Hamburger Beiträge
zur Angewandten Mathematik

An Error Estimate for the Variational Discretization of Semilinear Optimal Control Problems in the Presence of Pointwise Control and State Constraints

Morten Vierling

Nr. 2009-11
July 2009

An Error Estimate for the Variational Discretization of Semilinear Optimal Control Problems in the Presence of Pointwise Control and State Constraints

Morten Vierling*

Abstract: We consider a semilinear elliptic optimal control problem with pointwise control and state constraints. The problem is reformulated by means of $W^{1,p}(\Omega)$ instead of $C(\bar{\Omega})$, and a discretization of the state equation yields a sequence of optimal control problems. While the controls are not discretized, solutions of the first order necessary conditions for these problems can be computed. A linearized Slater condition, strict complementarity and a second order sufficient condition are assumed. Applying an Implicit Multifunction Theorem to the first order necessary conditions, we proof $O(h)$ convergence for a model problem in two space dimensions.

1 Introduction

We are interested in the numerical treatment of the following optimal control problem on a sufficiently smooth domain $\Omega \subset \mathbb{R}^n$, $n = 2, 3$

$$
\begin{aligned}
\min_{u\in L^2(\Omega),\, y\in C(\bar{\Omega})} \; & J(u, y) = \frac{1}{2}\|y - z\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} \\
& \text{subject to} \\
& y = S(u), \quad a \le u \le b, \quad y \ge 0\,,
\end{aligned}
\tag{1.1}
$$

with a desired state $z \in L^2(\Omega)$, a control $u \in L^2(\Omega)$, the state $y \in C(\bar{\Omega})$, a control-to-state operator $S \in C^2(L^2(\Omega), C(\bar{\Omega}))$, and the Tikhonov parameter $\alpha > 0$. We further assume $a, b \in L^\infty(\Omega)$, $a < b$ a.e., and by $U_{ad} = \{u \in L^2(\Omega) \mid a \le u \le b, \text{ a.e.}\}$ and $Y_{ad} = \{y \in C(\bar{\Omega}) \mid y \ge 0\}$ we denote the admissible sets for u and y. We further refer to $\mathcal{A}^a_u = \{x \in \Omega \mid u(x) = a\}$ as the active set of u with respect to a, and analogously introduce $\mathcal{A}^b_u$ and $\mathcal{A}_y$.

A lot of results are available for problem (1.1) in the situation of S being the solution operator of a linear or semilinear elliptic state equation. Second order sufficient conditions for the semilinear case were given in [CDLRT08]. The variational discretization considered in the present paper has first been proposed in [Hin05] for linear-quadratic control constrained problems. This approach has also been investigated in [DH07] including state and control constraints for the special case of a linear operator S. Error estimates for fully discretized

*Schwerpunkt Optimierung und Approximation, Universität Hamburg, Bundesstraße 55, 20146 Hamburg, Germany.

linear-quadratic problems can be found in [Mey08]. Both [DH07] and [Mey08] proved convergence of order $h^{1-\epsilon}$, $\epsilon > 0$ in two dimensions. Both results were improved in [HPUU09], where $O(h|\log h|)$ convergence was shown. Semilinear equations with finite dimensional control, and state constraints at a finite number of points were analyzed in [MTV08] using Robinson's Implicit Multifunction Theorem from [Rob80] in finite dimensions. The existence of a sequence of discrete solutions, converging towards a designated solution of (1.1), was discussed in a very general sense in [HM07], and in [ACT02] maximum norm estimates for the controls were derived in the absence of state constraints. We only mention that, to avoid dealing with the low regularity of the state constraint multipliers, regularizations of (1.1) have been proposed in [MRT06] and [CR08b] for linear operators S. Corresponding error estimates of order $h^{1-\epsilon}$, $\epsilon > 0$ were developed in [CR08a] in the situation of $\Omega \subset \mathbb{R}^2$.
The approach taken in the present paper, is to apply an Implicit Multifunction Theorem to the first order necessary conditions of (1.1). Given that a linearized Slater condition holds, these can be written in the form

$$(\mathbb{P}) \quad 0 \in \begin{pmatrix} \alpha u + S'^*(u)(y - z + K(y)) + N(u) \\ y - S(u) \\ C_{\geq 0}(\bar{\Omega}) - S(u) \end{pmatrix} \subset L^2(\Omega) \times C(\bar{\Omega}) \times C(\bar{\Omega})$$

with the normal cone

$$N(u) = \begin{cases} \left\{ v \in L^2(\Omega) \;\middle|\; \langle v, c - u \rangle_{L^2(\Omega)} \leq 0\,,\ \forall c \in U_{ad} \right\} & \text{if } u \in U_{ad} \\ \emptyset & \text{else} \end{cases}$$

and the cone

$$K(y) = \left\{ \mu \in C(\bar{\Omega})^* \;\middle|\; \langle \mu, \max(0, y) \rangle_{C(\bar{\Omega})^*, C(\bar{\Omega})} = 0 \wedge \forall c^+ \in Y_{ad} : \langle \mu, c^+ \rangle_{C(\bar{\Omega})^*, C(\bar{\Omega})} \leq 0 \right\}.$$

The set $C_{\geq 0}(\bar{\Omega})$ is the cone of pointwise nonnegative functions in $C(\bar{\Omega})$. By $C(\bar{\Omega})^*$ we denote the dual of $C(\bar{\Omega})$ and by $S'^*(u)$ the dual operator of the Fréchet derivative of S at u.
The idea is now to look at S as a parameter, and to investigate the dependence of solutions of $(\mathbb{P})$ on perturbations of that parameter. We consider a family $\{S_h\}$ of finite dimensional approximations to S. The set of indices h is an arbitrary but fixed, positive and strictly monotone sequence $\{h_n\}_{n\in\mathbb{N}}$ converging to zero, denoted $h \in \{h_n\}_{n\in\mathbb{N}}$. We further assume $S_{h_1} \neq S_{h_2}$ for $h_1 \neq h_2$. Using the convention $S_0 := S$, the set $\mathcal{P} = \{S_h\}_{h\geq 0}$ endowed with the metric $d_{\mathcal{P}}(S_{h_1}, S_{h_2}) = |h_1 - h_2|$ becomes a metric space, admitting exactly one convergent sequence. That space will be referred to as the parameter space.
The space $L^2(\Omega) \times C(\bar{\Omega})^2$ is not suitable for our approach, since the theory applied in Section 2 requires some regularity of the underlying spaces, namely the existence of a Fréchet smooth norm. We deal with this by formulating problem $(\mathbb{P})$ by means of a separable reflexive Banach space $W \subset C(\bar{\Omega})$, e.g. the Sobolev space $W^{1,p}(\Omega)$ with $n < p < \infty$. Since W is reflexive, it admits a Fréchet smooth equivalent norm, such that also the corresponding dual norm of W^* is Fréchet smooth (see for example [Die75] §9). For the rest of this paper, we consider W and W^* to be equipped with these smooth norms. The space W must be compatible with S and $\{S_h\}_{h>0}$ as

$$S_h : L^2(\Omega) \to W \subset C(\bar{\Omega}) \quad \forall h \geq 0\,. \tag{1.2}$$

We will also make use of the boundedness of the set of state-constraint multipliers in $C(\bar\Omega)^*$, so for technical reasons instead of $K(y)$ we consider

$$K_M(y) = \left\{ \mu \in K(y) \;\middle|\; \|\mu\|_{C(\bar\Omega)^*} \le M \right\}$$

for $M > 0$ sufficiently large.
The parameter dependent system then reads

$$(\mathbb{P}_h) \quad 0 \in F((u,y),S_h)) = \begin{pmatrix} \alpha u + S_h'^*(u)(y - z + K_M(y)) + N(u) \\ y - S_h(u) \\ W_{\ge 0} - S_h(u) \end{pmatrix} \subset L^2(\Omega) \times W^2 \,,$$

with a more regular state $y \in W$. The set $W_{\ge 0}$ is the cone of nonnegative functions in W.
Regarding problem $(\mathbb{P}_h)$, one observes, that the implementation proposed in [DH07] is applicable only if a and b are constant, or at least piecewise linear. Otherwise we have to discretize the bounds a, b first. This issue is addressed in Remark 3.15.
The objective of this paper is the application of an Implicit Multifunction Theorem, to obtain convergence of a sequence of solutions (u_h, y_h) of $(\mathbb{P}_h)$ towards each solution $(\bar u, \bar y)$ of $(\mathbb{P})$, that is sufficiently regular, i.e. that fulfills a second order sufficient condition and for that strict complementarity and the linearized Slater condition 1.1 hold. The order of convergence is determined by that of $S_h'(\bar u)$ in the operator norm, and the pointwise order of convergence of $S_h(\bar u)$, our main result being

$$\|u_h - \bar u\|_{L^2(\Omega)} + \|y_h - \bar y\|_W \le \frac{1}{\sigma} \left(\|(S_h'^*(\bar u) - S'^*(\bar u))(\bar y - z + \bar\mu)\|_{L^2(\Omega)} + 2\|S_h(\bar u) - S(\bar u)\|_W \right) .$$

For the example given in Section 4 the right hand side in the above estimate is $O(h)$. Note that, other than most authors, we do not assume uniform convergence of any order for S. With respect to the implementation observe that once the main result is stated the bound M becomes redundant and the same result holds for $M = \infty$.
Unfortunately, problem $(\mathbb{P})$ does not fulfill Robinson's condition for strong regularity, in fact the formulation given here does not even fit into Robinson's concept, so we cannot apply the results from [Rob80]. Note also, that uniqueness of the multipliers may not be given.
In Section 2, we therefore slightly generalize the Theorems 2.6 (Decrease Principle) and 3.1 (Implicit Multifunction Theorem) as well as Lemma 3.3 from [LZ99].
We then show in Section 3 that these results can be applied to $(\mathbb{P}_h)$. This approach is not aimed at showing uniqueness of a sequence of solutions u_h of $(\mathbb{P}_h)$ converging towards $\bar u$, but only at showing existence of such a sequence and some order of convergence. Under the given assumptions uniqueness of u_h can be recovered, this is however not carried out.
Finally, in Section 4 the abstract results are applied to an optimal control problem.
The following lemma concerns the relation between $(\mathbb{P})$ and $(\mathbb{P}_h)$ and the choice of M. The idea is to retain as much as possible of a given solution $(\bar u, \bar y)$ of $(\mathbb{P})$ when passing to $(\mathbb{P}_0)$. Under a linearized Slater assumption, we obtain boundedness of the multipliers $\mu \in K(\bar y)$ solving $(\mathbb{P})$, hence justifying the truncation of $K(y)$ into $K_M(y)$.

Assumption 1.1. There exists an admissible direction $d \in L^2(\Omega)$, such that $\bar u + d \in U_{ad}$ and

$$S(\bar u) + S'(\bar u)d \in \operatorname{int}(Y_{ad}) \,.$$

This assumption also ensures that $(\mathbb{P})$ holds at a given optimum $(\bar{u}, \bar{y})$ of (1.1).

Lemma 1.2. *The relation between $(\mathbb{P})$ and $(\mathbb{P}_h)$ is the following.*

1. *If $(\bar{u}, \bar{y})$ solves $(\mathbb{P})$ and also fulfills the Assumption 1.1, then the set of multipliers $\mu \in K(\bar{y})$ solving $(\mathbb{P})$ at $(\bar{u}, \bar{y})$ is bounded by some $B > 0$. Given the relation (1.2), $(\bar{u}, \bar{y})$ also solves $(\mathbb{P}_0)$ for $M = 3B$. Hence by the choice of M no relevant multipliers μ are lost when passing from $(\mathbb{P})$ to $(\mathbb{P}_h)$.*

2. *On the other hand every solution of $(\mathbb{P}_0)$ also is a solution to $(\mathbb{P})$.*

Proof. We only have to prove *1.* Given $\mu \in K(\bar{y})$ with $0 \in \alpha u + S_h'^*(\bar{u})(\bar{y} - z + \mu) + N(\bar{u})$, we have

$$\langle \alpha\bar{u} + S'^*(\bar{u})(\bar{y} - z + \mu), \tilde{u} - \bar{u}\rangle_{L^2(\Omega)} \geq 0 \quad \forall \tilde{u} \in U_{ad}$$

while the second line of $(\mathbb{P})$ says $\bar{y} = S(\bar{u})$. Now inserting $\tilde{u} = \bar{u} + d$ yields

$$-\langle \mu, S'(\bar{u})d\rangle_{C(\bar{\Omega})^*, C(\bar{\Omega})} \leq \langle \alpha\bar{u} + S'^*(\bar{u})(\bar{y} - z), d\rangle_{L^2(\Omega)} =: \tilde{M}\,.$$

Because $\operatorname{supp}(\mu) \subset \mathcal{A}_y$ and $\langle \mu, y^+\rangle \leq 0$ for all $y^+ \in Y_{ad}$ and because it follows from Assumption 1.1 that

$$S'(\bar{u})d \geq \delta > 0 \quad \text{on } \mathcal{A}_y$$

for some $\delta > 0$, we get

$$\delta|\langle \mu, 1\rangle_{C(\bar{\Omega})^*, C(\bar{\Omega})}| \leq \tilde{M}\,.$$

But on the other hand we have for any $\mu^- \in Y_{ad}^- = \left\{\mu \in C(\bar{\Omega})^* \mid \forall y \in Y_{ad} :\ \langle \mu, y\rangle \leq 0\right\}$

$$\|\mu^-\|_{C(\bar{\Omega})^*} = -\langle \mu^-, 1\rangle_{C(\bar{\Omega})^*, C(\bar{\Omega})}\,, \tag{1.3}$$

since if there were any $y \in C(\bar{\Omega})$ with $\|y\|_\infty \leq 1$ and $-\langle \mu^-, y\rangle_{C(\bar{\Omega})^*, C(\bar{\Omega})} > -\langle \mu^-, 1\rangle_{C(\bar{\Omega})^*, C(\bar{\Omega})}$ this would imply $\langle \mu^-, 1 - y\rangle_{C(\bar{\Omega})^*, C(\bar{\Omega})} > 0$, in contradiction to $\mu^- \in Y_{ad}^-$.
Thus finally one ends up with

$$\|\mu\|_{C(\bar{\Omega})^*} \leq \frac{\tilde{M}}{\delta} =: B\,.$$

∎

Remark 1.3. Equation (1.3) also implies, that in Y_{ad}^- weak* convergence entails convergence of the norms, which is why the sets $K_M(y)$ are weak* closed in $C(\bar{\Omega})^*$.

2 Implicit Multifunction Theorem

In this section we develop a slightly generalized Implicit Multifunction Theorem as in [LZ99]. The differentiability- and invertability-assumption of the classical Implicit Function Theorem is therein weakened to some condition on the subdifferential of a lower semicontinuous function, by making use of the following lemma.
Throughout this section we denote by $\hat{\partial} f(x)$ the Fréchet subdifferential of a lower semicontinuous function $f : X \to \mathbb{R}$ at $x \in X$, as defined and characterized in [LZ99] or more comprehensively in [Mor05]. Note, that if X allows for a Fréchet smooth norm, then there also exists a Fréchet smooth Lipschitz bump function on X.

Lemma 2.1 (Decrease Principle)**.** *Let X be a Banach space with a Fréchet smooth Lipschitz bump function, let $f : X \to \bar{\mathbb{R}}$ be a lower semicontinuous function bounded from below, and let $\bar{x} \in X$ as well as $r, \epsilon, \sigma > 0$. Suppose that for any $x \in B_r(\bar{x}) \cap \{x \in X \mid f(x) < f(\bar{x}) + \sigma r + \epsilon\}$, $\xi \in \hat{\partial} f(x)$ implies $\|\xi\|_{X^*} > \sigma > 0$. Then*

$$\inf_{x \in B_r(\bar{x})} f(x) \leq f(\bar{x}) - \sigma r \ .$$

Proof. Assume that for some $0 < \delta < \min(\sigma r, \epsilon/2)$

$$\inf_{x \in B_r(\bar{x})} f(x) > f(\bar{x}) - \sigma r + \delta \ . \tag{2.4}$$

Let $0 < \tau < r$, then we have

$$\lim_{\eta \to 0} \inf_{x \in B_\tau(\bar{x}) + B_\eta(0)} f(x) > f(\bar{x}) - \sigma r + \delta \ . \tag{2.5}$$

By the multidirectional mean-value inequality given in Theorem 2.5 from [LZ99] equation (2.5) implies the following. For every $\eta > 0$ we get $z \in B_\tau(\bar{x}) + B_\eta(0)$ and $z^* \in \hat{\partial} f(z)$, with

$$-\sigma r + \delta < \langle z^*, x - \bar{x} \rangle \quad \forall x \in B_\tau(\bar{x})$$

and

$$f(z) < f(\bar{x}) + \sigma r + \delta + \eta \, .$$

For a proof of the mean-value inequality see Theorem 2.6 in Chapter 3 of [CLSW98] . Choosing η sufficiently small now ensures $B_\tau(\bar{x}) + B_\eta(0) \subset B_r(\bar{x})$ and $f(z) < f(\bar{x}) + \sigma r + \epsilon$. Hence $\|z^*\| > \sigma$ and

$$\sigma r - \delta > \|z^*\|_{X^*} \tau > \sigma \tau \ .$$

Choosing τ sufficiently close to r yields a contradiction. The lemma follows from equation (2.4) hence being false for all sufficiently small $\delta > 0$. ∎

Lemma 2.1 is formulated as Theorem 2.6 in [LZ99] with the slightly stronger assumption, that $\xi \in \hat{\partial} f(x)$ implies $\|\xi\|_{X^*} > \sigma > 0$ for all $x \in B_r(\bar{x})$.
The next step is to generalize Theorem 3.1 from [LZ99]. This theorem deals with a lower semicontinuous function $f : X \times \mathcal{P} \to \bar{\mathbb{R}}$ on some smooth Banach space X and is concerned with solutions of

$$f(x, p) \leq 0 \, ,$$

depending on some parameter p out of a metric space P. For our purpose f will be the distance function $\mathrm{d}(0, F(x, p))$, measuring the distance between zero and the image $F(x, p)$ of a set valued mapping $F : X \times P \to 2^Y$, with another smooth enough Banach space Y. The distance is defined as usual

$$\forall y \in Y \, \forall \mathfrak{S} \subset Y : \ \mathrm{d}(y, \mathfrak{S}) = \inf_{s \in \mathfrak{S}} \|y - s\|_Y \ .$$

We further set $d(y, \emptyset) = \infty$ for all $y \in Y$, thus keeping $\mathrm{d}(y, F(x))$ well defined for all $x \in X$. The Theorem is formulated by means of the solution map $G : P \to X$

$$G(p) = \{x \in X \mid f(x, p) \leq 0\} \ .$$

The idea is to include the (very slight) generalization of the previous lemma by making use of the reduced assumptions on $\xi \in \hat{\partial} f(x)$. By $\hat{\partial}_x$ we denote the Fréchet subgradient with respect to the variable x.

Theorem 2.2 (Implicit Multifunction Theorem)**.** *Let X and Y be Banach spaces with Fréchet smooth Lipschitz bump functions, let (P, τ_P) be a topological space and let U be an open set in $X \times P$. Suppose that $f : U \to \bar{\mathbb{R}}$ satisfies*

1. *there exists $(\bar{x}, \bar{p}) \in U$ such that $f(\bar{x}, \bar{p}) \leq 0$;*
2. *$p \mapsto f(\bar{x}, p)$ is upper semicontinuous at $\bar{p}$;*
3. *for any p near $\bar{p}$, $x \mapsto f(x, p)$ is lower semicontinuous;*
4. *there exists $\epsilon > 0$ and $\sigma > 0$ such that, for any $(x, p) \in U$ with $0 < f(x, p) < \epsilon$, $\xi \in \hat{\partial}_x f(x, p)$ implies that $\|\xi\|_{X^*} > \sigma$.*

Then there exist open sets $W \subset X$ and $V \subset P$ containing $\bar{x}$ and $\bar{p}$ respectively, such that

1. *for any $p \in V$, $W \cap G(p) \neq \emptyset$;*
2. *for any $p \in V$ and $x \in W$,*
$$d(x, G(p)) \leq \frac{f_+(x, p)}{\sigma} ,$$
where $f_+(x, p) = \max(0, f(x, p))$.

Proof. The Proof is exactly the same as in [LZ99], but one has to choose r' sufficiently small to ensure $r'\sigma < \epsilon$. ∎

The fourth condition in Theorem 2.2 concerning $\hat{\partial}_x f$ is given a more easily manageable shape in the next lemma, whose proof is exactly the same as the one for Lemma 3.3 in [LZ99]. Before formulating its assertion, we need to clarify our notation.

Definition 2.3 (Projection)**.** For all $x \in X$, $y \in Y$ we define

$$\mathrm{pr}(y, F(x)) = \{\tilde{y} \in F(x) \mid \mathrm{d}(y, \tilde{y}) = \mathrm{d}(y, F(x))\} .$$

Definition 2.4 (Fréchet Normals)**.** Let X be an arbitrary Banach space and $\mathfrak{S} \subset X$. The Fréchet normal cone to $\mathfrak{S}$ at $\bar{x} \in \mathfrak{S}$ is defined as

$$\hat{N}(\bar{x}, \mathfrak{S}) = \left\{ x^* \in X^* \;\middle|\; \limsup_{x \in \mathfrak{S}\, x \to \bar{x}} \frac{\langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|_X} \leq 0 \right\} .$$

Definition 2.5 (Coderivative)**.** Let X and Y be Banach spaces and let $F : X \to 2^Y$ be a multifunction with closed graph and $y \in F(x)$. Then the Fréchet coderivative at (x, y) is defined as
$$\hat{D}^* F(x, y) y^* = \left\{ x^* \in X^* \;\middle|\; (x^*, -y^*) \in \hat{N}((x, y), \operatorname{graph} F) \right\}$$
where $\hat{N}((x, y), \operatorname{graph} F)$ denotes the Fréchet normal cone of the set $\operatorname{graph} F$ at the point (x, y). If F depends on some parameter p, we refer to the coderivative with respect to x as $D^* F(x, y; p)$.

Lemma 2.6. *Let X be a Banach space with Fréchet smooth Lipschitz bump functions, let Y be a Banach space with a Fréchet smooth norm, let $U \subset X$ be an open set and let $F : U \to 2^Y$ be a multifunction with closed graph, such that for any $x \in U$ there either holds $pr(0, F(x)) \neq \emptyset$ or $F(x) = \emptyset$. Denote by $\|\cdot\|'_Y$ the Fréchet derivative of the norm in Y, and let $f(x) = d(0, F(x))$ be lower semicontinuous on U. Suppose*

$$\text{for } x \in U \text{ with } 0 \notin F(x) \neq \emptyset \text{ we can choose } y \in pr(0, F(x)) \text{ such that}$$
$$\sigma \leq \inf\left\{ \|x^*\|_{X^*} \;\middle|\; x^* \in D^*F(x; y)(y^*), y^* = \|y\|'_Y \right\}.$$

Then $\xi \in \hat{\partial} f(x)$ implies that $\|\xi\|_{X^*} > \sigma$. Further the value of σ does not depend on the choice of $y \in pr(0, F(x))$.

Proof. If $F(x) = \emptyset$, then $\hat{\partial} f(x) = \emptyset$. If $F(x) \neq \emptyset$, let $\xi \in \hat{\partial} f(x)$ where $f(x) > 0$. By the definition of the subdifferential there exists a Fréchet smooth function g such that $g'(x) = \xi$ and $f - g$ attains a local minimum at x. Let $y \in \operatorname{pr}(0, F(x))$. Then $\|y\|_Y = f(x)$ and we have for x' sufficiently close to x

$$\begin{aligned} f(x) - g(x) &= \|y\|_Y - g(x) = \|y\|_Y + \boldsymbol{\delta}_{\operatorname{Graph} F}(x, y) - g(x) \\ &\leq f(x') - g(x') \leq \|y'\|_Y + \boldsymbol{\delta}_{\operatorname{Graph} F}(x', y') - g(x') \quad \forall y' \in Y, \end{aligned}$$

where $\boldsymbol{\delta}_{\operatorname{Graph} F}(x', y')$ denotes the indicator function of the set $\operatorname{Graph} F$ (i.e. $\boldsymbol{\delta}_{\operatorname{Graph} F}(x', y') = 0$ for $(x', y') \in \operatorname{Graph} F$ and ∞ otherwise). Hence the function

$$(x', y') \mapsto \|y'\|_Y + \boldsymbol{\delta}_{\operatorname{Graph} F}(x', y') - g(x')$$

attains a local minimum at (x, y). Note that $\|y\|_Y > 0$ and therefore $(x', y') \mapsto g(x') - \|y'\|_Y$ is differentiable at (x, y). Thus $(g'(x), -\|y\|'_Y) \in \hat{\partial}\boldsymbol{\delta}_{\operatorname{Graph} F}(x, y)$, and because $\hat{\partial}\boldsymbol{\delta}_{\operatorname{Graph} F}(x, y)$ is contained in the Fréchet normal cone it follows

$$\xi = g'(x) \in \hat{D}^*F(x, y)(\|y\|'_Y),$$

and finally $\|\xi\|_{X^*} > \sigma$. ∎

3 Application to the Optimal Control Problem

Now with respect to our original problem $(\mathbb{P}_h)$ we consider the spaces

$$X = L^2(\Omega) \times W \text{ and } Y = L^2(\Omega) \times W \times W$$

endowed with the Fréchet smooth norms

$$\|(u, y)\|_X = \sqrt{\|u\|^2_{L^2(\Omega)} + \|y\|^2_W} \text{ and } \|(u, y, v)\|_Y = \sqrt{\|u\|^2_{L^2(\Omega)} + \|y\|^2_W + \|v\|^2_W}.$$

The purpose of this section is to verify that Theorem 2.2 can be applied to $(\mathbb{P}_h)$ under reasonable assumptions on the family $\{S_h\}$ and the multipliers $\bar{\mu}$, $\bar{\lambda}$ that solve $(\mathbb{P})$ for some fixed solution $(\bar{u}, \bar{y})$.
We make the following assumptions concerning the convergence and stability of S_h.

Assumption 3.1. $S_h(u) \xrightarrow{h \to 0} S(u)$ in W for any fixed $u \in U_{ad}$.

Assumption 3.2. For any $h \geq 0$ there holds $S_h \in \mathcal{C}^2(L^2(\Omega), C(\bar{\Omega}))$ and S_h is differentiable as an operator from $L^2(\Omega)$ into W.

Assumption 3.3. $S_h'(u_h) \overset{h\to 0}{\longrightarrow} S'(u)$ in $\mathcal{L}(L^2(\Omega), W)$, for all sequences $u_h \overset{L^2(\Omega)}{\longrightarrow} u$ bounded in $L^\infty(\Omega)$.

We further simplify the notation by the following

Definition 3.4. For notational convenience, we introduce the function

$$\mathcal{F} : Q \subset L^2(\Omega) \times W \times C(\bar{\Omega})^* \times L^2(\Omega) \times W \times \mathcal{P} \longrightarrow L^2(\Omega) \times W \times W\,,$$

that indexes points in the image of $F((u, y), S_h)$ by

$$\mathcal{F}(u, y, \mu, \lambda, \nu; S_h) = \begin{pmatrix} \alpha u + S_h'^*(u)(y - z + \mu) + \lambda \\ y - S_h(u) \\ \nu - S_h(u) \end{pmatrix},$$

the domain of $\mathcal{F}$ is $Q = \{(u, y, \mu, \lambda, \nu) \mid \mu \in K_M(y),\ \lambda \in N(u),\ \nu \in W_{\geq 0}\} \times \mathcal{P}$.
We further denote by $DF^*(u, y, \mu, \lambda, \nu; S_h)$ the Fréchet coderivative with respect to (u, y)

$$DF^*((u, y), \mathcal{F}(u, y, \mu, \lambda, \nu; S_h); S_h)$$

at the point $\mathcal{F}(u, y, \mu, \lambda, \nu; S_h)$.

To apply Lemma 2.6 to $f(x, p) = f((u, y), S_h) = d(0, F((u, y), S_h))$, we have to prove the non-emptiness of $\mathrm{pr}(0, F((u, y), S_h))$ for non-empty F, and the lower semicontinuity of f with respect to x. Also, to make use of the lemma, one has to characterize the coderivative of F. Finally, the semicontinuity assumptions of Theorem 2.2 need to be verified.

Lemma 3.5. *Provided Assumptions 3.1 - 3.3 hold, the set-valued function F from $(\mathbb{P}_h)$ has the following properties.*

1. *$pr(0, F((u, y), S_h)) \neq \emptyset \vee F((u, y), S_h)) = \emptyset$.*
2. *$d(0, F(\,\cdot\,, S_h))$ is lower semicontinuous for any fixed S_h, $h \geq 0$.*
3. *$d(0, F((u, y),\,\cdot\,)$ is upper semicontinuous at S, for any given $u \in L^2(\Omega)$, $y \in W$.*
4. *The graph of $F(\,\cdot\,, S_h)$ is closed.*
5. *For admissible $u \in U_{ad}$ and $\|\mu\|_{C(\bar{\Omega})^*} \leq M/2$ the Fréchet coderivative*

$$DF^*(u, y, \mu, \lambda, \nu; S_h) : L^2(\Omega) \times W^* \times W^* \to L^2(\Omega) \times W^*\,,\ (\eta_1, \eta_2, \eta_3) \mapsto (u^*, y^*)$$

either has the shape

$$\begin{aligned} u^* - (\alpha Id + S_h''^*(u)(y - z + \mu))\eta_1 + S_h'^*(u)(\eta_2 + \eta_3) \quad &\in \quad N(u)\,, && (3.6)\\ y^* = S_h'(u)\eta_1 + \eta_2 \quad &\in \quad W^*\,, && (3.7) \end{aligned}$$

or is empty valued for (η_1, η_2, η_3)*. In particular it is empty valued for all but*

$$\begin{aligned}
-\eta_1 \in \quad & \left\{v \in L^2(\Omega) \;\middle|\; v \le 0 \text{ on } \mathcal{A}_u^b \wedge v \ge 0 \text{ on } \mathcal{A}_u^a\right\} \\
& \cap \left\{v \in L^2(\Omega) \mid v(x) = 0 \text{ if } \lambda(x) \neq 0\right\} \\
& \cap \left\{v \in L^2(\Omega) \;\middle|\; \langle \mu, S_h'(u)v\rangle_{C(\bar\Omega)^*, C(\bar\Omega)} = 0 \wedge S_h'(u)v \ge 0 \text{ on } \mathcal{A}_y\right\} \\
& =: C(u, y, \mu, \lambda, S_h)
\end{aligned} \tag{3.8}$$

as well as

$$\eta_3 \in \left\{w^* \in W^* \;\middle|\; \langle w^*, \nu\rangle_{W^*,W} = 0 \wedge \forall w^+ \in W_{\ge 0} : \langle w^*, w^+\rangle_{W^*,W} \ge 0\right\} \tag{3.9}$$

where $C(u, y, \mu, \lambda, S_h)$ *can be seen as a relaxation of the cone of critical directions* $C_{\bar u}$ *in [CDLRT08].*

Because we assumed $b > a$ a.e., we can also write for $u \in U_{ad}$

$$N(u) = \left\{v \in L^2(\Omega) \;\middle|\; v \ge 0 \text{ on } \mathcal{A}_u^b,\ v \le 0 \text{ on } \mathcal{A}_u^a,\ v = 0 \text{ otherwise}\right\}. \tag{3.10}$$

Proof. 1. Let $F((u, y), S_h) \neq \emptyset$. Then there exists a minimizing sequence

$$\mathbf{y}_k = \begin{pmatrix} \alpha u + S_h'^*(u)(y - z + \mu_k) + \lambda_k \\ y - S_h(u) \\ \nu_k - S_h(u) \end{pmatrix} \in F((u, y), S_h),$$

such that $\lim_{k\to\infty} d(\mathbf{y}_k, 0) = \inf_{\mathbf{y}\in F((u,y),S_h)} d(\mathbf{y}, 0)$. Now the sequences μ_k, ν_k and with μ_k also λ_k are bounded in their respective norms, and since bounded sets in $C(\bar\Omega)^*$ as well as in $L^2(\Omega)$ are relatively weakly* sequentially compact, we can extract a subsequence $(\mu_j, \lambda_j, \nu_j)$, converging weakly* towards some $(\tilde\mu, \tilde\lambda, \tilde\nu)$. Because $L^2(\Omega)$ is reflexive, weak and weak* convergence coincide. All three limits lie inside $F((u, y), S_h)$, because $K_M(y)$ is weak* closed (see Remark 1.3) and $N(u)$ and $W_{\ge 0}$ are closed and convex and hence weakly closed. The weak lower semicontinuity of the norms yields $d(0, \mathcal{F}(u, y, \tilde\mu, \tilde\lambda, \tilde\nu; S_h)) = \inf_{\mathbf{y}\in F((u,y),S_h)} d(0, \mathbf{y})$.

2. Suppose there exists a sequence $(u_k, y_k) \overset{k\to\infty}{\longrightarrow} (u, y)$ such that

$$\lim_{k\to\infty} d(F((u_k, y_k), S_h), 0) < d(F((u, y), S_h), 0), \tag{3.11}$$

in particular $F((u_k, y_k), S_h) \neq \emptyset$. Any sequence

$$\begin{pmatrix} \alpha u_k + S_h'^*(u_k)(y_k - z + \mu_k) + \lambda_k \\ y_k - S_h(u_k) \\ \nu_k - S_h(u_k) \end{pmatrix} \in \mathrm{pr}(0, F((u_k, y_k), S_h))$$

is bounded and hence $(\mu_k, \lambda_k, \nu_k)$ admits a subsequence with indices $\tilde k$ converging weakly towards $(\tilde\mu, \tilde\lambda, \tilde\nu)$. Because of the strong convergence of u_k, we finally have

$$\langle \tilde\lambda, c - u\rangle_{L^2(\Omega)} = \lim_{\tilde k\to\infty} \langle \lambda_{\tilde k}, c - u_{\tilde k}\rangle_{L^2(\Omega)} \le 0$$

for all $c \in U_{ad}$, that is $\tilde{\lambda} \in N(u)$. Because of the strong convergence of $y_{\tilde{k}}$ we have also

$$\langle \tilde{\mu}, \max(0, y) \rangle_{C(\bar{\Omega})^*, C(\bar{\Omega})} = \lim_{\tilde{k} \to \infty} \langle \mu_{\tilde{k}}, \max(0, y_{\tilde{k}}) \rangle = 0$$

and taking into account that $\langle \tilde{\mu}, c^+ \rangle_{C(\bar{\Omega})^*, C(\bar{\Omega})} \leq 0\,,\ \forall c^+ \in C_{\geq 0}(\bar{\Omega})$ we get $\tilde{\mu} \in K_M(y)$. Using Assumption 3.2 and hence $S_h'^*(u_{\tilde{k}})\mu_{\tilde{k}} \to S_h'^*(u)\tilde{\mu}$, the weak lower semicontinuity of the norm now yields

$$\liminf_{\tilde{k} \to \infty} d(F((u_{\tilde{k}}, y_{\tilde{k}}), S_h), 0) \geq d(F((u, y), S_h), 0)$$

in contradiction to (3.11)

3. Because of *1.* , there exist admissible multipliers μ, λ and ν such that

$$\begin{pmatrix} \alpha u + S'^*(u)(y - z + \mu) + \lambda \\ y - S(u) \\ \nu - S(u) \end{pmatrix} \in \operatorname{pr}(0, F((u, y), S))\,.$$

By fixing μ, λ and ν we get from Assumptions 3.1 and 3.3

$$\lim_{h \to 0} d(0, F((u, y), S_h)) \leq \mathrm{d}(0, F((u, y), S))\,.$$

4. Consider a sequence

$$\begin{pmatrix} u_k \\ y_k \\ \alpha u_k + S_h'^*(u_k)(y_k - z + \mu_k) + \lambda_k \\ y_k - S_h(u_k) \\ \nu_k - S_h(u_k) \end{pmatrix} \in \operatorname{graph}(F(\,\cdot\,, S_h))\,,$$

converging towards some (u, y, w_1, w_2, w_3). Due to Assumption 3.2, we have $w_2 = y - S_h(u)$ and $w_3 = \nu - S_h(u)$ for some $\nu \in W_{\geq 0}$. Using a weak* converging subsequence of μ_k, a consideration similar to *1.* shows that indeed

$$w_1 = \alpha u + S_h'^*(u)(y - z + \mu) + \lambda$$

for some $\mu \in K_M(y)$, $\lambda \in N(u)$.

5. Hence the Fréchet coderivative of F is well defined as in Definition 2.5.

To characterize the Fréchet normal cone (see Definition 2.4) at some point

$$\begin{pmatrix} \alpha u + S_h'^*(u)(y - z + \mu) + \lambda \\ y - S_h(u) \\ \nu - S_h(u) \end{pmatrix} \tag{3.12}$$

one can derive necessary conditions for $(u^*, y^*, -\eta_1, -\eta_2, -\eta_3)$ to belong to the Fréchet normal cone of $\operatorname{graph}(F(\,\cdot\,, S_h))$ at the point given by (3.12). By considering sequences inside the graph, that vary only in λ, μ or ν, respectively, one observes

$$\begin{aligned} \langle -\eta_1, \tilde{\lambda} - \lambda \rangle &\leq 0 \quad \forall \tilde{\lambda} \in N(u)\,, \\ \langle -S'(u)\eta_1, \tilde{\mu} - \mu \rangle &\leq 0 \quad \forall \tilde{\mu} \in K_M(y)\,, \\ \langle -\eta_3, \tilde{\nu} - \nu \rangle &\leq 0 \quad \forall \tilde{\nu} \in W_{\geq 0}\,. \end{aligned} \tag{3.13}$$

Since 0 as well as 2μ lie in $K_M(y)$, it follows that $\langle S'(u)\eta_1, \mu \rangle_{C(\bar{\Omega}),C(\bar{\Omega})^*} = 0$ and by the same reasoning for λ we get $\langle \eta_1, \lambda \rangle_{L^2(\Omega)} = 0$ and $\langle \eta_3, \nu \rangle_{W^*,W} = 0$. Now, by considering sequences in graph(F) that vary only in u or only in y converging towards

$$(\alpha u + S_h'^*(u)(y - z), y - S_h(u), -S_h(u)) \in F((u, y), S_h)\,,$$

and using the differentiability Assumptions 3.2, one gets

$$\begin{aligned} \langle -\eta_1, \alpha\, \mathrm{du} + S_h''^*(u)(y - z + \mu)\mathrm{du} \rangle + \langle -\eta_2 - \eta_3, -S_h'(u)\, \mathrm{du} \rangle + \langle u^*\,, \mathrm{du} \rangle &\leq 0 \quad \forall \mathrm{du} \in U_{ad} - u\,, \\ \langle -\eta_1, S_h'^*(u)\, \mathrm{dy} \rangle + \langle -\eta_2, \mathrm{dy} \rangle + \langle y^*, \mathrm{dy} \rangle &\leq 0 \quad \forall \mathrm{dy} \in W\,, \end{aligned}$$

yielding (3.6)-(3.7). Note, that as in Lemma 3.17 the operator $S_h''^*(u)(y - z + \mu)$ is selfadjoint due to Assumption 3.2. The relation (3.8) follows from (3.13), considered that $K_{M/2}(y) \subset K_M(y) - \mu$ because of $\|\mu\| \leq M/2$. Hence $\langle S_h'(u)\eta_1, \tilde{\mu} \rangle \geq 0$ for all $\tilde{\mu} \in K_{\frac{M}{2}}(y)$ implies $S_h'(u)\eta_1 \leq 0$ on $\mathcal{A}_y$. The necessity of (3.9) also follows from (3.13).

∎

The next lemma now makes sure, that the prerequisites for Lemma 2.6 hold, and thus also the prerequisite on $\hat{\partial} f$ in Theorem 2.2, provided that the following conditions apply to a given solution $(\bar{u}, \bar{y})$ of $(\mathbb{P})$.

Definition 3.6. For some given $(u, y) \in L^2(\Omega) \times W$, by $\mathcal{K}(u, y)$ we denote the set of multipliers $\mu \in K(y)$, that solve $(\mathbb{P})$. If $\mathcal{K}(u, y) \neq \emptyset$, then the multiplier λ_μ solving

$$\mathcal{F}(u, y, \mu, \lambda_\mu, \nu; S) = 0$$

is uniquely determined by $\mu \in \mathcal{K}(u, y)$ as in (3.16). We write $C(u, y, \mu, S) := C(u, y, \mu, \lambda_\mu, S)$.

We will make use of a second order sufficient condition.

Assumption 3.7. $\mathcal{K}(\bar{u}, \bar{y}) \neq \emptyset$ and for all $\mu \in \mathcal{K}(\bar{u}, \bar{y})$, $\lambda \in C(\bar{u}, \bar{y}, \mu, S) \setminus \{0\}$

$$\lambda(\alpha \mathrm{Id} + S''^*(\bar{u})(\bar{y} - z + \mu) + S'^*(\bar{u})S'(\bar{u}))\lambda > 0\,.$$

It was shown in [CDLRT08] that 3.7 is indeed sufficient for strict local optimality, for a class of semilinear problems including our example from Section 4. Further two strict complementarity conditions must be fulfilled, to discern active and inactive sets.

Assumption 3.8. For all $\mu \in \mathcal{K}(\bar{u}, \bar{y})$, the set

$$\left\{ x \in \Omega \;\middle|\; -\frac{1}{\alpha} S'^*(\bar{u})(\bar{y} - z + \mu)[x] \in \{a(x), b(x)\} \right\}$$

has Lebesgue measure zero.

Assumption 3.9. For all $\mu \in \mathcal{K}(\bar{u}, \bar{y})$ there holds

$$\mathrm{supp}(\mu) = \mathcal{A}_{\bar{y}} \,.$$

Of course some assumptions concerning the convergence of the second derivative are necessary.

Assumption 3.10. $S_h''(u_h) \stackrel{h\to 0}{\longrightarrow} S''(u)$ in $\mathcal{L}(L^2(\Omega)\times L^2(\Omega), C(\bar{\Omega}))$, for all sequences $u_h \stackrel{L^2(\Omega)}{\longrightarrow} u$ bounded in $L^\infty(\Omega)$.

The following assumption guarantees for the compactness of $S_h''^*(u)(y - z + \mu) : L^2(\Omega) \to L^2(\Omega)$ for $h > 0$.

Assumption 3.11. For $h > 0$ the operator $S_h''(u) : L^2(\Omega)^2 \to C(\bar{\Omega})$ is in fact the concatenation of some continuous linear operator $\pi_h : L^2(\Omega) \to V_h$ into some finite dimensional subspace of $L^2(\Omega)$ and some bilinear continuous operator $T_h : V_h^2 \to C(\bar{\Omega})$.

Lemma 3.12. *Let $(\bar{u}, \bar{y})$ solve $(\mathbb{P})$. Suppose that in addition to the prerequisites of Lemma 3.5 the Assumptions 3.7- 3.11 as well as the linearized Slater condition 1.1 hold at $(\bar{u}, \bar{y})$. Then there exists $\sigma > 0$ and $\epsilon > 0$ and an open set $U_x \times U_p \subset X \times \mathcal{P}$ containing $((\bar{u}, \bar{y}), S_0)$ such that for all $((u, y), S_h) \in U_x \times U_p$ with $0 < f((u, y), S_h) < \epsilon$ the following holds. Let $\mathcal{F}(u, y, \mu, \lambda, \nu; S_h) \in pr(0, F((u, y), S_h))$ and $\eta = \|\mathcal{F}(u, y, \mu, \lambda, \nu; S_h)\|_Y' \in L^2(\Omega) \times W^* \times W^*$ as in Lemma 2.6, then we have*

$$\|\hat{D}^* F(u, y, \mu, \lambda, \nu; S_h)\eta\|_{X^*} \geq \sigma \,.$$

Proof. Suppose the Lemma does not hold. Then there exists a sequence $(u_k, y_k, S_{h_k}) \to (\bar{u}, \bar{y}, S_0)$ with

$$f((u_k, y_k), S_{h_k}) \to 0 \tag{3.14}$$

and corresponding $\mu_k \in K_M(y_k)$, $\lambda_k \in N(u_k)$ and $\nu_k \in W_{\geq 0}$, such that

$$\mathcal{F}(u_k, y_k, \mu_k, \lambda_k, \nu_k; S_{h_k}) \in \mathrm{pr}(0, F((u_k, y_k), S_{h_k}))$$

and $\eta_k = \|\mathcal{F}(u_k, y_k, \mu_k, \lambda_k, \nu_k; S_{h_k})\|_Y'$, such that there exist

$$(u_k^*, y_k^*) \in \hat{D}^* F(u_k, y_k, \mu_k, \lambda_k, \nu_k; S_{h_k})\eta_k \text{ with } \lim_{k\to\infty} \|(u_k^*, y_k^*)\|_{X^*} = 0 \,. \tag{3.15}$$

We will show, that (3.15) contradicts $\|\eta_k\|_{Y^*} = 1$.
Since μ_k is bounded, a subsequence again denoted μ_k converges weak* against some $\bar{\mu} \in K_M(\bar{y})$ as in the proof of Lemma 3.5. Because of (3.14), the compactness of $S'^*(\bar{u})$ (compare Lemma 3.17) and Assumption 3.3

$$\alpha u_k + S_{h_k}'^*(u_k)(y_k - z + \mu_k) \to \alpha\bar{u} + S'^*(\bar{u})(\bar{y} - z + \bar{\mu})$$

strongly as $k \to \infty$ and hence we have

$$\lambda_k \stackrel{k\to\infty}{\longrightarrow} -\left[\bar{u} + S'^*(\bar{u})(\bar{y} - z + \bar{\mu})\right]_{\mathcal{A}_{\bar{u}}} = \bar{\lambda}_{\bar{\mu}} \,,$$

where by $[\,\cdot\,]_{\mathcal{A}_{\bar{u}}}$ we mean the operator that just cuts off any function to zero outside of $\mathcal{A}_{\bar{u}}$. Note that $\bar{\lambda}_{\bar{\mu}}$ is uniquely determined by $\bar{u}, \bar{y}$ and $\bar{\mu}$ via the $L^2(\Omega)$ minimization problem

$$\bar{\lambda}_{\bar{\mu}} = \arg\min_{\lambda\in N(\bar{u})} \left\|\alpha\bar{u} + S'^*(\bar{u})(\bar{y} - z + \bar{\mu}) + \lambda\right\|_{L^2(\Omega)}^2 \,, \tag{3.16}$$

namely

$$\bar\lambda_{\bar\mu} = \begin{cases} -\min\left(\alpha\bar u + S'^*(\bar u)(\bar y - z + \bar\mu), 0\right) & \text{on } \mathcal{A}^b_{\bar u} \\ -\max\left(\alpha\bar u + S'^*(\bar u)(\bar y - z + \bar\mu), 0\right) & \text{on } \mathcal{A}^a_{\bar u} \end{cases} .$$

The same holds true for λ_k with respect to u_k, y_k, μ_k, S_{h_k}. Because of Assumption 3.8 the multiplier $\bar\lambda_{\bar\mu}$ is a.e. non-zero on $\mathcal{A}_{\bar u}$.
Finally, due to (3.14) $\|\nu_k - \bar\nu\|_W \stackrel{k\to\infty}{\longrightarrow} 0$, setting $\bar\nu = S(\bar u)$, and we get

$$\mathcal{F}(\bar u, \bar y, \bar\mu, \bar\lambda_{\bar\mu}, \bar\nu; S_0) = 0 .$$

Now, let $\eta_k = (\eta_1^k, \eta_2^k, \eta_3^k)$. From the definition of η_k there follows $\|\eta_k\|_{Y^*} = 1$. Because of that, there exists a subsequence again denoted by η_k converging weakly in Y^* towards some $\bar\eta = (\bar\eta_1, \bar\eta_2, \bar\eta_3) \in Y^*$. This implies weak convergence of η_1^k towards $\bar\eta_1$ in $L^2(\Omega)$ and $\eta_2^k \rightharpoonup \bar\eta_2$ and $\eta_3^k \rightharpoonup \bar\eta_3$ in W^*. We show, that indeed $-\bar\eta_1$ is an admissible direction lying in $C(\bar u, \bar y, \bar\mu, S)$.
First, since $\lambda_k \to \bar\lambda_{\bar\mu}$, a subsequence $\lambda_{\tilde k}$ converges a.e. pointwise, and from (3.8) we know, that since $\bar\lambda_{\bar\mu}$ is non-zero on $\mathcal{A}_{\bar u}$, the multiplier $\eta_1^{\tilde k}$ tends towards zero pointwise on $\mathcal{A}_{\bar u}$. Its weak limit $\bar\eta_1$ thus equals zero a.e. on $\mathcal{A}_{\bar u}$ (compare Thm. 5.9, Ch. VI in [Els96]). The corresponding subsequence of η_k is again denoted by η_k.
Secondly, the convergence of $\langle \mu_k, S'_{h_k}(u_k)\eta_1^k\rangle$ towards $\langle \bar\mu, S'_0(\bar u)\bar\eta_1\rangle$ follows from Assumption 3.3 and the fact, that $S'_0(\bar u)$ is compact.
Thirdly, we have $\mu_k \stackrel{*}{\rightharpoonup} \bar\mu$; and for any open set $\mathcal{O} \subset \Omega$ with $\mathcal{O} \cap \operatorname{supp}(\bar\mu) \neq \emptyset$ there exists some $c_{\mathcal{O}} \in Y_{ad}$ with $\operatorname{supp}(c_{\mathcal{O}}) \cap \operatorname{supp}(\bar\mu)$ containing an open set, such that $\langle \bar\mu, c_{\mathcal{O}}\rangle = k_{\mathcal{O}} > 0$. Hence $\langle \mu_k, c_{\mathcal{O}}\rangle \to k_{\mathcal{O}}$ and $\mathcal{O} \cap \operatorname{supp}(\mu_k) \neq \emptyset$ for k sufficiently large. Also, as stated in Remark 1.3, $\|\mu_k\| \to \|\bar\mu\|$ and by the choice of M in Lemma 1.2 we can choose k large enough to ensure $\|\mu_k\| \le M/2$, and by Lemma 3.5 $S'_{h_k}(u_k)\eta_1^k \ge 0$ on $\mathcal{A}_{y_k}$.
Now for any $x \in \operatorname{supp}(\bar\mu)$, by considering a family of open sets $B_{\mathbb{R}^d}(x, 1/n)$, $n \in \mathbb{N}$ we obtain a subsequence y_i and some sequence $x_i \in \operatorname{supp}(\mu_i) \subset \mathcal{A}_{y_i}$ with $x_i \to x$, contradicting $(S'(\bar u)\bar\eta_1)(x) < 0$. By Assumption 3.9 now follows $S'(\bar u)\bar\eta_1 \ge 0$ on $\mathcal{A}_{\bar y}$.
Now that we have $-\bar\eta_1 \in C(\bar u, \bar y, \bar\mu, S)$ and $\|\mu_k\| \le M/2$ for large k, we apply the multiplier from (3.6) to $-\eta_1^k$, to obtain

$$\langle u_k^*, \eta_1^k\rangle_{L^2(\Omega)} \ge \eta_1^k(\alpha \mathrm{Id} + S''^*_{h_k}(u_k)(y_k - z + \mu_k))\eta_1^k - \eta_1^k S'^*_{h_k}(u_k)(\eta_2^k + \eta_3^k) ,$$

where we made use of (3.10). On the other hand one has from (3.7)

$$\langle y_k^*, S'_{h_k}(u_k)\eta_1^k\rangle_{W^*,W} = \eta_1^k S'^*_{h_k}(u_k) S'_{h_k}(u_k)\eta_1^k + \langle S'^*_{h_k}(u_k)\eta_2^k, \eta_1^k\rangle_{L^2(\Omega)} .$$

Combining both gives

$$\begin{aligned} \langle u_k^*, \eta_1^k\rangle_{L^2(\Omega)} + \langle y_k^*, S'_h(u_k)\eta_1^k\rangle_{W^*,W} \ge & \eta_1^k\Big(\alpha \mathrm{Id} + S''^*_h(u_k)(y_k - z + \mu_k) \\ & + S'^*_h(u_k)S'_h(u_k)\Big)\eta_1^k - \eta_1^k S'^*_h(u_k)\eta_3^k . \end{aligned} \tag{3.17}$$

We now show, that in fact $\|\eta_k\|_Y \to 0$, and start by showing $\|\eta_1^k\|_{L^2(\Omega)} \to 0$.
Assume that there exists a subsequence $\zeta_k = (\zeta_1^k, \zeta_2^k, \zeta_3^k)$ of η_k, such that $\lim_{k\to\infty} \|\zeta_1^k\|_{L^2(\Omega)} = \gamma > 0$. Because of (3.9) we know, that the support of the weak limit $\bar\zeta_3$ of ζ_3^k lies in $\mathcal{A}_{\bar y}$. On the other hand, $-S'_{h_k}(u_k)\zeta_1^k$ converges strongly in W towards $-S'^*_0(\bar u)\bar\zeta_1$. In fact $S'^*_0(\bar u)\bar\zeta_1 \ge 0$ on $\mathcal{A}_{\bar y}$ as was shown in the first part of this proof. Hence $\lim_{k\to\infty} -\langle \zeta_1^k, S'^*_{h_k}(u_k)\zeta_3^k\rangle \ge 0$.

Inserting this into (3.17) using $u_k, y_k \to 0$ yields

$$0 \geq \lim_{k\to\infty} \zeta_1^k(\alpha\mathrm{Id} + S''^*_{h_k}(u_k)(y_k - z + \mu_k) + S'^*_{h_k}(u_k)S'_{h_k}(u_k))\zeta_1^k\,,$$

and together with the convergence properties from Assumptions 3.3 and 3.10 we have

$$0 \geq \lim_{k\to\infty} \zeta_1^k(\alpha\mathrm{Id} + S''^*(\bar u)(\bar y - z + \bar\mu) + S'^*(\bar u)S'(\bar u))\zeta_1^k\,,$$

and using the compactness assured by Lemma 3.17, the second order sufficient condition 3.7 finally gives

$$\begin{aligned} 0 &\geq \alpha\gamma^2 + \bar\zeta_1(S''^*(\bar u)(\bar y - z + \bar\mu) + S'^*(\bar u)S'(\bar u))\bar\zeta_1 \\ &\geq \bar\zeta_1(\alpha\mathrm{Id} + S''^*(\bar u)(\bar y - z + \bar\mu) + S'^*(\bar u)S'(\bar u))\bar\zeta_1 > 0\,, \end{aligned}$$

for $-\bar\zeta_1 \in C(\bar u, \bar y, \bar\mu, S) \setminus \{0\}$. Hence $\bar\zeta_1 = 0$, but then it follows $0 \geq \alpha\gamma^2$, contradicting $\gamma > 0$. Thus $\lim_{k\to\infty} \|\eta_1^k\|_{L^2(\Omega)} = 0$ and by (3.7) also $\lim_{k\to\infty} \|\eta_2^k\|_{W^*} = 0$. It remains to show $\|\eta_3^k\|_{W^*} \overset{k\to\infty}{\longrightarrow} 0$. To this end we use the direction d from Assumption 1.1 by applying $d_k = d + \bar u - u_k$ to (3.6) and pass to the limit using $u_k^*, \eta_1^k, \eta_2^k \to 0$ to arrive at

$$0 \geq \lim_{k\to\infty} d_k S'^*_{h_k}(u_k)\eta_3^k\,. \tag{3.18}$$

The sequence $S'_{h_k}(u_k)d_k$ converges in W towards $S'(\bar u)d$ due to the convergence assumption on the derivatives 3.3.
Since $S'(\bar u)d > \delta > 0$ on $\mathcal{A}_{\bar y}$, there exists some $\epsilon > 0$, such that $S'(\bar u)d > 3\delta/4 > 0$ on $\mathcal{A}_{\bar y} + B_{\mathbb{R}^n}(0,\epsilon)$. Assuming the contrary easily gives a contradiction to $\bar\Omega$ being compact, as in the argument for K_ϵ below. Thus we have $S'_{h_k}(u_k)d_k > \delta/2$ on $\mathcal{A}_{\bar y} + B_{\mathbb{R}^n}(0,\epsilon)$ for all $k \geq K_{\delta/2}$.
Now the support of η_3^k lies in $\mathcal{A}_{S_{h_k}(u_k)}$. Because $y_k \overset{W}{\to} \bar y$ and because of (3.14) there holds $\|S_{h_k}(u_k) - \bar y\|_W \to 0$. Therefore, there exists K_ϵ, such that $\mathcal{A}_{S_{h_k}(u_k)} \subset \mathcal{A}_{\bar y} + B_{\mathbb{R}^n}(0,\epsilon)$ for all $k \geq K_\epsilon$. If not, there would exists some sequence $x_k \in \Omega\setminus(\mathcal{A}_{\bar y} + B_{\mathbb{R}^n}(0,\epsilon))$ with $(S_{h_k}(u_k))(x_k) = 0$, and some converging subsequence $x_{\tilde k} \to x$, for Ω is compact. But then $S(\bar u)(x) = \bar y(x) = 0$, contradicting $x \notin \mathcal{A}_{\bar y}$.
Hence $\langle \eta_3^k, S'_{h_k}(u_k)d_k\rangle \geq \frac{1}{2}\langle \eta_3^k, 1\rangle$ for $k > \max(K_{\delta/2}, K_\epsilon)$. Since after Lemma 3.18 $\langle \eta_3^k, 1\rangle \geq \|\eta_3^k\|_{W^*}/C$, we finally have from (3.18)

$$\lim_{k\to\infty} \|\eta_3^k\|_{W^*} = 0$$

and thus $\eta_k \to 0$, in contradiction to $\eta_k = \|\mathcal{F}(u_k, y_k, \mu_k, \lambda_k, \nu_k; S_{h_k})\|'$.

∎

Remark 3.13. The strict complementarity 3.8 assumed here can be dismissed completely if the second derivative $\alpha + S''^*(\bar u)(y - z + \bar\mu) + S'^*(\bar u)S'(\bar u)$ fulfills some sufficiently strong positive definiteness condition, as, for instance, being positive definite on all of $L^2(\Omega)$. The latter is true in the linear-quadratic case. If S is linear, assumption 3.7 becomes trivial.

The direction d from Assumption 1.1 plays an important role. In combination with the L^∞-convergence of S_h it ensures the stability of ($\mathbb{P}$) with respect to perturbations in S, namely the existence of admissible points for ($\mathbb{P}_h$) for sufficiently small $h > 0$, which was discussed in [CM02b].
The Lemmas 3.5 and 3.12 plugged into Lemma 2.6 and Theorem 2.2 now lead to our main result.

Theorem 3.14. *Consider a solution $(\bar{u}, \bar{y})$ of $(\mathbb{P})$ and multipliers $\bar{\mu}, \bar{\lambda}, \bar{\nu}$, such that*

$$\mathcal{F}(\bar{u}, \bar{y}, \bar{\mu}, \bar{\lambda}, \bar{\nu}; S) = 0\,.$$

Let the Assumptions 3.1-3.3 as well as 3.7-3.11 be fulfilled. Then there exists $h_0 > 0$ and $\sigma > 0$ as in Lemma 3.12 , such that for $0 < h < h_0$ problem $(\mathbb{P}_h)$ admits a solution (u_h, y_h) that fulfills

$$\|u_h - \bar{u}\|_{L^2(\Omega)} + \|y_h - \bar{y}\|_W \leq \frac{1}{\sigma}(\|(S_h'^*(\bar{u}) - S'^*(\bar{u}))(\bar{y} - z + \bar{\mu})\|_{L^2(\Omega)} + 2\|S_h(\bar{u}) - S(\bar{u})\|_W)\,.$$

Remark 3.15. The assertions of Theorem 3.14 remain true, if one allows for varying bounds $a_h \overset{L^2(\Omega)}{\longrightarrow} a_0 = a$ and $b_h \overset{L^2(\Omega)}{\longrightarrow} b_0 = b$ with $a_h \geq a$ and $b_h \leq b$ for all $h > 0$. The latter assumption is crucial to the lower semicontinuity of the function $f(x, p)$ with respect to $p = (S_h, a_h, b_h)$. One has to apply the following changes in the function F. Replace $N(u)$ by

$$N(u, a_h, b_h) = \begin{cases} \{v \in L^2(\Omega) \mid \langle v, c - u_{a_h}^{b_h} \rangle_{L^2(\Omega)} \leq 0\,,\ \forall c \in U_{ad}\} & \text{if } a \leq u \leq b \\ \emptyset & \text{else} \end{cases}\,,$$

with $u_{a_h}^{b_h} = \min(b_h, \max(a_h, u))$. Further, append these two lines to F from problem $(\mathbb{P}_h)$

$$\begin{pmatrix} L^2(\Omega)_{\geq 0} - (b_h - u) \\ L^2(\Omega)_{\geq 0} - (u - a_h) \end{pmatrix}$$

making F a set valued function into $L^2(\Omega) \times W^2 \times L^2(\Omega)^2$. This has been left out merely for the notational inconvenience involved. The modification of Lemma 3.5 and its proof is straightforward, as it is for Lemma 3.12. The only change to be made here is to choose

$$d_k = d + \bar{u} - u_k + [a_h - a]_{\mathcal{A}_{u_k}^{a_h}} + [b_h - b]_{\mathcal{A}_{u_k}^{b_h}}\,,$$

converging towards d as well. Again, for $\phi \in L^2(\Omega)$, $[\phi]_{\mathcal{A}}$ denotes the restriction $[\phi]_{\mathcal{A}}(x) = \phi(x)$ for $x \in \mathcal{A}$, zero otherwise. This is important because of the two additional multipliers $\eta_4^k, \eta_5^k \in L^2(\Omega)_{\geq 0}$.
Finally, the two terms $\|a_h - a\|_{L^2(\Omega)}$ and $\|b_h - b\|_{L^2(\Omega)}$ emerge on the right hand side of the error estimate given in the theorem.

Remark 3.16. Note also, that a slight alteration of our technique applies to purely control constrained problems, yielding

$$\|u_h - \bar{u}\|_{L^2(\Omega)} + \|y_h - \bar{y}\|_{L^2(\Omega)} \leq \frac{1}{\sigma}(\|(S_h'^*(\bar{u}) - S'^*(\bar{u}))(\bar{y} - z)\|_{L^2(\Omega)} + \|S_h(\bar{u}) - S(\bar{u})\|_{L^2(\Omega)})\,.$$

and thus an optimal order of convergence for the control u. To this end, just set $K(y) = \{0\} \in L^2(\Omega)$, eliminate the last line of F and replace $C(\bar{\Omega})$ and W by $L^2(\Omega)$ everywhere, making F a set valued function into $L^2(\Omega) \times L^2(\Omega)$. The proofs stay essentially the same and become, in fact, much simpler, as there are no multipliers in $C(\bar{\Omega})^*$ or W^* to consider. In this situation however, we could as well apply Robinson's Implicit Multifunction Theorem.

Proof of Theorem 3.14. The application of Theorem 2.2 is straightforward. As to the estimation of $f((\bar u,\bar y),S_h)$ we note, that the multipliers $\bar\mu$, $\bar\lambda$ and $\bar\nu$ are also admissible for $F((\bar u,\bar y),S_h)$. We then just apply

$$\mathrm{d}(0,F((\bar u,\bar y),S_h)) \le \frac{1}{\sigma}\|\mathcal{F}(\bar u,\bar y,\bar\mu,\bar\lambda,\bar\nu;S_h) - \underbrace{\mathcal{F}(\bar u,\bar y,\bar\mu,\bar\lambda,\bar\nu;S)}_{=0}\|_Y\,.$$

∎

The order of convergence asserted in Theorem 3.14, if any, is that of $S(\bar u)$ and $S'^*(\bar u)\mu$ at some fixed point $\bar u$ and for some fixed $\mu\in C(\bar\Omega)^*$. We make no use of uniform convergence. It holds in the semilinear case, but may not hold in other settings.
The following two Lemmas are used in the proof of Lemma 3.12.

Lemma 3.17. *Assume 3.2 and 3.11. Under the convergence Assumption from 3.10 the linear operator*

$$S_h''^*(u)(\mu) : L^2(\Omega)\to L^2(\Omega)$$

is compact and selfadjoint for any $h\ge 0$, $u\in L^\infty(\Omega)$ and $\mu\in C(\bar\Omega)^*$. In particular this holds for $\mu\in L^2(\Omega)$ or $\mu\in W$ via the canonical embeddings $W\subset L^2(\Omega)\subset C(\bar\Omega)^*$.
Assuming 3.3, the operator $S_h'(u): L^2(\Omega)\to W$ is compact as well for any $u\in L^\infty(\Omega)$, $h\ge 0$.

Proof. Since S_h is assumed to be $\mathcal{C}^2$ into $C(\bar\Omega)$ (Assumption 3.2), $S_h''(u) : L^2(\Omega)^2\to C(\bar\Omega)$ is symmetric in its two arguments, hence the selfadjointness. As to the compactness, the operators $S_h'(u)$ have finite dimensional image for $h>0$ and hence are compact. It then follows from Assumption 3.3, that $S'(u)$ can be approximated by compact operators and therefore is also compact.
Because of Assumption 3.11, the operator $S_h''^*(u)\mu$ is compact for $h>0$, since

$$\langle\mu, S_h''(u)(v,w)\rangle_{C(\bar\Omega)^*,C(\bar\Omega)} = \langle\pi_h^*T_h(\pi_h v,\,\cdot\,)^*\mu, w\rangle_{L^2(\Omega)}\,,$$

and π_h^* is compact, for π_h is compact. Note, that from the continuity of T_h there follows

$$\sup_{\|v_h\|_{V_h}=\|\tilde v_h\|_{V_h}=1}\langle\mu, T_h(v_h,\tilde v_h)\rangle \le C_\mu$$

and thus $\|T_h(\pi_h v,\,\cdot\,)^*\mu\|_{V_h^*}\le C_\mu$ for $\|v\|_{L^2(\Omega)}\le 1$. Thus

$$v\mapsto \pi_h^*T_h(\pi_h v,\,\cdot\,)^*\mu = S_h''^*(u)\mu$$

is compact. Now the convergence 3.10 yields

$$\|S_h''^*(u)\mu - S''^*(u)\mu\|_{\mathcal{L}(L^2(\Omega))} = \sup_{\|w\|=\|v\|=1}\langle\mu, S_h''(u)(v,w)-S''(u)(v,w)\rangle_{C(\bar\Omega)^*,C(\bar\Omega)} \xrightarrow{h\to 0} 0$$

and thus the compactness of $S''^*(u)\mu$. ∎

Lemma 3.18. *Consider some Banach space W, and let the inclusion $W\subset C(\bar\Omega)$ be continuous. Then there exists some $C>0$, such that for all*

$$w^*\in W^+_{\ge 0} = \{w^*\in W^* \mid \forall w\in W:\ w\ge 0 \Rightarrow \langle w^*,w\rangle_{W^*,W}\ge 0\}$$

there holds $C\langle w^*,1\rangle_{W^*,W}\ge\|w^*\|_{W^*}$.

Proof. There exists $C>0$, such that $C\|w\|_W\ge\|w\|_{C(\bar\Omega)}$. Assume now, that there exists some $w\in B_W(0,1)$ with $\langle w^*,w\rangle > C\langle w^*,1\rangle$. Then $\|w\|_{C(\bar\Omega)}\le C$ and hence $C-w\ge 0$. But $\langle w^*, C-w\rangle<0$, contradicting $w^*\in W^+_{\ge 0}$. ∎

4 Example

Consider some convex polygonal domain $\Omega \subset \mathbb{R}^2$, and let $S : L^2(\Omega) \to W^{2,2}(\Omega)$ be the solution operator of the equation

$$\begin{aligned} -\Delta y + y^3 &= u \quad \text{on } \Omega \\ y &= 0 \quad \text{on } \partial\Omega \,. \end{aligned} \tag{4.19}$$

We want to approximate the solution of

$$\begin{gathered} \min_{u \in L^2(\Omega),\, \hat{y} \in C(\bar{\Omega})} J(u, \hat{y}) = \|\hat{y} - z\|^2_{L^2(\Omega)} + \alpha \|u\|^2_{L^2(\Omega)} \\ \text{s.t.} \\ \hat{y} = \hat{S}(u) = S(u) + 1, \quad a \le u \le b, \quad \hat{y} \ge 0 \,, \end{gathered} \tag{4.20}$$

with a, b and z as in (1.1). An application of the standard Implicit Function Theorem as in Theorem 2.5 in [CM02a] shows, that $S \in C^2(L^2(\Omega), W^{1,4}(\Omega))$. We will choose W as $W^{1,4}(\Omega)$ endowed with an equivalent differentiable norm.

Lemma 4.1. *The operator S belongs to $S \in C^2(L^2(\Omega), W^{2,2}(\Omega))$ and its derivative $S'(u) : L^2(\Omega) \to W^{2,2}(\Omega)$, $\delta u \mapsto \delta y$ is the solution operator of*

$$\Delta \delta y + 3y(u)^2 \delta y = \delta u \text{ in } \Omega, \quad \delta y = 0 \text{ on } \partial\Omega \,,$$

while its second derivative takes the form

$$\delta^2 y = S''(u)\delta u_1 \delta u_2 = -S'(u) \left(6S(u)(S'(u)\delta u_1)(S'(u)\delta u_2)\right) \tag{4.21}$$

Proof. Existence of a unique solution in $W^{2,2}(\Omega)$ is standard. Consider the operator $A : W^{2,2}(\Omega) \cap W^{1,2}_0 \to L^2(\Omega)$

$$A(y) = \Delta y + y^3 \,,$$

which lies in $C^2(W^{2,2}(\Omega) \cap W^{1,2}_0, L^2(\Omega))$ with its derivative

$$A'(y)\delta y = \Delta \delta y + 3y^2 \delta y$$

being an isomorphism because of $y^2 \ge 0$. Now, since the map $G : L^2(\Omega) \times W^{2,2}(\Omega) \to L^2(\Omega)$

$$G(u, y) = A(y) - u$$

is twice continuously differentiable, the Implicit Function Theorem yields the first part of the lemma. The form of $S''(u)$ follows from the observation, that $A'(S(u)) \circ S'(u) = \mathrm{Id}_{L^2(\Omega)}$ and hence

$$A''(S(u)) \left(S'(u)\delta u_1, S'(u)\delta u_2\right) + A'(S(u))S''(u) \left(\delta u_1, \delta u_2\right) = 0 \,.$$

∎

The states y are discretized using some quasiuniform and geometrically conformal family of triangulations $\{\tau_h\}$ of Ω, $h \in \{h_n\}_{n \in \mathbb{N}}$

$$y_h \in V_h = \left\{ y \in C(\bar{\Omega}) \cap W^{1,2}_0 \mid \forall T \in \tau_h : \; y \text{ is linear on } T \right\} ,$$

with $h = \max_{T\in\mathcal{T}_h} \operatorname{diam}(T)$. The discretized equation now reads

$$\int_\Omega \nabla y_h \cdot \nabla\varphi + y_h^3\varphi \,\mathrm{d}x = \int_\Omega u\varphi \,\mathrm{d}x\,, \quad \forall \varphi \in V_h\,. \tag{4.22}$$

Analogously to the non-discretized equation one may investigate differentiability of $S_h : L^2(\Omega) \to W^{1,4}(\Omega)$ with $y_h = S_h(u)$ by the operator

$$A_h : (V_h, \|\cdot\|_{W^{1,4}(\Omega)}) \to (V_h, \|\cdot\|_{W^{1,\frac{4}{3}}(\Omega)})^*\,, \quad A_h(y_h) = -\Delta y_h + y_h^3\,.$$

In fact $A_h = (\imath_h^2)^* \circ A \circ \imath_h^1$ with the inclusions $\imath_h^1 : V_h \to W^{1,4}$, $\imath_h^2 : V_h \to W^{1,\frac{4}{3}}$ and the operator

$$A : W^{1,4}(\Omega) \to (W^{1,\frac{4}{3}})^*(\Omega)\,, \quad A(y) = \Delta y + y^3\,,$$

which lies in $C^2(W^{1,4}(\Omega), (W^{1,\frac{4}{3}})^*)$. The derivatives are

$$A'(y)\delta y = \Delta\delta y + 3y^2\delta y\,, \quad A''(y)\delta y_1\delta y_2 = 6y\delta y_1\delta y_2\,.$$

Hence A_h is is also twice continuously differentiable and

$$\langle A_h'(y_h)\delta y_h, \varphi\rangle = \int_\Omega \nabla\delta y_h \cdot \nabla\varphi + 3y_h^2\delta y_h\varphi \,\mathrm{d}x\,, \quad \forall\varphi \in V_h\,.$$

Because of $y_h^2 \geq 0$, $A_h'(y_h)$ is an isomorphism in V_h for any $y_h \in V_h$ and thus the application of the ordinary Implicit Function Theorem to

$$G_h : (V_h, \|\cdot\|_{W^{1,4}(\Omega)}) \times L^2(\Omega) \to (V_h, \|\cdot\|_{W^{1,\frac{4}{3}}(\Omega)})^*\,, \quad G(y_h, u) = A_h(y_h) - u$$

yields $S_h \in C^2(L^2(\Omega), W^{1,4}(\Omega))$, and Assumption 3.2 is fulfilled.

Lemma 4.2. *Let y and y_h be the solution of 4.19 and 4.22 respectively, for some fixed $u \in L^4(\Omega)$. Then there holds*

$$\|y - y_h\|_W \leq Ch\,.$$

Proof. From Theorem 1 in [CM02b] we already know

$$\|y - y_h\|_\infty \leq C_1 h\|y\|_{W^{2,2}(\Omega)}\,.$$

Introducing the auxiliary state $\tilde{y}$ as the solution of the linear equation

$$-\Delta\tilde{y} = u - y_h^3\,, \quad \tilde{y} = 0 \text{ on } \partial\Omega \tag{4.23}$$

we first obtain from the theory of linear elliptic equations

$$\|y - \tilde{y}\|_{W^{2,2}(\Omega)} \leq C_2\|y^3 - y_h^3\|_\infty \leq C_3\|y - y_h\|_\infty\,,$$

the second inequality following from $\|y\|_\infty$ and $\|y_h\|_\infty$ being bounded. Now the discretization of 4.23 yields an $\tilde{y}_h$, that equals y_h. Because of the $W^{1,4}$-stability of the discretization of the linear equation (4.23) (see for example Theorem 8.5.3 from [BS08]) and because of the stability property $\|y\|_\infty \leq C_4\|u\|_{L^4(\Omega)}$ we have

$$\|\tilde{y} - y_h\|_W \leq C_5 h(\|u\|_{L^4(\Omega)} + \|y_h^3\|_{L^4(\Omega)}) \leq C_6 h(\|u\|_{L^4(\Omega)} + \|u\|^3_{L^4(\Omega)})$$

∎

Lemma 4.3. *Let $\delta y = S'(u)\delta u$, then δy solves the equation*

$$\Delta\delta y + 3y(u)^2\delta y = \delta u \text{ in } \Omega, \quad \delta y = 0 \text{ on } \partial\Omega\,,$$

and let $\delta y_h = S_h'(u_h)\delta u$, and thus

$$\int_\Omega \nabla\delta y_h \cdot \nabla\varphi + 3y_h(u_h)^2\delta y_h\,\varphi dx = \int_\Omega \delta u\,\varphi dx \quad \forall\varphi \in V_h\,,$$

then there holds for $u_h = u$

$$\|\delta y_h - \delta y\|_{C(\bar\Omega)} \le Ch\|\delta u\|_{L^2(\Omega)}$$

and in case of $u_h \overset{L^2(\Omega)}{\longrightarrow} u$ with u_h bounded in $L^\infty(\Omega)$ we have $\delta y_h \overset{W}{\longrightarrow} \delta y$ uniformly for all $\|\delta u\|_{L^2(\Omega)} \le 1$.

Proof. Consider $u_h = u$, for all $h > 0$. We introduce the discrete state $\delta\tilde y_h$, satisfying

$$\int_\Omega \nabla\delta\tilde y_h \cdot \nabla\varphi + 3y(u)^2\delta\tilde y_h\,\varphi \mathrm{d}x = \int_\Omega \delta u\,\varphi \mathrm{d}x \quad \forall\varphi \in V_h\,,$$

then, on the one hand we have $\|\delta\tilde y_h - \delta y\|_{C(\bar\Omega)} \le Ch\|\delta u\|_{L^2(\Omega)}$ and on the other hand we have from the $W^{1,p}$ stability stated in Theorem 8.5.3 in [BS08]

$$\|\delta y_h - \delta\tilde y_h\|_W \le C\|y_h(u)^2 - y(u)^2\|_\infty\|\delta y_h\|_{L^2(\Omega)} \tag{4.24}$$

and hence finally

$$\|\delta y_h - \delta\tilde y_h\|_W \le Ch\|\delta u\|_{L^2(\Omega)}\,.$$

As for $u_h \to u$ in $L^2(\Omega)$, because $\|u_h\|_\infty$ is bounded, this implies $u_h \overset{*}{\rightharpoonup} u$ in $L^\infty(\Omega)$. Thus Theorem 9 from [CM02b] implies $S_h(u_h) \to S(u)$ in $C(\bar\Omega)$ and using (4.24) we obtain

$$\|\delta y_h - \delta\tilde y_h\|_W \to 0 \quad \text{uniformly for } \|\delta u\|_{L^2(\Omega)} \le 1\,.$$

The harder part is now to infer $\|\delta\tilde y_h - \delta y\|_W \to 0$ uniformly in δu. As a corollary to the $W^{1,p}$-stability one obtains

$$\|\delta\tilde y_h - \delta y\|_W \le C \inf_{v_h\in V_h} \|v_h - \delta y\|_W\,.$$

The usual estimates to the right hand side assume $\delta y \in W^{2,4}(\Omega)$, but here we only have $\delta y \in W^{2,2}(\Omega)$. Nevertheless, the element-wise estimate given in Theorem 16.2 from [CL91] yields for the interpolation operator I_h and any $T \in \tau_h$, $v \in W^{2,2}(\Omega)$

$$\|v - I_h v\|_{W^{1,q}(T)} \le C\text{meas}(T)^{\frac1q - \frac1p} h_T|v|_{W^{2,p}(T)}\,,$$

with h_T being the diameter of T and the constant C independent of h, v and T. Since the triangulations are quasi uniform, we have $\text{meas}(T) \ge ch_T^2$. For $\|v\|_{W^{2,2}(\Omega)} \le 1$ and h_T sufficiently small this implies

$$\|v - I_h v\|^4_{W^{1,4}(T)} \le \|v - I_h v\|^2_{W^{1,4}(T)} \le Ch_T|v|^2_{W^{2,2}(T)}\,,$$

and hence $\|v - I_h v\|_{W^{1,4}(\Omega)} \overset{h\to 0}{\longrightarrow} 0$ uniformly for all $\|v\|_{W^{2,2}(\Omega)} \le 1$. ∎

The convergence of the second derivative follows from the convergence of the first derivative, because of the structure of 4.21 and its discrete counterpart.
Now we finally have all ingredients assembled, to apply Theorem 3.14, yielding the following

Theorem 4.4. *Let $(\bar{u}, \bar{y})$ solve problem (4.20). Assume further, that the linearized Slater condition 1.1 as well as the strict complementarity Assumptions 3.8 - 3.9 hold at $(\bar{u}, \bar{y})$, and that $(\bar{u}, \bar{y})$ satisfies the second order sufficient condition 3.7. Then there exists $C > 0$, such that for h sufficiently small, there exist $\bar{u}_h$, $\bar{y}_h$ solving $(\mathbb{P}_h)$ with $\hat{S}_h = (S_h + 1)$ and*

$$\|\bar{u}_h - \bar{u}\|_{L^2(\Omega)} + \|\bar{y}_h - \bar{y}\|_{W^{1,4}(\Omega)} \leq Ch\,.$$

Together with $(\mathbb{P})$ and Assumption 1.1, Assumption 3.7 ensures strict local optimality of $(\bar{u}, \bar{y})$, as was shown in [CDLRT08]. Notes on the implementation of $(\mathbb{P}_h)$ can be found in [DH07]. The approach described there also applies to a nonlinear state equation.
Finally one has to deal with non-unique multipliers μ, λ insofar, as the strict complementarity has to hold for all of them. Therefore, in one last lemma, we demonstrate a case, in which the multipliers are in fact unique.

Lemma 4.5. *Let $(\bar{u}, \bar{y})$ solve $(\mathbb{P})$ with S defined by equation (4.19) . Suppose we can separate $\mathcal{A}_y$ and $\mathcal{A}_u$ by two disjoint open sets*

$$\mathcal{O}_{\bar{y}} \supset \mathcal{A}_{\bar{y}} \quad \textit{and} \quad \mathcal{O}_{\bar{u}} \supset \mathcal{A}_{\bar{u}}$$

with Lipschitz boundaries. Then the multipliers $\bar{\mu} \in \mathcal{K}(\bar{u}, \bar{y})$ and $\bar{\lambda}$ are unique.

Proof. One way to proof the lemma involves the necessary and sufficient conditions for uniqueness of Lagrange multipliers given in [Sha97]. Here the more direct approach is to make use of the fact that, given two different multipliers $\mu_1, \mu_2 \in \mathcal{K}(\bar{u}, \bar{y})$, the difference $S'^*(\bar{u})(\mu_1 - \mu_2)$ equals zero on the inactive set $\Omega \setminus \mathcal{A}_{\bar{u}}$ of $\bar{u}$. Construction of some $v \in L^2(\Omega)$, that equals zero on the active set $\mathcal{A}_{\bar{u}}$ but satisfies $\langle \mu_1 - \mu_2, S'(\bar{u})v \rangle_{C(\bar{\Omega})^*, C(\bar{\Omega})} \neq 0$ thus yields a contradiction. First one can assume w.l.o.g. $\mathcal{O}_{\bar{y}} \subset\subset \Omega$, because $\bar{y}$ is continuous and equals 0 on the boundary of Ω and -1 on $\mathcal{O}_{\bar{y}}$. Note also, that since Ω is bounded, there exists $R > 0$, such that $B(0, R) \supset \bar{\Omega}$. We proceed by choosing

$$\begin{aligned}
\delta y_v^1 \in \quad & C^\infty(\bar{\mathcal{O}}_{\bar{y}}) && \text{s.t. } \langle \mu_1 - \mu_2, \delta y_v^1 \rangle_{C(\bar{\Omega})^*, C(\bar{\Omega})} \neq 0 \\
\delta y_v^2 \in \quad & C^\infty(\bar{\mathcal{O}}_{\bar{u}}) && \delta y_v^2 \equiv 0 \\
\delta y_v^3 \in \quad & C^\infty(B(0, R) \setminus \Omega) && \delta y_v^3 \equiv 0\,.
\end{aligned}$$

Using an extension Theorem (e.g. Thm. 5, Ch. VI of [Ste70]), one can now extend this triple into an arbitrarily smooth function δy_v, defined on $\mathbb{R}^2$. This is now a strong solution of

$$\begin{aligned}
-\Delta \delta y_v + 3\bar{y}^2 \delta y_v &= v \quad \text{in } \Omega \\
\delta y_v &= 0 \quad \text{on } \partial\Omega\,,
\end{aligned}$$

with $v \in C(\bar{\Omega})$, because $\bar{y} \in C(\bar{\Omega})$. Furthermore the restriction $[v]_{\mathcal{A}_u}$ of v to $\mathcal{A}_u$ equals zero. ∎

References

[ACT02] Nadir Arada, Eduardo Casas and Fredi Tröltzsch. Error estimates for the numerical approximation of a semilinear elliptic control problem. *Comput. Optim. Appl.*, 23(2):201–229, 2002.

[BS08] Susanne C. Brenner and Ridgway L. Scott. *The mathematical theory of finite element methods. 3rd ed.* Texts in Applied Mathematics 15. New York, NY: Springer. xvii, 397 p., 2008.

[CDLRT08] Eduardo Casas, Juan Carlos De Los Reyes and Fredi Tröltzsch. Sufficient second-order optimality conditions for semilinear control problems with pointwise state constraints. *SIAM J. Optim.*, 19(2):616–643, 2008.

[CL91] Philippe G. Ciarlet and Jacques L. Lions, editors. *Handbook of numerical analysis. Volume II: Finite element methods (Part 1).* Amsterdam etc.: North-Holland. ix, 928 p., 1991.

[CLSW98] Francis H. Clarke, Yuri S. Ledyaev, Ron J. Stern and Peter R. Wolenski. *Nonsmooth analysis and control theory.* Graduate Texts in Mathematics 178. New York, NY: Springer. xii, 276 p., 1998.

[CM02a] Eduardo Casas and Mariano Mateos. Second order optimality conditions for semilinear elliptic control problems with finitely many state constraints. *SIAM J. Control Optimization*, 40(5):1431–1454, 2002.

[CM02b] Eduardo Casas and Mariano Mateos. Uniform convergence of the FEM applications to state constrained control problems. *Comput. Appl. Math.*, 21(1):67–100, 2002.

[CR08a] Svetlana Cherednichenko and Arnd Rösch. Error estimates for the discretization of elliptic control problems with pointwise control and state constraints. *Computational Optimization and Applications*, Online First, 2008.

[CR08b] Svetlana Cherednichenko and Arnd Rösch. Error estimates for the regularization of optimal control problems with pointwise control and state constraints. *Z. Anal. Anwend.*, 27(2):195–212, 2008.

[DH07] Klaus Deckelnick and Michael Hinze. A finite element approximation to elliptic control problems in the presence of control and state constraints. *Preprint, Hamburger Beitr. z. Ang. Math. 2007-01*, 2007.

[Die75] Joseph Diestel. *Geometry of Banach spaces. Selected topics.* Lecture Notes in Mathematics. 485. Berlin-Heidelberg-New York: Springer-Verlag. XI, 282 p., 1975.

[Els96] Jürgen Elstrodt. *Maß- und Integrationstheorie. (Measure and integration theory.).* Springer-Lehrbuch. Berlin: Springer. xvi, 398 p., 1996.

[Hin05] M. Hinze. A variational discretization concept in control constrained optimization: The linear-quadratic case. *Comput. Optim. Appl.*, 30(1):45–61, 2005.

[HM07] Michael Hinze and Christian Meyer. Stability of infinite dimensional control problems with pointwise state constraints. *Preprint No. 1236, Weierstraß Institute*, 2007.

[HPUU09] M. Hinze, R. Pinnau, M. Ulbrich and S. Ulbrich. *Optimization with PDE constraints.* Mathematical Modelling: Theory and Applications 23. Dordrecht: Springer. xi, 270 p., 2009.

[LZ99] Yuri S. Ledyaev and Qiji J. Zhu. Implicit multifunction theorems. *Set-Valued Analysis*, 7:209–238, 1999.

[Mey08] C. Meyer. Error estimates for the finite-element approximation of an elliptic control problem with pointwise state and control constraints. *Control and Cybernetics*, 37(1):51–83, 2008.

[Mor05] Boris S. Mordukhovich. *Variational analysis and generalized differentiation. I: Basic theory. II: Applications.* Grundlehren der Mathematischen Wissenschaften 330/331. Berlin: Springer. xxii, 579 p., xxii, 610 p., 2005.

[MRT06] Christian Meyer, Arnd Rösch and Fredi Tröltzsch. Optimal control of PDEs with regularized pointwise state constraints. *Comput. Optim. Appl.*, 33(2-3):209–228, 2006.

[MTV08] Pedro Merino, Fredi Tröltzsch and Boris Vexler. Error estimates for the finite element approximation of a semilinear elliptic control problem with state constraints and finite dimensional control space. *Preprint, DFG SPP1253-24-04*, 2008.

[Rob80] Stephen M. Robinson. Strongly regular generalized equations. *Math. Oper. Res.*, 5:43–62, 1980.

[Sha97] Alexander Shapiro. On uniqueness of Lagrange multipliers in optimization problems subject to cone constraints. *SIAM J. Optim.*, 7(2):508–518, 1997.

[Ste70] Elias M. Stein. *Singular integrals and differentiability properties of functions.* Princeton, N.J.: Princeton University Press. XIV, 287 p., 1970.