathbb{I}_u} v}\|_{L^s(\Gamma)} + \|\eta_{u,\chi_{\mathbb{I}_u} v}\|_{L^s(\Gamma)}) \le C_5\|w\|_{L^s(\Gamma)}$, and the result follows for $C = \max\{1, C_5\}$. ■

Algorithm 2 implements the semismooth Newton method to solve (P). As a straightforward consequence of [15, Theorem 3.13], Corollary 4.3, and Theorem 4.4 we conclude the convergence of this algorithm.

Corollary 4.5: Let $(\bar{u}, \bar{y}, \bar{\varphi}) \in U_{\rm ad} \times Y^2$ satisfy the first order optimality conditions (21)–(23), the strict complementarity condition $|\Sigma_{\bar{u}}| = 0$, and the second order sufficient optimality condition $J''(\bar{u})v^2 > 0$ for every $v \in C_{\bar{u}} \setminus \{0\}$. Then, there exists $\delta > 0$ such that for all $u_0 \in B_\delta(\bar{u})$ the sequence $\{u_j\}$ generated by Algorithm 2 is contained in the ball $B_\delta(\bar{u})$ and converges superlinearly to $\bar{u}$.

The radius of the basin of attraction δ depends on parameters related to the continuity properties of the involved functionals and its derivatives, the second order condition and the neighborhood in $L^s(\Gamma)$ for which the state equation is meaningful.

V. A NUMERICAL EXAMPLE AND SOME COMPUTATIONAL CONSIDERATIONS

Consider $\Omega = (0,1)^3$, $Ay = -\Delta y + y$, $a(x,y) = y^3 - \sin(2\pi x_1)\sin(\pi x_2)\cos(3\pi x_3)$, $g \equiv 0$, $L(x,y) = 0.5(y - y_d(x))^2$, with $y_d(x) = -512\prod_{i=1}^3 x_i(1-x_i)$, $\nu = 0.01$, $\alpha = 0$, and $\beta = 1$. We solve a finite element discretization of (P). Continuous piecewise linear functions are used for the state, the adjoint state, and the control. The Tichonov regularization term is discretized using the lumped mass matrix. In this way, the optimization algorithm for the discrete problem is exactly the discrete version of Algorithm 2.

The convergence history for $u_0 = 0$ is summarized in tables I and II for different mesh sizes. The expected superlinear convergence can be seen in the relative errors between consecutive

Algorithm 2: Semismooth Newton method for (P).

Initialize. Choose $u_0 \in \mathcal{A}$. Set $j = 0$.

for $j \geq 0$ **do**

Compute $y_j = G(u_j)$

Compute $\varphi_j = \Phi(u_j)$

Set $\mathbb{A}_j = \mathbb{A}_j^\beta \cup \mathbb{A}_j^\alpha$ and $\mathbb{I}_j = \Gamma \setminus \mathbb{A}_j$, where

$$\mathbb{A}_j^\beta = \{x \in \Gamma : y_j(x)\varphi_j(x) \geq \nu\beta\},$$
$$\mathbb{A}_j^\alpha = \{x \in \Gamma : y_j(x)\varphi_j(x) \leq \nu\alpha\}$$

Set $w_j(x) = -F(u_j)(x)$:

$$w_j(x) = \begin{cases} -u_j(x) + \beta & \text{if } x \in \mathbb{A}_j^\beta \\ -u_j(x) + \frac{1}{\nu}\varphi_j(x)y_j(x) & \text{if } x \in \mathbb{I}_j \\ -u_j(x) + \alpha & \text{if } x \in \mathbb{A}_j^\alpha \end{cases}$$

Compute $z_j = z_{u_j,\chi_{\mathbb{A}_j} w_j}$ and $\eta_j = \eta_{u_j,\chi_{\mathbb{A}_j} w_j}$

Solve the quadratic problem

$$(Q_j) \qquad \min_{v \in L^2(\mathbb{I}_j)} \mathbb{J}_j(v) := \frac{1}{2\nu} J''(u_j)(\chi_{\mathbb{I}_j} v)^2$$
$$- \int_{\mathbb{I}_j} (w_j + \frac{1}{\nu}[z_j\varphi_j + y_j\eta_j])v \,\mathrm{d}x$$

Name $v_{\mathbb{I}_j}$ its solution.

Set $u_{j+1} = u_j + \chi_{\mathbb{A}_j} w_j + \chi_{\mathbb{I}_j} v_{\mathbb{I}_j}$ and $j = j + 1$.

end

iterations, denoted δ_j. We also remark the mesh-independence of the convergence history, which is to be expected since we have obtained our results in the infinite-dimensional setting.

At each iteration we have to solve a nonlinear equation to compute y_j and solve an unconstrained quadratic problem to compute $v_{\mathbb{I}_j}$. We use Newton's method for the first task and the conjugate gradient method for the second one. Notice that $\mathbb{J}_j(v) = \frac{1}{2}(v, A_j v)_{L^2(\mathbb{I}_j)} - (b_j, v)_{L^2(\mathbb{I}_j)}$, where $b_j = \chi_{\mathbb{I}_j}(w_j + \frac{1}{\nu}[z_j\varphi_j + y_j\eta_j])$ and, for any $v \in L^2(\mathbb{I}_j)$,

$$A_j v = \chi_{\mathbb{I}_j}\left(v + \frac{1}{\nu}[z_{u_j,\chi_{\mathbb{I}_j} v}\varphi_j + \eta_{u_j,\chi_{\mathbb{I}_j} v} y_j]\right);$$

see eqs. (12) and (19)

We include in the tables the number of Newton iterations used to solve the nonlinear equation at each iteration. Each of these requires the factorization of the finite element matrix, and this number is a good measure of the global complexity of the method. In contrast, each of the conjugate gradient iterations used to solve (Q_j) requires the solution of two linear systems, but the matrix has been previously factorized in the last step of the nonlinear solve.

j	$J(u_j)$	δ_j	♯Newton	♯CG
0	4.7607853276096295e+00	7.3e-01	3	17
1	4.7590621154705985e+00	5.3e-01	3	12
2	4.7588905662088630e+00	1.1e-01	3	12
3	4.7588301468521248e+00	3.7e-04	3	12
4	4.7588301456859448e+00	7.9e-08	2	12
5	4.7588301456859456e+00	3.7e-15	2	12
6	4.7588301456859456e+00		1	

TABLE I

SOLUTION OF (P) FOR $h = 2^{-4}$.

j	$J(u_j)$	δ_j	♯Newton	♯CG
0	4.8308890801571112e+00	7.9e-01	3	16
1	4.8290362150750905e+00	5.8e-01	3	11
2	4.8288131518545896e+00	1.3e-01	3	12
3	4.8287240470263058e+00	7.3e-04	3	11
4	4.8287240439741863e+00	4.7e-06	2	11
5	4.8287240439742973e+00	6.3e-14	2	11
6	4.8287240439742973e+00		1	

TABLE II

SOLUTION OF (P) FOR $h = 2^{-5}$.

REFERENCES

[1] S. Amstutz and A. Laurain, "A semismooth Newton method for a class of semilinear optimal control problems with box and volume constraints," *Comput. Optim. Appl.*, vol. 56, no. 2, pp. 369–403, 2013.

[2] E. Casas, K. Chrysafinos, and M. Mateos, "Bilinear control of semilinear elliptic pdes: Convergence of a semismooth newton method," *arXiV*, vol. 2309.07554, 2023.

[3] E. Casas and M. Mateos, "Convergence analysis of the semismooth newton method for sparse control problems governed by semilinear elliptic equations," *arXiV*, vol. 2309.0593, 2023.

[4] E. Casas, "Boundary control of semilinear elliptic equations with pointwise state constraints," *SIAM J. Control Optim.*, vol. 31, no. 4, pp. 993–1006, 1993.

[5] E. Casas and F. Tröltzsch, "Second order analysis for optimal control problems: improving results expected from abstract theory," *SIAM J. Optim.*, vol. 22, no. 1, pp. 261–279, 2012.

[6] F. H. Clarke, *Optimization and Nonsmooth Analysis*. New York: John Wiley & Sons, 1983.

[7] J. C. de los Reyes and K. Kunisch, "A semi-smooth Newton method for control constrained boundary optimal control of the Navier-Stokes equations," *Nonlinear Anal.*, vol. 62, no. 7, pp. 1289–1316, 2005.

[8] A. Kröner, K. Kunisch, and B. Vexler, "Semismooth Newton methods for optimal control of the wave equation with control constraints," *SIAM J. Control Optim.*, vol. 49, no. 2, pp. 830–858, 2011.

[9] F. Mannel and A. Rund, "A hybrid semismooth quasi-Newton method for nonsmooth optimal control with PDEs," *Optim. Eng.*, vol. 22, no. 4, pp. 2087–2125, 2021.

[10] R. R. Mohler, *Bilinear control processes. With applications to engineering, ecology, and medicine*, ser. Math. Sci. Eng. Elsevier, Amsterdam, 1973, vol. 106.

[11] R. Nittka, "Regularity of solutions of linear second order elliptic and parabolic boundary value problems on Lipschitz domains," *J. Differ. Equations*, vol. 251, no. 4-5, pp. 860–880, 2011.

[12] J. Nocedal and S. J. Wright, *Numerical optimization*, ser. Springer Series in Operations Research. Springer-Verlag, New York, 1999.

[13] K. Pieper, "Finite element discretization and efficient numerical solution of elliptic and parabolic sparse control problems," Ph.D. dissertation, Technische Universität München, 2015. [Online]. Available: https://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:91-diss-20150420-1241413-1-4

[14] G. Stampacchia, "Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus," *Annales de l'Institut Fourier*, vol. 15, no. 1, pp. 189–257, 1965.

[15] M. Ulbrich, *Semismooth Newton methods for variational inequalities and constrained optimization problems in function spaces*, ser. MOS/SIAM Ser. Optim. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2011, vol. 11.

[16] M. Winkler, "Error estimates for the finite element approximation of bilinear boundary control problems," *Comput. Optim. Appl.*, vol. 76, no. 1, pp. 155–199, 2020.