START-Projekt Y 305-N18

Karl–Franzens–Universität Graz

On the Length of the Primal-Dual Path in Moreau-Yosida-based Path-following for State Constrained Optimal Control: Analysis and Numerics

Michael Hintermüller Anton Schiela

IFB-Report No. 50 September 2011

A–8010 GRAZ, MOZARTGASSE 14/II, AUSTRIA

Supported by the
Austrian Science Fund (FWF)
Der Wissenschaftsfonds.

On the Length of the Primal-Dual Path in Moreau-Yosida-based Path-following for State Constrained Optimal Control: Analysis and Numerics *

Anton Schiela & Michael Hintermüller

Abstract

We derive a-priori estimates on the length of the primal-dual path that results from a Moreau-Yosida approximation of the feasible set for state constrained optimal control problems. These bounds depend on the regularity of the state and the dimension of the problem. Comparison with numerical results indicates that these bounds are sharp and are attained for the case of a single active point.
AMS MSC 2000: 49M30, 49J52

Keywords: state constraints, PDE constrained optimization, path-following

1 Introduction

In recent years, path-following methods based on the Moreau-Yosida (or quadratic penalty) regularization of state constrained problems have received considerable attention. While general results on the convergence of this method can be derived under very mild assumptions, deriving estimates on the length of corresponding homotopy path, the "primal-dual path", and its asymptotic behavior is more delicate. In particular, numerical experience shows that this asymptotic behavior varies from problem to problem.

The purpose of this note is twofold. First, we present a-priori error estimates on the order of convergence of the primal-dual path that depend on the dimensionality of the problem and the smoothness of the solution. In comparison to the estimates that were derived in [9] we obtain an improvement in the rate, compared to [3] our results are based on a considerably weaker set of assumptions.

Second, we try to develop an understanding on the principles that govern the rate of convergence of the primal dual path. This will be accomplished by comparison of numerical and theoretical results. It will turn out that the topology of the active set plays a decisive role for the rate of convergence. For the "worst case",

*Supported by the DFG Research Center MATHEON "Mathematics for key technologies"

namely the case that the active set is a single touch point, our theoretical estimates coincide with the numerical observations.

To render the discussion concrete, we consider the primal-dual path-following method for a state constrained model problem in optimal control. The techniques presented here are, however, applicable in a much broader context. The main idea is to replace the problem:

$$\begin{aligned} \min_{y\in H^2(\Omega), u\in L^2(\Omega)} \frac{1}{2}\|y-y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} \\ \text{subject to} \quad -\Delta y - u = 0 \text{ in } \Omega \\ y = 0 \text{ in } \partial\Omega \\ y \le \psi \text{ in } \Omega \end{aligned} \tag{1}$$

(here Ω is a smoothly bounded domain in $\mathbb{R}^d$ for $d = 1, 2, 3$, $y_d \in L_2(\Omega)$, and ψ is a smooth, strictly positive function on $\overline{\Omega}$), by a family of problems

$$\begin{aligned} \min_{y\in H^2(\Omega), u\in L^2(\Omega)} \frac{1}{2}\|y-y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\|\max(y-\psi,0)\|^2_{L^2(\Omega)} \\ \text{subject to} \quad -\Delta y - u = 0 \text{ in } \Omega \\ y = 0 \text{ in } \partial\Omega \end{aligned} \tag{2}$$

and consider a sequence of solutions $x_\gamma := (y_\gamma, u_\gamma)$ of (2). It has been shown in [4] that this sequence converges to the original solution $x_* := (y_*, u_*)$ of (1) as γ tends to infinity.

Practical algorithms use a semi-smooth Newton method to solve discretizations of the subproblems (2) approximately or exactly. For this purpose the first order necessary conditions are derived for (2) which assert existence of an adjoint state $p_\gamma \in H^2(\Omega)$ such that

$$\begin{aligned} y_\gamma - y_d + \gamma\max(y_\gamma - \psi, 0) - \Delta p_\gamma = 0 \text{ in } \Omega \\ p_\gamma = 0 \text{ in } \partial\Omega \end{aligned} \tag{3}$$

$$\alpha u_\gamma - p_\gamma = 0 \text{ in } \Omega \tag{4}$$

$$\begin{aligned} -\Delta y_\gamma - u_\gamma = 0 \text{ in } \Omega \\ y_\gamma = 0 \text{ in } \partial\Omega. \end{aligned} \tag{5}$$

This can be compared with the first order necessary conditions for the original problem, which state existence of a measure valued Lagrangian multiplier $m \in \mathcal{M}(\overline{\Omega})$ and an adjoint state $p_* \in W^{1,q'}(\Omega)$ $(q' < d/(d-1))$, such that

$$\begin{aligned} y_* - y_d + m - \Delta p_* = 0 \text{ in } \Omega \\ p_* = 0 \text{ in } \partial\Omega \\ -\Delta y_* - u_* = 0 \text{ in } \Omega \\ y_* = 0 \text{ in } \partial\Omega \\ \alpha u_* - p_* = 0 \text{ in } \Omega \\ m \ge 0, \quad y_* \le \psi, \quad \langle m, y_* - \psi\rangle_{\mathcal{M}(\overline{\Omega})\times C(\overline{\Omega})} = 0 \text{ in } \overline{\Omega}. \end{aligned}$$

We observe that the function $\gamma \max(y_\gamma - \psi, 0)$ plays the role of m in the regularized setting.

Elimination of u_γ from (3)–(5) yields the system

$$F(x;\gamma) := \begin{cases} y - y_d + \gamma \max(y-\psi,0) - \Delta p &= 0 \text{ in } \Omega \\ p &= 0 \text{ in } \partial\Omega \\ -\Delta y - \alpha^{-1} p &= 0 \text{ in } \Omega \\ y &= 0 \text{ in } \partial\Omega, \end{cases} \tag{6}$$

which can be tackled by a semi-smooth Newton method as shown in [4].

2 Analysis of the Length of the Primal-Dual Path

In constrained optimization and in particular in state constrained optimal control (c.f. e.g. [1]) the existence of a strictly feasible point (a Slater point) is a standard assumption for the existence of Lagrange multipliers. In state constrained optimal control this assumption is used to show that the corresponding Lagrange multipliers are positive measures. We will assume existence of a Slater point throughout the paper:

Assumption 2.1. Assume that there is a constant $e > 0$, such that

$$\psi - \breve{y} > e \text{ on } \Omega$$

for some pair $(\breve{y}, \breve{u})$ that satisfies the state equation.

Existence of a strictly feasible point and smoothness of the state variable y will allow us to bound the length of the primal-dual path by a power of γ^{-1}, which is clearly a stronger result that mere convergence of the primal-dual path for $\gamma \to \infty$.

2.1 A-priori bounds for the Constraint Violation in L^1

In the following, denote by y_γ^+ the function $\max(y-\psi, 0)$. Our first aim is to show that $\gamma \left\| y_\gamma^+ \right\|_{L^1}$ is bounded uniformly for $\gamma \to \infty$. The following technique is well established by now, and used in various contexts (cf. e.g. [2, 8, 5]).

Lemma 2.2. *The expression $\gamma \left\| y_\gamma^+ \right\|_{L^1}$ is uniformly bounded for $\gamma \to \infty$.*

Proof. Let S be the solution operator of the PDE, i.e., the control to state mapping. We test (3) and (4) with a feasible direction (Sv, v) from the optimal control problem, and add them (taking into account that $\langle -\Delta p_\gamma, Sv\rangle = \langle p_\gamma, v\rangle$) to obtain

$$\langle u_\gamma, v\rangle + \langle y_\gamma - y_d, Sv\rangle + \gamma \langle y_\gamma^+, Sv\rangle = 0 \qquad \forall v \in L^2(\Omega).$$

Inserting $v := \breve{u}, S\breve{u} = \breve{y}$ we obtain

$$\alpha \langle u_\gamma, \breve{u}\rangle + \langle y_\gamma - y_d, \breve{y}\rangle - \gamma \langle y_\gamma^+, \breve{y}\rangle = 0.$$

Since $u_\gamma, y_\gamma, \breve{u}, \breve{y}, y_d$ are bounded in L^2 independent of γ we conclude

$$\gamma\langle y_\gamma^+, \breve{y}\rangle = \gamma \int_{\overline{\Omega}} \breve{y} y_\gamma^+ \, dt \leq c$$

and thus by non-negativity of y_γ^+ and positivity of $\breve{y}$:

$$\gamma \left\| y_\gamma^+ \right\|_{L^1} \leq \frac{\gamma\langle y_\gamma^+, \breve{y}\rangle}{\min\{\breve{y}\}} \leq \frac{c}{e}.$$

□

Let us discuss, which exponents s are to be expected for estimates of the form $\gamma^s \left\| y_\gamma^+ \right\|_{L^q} \leq c$ in generic situations. If $s \geq 1$, then we conclude that

$$\gamma \left\| y_\gamma^+ \right\|_{L^q} \leq c$$

hence, γy_γ^+ is either bounded in L^q or converges to 0 in L^q. If $q > 1$, one can conclude that a Lagrangian multiplier for the original problem exists in L^q:

Proposition 2.3. *If*

$$\left\| y_\gamma^+ \right\|_{L^q} \leq c\gamma^{-1}$$

for some $1 < q < \infty$, then there is a Lagrangian multiplier for the state constraints of problem (1) *which is contained in $L^q(\Omega)$.*

Proof. Since $\gamma \left\| y_\gamma^+ \right\|_{L^q} \leq c$, the function γy_γ^+ has a weak accumulation point $m \in L^q(\Omega)$, which is positive due to weak closedness of the positive cone in L^q. Moreover,

$$\langle m, y_* \rangle = \lim_{\gamma\to\infty} \langle \gamma y_\gamma^+, y_* \rangle = 0$$

by strong convergence of $y_\gamma \to y_* \leq \psi$. Then, by compactness, p_γ, the solution of (3), converges strongly in L^2 to some p_*, so that, since $u_\gamma \to u_*$ in $L^2(\Omega)$, also (4) is fulfilled in the limit. Hence, m is Lagrangian multiplier for the state constraints. □

Generically, however, one observes that these multipliers are only measures. So the case $s \geq 1$ for $p > 1$ will only appear in cases of exceptionally regular Lagrangian multipliers (or completely inactive state constraints), and thus, we generically expect $s < 1$ for $p > 1$. Consequently, Lemma 2.2 is a result as good, as we can expect.

2.2 Estimates, depending on the Constraint Violation in L^∞

We approach our aim via the value functional

$$V(\gamma) := J^\gamma(x_\gamma) = J(x_\gamma) + \frac{\gamma}{2} \left\| y_\gamma^+ \right\|^2.$$

It was shown in [4] that $\lim_{\gamma\to\infty} V(\gamma) = J(x_*)$. Here we show that the rate of convergence depends on $\|y_\gamma^+\|_\infty$.

Theorem 2.4. *We have the following estimate for the derivative of the value functional:*

$$0 \leq \frac{d}{d\gamma} V(\gamma) \leq \frac{c}{\gamma} \left\| y_\gamma^+ \right\|_{L^\infty}. \tag{7}$$

If $\left\| y_\gamma^+ \right\|_{L^\infty} \leq c\gamma^{-s}$*, then also*

$$J(x_*) - V(\gamma) \leq c\gamma^{-s} \tag{8}$$

and

$$\sqrt{\alpha} \left\| u_* - u_\gamma \right\|_U \leq c\gamma^{-s/2}. \tag{9}$$

Proof. In [4, Proposition 4.1] differentiability of V was shown and the expression

$$\frac{d}{d\gamma} V(\gamma) = \frac{1}{2} \left\| y_\gamma^+ \right\|_{L^2}^2$$

was derived. Since V is monotonically increasing, and by the estimate $\|v\|_{L^2}^2 \leq \|v\|_{L^1} \|v\|_{L^\infty}$, it follows

$$0 \leq \frac{d}{d\gamma} V(\gamma) \leq \left\| y_\gamma^+ \right\|_{L^1} \left\| y_\gamma^+ \right\|_{L^\infty}$$

By Lemma 2.2 we know that

$$\left\| y_\gamma^+ \right\|_{L^1} \leq \frac{c}{\gamma},$$

and (7) follows. If $\left\| y_\gamma^+ \right\|_{L^\infty} \leq c\gamma^{-s}$, then

$$0 \leq \frac{d}{d\gamma} V(\gamma) \leq c\gamma^{-1-s}$$

Let now $\gamma_1 > \gamma_2$ be given. Then, by the fundamental theorem of calculus (cf. e.g. [7, Thm. 25.18])

$$V(\gamma_1) - V(\gamma_2) = \int_{\gamma_2}^{\gamma_1} \frac{d}{d\gamma} V(\gamma)\, d\gamma \leq \int_{\gamma_2}^{\gamma_1} c\gamma^{-1-s}\, d\gamma = c(\gamma_2^{-s} - \gamma_1^{-s}).$$

Hence, since $\lim_{\gamma\to\infty} V(\gamma) = J(x_*)$,

$$J(x_*) - V(\gamma_2) = \lim_{\gamma_1\to\infty} V(\gamma_1) - V(\gamma_2) \leq c\gamma_2^{-s},$$

which yields (8). Finally, (9) follows straightforwardly from the uniform convexity of J^γ with respect to u and (8):

$$\begin{aligned}
\frac{\alpha}{2} \|u_* - u_\gamma\|_{L^2}^2 &\leq J^\gamma(x_*) + J^\gamma(x^\gamma) - 2J^\gamma\left(\frac{1}{2}x_* + \frac{1}{2}x^\gamma\right) \\
&\leq J^\gamma(x_*) + J^\gamma(x^\gamma) - 2J^\gamma(x^\gamma) = J^\gamma(x_*) - J^\gamma(x^\gamma) \\
&= J(x_*) - V(\gamma) \leq c\gamma^{-s}.
\end{aligned}$$

□

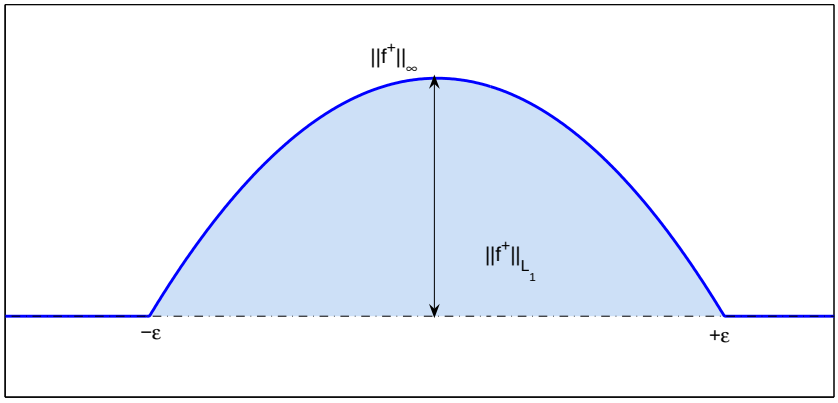

Figure 1: Geometrical idea of our considerations.

2.3 A Worst Case Estimate for the Constraint Violation in L^∞

The bottonline of the previous section is that we have to find an estimate for $\|y_\gamma^+\|_\infty$ as sharp as possible. This can be achieved by exploiting the smoothness of y.

The essence of our technique is a geometric idea, which we will explain for illustration at a simple example (cf. Figure 1). Let $f(t) = a(-t^2 + \varepsilon^2)$ be a concave parabola, and $f^+(t) := \max(f(t), 0)$ its positive part. Then $f(t) \geq 0$ for $t \in [-\varepsilon, +\varepsilon]$ with a maximum $\|f^+\|_\infty = f(0) = a\varepsilon^2$, and $\|f^+\|_{L^1} = \int_{-\varepsilon}^{\varepsilon} f(t)\,dx = a4/3\varepsilon^3$. Thus, we have

$$\left\|f^+\right\|_\infty = a\varepsilon^2 = a^{1/3}(3/4)^{2/3}(a4/3\varepsilon^3)^{2/3} \leq c\,\|f\|_{C^2}^{1/3}\left\|f^+\right\|_{L^1}^{2/3}.$$

Hence, from the boundedness of the second derivatives (by $2a$) one can conclude a relation between the L^1 norm and the L^∞ norm of a function with zero boundary values. The following proposition generalizes this observation.

To this end, we introduce following notation: let $m \in \mathbb{N}_0$. For $m < \beta \leq m+1$ let $C^\beta(\overline{\Omega})$ the subspace of $C^m(\overline{\Omega})$ of those functions which have Hölder continuous derivatives of order m. These spaces are equipped with the usual norms:

$$\|v\|_{C^\beta} := \|v\|_{C^m} + \|v^{(m)}\|_{C^{\beta-m}}.$$

In this setting we can show the following interpolation type estimate:

Proposition 2.5. *Let $\Omega \subset \mathbb{R}^d$ be bounded and open, $0 \leq y \in C^\beta(\overline{\Omega}), y \in L^q(\Omega)$, $1 \leq q < \infty$, $0 < \beta \leq 2$. Moreover, assume that $y = 0$ on $\partial\Omega$. Then*

$$\|y\|_{L^\infty} \leq c\,\|y\|_{C^\beta}^{1-\Theta}\,\|y\|_{L^q}^{\Theta} \tag{10}$$

with $\Theta = \frac{\beta}{\beta + q^{-1}d}$. The constant is independent of Ω.

Proof. Assume without loss of generality that $0 \in \Omega$ and $y(0) = \|y\|_{L^\infty}$. Denote by $B_r(0)$ the ball of radius r and center 0.

Since $y \in C^1$ for $\beta > 1$, and y necessarily obtains a maximum at 0, we conclude $\nabla y(0) = 0$ for $\beta \geq 1$. By the Hölder-continuity of y (and possibly of ∇y for $\beta > 1$) we infer that $y(t) > y(0) - \|y\|_{C^\beta} r^\beta$ for all $t \in B_r(0)$. In particular, $y(t)$ is positive for $t \in B(0, R)$ with

$$R = \left(\frac{y(0)}{\|y\|_{C^\beta}}\right)^{1/\beta} = \left(\frac{\|y\|_{L^\infty}}{\|y\|_{C^\beta}}\right)^{1/\beta}.$$

By the assumption $y = 0$ on $\partial\Omega$ it follows that $B_R(0) \subset \overline{\Omega}$. Hence, we can compute

$$\begin{aligned} \|y\|_{L^q}^q = \int_{\overline{\Omega}} |y(t)|^q \, dt &\geq c \int_{[0,R]} \left| y(0) - \|y\|_{C^\beta} r^\beta \right|^q r^{d-1} \, dr \qquad (11) \\ &= c \, \|y\|_{C^\beta}^q \int_{[0,R]} \left(\frac{y(0)}{\|y\|_{C^\beta}} - r^\beta \right)^q r^{d-1} \, dr \\ &= c \, \|y\|_{C^\beta}^q \int_{[0,R]} \left(R^\beta - r^\beta \right)^q r^{d-1} \, dr \\ &\geq c \, \|y\|_{C^\beta}^q \, R^{\beta q + d} = c \, \|y\|_{C^\beta}^{q - \frac{\beta q + d}{\beta}} \, \|y\|_{L^\infty}^{\frac{\beta q + d}{\beta}}. \end{aligned}$$

Solving this estimate for $\|y\|_{L^\infty}$, we conclude

$$\|y\|_{L^\infty} \leq c \, \|y\|_{L^q}^{\frac{\beta q}{\beta q + d}} \, \|y\|_{C^\beta}^{1 - \frac{\beta q}{\beta q + d}}.$$

□

Observe that (10) is only true for $\beta \leq 2$, which means that we can only use smoothness up to order 2. This corresponds to the fact that maximizers yield vanishing derivatives of first, but not of higher order.

A similar technique has been used in the context of interior point methods to show positive distance of the central path to the bounds [8]. Also, in [6, Lemma 4.7] similar techniques seem to be used, at least for the case $\beta \leq 1$ for the virtual control approach.

Corollary 2.6. *If y_γ is uniformly bounded in $C^\beta(\overline{\Omega})$ for $\gamma \to \infty$ and some $0 < \beta \leq 2$, we have the following estimate on the constraint violation:*

$$\|y_\gamma^+\|_\infty \leq c\gamma^{-s}, \text{ where } s = \frac{\beta}{\beta + d}. \qquad (12)$$

In particular, we have for every $\varepsilon > 0$:

$$\|y_\gamma^+\|_\infty \leq \begin{cases} c\gamma^{-2/3} & : \quad d = 1 \\ c\gamma^{-1/2+\varepsilon} & : \quad d = 2 \\ c\gamma^{-1/4+\varepsilon} & : \quad d = 3. \end{cases} \qquad (13)$$

Proof. The result (12) follows readily from Lemma 2.2 and Proposition 2.5.

To derive (13), we have to invoke standard regularity results for equations with right hand side measures. Let $q' = d/(d-1)$ for $d > 1$ and q' arbitrarily large otherwise. Then, since $\gamma \|y_\gamma +\|$ is uniformly bounded in $L^1(\Omega)$, it follows that u_γ is uniformly bounded in $W^{1,q'}(\Omega)$, implying that y_γ is uniformly bounded in $W^{3,q'}(\Omega)$. It follows by the Sobolev embedding theorems that y_γ is uniformly bounded in C^β for $\beta = 3-\varepsilon, 2-\varepsilon, 1-\varepsilon$ in the cases $d = 1, 2, 3$, respectively, and for every $\varepsilon > 0$. □

In our numerical experiments below, we will see that our technique yields sharp estimates, if the geometric situation (an elliptic paraboloid) modelled by the proof actually occurs. This is the case if the active set is a single point. In more regular situations, however, higher values of s are observed.

Remark 2.7. Under the assumption that y_γ is in C^2, uniformly, we get the heuristic bound $O(\gamma^{-2/5})$ for the constraint violation in the case $d = 3$.

Finally, we give a generic upper bound for s:

Proposition 2.8. *If $s \geq 1$, then problem (1) has a Lagrangian multiplier in $L^q(\Omega)$ for each $q < \infty$.*

Proof. This is a direct consequence of Proposition 2.3. □

2.4 The Length of the Primal-Dual Path

Combination of our estimates finally yields the following convergence estimate:

Theorem 2.9. *If y_γ is uniformly bounded in $C^\beta(\Omega)$ for $\gamma \to \infty$ and $0 < \beta \leq 2$, we have the following convergence estimate for the primal dual path:*

$$\sqrt{\alpha}\, \|u_* - u_\gamma\|_U \leq c\gamma^{-\frac{\beta}{2(\beta+d)}}. \tag{14}$$

In particular, we have for every $\varepsilon > 0$:

$$\sqrt{\alpha}\, \|u_* - u_\gamma\|_U \leq \begin{cases} c\gamma^{-1/3} & : \quad d = 1 \\ c\gamma^{-1/4+\varepsilon} & : \quad d = 2 \\ c\gamma^{-1/8+\varepsilon} & : \quad d = 3. \end{cases} \tag{15}$$

Proof. The result (14) follows from Theorem 2.4 and corollary 2.6. In the same way, (15) follows from (13). □

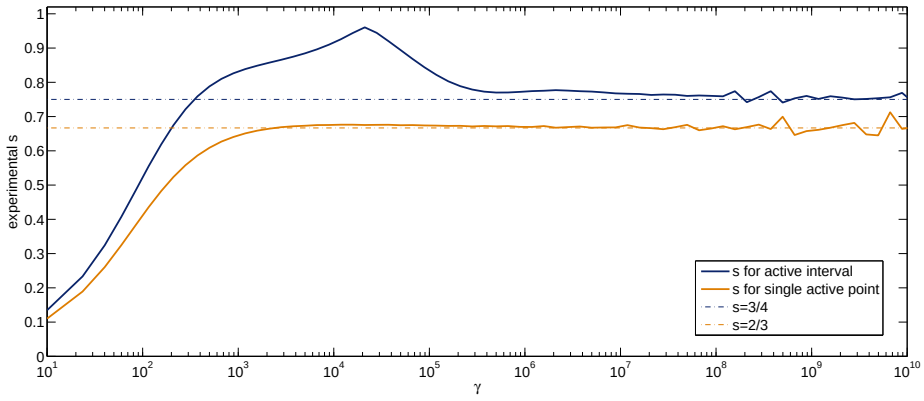

Figure 2: Experimental results on the exponent s of the relation $\left\|y_\gamma^+\right\|_{L^\infty} = O(\gamma^{-s})$ in a one-dimensional setting.

3 Comparison with Experimental Results

We conclude our paper with numerical experiments in 1d and 2d, which illustrate the close relationship between our theoretical estimates and the convergence in practice. However, they also indicate that, under additional geometric assumptions, the estimates may be refined.

To measure the exponent s we perform a numerical path-following method for the system (6), implemented in Matlab. We use a prescribed sequence of parameters γ_j. In our test problem we choose a positive, constant desired state $y_d = 10$, a regularization parameter $\alpha = 1$, and obtain different types of active sets by varying the spatially constant upper bounds ψ. Our discretization of y and p is done simply by classical finite differences, i.e, the 3-point and the 5-point star in the one-dimensional and the two-dimensional setting, respectively. We compute an estimate for s by the following formula:

$$s_j := \frac{\ln(\|y_{\gamma_j}^+\|_\infty) - \ln(\|y_{\gamma_{j+1}}^+\|_\infty)}{\ln \gamma_{j+1} - \ln \gamma_j}.$$

Observe that this formula is quite sensitive to perturbations of y, which explains in part the oscillations, seen in the plots.

Experimental results in 1d. In the one-dimensional case our computational domain is the unit interval, discretized uniformly by 10000 nodes. In our first setting (with $\psi \equiv 0.06$), the active set consists of a single point, and the optimal state y_* has a parabolic shape, i.e. a second derivative, which is bounded away from zero. Here, according to Figure 2, $s \approx 2/3$ behaves as predicted by Corollary 2.6.

In our second setting (with $\psi \equiv 0.01$), the active set is proper interval. Here we observe experimentally $s \approx 3/4$, which is a higher rate than predicted by by Corollary 2.6. However, this rate can be explained by the following heuristic model.

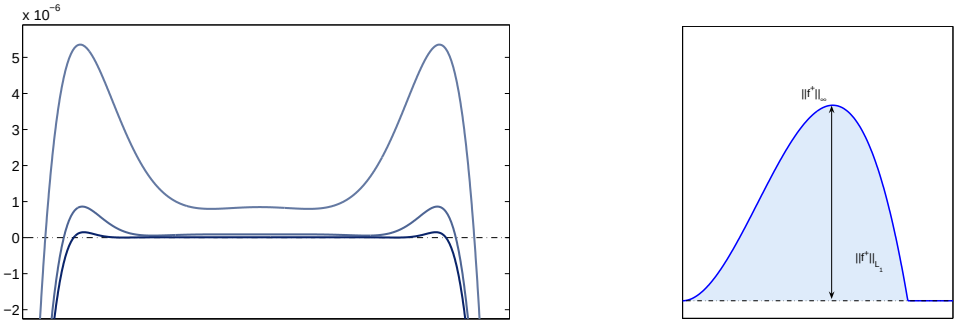

Figure 3: Geometric situation for y_γ near an active interval. Left: observed in numerical experiments for $\gamma = 10^7, 10^8$, and 10^9. Right: idealized model.

Since y_* is constant on the active interval $y_*^{(m)} = 0$ for all $m \in \mathbb{N}_0$ there. Consider on $[0, \infty[$ the functions $f(t) = a(-t^3 + \varepsilon t^2)$ and $f^+ = \max(f, 0)$. Obviously, $f^{(m)}(0) = 0$ for $m = 1, 2$. Clearly, f is positive on $[0, \varepsilon]$ and has a maximum at $t = 2/3\varepsilon$ with $f(2/3\varepsilon) = 10/27 a\varepsilon^3$, i.e. $\|f^+\|_\infty = ca\varepsilon^3$. Further, $\|f^+\|_{L^1} = \int_0^\varepsilon f(t)\,dt = a\varepsilon^4/12$. Hence, similarly as above, we may conclude that (using $6a = \|f\|_{C^3}$)

$$\left\|f^+\right\|_\infty \leq c \left\|f\right\|_{C^3}^{1/4} \left\|f^+\right\|_{L^1}^{3/4}.$$

These considerations cannot easily be carried over to y_γ^+ rigorously, since $y_\gamma' = 0$ need not hold on the active set of y_*, as assumed. However, as illustrated by Figure 3 the numerical experiments indicate that practically this situation occurs.

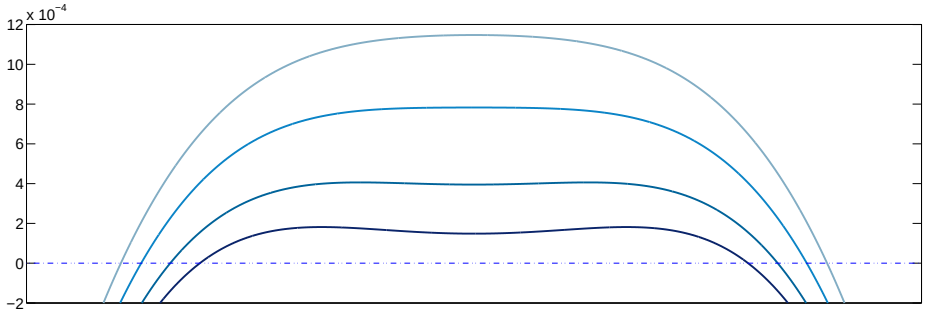

Figure 4: Transition from single maximum to double maximum of y_γ for $\gamma = 7.5 \cdot 10^{-5}, 5 \cdot 10^{-5}, 2.5 \cdot 10^{-5}$, and 10^{-5}.

Another interesting aspect is the high value of s in the range $\gamma \in [10^4, 10^5]$. A close look at the corresponding intermediate solutions shows that for these values y_γ has a very flat maximum, which explains this behavior. For larger values of γ this single maximum splits into two maxima, which can be nicely observed in Figure 4.

Experimental results in 2d. In the two dimensional case we use the unit square as computational domain and a uniform grid of 512×512 nodes. Compared to the

1d case, the geometric situation for $d = 2$ (and even more for $d = 3$) is much more complex, but still our analysis and numerical experiments suggest the conjecture that the rate of convergence of the primal-dual path will depend largely on the geometry of the active set, or equivalently on the structure of the Lagrangian multipliers.

If the active set is a single point (for $\psi \equiv 0.015$), we observe in Figure 5 that our estimates in Corollary 2.6 coincide with the numerical rate of convergence, namely $s \approx 0.5$. We attribute the oscillatory behavior of s to grid effects, because for large γ we have $y_\gamma^+ > 0$ only on a few nodes of the grid.

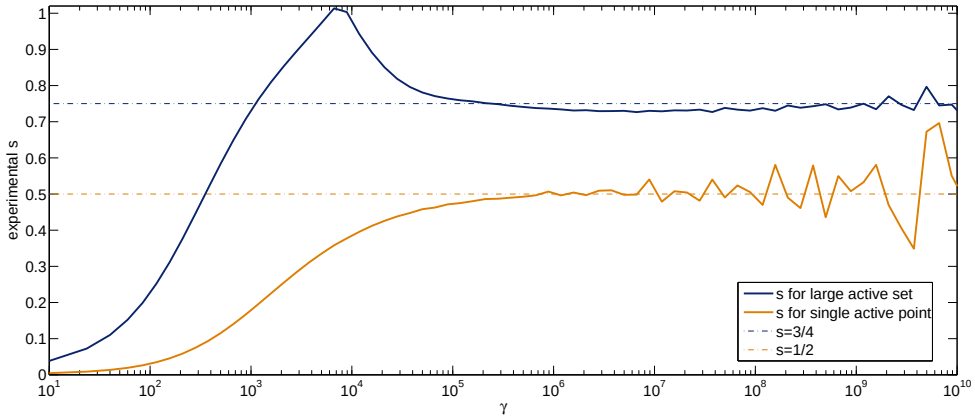

Figure 5: Experimental results on the exponent s of the relation $\left\|y_\gamma^+\right\|_{L^\infty} = O(\gamma^{-s})$ in a two-dimensional setting.

However, if the active set is a subset of Ω with non-empty interior (setting $\psi \equiv 0.001$), the radial integral estimate (11), including the factor r^{d-1} will lead to an over-estimation of $\|y_\gamma^+\|_\infty$. If the boundary of the active set is e.g. a submanifold of dimension $d_{\partial A} < d$, then one might rather expect a factor $r^{d-1-d_{\partial A}}$ in (11), since y_γ^+ then tends to be curved with respect to directions orthogonal to this submanifold. In this case one would expect results of the form

$$\|y_\gamma^+\|_\infty \leq c\gamma^{-s}, \text{ where } s = \frac{\beta}{\beta + (d - d_{\partial A})}. \tag{16}$$

In this light, and taking into account the insight gained from the one-dimensional case that close to a flat surface formula (16) can be used with $\beta = 3$, one might expect that in the case of an active set with non-empty interior (i.e. $d_{\partial A} = 1$)

$$s \approx \frac{3}{3 + (2 - 1)} = \frac{3}{4}.$$

Figure 5 verifies that this rate is actually observed in numerical practice. The "hump" in the pre-asymptotic phase is again explained by a very flat maximum of y_γ in this range of parameters.

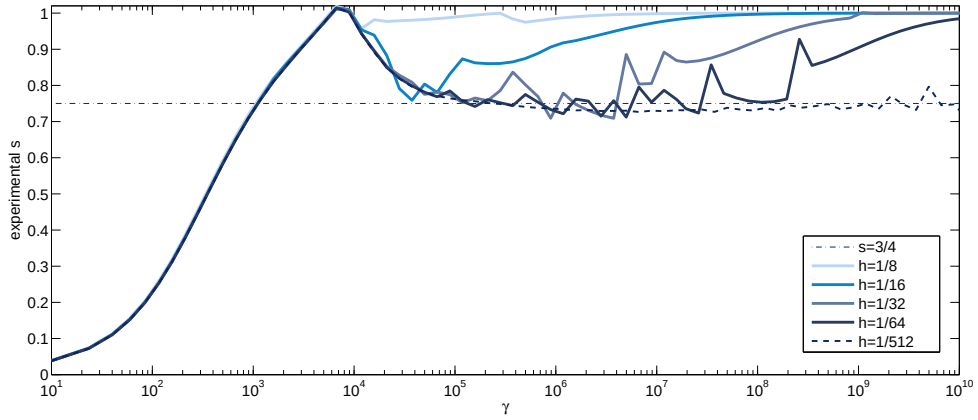

Figure 6: Experimental results on the exponent in a two-dimensional setting with large active set for various coarse grids. For comparison the asymptotics for a fine grid $h = 1/512$ is added as a dashed line.

Effect of a fixed discretization on the rate. If on a fixed grid y_γ is computed to very high values of γ it can be observed (see Figure 6) that a rate $s = 1$ finally occurs. This can simply be explained by the fact that in finite dimensional spaces all norms are equivalent, so that for uniformly bounded discrete $\gamma\|(y_\gamma^+)_h\|_{L^1}$ we eventually observe the upper bound

$$\|(y_\gamma^+)_h\|_{L^\infty(\Omega)} \le c(h)\|(y_\gamma^+)_h\|_{L^1(\Omega)} \le c(h)\gamma^{-1} \quad \Leftrightarrow \quad s = 1.$$

If such a behavior is observed in practice, it is a clear indication that the problem has been "oversolved" numerically.

4 Conclusion

Numerical experiments indicate that the asymptotics on the constraint violation derived in this note are sharp, if the active set is a single point, at least in one- and two-dimensional problems. Moreover, by heuristic arguments we were able to explain the higher rates of convergence, observed numerically in the case of larger active sets. It might be a topic of future research to render these considerations rigorous.

From a practical point of view, the observation that s generically varies in an interval $s \in [s_{\text{point}}, 1[$, so that the rate of convergence varies in $[s_{\text{point}}/2, 1/2[$, may help in the construction of adaptive algorithms, which try to balance algebraic errors and discretization errors.

References

[1] E. Casas. Control of an elliptic problem with pointwise state constraints. *SIAM J. Control Optim.*, 24(6):1309–1318, 1986.

[2] K. Deckelnick and M. Hinze. Convergence of a finite element approximation to a state-constrained elliptic control problem. *SIAM J. Numer. Anal.*, 45(5):1937–1953, 2007.

[3] M. Hintermüller and M. Hinze. A note on an optimal parameter adjustment in a Moreau-Yosida based approach to state constrained elliptic control problems. *SIAM J. Numer. Anal.*, 47:1666–1683, 2009.

[4] M. Hintermüller and K. Kunisch. Feasible and non-interior path-following in constrained minimization with low multiplier regularity. *SIAM J. Control Optim.*, 45(4):1198–1221, 2006.

[5] M. Hintermüller and K. Kunisch. PDE-constrained optimization subject to pointwise constraints on the control, the state, and its derivative. *SIAM J. Optim.*, 20(3):1133–1156, 2009.

[6] K. Krumbiegel. *Numerical concepts and error analysis for elliptic Neumann boundary control problems with pointwise state and control constraints.* PhD thesis, Universität Duisburg-Essen, 2009.

[7] E. Schechter. *Handbook of Analysis and its Foundations.* Academic Press, 1997.

[8] A. Schiela. Barrier methods for optimal control problems with state constraints. *SIAM J. Optim.*, 20(2):1002–1031, 2009.

[9] M. Ulbrich. *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces.* MPS-SIAM Series on Optimization. Cambridge University Press, 2011.