SIAM J. OPTIM.
Vol. 34, No. 4, pp. 3681–3698

SUPERLINEAR CONVERGENCE OF A SEMISMOOTH NEWTON METHOD FOR SOME OPTIMIZATION PROBLEMS WITH APPLICATIONS TO CONTROL THEORY*

EDUARDO CASAS†

Abstract. In this paper, we formulate a semismooth Newton method for an abstract optimization problem and prove its superlinear convergence by assuming that the no-gap second order sufficient optimality condition and the strict complementarity condition are fulfilled at the local minimizer. Many control problems fit this abstract formulation. In particular, we apply this abstract result to distributed control problems of a semilinear elliptic equation, to boundary bilinear control problems associated with a semilinear elliptic equation, and to distributed control of a semilinear parabolic equation.

Key words. semismooth Newton method, optimal control, second order optimality conditions, strict complementarity condition

MSC codes. 49K20, 49M05, 49M15

DOI. 10.1137/24M1644286

1. Introduction.

Let $(X,\mathcal{S},\mu)$ be a measure space with $0<\mu(X)<\infty$. In this paper, we prove the superlinear convergence of a semismooth Newton method to solve the following abstract optimization problem:

$$\text{(P)} \qquad \min_{\alpha\le u(x)\le\beta \text{ a.e.}[\mu]} \mathcal{J}(u)+\frac{\kappa}{2}\|u\|^2_{L^2(X)}$$

where $\kappa>0$, $-\infty\le\alpha<\beta\le+\infty$, and $\mathcal{J}:L^p(X)\longrightarrow\mathbb{R}$ is a function of class C^2 for some $p\in[2,+\infty)$. Many optimal control problems fall within this abstract formulation: distributed or boundary control problems, and bilinear control problems associated with nonlinear elliptic or parabolic equations. The analysis of semismooth Newton methods has a long history, but there are only a few papers in the framework of optimal control problems of partial differential equations where this method is proved to converge superlinearly. The reader is referred to [8, 11, 13] and [14, Chapter 8] for the case of convex control problems. In this case, the well-posedness and superlinear convergence of the method are established. The situation is more delicate for nonlinear state equations [1, 9, 11], where a strong second order condition is assumed to prove the convergence of the algorithm. We also mention [14, Chapter 10] where the semismooth Newton method is applied to solve a control problem associated with the Navier-Stokes equations in dimension $n=2$. In this chapter, the superlinear convergence is proved assuming that the second derivative of the cost functional at the local minimizer is coercive on the tangent space of the strongly active constraints and that the control constraint is reduced to $u\ge 0$.

The aim of this paper is to prove the superlinear convergence of the algorithm to local solutions $\bar u$ assuming that the no-gap second order sufficient optimality condition and the strict complementarity condition are fulfilled at $\bar u$. These are the usual

*Received by the editors March 5, 2024; accepted for publication (in revised form) August 26, 2024; published electronically November 19, 2024.
https://doi.org/10.1137/24M1644286
Funding: The author was partially supported by MCIN/AEI/10.13039/501100011033/ under research project PID2020-114837GB-I00.
†Departmento de Matemática Aplicada y Ciencias de la Computación, E.T.S.I. Industriales y de Telecomunicación, Universidad de Cantabria, 39005 Santander, Spain (eduardo.casas@unican.es).

assumptions required to prove the superlinear convergence of numerical algorithms in finite-dimensional optimization problems; see, for instance, [10, Chapters 17 and 18].

The plan of this paper is as follows. In section 2, we formulate the hypotheses on the problem (P), propose a semismooth Newton method and prove its superlinear convergence. In section 3, we formulate and establish superlinear convergence of this algorithm for three different control problems: distributed control of a semilinear elliptic equation; boundary bilinear control associated with a semilinear elliptic equation; and distributed control of a semilinear parabolic equation. This selection of control problems is made only to show the generality of the abstract result. Many other cases could be included in this abstract formulation.

2. A semismooth Newton method to solve (P).

This section is divided into two parts. In the first part, we establish the assumptions on (P) and carry out the first and second order analysis of this problem. In the second part, we formulate a semismooth Newton method to compute a local solution of (P) and prove its superlinear convergence.

2.1. Analysis of (P).

Let us fix some notation and make the assumptions on problem (P). We set

$$\mathcal{U}_{ad} = \{u \in L^p(X) : \alpha \le u(x) \le \beta \text{ a.e.}[\mu]\},$$

where $-\infty \le \alpha < \beta \le +\infty$ and additionally

$$\text{if } p > 2, \quad \text{then} \quad -\infty < \alpha < \beta < +\infty. \tag{2.1}$$

Let $\mathcal{A}$ be an open subset of $L^p(X)$ such that $\mathcal{U}_{ad} \subset \mathcal{A}$. The function $\mathcal{J} : \mathcal{A} \longrightarrow \mathbb{R}$ is of class C^2 and satisfies the following hypotheses:

(H1) There exists a C^1 mapping $\Phi : \mathcal{A} \longrightarrow L^\infty(X)$ such that

$$\mathcal{J}'(u)v = \int_X \Phi(u)v \,\mathrm{d}\mu \quad \forall u \in \mathcal{A} \ \text{ and } \ \forall v \in L^p(X). \tag{2.2}$$

(H2) For every $u \in \mathcal{A}$ the linear mapping $\Phi'(u) : L^p(X) \longrightarrow L^\infty(X)$ has an extension to a compact operator $\Phi'(u) : L^2(X) \longrightarrow L^p(X)$ satisfying the following: for all $\varepsilon > 0$ there exists $\rho > 0$ with $B_\rho(u) \subset \mathcal{A}$ such that

$$\|[\Phi'(w) - \Phi'(u)]v\|_{L^p(X)} \le \varepsilon \|v\|_{L^2(X)} \quad \forall w \in B_\rho(u) \text{ and } \forall v \in L^2(X). \tag{2.3}$$

Above and along this paper, $B_\rho(u)$ denotes the open ball of $L^p(X)$ centered at u with radius ρ.

Remark 2.1.

i) The assumption (2.1) is used in the second order analysis, where some Taylor expansions have to be performed in an $L^2(X)$ neighborhood of some point $\bar{u}$.
ii) In the case $p = 2$, (2.3) is a consequence of **(H1)**, hence hypothesis **(H2)** only assumes that the linear mapping $\Phi'(u) : L^2(X) \longrightarrow L^2(X)$ is compact.
iii) As a consequence of hypotheses **(H1)** and **(H2)**, we infer that for every $u \in \mathcal{A}$ the bilinear form $\mathcal{J}''(u) : L^p(X)^2 \longrightarrow \mathbb{R}$ has a continuous extension to $L^2(X)^2$ such that the weak convergence $(v_k, w_k) \rightharpoonup (v, w)$ in $L^2(X)^2$ implies

$$\mathcal{J}''(u)(v_k, w_k) = \int_X [\Phi'(u)v_k]w_k \,\mathrm{d}\mu \to \int_X [\Phi'(u)v]w \,\mathrm{d}\mu = \mathcal{J}''(u)(v, w). \tag{2.4}$$

We define the function $J:\mathcal{A}\longrightarrow\mathbb{R}$ by $J(u)=\mathcal{J}(u)+\frac{\kappa}{2}\|u\|^2_{L^2(X)}$. Then, (P) can be written as follows:

$$\text{(P)} \qquad \min_{u\in\mathcal{U}_{ad}} J(u).$$

From (2.1) we infer that $\bar u$ is a (strict) local minimizer of (P) in the $L^p(X)$-sense if and only if it is a (strict) local minimizer in the $L^2(X)$ sense. From now on, a local minimizer will be understood in the $L^2(X)$ sense.

THEOREM 2.2. *If $\bar u\in\mathcal{U}_{ad}$ is a local minimizer of* (P)*, then the following identity holds:*

$$\bar u(x)=\text{Proj}_{[\alpha,\beta]}\Big(-\frac{1}{\kappa}\Phi(\bar u)(x)\Big) \quad a.e.\,[\mu]. \tag{2.5}$$

Proof. Using the convexity of $\mathcal{U}_{ad}$ and (2.2) we get

$$\int_X(\Phi(\bar u)+\kappa\bar u)(u-\bar u)\,\mathrm{d}\mu=J'(\bar u)(u-\bar u)\ge 0 \quad \forall u\in\mathcal{U}_{ad},$$

that is equivalent to (2.5). ∎

Associated to a local minimizer $\bar u$ we define the cone of critical directions $C_{\bar u}$ as the set of elements $v\in L^2(X)$ satisfying

$$v(x)\begin{cases}\ge 0 & \text{if } \bar u(x)=\alpha,\\ \le 0 & \text{if } \bar u(x)=\beta,\\ =0 & \text{if } \kappa\bar u(x)+\Phi(\bar u)(x)\ne 0,\end{cases} \qquad \text{a.e. } [\mu].$$

It is well known that a local minimizer of (P) satisfies the second order necessary condition: $J''(\bar u)v^2\ge 0$ $\forall v\in C_{\bar u}$; see, for instance, [5, Theorem 2.4 and Remark 2.5]. Conversely, if $\bar u\in\mathcal{U}_{ad}$ satisfies (2.5) and $J''(\bar u)v^2>0$ $\forall v\in C_{\bar u}\setminus\{0\}$, then there exist $\varepsilon>0$ and $\nu>0$ such that

$$J(\bar u)+\frac{\nu}{2}\|u-\bar u\|^2_{L^2(X)}\le J(u) \quad \forall u\in\mathcal{U}_{ad} \text{ with } \|u-\bar u\|_{L^2(X)}\le\varepsilon. \tag{2.6}$$

See [5, Theorem 2.6] for the proof. Furthermore, it was proved in [6, Corollary 2.6] that $\varepsilon>0$ can be selected such that there is no other element $u\in\mathcal{A}$ with $\|u-\bar u\|_{L^2(X)}<\varepsilon$ satisfying the optimality condition (2.5).

(H3) In the rest of this section, $\bar u$ will denote a local minimizer of (P) satisfying the following assumptions

$$J''(\bar u)v^2>0 \quad \forall v\in C_{\bar u}\setminus\{0\}, \tag{2.7}$$

$$\mu\big(\{x\in X:\bar u(x)\in\{\alpha,\beta\} \text{ and } \kappa\bar u(x)+\Phi(\bar u)(x)=0\}\big)=0. \tag{2.8}$$

We refer to (2.8) as the strict complementarity condition. Now, for $\tau\ge 0$ we define

$$E^\tau_{\bar u}=\{v\in L^2(X):v(x)=0 \text{ if } |\kappa\bar u(x)+\Phi(\bar u)(x)|>\tau\}. \tag{2.9}$$

The following result is crucial in the proof of the semismooth Newton method defined later.

THEOREM 2.3. *Let $\bar u\in\mathcal{U}_{ad}$ satisfy* (2.7) *and* (2.8). *Then, there exist $\delta>0$ and $\tau>0$ such that*

$$J''(\bar u)v^2\ge\delta\|v\|^2_{L^2(X)} \quad \forall v\in E^\tau_{\bar u}. \tag{2.10}$$

Proof. First, we observe that (2.8) implies that $C_{\bar u} = E^0_{\bar u}$. Now, we proceed by contradiction. Assume that for every integer $k \ge 1$ there exists $v_k \in E^{\frac{1}{k}}_{\bar u}$ such that $J''(\bar u)v_k^2 < \frac{1}{k}\|v_k\|^2_{L^2(X)}$. Dividing v_k by its $L^2(X)$-norm and taking a subsequence we infer

$$v_k \in E^{\frac{1}{k}}_{\bar u} \text{ and } J''(\bar u)v_k^2 < \frac{1}{k} \ \ \forall k \ge 1, \text{ and } v_k \rightharpoonup v \text{ in } L^2(X) \text{ as } k \to \infty. \tag{2.11}$$

Let us take $\varepsilon > 0$ arbitrarily. It is obvious that $E^\varepsilon_{\bar u}$ is a closed subspace of $L^2(X)$ and $\{v_k\}_{k>\frac{1}{\varepsilon}} \subset E^\varepsilon_{\bar u}$. Therefore, $v \in E^\varepsilon_{\bar u}$ holds. Since $E^0_{\bar u} = \cap_{\varepsilon>0} E^\varepsilon_{\bar u}$, we infer that $v \in E^0_{\bar u}$. Hence, (2.7) implies that $J''(\bar u)v^2 > 0$ unless $v = 0$. But, (2.11) and (2.4) lead to

$$J''(\bar u)v^2 \le \liminf_{k\to\infty} J''(\bar u)v_k^2 \le \limsup_{k\to\infty} J''(\bar u)v_k^2 \le 0.$$

Therefore, we have that $v = 0$. Using (2.4) again and the above inequalities we get that $\mathcal{J}''(\bar u)v_k^2 \to 0$ and $J''(\bar u)v_k^2 \to 0$. These convergences and the fact that $\|v_k\|_{L^2(X)} = 1$ yield

$$\kappa = \lim_{k\to\infty} \kappa\|v_k\|^2_{L^2(X)} = \lim_{k\to\infty} (J''(\bar u)v_k^2 - \mathcal{J}''(\bar u)v_k^2) = 0.$$

This contradicts our assumption $\kappa > 0$. ∎

2.2. A semismooth Newton method. Let $\psi : \mathbb{R} \longrightarrow \mathbb{R}$ be the function given by $\psi(t) = \mathrm{Proj}_{[\alpha,\beta]}(-\frac{1}{\kappa}t)$. We also define the mappings $\Psi : L^\infty(X) \longrightarrow L^\infty(X)$ and $\mathcal{F} : \mathcal{A} \longrightarrow L^p(X)$ by $\Psi(y)(x) = \psi(y(x))$ and $\mathcal{F}(u) = u - \Psi(\Phi(u))$. From (2.5) we infer that any local solution of (P) solves the equation $\mathcal{F}(u) = 0$. Moreover, as claimed after (2.6), if $\bar u$ satisfies (2.7), then it is the unique solution of this equation in a $L^p(X)$ ball around $\bar u$. Unfortunately, $\mathcal{F}$ is not Fréchet differentiable in $\mathcal{A}$ due to the lack of differentiability of ψ. However, $\mathcal{F}$ is semismooth. Let us recall the definition of semismoothness. To this end, we follow [14, Definition 3.1]. A slightly different approach using the concept of slant differentiability can be found in [8].

Definition 2.4. *Given two Banach spaces $\mathcal{X}$ and $\mathcal{Y}$, an open subset $\mathcal{V}$ of $\mathcal{X}$, a continuous function $\mathcal{H} : \mathcal{V} \longrightarrow \mathcal{Y}$, and a set-valued mapping $\partial\mathcal{H} : \mathcal{V} \rightrightarrows \mathcal{L}(\mathcal{X}, \mathcal{Y})$ such that $\partial\mathcal{H}(u) \ne \emptyset$ for every $u \in \mathcal{V}$, we say that $\mathcal{H}$ is $\partial\mathcal{H}$-semismooth at $\bar u \in \mathcal{V}$ if*

$$\lim_{v\to 0} \sup_{M\in\partial\mathcal{H}(\bar u+v)} \frac{\|\mathcal{H}(\bar u + v) - \mathcal{H}(\bar u) - Mv\|_{\mathcal{Y}}}{\|v\|_{\mathcal{X}}} = 0. \tag{2.12}$$

The multifunction $\partial\mathcal{H}$ is called the generalized derivative of $\mathcal{H}$.

In order to solve the equation $\mathcal{H}(u) = 0$, the semismooth Newton method generates a sequence according to Algorithm 1.

The proof of the following convergence theorem can be found in [14, Theorem 3.13]. See also [8, Theorem 1.1].

Algorithm 1. Semismooth Newton method.

```
Initialize. Choose u_0 ∈ V. Set j = 0.
repeat
    Choose M_j ∈ ∂H(u_j) and solve M_j v_j = −H(u_j).
    Set u_{j+1} = u_j + v_j and j = j + 1.
until convergence
```

THEOREM 2.5. *Suppose that $\mathcal{H}:\mathcal{V}\longrightarrow\mathcal{Y}$ is $\partial\mathcal{H}$-semismooth at $\bar u\in\mathcal{V}$ solution of $\mathcal{H}(u)=0$. Then, there exists $\delta>0$ such that for all $u_0\in\mathcal{V}$ with $\|u_0-\bar u\|_{\mathcal{X}}<\delta$ the sequence $\{u_j\}_{j\geq 0}$ generated by the semismooth Newton method* 1 *converges superlinearly to $\bar u$ if the operators $M_j\in\partial\mathcal{H}(u_j)$ are invertible and there exists $C_{\mathcal{H}}>0$ such that*

$$\|M_j^{-1}\|_{\mathcal{L}(\mathcal{Y},\mathcal{X})}\leq C_{\mathcal{H}}\ \forall j\geq 0. \tag{2.13}$$

Let us prove that Algorithm 1 can be applied to our equation $\mathcal{F}(u)=0$ and the superlinear convergence holds. First, we check that $\mathcal{F}$ is semismooth in $\mathcal{A}$ with respect to some generalized derivative that we define below. Foremost, we observe that ψ is a Lipschitz function with $\frac{1}{\kappa}$ as Lipschitz constant and its Clarke's subdifferential is given by

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

Now, we define

$$\begin{aligned}\partial\mathcal{F}(u)=\{&M\in\mathcal{L}(L^p(X),L^p(X)):(Mv)(x)=v(x)-h(\Phi(u)(x))[\Phi'(u)v](x),\\ &\text{where } h:\mathbb{R}\longrightarrow\mathbb{R} \text{ is Lebesgue measurable and } h(\Phi(u)(x))\in\partial\psi(\Phi(u)(x))\}.\end{aligned}$$

It is obvious that $\partial\mathcal{F}(u)$ is not empty for every $u\in\mathcal{A}$. Since ψ is Lipschitz and $\Phi:\mathcal{A}\longrightarrow L^\infty(X)$ is of class C^1 and, hence, locally Lipschitz, we infer from the results of [14, section 3.3] that $\mathcal{F}$ is $\partial\mathcal{F}$-semismooth in $\mathcal{A}$.

Let us define the function $g:\mathbb{R}\longrightarrow\mathbb{R}$ by

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

It is obvious that $g(t)\in\partial\psi(t)$ for every $t\in\mathbb{R}$.

To implement the semismooth Newton method we select for every $u\in\mathcal{A}$ the element $M_u\in\partial\mathcal{F}(u)$ defined by $M_uv=v-h_u\cdot\Phi'(u)v$, where $h_u(x)=g(\Phi(u)(x))$. We have the following property of these linear operators.

THEOREM 2.6. *Let $\bar u$ satisfy* (2.5), (2.7), *and* (2.8). *Then, there exist $\rho>0$ and $C>0$ such that $M_u:L^p(X)\longrightarrow L^p(X)$ is an isomorphism and $\|M_u^{-1}\|\leq C$ for all $u\in B_\rho(\bar u)\subset\mathcal{A}$, where $B_\rho(\bar u)$ denotes the ball of center $\bar u$ and radius ρ in $L^p(X)$.*

Proof. We split the proof into two steps. First, we prove that M_u is an isomorphism for every u in a certain ball around $\bar u$ and then we prove that $\|M_u^{-1}\|\leq C$ for all u in such a ball.

Step 1. *M_u is an isomorphism.* Let u be an element of $\mathcal{A}$. Since $h_u:X\longrightarrow\mathbb{R}$ is measurable, $\|h_u\|_{L^\infty(X)}\leq\frac{1}{\kappa}$, and $\Phi'(u):L^p(X)\longrightarrow L^\infty(X)$ is a continuous linear mapping, we deduce the continuity of $M_u:L^p(X)\longrightarrow L^p(X)$. To prove that M_u is an isomorphism it is enough to show that M_u is bijective and apply the open mapping theorem. Given $w\in L^p(X)$ arbitrary we will prove that there exists a unique element $\bar v\in L^p(X)$ such that $M_u\bar v=w$. Let us define the measurable sets

$$\mathbb{I}_u=\left\{x\in X:-\frac{1}{\kappa}\Phi(u)(x)\in(\alpha,\beta)\right\}\quad\text{and}\quad\mathbb{A}_u=X\setminus\mathbb{I}_u.$$

From the definition of h_u we get that $M_u\bar v = w$ holds if and only if

$$\begin{cases} \bar v(x) = w(x) & \text{if } x \in \mathbb{A}_u,\\ \bar v(x) + \dfrac{1}{\kappa}[\Phi'(u)\bar v](x) = w(x) & \text{if } x \in \mathbb{I}_u. \end{cases} \tag{2.14}$$

Denoting by $\chi_{\mathbb{I}_u}$ and $\chi_{\mathbb{A}_u}$ the characteristic functions of $\mathbb{I}_u$ and $\mathbb{A}_u$, respectively, we have that $v = \chi_{\mathbb{I}_u} v + \chi_{\mathbb{A}_u} v$ for every $v \in L^p(X)$. The first equation of (2.14) defines $\bar v$ univocally in $\mathbb{A}_u$. Using this fact we infer that the second equation is equivalent to

$$[\chi_{\mathbb{I}_u}\bar v](x) + \frac{1}{\kappa}[\Phi'(u)(\chi_{\mathbb{I}_u}\bar v)](x) = [\chi_{\mathbb{I}_u} w](x) - \frac{1}{\kappa}[\Phi'(u)(\chi_{\mathbb{A}_u} w)](x) \text{ if } x \in \mathbb{I}_u. \tag{2.15}$$

For every measurable function $v : \mathbb{I}_u \longrightarrow \mathbb{R}$, $\chi_{\mathbb{I}_u} v$ denotes the extension of v to X by zero. Now, we define the linear quadratic mapping $H : L^2(\mathbb{I}_u) \longrightarrow \mathbb{R}$ by

$$H(v) = \frac{1}{2\kappa} J''(u)(\chi_{\mathbb{I}_u} v)^2 - \int_{\mathbb{I}_u} \big([\chi_{\mathbb{I}_u} w](x) - \frac{1}{\kappa}[\Phi'(u)(\chi_{\mathbb{A}_u} w)](x)\big)[\chi_{\mathbb{I}_u} v](x)\,\mathrm{d}\mu.$$

From (2.2) we get that $\chi_{\mathbb{I}_u}\bar v$ satisfies (2.15) if and only if $H'(\bar v) = 0$. Let us recall that hypothesis **(H2)** yields $\Phi'(u)v \in L^p(X)$ for every $v \in L^2(X)$. Therefore, since $w \in L^p(X)$, if $\chi_{\mathbb{I}_u}\bar v \in L^2(X)$ solves the equality (2.15), then $\chi_{\mathbb{I}_u}\bar v \in L^p(X)$ holds. If we prove that H is strictly convex and coercive we infer the existence and uniqueness of a point $\chi_{\mathbb{I}_u}\bar v$ satisfying (2.15). To this end, it is enough to show the existence of $\nu > 0$ such that $J''(u)(\chi_{\mathbb{I}_u} v)^2 \ge \nu\|v\|^2_{L^2(\mathbb{I}_u)}$. Let δ be the constant introduced in (2.10). From (2.3) with $\varepsilon = \frac{\delta}{2\mu(X)^{\frac{p-2}{2p}}}$ we deduce the existence of $\rho_1 > 0$ such that $B_{\rho_1}(\bar u) \subset \mathcal{A}$ and

$$\|[\Phi'(u) - \Phi'(\bar u)]v\|_{L^p(X)} \le \frac{\delta}{2\mu(X)^{\frac{p-2}{2p}}}\|v\|_{L^2(X)} \ \forall u \in B_{\rho_1}(\bar u) \text{ and } \forall v \in L^2(X). \tag{2.16}$$

Now, using (2.10), taking $u \in B_{\rho_1}(\bar u)$, and assuming that $\chi_{\mathbb{I}_u} v \in E^\tau_{\bar u}$, the following inequalities hold:

$$\begin{aligned} J''(u)(\chi_{\mathbb{I}_u} v)^2 &\ge J''(\bar u)(\chi_{\mathbb{I}_u} v)^2 - |[J''(u) - J''(\bar u)](\chi_{\mathbb{I}_u} v)^2| \\ &= J''(\bar u)(\chi_{\mathbb{I}_u} v)^2 - \left|\int_X \big([\Phi'(u) - \Phi'(\bar u)](\chi_{\mathbb{I}_u} v)\big)\chi_{\mathbb{I}_u} v\,\mathrm{d}\mu\right| \\ &\ge \delta\|v\|^2_{L^2(\mathbb{I}_u)} - \|[\Phi'(u) - \Phi'(\bar u)](\chi_{\mathbb{I}_u} v)\|_{L^2(X)}\|\chi_{\mathbb{I}_u} v\|_{L^2(X)} \ge \frac{\delta}{2}\|v\|^2_{L^2(\mathbb{I}_u)}. \end{aligned}$$

Let us prove that $\chi_{\mathbb{I}_u} v \in E^\tau_{\bar u}$ for every $v \in L^2(\mathbb{I}_u)$ if we take u close enough to $\bar u$. Since $\Phi : \mathcal{A} \longrightarrow L^\infty(X)$ is of class C^1, it is locally Lipschitz. Therefore, there exists $\rho_2 \in (0, \rho_1]$ and a constant $L_{\bar u}$ such that $B_{\rho_2}(\bar u) \subset \mathcal{A}$ and

$$\|\Phi(u_2) - \Phi(u_1)\|_{L^\infty(X)} \le L_{\bar u}\|u_2 - u_1\|_{L^p(X)} \quad \forall u_1, u_2 \in B_{\rho_2}(\bar u). \tag{2.17}$$

Setting $\rho = \min\{\rho_2, \frac{\tau}{L_{\bar u}}\}$ we deduce from the above inequality that

$$B_\rho(\bar u) \subset \mathcal{A} \quad \text{and} \quad \|\Phi(u) - \Phi(\bar u)\|_{L^\infty(X)} < \tau \quad \forall u \in B_\rho(\bar u). \tag{2.18}$$

Given $x \in X$, if $\kappa\bar u(x) + \Phi(\bar u)(x) > \tau$ holds, then (2.5) implies that $\bar u(x) = \alpha$. Thus we have

$$-\frac{1}{\kappa}\Phi(\bar u)(x) < \alpha - \frac{\tau}{\kappa}.$$

This inequality and (2.18) lead to

$$-\frac{1}{\kappa}\Phi(u)(x) \le -\frac{1}{\kappa}\Phi(\bar u)(x) + \frac{1}{\kappa}\|\Phi(u)-\Phi(\bar u)\|_{L^\infty(X)} < \alpha.$$

This yields $x \in \mathbb{A}_u$ and, consequently, $[\chi_{\mathbb{I}_u} v](x) = 0$. Arguing in a similar way we deduce that $[\chi_{\mathbb{I}_u} v](x) = 0$ if $\kappa\bar u(x) + \Phi(\bar u)(x) < -\tau$. Hence, $\chi_{\mathbb{I}_u} v \in E^\tau_{\bar u}$ holds. Therefore, H is strictly convex and coercive and, consequently, M_u is an isomorphism.

Step 2. $\exists C > 0$ *such that* $\|M_u^{-1}\| \le C\ \forall u \in B_\rho(\bar u)$. Multiplying the identity (2.15) by $\kappa\chi_{\mathbb{I}_u}\bar v$ and integrating in X we get

$$\begin{aligned}
\frac{\delta}{2}\|\bar v\|^2_{L^2(\mathbb{I}_u)} &\le J''(u)(\chi_{\mathbb{I}_u}\bar v)^2 = \kappa\|\bar v\|^2_{L^2(\mathbb{I}_u)} + \int_X [\Phi'(u)(\chi_{\mathbb{I}_u}\bar v)]\chi_{\mathbb{I}_u}\bar v \,\mathrm{d}\mu \\
&= \int_X \big(\kappa w - [\Phi'(u)(\chi_{\mathbb{A}_u} w)]\big)\chi_{\mathbb{I}_u}\bar v\,\mathrm{d}\mu \le \Big(\kappa\|w\|_{L^2(X)} + \|\Phi'(u)(\chi_{\mathbb{A}_u} w)\|_{L^2(X)}\Big)\|\bar v\|_{L^2(\mathbb{I}_u)} \\
&\le \Big(\kappa\mu(X)^{\frac{p-2}{2p}}\|w\|_{L^p(X)} + \mu(X)^{\frac{1}{2}}\|\Phi'(u)\chi_{\mathbb{A}_u} w\|_{L^\infty(X)}\Big)\|\bar v\|_{L^2(\mathbb{I}_u)}.
\end{aligned}$$

The Lipschitz property (2.17) implies that

$$\|\Phi'(u)v\|_{L^\infty(X)} \le L_{\bar u}\|v\|_{L^p(X)} \quad \forall u \in B_\rho(\bar u) \text{ and } \forall v \in L^p(X). \tag{2.19}$$

Inserting this inequality in the above estimate it follows that

$$\frac{\delta}{2}\|\bar v\|_{L^2(\mathbb{I}_u)} \le \Big(\kappa\mu(X)^{\frac{p-2}{2p}} + \mu(X)^{\frac{1}{2}}L_{\bar u}\Big)\|w\|_{L^p(X)}.$$

Thus, we get $\|\bar v\|_{L^2(\mathbb{I}_u)} \le C_1\|w\|_{L^p(X)}$ for $C_1 = \frac{2}{\delta}(\kappa\mu(X)^{\frac{p-2}{2p}} + \mu(X)^{\frac{1}{2}}L_{\bar u})$. We use this inequality to estimate $\|\bar v\|_{L^p(\mathbb{I}_u)}$. First, we observe that hypothesis **(H2)** along with (2.16) implies

$$\begin{aligned}
\|\Phi'(u)(\chi_{\mathbb{I}_u}\bar v)\|_{L^p(X)} &\le \|\Phi'(\bar u)(\chi_{\mathbb{I}_u}\bar v)\|_{L^p(X)} + \|[\Phi'(u)-\Phi'(\bar u)](\chi_{\mathbb{I}_u}\bar v)\|_{L^p(X)} \\
&\le \Big(\|\Phi'(\bar u)\|_{\mathcal{L}(L^2(X),L^p(X))} + \frac{\delta}{2\mu(X)^{\frac{p-2}{2p}}}\Big)\|\chi_{\mathbb{I}_u}\bar v\|_{L^2(X)} = C_2\|\chi_{\mathbb{I}_u}\bar v\|_{L^2(X)},
\end{aligned}$$

where C_2 is independent of u. Now, combining (2.15), the above inequality, and (2.19) we infer that

$$\begin{aligned}
\|\chi_{\mathbb{I}_u}\bar v\|_{L^p(X)} &\le \frac{1}{\kappa}\|\Phi'(u)(\chi_{\mathbb{I}_u}\bar v)\|_{L^p(X)} + \|\chi_{\mathbb{I}_u} w\|_{L^p(X)} + \frac{1}{\kappa}\|\Phi'(u)(\chi_{\mathbb{A}_u} w)\|_{L^p(X)} \\
&\le \frac{C_2}{\kappa}\|\chi_{\mathbb{I}_u}\bar v\|_{L^2(X)} + \left(1 + \frac{1}{\kappa}\mu(X)^{\frac{1}{p}}L_{\bar u}\right)\|w\|_{L^p(X)} \\
&\le \Big(\frac{C_1C_2}{\kappa} + 1 + \frac{1}{\kappa}\mu(X)^{\frac{1}{p}}L_{\bar u}\Big)\|w\|_{L^p(X)} = C_3\|w\|_{L^p(X)}.
\end{aligned}$$

Finally, setting $C = 1 + C_3$ and using the first identity of (2.14) we obtain

$$\|\bar v\|_{L^p(X)} \le \|\bar v\|_{L^p(\mathbb{A}_u)} + \|\bar v\|_{L^p(\mathbb{I}_u)} \le \|w\|_{L^p(\mathbb{A}_u)} + C_3\|w\|_{L^p(X)} \le C\|w\|_{L^p(X)}.$$

This proves that $\|M_u^{-1}w\|_{L^p(X)} \le C\|w\|_{L^p(X)}$ for all $u \in B_\rho(\bar u)$ and all $w \in L^p(X)$, which concludes the proof. □

Semismooth Newton's method for problem (P) is detailed in Algorithm 2. Theorems 2.5 and 2.6 yield the following result on its convergence.

Algorithm 2. Semismooth Newton method for (P).

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

repeat

Compute $\varphi_j = \Phi(u_j)$.

Set $\mathbb{I}_j = \{x \in X : \alpha < -\frac{1}{\kappa}\varphi_j(x) < \beta\}$, $\mathbb{A}_j^\alpha = \{x \in X : -\frac{1}{\kappa}\varphi_j(x) \le \alpha\}$, $\mathbb{A}_j^\beta = \{x \in X : -\frac{1}{\kappa}\varphi_j(x) \ge \beta\}$, $\mathbb{A}_j = \mathbb{A}_j^\alpha \cup \mathbb{A}_j^\beta$.

Compute

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

Compute $\eta_j = \chi_{\mathbb{I}_j} w_j - \frac{1}{\kappa}\Phi'(u_j)(\chi_{\mathbb{A}_j} w_j)$.

Solve the quadratic problem

$$(Q_j) \min_{v \in L^2(\mathbb{I}_j)} H_j(v) := \frac{1}{2\kappa}\int_X (\Phi'(u_j)(\chi_{\mathbb{I}_j} v) + \kappa\chi_{\mathbb{I}_j} v)\chi_{\mathbb{I}_j} v \,\mathrm{d}\mu - \int_{\mathbb{I}_j} \eta_j v \,\mathrm{d}\mu$$

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

Set $v_j = \chi_{\mathbb{A}_j} w_j + \chi_{\mathbb{I}_j} v_{\mathbb{I}_j}$

Set $u_{j+1} = u_j + v_j$ and $j = j + 1$.

until *convergence*

COROLLARY 2.7. *Let $\bar{u} \in \mathcal{U}_{ad}$ satisfy* (2.5) *and assume that the second order sufficient condition* (2.7) *and the strict complementarity condition* (2.8) *hold at $\bar{u}$. Then, there exists $\rho > 0$ such that for all $u_0 \in B_\rho(\bar{u}) \subset \mathcal{A}$ the sequence generated by Algorithm* 2 *converges superlinearly to $\bar{u}$.*

Remark 2.8. Note that the assumptions (2.7) and (2.8) are not used directly in the proof of Theorem 2.6. We simply use (2.10), which is a consequence of (2.7) and (2.8); see Theorem 2.3. Therefore, the previous corollary remains valid if we replace the assumptions (2.7) and (2.8) with (2.10). However, the assumptions (2.7) and (2.8) seem to be more natural, easier to verify and are the hypotheses formulated for the analysis of numerical algorithms in finite-dimensional optimization. If (2.8) does not hold, then (2.10) is very far from the necessary second order conditions.

3. Application of Algorithm 2 to solve some control problems. In this section, we show how Algorithm 2 can be applied to solve some control problems. Along this section Ω will denote an open, connected, and bounded subset of $\mathbb{R}^n$ with $1 \le n \le 3$ and a Lipschitz boundary Γ. In case $n = 1$ we assume that Ω is a real interval (a, b) with $-\infty < a < b < +\infty$, and $\Gamma = \{a, b\}$. In Ω we consider a partial differential operator A defined by

$$Ay = -\sum_{i,j=1}^n \partial_{x_j}[a_{ij}\partial_{x_i} y] + a_0 y$$

with $a_0, a_{ij} \in L^\infty(\Omega)$ for $1 \le i, j \le n$, $a_0 \ge 0$. We also assume that there exists $\Lambda > 0$ such that

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq \Lambda|\xi|^2 \text{ for a.e. } x\in\Omega \text{ and all } \xi\in\mathbb{R}^n.$$

3.1. A semilinear elliptic control problem with distributed control. The first control problem is formulated as follows:

$$(\mathrm{P}_1) \qquad \min_{u\in\mathcal{U}_{ad}} J(u) := \mathcal{J}(u) + \frac{\kappa}{2}\int_\Omega u(x)^2\,\mathrm{d}x,$$

where $\kappa > 0$,

$$\mathcal{J}(u) = \int_\Omega L(x, y_u(x))\,\mathrm{d}x \quad \text{ and } \quad \mathcal{U}_{ad} = \{u\in L^2(\Omega) : \alpha \leq u(x) \leq \beta \text{ for a.a. } x\in\Omega\}$$

with $-\infty \leq \alpha < \beta \leq +\infty$. Above y_u denotes the state associated to the control u related by the following semilinear elliptic state equation:

$$\text{(3.1)} \qquad \begin{cases} Ay_u + f(x, y_u) = u & \text{in } \Omega,\\ y_u = 0 \text{ on } \Gamma. \end{cases}$$

The following assumptions are made on f and L:

(A1) We assume that $f : \Omega\times\mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the second variable satisfying the following conditions for almost all $x\in\Omega$:

- $\exists \bar p > \dfrac{n}{2}$ such that $f(\cdot, 0) \in L^{\bar p}(\Omega)$,
- $\dfrac{\partial f}{\partial y}(x, y) \geq 0 \quad \forall y\in\mathbb{R}$,
- $\forall M > 0\ \exists C_M > 0$ such that $\sum_{j=1}^{2} |\frac{\partial^j f}{\partial y^j}(x, y)| \leq C_{f,M}$ for all $|y| \leq M$,
- $\forall\varepsilon > 0$ and $\forall M > 0\ \exists\delta > 0$ such that $|\frac{\partial^2 f}{\partial y^2}(x, y_1) - \frac{\partial^2 f}{\partial y^2}(x, y_2)| \leq \varepsilon$ for all $|y_1|, |y_2| \leq M$ with $|y_1 - y_2| \leq \delta$.

(A2) For the cost functional we suppose that $L : \Omega\times\mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the second variable satisfying the following conditions for almost all $x\in\Omega$:

- $L(\cdot, 0) \in L^1(\Omega)$ and $\forall M > 0\ \exists\Psi_{L,M} \in L^{\bar p}(\Omega)$ and $C_{L,M} > 0$ such that
$$\left|\frac{\partial L}{\partial y}(x, y)\right| \leq \Psi_{L,M}(x) \text{ and } \left|\frac{\partial^2 L}{\partial y^2}(x, y)\right| \leq C_{L,M} \text{ for all } |y| \leq M,$$
- $\forall\varepsilon > 0$ and $\forall M > 0\ \exists\delta > 0$ such that $|\frac{\partial^2 L}{\partial y^2}(x, y_1) - \frac{\partial^2 L}{\partial y^2}(x, y_2)| \leq \varepsilon$ for all $|y_1|, |y_2| \leq M$ with $|y_1 - y_2| < \delta$.

Let us consider the Banach space $Y = H_0^1(\Omega)\cap C(\bar\Omega)$. Under the above assumptions, it is known that (3.1) has a unique solution $y_u \in Y$ for every $u\in L^2(\Omega)$. The mapping $G : L^2(\Omega) \longrightarrow Y$, given by $G(u) = y_u$, is of class C^2. Furthermore, for all $u, v \in L^2(\Omega)$, $z_{u,v} = G'(u)v$ is the unique solution to

$$\begin{cases} Az + \dfrac{\partial f}{\partial y}(x, y_u)z = v & \text{in } \Omega,\\ z = 0 & \text{on } \Gamma, \end{cases}$$

and, given $v_1, v_2 \in L^2(\Omega)$, $z_{u,(v_1,v_2)} = G''(u)(v_1, v_2)$ is the unique solution to

$$\begin{cases} Az + \dfrac{\partial f}{\partial y}(x, y_u)z = -\dfrac{\partial^2 f}{\partial y^2}(x, y_u) z_{v_1} z_{v_2} & \text{in } \Omega, \\ z = 0 \quad \text{on } \Gamma, \end{cases}$$

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

Further, for every $u \in L^2(\Omega)$ the adjoint state equation

$$\begin{cases} A^*\varphi + \dfrac{\partial f}{\partial y}(x, y_u)\varphi = \dfrac{\partial L}{\partial y}(x, y_u) & \text{in } \Omega, \\ \varphi = 0 \text{ on } \Gamma \end{cases}$$

has a unique solution $\varphi_u \in Y$. The mapping $\Phi : L^2(\Omega) \longrightarrow Y$ defined by $\Phi(u) = \varphi_u$ is of class C^1 and $\eta_v = \Phi'(u)v \in Y$ is the solution of the linear equation

$$\begin{cases} A^*\eta_v + \dfrac{\partial f}{\partial y}(x, y_u)\eta_v = \left(\dfrac{\partial^2 L}{\partial y^2}(x, y_u) - \varphi_u \dfrac{\partial^2 f}{\partial y^2}(x, y_u) \right) z_v & \text{in } \Omega, \\ \eta_v = 0 \quad \text{on } \Gamma. \end{cases}$$

We also have that the functional $\mathcal{J} : L^2(\Omega) \longrightarrow \mathbb{R}$ is of class C^2 and for every $u, v, v_1, v_2 \in L^2(\Omega)$ the following identities hold:

$$\begin{aligned} \mathcal{J}'(u)v &= \int_\Omega \varphi_u v \, \mathrm{d}x = \int_\Omega \Phi(u)v \, \mathrm{d}x, \\ \mathcal{J}''(u)(v_1, v_2) &= \int_\Omega \left(\frac{\partial^2 L}{\partial y^2}(x, y_u) - \varphi_u \frac{\partial^2 f}{\partial y^2}(x, y_u) \right) z_{v_1} z_{v_2} \, \mathrm{d}x \\ &= \int_\Omega \eta_{v_1} v_2 \, \mathrm{d}x = \int_\Omega \eta_{v_2} v_1 \, \mathrm{d}x, \end{aligned}$$

where $\varphi_u = \Phi(u)$, $z_{v_i} = G'(u)v_i$, and $\eta_{v_i} = \Phi'(u)v_i$ for $i = 1, 2$.

Any local minimizer $\bar{u}$ of (P_1) satisfies the identity

$$\bar{u}(x) = \mathrm{Proj}_{[\alpha,\beta]}\left(-\frac{1}{\kappa}\bar{\varphi}(x) \right) = \mathrm{Proj}_{[\alpha,\beta]}\left(-\frac{1}{\kappa}\Phi(\bar{u})(x) \right) \text{ for a.a. } x \in \Omega$$

or, equivalently, $\mathcal{F}(\bar{u}) = 0$ with $\mathcal{F}(u) = u - \Psi(\Phi(u))$, Ψ as defined at the beginning of subsection 2.2.

The reader is referred to [4] for the proof of the above statements.

We apply Algorithm 2 to compute a local minimizer $\bar{u}$. To this end we identify the following elements: $X = \Omega$, μ is the Lebesgue measure, $p = 2$, $\mathcal{A} = L^2(\Omega)$. Since $Y \subset L^\infty(\Omega)$ and it is compactly embedded in $L^2(\Omega)$, hypotheses **(H1)** and **(H2)** hold. Then, Algorithm 3 is the application of Algorithm 2 to (P_1).

Under assumptions (2.7) and (2.8), Corollary 2.7 implies that Algorithm 3 is locally and superlinearly convergent to the local minimizer $\bar{u}$.

3.2. A semilinear elliptic bilinear control problem with boundary control. In this section, we consider the following control problem:

$$(\mathrm{P}_2) \qquad \min_{u \in \mathcal{U}_{ad}} J(u) := \mathcal{J}(u) + \frac{\kappa}{2}\int_\Gamma u(x)^2 \, \mathrm{d}x,$$

where $\kappa > 0$,

$$\mathcal{J}(u) = \int_\Omega L(x, y_u(x)) \, \mathrm{d}x \quad \text{and} \quad \mathcal{U}_{ad} = \{u \in L^2(\Gamma) : \alpha \le u(x) \le \beta \text{ for a.a. } x \in \Gamma\}$$

Algorithm 3. Semismooth Newton method for (P_1).

Initialize. Choose $u_0 \in L^2(\Omega)$. Set $j = 0$.

repeat

Compute $y_j = G(u_j)$ solving the nonlinear equation

$$Ay_j + f(x, y_j) = u_j \text{ in } \Omega,\ y_j = 0 \text{ in } \Gamma$$

Compute $\varphi_j = \Phi(u_j)$ solving the linear equation

$$A^*\varphi_j + \frac{\partial f}{\partial y}(x, y_j)\varphi_j = \frac{\partial L}{\partial y}(x, y_j) \text{ in } \Omega,\ \varphi_j = 0 \text{ in } \Gamma$$

Set $\mathbb{I}_j = \{x \in \Omega : \alpha < -\frac{1}{\kappa}\varphi_j(x) < \beta\}$, $\mathbb{A}_j^\alpha = \{x \in \Omega : -\frac{1}{\kappa}\varphi_j(x) \le \alpha\}$, $\mathbb{A}_j^\beta = \{x \in \Omega : -\frac{1}{\kappa}\varphi_j(x) \ge \beta\}$, $\mathbb{A}_j = \mathbb{A}_j^\alpha \cup \mathbb{A}_j^\beta$.

Compute

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

Compute $\eta_j = \eta_{\chi_{\mathbb{A}_j} w_j}$ solving the linear equations

$$\begin{aligned} Az_j + \frac{\partial f}{\partial y}(x, y_j)z_j &= \chi_{\mathbb{A}_j} w_j \text{ in } \Omega,\ z_j = 0 \text{ on } \Gamma, \\ A^*\eta_j + \frac{\partial f}{\partial y}(x, y_j)\eta_j &= \left(\frac{\partial^2 L}{\partial y^2}(x, y_j) - \varphi_j \frac{\partial^2 f}{\partial y^2}(x, y_j)\right) z_j \text{ in } \Omega,\ \eta_j = 0 \text{ on } \Gamma. \end{aligned}$$

Solve the quadratic problem

$$(Q_j) \qquad \min_{v \in L^2(\mathbb{I}_j)} H_j(v) := \frac{1}{2\kappa}\mathcal{J}''(u_j)(\chi_{\mathbb{I}_j} v)^2 + \int_{\mathbb{I}_j} (\frac{1}{2}v - w_j + \frac{1}{\kappa}\eta_j)v \, \mathrm{d}\mu$$

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

Set $v_j = \chi_{\mathbb{A}_j} w_j + \chi_{\mathbb{I}_j} v_{\mathbb{I}_j}$

Set $u_{j+1} = u_j + v_j$ and $j = j + 1$.

until *convergence*

with $0 \le \alpha < \beta \le \infty$. We assume that $\beta < \infty$ if $n = 3$. Here y_u is the solution of the state equation

$$\begin{cases} Ay + f(x, y) = 0 & \text{in } \Omega, \\ \partial_{n_A} y + uy = g & \text{on } \Gamma. \end{cases} \tag{3.2}$$

Regarding this problem, the reader is referred to [2, 3] for all unproven claims set out below.

We suppose that the functions L and f satisfy the same assumptions (A1) and (A2) introduced in section 3.1. Additionally we assume that

(A3) $n=2$ or 3, $a_0 \not\equiv 0$, $g \in L^q(\Gamma)$ with $q > n-1$ and, without loss of generality, $q \le n$.

We set $p=2$ if $n=2$ and $p=q$ if $n=3$. Under assumptions (A1), (A2), and (A3), there exists an open set $\mathcal{A}$ of $L^p(\Gamma)$ such that $\mathcal{A} \supset \{u \in L^p(\Gamma) : u \ge 0\}$ and (3.2) has a unique solution $y_u \in Y = H^1(\Omega) \cap C(\bar\Omega)$ for every $u \in \mathcal{A}$. Moreover, 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 $\forall v, v_1, v_2 \in L^p(\Gamma)$ the functions $z_{u,v} = G'(u)v$ and $z_{u,(v_1,v_2)} = G''(u)(v_1, v_2)$ are the unique solutions of the equations:

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

$$\begin{cases} Az + \dfrac{\partial f}{\partial y}(x, y_u)z = -\dfrac{\partial^2 f}{\partial y^2}(x, y_u)z_{u,v_1}z_{u,v_2} \ \text{ in } \Omega, \\ \partial_{n_A} z + uz = -v_1 z_{u,v_2} - v_2 z_{u,v_1} \ \text{ on } \Gamma, \end{cases}$$

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

An application of the chain rule implies that the function $\mathcal{J} : \mathcal{A} \longrightarrow \mathbb{R}$ is of class C^2 and its derivatives are given by the expressions:

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

$$\mathcal{J}''(u)(v_1, v_2) = \int_\Omega \left[\frac{\partial^2 L}{\partial y^2}(x, y_u) - \varphi_u \frac{\partial^2 f}{\partial y^2}(x, y_u)\right] z_{u,v_1} z_{u,v_2} \,\mathrm{d}x - \int_\Gamma [v_1 z_{u,v_2} + v_2 z_{u,v_1}]\varphi_u \,\mathrm{d}x$$

for all $u \in \mathcal{A}$ and all $v, v_1, v_2 \in L^p(\Gamma)$, 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

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

Now, we introduce the mapping $S : \mathcal{A} \longrightarrow Y$ given by $S(u) = \varphi_u$. This mapping is of class C^1 and for all $u \in \mathcal{A}$ and $v \in L^p(\Gamma)$ the function $\eta_{u,v} = S'(u)v$ is the unique solution of

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

where $z_{u,v} = G'(u)v$. Now, we define the mapping $\Phi : \mathcal{A} \longrightarrow L^\infty(\Gamma)$ by $\Phi(u) = -(y_u\varphi_u)_{|\Gamma} = -[G(u)S(u)]_{|\Gamma}$. We have that Φ is of class C^1 and the following alternative expressions for the derivatives of $\mathcal{J}$ hold:

$$\mathcal{J}'(u)v = \int_\Gamma \Phi(u)v \,\mathrm{d}x,$$

$$\mathcal{J}''(u)(v_1, v_2) = \int_\Gamma [\Phi'(u)v_1]v_2 \,\mathrm{d}x = \int_\Gamma [\Phi'(u)v_2]v_1 \,\mathrm{d}x = -\int_\Gamma (\varphi_u z_{u,v_1} + y_u \eta_{u,v_1})v_2 \,\mathrm{d}x = -\int_\Gamma (\varphi_u z_{u,v_2} + y_u \eta_{u,v_2})v_1 \,\mathrm{d}x.$$

Looking at the equations satisfied by $z_{u,v}$ and $\eta_{u,v}$ we immediately infer that $G'(u)$ and $S'(u)$ are linear and continuous mappings form $L^2(\Gamma)$ to $H^1(\Omega)$. Moreover, using that $y_u, \varphi_u \in Y$ for every $u \in L^p(\Gamma)$, we get that $[G'(u)v]S(u) + G(u)[S'(u)v] = z_{u,v}\varphi_u + y_u\eta_{u,v} \in H^1(\Omega)$. Hence, $\Phi'(u)v = -([G'(u)v]S(u) + G(u)[S'(u)v])_{|\Gamma}$ defines a linear and continuous mapping from $L^2(\Gamma)$ to $H^{\frac{1}{2}}(\Gamma)$. Since $p \le n$, then $H^{\frac{1}{2}}(\Gamma)$ is compactly embedded in $L^p(\Gamma)$ and, consequently, $\Phi'(u) : L^2(\Gamma) \longrightarrow L^p(\Gamma)$ is a compact linear mapping.

(P_2) has at least a global minimizer. Moreover, any local minimizer satisfies the identity

$$\bar{u}(x) = \mathrm{Proj}_{[\alpha,\beta]}\Big(\frac{1}{\kappa}\bar{y}(x)\bar{\varphi}(x)\Big) = \mathrm{Proj}_{[\alpha,\beta]}\Big(-\frac{1}{\kappa}\Phi(\bar{u})(x)\Big) \text{ for a.a. } x \in \Gamma,$$

where $\bar{y} = G(\bar{u})$ and $\bar{\varphi} = S(\bar{u})$.

Problem (P_2) fits into the abstract framework of (P) by taking $X = \Gamma$ and μ equal to the $n-1$-dimensional measure on Γ. As a straightforward consequence of the above statements we get that hypotheses **(H1)** and **(H2)** hold. A detailed proof of inequality (2.3) in the case $n = 3$ can be found in [3, Lemma A.10]. Under assumptions (2.7) and (2.8), Corollary 2.7 implies that the semismooth Newton method applied to (P_2) is locally and superlinearly convergent to the local minimizer $\bar{u}$.

3.3. A semilinear parabolic control problem with distributed control. Now, we analyze the following control problem:

$$(\mathrm{P}_3) \qquad \min_{u \in \mathcal{U}_{ad}} J(u) := \mathcal{J}(u) + \frac{\kappa}{2}\int_Q u(x,t)^2 \,\mathrm{d}x\,\mathrm{d}t \quad (\kappa > 0),$$

where $\mathcal{U}_{ad} = \{u \in L^2(Q) : \alpha \le u(x,t) \le \beta \text{ for a.a. } (x,t) \in Q\}$, $-\infty < \alpha < \beta < +\infty$ and

$$\mathcal{J}(u) = \int_Q L(x,t,y_u(x,t))\,\mathrm{d}x\,\mathrm{d}t.$$

Here y_u denotes the solution of the state equation

$$(3.3) \qquad \begin{cases} \dfrac{\partial y}{\partial t} + Ay + f(x,t,y) = u & \text{in } Q = \Omega \times (0,T), \\ y = 0 \quad \text{on } \Sigma = \Gamma \times (0,T), \ y(0) = y_0 & \text{in } \Omega. \end{cases}$$

We make the following assumptions on the data of (P_3):

(A4) $y_0 \in L^\infty(\Omega)$.

(A5) We assume that $f : Q \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the second variable satisfying the following conditions for almost all $(x,t) \in Q$:

- $\exists q, r \ge 2$ such that $\dfrac{1}{r} + \dfrac{n}{2q} < 1$ and $f(\cdot,\cdot,0) \in L^r(0,T;L^q(\Omega))$,
- $\exists C_f \in \mathbb{R}$ such that $\dfrac{\partial f}{\partial y}(x,t,y) \ge C \quad \forall y \in \mathbb{R}$,
- $\forall M > 0,\ \exists C_M > 0$ such that $\sum_{j=1}^2 |\frac{\partial^j f}{\partial y^j}(x,t,y)| \le C_{f,M}$ for all $|y| \le M$,
- $\forall \varepsilon > 0$ and $\forall M > 0$ $\exists \delta > 0$ such that $|\frac{\partial^2 f}{\partial y^2}(x,t,y_1) - \frac{\partial^2 f}{\partial y^2}(x,t,y_2)| \le \varepsilon$ for all $|y_1|, |y_2| \le M$ with $|y_1 - y_2| \le \delta$.

(A6) For the cost functional we suppose that $L: Q \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the second variable satisfying the following conditions for almost all $(x,t) \in Q$:

- $L(\cdot,\cdot,0) \in L^1(Q)$ and $\forall M > 0$ $\exists \Psi_{L,M} \in L^r(0,T;L^q(\Omega))$ and $C_{L,M} > 0$ satisfying
$$\left|\frac{\partial L}{\partial y}(x,t,y)\right| \le \Psi_{L,M}(x,t) \quad \text{and} \quad \left|\frac{\partial^2 L}{\partial y^2}(x,t,y)\right| \le C_{L,M} \text{ for all } |y| \le M,$$
- $\forall \varepsilon > 0$ and $\forall M > 0$, $\exists \delta > 0$ such that $|\frac{\partial^2 L}{\partial y^2}(x,t,y_1) - \frac{\partial^2 L}{\partial y^2}(x,t,y_2)| \le \varepsilon$ for all $|y_1|, |y_2| \le M$ with $|y_1 - y_2| < \delta$.

The reader is referred to [7] for the unproven statements of this section.

Let us consider the Banach space $Y = W(0,T) \cap L^\infty(Q)$, where $W(0,T) = L^2(0,T;H_0^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega))$. Under the above assumptions, it is known that (3.3) has a unique solution $y_u \in Y$ for every $u \in L^p(Q)$ with $p = 2$ if $n = 1$ and $p > 1 + \frac{n}{2}$ if $n = 2$ or 3. The mapping $G: L^p(Q) \longrightarrow Y$, given by $G(u) = y_u$, is of class C^2. Furthermore, for all $u, v \in L^p(Q)$, $z_{u,v} = G'(u)v$ is the unique solution to

$$\text{(3.4)} \qquad \begin{cases} \dfrac{\partial z}{\partial t} + Az + \dfrac{\partial f}{\partial y}(x,t,y_u)z = v & \text{in } Q, \\ z = 0 \ \text{ on } \Sigma, \quad z(0) = 0 & \text{in } \Omega, \end{cases}$$

and, given $v_1, v_2 \in L^p(Q)$, $z_{u,(v_1,v_2)} = G''(u)(v_1,v_2)$ is the unique solution to

$$\text{(3.5)} \qquad \begin{cases} \dfrac{\partial z}{\partial t} + Az + \dfrac{\partial f}{\partial y}(x,t,y_u)z = -\dfrac{\partial^2 f}{\partial y^2}(x,t,y_u)z_{v_1}z_{v_2} & \text{in } Q, \\ z = 0 \ \text{ on } \Sigma, \quad z(0) = 0 & \text{in } \Omega, \end{cases}$$

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

From assumption (A5) we infer that $\frac{\partial f}{\partial y}(\cdot,\cdot,y_u) \in L^\infty(Q)$ for all $u \in L^p(Q)$. Hence, it is well known that (3.4) has a unique solution $z_v \in H^1(Q)$ for every $v \in L^2(Q)$; see, for instance, [12, section III.2]. Hence, $G'(u): L^2(Q) \longrightarrow H^1(Q)$ is a continuous linear extension of the mapping $G'(u): L^p(Q) \longrightarrow Y$. Moreover, taking into account that $H^1(Q) \subset L^4(Q)$ for $1 \le n \le 3$, and using again assumption (A5), we deduce that $\frac{\partial^2 f}{\partial y^2}(\cdot,\cdot,y_u)z_{v_1}z_{v_2} \in L^2(Q)$. Consequently, we have that $z_{u,(v_1,v_2)} \in H^1(Q)$ and $G''(u): L^2(Q) \times L^2(Q) \longrightarrow H^1(Q)$ is a bilinear continuous extension of $G''(u): L^p(Q) \times L^p(Q) \longrightarrow Y$.

Further, for every $u \in L^p(\Omega)$ the adjoint state equation

$$\text{(3.6)} \qquad \begin{cases} -\dfrac{\partial \varphi}{\partial t} + A^*\varphi + \dfrac{\partial f}{\partial y}(x,t,y_u)\varphi = \dfrac{\partial L}{\partial y}(x,t,y_u) & \text{in } Q, \\ \varphi = 0 \text{ on } \Sigma \quad \varphi(T) = 0 & \text{in } \Omega, \end{cases}$$

has a unique solution $\varphi_u \in Y$. We define the mapping $\Phi: L^p(Q) \longrightarrow Y$ by $\Phi(u) = \varphi_u$.

From the chain rule we infer that the functional $\mathcal{J}: L^p(Q) \longrightarrow \mathbb{R}$ is of class C^2 and for every $u, v, v_1, v_2 \in L^p(Q)$ the following identities hold:

$$\text{(3.7)} \qquad \mathcal{J}'(u)v = \int_Q \varphi_u v \,\mathrm{d}x\,\mathrm{d}t = \int_Q \Phi(u)v \,\mathrm{d}x\,\mathrm{d}t,$$

$$\text{(3.8)} \qquad \mathcal{J}''(u)(v_1,v_2) = \int_Q \left(\frac{\partial^2 L}{\partial y^2}(x,t,y_u) - \varphi_u \frac{\partial^2 f}{\partial y^2}(x,t,y_u)\right) z_{v_1}z_{v_2} \,\mathrm{d}x\,\mathrm{d}t,$$

where $z_{v_i} = G'(u)v_i$. From the above comments on the extensions of $G'(u)$ and $G''(u)$ we deduce that $\mathcal{J}'(u)$ and $\mathcal{J}''(u)$ can be extended to continuous linear and bilinear forms on $L^2(Q)$ given by the same integral expressions written above.

The following theorem provides some important properties of Φ.

THEOREM 3.1. *The mapping Φ enjoys the following properties:*

(i) *Φ is of class C^1 and $\eta_v = \Phi'(u)v \in Y$ is the solution of the linear equation*

$$\begin{cases} -\dfrac{\partial \eta_v}{\partial t} + A^*\eta_v + \dfrac{\partial f}{\partial y}(x,t,y_u)\eta_v = \left(\dfrac{\partial^2 L}{\partial y^2}(x,t,y_u) - \varphi_u \dfrac{\partial^2 f}{\partial y^2}(x,t,y_u)\right) z_v \text{ in } Q, \\ \eta_v = 0 \quad \text{on } \Sigma, \eta_v(T) = 0 \quad \text{in } \Omega. \end{cases} \tag{3.9}$$

(ii) *For $1 + \frac{n}{2} < p < 4$ and $u \in L^p(Q)$ the linear mapping $\Phi'(u) : L^p(Q) \longrightarrow Y$ has a unique extension to a compact operator $\Phi'(u) : L^2(Q) \longrightarrow Y$ and, for every $v \in L^2(Q)$, $\eta_v = \Phi'(u)v \in H^1(Q)$ solves* (3.9).

(iii) *Taking p as in* (ii) *and given $u \in L^p(Q)$, for every $\varepsilon > 0$ there exists $\rho > 0$ such that*

$$\|[\Phi'(w) - \Phi'(u)]v\|_{L^p(Q)} \le \varepsilon \|v\|_{L^2(Q)} \quad \forall w \in B_\rho(u) \text{ and } \forall v \in L^2(Q), \tag{3.10}$$

where $B_\rho(u)$ denotes the ball with respect to the $L^p(Q)$-norm.

(iv) *The following identities hold:*

$$\mathcal{J}''(u)(v_1, v_2) = \int_Q \eta_{v_1} v_2 \, dx\, dt = \int_Q \eta_{v_2} v_1 \, dx\, dt, \tag{3.11}$$

where $\eta_{v_i} = \Phi'(u)v_i$ for $i = 1, 2$.

The proof of this theorem is given in the appendix.

(P_3) has at least a global solution. Moreover, any local solution satisfies the identity

$$\bar{u}(x,t) = \text{Proj}_{[\alpha,\beta]}\Big(-\frac{1}{\kappa}\bar{\varphi}(x,t)\Big) = \text{Proj}_{[\alpha,\beta]}\Big(-\frac{1}{\kappa}\Phi(\bar{u})(x,t)\Big) \text{ for a.a. } (x,t) \in Q.$$

Problem (P_3) falls into the abstract framework for (P). It is enough to set $X = Q$, μ equal to the Lebesgue measure in Q, $p = 2$ if $n = 1$ and $1 + \frac{n}{2} < p < 4$ if $n = 2$ or 3, and $\mathcal{A} = L^p(Q)$. Since the continuous embeddings $Y \subset L^\infty(Q) \subset L^p(Q)$ hold, then hypotheses **(H1)** and **(H2)** are consequences of Theorem 3.1. Therefore, under the assumptions (2.7) and (2.8) the superlinear convergence of the semismooth Newton method follows from Corollary 2.7.

Remark 3.2. Often the control is located in a small region of the domain Ω or the boundary Γ. Assume that ω is a measurable subset of Ω or Γ with nonzero measure. The case of an elliptic control problem with controls located in ω fits into the abstract framework by setting $X = \omega$ and $\Phi(u) = \varphi_{u|_\omega}$, or $\Phi(u) = -(y_u\varphi_u)_{|_\omega}$ in the case of a boundary bilinear control. For parabolic control problems we set $X = \omega \times (0,T)$.

Appendix A. Proof of Theorem 3.1.

Proof of (i). We apply the implicit function theorem. To this end we introduce the space $Y_\varphi = \{\varphi \in Y : -\frac{\partial \varphi}{\partial t} + A^*\varphi \in L^\infty(Q) \text{ and } \varphi(T) = 0\}$. Endowed with the graph norm this is a Banach space. Now, we define the mapping

$$F : Y_\varphi \times L^p(Q) \longrightarrow L^\infty(Q),$$
$$F(\varphi, u) = -\frac{\partial \varphi}{\partial t} + A^*\varphi + \frac{\partial f}{\partial y}(\cdot,\cdot,y_u)\varphi - \frac{\partial L}{\partial y}(\cdot,\cdot,y_u).$$

Since the mapping $G : L^p(Q) \longrightarrow L^\infty(Q)$ is of class C^1, it is a straightforward application of the chain rule and the assumptions (A5) and (A6) that F is of class C^1 and

$$\frac{\partial F}{\partial \varphi}(\varphi, u)\eta = -\frac{\partial \eta}{\partial t} + A^*\eta + \frac{\partial f}{\partial y}(\cdot,\cdot,y_u)\eta,$$
$$\frac{\partial F}{\partial u}(\varphi, u)v = \frac{\partial^2 f}{\partial y^2}(\cdot,\cdot,y_u)z_v\varphi - \frac{\partial^2 L}{\partial y^2}(\cdot,\cdot,y_u)z_v,$$

where $z_v = G'(u)v$. It is obvious that $\frac{\partial F}{\partial \varphi}(\varphi, u) : Y_\varphi \longrightarrow L^\infty(Q)$ is an isomorphism. Moreover, we have that $\Phi(u) \in Y_\varphi$ for every $u \in L^p(Q)$ and $F(\Phi(u), u) = 0$. Then, the implicit function theorem implies that Φ is of class C^1 and using the expressions for $\frac{\partial F}{\partial \varphi}(\varphi, u)$ and $\frac{\partial F}{\partial u}(\varphi, u)$ we get that $\eta_v = \Phi'(u)v$ is the solution of (3.7).

Proof of (ii). Given $v \in L^2(Q)$ we know that $z_v \in H^1(Q)$ and the embedding $H^1(Q) \subset L^p(Q)$ is compact for $p < 4$. Hence, the operator $G'(u) : L^2(Q) \longrightarrow L^p(Q)$ is compact. From (3.9) and this compactness property we deduce that the linear mapping $\Phi'(u) : L^2(Q) \longrightarrow Y$ is compact as well.

Proof of (iii). Since $G, \Phi : L^p(Q) \longrightarrow Y$ are continuous mappings, given $u \in L^p(Q)$ and $\varepsilon_1 > 0$ there exists $\rho_1 > 0$ such that

$$\|y_w - y_u\|_Y + \|\varphi_w - \varphi_u\|_Y < \varepsilon_1 \quad \forall w \in B_{\rho_1}(u) \subset L^p(Q). \tag{A.1}$$

This leads to the existence of $M_1 > 0$ such that

$$\|y_w\|_{L^\infty(Q)} + \|\varphi_w\|_{L^\infty(Q)} \le M_1 \quad \forall w \in B_{\rho_1}(u). \tag{A.2}$$

Using this, we infer from (3.4) and assumption (A5) that

$$\|z_{w,v}\|_{L^p(Q)} = \|G'(w)v\|_{L^p(Q)} \le C_1\|G'(w)v\|_{H^1(Q)} \le C_2\|v\|_{L^2(Q)} \tag{A.3}$$

for all $(w, v) \in B_{\rho_1}(u) \times L^2(Q)$. Now, for every $w \in L^p(Q)$ we denote

$$R(w) = \frac{\partial^2 L}{\partial y^2}(\cdot,\cdot,y_w) - \varphi_w \frac{\partial^2 f}{\partial y^2}(\cdot,\cdot,y_w).$$

With assumptions (A5) and (A6) and (A.2) we deduce for every $w \in B_{\rho_1}(u)$ and all $v \in L^2(Q)$ that

$$\|R(w)\|_{L^\infty(Q)} \le M_2 = C_{L,M_1} + M_1 C_{f,M_1} \tag{A.4}$$

and with (A.3)

$$\|R(w)z_{w,v}\|_{L^p(Q)} \le M_2 C_2 \|v\|_{L^2(Q)}. \tag{A.5}$$

Then, from (3.9) we get that

$$\|\eta_{w,v}\|_Y = \|\Phi'(w)v\|_Y \le C_3\|R(w)z_{w,v}\|_{L^p(Q)} \le C_3 M_2 C_2 \|v\|_{L^2(Q)} \tag{A.6}$$

for every $w \in B_{\rho_1}(u)$ and all $v \in L^2(Q)$.

Now, setting $z = z_{u,v} - z_{w,v} = G'(u)v - G'(w)v$ and subtracting the equations satisfied by $z_{u,v}$ and $z_{w,v}$ we get

$$\begin{cases} \dfrac{\partial z}{\partial t} + Az + \dfrac{\partial f}{\partial y}(x,t,y_u)z = \Big[\dfrac{\partial f}{\partial y}(x,t,y_w) - \dfrac{\partial f}{\partial y}(x,t,y_u)\Big] z_{w,v} \ \text{ in } Q, \\ z = 0 \ \text{ on } \Sigma, \quad z(0) = 0 \ \text{ in } \Omega. \end{cases}$$

Using assumption (A5), (A.1), (A.3), and (A.4) we get for every $w \in B_{\rho_1}(u)$ and all $v \in L^2(Q)$ that

$$\begin{aligned}\|z\|_{L^p(Q)} &\le C_1\|z\|_{H^1(Q)} \le C_4\Big\|\frac{\partial f}{\partial y}(\cdot,\cdot,y_w) - \frac{\partial f}{\partial y}(\cdot,\cdot,y_u)\Big\|_{L^\infty(Q)}\|z_{w,v}\|_{L^p(Q)}\\ &\le C_4C_{f,M_1}\|y_w - y_u\|_{L^\infty(Q)}C_2\|v\|_{L^2(Q)} \le C_2C_4C_{f,M_1}\varepsilon_1\|v\|_{L^2(Q)}.\end{aligned}$$

Setting $C_5 = C_2C_4C_{f,M_1}$ we obtain

$$\|z_{u,v} - z_{w,v}\|_{L^p(Q)} \le C_5\varepsilon_1\|v\|_{L^2(Q)} \ \forall w \in B_{\rho_1}(u) \text{ and } \forall v \in L^2(Q). \tag{A.7}$$

Finally, we set $\eta = \eta_{u,v} - \eta_{w,v} = [\Phi'(u) - \Phi'(w)]v$. Subtracting the corresponding equations we get

$$\begin{aligned}-\frac{\partial \eta}{\partial t} + A^*\eta + \frac{\partial f}{\partial y}(x,t,y_u)\eta = &\Big[\frac{\partial f}{\partial y}(\cdot,\cdot,y_w) - \frac{\partial f}{\partial y}(\cdot,\cdot,y_u)\Big]\eta_{w,v}\\ &+ [R(u) - R(w)]z_{u,v} + R(w)(z_{u,v} - z_{w,v}).\end{aligned}$$

Then, using (A.3), assumption (A5), (A.4), (A.6), and (A.7) we get

$$\begin{aligned}\|\eta\|_Y \le C_3\Big\{&\Big\|\frac{\partial f}{\partial y}(\cdot,\cdot,y_w) - \frac{\partial f}{\partial y}(\cdot,\cdot,y_u)\Big\|_{L^\infty(Q)}\|\eta_{w,v}\|_{L^p(Q)}\\ &+ \|R(u) - R(w)\|_{L^\infty(Q)}\|z_{u,v}\|_{L^p(Q)} + \|R(w)\|_{L^\infty(Q)}\|z_{u,v} - z_{w,v}\|_{L^p(Q)}\Big\}\\ \le C_3\Big\{&C_{f,M_1}\varepsilon_1C_3M_2C_2|Q|^{\frac{1}{p}}\|v\|_{L^2(Q)}\\ &+ \|R(u) - R(w)\|_{L^\infty(Q)}C_2\|v\|_{L^2(Q)} + M_2C_5\varepsilon_1\|v\|_{L^2(Q)}\Big\}.\end{aligned}$$

From assumptions (A5) and (A6) and inequality (A.1) we infer the existence of $\rho \in (0, \rho_1]$ such that

$$\|R(u) - R(w)\|_{L^\infty(Q)} \le \varepsilon_1 \quad \forall w \in B_\rho(u).$$

Inserting this inequality above and selecting ε_1 small enough we deduce (3.10).

Proof of (iv). Equalities (3.11) are a straightforward consequence of (3.8) and (3.9).

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., 56 (2013), pp. 369–403.

[2] E. Casas, K. Chrysafinos, and M. Mateos, *Semismooth Newton method for boundary bilinear control*, IEEE Control Syst. Lett., 7 (2023), pp. 3549–3554.

[3] E. Casas, K. Chrysafinos, and M. Mateos, *Semismooth Newton Method for Boundary Bilinear Control*, preprint, https://arxiv.org/abs/2403.01135v1, 2024.

[4] E. Casas and M. Mateos, *Convergence Analysis of the Semismooth Newton Method for Sparse Control Problems Governed by Semilinear Elliptic Equations*, preprint, https://arxiv.org/abs/2309.05393, 2024.

[5] E. Casas and F. Tröltzsch, *A general theorem on error estimates with application to elliptic optimal control problems*, Comput. Optim. Appl., 53 (2012), pp. 173–206.

[6] E. Casas and F. Tröltzsch, *Second order analysis for optimal control problems: Improving results expected from abstract theory*, SIAM J. Optim., 22 (2012), pp. 261–279, https://doi.org/10.1137/110840406.

[7] E. Casas and F. Tröltzsch, *Second order optimality conditions for weak and strong local solutions of parabolic optimal control problems*, Vietnam J. Math., 44 (2016), pp. 181–202.

[8] M. Hintermüller, K. Ito, and K. Kunisch, *The primal-dual active set strategy as a semismooth Newton method*, SIAM J. Optim., 13 (2003), pp. 865–888, https://doi.org/10.1137/S1052623401383558.

[9] F. Mannel and A. Rund, *A hybrid semismooth quasi-Newton method for nonsmooth optimal control with PDEs*, Optim. Eng., 22 (2021), pp. 2087–2125.

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

[11] K. Pieper, *Finite Element Discretization and Efficient Numerical Solution of Elliptic and Parabolic Sparse Control problems*, Ph.D. thesis, Technische Universität München, München, Germany, 2015.

[12] R. E. Showalter, *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*, Math. Surv. and Monogr. 49, American Mathematical Society, Providence, RI, 1997.

[13] G. Stadler, *Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices*, Comput. Optim. Appl., 44 (2009), pp. 159–181.

[14] M. Ulbrich, *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*, MOS-SIAM Ser. Optim. 11, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Philadelphia, 2011, https://doi.org/10.1137/1.9781611970692.