

THE LENGTH OF THE PRIMAL-DUAL PATH IN MOREAU–YOSIDA-BASED PATH-FOLLOWING METHODS FOR STATE CONSTRAINED OPTIMAL CONTROL*

MICHAEL HINTERMÜLLER[†], ANTON SCHIELA[‡], AND WINNIFRIED WOLLNER[§]

Abstract. A priori estimates of the length of the primal-dual path resulting from a Moreau–Yosida approximation of the feasible set for state constrained optimal control problems are derived. These bounds depend on the regularity of the state and the dimension of the problem. Numerical results indicate that the bounds are indeed sharp and are typically attained in cases where the active set consists of isolated active points. Further conditions on the multiplier approximation are identified which guarantee higher convergence rates for the feasibility violation due to the Moreau–Yosida approximation process. Numerical experiments show again that the results are sharp and accurately predict the convergence behavior.

Key words. Moreau–Yosida regularization, PDE constrained optimization, path-following, pointwise state constraints, regularization error

AMS subject classifications. 49M30, 49J52, 49J20

DOI. 10.1137/120866762

1. Introduction. In recent years, path-following methods based on the Moreau–Yosida regularization of state constrained problems have received considerable attention. While general results on the convergence of this method can be derived under fairly mild assumptions, deriving estimates on the length of the corresponding homotopy path, the “primal-dual path,” and its asymptotic behavior is more delicate. In particular, numerical experience shows that the asymptotic behavior varies from problem to problem. This has lead to the consideration of a posteriori error estimators for the regularization error; see, e.g., [13, 16].

The purpose of this paper is twofold. First, we present a priori error estimates on the order of convergence of the primal-dual path that depend on the dimension of the problem and the smoothness of the solution. In comparison to the estimates that were derived in [15] or [8], we obtain an improvement in the rate. In fact, we show that the convergence is dominated by the maximal feasibility violation, for which we derive rates. In view of this, in [15, Example 8.38] for H^2 -regular solutions in two spatial dimensions the convergence order $O(\gamma^{-15/52+\epsilon})$ is obtained, whereas in [8, section 3.1] $O(\gamma^{-1/4})$ is derived. Our estimates yield $O(\gamma^{-1/3+\epsilon})$. Here γ denotes the Moreau–Yosida regularization parameter as stated in problem (1.2). We also

*Received by the editors February 21, 2012; accepted for publication (in revised form) October 3, 2013; published electronically January 14, 2014.

<http://www.siam.org/journals/siopt/24-1/86676.html>

[†]Institut für Mathematik, Humboldt-Universität zu Berlin, 10099 Berlin, Germany (hint@math.hu-berlin.de). This author acknowledges support by the DFG-Research Center DFG Research Center MATHEON “Mathematics for Key Technologies,” the DFG-SPP 1253 “Optimization with PDEs,” the START-Project Y305 “Interfaces and Free Boundaries” administered by the Austrian Science Fund FWF and the FWF-SFB F32 “Mathematical Optimization and Its Application in Biomedical Sciences” at the Karl-Franzens University of Graz through subproject F32-04 “FREELEVEL.”

[‡]Institut für Mathematik, Technische Universität Hamburg-Harburg, Hamburg, Germany (anton.schiela@tuhh.de). This author’s research was supported by the DFG Research Center MATHEON “Mathematics for Key Technologies.”

[§]Department of Mathematics, University of Hamburg, 20146 Hamburg, Germany (winnifried.wollner@uni-hamburg.de).

note that the techniques of [8], when augmented by a bootstrapping argument, also have the potential to yield the same rate as ours; see section 2.5. However, we point out that our results are independent of any error estimates for the discretization. Finally, we note that in some cases, in particular for functions which are at most H^2 -regular, one may derive our worst case estimates also from [4, 11]. However, none of these approaches yield sharp results in case of the best possible solution regularity of $W^{3,q'}(\Omega)$ with $q' < d/(d-1)$. Moreover, none of these approaches is able to fully explain numerical observations by theoretical results.

Second, we develop an understanding of the principles that govern the rate of convergence of the primal-dual path. This will be accomplished by a comparison of numerical and theoretical results. In fact, it turns out that the topology of the active set plays a decisive role for the rate of convergence. In the “worst case” scenario, which is the case of the active set being a single touch point (or a set of isolated touch points), our theoretical worst case estimates coincide with numerical observations. In cases where the active set is a submanifold of the domain or even has nonempty interior, we provide a mathematical explanation for the numerically observed rates.

In order to render our discussion concrete, we consider the primal-dual path-following method for a state constrained model problem in optimal control. We emphasize, however, that the techniques presented here are applicable in a much broader context; see, e.g., [17] for constraints on the gradient of the state. Thus, for our specific discussion we consider throughout the model problem (1.1) below. Applying the Moreau–Yosida approximation to the indicator function of the set related to the pointwise inequality constraint in

$$(1.1) \quad \begin{aligned} &\text{minimize} \quad J(y, u) = \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega)}^2 \quad \text{over } y \in H^1(\Omega), u \in L^2(\Omega) \\ &\text{subject to} \quad -\Delta y - u = 0 \text{ in } \Omega, \quad y = 0 \text{ on } \partial\Omega, \\ &\quad \quad \quad y \leq \psi \text{ in } \Omega, \end{aligned}$$

results in the family of problems

$$(1.2) \quad \begin{aligned} &\text{minimize} \quad J^\gamma(y, u) = \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega)}^2 + \frac{\gamma}{2} \|\max(y - \psi, 0)\|_{L^2(\Omega)}^2 \\ &\quad \quad \quad \text{over } y \in H^1(\Omega), u \in L^2(\Omega) \\ &\text{subject to} \quad -\Delta y - u = 0 \text{ in } \Omega, y = 0 \text{ on } \partial\Omega \end{aligned}$$

for $\gamma > 0$. Above, we assume for simplicity that Ω is a smooth, 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 $\bar{\Omega}$ and $\alpha > 0$ some given constant. Moreover, we refer to y as the state and to u as the control variable, respectively.

By standard arguments it follows that both problems have a unique solution since J and J^γ are coercive, weakly lower semicontinuous, and strictly convex. Let $x_\gamma := (y_\gamma, u_\gamma)$ denote the unique solution of (1.2). It was shown in [9, Proposition 2.1] that the sequence (x_γ) converges strongly in $H_0^1(\Omega) \times L^2(\Omega)$ to the original solution $x_* := (y_*, u_*)$ of (1.1) as γ tends to infinity. Given the structure of our model problem, we again remark that the following analysis does not depend on the special form of the (partial differential) equation or the inequality constraint in (1.1), but rather on the fact that the feasibility violation, i.e., $\max(y_\gamma - \psi, 0)$ in our case, is not only in $L^2(\Omega)$ but, in fact, is in $C^\beta(\bar{\Omega})$ for some $0 < \beta \leq 2$.

Using slightly nonstandard notation we define the space $C^\beta(\overline{\Omega})$ as follows. Let $m \in \mathbb{N}_0$. For $m < \beta \leq m+1$ denote by $C^\beta(\overline{\Omega})$ the subspace of $C^m(\overline{\Omega})$ whose elements have m th derivatives that are Hölder continuous with exponent $\beta - m$. These spaces are equipped with the usual norms

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

Practical algorithms use a semismooth Newton method to solve discretizations of the problems (1.2) approximately or exactly; see, e.g., [9, 10]. For this purpose, the first order necessary conditions are derived for (1.2) which assert the existence of an adjoint state $p_\gamma \in H^2(\Omega)$ such that

$$(1.3) \quad y_\gamma - y_d + \gamma \max(y_\gamma - \psi, 0) - \Delta p_\gamma = 0 \text{ in } \Omega, \quad p_\gamma = 0 \text{ on } \partial\Omega,$$

$$(1.4) \quad \alpha u_\gamma - p_\gamma = 0 \text{ in } \Omega,$$

$$(1.5) \quad -\Delta y_\gamma - u_\gamma = 0 \text{ in } \Omega, \quad y_\gamma = 0 \text{ on } \partial\Omega.$$

The system (1.3)–(1.5) approximates the first order necessary (and in our case also sufficient) conditions for the original problem (1.1), which yield the existence of a measure-valued Lagrange multiplier $m_* \in \mathcal{M}(\overline{\Omega})$ and an adjoint state $p_* \in W^{1,q'}(\Omega)$ with $q' < d/(d-1)$ such that

$$\begin{aligned} y_* - y_d + m_* - \Delta p_* &= 0 \text{ in } \Omega, \quad p_* = 0 \text{ on } \partial\Omega, \\ -\Delta y_* - u_* &= 0 \text{ in } \Omega, \quad y_* = 0 \text{ on } \partial\Omega, \\ \alpha u_* - p_* &= 0 \text{ in } \Omega, \\ m_* &\geq 0, \quad y_* \leq \psi, \quad \langle m_*, y_* - \psi \rangle_{\mathcal{M}(\overline{\Omega}), C(\overline{\Omega})} = 0 \text{ in } \overline{\Omega}, \end{aligned}$$

where $\langle \cdot, \cdot \rangle_{\mathcal{M}(\overline{\Omega}), C(\overline{\Omega})}$ denotes the duality pairing between $C(\overline{\Omega})$ and its dual space $\mathcal{M}(\overline{\Omega})$, the regular Borel measures; see, e.g., [3, Theorem 2]. We observe that in (1.3) the function $\gamma \max(y_\gamma - \psi, 0)$ plays the role of m_* .

The elimination of u_γ from (1.3)–(1.5) yields the system

$$(1.6) \quad F(x; \gamma) := \begin{cases} y - y_d + \gamma \max(y - \psi, 0) - \Delta p = 0 \text{ in } \Omega, & p = 0 \text{ on } \partial\Omega, \\ -\Delta y - \alpha^{-1}p = 0 \text{ in } \Omega, & y = 0 \text{ on } \partial\Omega, \end{cases}$$

which can be addressed by a semismooth Newton method as shown in [9].

Note that the state equation (1.5) gives rise to a continuous solution operator $S: L^2(\Omega) \rightarrow H_0^1(\Omega)$. Under our regularity assumption, on the domain this operator is also continuous as a mapping from $L^2(\Omega)$ to $H_0^1(\Omega) \cap H^2(\Omega)$.

2. Analysis of the length of the primal-dual path. Our analysis proceeds in three main steps. First, we derive uniform L^1 -bounds on the constraint violation. In a second step we conclude that, given these L^1 -bounds, the length of the primal-dual path depends on L^∞ -bounds on the constraint violation. Then, finally, these latter bounds are derived by exploiting the Hölder continuity of the states.

2.1. A priori bounds for the constraint violation in L^1 . Abusing notation, in what follows we write y_γ^+ for $\max(y_\gamma - \psi, 0)$. Our first aim is to show that $\gamma \|y_\gamma^+\|_{L^1}$ is bounded uniformly as $\gamma \rightarrow \infty$. The following technique is well established by now and has been used in various contexts (see, e.g., [5, 10, 14]). Below, c denotes a generic constant which may take different values at different occasions.

From (1.2) one readily finds that u_γ, y_γ, y_d are bounded in $L^2(\Omega)$ independently of γ since

$$\|y_\gamma\|_{H^1}^2 \leq c\|u_\gamma\|_{L^2}^2 \leq cJ^\gamma(y_\gamma, u_\gamma) \leq J^\gamma(0, 0) = J(0, 0).$$

Moreover, it is also clear that $\gamma\|y_\gamma^+\|_{L^2}^2 \leq J(0, 0)$. Moreover, by [9, Proposition 2.1] it holds that $(y_\gamma, u_\gamma) \rightarrow (y_*, u_*)$ strongly in $H_0^1(\Omega) \times L^2(\Omega)$, and we conclude that $\gamma\|y_\gamma^+\|_{L^2}^2 \rightarrow 0$ because of

$$J(y_*, u_*) \leftarrow J(y_\gamma, u_\gamma) \leq J^\gamma(y_\gamma, u_\gamma) \leq J^\gamma(y_*, u_*) = J(y_*, u_*).$$

By the assumed properties of the solution operator S this also implies convergence of $y_\gamma \rightarrow y_*$ in $H^2(\Omega)$. The following result readily follows from [10], but it is proven here in order to keep the paper self-contained.

LEMMA 2.1. *The expression $\gamma\|y_\gamma^+\|_{L^1}$ is uniformly bounded for $\gamma \rightarrow \infty$.*

Proof. Let S be the solution operator, i.e., the control-to-state mapping, of the state equation which is the partial differential equation constraint in (1.2). We test (1.3) and (1.4) by a feasible direction $(Sv, v) \in H^1(\Omega) \times L^2(\Omega)$ and add the resulting two equations (taking into account that $\langle -\Delta p_\gamma, Sv \rangle = (p_\gamma, v)$, where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $H^1(\Omega)$ and its dual $H^1(\Omega)^*$ and $(\cdot, \cdot)_{L^2}$ is the usual inner product in $L^2(\Omega)$). This yields

$$\alpha(u_\gamma, v)_{L^2} + (y_\gamma - y_d, Sv)_{L^2} - \gamma(y_\gamma^+, Sv)_{L^2} = 0 \quad \forall v \in L^2(\Omega).$$

Inserting $v := u_\gamma, Su_\gamma = y_\gamma$ we obtain

$$\alpha(u_\gamma, u_\gamma)_{L^2} + (y_\gamma - y_d, y_\gamma)_{L^2} - \gamma(y_\gamma^+, y_\gamma)_{L^2} = 0.$$

From the uniform boundedness of y_γ and u_γ we infer that

$$\gamma(y_\gamma^+, y_\gamma)_{L^2} = \gamma \int_\Omega y_\gamma y_\gamma^+ d\omega \leq c$$

for some constant $c > 0$ independent of γ . Hence, by the nonnegativity of y_γ^+ and $0 < \underline{\psi} := \min_{\overline{\Omega}} \psi \leq \psi \leq y_\gamma$ on $\{y_\gamma^+ \neq 0\}$ we get

$$\gamma \underline{\psi} \|y_\gamma^+\|_{L^1} \leq \gamma \int_\Omega y_\gamma^+(x) \underline{\psi} dx \leq \gamma(y_\gamma^+, y_\gamma)_{L^2} \leq c.$$

This concludes the proof. \square

Given $1 \leq q \leq \infty$, one may be interested in whether there exists an exponent $s > 0$ which yields an estimate of the form $\gamma^s \|y_\gamma^+\|_{L^q} \leq c$ in certain generic situations. In order to shed some light on this, assume for the moment that such an estimate would hold true. If $s \geq 1$, then we find that

$$\gamma \|y_\gamma^+\|_{L^q} \leq \gamma^s \|y_\gamma^+\|_{L^q} \leq c \quad \text{for } \gamma \geq 1.$$

Hence, γy_γ^+ is at least bounded in $L^q(\Omega)$ and converges to 0 in $L^q(\Omega)$ if $s > 0$. In the case where we also have $q > 1$, one infers that the original problem admits a Lagrange multiplier m_* in $L^q(\Omega)$. This is the content of the following proposition.

PROPOSITION 2.2. *Assume that $\|y_\gamma^+\|_{L^q} \leq c\gamma^{-1}$ holds true for some $1 < q < \infty$. Then there exists a Lagrange multiplier for the pointwise inequality constraints on the state in problem (1.1) which is an element of $L^q(\Omega)$.*

Proof. Since $\gamma \|y_\gamma^+\|_{L^q} \leq c$, the function γy_γ^+ has a weak accumulation point $m_* \in L^q(\Omega)$, which is nonnegative due to weak closedness of the nonnegative cone in $L^q(\Omega)$. Moreover, with $q' = \frac{q}{q-1}$ (and along a subsequence denoted here by (γ) as well) it holds that

$$\begin{aligned} |(m_*, y_*)_{L^2}| &= \lim_{\gamma \rightarrow \infty} |(\gamma y_\gamma^+, y_*)_{L^2}| = \lim_{\gamma \rightarrow \infty} |(\gamma y_\gamma^+, y_\gamma)_{L^2} + (\gamma y_\gamma^+, y_\gamma - y_*)_{L^2}| \\ &\leq \lim_{\gamma \rightarrow \infty} \gamma \|y_\gamma^+\|_{L^2}^2 + \lim_{\gamma \rightarrow \infty} \gamma \|y_\gamma^+\|_{L^q} \|y_\gamma - y_*\|_{L^{q'}} = 0 \end{aligned}$$

by strong convergence of $u_\gamma \rightarrow u_*$ in $L^2(\Omega)$, and thus $y_\gamma \rightarrow y_*$ in $L^\infty(\Omega)$ as well as $\gamma \|y_\gamma^+\|_{L^2}^2 \rightarrow 0$; see [10, Proposition 2.1] and our related comment at the beginning of this section. From the adjoint equation (1.3) one readily infers uniform boundedness of (p_γ) in $H^1(\Omega)$. Then, by compactness, p_γ converges strongly in $L^2(\Omega)$ to some p_* as $\gamma \rightarrow \infty$ (along a subsequence again denoted by (γ)). Consequently, $u_\gamma \rightarrow u_*$ in $L^2(\Omega)$, as well, and (1.4) is satisfied in the limit. Hence, m_* is the Lagrange multiplier for the pointwise inequality constraints on the state. \square

In general, one observes that the multipliers m_* are measures only. Thus, the case $s \geq 1$ for $q > 1$ only appears in exceptionally regular situations with respect to Lagrange multipliers (or completely inactive state inequality constraints). Hence, we generically expect $s < 1$ for $q > 1$. Consequently, the result of Lemma 2.1 appears to be optimal.

2.2. Estimates depending on the constraint violation in L^∞ . We approach our aim via the value functional

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

It was shown in [9] that $\lim_{\gamma \rightarrow \infty} V(\gamma) = J(x_*)$. Here we show that the rate of convergence depends on $\|y_\gamma^+\|_{L^\infty}$.

THEOREM 2.3. *The value functional V is differentiable and the following estimate for the derivative holds true:*

$$(2.1) \quad 0 \leq \frac{d}{d\gamma} V(\gamma) \leq \frac{c}{\gamma} \|y_\gamma^+\|_{L^\infty}.$$

If $\|y_\gamma^+\|_{L^\infty} \leq c\gamma^{-s}$ for some $s > 0$, uniformly as $\gamma \rightarrow \infty$, then one has

$$(2.2) \quad 0 \leq \frac{d}{d\gamma} V(\gamma) \leq c\gamma^{1-s},$$

$$(2.3) \quad 0 \leq J(x_*) - V(\gamma) \leq c\gamma^{-s},$$

$$(2.4) \quad \sqrt{\alpha} \|u_* - u_\gamma\|_{L^2} \leq c\gamma^{-s/2}.$$

Proof. In [9, Proposition 4.1] differentiability of V was shown and the expression

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

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

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

Due to Lemma 2.1 we have that $\gamma \|y_\gamma^+\|_{L^1} \leq c$. Hence, (2.1) follows.

Now assume that $\|y_\gamma^+\|_{L^\infty} \leq c\gamma^{-s}$ holds true. Then we obtain

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

Let $\gamma_1 > \gamma_2$ be given. Then, from the fundamental theorem of calculus we infer that

$$0 \leq V(\gamma_1) - V(\gamma_2) = \int_{\gamma_2}^{\gamma_1} \frac{d}{d\gamma} V(\gamma) d\gamma \leq \int_{\gamma_2}^{\gamma_1} c\gamma^{-1-s} d\gamma = c(\gamma_2^{-s} - \gamma_1^{-s}).$$

Since $\lim_{\gamma \rightarrow \infty} V(\gamma) = J(x_*)$ this estimate yields

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

which implies (2.3). Finally, (2.4) readily follows from the uniform convexity of J^γ with respect to u in $L^2(\Omega)$ and (2.3). Indeed, we have

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

which concludes the proof. \square

2.3. A worst case estimate for the constraint violation in L^∞ . The bottom line of the previous section is that we have to find an estimate for $\|y_\gamma^+\|_{L^\infty}$, which is as sharp as possible. This can be achieved by exploiting the smoothness of y_γ .

The essence of our technique is a geometric idea, which we explain by means of a simple example for illustration purposes; cf. Figure 2.1. Let $f(\omega) = a(-\omega^2 + \varepsilon^2)$ be a concave parabola, and $f^+(\omega) := \max(f(\omega), 0)$ its positive part. Then $f(\omega) \geq 0$ for $\omega \in [-\varepsilon, +\varepsilon]$ with a maximum $\|f^+\|_\infty = f(0) = a\varepsilon^2$, and $\|f^+\|_{L^1} = \int_{-\varepsilon}^{\varepsilon} f(\omega) d\omega = \frac{4}{3}a\varepsilon^3$. Thus, we have

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

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.

PROPOSITION 2.4. *Let $G \subset \mathbb{R}^d$ be bounded and open, $0 \leq y \in C^\beta(\overline{G})$, with $0 < \beta \leq 2$, and thus, in particular, $y \in L^1(G)$. Moreover, assume that $y = 0$ on ∂G . Then*

$$(2.5) \quad \|y\|_{L^\infty} \leq c(d, \beta) \|y\|_{C^\beta}^{1-\Theta} \|y\|_{L^1}^\Theta$$

with $\Theta = \frac{\beta}{\beta+d}$. The constant $c(d, \beta)$ depends only on β and on the dimension d , but not on the geometry of G .

Proof. Without loss of generality, assume that $0 \in G$, $y(0) = \|y\|_{L^\infty}$, and denote by $B_r(0)$ the ball of radius r and with center 0.

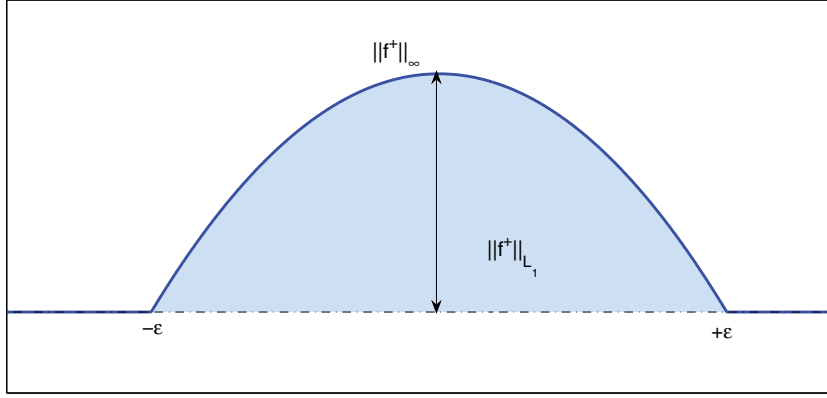


FIG. 2.1. Geometrical idea of our considerations.

If $\beta \leq 1$, then by the definition of Hölder continuity, it follows that $y(\omega) \geq y(0) - \|y\|_{C^\beta} r^\beta$ for all $\omega \in B_r(0)$. If $\beta > 1$, then y is once continuously differentiable and attains a maximum at 0. Hence, we conclude that $\nabla y(0) = 0$ for $\beta > 1$. Moreover, we can compute the value at any point $\omega \in B_r(0)$ using the fundamental theorem of calculus along the line $[0, \omega]$. Using the Hölder continuity of the first derivative, one then obtains $y(\omega) \geq y(0) - \|y\|_{C^\beta} r^\beta$ for all $\omega \in B_r(0)$ also in this case. In particular, $y(\omega)$ is positive for $\omega \in B_R(0)$ with

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

From the assumption $y = 0$ on ∂G it follows that $B_R(0) \subset \overline{G}$. Using a transformation to spherical coordinates, and, in the last line, our definition of R , we compute

$$\begin{aligned}
 (2.6) \quad \|y\|_{L^1} &= \int_{\overline{G}} |y(\omega)| \, d\omega \geq d \cdot \kappa_d \int_{[0, R]} |y(0) - \|y\|_{C^\beta} r^\beta| r^{d-1} \, dr \\
 &= d \cdot \kappa_d \|y\|_{C^\beta} \int_{[0, R]} \left(\frac{y(0)}{\|y\|_{C^\beta}} - r^\beta \right) r^{d-1} \, dr \\
 &= d \cdot \kappa_d \|y\|_{C^\beta} \int_{[0, R]} (R^\beta - r^\beta) r^{d-1} \, dr \\
 &= d \cdot \kappa_d \left(\frac{1}{d} - \frac{1}{d+\beta} \right) \|y\|_{C^\beta} R^{\beta+d} \\
 (2.7) \quad &= \kappa_d \frac{\beta}{d+\beta} \|y\|_{C^\beta}^{1-\frac{\beta+d}{\beta}} \|y\|_{L^\infty}^{\frac{\beta+d}{\beta}}.
 \end{aligned}$$

Here $\kappa_d = \text{Vol}(B_1)$ is the volume of the d -dimensional unit sphere. Solving for $\|y\|_{L^\infty}$, we obtain

$$\|y\|_{L^\infty} \leq \underbrace{\left(\kappa_d \frac{\beta}{d+\beta} \right)^{-\frac{\beta}{\beta+d}}}_{=: c(d, \beta)} \|y\|_{L^1}^{\frac{\beta}{\beta+d}} \|y\|_{C^\beta}^{1-\frac{\beta}{\beta+d}}$$

as stated in the assertion. \square

Remark 2.5. Observe that (2.5) is true only 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 necessarily of higher, order.

For $1 \leq \beta \leq 2$ our argumentation is only valid due to our assumption $y|_{\partial G} = 0$. This allows us to exploit that the maximizer of y exists in the open set G . If this assumption is dropped, then our result still holds for $\beta \leq 1$ under a cone condition on G . This is of interest when boundary conditions different from Dirichlet conditions are considered.

It is of crucial importance that $c(d, \beta)$ in (2.5) is independent of the geometry of G , since we will apply Proposition 2.4 with G depending on y_γ . This is not the case for standard inequalities of similar form, such as the Gagliardo–Nirenberg inequality.

Also, in [11, Lemma 4.7] similar techniques have been employed for the case $\beta \leq 1$ in the context of the virtual control approach, which yield the same worst case bounds on $\|y_\gamma^+\|_\infty$ in this restricted range of β .

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

$$(2.8) \quad \|y_\gamma^+\|_{L^\infty} \leq c\gamma^{-s}, \text{ where } s = \frac{\beta}{\beta + d}.$$

In particular, for the example problem (1.1) we have for every $\varepsilon > 0$:

$$(2.9) \quad \|y_\gamma^+\|_{L^\infty} \leq \begin{cases} c\gamma^{-2/3} & : d = 1, \\ c\gamma^{-1/2+\varepsilon} & : d = 2, \\ c\gamma^{-1/4+\varepsilon} & : d = 3. \end{cases}$$

Proof. Let $G := \{\omega \in \overline{\Omega} : y_\gamma^+(\omega) > 0\}$. Since $y_\gamma \in C^\beta(\overline{\Omega})$ we have by restriction that $y_\gamma^+|_G = y_\gamma|_G \in C^\beta(G)$ with $\|y_\gamma^+\|_{C^\beta(G)} \leq \|y_\gamma\|_{C^\beta(\overline{\Omega})}$. In this way, the result (2.8) follows readily from Proposition 2.4, applied to y_γ^+ on G , and from Lemma 2.1.

For deriving (2.9), we have to invoke standard regularity results for partial differential equations with measures in the right-hand side; see, e.g., [3, Theorem 4]. Let $q' = d/(d-1)$ for $d > 1$ and q' arbitrarily large if $d = 1$. Then, since $\gamma\|y_\gamma^+\|_{L^1}$ is uniformly bounded, it follows that u_γ is uniformly bounded in $W^{1,q'}(\Omega)$. This implies that y_γ is uniformly bounded in $W^{3,q'}(\Omega)$. It follows from 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 $\varepsilon > 0$. \square

In our numerical experiments below, we shall see that our technique yields sharp estimates in geometric situations where the graph of the state has a shape similar to an elliptic paraboloid, as modeled in our proof above. Such a configuration occurs when the active set is indeed a single point, only. In more regular situations, however, larger values of s are observed. We provide some further insight into this in section 2.6.

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

Finally, we establish a generic upper bound on s .

PROPOSITION 2.8. *If $s \geq 1$, then problem (1.1) has a Lagrangian multiplier in $L^q(\Omega)$ for each $1 \leq q < \infty$.*

Proof. This is a direct consequence of Proposition 2.2. \square

2.4. The length of the primal-dual path. When combining our estimates above, we obtain the following convergence estimate.

THEOREM 2.9. *If y_γ is uniformly bounded in $C^\beta(\Omega)$ for $\gamma \rightarrow \infty$ and some $0 < \beta \leq 2$, then the primal-dual path satisfies the following convergence estimate:*

$$(2.10) \quad 0 \leq J(x_*) - V(\gamma) \leq c\gamma^{-\frac{\beta}{\beta+d}}$$

and

$$(2.11) \quad \sqrt{\alpha} \|u_* - u_\gamma\|_{L^2} \leq c\gamma^{-\frac{\beta}{2(\beta+d)}}.$$

In particular, for every $\varepsilon > 0$ it holds that

$$(2.12) \quad \sqrt{\alpha} \|u_* - u_\gamma\|_{L^2} \leq \begin{cases} c\gamma^{-1/3} & : d = 1, \\ c\gamma^{-1/4+\varepsilon} & : d = 2, \\ c\gamma^{-1/8+\varepsilon} & : d = 3. \end{cases}$$

In both situations, the constant $c > 0$ is independent of γ .

Proof. The results (2.10) and (2.11) follow from Theorem 2.3 and Corollary 2.6. In the same way, (2.12) follows from (2.9). \square

Remark 2.10. It is of interest to compare these results with corresponding estimates obtained for interior point path-following methods [14], where a parameter μ is sent to 0. There, regardless of the dimension and regularity of the problem, one can show $J(x_*) - V(\mu) \leq c\mu$ and $\sqrt{\alpha} \|u_* - u_\gamma\|_{L^2} \leq c\mu^{1/2}$. Thus, to obtain comparable guaranteed accuracy, one has to choose $\gamma \sim \mu^{-2}$ in two dimensions, and $\gamma \sim \mu^{-4}$ in three dimensions (and not straightforwardly $\gamma \sim \mu^{-1}$).

We note that the dimension and regularity of the problem enter the analysis of interior point methods in a different way. They influence how the minimal distance between homotopy solutions and the bounds (i.e., $\min_{\omega \in \bar{\Omega}} \psi(\omega) - y_\mu(\omega)$) depends on μ . Indeed, similar techniques, as in the proof of Proposition 2.4, have already been applied in [14] to show that this minimal distance is nonzero for $\mu > 0$ and for an appropriate choice of the barrier functional.

In summary, the same quantities that govern the length of the Moreau–Yosida path affect the choice of barrier functional in interior point methods.

2.5. A bootstrapping argument. By using bootstrapping arguments we now come back to the possible improvement of the results obtained in [8]. For this purpose, we derive a bound on the L^∞ -norm based on a boundedness result for the L^2 -norm of the feasibility violation. The result will be the same as in the previous section where we used the boundedness of the L^1 norm, but it shows that the estimate cannot be further improved by using the bootstrapping technique.

PROPOSITION 2.11. *Let $G \subset \mathbb{R}^d$ be bounded and open, $0 \leq y \in C^\beta(\bar{G})$, $y \in L^2(G)$, and $0 < \beta \leq 2$. Moreover, assume that $y = 0$ on ∂G . Then*

$$(2.13) \quad \|y\|_{L^\infty} \leq c_2(d, \beta) \|y\|_{C^\beta}^{1-\Theta} \|y\|_{L^2}^\Theta$$

with $\Theta = \frac{2\beta}{2\beta+d}$. The constant $c_2 = c(d, \beta)$ is independent of y and G .

Proof. The proof is analogue to the one of Proposition 2.4 with the following replacement for (2.6):

$$\begin{aligned}
 (2.14) \quad \|y\|_{L^2}^2 &= \int_G |y(\omega)|^2 d\omega \geq d \cdot \kappa_d \int_{[0,R]} |y(0) - \|y\|_{C^\beta} r^\beta|^2 r^{d-1} dr \\
 &= d \cdot \kappa_d \|y\|_{C^\beta}^2 \int_{[0,R]} \left(\frac{y(0)}{\|y\|_{C^\beta}} - r^\beta \right)^2 r^{d-1} dr \\
 &= d \cdot \kappa_d \|y\|_{C^\beta}^2 \int_{[0,R]} (R^\beta - r^\beta)^2 r^{d-1} dr \\
 &\geq d \cdot \kappa_d \left(\frac{1}{d} - \frac{2}{\beta+d} + \frac{1}{2\beta+d} \right) \|y\|_{C^\beta}^2 R^{2\beta+d} \\
 &= d \cdot \kappa_d \left(\frac{1}{d} - \frac{2}{\beta+d} + \frac{1}{2\beta+d} \right) \|y\|_{C^\beta}^{2-\frac{2\beta+d}{\beta}} \|y\|_{L^\infty}^{\frac{2\beta+d}{\beta}}.
 \end{aligned}$$

Here $\kappa_d = \text{Vol}(B_1)$ is the volume of the d -dimensional unit sphere. Solving for $\|y\|_{L^\infty}$, we conclude that

$$\|y\|_{L^\infty} \leq c(d, \beta) \|y\|_{L^2}^{\frac{2\beta}{2\beta+d}} \|y\|_{C^\beta}^{1-\frac{2\beta}{2\beta+d}}$$

holds with a constant independent of y and G . \square

Starting from the feasibility violation in $L^2(\Omega)$ and using a bootstrapping argument we next establish the same results as in Corollary 2.6 for the L^∞ -norm of the feasibility violation.

THEOREM 2.12. *Let y_γ be uniformly bounded in $C^\beta(\overline{\Omega})$ for $\gamma \rightarrow \infty$ and some $0 < \beta \leq 2$. Then the following estimate on the constraint violation holds true:*

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

Proof. By Theorem 2.3, and, in particular, (2.1), we have that

$$(2.15) \quad \gamma \|y_\gamma^+\|_{L^2}^2 \leq c_0 \|y_\gamma^+\|_{L^\infty}.$$

Moreover, due to the convergence of $\gamma \|y_\gamma^+\|_{L^2}^2 \rightarrow 0$, there exists a constant C_0 such that for all $\gamma > 0$,

$$\gamma \|y_\gamma^+\|_{L^2}^2 \leq C_0.$$

In particular, with the constant C_0 and $\theta_0 = 0$ it holds that

$$(2.16) \quad \|y_\gamma^+\|_{L^2}^2 \leq C_0 \gamma^{-1-\theta_0}.$$

From Proposition 2.11 one can then obtain convergence rates for $\|y_\gamma^+\|_{L^\infty}$, and these in turn allow us to improve upon the feasibility violation in L^2 using (2.15). By repeating these steps we will obtain a sequence of elements $C_n \in \mathbb{R}$ and $\theta_n \geq 0$ such that

$$\|y_\gamma^+\|_{L^2}^2 \leq C_n \gamma^{-1-\theta_n}.$$

To obtain explicit formulas for C_n and θ_n we assume for the purpose of induction that the above formula holds for some $n \geq 0$. Then by Proposition 2.11 and the

assumption above, we obtain

$$\begin{aligned}\|y_\gamma^+\|_{L^\infty} &\leq \underbrace{c_2 \|y_\gamma^+\|_{C^\beta}^{1-\frac{2\beta}{2\beta+d}}}_{\widehat{c}} \|y_\gamma^+\|_{L^2}^{\frac{2\beta}{2\beta+d}} \\ &\leq \widehat{c} C_n^{\frac{\beta}{2\beta+d}} \gamma^{(-1-\theta_n)\frac{\beta}{2\beta+d}}.\end{aligned}$$

Combining this estimate with our initial bound $\gamma \|y_\gamma^+\|_{L^2}^2 \leq c_0 \|y_\gamma^+\|_{L^\infty}$ yields the improved rate

$$\|y_\gamma^+\|_{L^2}^2 \leq c_0 \widehat{c} C_n^{\frac{\beta}{2\beta+d}} \gamma^{-1-\theta_{n+1}}$$

with $\theta_{n+1} = (1 + \theta_n) \frac{\beta}{2\beta+d}$. Now we set $C_{n+1} := c_0 \widehat{c} C_n^{\frac{\beta}{2\beta+d}}$ and deduce the following formulas for C_n and θ_n , where $q := \frac{\beta}{2\beta+d} < 1$:

$$C_n = \left((c_0 \widehat{c})^{\sum_{k=0}^{n-1} q^k} \right) C_0^{(q^n)}, \quad \theta_n = \sum_{k=1}^n q^k.$$

As $q < 1$, both sequences converge and we infer the existence of $C_\infty \in \mathbb{R}$ and

$$\theta_\infty = \sum_{k=1}^{\infty} q^k = \frac{1}{1-q} - 1 = \frac{2\beta+d}{\beta+d} - 1 = \frac{\beta}{\beta+d}.$$

By definition it holds that

$$\|y_\gamma^+\|_{L^2}^2 \leq C_\infty \gamma^{-1-\theta_\infty}.$$

An additional application of Proposition 2.11 then yields

$$\|y_\gamma^+\|_{L^\infty} \leq \widehat{c} \|y_\gamma^+\|_{L^2}^{\frac{2\beta}{2\beta+d}} \leq \widehat{c} C_\infty^{\frac{\beta}{2\beta+d}} \gamma^{(-1-\theta_\infty)\frac{\beta}{2\beta+d}} = c \gamma^{-\frac{\beta}{\beta+d}}$$

and thus the assertion. \square

2.6. Structural assumptions and refined results. In the numerical section below we shall see that the rate of convergence of the primal-dual path seems to depend on the structure of the Lagrange multiplier at the optimal solution. For example, it can be observed that rates differ between line measures and point measures. In order to clarify this issue, in this section we provide an in-depth discussion.

The following structural assumptions are motivated by the observation that Lagrangian multipliers for state constraints are frequently concentrated along smooth curves or more generally on smooth submanifolds of Ω . This may happen, for example, if the active set itself is a smooth submanifold of Ω or a smoothly bounded subset of Ω .

We also note that the following regularity, Assumption 2.13, applied to an arbitrary measure m , is widely used in the theory of Sobolev spaces. It is the simplest instance of a regularity condition for measures and is used in refined theories of Sobolev spaces, as those treated in the monograph [12], for example. This assumption holds, for instance, when m is a measure with support on a δ -dimensional submanifold in Ω and it is represented by a bounded function. In order to emphasize the role of γy_γ^+ as a measure approximation, we abbreviate $m_\gamma := \gamma y_\gamma^+$.

Assumption 2.13. For $\omega \in \Omega$ denote by $B_R(\omega)$ the open ball of radius $R > 0$ and with center ω . Let $\delta \geq 0$ be given. Assume that there exists $K < \infty$, independent of ω, R, δ , and γ , such that the following inequality holds:

$$(2.17) \quad \int_{B_R(\omega)} dm_\gamma \leq KR^\delta$$

for all sufficiently small $R > 0$.

Let us motivate our requirement (2.17). The exponent δ measures the actual amount of singularity of the measure. For instance, if m is induced by a bounded function on Ω , then $\delta = d$. On the other hand, if m is a line measure, induced by a bounded function on this line, then $\delta = 1$. Noninteger values of δ may arise in cases where the measures are induced by L^q -functions on submanifolds of $\bar{\Omega}$.

THEOREM 2.14. *Suppose that Assumption 2.13 holds true for some $\delta \geq 0$. Then the estimate for the constraint violation can be refined as*

$$(2.18) \quad \|y_\gamma^+\|_\infty \leq c(d, \beta, K) \gamma^{-s} \|y_\gamma\|_{C^\beta}^{1-s} \quad \text{with } s = \frac{\beta}{\beta + d - \delta}.$$

Proof. Assume that y_γ attains its maximum at ω_γ . As in Proposition 2.4 we choose $R = (\|y_\gamma^+\|_\infty / \|y_\gamma\|_{C^\beta})^{1/\beta}$ and compute

$$K \gamma^{-1} R^\delta \geq \int_{B_R(\omega_\gamma)} y_\gamma^+ dx \geq \tilde{c}(d, \beta) \|y_\gamma\|_{C^\beta} R^{\beta+d} = \tilde{c}(d, \beta) \|y_\gamma\|_{C^\beta}^{1-\frac{\beta+d}{\beta}} \|y_\gamma^+\|_\infty^{\frac{\beta+d}{\beta}}.$$

We solve the above inequality for $\|y_\gamma^+\|_\infty$ and set $s = \frac{\beta}{\beta+d-\delta}$ to obtain

$$\|y_\gamma^+\|_\infty \leq c(d, \beta, K) \gamma^{-s} \|y_\gamma\|_{C^\beta}^{1-s},$$

which concludes the proof. \square

Our assumption yields additional regularity. In particular we obtain higher regularity of p_γ and p_* (and thus of the constraint violation) if the exponent δ in (2.17) is large.

PROPOSITION 2.15. *Suppose that Assumption 2.13 holds true for some $0 \leq \delta \leq d$. Then (2.17) also holds for the Lagrange multiplier m_* , i.e., $\int_{B_R(\omega)} dm_* \leq KR^\delta$. Moreover, for any*

$$l \in \{0, 1\} \quad \text{and} \quad q' \in \left[\max \left(1, \frac{d-\delta}{2-l} \right), d \right],$$

one has $p_\gamma \in W^{l,q}(\Omega)$ uniformly, where $1/q + 1/q' = 1$. The same holds true for p_* .

Proof. Due to weak*-convergence $m_\gamma \rightharpoonup^* m_*$ as $\gamma \rightarrow +\infty$ (see [9, Proposition 2.1]), we have

$$\int_{B_R(\omega)} dm_\gamma \rightarrow \int_{B_R(\omega)} dm_*,$$

which implies that (2.17) also holds for m_* .

In the following we apply a result of [12, section 1.4.5]. In fact, let m be a measure that satisfies (2.17) for certain δ and K , and denote by $L^1(\Omega, m)$ the space of all functions that are integrable with respect to the measure m . Then [12, section 1.4.5] asserts the existence of a continuous embedding $W^{2-l,q'}(\Omega) \hookrightarrow L^1(\Omega, m)$ for all l and

q' that satisfy our assumptions. This can be viewed as a generalization and unification of well-known Sobolev embedding and trace theorems.

The application of this result to $m = m_\gamma$ and $m = m_*$ yields the conclusion that m_γ and m_* are continuous linear functionals on $W^{2-l,q'}(\Omega)$ via the mappings

$$v \rightarrow \int_{\Omega} v \, dm_\gamma \quad \text{and} \quad v \rightarrow \int_{\Omega} v \, dm_* \quad \text{for } v \in W^{2-l,q'}(\Omega),$$

respectively, and thus they are elements of $W^{2-l,q'}(\Omega)^*$.

The differential operator $A = -\Delta : W^{2-l,q'}(\Omega) \rightarrow (W^{l,q}(\Omega))^*$ is an isomorphism due to our regularity assumptions. Thus, $A^* : W^{l,q}(\Omega) \rightarrow (W^{2-l,q'}(\Omega))^*$ is an isomorphism as well. This follows from the general facts that $(T^*)^{-1} = (T^{-1})^*$ and $\|T\| = \|T^*\|$ for any continuous linear mapping T .

Hence, $p_\gamma = (A^*)^{-1}(y - y_d + m_\gamma)$ and $p_* = (A^*)^{-1}(y - y_d + m_*)$ are uniformly bounded in $W^{l,q}(\Omega)$. \square

COROLLARY 2.16. *Suppose that Assumption 2.13 holds true in each of the following cases. Then (2.2), (2.3), and (2.4) in Theorem 2.3 hold with the following values of s :*

- (i) (line measure in 2d) If $d = 2, \delta = 1$, then $s = 2/3$.
- (ii) (surface measure in 3d) If $d = 3, \delta = 2$, then $s = 2/3$.
- (iii) (line measure in 3d) If $d = 3, \delta = 1$, then $s = 1/2 - \varepsilon$ for each $\varepsilon > 0$.

Proof. From our results of Proposition 2.15 we obtain p_γ uniformly bounded in $W^{1,q}(\Omega)$ for arbitrary $q < \infty$ in the cases (i) and (ii), and thus $y \in W^{3,q}(\Omega) \hookrightarrow C^2(\overline{\Omega})$ uniformly. Hence, we can apply Theorem 2.14 with $\beta = 2$.

In case (iii) we obtain $p_\gamma \in L^q(\Omega)$ for arbitrary $q < \infty$, only, and thus $y \in W^{2,q}(\Omega) \hookrightarrow C^{2-\varepsilon}(\overline{\Omega})$ for all $\varepsilon > 0$. Theorem 2.14 with $\beta = 2 - \varepsilon$ then yields the result. \square

Comparing this corollary with Corollary 2.6 we see that we obtain the same rates for d and δ as for $d - \delta$ in Corollary 2.6. This confirms that problems with higher regularity of the Lagrange multipliers essentially behave like lower dimensional problems.

Remark 2.17. The refined estimates given here for (2.3) and (2.4) in Theorem 2.3 rely on the uniform boundedness of the constant K for $\gamma \rightarrow \infty$. Thus, from an analytic point of view it would be desirable to have a converse to Proposition 2.15: Higher regularity of m_* should imply similar regularity of m_γ for sufficiently large γ . The authors have little hope that such a converse can be found under reasonable assumptions, because it is well conceivable that the topology of the active set may change arbitrarily often during the path-following procedure in pathological situations (cf., e.g., [7]).

From an algorithmic point of view, however, the local bound (2.2) is more interesting, because it gives local information on the *sensitivity of the path* with respect to γ . This result requires only that a local version of Assumption 2.13 holds, namely for values of γ that are close to the current γ used in the path-following scheme.

In summary, the results in this section give insight into the connection of two phenomena, which are observed in practice: The local exponent $s = s(\gamma)$ (governing the sensitivity of the path) and the regularity of m_γ . We will observe in the numerical experiments that s indeed depends on γ in some preasymptotic phase, but seems to approach a constant value for $\gamma \rightarrow \infty$, at least in the considered examples.

2.6.1. Active sets with nonempty interior. Finally, we show that further improvement of the convergence rate is possible, when the active set has a nonempty

interior. Here, we confine our discussion to the one-dimensional (1d)-case, but augment our findings by comments concerning the general case.

As a motivation for the following proof, consider the typical setting for $d = 1$ with an active set which has a nonempty interior; for numerical details we refer to section 3. Since y_* is constant on the active interval the derivatives $y_*^{(m)}$ satisfy $y_*^{(m)} = 0$ for all $m \in \mathbb{N}_0$ on this set. Consider the functions $f(\omega) = a(-\omega^3 + \varepsilon\omega^2)$ and $f^+ = \max(f, 0)$ on $[0, \infty[$; see the left plot of Figure 2.2 for a graphical illustration.

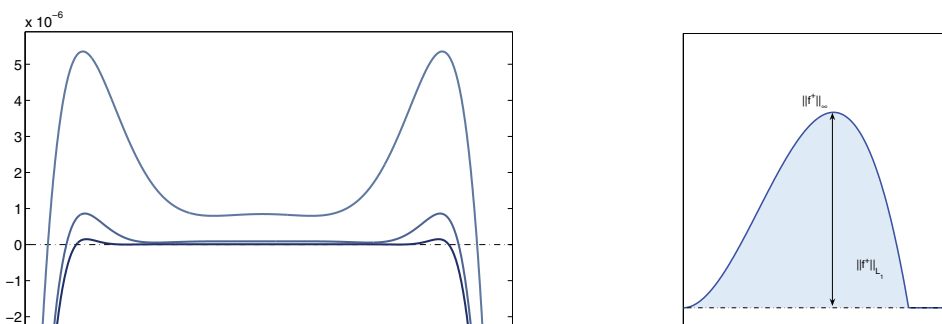


FIG. 2.2. Geometric situation for y_γ near an active interval. Left: Observed in numerical experiments for $\gamma = 10^7, 10^8$, and 10^9 . Right: Idealized model.

Obviously, $f^{(m)}(0) = 0$ for $m = 1, 2$. Clearly, f is positive on $[0, \varepsilon]$ and has a maximum at $\omega = 2/3\varepsilon$ with $f(2/3\varepsilon) = 10/27a\varepsilon^3$, thus, $\|f^+\|_\infty = ca\varepsilon^3$. Further, $\|f^+\|_{L^1} = \int_0^\varepsilon f(\omega) d\omega = a\varepsilon^4/12$. Hence, similarly as above and using $6a = \|f\|_{C^3}$, we conclude that

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

THEOREM 2.18. *Let $d = 1$, let Ω be a bounded interval, and assume that for all sufficiently large γ the support of y_γ^+ consists of finitely many intervals of positive length. Assume that the minimal length of these intervals is bounded from below by $a_0 > 0$, which is independent of γ . Then, for each $\varepsilon > 0$ one has*

$$\|y_\gamma^+\|_\infty \leq c\gamma^{-s} \text{ with } s = \frac{3}{4} - \varepsilon.$$

Proof. Since our discussion is limited to the 1d case, we obtain that y_γ is bounded in $W^{3,\infty}(\Omega)$ by a constant M as $\gamma \rightarrow \infty$. Moreover, we have shown that $\|y_\gamma^+\|_\infty \leq c\gamma^{-s}$ for $s = 2/3$. A bootstrapping technique with respect to s allows us to increase s to any value lower than $3/4$.

As a special case of the Gagliardo–Nirenberg inequalities (cf., e.g., [18, Appendix]) it follows for a fixed domain (say, the unit interval I) that

$$\|\nabla^j f\|_{L^\infty(I)} \leq K \|f\|_{W^{m,\infty}(I)}^\theta \|f\|_{L^\infty(I)}^{1-\theta},$$

where $\theta = j/m$. Transforming I to an interval of arbitrary positive length a through $x \mapsto ax + x_0$ shows that the constant K solely depends on the length a . The constant degenerates only if $a \rightarrow 0$ or $a \rightarrow \infty$.

Let I_γ denote the subinterval where y_γ^+ attains its maximum. By assumption, the length of I_γ is bounded from below by a_0 and from above by the length of Ω . Using

our convention that $C^\beta(\Omega)$, for integral $\beta > 1$, contains elements whose derivatives up to the order $(\beta - 1)$ are Lipschitz continuous. We, hence, get

$$\|y_\gamma^+\|_{C^2(I_\gamma)} \leq K(a_0) \|y_\gamma^+\|_{W^{3,\infty}(I_\gamma)}^{2/3} \|y_\gamma^+\|_{L^\infty(I_\gamma)}^{1/3} \leq K(a_0) M^{2/3} \gamma^{-s/3}.$$

Employing Theorem 2.14 we obtain with $s_0 = \beta/(1 + \beta) = 2/3$ that

$$\|y_\gamma^+\|_\infty \leq \|y_\gamma^+\|_{C^2(I_\gamma)}^{1-s_0} \gamma^{-s_0} \leq c \gamma^{-(\frac{2}{3}\frac{1}{3}) - \frac{2}{3}}.$$

Thus, we conclude that

$$\|y_\gamma^+\|_\infty \leq c \gamma^{-s} \implies \|y_\gamma^+\|_\infty \leq c \gamma^{-\frac{2}{3} - \frac{2}{3}}.$$

Consequently, starting with $s = s_0 = 2/3$ we obtain a sequence (s_n) whose elements are defined recursively by

$$(2.19) \quad s_{n-1} \mapsto s_n := \frac{s_{n-1}}{9} + \frac{2}{3}.$$

A short computation yields that this sequence converges monotonically to the fixed point $s = 3/4$ of (2.19), which shows our result. \square

Remark 2.19. When one tries to extend this result to higher dimensions, one has to find an analogue to the stated minimal length assumption, in the sense that the maximum of y_γ^+ lies in the interior of a subset of $\text{supp } y_\gamma^+$, the support of y_γ^+ , that can be transformed to a fixed domain, say the unit square, such that the transformations do not degenerate for $\gamma \rightarrow \infty$. This will lead to uniform cone conditions on the boundaries of the active sets, independent of γ . While this conceptual idea is rather clear, the mathematical realization, however, is technical and, in our opinion, yields no further insight. Hence we refrain from pursuing this direction here.

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 behavior in practice.

In order to measure the exponent s we perform a numerical path-following method for the system (1.6), implemented in MATLAB for the case $d = 1$. For this purpose, we use a prescribed sequence of parameters γ_j . In our test problems we choose a constant desired state $y_d \equiv 10$, the control cost $\alpha = 1$, and obtain different types of active sets by varying the constant upper bound ψ . The discretization of the state y and the adjoint state p is based on classical finite differences, i.e., the 3-point stencil in the one-dimensional setting. In the 2d-case we use the optimization toolbox **DOpElib** [6] based on the finite element library **deal.i** [1, 2]. For obtaining a high resolution of the quantity $\|y_\gamma^+\|_\infty$ we use anisotropic mesh refinement and locally refined grids with piecewise bilinear finite elements.

A local estimate for the values of s close to γ_j is then computed by

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

Observe that this formula is quite sensitive to perturbations of y , which partially explains the slightly oscillatory behavior observed in the plots.

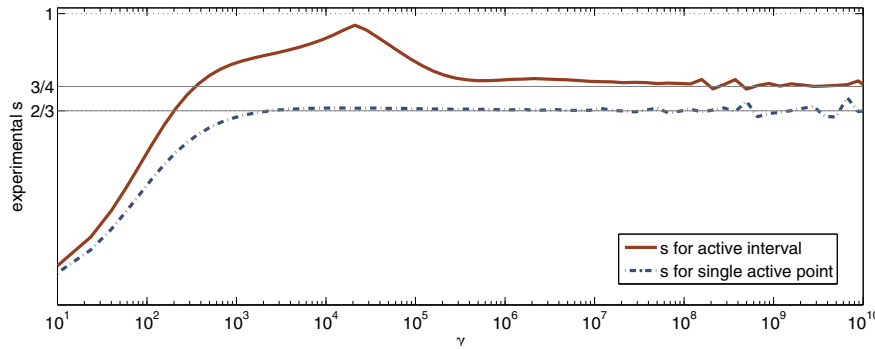


FIG. 3.1. Experimental results on the exponent s of the relation $\|y_\gamma^+\|_{L^\infty} = O(\gamma^{-s})$ in a 1d setting.

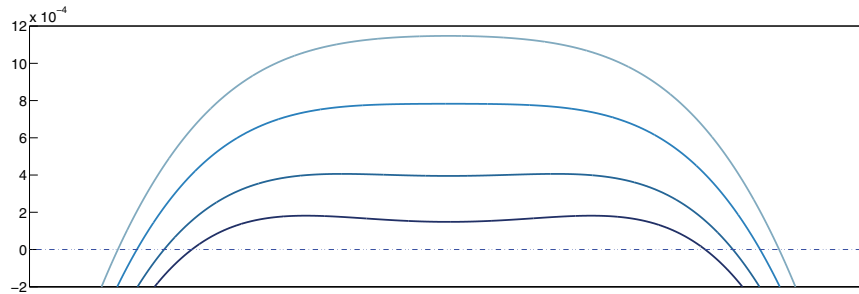


FIG. 3.2. Transition from single maximum to double maximum of y_γ for $\gamma = 7.5 \cdot 10^{-5}, 5 \cdot 10^{-5}, 2.5 \cdot 10^{-5}$, and 10^{-5} .

3.1. Experimental results in the 1d case. In the 1d 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 3.1, $s \approx 2/3$ as predicted by Corollary 2.6.

In our second setting (with $\psi \equiv 0.01$), the active set is a proper interval. Here we observe experimentally $s \approx 3/4$, which is a higher rate than predicted by Corollary 2.6. However, this rate can be explained by the results of Theorem 2.18.

Another interesting aspect is the large value of s in the range $\gamma \in [10^4, 10^5]$. A close look at the corresponding intermediate solutions shows that for the associated y_γ have a very flat maximum, which explains this behavior. In terms of Assumption 2.13 this means $\delta \approx d$ for an intermediate range of γ . For larger values of γ this single maximum splits into two maxima, which can be nicely observed in Figure 3.2. For larger values of γ the qualitative behavior of y_γ^+ can be observed in Figure 2.2.

3.2. Experimental results in the 2d case. 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 Lagrange multipliers.

If the active set is a single point (for $\psi \equiv 0.04$), we observe in Figure 3.3 that our estimates of Corollary 2.6 coincide with the numerical rate of convergence, namely

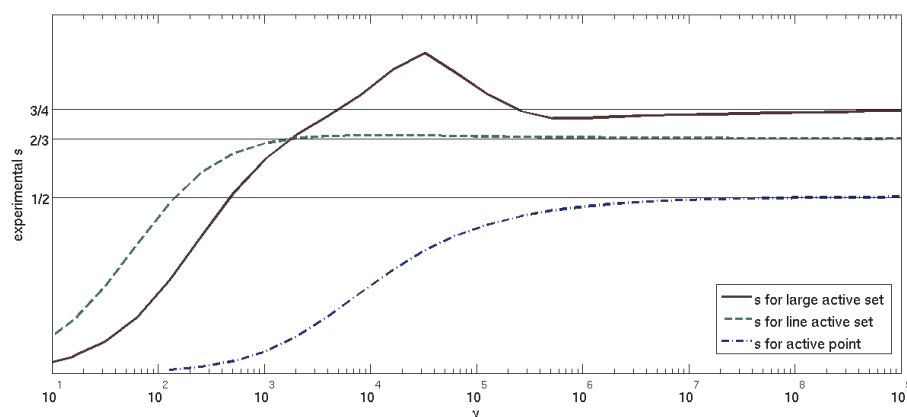


FIG. 3.3. Experimental results on the exponent s of the relation $\|y_\gamma^+\|_{L^\infty} = O(\gamma^{-s})$ in a 2d setting.

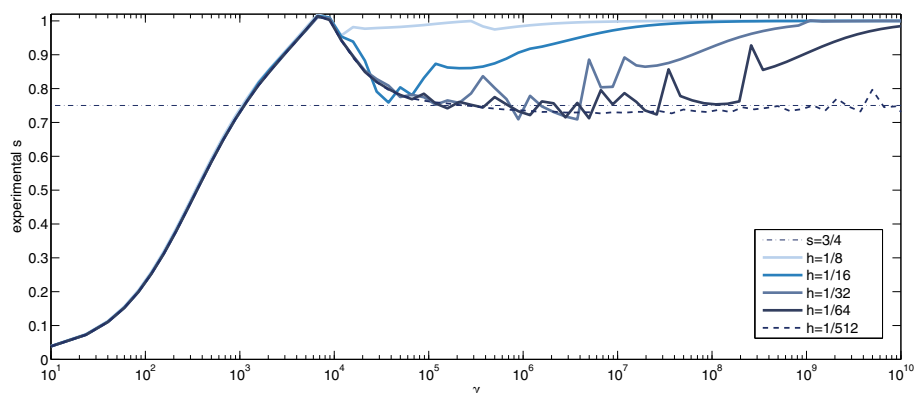


FIG. 3.4. Experimental results on the exponent in a 2d setting with large active set for various coarse grids. For comparison the asymptotics for a fine grid $h = 1/512$ is added as a dashed line.

$s \approx 0.5$. We observe the predicted rate of $s \approx 0.75$ in the case of the active set with nonempty interior (for $\psi \equiv 0.01$).

The computations were done using the optimization toolkit **D0pElib** based upon **deal.ii**, using adaptive mesh refinement to accurately capture the behavior of $\|y_\gamma^+\|_\infty$. On the final mesh, which yielded a total of approximately 1.5 million unknowns, the smallest elements had a diameter of 2^{-12} .

In order to obtain numerical results for a line measure we use a slightly different numerical setting. In fact, we chose homogeneous Neumann boundary conditions on two opposing boundary faces (i.e., $\{\omega_2 \in \{0, 1\}\}$) of the unit square and homogeneous Dirichlet conditions on the other two faces (i.e., $\{\omega_1 \in \{0, 1\}\}$). For obtaining a high resolution for the line measure case an anisotropic mesh refinement is employed such that the mesh is refined more in the ω_1 -direction yielding an aspect ratio of 1:128 with a total number of 1065090 degrees of freedom.

3.3. Effect of a fixed discretization on the rate. On a fixed mesh, when y_γ is computed up to very large values of γ , then it can be observed (see Figure 3.4) that

a rate of $s = 1$ 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 (indicated by subscript h) $\gamma \|(y_\gamma^+)_h\|_{L^1}$ we eventually observe the upper bound

$$\|(y_\gamma^+)_h\|_{L^\infty} \leq c(h) \|(y_\gamma^+)_h\|_{L^1} \leq c(h) \gamma^{-1} \iff s = 1.$$

If such a behavior is observed in practice, it is a clear indication that the problem has been “oversolved” numerically, or that the multiplier m_* is in $L^q(\Omega)$; see Proposition 2.8.

4. Conclusion. Numerical experiments indicate that, at least in 1- and 2d problems, the asymptotics on the constraint violation derived in this note are sharp, when the active set is a singleton. Moreover, we provide arguments concerning the implication of the regularity of the multiplier on the observed convergence of the constraint violation. Numerical experiments indicate that the bounds obtained in this paper are sharp in all these cases.

From a practical point of view, the observation that one has

$$\|y_\gamma^+\|_{L^\infty} \leq c\gamma^{-s},$$

where s generically varies in an interval $s \in [s_{\text{point}}, 1[$ in turn implies a rate of convergence for the error in the control which varies between $[s_{\text{point}}/2, 1/2[$. This may help in the construction of adaptive algorithms, which try to balance algebraic errors and discretization errors and clearly shows that an a posteriori estimation of this error is necessary as the rates depend on the a priori unknown behavior of the feasibility violation.

REFERENCES

- [1] W. BANGERTH, R. HARTMANN, AND G. KANSCHAT, *deal.II—a general purpose object oriented finite element library*, ACM Trans. Math. Software, 33 (2007), pp. 24/1–24/27.
- [2] W. BANGERTH, T. HEISTER, G. KANSCHAT, ET AL., *deal.II Differential Equations Analysis Library*, Technical reference, available online at <http://www.dealii.org>.
- [3] E. CASAS, *Control of an elliptic problem with pointwise state constraints*, SIAM J. Control Optim., 24 (1986), pp. 1309–1318.
- [4] S. CHEREDNICHENKO, K. KRUMBIEGEL, AND A. RÖSCH, *Error estimates for the Lavrentiev regularization of elliptic optimal control problems*, Inverse Problems, 24 (2008), 055003.
- [5] K. DECKELNICK AND M. HINZE, *Convergence of a finite element approximation to a state-constrained elliptic control problem*, SIAM J. Numer. Anal., 45 (2007), pp. 1937–1953.
- [6] C. GOLL, T. WICK, AND W. WOLLNER, *The differential equation and optimization environment: DOpElib*, available online at <http://www.dopelib.net>.
- [7] A. HERMANT, *Homotopy algorithm for optimal control problems with a second-order state constraint*, Appl. Math. Optim., 61 (2010), pp. 85–127.
- [8] M. HINTERMÜLLER AND M. HINZE, *Moreau–Yosida regularization in state constrained elliptic control problems: Error estimates and parameter adjustment*, SIAM J. Numer. Anal., 47 (2009), pp. 1666–1683.
- [9] M. HINTERMÜLLER AND K. KUNISCH, *Feasible and noninterior path-following in constrained minimization with low multiplier regularity*, SIAM J. Control Optim., 45 (2006), pp. 1198–1221.
- [10] M. HINTERMÜLLER AND K. KUNISCH, *PDE-constrained optimization subject to pointwise constraints on the control, the state, and its derivative*, SIAM J. Optim., 20 (2009), pp. 1133–1156.
- [11] K. KRUMBIEGEL, *Numerical Concepts and Error Analysis for Elliptic Neumann Boundary Control Problems with Pointwise State and Control Constraints*, Ph.D. thesis, Universität Duisburg-Essen, Essen, Germany, 2009.
- [12] V. MAZ'YA, *Sobolev Spaces: with Applications to Elliptic Partial Differential Equations*, Springer, Heidelberg, 2011.

- [13] R. RANNACHER, B. VEXLER, AND W. WOLLNER, *A posteriori error estimation in PDE-constrained optimization with pointwise inequality constraints*, in *Constrained Optimization and Optimal Control for Partial Differential Equations*, International Series of Numerical Mathematics 160, Springer, Basel, 2011, pp. 349–373.
- [14] A. SCHIELA, *Barrier methods for optimal control problems with state constraints*, SIAM J. Optim., 20 (2009), pp. 1002–1031.
- [15] M. ULBRICH, *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems*, MOS-SIAM Ser. Optim., SIAM, Philadelphia, 2011.
- [16] W. WOLLNER, *Adaptive Methods for PDE-based Optimal Control with Pointwise Inequality Constraints*, Ph.D. thesis, Mathematisch-Naturwissenschaftliche Gesamtfakultät, Universität Heidelberg, Heidelberg, Germany, 2010.
- [17] W. WOLLNER, *A priori error estimates for optimal control problems with constraints on the gradient of the state on nonsmooth polygonal domains*, in *Control and Optimization with PDE Constraints*, K. Bredies, C. Clason, K. Kunisch, and G. von Winckel, eds., International Series of Numerical Mathematics 164, Springer, Basel, 2013, pp. 193–215.
- [18] E. ZEIDLER, *Nonlinear Functional Analysis and its Applications*, II/B, Springer, New York, 1990.