Deutsche Forschungsgemeinschaft

Priority Program 1253

Optimization with Partial Differential Equations

Michael Hintermüller and Michael Hinze

Moreau-Yosida Regularization in State Constrained Elliptic Control Problems: Error Estimates and Parameter Adjustment

March 2008

Preprint-Number SPP1253-08-04

---

http://www.am.uni-erlangen.de/home/spp1253

Moreau-Yosida regularization in state constrained elliptic control problems: error estimates and parameter adjustment

Michael Hintermüller* Michael Hinze†

Abstract: An adjustment scheme for the regularization parameter of a Moreau-Yosida-based regularization, or relaxation, approach to the numerical solution of pointwise state constrained elliptic optimal control problems is introduced. The method utilizes error estimates of an associated finite element discretization of the regularized problems for the optimal selection of the regularization parameter in dependence on the mesh size of discretization and error estimates for the approximation error due to regularization. The theoretical results are verified numerically.

Mathematics Subject Classification (2000): 49M15, 49M37, 65K05, 65N12, 65N30, 90C33

Keywords: Elliptic optimal control problem, error estimates, Moreau-Yosida-based regularization, pointwise state constraints.

1 Introduction

In this paper we are interested in the numerical analysis of a Moreau-Yosida-based regularization of PDE-constrained optimization problems subject to pointwise state constraints. This regularization scheme along with a primal-dual path-following algorithm was considered in [9] in order to approximate the measure-valued Lagrange multiplier of the original state constrained problem by a sequence of regular multipliers. Advantages of such a regularization procedure are the availability of highly efficient solvers for the regularized problem (such as semismooth Newton methods) and the induced numerical stability with respect to the mesh size of discretization. This latter property, however, relies on an appropriate tuning of the regularization parameter. Optimal adjustment strategies link the regularization parameter to the mesh size in order to balance the regularization and discretization errors properly. As such parameter selection rules are currently not available for the type of regularization considered here, the aim of this work is to close this gap and, thus, to allow for fine tuned implementations of corresponding solution algorithms.

In order to set up a model problem class for our subsequent discussion, let $\Omega \subset \mathbb{R}^d$ ($d = 2, 3$) be a bounded domain with a smooth boundary $\partial\Omega$ and consider the second order linear

*Department of Mathematics, University of Sussex, Falmer, Brighton, UK, and Institute for Mathematics and Scientific Computing, University of Graz, Graz, Austria

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

elliptic differential operator (in divergence form)

$$Ay := -\sum_{i,j=1}^{d} \partial_{x_j}\left(a_{ij}y_{x_i}\right) + \sum_{i=1}^{d} b_i y_{x_i} + cy,$$

where, for simplicity, the coefficients a_{ij}, b_i and c are assumed to be smooth functions in $\bar{\Omega}$. Let $a(\cdot,\cdot)$ denote the bilinear form associated with the differential operator A. For some constant $c_1 > 0$, it is assumed to satisfy

$$a(v,v) \geq c_1\|v\|_{H^1}^2 \qquad \text{for all } v \in H^1(\Omega). \tag{1.1}$$

Next let $f \in (H^1(\Omega))'$. Then it follows that the elliptic boundary value problem

$$\begin{aligned} Ay &= f \quad \text{in } \Omega, \\ \textstyle\sum_{i,j=1}^{d} a_{ij}u_{x_i}\nu_j &= 0 \quad \text{on } \partial\Omega \end{aligned} \tag{1.2}$$

admits a unique solution $y \in H^1(\Omega)$, which we denote by $y = \mathcal{G}(f)$. Here, ν is the outward unit normal to $\partial\Omega$. Furthermore, if $f \in L^2(\Omega)$, then the solution y belongs to $H^2(\Omega)$ and satisfies $\|y\|_{H^2} \leq C\|f\|$, where $\|\cdot\|$ denotes the $L^2(\Omega)$-norm.

The Hilbert space of controls is denoted by $(U, (\cdot,\cdot)_U)$ and is identified with its dual. Further, $B : U \to L^2(\Omega)$ is a linear, continuous operator which models the impact of the control action. Subsequently we study the following model problem class:

$$\begin{aligned} &\text{minimize} \quad J(w) = \tfrac{1}{2}\textstyle\int_\Omega |\mathcal{G}(Bw) - y_0|^2 + \tfrac{\alpha}{2}\|w\|_U^2 \quad \text{over} \quad w \in U \\ &\text{subject to } \mathcal{G}(Bw) \leq b \text{ a.e. in } \Omega, \end{aligned} \tag{1.3}$$

where $\alpha > 0$, $y_0 \in H^1(\Omega)$ and $b \in W^{2,\infty}(\Omega)$ are given. Here and throughout, "a.e." stands for "almost everywhere". We also invoke the following constraint qualification (Slater condition):

$$\text{There exist } \hat{u} \in U, \tau > 0: \quad \mathcal{G}(B\hat{u}) \leq b - \tau \text{ a.e. in } \Omega. \tag{1.4}$$

If B is invertible, then our model problem satisfies the Slater condition automatically.

Standard techniques guarantee that problem (1.3) admits a unique solution $u \in U$. Moreover, from [4, Theorem 2] we deduce that there exist unique functions $\lambda \in \mathcal{M}(\bar{\Omega})$ and $p \in L^2(\Omega)$ satisfying, together with $y = \mathcal{G}(Bu)$, the dual system

$$\int_\Omega pAv = \int_\Omega (y - y_0)v + \int_{\bar{\Omega}} v\,d\lambda \qquad \forall v \in H^2(\Omega) \text{ with } \sum_{i,j=1}^{d} a_{ij}v_{x_i}\nu_j = 0 \text{ on } \partial\Omega, \tag{1.5}$$

$$B^*p + \alpha u = 0 \quad \text{in } U, \tag{1.6}$$

$$\lambda \geq 0,\ y \leq b \text{ a.e. in } \Omega \text{ and } \int_{\bar{\Omega}} (b - y)\,d\lambda = 0. \tag{1.7}$$

Here, $\mathcal{M}(\bar{\Omega})$ denotes the space of Radon measures, which is defined as the dual space of $C_0(\bar{\Omega})$, and B^* is the adjoint of B. It is endowed with the norm

$$\|\lambda\|_{\mathcal{M}(\bar{\Omega})} = \sup_{f \in C_0(\bar{\Omega}), |f| \leq 1} \int_{\bar{\Omega}} f\,d\lambda.$$

Subsequently we also use $\langle \lambda, b-y\rangle$ instead of $\int_{\bar\Omega}(b-y)d\lambda$.

A finite element analysis of problem (1.3) was carried out in [7] (compare also [6]) yielding the following error bounds:

$$\|u-u_h\|_U,\ \|y-y_h\|_{H^1} = \begin{cases} O(h^{\frac{1}{2}}), & \text{if } d=2,\\ O(h^{\frac{1}{4}}), & \text{if } d=3, \end{cases} \tag{1.8}$$

where u_h denotes the variational discrete control (see [11]) and y_h the associated piecewise linear and continuous discrete state. If, in addition, $Bu \in W^{1,s}(\Omega)$ then

$$\|u-u_h\|_U,\ \|y-y_h\|_{H^1} \le Ch^{\frac{3}{2}-\frac{d}{2s}}\sqrt{|\log h|}.$$

As noted earlier, our aim is to investigate a finite element approximation of a Moreau-Yosida-based regularization (or, alternatively, augmented Lagrangian-type penalization) technique for the numerical solution of (1.3) and to provide optimal adjustment strategies for the regularization/penaltization parameter with respect to a given finite element mesh. We emphasize that our subsequent results readily carry over, e.g., to the case of Dirichlet (rather then Neumann) boundary conditions. From here onwards and without loss of generality it is convenient to set $b\equiv 0$.

The Moreau-Yosida-regularized version of (1.3) reads

$$\min_{w\in U} J(w) = \frac{1}{2}\int_\Omega |\mathcal{G}(Bw)-y_0|^2 + \frac{\alpha}{2}\|w\|_U^2 + \frac{1}{2\gamma}\int_\Omega |\max(0,\bar\lambda+\gamma\,\mathcal{G}(Bw))|^2, \tag{1.9}$$

where $\gamma>0$ denotes the regularization (or penalization) parameter and $\bar\lambda \ge 0$, $\bar\lambda \in L^2(\Omega)$ a fixed shift-parameter. The problem (1.9) admits a unique solution $u^\gamma \in U$; see [9]. Furthermore, there exists a unique $p^\gamma \in H^2(\Omega)$ satisfying the adjoint system

$$a(v,p^\gamma) = \int_\Omega (y^\gamma - y_0)v + \int_\Omega (\bar\lambda+\gamma\, y^\gamma)^+ v \qquad \forall v\in H^1(\Omega), \tag{1.10}$$

$$B^* p^\gamma + \alpha u^\gamma = 0 \text{ in } U, \tag{1.11}$$

where $y^\gamma = \mathcal{G}(Bu^\gamma)$ and $(\cdot)^+ = \max(0,\cdot)$ in the pointwise sense. If $\bar\lambda \equiv 0$ and y^γ is feasible for (1.3), then $y^\gamma = y$ and $u^\gamma = u$. From our subsequent results we conclude that, given a mesh size of discretization h, there is an upper bound for γ at which the error due to regularization is of the order of the discretization error. Increasing γ beyond this h-dependent threshold does not improve the overall approximation error and, as (1.9) is harder to solve the larger γ becomes, would result in unnecessary extra work in the solution procedure.

The rest of the paper is organized as follows: In Section 2 we investigate the dependence of problem (1.9) on the parameter γ. For the finite element analysis developed in this work we provide uniform bounds with respect to γ on $|y^\gamma|_2$, $\frac{1}{\gamma}|p^\gamma|_2$, and we prove the decay rate

$$\|(y^\gamma)^+\| \le C\omega(\gamma^{-1})\gamma^{-1/2},$$

with $\omega(z)\downarrow 0$ for $z\downarrow 0$, as well as the error estimate

$$\|y^\gamma - y\|_{H^1} + \|u^\gamma - u\|_U \le C\|u^\gamma - u\|_U \le \frac{C}{\sqrt{\alpha}}\left(h^{1-\frac{d}{p}} + \gamma^{-\frac{1}{2}}h^{-\frac{d}{2}}\right)^{\frac{1}{2}} + \frac{C}{\sqrt{\alpha\gamma}}\|\bar\lambda\|,$$

where y and u denote the solutions of (1.3) and $0 < h \le 1$ is arbitrary. Here and below we use the seminorm $|y|_2 = \|A_0 y\|$, where A_0 denotes the leading part of the differential operator A, i.e.,

$$A_0 = -\sum_{i,j=1}^{d} \partial_{x_j}(a_{ij} y_{x_i}).$$

In Section 3 we present the finite element analysis of problem (1.9). Among other aspects we prove the error bounds

$$\|y^\gamma - y_h^\gamma\|_{H^1} + \|u^\gamma - u_h^\gamma\|_U \le C\left(h + \gamma h^2\right),$$

and also the following estimate which is uniform with respect to γ:

$$\|y^\gamma - y_h^\gamma\|_{H^1} + \|u^\gamma - u_h^\gamma\|_U \le C h^{1-\frac{d}{4}}.$$

Here and below, y_h^γ, u_h^γ denote the finite element approximations to y^γ and u^γ, respectively. We note that the latter estimate is in the spirit of (1.8). In Section 4 we discuss the overall errors

$$\|y - y_h^\gamma\|_{H^1} \sim \|y - y^\gamma\|_{H^1} + \|y^\gamma - y_h^\gamma\|_{H^1} \text{ and } \|u - u_h^\gamma\|_U \sim \|u - u^\gamma\|_U + \|u^\gamma - u_h^\gamma\|_U$$

and propose strategies for adjusting γ to h. In Section 4 we present numerical results which confirm our theoretical findings.

We point out that error estimates for the Moreau-Yosida regularization of pointwise state constraints are currently not available in the literature. In that respect, the present contribution closes this gap and allows a numerically fine-tuned implementation of the primal-dual path following algorithm in [9]. Further, our analysis is unconditional. This is in contrast to currently available results, e.g., for Lavrentiev regularization of pointwise state constraints [15, 14]. Moreover, compared to the latter strategy (or techniques relying on interior-point treatments) our Moreau-Yosida based approach does not rely on a pointwise regularization rather it relaxes the original problem by using an L^2-averaging of the constraint violation. Compared to pointwise approaches our averaging technique therefore results in weaker requirements.

2 Analysis of the Moreau-Yosida regularized problem (1.9)

For our error analysis in the subsequent section we need several results concerning the boundedness and convergence behavior of $(y^\gamma, u^\gamma, p^\gamma)$ as $\gamma \to \infty$. Throughout we use the notation

$$J^\gamma(v) = J(v) + \frac{1}{2\gamma}\int_\Omega \left((\bar\lambda + \gamma\,\mathcal{G}(Bv))^+\right)^2.$$

We start by observing that for $\gamma \ge 1$

$$J(u^\gamma) \le J^\gamma(u^\gamma) \le J^\gamma(u) \le J(u) + \frac{1}{2\gamma}\|\bar\lambda\|^2 \le J(u) + \frac{1}{2}\|\bar\lambda\|^2 =: C_u \tag{2.1}$$

Hence, $\frac{1}{2\gamma}\int_\Omega \left((\bar\lambda + \gamma\,\mathcal{G}(Bu^\gamma))^+\right)^2$ is uniformly bounded. Moreover, in [9, Proposition 2.1] it was shown that $u^\gamma \to u$ strongly in $L^2(\Omega)$. Hence, (2.1) implies

$$\|(y^\gamma)^+\| = \omega(\gamma^{-1})\,\gamma^{-1/2} \tag{2.2}$$

with

$$\omega(\gamma^{-1}) = 2\max\left(\frac{1}{\gamma}\|\bar{\lambda}\|^2, (J(u) - J(u^\gamma))^+\right)^{1/2}$$

which satisfies $\omega(z) \downarrow 0$ as $z \downarrow 0$. Moreover, (2.1) yields

$$\max(\|y^\gamma\|^2, \alpha\|u^\gamma\|_U^2) \leq 2C_u. \tag{2.3}$$

Since $y^\gamma = \mathcal{G}(Bu^\gamma)$ for all γ, from elliptic regularity, (2.3) and $\|Bu\| \leq C_B$ for some constant $C_B \geq 0$ independent of γ we infer

$$|y^\gamma|_2 = \|A_0 y^\gamma\| \leq C_0 \tag{2.4}$$

with some positive constant C_0 independent of γ. For the adjoint state p^γ we obtain

$$|p^\gamma|_2 \leq \sqrt{2C_u}\left(1 + \omega(\gamma^{-1})\sqrt{\gamma}\right) + \|y_0\|_{L^2} \leq C_0^*\left(1 + \omega(\gamma^{-1})\sqrt{\gamma}\right) \tag{2.5}$$

with a positive constant C_0^* independent of γ. In fact, for (2.5) we use the adjoint equation (1.10) together with (2.2) and (2.3).

Next we estimate the distance between (y, u) and (y^γ, u^γ).

Theorem 2.1. *Let u denote the solution of* (1.3) *and u^γ the solution of* (1.9)*. Then we have*

$$\alpha\|u - u^\gamma\|_U^2 + \|y - y^\gamma\|^2 + \gamma\|(y^\gamma)^+\|^2 \leq \frac{1}{\gamma}\|\bar{\lambda}\|^2 + \langle\lambda, y^\gamma\rangle, \tag{2.6}$$

and for the feasibility violation there holds

$$\|(y^\gamma)^+\| \leq \sqrt{\frac{2}{\gamma}}\max\left(\frac{\|\bar{\lambda}\|^2}{\gamma}, |\langle\lambda, (y^\gamma)^+\rangle|\right)^{1/2}. \tag{2.7}$$

Proof. The first order optimality systems for the original and the Moreau-Yosida-regularized problem yield

$$\alpha(u - u^\gamma) = B^*(p^\gamma - p).$$

Multiplying by $u - u^\gamma$ and using the respective first order system we get

$$\begin{aligned}
\alpha\|u - u^\gamma\|_U^2 &= \int_\Omega B(u - u^\gamma)(p^\gamma - p) = a(y - y^\gamma, p^\gamma - p) \\
&= -\|y - y^\gamma\|^2 + \langle\lambda, y^\gamma\rangle + \int_\Omega (\bar{\lambda} + \gamma y^\gamma)^+(y - y^\gamma) \\
&= -\|y - y^\gamma\|^2 + \langle\lambda, y^\gamma\rangle - \gamma\|(y^\gamma)^+\|^2_{\Omega_\gamma^+(\bar{\lambda})} + \int_{\Omega_\gamma^+(\bar{\lambda})} \bar{\lambda} y^\gamma,
\end{aligned} \tag{2.8}$$

where $\Omega_\gamma^+(\bar{\lambda}) := \{\bar{\lambda} + \gamma y^\gamma > 0\}$. Next we observe

$$-y^\gamma < \frac{\bar{\lambda}}{\gamma} \quad \text{a.e. in } \Omega_\gamma^+(\bar{\lambda}). \tag{2.9}$$

Due to $\bar{\lambda} \geq 0$ a.e. in Ω, we further note that

$$\Omega_\gamma^+(\bar{\lambda}) \supseteq \Omega_\gamma^+(0) = \{y^\gamma > 0\}. \tag{2.10}$$

Using (2.9) and (2.10) we continue (2.8) and obtain

$$\alpha\|u-u^\gamma\|_U^2+\|y-y^\gamma\|^2+\gamma\|(y^\gamma)^+\|^2\le\frac{1}{\gamma}\|\bar\lambda\|^2+\langle\lambda,y^\gamma\rangle \tag{2.11}$$

Next recall that $y^\gamma\in\mathcal{C}_0(\bar\Omega)$. Hence, we also have $(y^\gamma)^+\in\mathcal{C}_0(\bar\Omega)$. Together with $\lambda\in\mathcal{M}(\bar\Omega)$ and $\lambda\ge 0$ we further estimate (2.11):

$$\alpha\|u-u^\gamma\|_U^2+\|y-y^\gamma\|^2+\gamma\|(y^\gamma)^+\|^2\le\frac{1}{\gamma}\|\bar\lambda\|^2+\langle\lambda,(y^\gamma)^+\rangle \tag{2.12}$$

From this we obtain

$$\|(y^\gamma)^+\|\le\sqrt{\frac{2}{\gamma}}\max\left(\frac{\|\bar\lambda\|^2}{\gamma},|\langle\lambda,(y^\gamma)^+\rangle|\right)^{1/2} \tag{2.13}$$

which proves the assertion. □

We end this section by studying sufficient conditions for

$$\|(y^\gamma)^+\|=\mathcal{O}(\gamma^{-1})\text{ as }\gamma\to\infty.$$

Theorem 2.2. *Let u denote the solution of (1.3) and u^γ the solution of (1.9). If $y_0\ge 0$ a.e. in Ω, then $\|(y^\gamma)^+\|=\mathcal{O}(\gamma^{-1})$ as $\gamma\to\infty$.*

Proof. The optimal solution $(u^\gamma,y^\gamma=\mathcal{G}(Bu^\gamma))$ together with the corresponding adjoint state satisfy

$$\begin{aligned} a(v,p^\gamma) &= \int_\Omega(y^\gamma-y_0)v+\int_\Omega(\bar\lambda+\gamma y^\gamma)^+v\quad\forall v\in H^1(\Omega), \qquad (2.14)\\ a(y^\gamma,w) &= \int_\Omega Bu^\gamma w\quad\forall w\in H^1(\Omega). \end{aligned}$$

Using $v=y^\gamma$ in the first and $w=p^\gamma$ in the second equation, respectively, together with (1.11) and subtracting yield

$$0=\|y^\gamma\|^2-\int_\Omega y_0y^\gamma+\int_{\Omega_+^\gamma(\bar\lambda)}\bar\lambda y^\gamma+\gamma\int_{\Omega_+^\gamma(\bar\lambda)}(y^\gamma)^2+\alpha\|u^\gamma\|_U^2 \tag{2.15}$$

Now, our assumption $y_0\ge 0$ yields

$$\gamma\|(y^\gamma)^+\|^2\le\int_\Omega y_0y^\gamma+\gamma^{-1}\int_{\Omega_+^\gamma(\bar\lambda)\setminus\Omega_+(0)}\bar\lambda^2\le\int_\Omega y_0(y^\gamma)^++\gamma^{-1}\int_{\Omega_+^\gamma(\bar\lambda)\setminus\Omega_+(0)}\bar\lambda^2.$$

From this we conclude

$$\|(y^\gamma)^+\|=\mathcal{O}(\gamma^{-1})\text{ as }\gamma\to\infty.$$

□

In view of the proof of Theorem 2.2 we also have the following result. By w^- we denote the negative part of w, i.e., $w=w^+-w^-$.

Theorem 2.3. *Let u denote the solution of* (1.3) *and u^γ the solution of* (1.9). *If there exists $\epsilon > 0$ such that*

$$-\int_\Omega y_0^- y - \|y\|^2 - \alpha\|u\|^2 \leq -\epsilon,$$

then $\|(y^\gamma)^+\| = \mathcal{O}(\gamma^{-1})$ as $\gamma \to \infty$.

Proof. Similar as in [9] we get $u^\gamma \to u$ strongly in $L^2(\Omega)$ and $y^\gamma \to y$ strongly in $H^1(\Omega)$. Hence, for sufficiently large γ, we conclude

$$-\int_\Omega y_0^- y^\gamma - \|y^\gamma\|^2 - \alpha\|u^\gamma\|^2 \leq -\frac{\epsilon}{2} < 0.$$

Then (2.15) yields

$$\gamma\|(y^\gamma)^+\|^2 \leq \int_\Omega y_0(y^\gamma)^+ + \gamma^{-1}\int_{\Omega_+^\gamma(\bar\lambda)\setminus\Omega_+(0)} \bar\lambda^2.$$

From this we conclude

$$\|(y^\gamma)^+\| = \mathcal{O}(\gamma^{-1}) \text{ as } \gamma \to \infty.$$

□

3 Finite element discretization error analysis for (1.9)

Let $\mathcal{T}_h$ be a triangulation of Ω with maximum mesh size $h := \max_{T\in\mathcal{T}_h} \operatorname{diam}(T)$ and vertices $x_1, \dots, x_m$. We suppose that $\bar\Omega$ is the union of the elements of $\mathcal{T}_h$ so that element edges lying on the boundary are curved. In addition, we assume that the triangulation is quasi-uniform in the sense that there exists a constant $\kappa > 0$ (independent of h) such that each $T \in \mathcal{T}_h$ is contained in a ball of radius $\kappa^{-1}h$ and contains a ball of radius κh. Let us define the space of linear finite elements,

$$X_h := \{v_h \in C_0(\bar\Omega) \,|\, v_h \text{ is a linear polynomial on each } T \in \mathcal{T}_h\}$$

with the appropriate modification for boundary elements. We use the following approximation and inverse properties:

(a) According to [5, Theorem, Corollary] combined with [3, Theorem (4.4.20)] the L^2-projection $\Pi_h : W^{1,p}(\Omega) \to X_h$ ($d < p \leq \infty$) satisfies

$$\|v - \Pi_h v\|_{L^\infty} \leq Ch^{1-\frac{d}{p}}\|v\|_{W^{1,p}} \tag{3.16}$$

and is stable in L^2.

(b) Due to [3, Lemma (4.5.3)], for $v_h \in X_h$ we have

$$\|v_h\|_{L^\infty} \leq Ch^{-\frac{d}{2}}\|v_h\|. \tag{3.17}$$

In what follows it is convenient to introduce a discrete approximation of the operator $\mathcal{G}$. In fact, for a given function $v \in L^2(\Omega)$ we denote by $z_h = \mathcal{G}_h(v) \in X_h$ the solution of the discrete Neumann problem

$$a(z_h, v_h) \;=\; \int_\Omega v v_h \quad \text{for all } v_h \in X_h.$$

It is well-known [16] that for all $v \in L^2(\Omega)$

$$\|\mathcal{G}(v) - \mathcal{G}_h(v)\| \leq Ch^2\|v\|, \tag{3.18}$$

$$\|\mathcal{G}(v) - \mathcal{G}_h(v)\|_{L^\infty} \leq Ch^{2-\frac{d}{2}}\|v\|. \tag{3.19}$$

The estimate (3.19) can be improved provided one strengthens the assumption on v.

3.1 Estimate for $\|y^\gamma - y\|_{H^1} + \|u^\gamma - u\|_U$

We recall that by (2.12)

$$\alpha\|u - u^\gamma\|_U^2 + \|y - y^\gamma\|^2 + \gamma\|(y^\gamma)^+\|^2 \leq \frac{1}{\gamma}\|\bar{\lambda}\|^2 + \langle\lambda, (y^\gamma)^+\rangle.$$

Since the right-hand-side of this estimate is uniformly bounded in γ, we obtain $\|(y^\gamma)^+\| = \mathcal{O}(\gamma^{-\frac{1}{2}})$ as $\gamma \to \infty$. The decay of $\|y^\gamma - y\|^2 + \alpha\|u^\gamma - u\|_U^2$ with respect to γ can be estimated by that of $\|(y^\gamma)^+\|_{\mathcal{C}_0(\bar{\Omega})}$. To the best of the authors knowledge, an estimate of the latter term has not yet been obtained. As a partial result in this direction we have the following lemma which bridges the gap between the averaged L^2-type penalization of the constraint violation and the required uniform bound for $(y^\gamma)^+$.

Lemma 3.1. *Let $0 < h \leq 1$. Then, for some $p > d$ we have*

$$\|(y^\gamma)^+\|_{\mathcal{C}_0(\bar{\Omega})} \leq C\left(h^{1-\frac{d}{p}} + \gamma^{-\frac{1}{2}}h^{-\frac{d}{2}}\right), \tag{3.20}$$

where the positive constant is independent of γ and h.

Proof. We use (3.16), (3.17) and (2.2) to obtain

$$\begin{aligned}
\|(y^\gamma)^+\|_{\mathcal{C}_0(\bar{\Omega})} &\leq \|(y^\gamma)^+ - \Pi_h(y^\gamma)^+\|_{\mathcal{C}_0(\bar{\Omega})} + \|\Pi_h(y^\gamma)^+\|_{\mathcal{C}_0(\bar{\Omega})} \leq \\
&\leq C\left\{h^{1-\frac{d}{p}}\|(y^\gamma)^+\|_{1,p} + h^{-\frac{d}{2}}\|\Pi_h(y^\gamma)^+\|\right\} \leq C\left\{h^{1-\frac{d}{p}}\|u^\gamma\|_U + h^{-\frac{d}{2}}\|(y^\gamma)^+\|\right\} \leq \\
&\leq C\left(h^{1-\frac{d}{p}} + \gamma^{-\frac{1}{2}}h^{-\frac{d}{2}}\right).
\end{aligned}$$

This concludes the proof. $\square$

From

$$c_1\|y - y^\gamma\|_{H^1}^2 \leq a(y - y^\gamma, y - y^\gamma) = \int_\Omega B(u - u^\gamma)(y - y^\gamma) \leq C\|u - u^\gamma\|_U\|y^\gamma - y\|_{H^1}$$

we immediately infer

$$\|y^\gamma - y\|_{H^1} + \|u^\gamma - u\|_U \leq C\|u^\gamma - u\|_U \leq \frac{C}{\sqrt{\alpha}}\left(h^{1-\frac{d}{p}} + \gamma^{-\frac{1}{2}}h^{-\frac{d}{2}}\right)^{\frac{1}{2}} + \frac{C}{\sqrt{\alpha\gamma}}\|\bar{\lambda}\|. \tag{3.21}$$

Under additional regularity assumptions, the bound in Lemma 3.1 also holds for $p = \infty$, i.e.,

$$\|(y^\gamma)^+\|_{\mathcal{C}_0(\bar{\Omega})} \leq C\left(h + \gamma^{-\frac{1}{2}}h^{-\frac{d}{2}}\right).$$

Remark 3.2. If the assumptions of Theorem 2.2 or Theorem 2.3 hold true, then the estimate (3.21) improves:

$$\|y^\gamma - y\|_{H^1} + \|u^\gamma - u\|_U \leq C\|u^\gamma - u\|_U \leq \frac{C}{\sqrt{\alpha}}\left(h^{1-\frac{d}{p}} + \gamma^{-1}h^{-\frac{d}{2}}\right)^{\frac{1}{2}} + \frac{C}{\sqrt{\alpha\gamma}}\|\bar{\lambda}\|. \tag{3.22}$$

3.2 Estimate for $\|y^\gamma - y_h^\gamma\|_{H^1} + \|u^\gamma - u_h^\gamma\|_U$

Problem (1.9) is now approximated by the following sequence of control problems depending on the mesh parameter h:

$$\min_{u \in U} J_h(u) := \frac{1}{2}\int_\Omega |\mathcal{G}_h(Bu) - y_0|^2 + \frac{\alpha}{2}\|u\|_U^2 + \frac{1}{2}\int_\Omega |(\bar\lambda + \gamma\,\mathcal{G}_h(Bu))^+|^2. \tag{3.23}$$

Problem (3.23) represents a convex infinite-dimensional optimization problem of a structure similar to problem (1.9). It admits a unique solution u_h^γ with corresponding state $y_h^\gamma \in X_h$. Furthermore, there exist a unique function $p_h^\gamma \in X_h$ satisfying

$$a(v_h, p_h^\gamma) = \int_\Omega (y_h^\gamma - y_0 + (\bar\lambda + \gamma\, y_h^\gamma)^+) v_h \text{ for all } v_h \in X_h, \text{ and} \tag{3.24}$$

$$\alpha u_h^\gamma + B^* p_h^\gamma = 0 \text{ in } U. \tag{3.25}$$

Let us first prove an error estimate which is optimal with respect to the approximation order of the finite element space, but it depends on the relaxation parameter γ.

Theorem 3.3. *Let u^γ denote the solution of* (1.9) *with $y^\gamma = \mathcal{G}(Bu^\gamma)$, and u_h^γ the solution of* (3.23)*. Then there exists $h_0 > 0$ and a constant independent of γ and h such that*

$$\|u^\gamma - u_h^\gamma\|_U \le \frac{C}{\alpha}\gamma h^2 \textit{ for all } 0 < h \le h_0. \tag{3.26}$$

Proof. Let $y^h, p^h, p^{hh} \in X_h$ denote the solutions to

$$a(y^h, v_h) = (Bu^\gamma, v_h) \text{ for all } v_h \in X_h,$$

$$a(v_h, p^h) = \int_\Omega (y^\gamma - y_0 + (\bar\lambda + \gamma\, y^\gamma)^+ v_h \text{ for all } v_h \in X_h,$$

$$a(v_h, p^{hh}) = \int_\Omega (y^h - y_0 + (\bar\lambda + \gamma\, y^h)^+ v_h \text{ for all } v_h \in X_h.$$

Subtracting (1.11) from (3.25) yields

$$\begin{aligned}
\alpha\|u^\gamma - u_h^\gamma\|_U^2 &= \int_\Omega (p_h^\gamma - p^\gamma) B(u^\gamma - u_h^\gamma) = \\
&= \int_\Omega (p^\gamma - p^h) B(u_h^\gamma - u^\gamma) + \int_\Omega (p^h - p^{hh}) B(u_h^\gamma - u^\gamma) + \int_\Omega (p^{hh} - p_h^\gamma) B(u_h^\gamma - u^\gamma) = \\
&= (1) + (2) + (3).
\end{aligned}$$

Since $p^\gamma \in H^2(\Omega)$ it is straightforward to show

$$(1) \le C\|p^\gamma - p^h\|\|u^\gamma - u_h^\gamma\|_U \le Ch^2|p^\gamma|_2\|u^\gamma - u_h^\gamma\|_U.$$

Furthermore,

$$\begin{aligned}
(2) = a(y_h^\gamma - y^h, p^h - p^{hh}) &= \int_\Omega (y^\gamma - y^h)(y_h^\gamma - y^h) + ((\bar\lambda + \gamma\, y^\gamma)^+ - (\bar\lambda + \gamma\, y^h)^+)(y_h^\gamma - y^h) \le \\
&\le C(1+\gamma)\|y^\gamma - y^h\|\|u^\gamma - u_h^\gamma\|_U \le Ch^2(1+\gamma)|y^\gamma|_2\|u^\gamma - u_h^\gamma\|_U,
\end{aligned}$$

and

$$(3) = a(y_h^\gamma - y^h, p^{hh} - p_h^\gamma) = \int_\Omega (y^h - y_h^\gamma)(y_h^\gamma - y^h) + \overbrace{\left((\bar\lambda + \gamma\, y^h)^+ - (\bar\lambda + \gamma\, y_h^\gamma)^+\right)(y_h^\gamma - y^h)}^{\leq 0 \text{ by Lemma } A.1} \leq 0,$$

such that

$$\|u^\gamma - u_h^\gamma\|_U \leq \frac{C}{\alpha} h^2 \{(1+\gamma)|y^\gamma|_2 + |p^\gamma|_2\}.$$

Using the bounds (2.5) and (2.4) on $|p^\gamma|_2$ and $|y^\gamma|_2$, respectively, proves the assertion. $\square$

Since

$$\|y^\gamma - y_h^\gamma\|_{H^1} \leq \|y^\gamma - y^h\|_{H^1} + \|y^h - y_h^\gamma\|_{H^1} \leq C\left(h\|u_\gamma\|_U + \|u^\gamma - u_h^\gamma\|_U\right)$$

by the Lipschitz continuity of $\|y^h - y_h^\gamma\|_{H^1}$ with respect to u, we obtain with the help of (3.26)

$$\|y^\gamma - y_h^\gamma\|_{H^1} + \|u^\gamma - u_h^\gamma\|_U \leq C\left(h + \frac{\gamma}{\alpha}h^2\right). \tag{3.27}$$

Next we prove an error estimate in h which is independent of γ. For this purpose we first prove a uniform (in γ) L^1-bound for $(\bar\lambda + \gamma y^\gamma)^+$ and $(\bar\lambda + \gamma y_h^\gamma)^+$.

Lemma 3.4. *Let Assumption 1.4 be satisfied, and let u^γ, u_h^γ denote the unique solutions to* (1.9) *and* (3.23)*, respectively, with associated states y^γ, y_h^γ. Then*

$$\|\max((\bar\lambda + \gamma y^\gamma)^+, (\bar\lambda + \gamma y_h^\gamma)^+)\|_{L^1(\Omega)} \leq C$$

with some positive constant C independent of γ and of h.

Proof. Using $\hat u$ of Assumption 1.4 as a test function in (1.11) yields

$$\alpha(u^\gamma, \hat u) + (B^*\mathcal{G}^*(\mathcal{G}(Bu^\gamma) - y_0), \hat u) = 0,$$

where the action of the operator $\mathcal{G}^*$ is defined in (2.14). Thus,

$$\int_\Omega \tau(\bar\lambda + \gamma y^\gamma)^+ dx \leq -\int_\Omega (\bar\lambda + \gamma y^\gamma)^+ \mathcal{G}(B\hat u) dx = \alpha(u^\gamma, \hat u) + \int_\Omega (y^\gamma - y_0)\mathcal{G}(B\hat u) dx \leq C$$

independent of γ due to the continuity of $B, \mathcal{G}$, and the uniform bounds on $\|u^\gamma\|_U$ with respect to γ. This proves the claim for $\|(\bar\lambda + \gamma y^\gamma)^+\|_{L^1(\Omega)}$. Since $\|\mathcal{G}_h(B\hat u) - \mathcal{G}(B\hat u)\|_\infty \to 0$ for $h \to 0$ by (3.19) and the solutions u_h^γ to (3.23) are uniformly bounded in U with respect to γ and h, similar arguments give the desired estimate also for $\|(\bar\lambda + \gamma y_h^\gamma)^+\|_{L^1(\Omega)}$ with a possibly smaller constant $0 < \tilde\tau \leq \tau$. $\square$

We are now prepared to prove a γ-independent error estimate.

Theorem 3.5. *Let u^γ denote the solution of* (1.9) *with $y^\gamma = \mathcal{G}(Bu^\gamma)$, and u_h^γ the solution to* (3.23) *with $y_h^\gamma = \mathcal{G}_h(Bu_h^\gamma)$. Then there exist $h_0 \in (0,1]$ and a constant independent of γ and h such that*

$$\|u^\gamma - u_h^\gamma\|_U + \|y^\gamma - y_h^\gamma\|_{H^1} \leq Ch^{1-\frac{d}{4}} \text{ for all } 0 < h \leq h_0. \tag{3.28}$$

Proof. Let $y^h, p^h \in X_h$ denote the finite element approximations defined at the beginning of the proof of Theorem 3.3. Next we multiply the difference of (1.11) and (3.25) by $u^\gamma - u_h^\gamma$. This gives

$$\alpha\|u^\gamma - u_h^\gamma\|_U^2 = \int_\Omega (p^\gamma - p_h^\gamma)B(u_h^\gamma - u^\gamma)dx =$$

$$= \int_\Omega (p^\gamma - p_h)B(u_h^\gamma - u^\gamma)dx + \int_\Omega (p^h - p_h^\gamma)B(u_h^\gamma - u^\gamma)dx =: (1) + (2).$$

We proceed by estimating

$$(1) \le \frac{\alpha}{2}\|u_h^\gamma - u^\gamma\|_U^2 + \frac{C}{\alpha}\|p^\gamma - p^h\|^2 \le \frac{\alpha}{2}\|u_h^\gamma - u^\gamma\|_U^2 + \frac{C}{\alpha}h^{4-d}\left(\|y^\gamma - y_0\|^2 + \|(\bar\lambda + \gamma y^\gamma)^+\|_{L^1}^2\right),$$

where we have used [1] to estimate the finite element error $\|p^\gamma - p^h\|$. Further, using the definition of the auxilliary functions y^h, p^h and the optimality conditions we get

$$(2) = a(y_h^\gamma - y^h, p^h - p_h^\gamma) = \int_\Omega (y^\gamma - y_h^\gamma)(y_h^\gamma - y^h) + ((\bar\lambda + \gamma y^\gamma)^+ - (\bar\lambda + \gamma y_h^\gamma)^+)(y_h^\gamma - y^h)dx =$$

$$= -\|y^\gamma - y_h^\gamma\|^2 + \int_\Omega (y^\gamma - y_h^\gamma)(y^\gamma - y^h) + ((\bar\lambda + \gamma y^\gamma)^+ - (\bar\lambda + \gamma y_h^\gamma)^+)(y_h^\gamma - y^h)dx \le$$

$$\le -\frac{1}{2}\|y^\gamma - y_h^\gamma\|^2 + \frac{1}{2}\|y^\gamma - y^h\|^2 + \int_\Omega ((\bar\lambda + \gamma y^\gamma)^+ - (\bar\lambda + \gamma y_h^\gamma)^+)(y_h^\gamma - y^h)dx =$$

$$= -\frac{1}{2}\|y^\gamma - y_h^\gamma\|^2 + \frac{1}{2}\|y^\gamma - y^h\|^2 + \underbrace{\int_\Omega ((\bar\lambda + \gamma y^\gamma)^+ - (\bar\lambda + \gamma y_h^\gamma)^+)(y_h^\gamma - y^\gamma)dx}_{\le 0 \text{ by Lemma } A.1} +$$

$$+ \int_\Omega (\bar\lambda + \gamma y^\gamma)^+(y^\gamma - y^h)dx + \int_\Omega (\bar\lambda + \gamma y_h^\gamma)^+(y^h - y^\gamma)dx \le$$

$$\le -\frac{1}{2}\|y^\gamma - y_h^\gamma\|^2 + \frac{1}{2}\|y^\gamma - y^h\|^2 + \max(\|(\bar\lambda + \gamma y^\gamma)^+\|_{L^1}, \|(\bar\lambda + \gamma y_h^\gamma)^+\|_{L^1})\|y^\gamma - y^h\|_\infty.$$

Combining (1) and (2) we obtain with the help of (3.18) and (3.19)

$$\alpha\|u^\gamma - u_h^\gamma\|_U^2 + \frac{1}{2}\|y^\gamma - y_h^\gamma\|^2 \le$$

$$\le C\left(\frac{1}{\alpha}h^{4-d}\left(\|y^\gamma - y_0\|^2 + \|(\bar\lambda + \gamma y^\gamma)^+\|_{L^1}^2\right) + h^2\|u^\gamma\|_U^2 + \right.$$

$$\left. + h^{2-\frac{d}{2}}\max(\|(\bar\lambda + \gamma y^\gamma)^+\|_{L^1}, \|(\bar\lambda + \gamma y_h^\gamma)^+\|_{L^1})\right).$$

Using this estimate and again

$$\|y^\gamma - y_h^\gamma\|_{H^1} \le \|y^\gamma - y^h\|_{H^1} + \|y^h - y_h^\gamma\|_{H^1} \le C\left\{h\|u_\gamma\|_U + \|u^\gamma - u_h^\gamma\|_U\right\},$$

we finally get the desired result, since $h_0 \le 1$. $\square$

Let us recall that by [7, Lemma 2.1] for $v \in W^{1,s}(\Omega)$, with $1 < s < \frac{d-1}{d}$, we have

$$\|\mathcal{G}(v) - \mathcal{G}_h(v)\|_\infty \le Ch^{3-\frac{d}{s}}\|v\|_{W^{1,s}}. \tag{3.29}$$

Now let us assume that Bu^γ is uniformly bounded in $W^{1,s}(\Omega)$ for some $1 < s < \frac{d-1}{d}$. Then we deduce the following result from the proof of the previous theorem.

Corollary 3.6. *Let the assumptions of Theorem 3.5 hold true. Then*

$$\|u^\gamma - u_h^\gamma\|_U + \|y^\gamma - y_h^\gamma\|_{H^1} \le Ch^{\frac{3}{2}-\frac{d}{2s}}\sqrt{|\log h|} \text{ for all } 0 < h \le h_0. \tag{3.30}$$

4 Numerical verification and parameter selection

We end this paper by illustrating our theoretical findings by numerical results. In our examples below we use homogeneous Dirichlet (rather then Neumann) boundary conditions. As mentioned earlier, our theory covers this case as well. In all test runs, in order to compactify the computations, we used a mass lumping technique which preserves the approximation order. Throughout we have $\Omega = (0,1)^2$, $U = L^2(\Omega)$, $B = \text{id}$, and $\bar\lambda \equiv 0$.

Example 1. Our first example is taken from [9]. The data are as follows: $y_0(x_1, x_2) = 10(\sin(2\pi x_1) + x_2)$, $b \equiv 0.01$, and $\alpha = 0.1$. We first solve the problem on a very fine mesh in order to generate a reference solution which we then restrict onto the respective (coarser) mesh for computing the relevant error quantity.

Figure 1 depicts the convergence behavior of $\|(y_h^\gamma)^+\|$ and of u_h^γ for various mesh sizes and γ-values, respectively. Clearly, we have $\|(y_h^\gamma)^+\| = \mathcal{O}(\gamma^{-1})$. In fact, it turns out that the assumption of Theorem 2.3 is satsified. Hence, the estimate of Remark 3.2 and its analogue for $p = \infty$ are applicable. This gives

$$\|u - u_h^\gamma\|_U^2 \le C(h + \gamma^{-1}h^{-1} + h). \tag{4.31}$$

Hence, $\gamma = h^{-2}$ is optimal as it produces an overall error of the order h. Considering the bound in the right hand side of (4.31) as a function of γ, we obtain

$$\|u - u_h^\gamma\|_U = \mathcal{O}(\gamma^{-1/4}).$$

Our numerical results depicted in the right plot of Figure 1 indicate an even better rate of convergence at the order of $\mathcal{O}(\gamma^{-1/2})$ (compare with the slope of the solid line). In fact, in our numerical tests we found that $\|(y_h^\gamma)^+\|_{L^\infty} = \mathcal{O}(\gamma^{-1})$ which explains the improved rate for $(u_h^\gamma)_\gamma$. This shows that there are situations where the estimate of Lemma 3.1 is too pessimistic. The convergence results of Figure 1 further show that the L^2-error in u_h^γ levels off as γ increases. Also, the corresponding γ-threshold depends on the mesh size of discretization and is related to the point where the error due to relaxation equals the discretization error. Moreover, a reduction in h results in an increase of the threshold, as predicted by our theory.

Example 2. Our second example is constructed such that an explicit solution is available and such that the multiplier can be decomposed analytically into a regular part and a singular part concentrated on the boundary of the active set $\mathcal{A} := \{x \in \Omega : y(x) = 0\}$ at the optimal solution. For later use we also define the corresponding inactive set $\mathcal{I} := \Omega \setminus \mathcal{A}$. Moreover, this example violates the constraint qualification (1.4). Hence, existence of λ and p cannot be argued by [4]. Rather one has to use weaker constraint qualifications. Such a theoretical

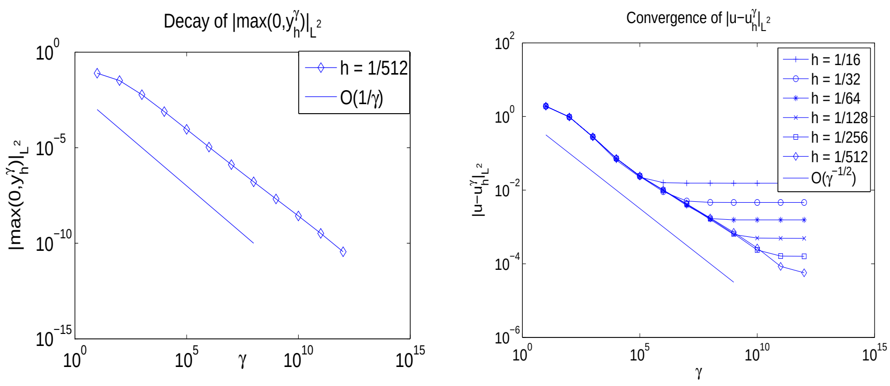

Figure 1: Example 1. Decay of $\|(y_h^\gamma)^+\|$ (left) and the convergence of $\|u - u_h^\gamma\|_U$ (right).

investigation, however, is beyond the scope of the present work. Here we study the numerical effect of such a situation, only.

For the construction of the optimal state we define

$$\hat{y}(x_1, x_2) = \sin(3\pi x_1)\sin(x_2)\sin(x_2 - 1)\sin(1 + 2\pi x_1 + 2\pi x_2^2).$$

Then the optimal state is given by

$$y(x_1, x_2) := -\max(0, \hat{y}(x_1, x_2))^4,$$

and the corresponding optimal control is $u := -\Delta y$. The associated adjoint state is $p := -\alpha u$ with α = 1e-3. Further, we introduce $\hat{\lambda} := \Delta p - y$. Note that on $\partial\mathcal{A}$ p admits only a generalized second derivative containing an element of $\partial(\hat{y})^+$, where we have used ∂ for representing the subdifferential of a convex function. Let $g^{\mathcal{A}}$ denote an arbitrarily fixed element of the generalized derivative at $\partial\mathcal{A}$ involved in Δp. Then one can show that $g^{\mathcal{A}}$ is nonnegative. This allows to decompose $\hat{\lambda}$ into a regular part $\hat{\lambda}_r = \hat{\lambda}|_{\operatorname{int}\mathcal{A}} + \hat{\lambda}|_{\mathcal{I}}$ and a singular part concentrated on $\partial\mathcal{A}$, i.e., $\lambda_s = g^{\mathcal{A}}$. For $\lambda|_{\mathcal{I}} := \hat{\lambda}|_{\mathcal{I}}$ it can be shown that it vanishes as a consequence of the complementarity system (1.7). For the determination of $\lambda|_{\operatorname{int}\mathcal{A}}$ we first define the desired state y_0 by

$$y_0|_{\operatorname{int}\mathcal{A}} = -2\min(0, \hat{\lambda}|_{\operatorname{int}\mathcal{A}}) + f|_{\operatorname{int}\mathcal{A}},$$

where $f(x_1, x_2) := 0.001 \cdot (2 + 7.5(x_1 + x_2))^4$, and

$$y_0|_{\mathcal{I}} = y|_{\mathcal{I}} - (\Delta p)|_{\mathcal{I}}. \tag{4.32}$$

As the right hand side in (4.32) is continuous, we define $y_0|_{\partial\mathcal{A}}$ by continuous extension. Then, $\lambda|_{\operatorname{int}\mathcal{A}}$ is given by

$$\lambda|_{\operatorname{int}\mathcal{A}} = (\Delta p)|_{\operatorname{int}\mathcal{A}} + z|_{\operatorname{int}\mathcal{A}} - y|_{\operatorname{int}\mathcal{A}}$$

and

$$\lambda|_{\partial\mathcal{A}} = g^{\mathcal{A}}.$$

In Figure 2 we show the optimal state (left), the optimal control (middle) and the regular part of the associated Lagrange multiplier (right) for $h = 1/128$. Combing the plots of y and λ we find that the active set structure is rather involved.

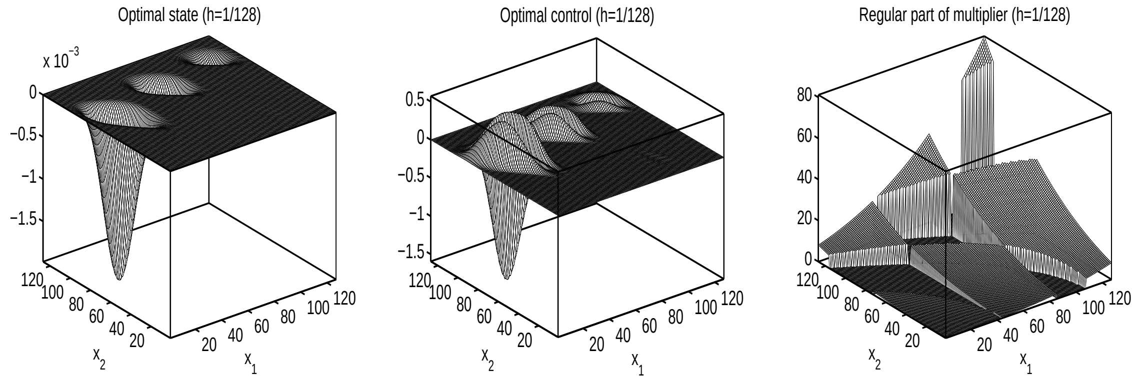

Figure 2: Example 2. Optimal state (left), optimal control (middle) and the regular part of the associated multiplier (right).

Figure 3 depicts the convergence behavior of $\|(y_h^\gamma)^+\|$ and of u_h^γ for various mesh sizes and γ-values, respectively. As for example 1, we have $\|(y_h^\gamma)^+\| = \mathcal{O}(\gamma^{-1})$. In fact, again it turns out that the assumption of Theorem 2.3 is satsified and (4.31) is available yielding $\|u - u_h^\gamma\|_U = \mathcal{O}(\gamma^{-1/2})$. All other conclusions are similar to the ones for example 1. Hence,

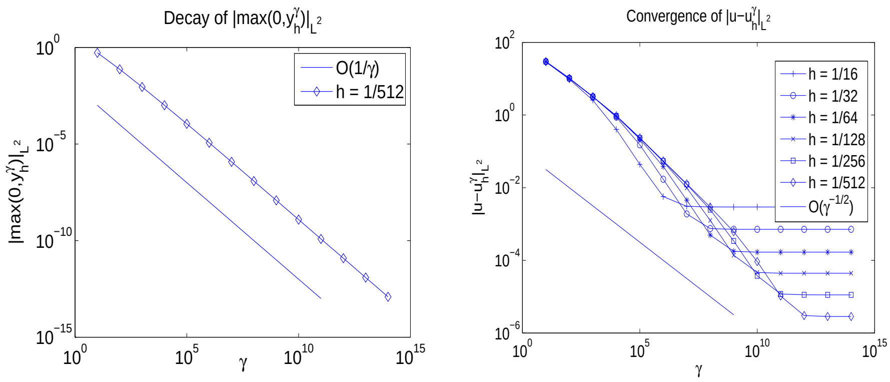

Figure 3: Example 2. Decay of $\|(y_h^\gamma)^+\|$ (left) and the convergence of $\|u - u_h^\gamma\|_U$ (right).

the lack of a Slater point (i.e., failure of (1.4) to hold true) did not cause any numerical instabilities in our test runs.

Higher order rate for $\mathbf{y_h^\gamma}$. From the results depicted in Figure 4 we observe that the convergence of $\|y - y_h^\gamma\|_{L^2}$ is (almost) of the order $\mathcal{O}(\gamma^{-1})$. This aspect is not covered by the theory of Sections 2 and 3. Rather we offer the following explanation: In [9] it was shown that under the assumption that the set

$$S_0^\gamma := \{x \in \Omega : \bar{\lambda}(x) + \gamma y(x) = 0\}$$

has measure zero, y^γ and u^γ (and hence p^γ) are strongly differentiable (in $H^2(\Omega) \cap H_0^1(\Omega)$ and $L^2(\Omega)$) with respect to γ. Let $\dot{y}^\gamma$ and $\dot{u}^\gamma$ denote the corresponding derivatives. We point

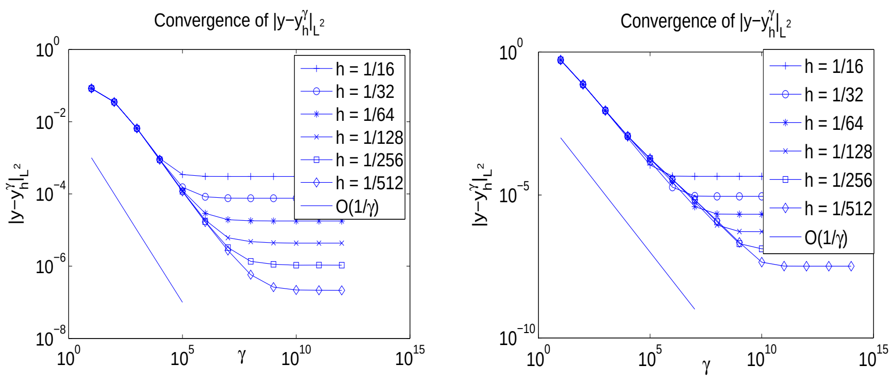

Figure 4: Convergence of $\|y - y_h^\gamma\|_{L^2}$ for example 1 (left) and example 2 (right).

out that in the sequel we only argue for homogeneous Dirichlet boundary conditions as this reflects our numerical example. Other boundary conditions (such as the Neumann condition discussed earlier) can be covered with little modification. In [9] it was further shown that the above derivatives satisfy a system of sensitivity equations (see Proposition 2.4 and Corollary 2.1 in [9]):

$$\begin{aligned} -\Delta \dot{y}^\gamma - \dot{u}^\gamma &= 0, && (4.33)\\ -\Delta \dot{p}^\gamma - \chi_{S^\gamma}(y^\gamma + \gamma \dot{y}^\gamma) - \dot{y}^\gamma &= 0, && (4.34)\\ \alpha \dot{u}^\gamma + \dot{p}^\gamma &= 0, && (4.35) \end{aligned}$$

where $S^\gamma := \{x \in \Omega : \bar{\lambda}(x) + \gamma y^\gamma(x) > 0\}$ and χ_S denotes the characteristic function of a set $S \subset \Omega$. Replacing $\dot{p}^\gamma$ by $-\alpha \dot{u}^\gamma$, reducing the remaining two equations to one and testing with $\dot{y}^\gamma \in H^2(\Omega) \cap H_0^1(\Omega)$ results in

$$\alpha \|\Delta \dot{y}^\gamma\|^2 + \|\dot{y}^\gamma\|^2 + \gamma \|\dot{y}^\gamma\|_{S^\gamma}^2 = (\chi_{S^\gamma} y^\gamma, \dot{y}^\gamma). \tag{4.36}$$

Assuming $\|(y^\gamma)^+\| = \mathcal{O}(\kappa(\gamma))$ with a continuous function κ satisfying $\kappa(\gamma) \downarrow 0$ as $\gamma \to \infty$, from (4.36) we immediately derive

$$\begin{aligned} \|\dot{y}^\gamma\| &\le \|(y^\gamma)^+\| = \mathcal{O}(\kappa(\gamma)), && (4.37)\\ \|\Delta \dot{y}^\gamma\|^2 &\le \alpha^{-1} \|\dot{y}^\gamma\| \, \|(y^\gamma)^+\| = \mathcal{O}(\kappa(\gamma)^2), && (4.38)\\ \|\dot{y}^\gamma\|_{S^\gamma} &\le \gamma^{-1} \|(y^\gamma)^+\| = \mathcal{O}(\gamma^{-1}\kappa(\gamma)). && (4.39) \end{aligned}$$

Note that we (at least) have $\kappa(\gamma) = \gamma^{-1/2}$. In our numerical example we rather observe $\kappa(\gamma) = \gamma^{-1}$ (see also Theorem 2.2 and 2.3 for theoretical investigations). Now, assuming that, when increasing γ, y^γ changes most in the region where it violates the pointwise inequality constraint $y \le 0$, we invoke the assumption

$$\|\dot{y}^\gamma\|_{S_c^\gamma} \le C \|\dot{y}^\gamma\|_{S^\gamma} \tag{4.40}$$

for some positive constant C independent of γ. Above S_c^γ denotes the complement of S^γ in Ω. We point out that we could weaken (4.40) by allowing some dependence of C on γ as long as $C\kappa(\gamma) = \mathcal{O}(\gamma^{-t})$ for some $t > 0$. But for the sake of simplicity, we keep the γ-independent formulation. Then, still assuming that $\text{meas}(S_0^\gamma) = 0$ for all $\gamma > 0$, we obtain

$$\|y^\gamma - y\| \leq \lim_{\tau \to \infty} \int_\gamma^\tau \|\dot{y}^s\| ds \leq \hat{C} \lim_{\tau \to \infty} \int_\gamma^\tau s^{-1} \kappa(\gamma) ds, \tag{4.41}$$

with some positive constant $\hat{C}$ depending on C. In the case where $\kappa(\gamma) = \gamma^{-1/2}$, we obtain

$$\|y^\gamma - y\| = \mathcal{O}(\gamma^{-1/2}), \tag{4.42}$$

and

$$\|y^\gamma - y\| = \mathcal{O}(\gamma^{-1}), \tag{4.43}$$

for $\kappa(\gamma) = \gamma^{-1}$. The latter case appears to be reflected in Figure 4.

Conclusions

In this paper we develop error estimates for the Moreau-Yosida regularization of state constrained optimal control problems. A critical tool is an estimate for the pointwise violation of the constraint due to the involved relaxation in $L^2(\Omega)$. Based on this result and an estimate for the finite element discretization error an estimate for the overall error is obtained which allows an optimal regularization parameter adjustment. The theoretical findings are verified numerically. Further, based on our numerical results, we also argue estimates for the state in the L^2-norm. Our second numerical example shows that even in cases where the Slater condition fails to hold, our error estimates seem to remain true. A thorough analytical investigation of this situation under a weaker constraint qualification remains subject to future work.

Acknowledgements

The first author gratefully acknowledges support by the FWF under START-Program Y305 "Interfaces and Free Boundaries". The second author gratefully acknowledges the support of the DFG Priority Program 1253 entitled Optimization With Partial Differential Equations.

A Projection

Lemma A.1. *Let $(X, (\cdot, \cdot)$ denote a real Hilbert space, $S \subseteq X$ a convex and closed subset, and $P : X \to S$ the orthogonal projection in X onto S. Then*

$$\|P(s) - P(t)\|^2 \leq (s - t, P(s) - P(t)) \text{ for all } s, t \in X.$$

Proof. $P : X \to S$ is the orthogonal projection in X onto S iff

$$(a - P(a), b - P(a)) \leq 0 \text{ for all } b \in S.$$

Use now $a = s, b = P(t)$ and $a = t, b = P(s)$, and add the resulting inequalities. This proves the assertion. □

We have the following immediate consequence of Lemma A.1.

Corollary A.2. *In Lemma A.1, let $X := L^2(\Omega)$ and $S := \{y \geq 0 \text{ a.e. in } \Omega\}$. Then for $y \in L^2(\Omega)$*

$$P(y)(x) = y^+(x) := \max(0, y(x)) \text{ a.e. in } \Omega.$$

References

[1] Casas, E.: *L^2 estimates for the finite element method for the Dirichlet problem with singular data*, Numer. Math. **47**, 627–632 (1985).

[2] Bernardi, C.: *Optimal finite element interpolation on curved domains*, SIAM J. Numer. Anal. **26**, 1212–1240 (1989).

[3] Brenner, S., Scott, R.: *The mathematical theory of finite elements*, Springer New York, second editon (2002).

[4] Casas, E.: *Control of an elliptic problem with pointwise state constraints*, SIAM J. Cont. Optim. **24**, 1309–1322 (1986).

[5] Douglas, J., Dupont, T., Wahlbin, L.: *The stability in L^q of the L^2-projection into finite element function spaces*, Numer. Math. **23**, 193-197 (1975).

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

[7] Deckelnick, K., Hinze, M.: *A finite element approximation to elliptic control problems in the presence of control and state constraints*, Preprint HBAM2007-01, Hamburger Beiträge zur Angewandten Mathematik, Universität Hamburg (2007).

[8] Hintermüller, M., Kunisch, K.: *Path following methods for a class of constrained minimization methods in function spaces*, SIAM J. Optim. **17**, 159-187 (2006).

[9] Hintermüller, M., Kunisch, K.: *Feasible and non-interior path following in constrained minimization with low multiplier regularity*, SIAM J. Control Optim. **45**, 1198-1221 (2006).

[10] Hintermüller, M.: *Mesh-independence and fast local convergence of a primal-dual active-set method for mixed control-state constrained elliptic problems*, ANZIAM Journal **49** (1), 1-38 (2007).

[11] Hinze, M.: *A variational discretization concept in control constrained optimization: the linear-quadratic case*, J. Computational Optimization and Applications **30**, 45-63 (2005).

[12] Hinze, M., Meyer, C.: *Numerical analysis of Lavrentiev-regularized state constrained elliptic control problems*, WIAS Preprint 1168, Berlin (2006).

[13] Morrey Jr., C. B.: *Multiple Integrals in the Calculus of Variations*, Springer, New York, 1966.

[14] Tröltzsch, F.: *Regular Lagrange multipliers for control problems with mixed pointwise control-state constraints*, SIAM J. Optim. **15**, 616–634 (2005).

[15] Meyer, C., Prüfert, U., Tröltzsch, F.: *On two numerical methods for state-constrained elliptic control problems*, Technical Report 5-2005, Institut für Mathematik, TU Berlin (2005).

[16] Scott, R.: *Optimal L^∞ estimates for the finite element method on irregular meshes*, Math. Comp. **30**, 681–897 (1976).

[17] Scott, L.R., Zhang, S.: *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp. **54**, 483–493 (1990).