STATIONARY OPTIMAL CONTROL PROBLEMS WITH POINTWISE STATE CONSTRAINTS

M. HINTERMÜLLER AND KARL KUNISCH

Abstract. First order optimality conditions for stationary pointwise state constrained optimal control problems are considered. It is shown that the Lagrange multiplier associated with the pointwise inequality state constraint is a regular Borel measure only, in general. In case of sufficiently smooth data and under a regularity assumption on the active set, the Lagrange multiplier can be decomposed into a regular L^2-part concentrated on the active set and a singular part, which is concentrated on the interface between the active and the inactive set. In a second part of the paper numerical solution strategies are reviewed. These methods fall into two classes: the first class which includes interior-point methods as well as active set strategies is purely finite dimensional. The second class, however, admits an analysis in function space. The latter methods typically rely on regularization. In this respect, Moreau-Yosida-based and Lavrentiev-based techniques are discussed. The paper ends by a numerical comparison of the presented solution algorithms.

Date: October 4, 2006.

1991 *Mathematics Subject Classification.* 49M15,49M37,65K05,90C33.

Key words and phrases. Active set methods, first order optimality conditions, interior point methods, Lavrentiev regularization, Moreau-Yosida regularization, path-following, semismooth Newton, pointwise state constraints, numerical methods.

The first author (M.H.) acknowledges support by the Austrian Ministry of Education, Science and Culture (bm:bwk) and the Austrian Science Fund FWF under START-grant Y305 "Interfaces and Free Boundaries".

1. Introduction

State-constrained optimal control problems involving partial differential equations (PDEs) occur in many practical applications. For instance, in the control of heat phenomena, where the underlying PDE is the well-known heat equation, one might be interested in keeping the temperature, i.e., the state of the system, within a given reference domain below a certain threshold value. If this requirement cannot be fulfilled then unwanted phase-transitions might occur. In this context, different control actions are conceivable. As an example, the control may be due to the presence of a cooling device on the boundary or, in special cases, within the domain. Or, in elasticity theory the displacement (state) within a (linear) elastic medium is determined as the solution of the Navier-Lamé equations, where the source term is a given body force. By controlling the body force or, alternatively the boundary traction, one may be interested in steering the system toward some desired state. This control-induced displacement, however, can be limited by the presence of a rigid obstacle. Hence, again one has to cope with a pointwise state constraint which, in this case, acts on the boundary (or parts thereof) of the elastic medium. Besides the two applications mentioned above there are many more instances as for instance in fluid dynamics or financial mathematics.

It was observed in, e.g., [C, CRZ] that the Lagrange multiplier associated to the pointwise inequality state-constraint exists only as a measure, in general. This fact has an impact on both, the analytical level when characterizing first order optimality of a stationary point of the optimal control problem and on the numerical level when discretizing the problem. To be more specific, let us consider the following model problem:

$$(\mathcal{P}) \qquad \begin{cases} \min\ J(y,u) = \dfrac{1}{2}\displaystyle\int_\Omega (y-z)^2\ dx + \dfrac{\alpha}{2}\int_\Omega\ u^2\ dx\ , \\ -\Delta y = u \text{ in } \Omega\ ,\ y = 0\ \text{ on } \partial\Omega\ , \\ y \le \psi \text{ a.e. in } \Omega\ , \\ (y,u) \in \mathcal{W} \times L^2(\Omega)\ , \end{cases}$$

where $\mathcal{W} = H^2(\Omega) \cap H_0^1(\Omega)$, $\Omega \subset \mathbb{R}^d$, $d \le 3$, is the underlying sufficiently smooth and bounded domain, and $\alpha > 0$. Further data regularity will be specified below. We call y the state and u the control variable, respectively. The function ψ denotes the bound constraint on the state, i.e., $y \le \psi$ has to hold pointwise almost everywhere (a.e.) in Ω. Then the first order optimality condition of $(\mathcal{P})$ involves the *complementarity system*

$$y \le \psi \text{ a.e. in } \Omega, \quad \langle \lambda, \tilde{y} - y \rangle_{\mathcal{C}^*,\mathcal{C}} \ge 0 \text{ for all } \tilde{y} \le \psi,\ \tilde{y} \in \mathcal{W}, \tag{1.1}$$

where λ represents the aforementioned Lagrange multiplier and $\mathcal{W} \subset \mathcal{C}(\Omega)$ is used. Obviously, the poor multiplier regularity does not admit a pointwise representation of the complementarity system, which is frequently crucial for numerical algorithms.

In fact, solution techniques and their (local) convergence behavior often hinge on the multiplier regularity. Classical active set methods, for instance, require a pointwise (almost everywhere) interpretation of λ for the active set estimation. Here we refer to section 5 on numerical approaches for more details on this issue. In that section we will also discuss regularization techniques which finally allow such a pointwise interpretation of the Lagrange multipliers and the development

of function space oriented active-set algorithms. Further, in the case of pointwise constraints, techniques like the projected gradient methods will not work without modification since the sum of the iteration variable and the gradient of the objective, which coincides with the negative multiplier, is needed for the update. Since they have different regularity properties this is not feasible in general. An analogous comment applies for projected Newton techniques.

Recently it was found that semismooth Newton methods are highly efficient in solving certain classes of constrained optimization problems in function space [CNQ, HIK]. These methods rely on a pointwise almost everywhere interpretation of the complementarity system involved in the first order optimality characterization and smoothing properties of the control-to-adjoint-state mapping. For instance, in the context of classes of control-constrained (rather than state-constrained) optimal control problems of the type

$$(\mathcal{P}_c) \qquad \begin{cases} \min\ J(y,u) = \dfrac{1}{2}\displaystyle\int_\Omega (y-z)^2\ dx + \dfrac{\alpha}{2}\int_\Omega\ u^2\ dx\ , \\ -\Delta y = u \text{ in } \Omega\ ,\ y = 0\ \text{ on } \partial\Omega\ , \\ u \le \psi \text{ a.e. in } \Omega\ , \\ (y,u) \in H_0^1(\Omega) \times L^2(\Omega)\ , \end{cases}$$

one has $\lambda \in L^2(\Omega)$ and (1.1) is replaced by

$$u \le \psi, \quad \lambda \ge 0, \quad \lambda(u-\psi) = 0 \quad \text{a.e. in } \Omega. \tag{1.2}$$

Above $u \in L^2(\Omega)$ denotes the optimal control variable. Then the pointwise interpretation of the complementarity system allows to express (1.2) equivalently as

$$\lambda - \max\left(0, \lambda + c(u-\psi)\right) = 0, \tag{1.3}$$

for some arbitrarily fixed $c > 0$, and the smoothing of the control-to-adjoint-state operator implies that for the choice $c = \alpha$ the mapping

$$\theta : u \mapsto \lambda(u) + c(u-\psi)$$

can be considered as $\theta : L^2(\Omega) \to L^q(\Omega)$ with $q > 2$. The norm gap between $L^q(\Omega)$ and the space $L^2(\Omega)$, in which the inequality (1.2) is posed, is crucial in proving generalized differentiability of

$$u \mapsto \max(0, \theta(u));$$

and in arguing well-definedness and locally superlinear convergence of the generalized (semismooth) Newton method for solving the underlying nonsmooth first order optimality system; see [HIK] for details. Again, the low multiplier regularity in state-constrained problems prevents the pointwise interpretation and/or the smoothing of the control-to-adjoint-state mapping. However, in our sections 3 and 5 we discuss a regularization technique which yields Lagrange multipliers in $L^2(\Omega)$ and which allows a function space analysis of semismooth Newton or, equivalently, primal-dual active-set methods.

An approach for solving state constrained optimal control problems which does not rely on the use of multipliers was introduced in [HR]. This method operates with the interface (boundary) between the active set $\{y = \psi\}$ and the inactive set $\{y < \psi\}$ as the optimization variable, and the constrained minimization problem is transformed into a shape optimization problem. Since the interface allows a unique identification of the inactive region, the multiplier itself is not an issue. While this

technique is appealing due to its favorable analytical properties, the implementation is rather technical.

In the next section we continue these notes by reviewing the first order optimality theory for a state-constrained model problem of the type ($\mathcal{P}$). We focus on the case where the governing equation is strongly elliptic. It turns out that under certain regularity assumptions on the active respectively inactive sets at the optimal solution, the Lagrange multiplier pertinent to the pointwise inequality constraint can be decomposed into a regular L^2-part and a singular measure-valued part which is concentrated on the boundary between the active and inactive sets. At the end of section 2 we briefly compare the first order conditions of ($\mathcal{P}$) and ($\mathcal{P}_c$) with respect to their analytical properties.

In the subsequent sections 3 and 4 we discuss a *Moreau-Yosida*-based regularization technique for state-constrained problems. This regularization depends on a scalar relaxation/regularization parameter which introduces a so-called *primal-dual path* which consists of the family of optimal solutions of the regularized problems together with the corresponding adjoint states and induced regular Lagrange multipliers. Further it induces a primal-dual *path value-functional*. We then study regularity and differentiability properties of the path with respect to the relaxation parameter. It turns out that under a strict complementarity assumption the value functional of the regularized problems is twice continuously differentiable and exhibits monotonicity properties. As a result good low-parametric models of the path value-functional can be found based on the structure of the relaxation term. These models are subsequently used for driving the path parameter to its limit, i.e., to find a solution of the original (less regular) problem. This procedure has several analytical as well as numerical benefits such as sufficiently regular subproblems for which standard methods (like semismooth Newton algorithms) converge rapidly in function space setting, a simple path-structure such that one can find good approximating models for the primal-dual path value functional, controlled path-parameter updates based on model functions to avoid ill-conditioning, and wide applicability.

In section 5 we introduce two solution paradigms: one based on finite dimensional methods and the other one allowing a function space analysis. In the context of purely finite dimensional approaches we discuss primal-dual path-following interior-point methods (see, e.g., [FGW, MTY, V2, Wr, Y]) and a primal-dual active-set method [BHHK, BK3]. Both methods are considered to be extremely efficient on the discrete level, but a function space theory is not available. As far as function space related techniques are concerned we address an inexact path following concept based on the Moreau-Yosida regularization addressed above. The development is due to [HK2]. We also briefly mention a technique based on a Lavrentiev-regularization [PTW, Tr] of the state-constrained problem.

Let us emphasize here that in the case of regular Lagrange multipliers (e.g., for ($\mathcal{P}_c$)) our work [BHHK, HIK] indicates that semismooth Newton and primal-dual active set methods are superior to path-following strategies. This includes a wide class of pointwise control constraints in the optimal control of partial differential equations.

Finally we point out that these notes focus on the elliptic case. However, some of the topics that we treat here were also considered for time dependent problems. Here we only mention [CRZ, HH, RK] and the references therein.

2. Problem Statement and Optimality Conditions

Unless specified otherwise these notes will focus on the problem

$$(\mathcal{P}) \qquad \begin{cases} \min\ J(y,u) = \dfrac{1}{2}\displaystyle\int_\Omega (y-z)^2\ dx + \dfrac{\alpha}{2}\int_\Omega\ u^2\ dx\ , \\ -\Delta y = u \text{ in } \Omega\ ,\ y = 0\ \text{ on } \partial\Omega\ , \\ y \le \psi \text{ a.e. in } \Omega\ , \\ (y,u) \in H_0^1(\Omega) \times L^2(\Omega)\ . \end{cases}$$

The results which we shall discuss for $(\mathcal{P})$ hold equally well for the case when the Laplacian is replaced by a strongly elliptic differential operator of second order, or when the distributed control is replaced by Neumann boundary control.

Above Ω denotes a bounded domain in $\mathbb{R}^d,\ d \in \{1,2,3\}$, with $\mathcal{C}^{1,1}$ boundary $\partial\Omega$. It is assumed throughout that $\alpha > 0,\ z \in L^2(\Omega),\ \psi \in H^2(\Omega)$, and that $\psi > 0$ on $\partial\Omega$. Note that for the control $\tilde{u} = -\Delta\psi - 1$ the corresponding state $\tilde{y} = y(\tilde{u})$ satisfies $\tilde{y} < \psi$ on $\bar{\Omega}$. This is a consequence of the maximum principle and the assumption that $\psi > 0$ on $\partial\Omega$. Hence $(\tilde{y}, \tilde{u})$ is a feasible pair for $(\mathcal{P})$.

Due to the regularity requirements on Ω every solution y to $-\Delta y = u,\ y = 0$ on $\partial\Omega$, with $u \in L^2(\Omega)$, satisfies $y \in \mathcal{W}$ where $\mathcal{W} = H^2(\Omega) \cap H_0^1(\Omega)$. We recall that $\mathcal{W} \subset \mathcal{C}(\Omega)$, the space of continuous function on $\overline{\Omega}$, if $d \le 3$. It is standard to argue the existence of a unique solution $(y^*, u^*) \in \mathcal{W} \times L^2(\Omega)$ to $(\mathcal{P})$. Together with J we also consider the reduced cost functional $\hat{J} : L^2(\Omega) \to \mathbb{R}$ defined by $\hat{J}(u) = J(y(u), u)$. A first order optimality condition is given next. We denote by $\mathcal{M}(\Omega)$ the space of real regular Borel measures on Ω and recall that $\mathcal{M}(\Omega)$ can be identified with the dual of $\mathcal{C}(\Omega)$.

Proposition 2.1. *The pair $(y^*, u^*) \in \mathcal{W} \times L^2(\Omega)$ is the solution to $(\mathcal{P})$ if and only if there exists $p^* \in L^2(\Omega)$ and $\lambda^* \in \mathcal{M}(\Omega)$ such that*

$$-\Delta y^* = u^* \quad \text{in } \Omega\ ,\ y^* = 0 \quad \text{on } \partial\Omega\ , \tag{2.1a}$$

$$(p^*, -\Delta y)_\Omega + \langle \lambda^*, y \rangle_{\mathcal{C}^*,\mathcal{C}} = (z - y^*, y)_\Omega \quad \text{for all } y \in \mathcal{W}\ , \tag{2.1b}$$

$$\langle \lambda^*, y - y^* \rangle_{\mathcal{C}^*,\mathcal{C}} \le 0 \text{ for all } y \in \mathcal{C}(\Omega),\ y \le \psi\ , \tag{2.1c}$$

$$p^* = \alpha\ u^*, \tag{2.1d}$$

$$y^* \le \psi\ . \tag{2.1e}$$

Moreover uniqueness of the pair (y^*, u^*) implies uniqueness of (p^*, λ^*).

We note that as a consequence of (2.1a) we have $y^* \in \mathcal{W}$ and hence (2.1c) is well-defined. Above $(\cdot,\cdot)_\Omega$, below also written as $(\cdot,\cdot)$, denotes the $L^2(\Omega)$-inner product on Ω and $\langle \cdot,\cdot \rangle_{\mathcal{C}^*,\mathcal{C}}$ stands for the duality pairing between $\mathcal{M}(\Omega)$ and $\mathcal{C}(\Omega)$. Proposition 2.1 can be verified with techniques developed in [BK1, C], for example. For convenience of the reader we include a proof.

Proof. We require some notation. Let $\mathcal{T} : L^2(\Omega) \to \mathcal{C}(\Omega)$ denote the operator which assigns to $u \in L^2(\Omega)$ the solution $y(u) \in \mathcal{W} \subset \mathcal{C}(\Omega)$ of $-\Delta y = u$ with Dirichlet boundary conditions. Further let

$$K = \{\ y \in \mathcal{C}(\Omega)\ \mid\ y \le \psi\ \}$$

and $I_K : \mathcal{C}(\Omega) \to \mathbb{R} \cup \{+\infty\}$ be defined as

$$I_K = \begin{cases} 0 & \text{for } y \in K \\ +\infty & \text{for } y \notin K \ . \end{cases}$$

Recall that an element $\lambda \in \mathcal{M}(\Omega)$ is in the subdifferential $\partial I_K(y^*)$ of I_K at y^* if and only if

$$y^* \in K \text{ and } \langle \lambda, y - y^* \rangle_{\mathcal{C}^*, \mathcal{C}} \leq 0 \text{ for all } y \in K \ .$$

By properties of subdifferential calculus and due to the strict convexity of $\hat{J}$ and the convexity of I_K, $(y^*, u^*) = (y(u^*), u^*)$ is a solution to $(\mathcal{P})$ if and only if

$$0 \in \partial \left(\hat{J}(u^*) + I_K(\mathcal{T} u^*) \right) \ . \tag{2.2}$$

We refer to [BP] for convex analysis results which are used in the following arguments. Since $\hat{J}$ is defined on all of $L^2(\Omega)$, the differential inclusion (2.2) is equivalent to

$$0 \in \partial \hat{J}(u^*) + \mathcal{T}^* \partial I_K(\mathcal{T} u^*) \ .$$

This latter condition is equivalent to the existence of

$$\lambda^* \in \partial I_K(\mathcal{T} u^*) \tag{2.3}$$

such that

$$0 \in \partial \hat{J}(u^*) + \mathcal{T}^* \lambda^* \ . \tag{2.4}$$

Note that (2.4) can be expressed as

$$\begin{aligned} 0 &= (\mathcal{T} u^* - z, \mathcal{T}(u - u^*)) + \alpha \, (u^* - u_d, u - u^*) + (\mathcal{T}^* \lambda^*, u - u^*) \\ &= (-p^* + \alpha(u^* - u_d), u - u^*) \end{aligned} \tag{2.5}$$

for all $u \in L^2(\Omega)$, where

$$p^* = -\mathcal{T}^* \lambda^* + \mathcal{T}^*(z - \mathcal{T} u^*) \ . \tag{2.6}$$

The equality $y^* = \mathcal{T} u^*$ and $\mathcal{T} u^* \in K$ are equivalent to (2.1a), (2.1e), further (2.3), (2.5) and (2.6) are equivalent to (2.1b),(2.1c),(2.1d) and (2.1e). This concludes the proof. □

Under appropriate assumptions on the active set associated to the constraint $y \leq \psi$ the structure of the Lagrange multiplier can further be analyzed. Understanding this structure is useful to interpret some of the difficulties that arise in numerical realization of state constrained optimal control problems, and it is also of independent interest. In particular it will follow that the Lagrange multiplier is not in $L^2(\Omega)$, in general. Concerning the regularity assumption on ψ we used $\psi \in H^2(\Omega)$ in the argument that $(\mathcal{P})$ admits at least one feasible element. In the following theorem we shall require additional regularity of ψ. But in the subsequent sections it will suffice that $\psi \in L^2(\Omega)$ and that $(\mathcal{P})$ admits a feasible element. We denote by

$$\mathcal{A} = \{ \ x \in \overline{\Omega} \mid y^*(x) = \psi(x) \ \} \text{ and } \mathcal{I} = \Omega \backslash \mathcal{A} \ ,$$

the *active* and *inactive* sets corresponding to the solution (y^*, u^*).
We shall utilize the following assumption:

$$(A1) \qquad \begin{cases} \mathcal{A} = \bigcup_{i=1}^{\ell} \mathcal{A}_i,\ \overline{\overset{\circ}{\mathcal{A}_i}} = \mathcal{A}_i, \mathcal{A} \cap \partial\Omega = \emptyset, \\ \mathcal{A}_i, i = 1, \cdots, \ell \text{ are pairwise disjoint }, \\ \mathcal{A}_i \text{ is connected with } \mathcal{C}^{1,1} \text{ boundary for each } i. \end{cases}$$

In the situation of (A1) we set $\Gamma = \partial\mathcal{A}$ and let $n_{\mathcal{I}}$ and $n_{\mathcal{A}}$ denote the outer normal vectors to $\mathcal{I}$ and $\mathcal{A}$.

Theorem 2.1. *Assume $\psi \in H^4(\Omega)$ and that (A1) holds. Then $p^* \in H_0^1(\Omega)$, $p^*_{|\mathcal{A}} \in H^2(\overset{\circ}{\mathcal{A}})$, and $p^*_{|\mathcal{I}} \in H^2(\mathcal{I})$. Moreover the pair (p^*, λ^*) of Proposition 2.1 is characterized by*

$$p^* = -\alpha\Delta\psi \quad \text{on } \mathcal{A}\ , \tag{2.7a}$$

$$-\Delta p^* = -(y^* - z) \text{ in } \mathcal{I},\ p^* = 0 \text{ on } \partial\Omega,\ p^* = -\alpha\Delta\psi \text{ on } \Gamma\ , \tag{2.7b}$$

$$\lambda^* = \mu^* + \mu^*_\Gamma \text{ with } \mu^* \in L^2(\Omega)\ ,\ \mu^*_\Gamma \in H^{1/2}(\Gamma)\ , \tag{2.7c}$$

where

$$\mu^* = \begin{cases} 0 & \text{on } \mathcal{I} \\ z - \psi - \alpha(\Delta^2\psi + \Delta u_d) & \text{on } \mathcal{A} \end{cases} \quad \text{and } \mu^*_\Gamma = -\frac{\partial p^*_{|\mathcal{I}}}{\partial n_{\mathcal{I}}} + \alpha\frac{\partial(\Delta\psi + u_d)}{\partial n_{\mathcal{A}}}$$

with $\mu^* \geq 0$ in Ω and $\mu^*_\Gamma \geq 0$ on Γ.

Proof. On $\mathcal{A} = \overline{\overset{\circ}{\mathcal{A}}}$ we have $y^* = \psi$ and hence $u^* = -\Delta\psi$. By (2.1d) it follows that $p^* = -\alpha\Delta\psi$ on $\mathcal{A}$. Thus $p^*_{|\mathcal{A}} \in H^2(\mathcal{A})$. Equation (2.1b) further implies that

$$\lambda^* = z - \psi + \Delta p^* = z - \psi - \alpha\Delta^2\psi \text{ in } \overset{\circ}{\mathcal{A}}\ . \tag{2.8}$$

From equation (2.1c) one deduces that

$$\lambda^* = 0 \text{ on } \mathcal{I} \cup \partial\Omega\ . \tag{2.9}$$

Combining (2.8) and (2.9) it follows that λ^* can be expressed as

$$\lambda^* = \mu^* + \mu^*_\Gamma,$$

with $\mu^* \in L^2(\Omega)$, given by the values of λ^* in $\overset{\circ}{\mathcal{A}}$ and $\mathcal{I}$, and $\mu^*_\Gamma \in \mathcal{M}(\Gamma)$ is a measure concentrated on Γ. Utilizing (2.1b) and the assumption that $\partial\Omega$ is $\mathcal{C}^{1,1}$ smooth, one argues that p^* is H^2 in a neighborhood of $\partial\Omega$ and in particular that

$$p^* = 0 \text{ on } \partial\Omega. \tag{2.10}$$

Recall next that $W_0^{1,q'}(\Omega) \subset \mathcal{C}(\Omega)$ for $q' > n$ and that $\mathcal{W}$ is dense in $W_0^{1,q'}(\Omega)$. The specific choice $q' = 4$ is convenient for the purpose of this proof. It then follows that the functional

$$y \to (z - y^*, y)_\Omega - \langle\lambda^*, y\rangle_{\mathcal{C}^*,\mathcal{C}} = (z - y^*, y)_\Omega - (\mu^*, y)_\Omega - \langle\mu^*_\Gamma, y\rangle_{\mathcal{C}^*(\mathcal{I}),\mathcal{C}(\mathcal{I})}$$

can be extended uniquely as a continuous linear functional from $\mathcal{W}$ to $W_0^{1,q'}(\Omega)$. From (2.1b) we conclude that $p^* \in W_0^{1,q}(\Omega)$, where q denotes the conjugate of q' (see [T], pg 46, 201). Moreover, again from (2.1b) we obtain that

$$-\Delta p^* = z - y^* \text{ in } \mathcal{I}. \tag{2.11}$$

Since $(z-y^*)_{|\mathcal{I}} \in L^2(\mathcal{I})$ we have $\Delta p^*_{|\mathcal{I}} \in L^2(\mathcal{I})$ which, together with $p^* \in W^{1,q}(\mathcal{I})$, implies that the generalized Green's formula ([T], pg 100) is applicable in the calculations below, with $\frac{\partial u}{\partial n_\mathcal{I}}$ interpreted as element of $\left(W^{\frac{1}{q},q'}(\partial\mathcal{I})\right)^*$. Let us denote by $p^*_\mathcal{I}$ and $p^*_\mathcal{A}$ the restrictions of p^* to $\mathcal{I}$ and $\overset{\circ}{\mathcal{A}}$. We now characterize the boundary condition of p^* on Γ as well as the value for μ^*_Γ. From (2.1b) we find for every $\varphi \in \mathcal{W}$,

$$(p^*, -\Delta\varphi)_\Omega = (z - y^* - \mu^*, \varphi)_\Omega - \langle \mu^*_\Gamma, y\rangle_{\mathcal{C}^*(\Gamma),\mathcal{C}(\Gamma)}. \tag{2.12}$$

Utilizing Green's formula on $\overset{\circ}{\mathcal{A}}$ and the generalized Green's formula on $\mathcal{I}$, as well as (2.12) we find

$$\begin{aligned}(p^*, -\Delta\varphi)_\Omega &= (p^*, -\Delta\varphi)_\mathcal{I} + (p^*, -\Delta\varphi)_\mathcal{A} \\ = (-\Delta p^*, \varphi)_\mathcal{I} + (-\Delta p^*, \varphi)_\mathcal{A} &+ \int_\Gamma \frac{\partial p^*_\mathcal{A}}{\partial n_\mathcal{A}}\,\varphi d\gamma + \langle \frac{\partial p^*_\mathcal{I}}{\partial n_\mathcal{I}}, \varphi\rangle_{W^{\frac{1}{q},q'}(\Gamma)^*, W^{\frac{1}{q},q'}(\Gamma)} \\ - \int_\Gamma (p^*_\mathcal{A} - p^*_\mathcal{I})\,\frac{\partial\varphi}{\partial n_\mathcal{A}}\,d\gamma &= (z - y^* - \mu^*, \varphi)_\Omega - \langle \mu^*_\Gamma, \varphi\rangle_{\mathcal{C}^*(\Gamma),\mathcal{C}(\Gamma)}\ .\end{aligned}$$

Here we utilize the facts that $H^2(\Omega) \subset H^{1,4}(\Omega)$ for $n \le 3$, $p^*_\mathcal{I}|_{\partial\mathcal{I}} \in L^{8/5}(\partial\mathcal{I})$, ([T], pg. 70), and that $\frac{\partial\varphi}{\partial n} \in L^4(\partial\Omega \cup \Gamma)$, ([T], pg 72). From the above equality we deduce that

$$p^*_\mathcal{A} = p^*_\mathcal{I} \ \text{ on } \ \Gamma. \tag{2.13}$$

Note that $\partial\mathcal{I}$ is $\mathcal{C}^{1,1}$ by (A1). Together with (2.11), (2.13), and the fact that $p^* = -\alpha\Delta\psi$ on Γ, standard regularity theory for elliptic equations implies that $p^*_\mathcal{I} \in H^2(\mathcal{I})$. Consequently $\frac{\partial p^*_\mathcal{I}}{\partial n_\mathcal{I}} \in H^{1/2}(\partial\mathcal{I})$ and referring once again to the equality above (2.13) we find

$$\mu^*_\Gamma = \frac{\partial p^*_\mathcal{A}}{\partial n_\mathcal{A}} - \frac{\partial p^*_\mathcal{I}}{\partial n_\mathcal{I}} = -\frac{\partial p^*_\mathcal{I}}{\partial n_\mathcal{I}} + \alpha\frac{\partial\,\Delta\psi}{\partial n_\mathcal{A}}, \tag{2.14}$$

in $H^{1/2}(\Gamma)$. From (2.13) it follows that $p^* \in H^1_0(\Omega)$. Finally $\mu^* \ge 0$ a.e. in Ω and $\mu^*_\Gamma \ge 0$ a. e. on Γ due to (2.1c). □

In Theorem 2.1 information on the structure of the Lagrange multiplier is obtained under the assumption that structural information on the active set is available. It would certainly be of interest to find conditions, e.g., on z and ψ such that (A1) holds. In [BK2] the structure of the Lagrange multiplier is also discussed for the case when the active set is a curve in $\Omega \subset \mathbb{R}^2$.

For the sake of comparison we next consider the case of pointwise *control* constraints. In this situation the Lagrange multiplier associated to the inequality constraint is $L^2(\Omega)$ regular and the adjoint state possesses extra regularity, which can be used for superlinear convergence of Newton's method.

We treat the bilateral control constrained problem

$$(\mathcal{P}_c) \qquad \begin{cases} \min\ J(y,u) = \dfrac{1}{2}\displaystyle\int_\Omega (y-z)^2\,dx + \frac{\alpha}{2}\int\ u^2\,dx\ , \\ -\Delta y = u \text{ in } \Omega\ ,\ y = 0\ \text{ on } \partial\Omega\ , \\ \varphi \le u \le \psi \text{ a.e. in } \Omega\ , \\ (y,u) \in H^1_0(\Omega) \times L^2(\Omega)\ , \end{cases}$$

where $z \in L^2(\Omega)$, $\alpha > 0$, and $\varphi < \psi$ with $\varphi, \psi \in L^2(\Omega)$. By well-known arguments $(\mathcal{P}_c)$ admits a unique solution $(y^*, u^*) \in \mathcal{W} \times L^2(\Omega)$. A first order optimality condition is given next.

Theorem 2.2. *The pair $(y^*, u^*) \in \mathcal{W} \times L^2(\Omega)$ is the solution to $(\mathcal{P}_c)$ if and only if there exists $p^* \in \mathcal{W}$ and $\lambda^* \in L^2(\Omega)$ such that*

$$-\Delta y^* = u^* \text{ in } \Omega, \quad y^* = 0 \text{ on } \partial\Omega \tag{2.15a}$$

$$-\Delta p^* = -(y^* - z) \text{ in } \Omega, \quad p^* = 0 \text{ on } \partial\Omega \tag{2.15b}$$

$$\alpha u^* + \lambda^* = p^* \text{ in } \Omega, \tag{2.15c}$$

$$\lambda^* \geq 0 \text{ on } \mathcal{A}_\psi = \{u^* = \psi\}, \quad \lambda^* \leq 0 \text{ on } \mathcal{A}_\varphi = \{u^* = \varphi\} \tag{2.15d}$$

$$\lambda^*(u^* - \psi)(u^* - \varphi) = 0. \tag{2.15e}$$

In (2.15d) the notation $\{u^* = \psi\}$ is an abbreviation for $\{x \in \Omega : u^*(x) = \psi(x)\}$. Note that (2.15d)-(2.15e) can equivalently be expressed as

$$\lambda^* = \max(0, \lambda^* + c(u^* - \psi)) + \min(0, \lambda^* + c(u^* - \varphi)) \tag{2.16}$$

for any $c > 0$. Here the max- and min-operations are interpreted pointwise. In the case of an unilateral constraint $u \leq \psi$ the terms involving φ in (2.15) and (2.16) are simply dropped.

Proof. Let $\mathcal{T} : L^2(\Omega) \to \mathcal{W}$ denote the control to state operator $u \to y(u)$. Then u^* is a solution to $(\mathcal{P}_c)$ if and only if for the uniformly convex reduced functional $\hat{J}$

$$(\hat{J}'(u^*), u - u^*) = (\alpha u^* + \mathcal{T}^*(\mathcal{T}u^* - z), u - u^*) \geq 0, \quad \text{for all } \varphi \leq u \leq \psi.$$

Setting $p^* = -\mathcal{T}^*(\mathcal{T}u^* - z)$, which is (2.15b), this is equivalent to

$$(\alpha u^* - p^*, u - u^*) \geq 0, \text{ for al } \varphi \leq u \leq \psi. \tag{2.17}$$

Therefore

$$\alpha u^* = p^* \text{ on } \mathcal{I},$$

where $\mathcal{I} = \Omega \setminus (\mathcal{A}_\psi \cup \mathcal{A}_\varphi)$. In particular (2.15c) holds on $\mathcal{I}$. From (2.17) we further deduce that

$$\alpha u^* - p^* \leq 0 \text{ on } \mathcal{A}_\psi,$$

and hence, setting $\lambda^* = p^* - \alpha u^* \geq 0$ on $\mathcal{A}_\psi$, we have

$$\alpha u^* + \lambda^* = p^* \text{ on } \mathcal{A}_\psi.$$

Similarly $p^* - \alpha u^* \leq 0$ on $\mathcal{A}_\varphi$ and setting $\lambda^* = p^* - \alpha u^* \leq 0$ on $\mathcal{A}_\varphi$ we have shown (2.15c) - (2.15e). □

3. Regularization and its Path

As explained in section 2 the Lagrange multiplier associated to the state constrained $y \leq \psi$ is not $L^2(\Omega)$ - regular in general, and the representation of the complementarity condition as in (2.16) is therefore not possible. We shall see in section 4 that the representation of the complementarity condition by means of a properly chosen complementarity function is of paramount importance for an efficient numerical realization of constrained optimal control problems. The choice of the max (and min) function has proven to be very efficient in computations.

We therefore introduce a family of regularized problems $(\mathcal{P}_\gamma)$ which asymptotically approximate $(\mathcal{P})$, as $\gamma \to \infty$. Due to its relation to the generalized Moreau-Yosida regularization of the indicator function of the set $\{y \leq \psi\}$ (see, e.g., [BHHK]) we call $(\mathcal{P}_\gamma)$ the *Moreau-Yosida-based* regularization of $(\mathcal{P})$. To control the size of the regularization parameter, it will be convenient to consider the path associated to $(\mathcal{P}_\gamma)$. This path consists of the solutions to $(\mathcal{P}_\gamma)$ considered as a function of γ. Regularity properties of the path will be analyzed. For $\bar\lambda \in L^2(\Omega)$ and $\gamma \in \mathbb{R}^+$ we consider

$$(\mathcal{P}_\gamma) \qquad \begin{cases} \min\ J(y,u) + \frac{1}{2\gamma}\int_\Omega |(\bar\lambda + \gamma(y-\psi))^+|^2\ dx =: J(y,u;\gamma), \\ -\Delta y = u \text{ in } \Omega\ ,\ y = 0\ \text{ on } \partial\Omega\ , \\ (y,u) \in H_0^1(\Omega) \times L^2(\Omega)\ , \end{cases}$$

where $(\cdot)^+ = \max(0,\cdot)$. The role of $\bar\lambda$ will be addressed in section 4. Here we consider $\bar\lambda$ to be fixed and focus on $(\mathcal{P}_\gamma)$ as $\gamma \to \infty$. For every $\gamma > 0$ there exists a unique solution $(y_\gamma, u_\gamma) = (y(u_\gamma), u_\gamma)$ to $(\mathcal{P}_\gamma)$ which is characterized by the unique solution to

$$(3.1) \qquad \begin{cases} -\Delta y_\gamma = u_\gamma \text{ in } \Omega\ ,\ y_\gamma = 0\ \text{ on } \partial\Omega\ , \\ -\Delta p_\gamma + \lambda_\gamma = -(y_\gamma - z) \text{ in } \Omega\ ,\ p_\gamma = 0\ \text{ on } \partial\Omega\ , \\ \alpha u_\gamma = p_\gamma \text{ in } \Omega\ , \\ \lambda_\gamma = (\bar\lambda + \gamma(y_\gamma - \psi))^+, \end{cases}$$

where $(y_\gamma, u_\gamma, p_\gamma, \lambda_\gamma) \in \mathcal{W} \times L^2(\Omega) \times \mathcal{W} \times L^2(\Omega)$. We refer to

$$\mathcal{C} = \{(\mathbf{x}_\gamma, p_\gamma, \lambda_\gamma) \in\ \mathcal{W} \times L^2(\Omega) \times L^2(\Omega) \times \mathcal{W}^* : \gamma \in (0,\infty)\}$$

as the *primal-dual path* associated to $(\mathcal{P}_\gamma)$. Here $\mathbf{x}_\gamma = (y_\gamma, u_\gamma)$. For $r \geq 0$ we further set

$$\mathcal{C}_r = \{(\mathbf{x}_\gamma, p_\gamma, \lambda_\gamma) \in\ \mathcal{W} \times L^2(\Omega) \times L^2(\Omega) \times \mathcal{W}^* : \gamma \in [r,\infty)\}.$$

Proposition 3.1. *For every $r > 0$ the path $\mathcal{C}_r$ is bounded. Moreover $(\mathbf{x}_\gamma, p_\gamma, \lambda_\gamma) \rightharpoonup (\mathbf{x}^*, p^*, \lambda^*)$ in $\mathcal{W} \times L^2(\Omega) \times L^2(\Omega) \times \mathcal{W}^*$ and $\mathbf{x}_\gamma \to \mathbf{x}^*$ in $\mathcal{W} \times L^2(\Omega)$ as $\gamma \to \infty$.*

If $\bar\lambda = 0$, then the path $\mathcal{C}_0$ is bounded in $\mathcal{W} \times L^2(\Omega) \times L^2(\Omega) \times \mathcal{W}^*$ and $\lim_{\gamma \to 0^+}(\mathbf{x}_\gamma, p_\gamma, \lambda_\gamma) = (\tilde{\mathbf{x}}, \tilde p, 0)$, where $\tilde{\mathbf{x}}$ is the solution to $(\mathcal{P})$ without inequality constraint and $\tilde p$ is the adjoint state.

The proof for this and the following proposition is given in [HK2].

Proposition 3.2. *The primal-dual path $\mathcal{C}_r$ is globally Lipschitz continuous for every $r > 0$, and $\gamma \mapsto \lambda_\gamma$ is locally Lipschitz continuous from $(0,\infty)$ to $L^2(\Omega)$. If $\bar\lambda = 0$ then $\mathcal{C}_0$ is globally Lipschitz continuous.*

Henceforth we set for $\gamma > 0$

$$S_\gamma = \{x \in \omega : \bar\lambda(x) + \gamma(y_\gamma - \psi)(x) > 0\}$$

and

$$g(\gamma) = \bar\lambda + \gamma(y_\gamma - \psi).$$

Since $\gamma \mapsto (\mathbf{x}_\gamma, p_\gamma)$ is locally Lipschitz continuous from $(0,\infty)$ to $\mathcal{W} \times L^2(\Omega) \times \mathcal{W}$ by Proposition 3.2 and (3.1),

$$\left(\frac{1}{\bar\gamma - \gamma}(\mathbf{x}_{\bar\gamma} - \mathbf{x}_\gamma), \frac{1}{\bar\gamma - \gamma}(p_{\bar\gamma} - p_\gamma)\right)$$

admits an weak accumulation point $(\dot{\mathbf{x}}_\gamma^+, \dot p_\gamma^+)$ in $\mathcal{W} \times L^2(\Omega) \times \mathcal{W}$ as $\bar\gamma \to \gamma$. Consequently

$$\frac{1}{\bar\gamma - \gamma}\left(g(\bar\gamma) - g(\gamma)\right) = \frac{\bar\gamma}{\bar\gamma - \gamma}(y_{\bar\gamma} - y_\gamma) + (y_\gamma - \psi)$$

has $\dot g(\gamma) = y_\gamma - \psi + \gamma \dot y_\gamma^+$ as strong accumulation point in $H^1(\Omega)$.

Proposition 3.3. *Let $\gamma > 0$ and set*

$$S_\gamma^+ = S_\gamma \cup \{x : \bar\lambda(x) + \gamma(y_\gamma - \psi)(x) = 0 \wedge \dot g(\gamma)(x) \geq 0\}.$$

Then $(\dot{\mathbf{x}}_\gamma, \dot p_\gamma) = (y_\gamma, u_\gamma, p_\gamma)$ satisfies

$$\text{(3.2)} \qquad \begin{cases} -\Delta \dot y_\gamma = \dot u_\gamma \text{ in } \Omega\ ,\ \dot y_\gamma = 0 \quad \text{on } \partial\Omega\ , \\ -\Delta \dot p_\gamma = -\dot y_\gamma - (y_\gamma - \psi + \gamma \dot y_\gamma)\chi_{S_\gamma^+} \text{ in } \Omega\ ,\ \dot p_\gamma = 0 \quad \text{on } \partial\Omega\ , \\ \alpha \dot u_\gamma = \dot p_\gamma \text{ in } \Omega\ . \end{cases}$$

Proof. The first and third equations follow from the corresponding equations in (3.1). To obtain the second equation we first observe that

$$\left|\frac{g(\bar\gamma)^+(x) - g(\gamma)^+(x)}{\bar\gamma - \gamma}\right| \leq \left|\frac{g(\bar\gamma)(x) - g(\gamma)(x)}{\bar\gamma - \gamma}\right|,$$

and the right hand side is uniformly bounded in $x \in \Omega$ and $\bar\gamma$ in a neighborhood of γ. Hence by Lebesgue's bounded convergence theorem

$$\lim_{\bar\gamma \to \gamma^+} \int_\Omega \left|\frac{g(\bar\gamma)^+ - g(\gamma)^+}{\bar\gamma - \gamma} - \dot g(\gamma)\chi_{S_\gamma^+}\right|^2 dx = 0.$$

Using this fact and passing to the limit in the second and fourth equations of (3.1) we obtain the second equation of (3.2). □

We now set

$$S_\gamma^0 = \{x \in \omega : \bar\lambda(x) + \gamma(y_\gamma - \psi)(x) = 0\}.$$

Corollary 3.1. *If $\operatorname{meas}(S_\gamma^0) = 0$, then $\gamma \mapsto (y_\gamma, u_\gamma, p_\gamma)$ is differentiable from $(0,\infty)$ to $\mathcal{W} \times H_0^1(\Omega) \times H_0^1(\Omega)$, and the derivative satisfies (3.2) with $\chi_{S_\gamma^+}$ replaced by $\chi_{S_\gamma^0}$.*

Proof. Let $(\delta\mathbf{x}, \delta p) = (\delta y, \delta u, \delta p)$ be the difference of two weak accumulation points in $\mathcal{W} \times L^2(\Omega) \times \mathcal{W}$ of

$$\left(\frac{1}{\bar\gamma - \gamma}(x_{\bar\gamma} - x_\gamma), \frac{1}{\bar\gamma - \gamma}(p_{\bar\gamma} - p_\gamma)\right)$$

as $\bar\gamma \to \gamma$. By (3.2) we have

$$\begin{cases} -\Delta\delta y = \frac{1}{\alpha}\delta p \ \text{ in } \Omega\ ,\ \delta y_\gamma = 0 \ \text{ on } \partial\Omega\ , \\ -\Delta\delta p = -\delta y - \gamma\ \delta y\ \chi_{S_\gamma} \ \text{ in } \Omega\ ,\ \delta p_\gamma = 0 \ \text{ on } \partial\Omega. \end{cases}$$

Taking the inner product with δp in the first and with δy in the second equation, and subtracting we find

$$0 = \frac{1}{\alpha}|\delta p|^2_{L^2(\Omega)} + |\delta y|^2_{L^2(\Omega)} + \gamma|\chi_{S_\gamma}\delta y|^2_{L^2(\Omega)},$$

which implies that $(\delta y, \delta u, \delta p) = 0$, and thus the weak accumulation point for $\bar\gamma \to \gamma$ is unique. A similar argument holds for $\bar\gamma \to \gamma^-$ and hence we need not differentiate between right and left accumulation points. To verify strong differentiability of $\gamma \to (\mathbf{x}_\gamma, p_\gamma)$ from $(0,\infty)$ to $\mathcal{W} \times H^1_0(\Omega) \times H^1_0(\Omega)$ we note that the embedding of $\mathcal{W}$ to $H^1_0(\Omega)$ is compact. Hence $\gamma \to p_\gamma$ and $\gamma \to u_\gamma$ are differentiable from $(0,\infty)$ to $H^1_0(\Omega)$ and by (3.1) $\gamma \to y_\gamma$ is differentiable from $(0,\infty)$ to $\mathcal{W}$. □

4. The Value Function and its Models

In this section we focus on the optimal value functional associated to $(\mathcal{P}_\gamma)$. We study its smoothness and monotonicity properties which will provide useful information for tuning γ in an iterative numerical procedure. We shall also describe approximations of V by low-parametric families of *model functions*.

Definition 4.1. *The functional*

$$\gamma \mapsto V(\gamma) = J(x_\gamma) + \frac{1}{2\gamma}\int_\omega |(\bar\lambda + \gamma(y_\gamma - \psi))^+|^2 dz$$

defined on $(0,\infty)$ *is called the* primal-dual-path value functional.

Proposition 4.1. *The value functional V is differentiable with*

$$\dot V(\gamma) = -\frac{1}{2\gamma^2}\int_\Omega |(\bar\lambda + \gamma(y_\gamma - \psi))^+|^2 + \frac{1}{\gamma}\int_\Omega (\bar\lambda + \gamma(y_\gamma - \psi))^+(y_\gamma - \psi).$$

The proofs of this and the following proposition are given in [HK2]. Concerning the second derivative of V we have:

Proposition 4.2. *Let $\dot y_\gamma$ denote an accumulation point of $\frac{1}{\bar\gamma - \gamma}(y_{\bar\gamma} - y_\gamma)$ as $\bar\gamma \to \gamma$. Then for any subsequence $\{\gamma_n\}$ realizing this accumulation point*

$$\begin{aligned} \lim_{\bar\gamma_n \to \gamma} \frac{1}{\bar\gamma_n - \gamma}(\dot V(\bar\gamma_n) - \dot V(\gamma)) = \frac{1}{\gamma^3}\int_\Omega |(\bar\lambda + \gamma(y_\gamma - \psi))^+|^2 - \\ \frac{2}{\gamma^2}\int_\Omega (\bar\lambda + \gamma(y_\gamma - \psi))^+(y_\gamma - \psi) + \\ \frac{1}{\gamma}\int_\Omega (y_\gamma - \psi)(y_\gamma - \psi + \gamma\dot y_\gamma)\chi_{S^+_\gamma}. \end{aligned} \tag{4.1}$$

If $\operatorname{meas}(S^\circ_\gamma) = 0$, *then $\gamma \to V(\gamma)$ is twice differentiable at γ and the second derivative is given by the right hand side in* (4.1) *with $\chi_{S^+_\gamma}$ replaced by χ_{S_γ}.*

4.1. **Case $\bar{\lambda} = 0$.** In this case

$$\dot{V}(\gamma) = \frac{1}{2}\int_\Omega |(y_\gamma - \psi)^+|^2$$

and $\gamma \to V(\gamma)$ is increasing. By Proposition 3.1, moreover, $V(0)$ equals the value of the cost in $(\mathcal{P})$ without the constraint $y \le \psi$ and $V(\infty)$ is the value of the cost in $(\mathcal{P})$.

Note that unless the solution $(\tilde{y}, \tilde{u})$ to $(\mathcal{P})$ without state constraint already satisfies $\tilde{y} \le \psi$, we have

$$\dot{V}(\gamma) = \frac{1}{2}\int_\omega |(y_\gamma - \psi)^+|^2 > 0 \quad \text{for} \quad \bar{\lambda} = 0. \tag{4.2}$$

In fact, if $\dot{V}(\gamma) = 0$ for some $\gamma > 0$, then $y_\gamma \le \psi$, i.e., y_γ is feasible. Thus, $\lambda_\gamma = 0$ and, from (3.1), we find that $(y_\gamma, u_\gamma, \lambda_\gamma) = (\tilde{y}, \tilde{u}, 0)$ is the solution to $(\mathcal{P})$ without the state constraint, which was ruled out.

In what follows we use $|\cdot|_{L^2} = |\cdot|_{L^2(\Omega)}$.

Proposition 4.3. *The mapping $\gamma \to \dot{V}(\gamma)$ is monotonically decreasing.*

Proof. For $\bar{\gamma} > \gamma > 0$ we have

$$\begin{aligned} J(y_\gamma, u_\gamma) + \frac{\gamma}{2}|(y_\gamma - \psi)^+|^2_{L^2} &\le J(y_{\bar{\gamma}}, u_{\bar{\gamma}}) + \frac{\gamma}{2}|(y_{\bar{\gamma}} - \psi)^+|^2_{L^2} \\ &\le J(y_{\bar{\gamma}}, u_{\bar{\gamma}}) + \frac{\bar{\gamma}}{2}|(y_{\bar{\gamma}} - \psi)^+|^2_{L^2} \le J(y_\gamma, u_\gamma) + \frac{\bar{\gamma}}{2}|(y_\gamma - \psi)^+|^2_{L^2} \end{aligned}$$

and hence

$$J(y_\gamma, u_\gamma) - J(y_{\bar{\gamma}}, u_{\bar{\gamma}}) \le \frac{\gamma}{2}(|(y_{\bar{\gamma}} - \psi)^+|^2_{L^2} - |(y_\gamma - \psi)^+|^2_{L^2})$$

and further

$$J(y_{\bar{\gamma}}, u_{\bar{\gamma}}) - J(y_\gamma, u_\gamma) \le \frac{\bar{\gamma}}{2}(|(y_\gamma - \psi)^+|^2_{L^2} - |(y_{\bar{\gamma}} - \psi)^+|^2_{L^2})$$

which implies

$$0 \le \frac{\bar{\gamma} - \gamma}{2}(|(y_\gamma - \psi)^+|^2_{L^2} - |(y_{\bar{\gamma}} - \psi)^+|^2_{L^2}) = (\bar{\gamma} - \gamma)(\dot{V}(\gamma) - \dot{V}(\bar{\gamma})).$$

□

Note that Proposition 4.3 implies that $\ddot{V}(\gamma) \le 0$ whenever the second derivative of V exists at γ. A class of functions which satisfies the above properties of V is given by

$$m(\gamma) = C_1 - \frac{C_2}{(D + \gamma)^r}, \tag{4.3}$$

with $C_1 \in \mathbb{R}$, $C_2 > 0$, $D > 0$, $r > 0$. In fact, $\dot{m} > 0$, $\ddot{m} < 0$, and $m(0)$, $m(\gamma)$ for $\gamma \to \infty$ are well defined. In our use of m for path following algorithms, (C_1, C_2, D) will be treated differently from r. While r will be chosen as a fixed number, (C_1, C_2, D) will be updated in an iterative procedure.

4.2. **Case $\bar\lambda \geq 0$.** In case there exist $\bar\lambda \in L^2(\Omega)$ and $\gamma > 0$ such that

$$\bar\lambda \geq 0 \quad \text{and} \quad y_\gamma \leq \psi \tag{4.4}$$

we have $\dot V(\gamma) \leq 0$ and, unless the solution $(\tilde y, \tilde u)$ to $(\mathcal{P})$ without pointwise inequality constraint satisfies $\tilde y \leq \psi$, we have $\dot V(\gamma) < 0$. In fact, if $\dot V(\gamma) = 0$, then

$$0 = \dot V(\gamma) \leq -\frac{1}{2\gamma^2} \int_\omega |(\bar\lambda + \gamma(y_\gamma - \psi))^+|^2 \leq 0,$$

and hence $\lambda_\gamma = 0$. Therefore $(y_\gamma, u_\gamma, \lambda_\gamma) = (\tilde y, \tilde u, 0)$ is the solution of (2.1), and hence the solution $(\tilde y, \tilde u)$ to $(\mathcal{P})$ without pointwise inequality constraint, which is excluded.

Let us also observe that $\lim_{\gamma\to\infty} V(\gamma)$ equal to the value of the objective of $(\mathcal{P})$, and $\lim_{\gamma\to 0} V(\gamma) = \infty$.

Proposition 4.4. *Assume that* $\operatorname{meas}(S_\gamma^\circ) = 0$ *and that* (4.4) *holds. Then* $\ddot V(\gamma) \geq 0$ *for* $\gamma \geq 0$.

Proof. From (4.1) with $S_\gamma^+ = S_\gamma$ we have

$$\ddot V(\gamma) \geq \frac{1}{\gamma} \int_{S_\gamma} (y_\gamma - \psi)^2 + \int_{S_\gamma} (y_\gamma - \psi)\dot y_\gamma.$$

Using (3.2) we deduce that

$$\alpha|\dot u_\gamma|^2 + |\dot y_\gamma|^2 = -(y_\gamma - \psi + \gamma\dot y_\gamma, \chi_{S_\gamma}\dot y_\gamma),$$

which implies that

$$\gamma|\dot y_\gamma|_{L^2(S_\gamma)} \leq |y_\gamma - \psi|_{L^2(S_\gamma)}.$$

It follows that $\ddot V(\gamma) \geq 0$.

□

A class of model functions, which satisfy the above properties of V is given by

$$m(\gamma) = C_1 - \frac{C_2}{(D+\gamma)^r} + \frac{B}{\gamma^r}, \tag{4.5}$$

with $C_1 \in \mathbb{R},\ B \geq C_2 > 0,\ D > 0$ and $r \in (0, 1]$. In fact, $m(0) = \infty$, $\lim_{\gamma\to\infty} m(\gamma) = C_1$, $\dot m(\gamma) < 0$ and $\ddot m(\gamma) > 0$.

Remark 4.1. From (3.1) we find

$$\begin{aligned} |\nabla(y_\gamma - \psi)^+|_{L^2}^2 &= (\tfrac{1}{\alpha}p_\gamma - \psi, (y_\gamma - \psi)^+) \\ &= -\tfrac{1}{\alpha}(\lambda_\gamma + y_\gamma - z, (-\Delta)^{-1}(y_\gamma - \psi)^+) - (\psi, (y_\gamma - \psi)^+) \leq 0 \end{aligned}$$

provided that $\psi \geq 0$ and $(\bar\lambda + \gamma(y_\gamma - \psi))^+ + y_\gamma - z \geq 0$. In this case $y_\gamma \leq \psi$ and (4.4) holds.

5. Numerical solution techniques

We continue our survey by addressing available solution strategies for state constrained optimal control problems. In what follows we focus on a review of numerical algorithms for the efficient solution of the underlying problem class. Of course, the discretization of the problems and its convergence analysis (including the proof of optimal rates with respect to the mesh-size of dicretization) are further important issues in numerical analysis and, in many cases, are the subject of ongoing research

efforts. A comprehensive discussion of the latter issues is beyond the scope of this review.

Concerning the development of solution algorithms, roughly two approaches are conceivable. First, one may consider discretizing the problem and then applying suitable algorithms from nonlinear programming. Conveniently, one can rely on the available finite dimensional convergence analysis of the chosen method. However, in many cases upon refining the mesh-size of discretization, one will experience an increase of iterations until the successful termination of the algorithm. This *mesh dependence* typically indicates a lack of function space properties of the selected solution method. Therefore, in contrast to the first approach, one is interested in developing algorithms in infinite dimensions and study their convergence behavior analytically. In the case of a successful analysis in function space, one can hope for a *mesh independent* convergence behavior of the algorithm.

5.1. **Finite dimensional methods.** We start by briefly reviewing some of the finite dimensional approaches proposed in the literature and highlight their properties. These methods are considered to be efficient (i.e., locally at least superlinearly convergent) for solving the discretized problem on a fixed mesh.

For this purpose we introduce the discretized version of $(\mathcal{P})$. We proceed in general terms with finite differences and finite element realizations in mind. Assume first that Ω is endowed with a uniform grid Ω_h with mesh-size h. Next let z_h, y_h, u_h, $\psi_h \in \mathbb{R}^N$ be finite-dimensional approximations to z, y, u, and ψ, respectively. Further let $L_h \in \mathbb{R}^{N\times N}$ stand for a symmetric positive-definite approximation to $-\Delta$ (including the homogenous Dirichlet boundary conditions). Then the discretized control constrained problem is given by

$$(P_h) \qquad \begin{cases} \text{minimize} & J_h(y_h, u_h) \\ \text{subject to} & L_h y_h = u_h,\ y_h \le \psi_h\,, \end{cases}$$

where

$$J_h(y_h, u_h) = \frac{1}{2}(y_h - z_h)^\top M_{1h}(y_h - z_h) + \frac{\alpha}{2} u_h^\top M_{2h} u_h \tag{5.1}$$

and $M_{1h} \in \mathbb{R}^{N\times N}$ is positive-semidefinite and $M_{2h} \in \mathbb{R}^{N\times N}$ is positive-definite. The matrices M_{1h} and M_{2h} result from the numerical integration of the cost functional J. For finite difference approximation with integration based on the trapezoidal rule, M_{1h} and M_{2h} are positive-definite diagonal matrices.

Concerning the matrices L_h, M_{1h}, and M_{2h} we henceforth utilize the following assumption:

$$(A) \qquad \begin{cases} L_h \text{ is sparse banded, symmetric, and positive-definite,} \\ M_{1h} \text{ is sparse banded and positive-semidefinite,} \\ M_{2h} \text{ is sparse banded and positive-definite}\,. \end{cases}$$

In a 2D-setting, for instance, we remark that the discretization of $-\Delta$ by the well-known five-point star scheme implies that L_h is a symmetric positive-definite banded M-matrix. Applying the trapezoidal rule for approximating J implies that M_{1h} and M_{2h} are positive-definite and diagonal. In what follows, for convenience we omit h for problem variables and keep it for problem data throughout this section.

5.1.1. *Primal-dual active set strategy.* The first method which we outline here is, as far as state constraints are concerned, a finite dimensional approach. It is based on the following first order necessary and sufficient condition for (P_h):

$$L_h y = u, \tag{5.2a}$$
$$L_h p + \lambda = M_{1h}(z_h - y), \tag{5.2b}$$
$$\alpha M_{2h} u = p, \tag{5.2c}$$
$$\lambda = \max(0, \lambda + c(y - \psi_h)), \tag{5.2d}$$

where $c > 0$ is an arbitrarily fixed real and the max-operation is understood componentwise. Note that (5.2d) is equivalent to the *complementarity system*

$$\lambda \geq 0, \quad y \leq \psi_h, \quad \lambda^\top (y - \psi_h) = 0. \tag{5.3}$$

In order to derive the method, let $(y^*, u^*, p^*, \lambda^*)$ denote the solution of (5.2). Further we define the *active* and *inactive* sets at the optimal solution as

$$A^* = \{i \in \{1, \dots, N\} : y_i^* = \psi_{h,i}\}, \tag{5.4a}$$
$$I^* = \{i \in \{1, \dots, N\} : y_i^* < \psi_{h,i}\}. \tag{5.4b}$$

Next we study the expression in the max-operation in (5.2d). We have

$$\lambda_i^* + c(y_i^* - \psi_{h,i}) \geq 0 \quad \text{on } A^*, \tag{5.5a}$$
$$\lambda_i^* + c(y_i^* - \psi_{h,i}) < 0 \quad \text{on } I^*. \tag{5.5b}$$

Moreover, if *strict complementarity* holds true at the optimal solution, then

$$\lambda^* + (\psi_h - y^*) > 0.$$

This yields

$$\lambda_i^* + c(y_i^* - \psi_{h,i}) > 0 \quad \text{on } A^* \tag{5.6}$$

instead of (5.5a). Now, given an estimate (y_{n-1}, λ_{n-1}) of (y^*, λ^*) we use (5.5b) and (5.6) as a prediction strategy of the correct active and inactive sets at the solution, i.e., we define the corresponding set estimates by

$$A_n = \{i \in \{1, \dots, N\} : \lambda_{n-1,i} + c(y_{n-1,i} - \psi_{h,i}) > 0\}, \tag{5.7a}$$
$$I_n = \{i \in \{1, \dots, N\} : \lambda_{n-1,i} + c(y_{n-1,i} - \psi_{h,i}) \leq 0\}. \tag{5.7b}$$

This yields the following algorithm.

Primal-dual active-set strategy (PDAS).

(1) Initialization: choose $y_0, \lambda_0, u_0 \in \mathbb{R}^N$, $c > 0$, and set $n = 1$.
(2) Determine the subset of active/inactive indices according to (5.7).
(3) If $n \geq 2$ and $A_n = A_{n-1}$, then STOP; otherwise go to step 4.
(4) Find $(y_n, u_n, p_n, \lambda_n)$ as the solution to

$$\begin{aligned} &L_h y_n = u_n, \\ &L_h p_n + \lambda_n + M_{1h}(y_n - z_{dh}) = 0, \\ &p_n = \alpha M_{2h} u_n, \\ &y_{n,i} = \psi_{h,i} \text{ for } i \in A_n, \ \lambda_{n,i} = 0 \text{ for } i \in I_n. \end{aligned}$$

(5) Set $n = n + 1$ and go to step 2.

We summarize some properties of PDAS:

- The iterates can be infeasible (both primal and dual variables).
- The algorithm does not rely on a globalization strategy.

- For $n > 1$ the algorithm is independent of c.
- If the algorithm stops in step 3, then the exact solution of the discretized problem is obtained.
- Utilizing $y_i = \psi_{h,i}$ for $i \in A_n$ the primal system in step 4 need only be solved for $i \in I_n$.
- If $(\bar{y}_h, \bar{u}_h) = \operatorname{argmin}\{J_h(y_h, u_h) : A_h y_h = u_h\}$ satisfies $\bar{y}_h \leq \psi_h$ and $A_0 = \emptyset$, then PDAS computes the solution in one step.
- The algorithm has the property that from one iteration to the next many coordinates of the discretized control or state vector can move from A_{n-1} to I_n and vice versa. Numerically it turns out that changes from active to inactive sets occur primarily along the boundary between active and inactive sets. This is due to the fact that λ for PDAS is the discretization of the measure $\lambda^* \in \mathcal{M}(\Omega)$ whose singular part is concentrated at the boundary of the active set [BK2]. For control constrained problems the algorithm behaves even more efficiently: from one iteration to the next it makes changes form active to inactive not only near the boundary but also in the interior of the currently active or inactive sets, typically resulting in convergence within a few number of – mesh independent – steps; see [HU].
- The iterates y_n are mostly feasible and the active sets A_n typically approximate A^* from outside. This approximation is typically monotone with respect to the cardinality of the active set but non-monotone in the set-wise sense.
- It turns out that PDAS is equivalent to a semismooth Newton method in $\mathbb{R}^N$; see the next section for a related issue in function space. Thus, applying the semismooth Newton theory in $\mathbb{R}^N$ developed in, e.g., [HIK] one can prove a locally superlinear convergence of PDAS. A conditional global convergence result can be found in [BK3].
- For control constrained optimal control problems (e.g., of the type $(\mathcal{P}_c)$ the analogue of PDAS is again equivalent to a semismooth Newton method. Moreover, it is shown in [HIK] that its function space version converges locally at a superlinear rate.

5.1.2. *Primal-dual path-following interior-point methods.* The second method which we outline here is a primal-dual path-following interior-point method based on a predictor-corrector strategy; see [BG, BPR, LMS, Meh, MTY, V2, Wr, VY, Y, Z2] and the references therein for details and related approaches. Due to the specific structure of optimal control problems, our linear algebra takes care of the fact that the vector of unknowns can be decomposed into the state variable y_h and the control variable u_h. The interior-point algorithm described here is a (large neighborhood) modification of the algorithm in [MTY]. See also [BPR], where the corresponding convergence analysis can be found. It will be exemplarily presented for state constrained problems. However, adaptation to control constraints is straightforward. An alternative, highly efficient predictor-corrector interior-point technique can be found, e.g., in [LMS, Meh]. It is tailored to optimal control problems in [BHHK].

As there is no function space version of primal-dual path-following interior-point methods for state constraints available up-to-date, our starting point is the discretized, finite-dimensional version (P_h) of $(\mathcal{P})$. Introducing a vector of slack variables denoted by $w \in \mathbb{R}^N$ in the inequality constraint in (P_h), we obtained the

following modification:

$$(P_h^{\mathrm{mod}}) \qquad \begin{cases} \text{minimize} & J_h(y,u) \\ \text{subject to} & B_h y - h^2 u = 0, \\ & y + w = \psi_h, \\ & w \geq 0\,, \end{cases}$$

with $B_h = h^2 L_h$. Standard arguments yield the existence of an optimal solution of (P_h^{mod}), which is characterized by the following system equivalent to the first-order Karush–Kuhn–Tucker conditions:

$$B_h w + h^2 u - B_h \psi_h = 0, \tag{5.8a}$$

$$w \geq 0, \tag{5.8b}$$

$$M_{1h} w + B_h p - \lambda - M_{1h}(\psi_h - z_h) = 0, \tag{5.8c}$$

$$\alpha M_{2h} u + h^2 p = 0, \tag{5.8d}$$

$$\lambda \geq 0, \tag{5.8e}$$

$$\Lambda W e = 0\,, \tag{5.8f}$$

where we use standard interior-point notation when writing Λ for the diagonal $N \times N$-matrix $\mathrm{diag}(\lambda_1, \ldots, \lambda_N)$ and similarly for W. Further, e denotes the vector of all ones in $\mathbb{R}^N$.

The defining equations for a point on the primal-dual interior-point central path are obtained by replacing the last equation (5.8f) in (5.8) by

$$\Lambda W e = \mu e, \tag{5.9}$$

with μ being some positive scalar parameter. By similar arguments as before it can be seen that the resulting $(4N \times 4N)$-system admits a unique solution. Now, the primal-dual central-path is the set

$$\{(w,u,p,\lambda) \in \mathbb{R}^{4N} : (5.8\mathrm{a}) - (5.8\mathrm{e}) \text{ and } (5.9) \text{ are satisfied for } \mu \geq 0\}.$$

Let us suppose that we have decided on a target value for μ, that $(w,u,p,\lambda) \in \mathbb{R}^{4N}$ satisfies $w, \lambda > 0$, and that $(w + \Delta w, \ldots, \lambda + \Delta\lambda)$ denotes the point on the primal-dual central path corresponding to μ. Thus, we obtain the following system for the increments $(\delta w, \ldots, \delta\lambda)$ (see [BHHK] for details):

$$(S(w,\lambda)) \qquad \begin{bmatrix} \alpha M_{2h} & & h^2 I & \\ & M_{1h} & B_h & -I \\ h^2 I & B_h & & \\ & \Lambda & & W \end{bmatrix} \begin{bmatrix} \delta u \\ \delta w \\ \delta p \\ \delta\lambda \end{bmatrix} = \begin{bmatrix} \beta_2 \\ \beta_1 \\ \alpha_1 \\ \gamma_1 \end{bmatrix},$$

with

$$\alpha_1 = B_h(\psi_h - w) - h^2 u, \quad \beta_1 = M_{1h}(\psi_h - z_h - w) - B_h p + \lambda,$$
$$\beta_2 = \alpha M_{2h} u - h^2 p, \quad \gamma_1 = \mu e - W\Lambda e - \delta W\, \delta\Lambda e\,.$$

In path-following methods it is typically required that the iterates stay within a neighborhood of the central path. For this purpose, we introduce the (large) neighborhood of the central path:

$$\mathcal{N}(\nu) = \{(w,\lambda,\mu) \in \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}_+ | \nu\mu e \leq \Lambda W e \leq \nu^{-1}\mu e\}\,,$$

where $0 < \nu < 1$ is a given constant. Typical values for ν are $\nu = 0.01$ and $\nu = 0.001$. In our experiments [BHHK] these values give essentially the same results. It is well-known that small neighborhoods usually give better complexity

estimates when compared to large neighborhood path-following methods. On the other hand, the large neighborhood variants typically outperform the small neighborhood methods on the numerical level.

Next we describe the primal-dual path-following predictor-corrector interior-point algorithm (Mizuno-Todd-Ye-type). The subscript n denotes the iteration level.

PRIMAL-DUAL PATH-FOLLOWING INTERIOR-POINT METHOD (PDIP).

(1) *Initialization*: Choose $0 < \nu < 1$ and $(w_o, \lambda_o, \mu_o) > 0$, with $(w_o, \lambda_o, \mu_o) \in \mathcal{N}(\nu)$, a stopping tolerance $\varepsilon_s > 0$ and $e_f \in \mathbb{N}$. Set $n := 0$.

(2) *Check stopping criteria*: If $\max(\mathrm{res}^p, \mathrm{res}^d) \leq \varepsilon_s$, with

$$\mathrm{res}^p = h\|\alpha_{1,n}\|_2 \quad \text{and} \quad \mathrm{res}^d = h\|(\beta_{1,n}, \beta_{2,n})^\top\|_2,$$

and $f_n = \max\{-\log_{10}[(J_n - \phi_n)/(1+|J_n|)], 0\} \geq e_f$, then STOP; otherwise go to step 3.

(3) *Corrector step* $\delta u_c, \delta w_c, \delta p_c, \delta\lambda_c$: Solve $S(w_n, \lambda_n)$ with right-hand side $(0, 0, 0, \mu_n e - W_n\Lambda_n e)^\top$. Compute $\tau_c \in (0, 1]$ such that

$$((w_n, \lambda_n) + \tau_c(\delta w_c, \delta\lambda_c), \mu_n) \in \mathcal{N}(\nu)\,,$$

and put

$$(u_{n+\frac{1}{2}}, w_{n+\frac{1}{2}}, p_{n+\frac{1}{2}}, \lambda_{n+\frac{1}{2}}) = (u_n, w_n, p_n, \lambda_n) + \tau_c(\delta u_c, \delta w_c, \delta p_c, \delta\lambda_c)\,.$$

(4) *Predictor step* $\delta u_a, \delta w_a, \delta p_a, \delta\lambda_a$: Solve $S(w_{n+\frac{1}{2}}, \lambda_{n+\frac{1}{2}})$ with right-hand side $(\beta_{2,n}, \beta_{1,n}, \alpha_{1,n}, -\delta W_c\delta\Lambda_c e)^\top$. Compute τ_a, the largest value in $(0, 1)$ such that $((w_{n+\frac{1}{2}}, \lambda_{n+\frac{1}{2}}) + \tau_a(\delta w_a, \delta\lambda_a), (1-\tau_a)\mu_n) \in \mathcal{N}(\nu)$. Put $\mu_{n+1} = (1-\tau_a)\mu_n$ and

$$\begin{aligned}(u_{n+1}, w_{n+1}, p_{n+1}, \lambda_{n+1}) \quad = \quad & (u_{n+\frac{1}{2}}, w_{n+\frac{1}{2}}, p_{n+\frac{1}{2}}, \lambda_{n+\frac{1}{2}}) \\ & +\tau_a(\delta u_a, \delta w_a, \delta p_a, \delta\lambda_a)\,.\end{aligned}$$

Set $n = n + 1$, and go to step 2.

Let us briefly comment on the stopping criteria in step 2. The first two criteria involving res^p and res^d check for smallness of the residuals of the state and the adjoint equation, respectively. The third criterion $f_n \geq e_f$ checks the number of digits of coincidence between the primal objective value $J_n = J_h(y_n, u_n)$ and the dual objective value

$$\phi_n = -J_n - p_n^\top B_h z_h - \lambda_n^\top(z_h - \psi_h).$$

Note that by standard duality theory the difference between the optimal primal and dual objective values vanishes. The systems in step 3 and 4 are reduced in advance by choosing specific pivots. This process results in

$$(\Lambda + WM_{1h} + \alpha WL_hM_{2h}L_h)\Delta w = r^s\,, \tag{5.10}$$

where $r^s \in \mathbb{R}^N$ denotes the appropriate right-hand side. Note that this system could easily be symmetrized. Finally, due to (A) the key assumptions of [BPR] for proving convergence and complexity results for Algorithm PDIP are satisfied.

Considering the continuous version of PDIP we note that μ acts as a regularization parameter, if y is bounded away from the constraint ψ. There is no analog for that in PDAS, see, however, algorithm PDAS_γ in the following subsection.

5.2. **Function space approaches.** Next we discuss recent methods based on appropriate regularization of the measure-valued Lagrange multiplier for state constrained problems. The techniques take into account the function space properties of the underlying problems and therefore admit a function space analysis. In fact, these methods respect the low Lagrange multiplier and adjoint regularity, and are based on using a family of approximating problems whose Lagrange multipliers and adjoints are more regular.

5.2.1. *Inexact feasible and non-interior path-following.* We start our discussion with a recent path-following method which is based on the regularized state constrained problem $(\mathcal{P}_\gamma)$; see [HK2]. For convenience we recall the regularized problem:

$$(\mathcal{P}_\gamma) \qquad \begin{cases} \min\ J(y,u) + \frac{1}{2\gamma}\int_\Omega |(\bar\lambda + \gamma(y-\psi))^+|^2\ dx =: J(y,u;\gamma), \\ -\Delta y = u \text{ in } \Omega\ ,\ y = 0\ \text{ on } \partial\Omega\ , \\ (y,u) \in H_0^1(\Omega)\times L^2(\Omega)\ , \end{cases}$$

where $J(y,u) = \|y-z\|^2_{L^2(\Omega)} + (\alpha/2)\|u\|^2_{L^2(\Omega)}$. The algorithm is composed of an outer iteration for adapting the regularization/penalization parameter γ and an inner iteration which (approximately) solves $(\mathcal{P}_\gamma)$ for fixed γ. To be more specific, its inner loop utilizes a locally superlinearly convergent (in function space) algorithm for solving the regularized path problem. The outer loop employs a γ-update strategy based on our model functions (4.3), respectively (4.5).

The active set detection in the inner loop now serves the purpose of a decomposition of the argument of the max-operator in the objective functional of $(\mathcal{P}_\gamma)$ into a positive and a non-negative part, i.e., given some estimate y^{l-1} of y_γ we set

$$A^l := \{x\in\Omega : \bar\lambda(x) + \gamma(y^{l-1}(x) - \psi(x)) > 0\}, \tag{5.11a}$$

$$I^l := \Omega \setminus A^l. \tag{5.11b}$$

The role of $\bar\lambda \geq 0$ in (5.11) is as in section 4. Here, in addition, we require that $\bar\lambda \in L^p(\Omega)$ with $p > 2$ sufficient large as it will be specified below. In order to motivate the choice (5.11), now we establish a connection between (5.11) and generalizations of Newton's method. For this purpose observe first that $y \mapsto \max(0, \bar\lambda - \gamma(y-\psi))$ is nondifferentiable on $I_0 := \{x\in\Omega : \bar\lambda(x) + \gamma(y(x)-\psi(x)) = 0\}$. With the aim of defining a generalized derivative of $F_{\max} : L^p(\Omega) \to L^2(\Omega)$, with $p > 2$ and $F_{\max}(y) = \max(0,y)$, we define the mapping $G_{\max} \in \mathcal{L}(L^p(\Omega), L^\infty(\Omega))$ by

$$G_{\max}(\omega)(x) := \begin{cases} 0 & \text{if } \omega(x) \leq 0, \\ 1 & \text{if } \omega(x) > 0. \end{cases} \tag{5.12}$$

In [HIK] it was shown that $\max(0,\cdot) : L^p(\Omega) \to L^q(\Omega)$ is *Newton* differentiable if and only if $1 \leq q < p \leq +\infty$, i.e., it satisfies the following *semismoothness* property:

$$\|\max(0, y+s) - \max(0,y) - G_{\max}(y+s)s\|_{L^q} = o(\|s\|_{L^p}) \tag{5.13}$$

for $\|s\|_{L^p} \to 0$. Hence, $F_{\max}$, as defined above, is Newton differentiable in the sense of (5.13), and the operator $G_{\max}$ serves as a particular generalized derivative. In appendix A we recall the general definition of Newton differentiability.

Next we apply this calculus to the first order necessary and sufficient conditions of $(\mathcal{P}_\gamma)$:

$$-\Delta y = u, \tag{5.14a}$$

$$-\Delta p + (\bar\lambda + \gamma(y-\psi))^+ = -J_y(y,u), \tag{5.14b}$$

$$\alpha u - p = 0. \tag{5.14c}$$

Due to (5.14a) we have $y = y(u) = (-\Delta)^{-1}u$, where $(-\Delta)^{-1}$ denotes the solution operator of $-\Delta y = u$ on Ω, $y \in H_0^1(\Omega)$, for given $u \in L^2(\Omega)$. Hence, (5.14) can be condensed into

$$-\alpha\Delta u + (\bar\lambda + \gamma((-\Delta)^{-1}u - \psi))^+ + (-\Delta)^{-1}u = z.$$

Thus, solving (5.14) is equivalent to finding a root of $F : L^2(\Omega) \to L^2(\Omega)$,

$$F(v) = \alpha v + (-\Delta)^{-2}v + (\bar\lambda + \gamma((-\Delta)^{-2}v - \psi))^+ - z. \tag{5.15}$$

In order to see the connection between (5.11) and a generalized (or semismooth) Newton step based on (5.13) for $F(v) = 0$, let $v^{l-1} \in L^2(\Omega)$ denote the current iterate. Then the generalized derivative of F at v^{l-1} is given by

$$G(v^{l-1})s = \alpha s + (-\Delta)^{-2}s + \gamma\chi_{A(v^{l-1})}(-\Delta)^{-2}s,$$

with

$$A(v^{l-1}) = \{x \in \Omega : (\bar\lambda + \gamma((-\Delta)^{-2}v^{l-1} - \psi))(x) > 0\},$$

for $s \in L^2(\Omega)$. For a short proof we only focus on the max-term in F since the rest is linear. We set $F_m(v) := (\bar\lambda + \gamma((-\Delta)^{-2}v - \psi))^+$. Then

$$\begin{aligned}
&\frac{1}{\|s\|_{L^2}}\|F_m(v^{l-1}+s) - F_m(v^{l-1}) - \gamma\chi_{A(v^{l-1}+s)}(-\Delta)^{-2}s\|_{L^2} \\
&\le \frac{C}{\|(-\Delta)^{-2}s\|_{H_0^1}}\|F_{\max}(\bar\lambda + \gamma((-\Delta)^{-2}(v^{l-1}+s) - \psi)) \\
&\qquad - F_{\max}(\bar\lambda + \gamma((-\Delta)^{-2}v^{l-1} - \psi)) \\
&\qquad - \gamma G_{\max}(\bar\lambda + \gamma((-\Delta)^{-2}(v^{l-1}+s) - \psi))(-\Delta)^{-2}s\|_{L^2} \\
&= \frac{C}{\|\hat s\|_{H_0^1}}\|F_{\max}(\bar\lambda + \gamma(\hat v^{l-1} + \hat s - \psi)) - F_{\max}(\bar\lambda + \gamma(\hat v^{l-1} - \psi)) \\
&\qquad - \gamma G_{\max}(\bar\lambda + \gamma(\hat v^{l-1} + \hat s - \psi))\hat s\|_{L^2}\hat s \\
&\le \frac{C}{\|\hat s\|_{L^p}}\|F_{\max}(\bar\lambda + \gamma(\hat v^{l-1} + \hat s - \psi)) - F_{\max}(\bar\lambda + \gamma(\hat v^{l-1} - \psi)) \\
&\qquad - \gamma G_{\max}(\bar\lambda + \gamma(\hat v^{l-1} + \hat s - \psi))\hat s\|_{L^2} \\
&= o(\|\hat s\|_{L^p}),
\end{aligned}$$

where $C > 0$ is a constant independent of $\hat s$, $\hat v^{l-1} = (-\Delta)^{-2}v^{l-1} \in H_0^1(\Omega)$, $\hat s = (-\Delta)^{-2}s \in H_0^1(\Omega)$, $H_0^1(\Omega) \subset L^p(\Omega)$ with $p = \infty$ for $n \le 2$ and $p = \frac{2n}{n-2} > 2$ for $n > 2$, (5.13), the chain rule and the fact that $\bar\lambda - \gamma\psi$ is fixed in $L^p(\Omega)$. Note that $G : L^2(\Omega) \to L^2(\Omega)$ satisfies

$$\langle G(u)s, s\rangle \ge \alpha\|s\|_{L^2}^2 \quad \text{for all } s \in L^2(\Omega).$$

Hence, $G(\cdot)^{-1} \in \mathcal{L}(L^2(\Omega))$ exists and is uniformly bounded.

Given v^{l-1}, the next iterate of Newton's method is defined as the root of $\ell(\cdot; v^{l-1}) := F(v^{l-1}) + G(v^{l-1})(\cdot - v^{l-1})$, i.e.,

$$\ell(v^l; v^{l-1}) = 0.$$

This yields

$$\alpha v^l + (-\Delta)^{-2} v^l + \chi_{A(v^{l-1})}(\bar\lambda + \gamma((-\Delta)^{-2} v^l - \psi)) - z = 0.$$

Setting $u^l := (-\Delta)^{-1} v^l$, $p^l = \alpha u^l$ and $y^l := (-\Delta)^{-1} u^l$ we obtain $A(v^l) = A^{l+1}$, with the latter set according to (5.11a). We therefore conclude that the active set strategy based on the selection rule (5.11) is equivalent to a semismooth Newton for solving $F(v) = 0$ where the operator F is given by (5.15).

Inner iteration: An algorithm for solving $(\mathcal{P}_\gamma)$. Here we adopt the primal-dual active set strategy as proposed in section 5.1.1. As it turned out in the previous section, the method is equivalent to a semismooth Newton algorithm, and, using the techniques in [HIK] (see appendix A for details), it can be shown that it converges locally at a q-superlinear rate in function space.

Primal-Dual Active-set Strategy for $(\mathcal{P}_\gamma)$ – (PDAS_γ).

(i) Choose $\bar\lambda \geq 0$ and $\mathbf{x}^0 := (y^0, u^0) \in \mathcal{W} \times L^2(\Omega) =: X$; set $l = 0$.

(ii) Determine the active and inactive sets

$$\begin{aligned} A^{l+1} &:= \{x \in \Omega : \bar\lambda(x) + \gamma(y^l(x) - \psi(x)) > 0\}, \\ I^{l+1} &:= \Omega \setminus A^{l+1}. \end{aligned}$$

(iii) Compute the solution $\mathbf{x}^{l+1} := (y^{l+1}, u^{l+1})$ with associated adjoint state p^{l+1} of

$$\begin{aligned} \text{minimize } & \langle J'(\mathbf{x}^l) + \frac{1}{2}\langle J''(\mathbf{x}^l)(\mathbf{x} - \mathbf{x}^l), \mathbf{x} - \mathbf{x}^l\rangle_{X^*,X} \\ & \qquad + \frac{1}{2\gamma}\|(\bar\lambda + \gamma(y - \psi))^+\|^2_{L^2(A^{l+1})} \quad \text{over } \mathbf{x} \in X \\ \text{subject to } & -\Delta y = u \text{ in } L^2(\Omega). \end{aligned}$$

(iv) Compute

$$\lambda^{l+1} = \begin{cases} 0 & \text{on } I^{l+1}, \\ \bar\lambda + \gamma(y^{l+1} - \psi) & \text{on } A^{l+1}, \end{cases}$$

set $l = l + 1$, and go to (ii).

The first order optimality system of the minimization problem in step (iii) is given by

$$\begin{aligned} & -\Delta y^{l+1} = u^{l+1} \text{ in } L^2(\Omega), \\ & J''(\mathbf{x}^l)\mathbf{x}^{l+1} - \Delta p^{l+1} + \big(\gamma(y^{l+1} - \psi)\chi_{A^{l+1}}, 0\big) = \\ & \qquad - J'(\mathbf{x}^l) + J''(\mathbf{x}^l)\mathbf{x}^l - \big(\bar\lambda\chi_{A^{l+1}}, 0\big) \text{ in } X^*. \end{aligned}$$

Note that this system corresponds to a linearization of (5.14) at $\mathbf{x}^l = (y^l, u^l)$. Consequently, step (iii) is identical to the solution of the linear system within an iteration of a semismooth Newton method for solving (5.14).

Outer iteration: Inexact solutions and γ-update. Next we turn to the outer iteration. First notice that for small γ there is no need for highly accurate solutions of the regularized problem $(\mathcal{P}_\gamma)$, since we expect (y_γ, u_γ) to be only a coarse approximation of (y^*, u^*). Motivated by [HK1] we therefore propose a procedure requiring approximate solutions of the path problem lying in some neighborhood of the path only. For this purpose we introduce the residuals

$$\begin{aligned} r_{\mathbf{x}}(\mathbf{x}) &= \| -\Delta y - u\|_{L^2}, \\ r_p(\mathbf{x}, p, \lambda) &= \|J'(\mathbf{x}) - \Delta p + (\lambda, 0)\|_{X^*}, \\ r_\lambda(y, \lambda) &= \|\lambda - (\bar\lambda + \gamma(y - \psi))^+\|_{\mathcal{W}^*} \end{aligned}$$

and define the neighborhood

$$\mathcal{N}(\gamma, r) = \left\{ (\mathbf{x}, p, \lambda) \in Z : \max\{r_{\mathbf{x}}(\mathbf{x}), r_p(\mathbf{x}, p, \lambda), r_\lambda(y, \lambda)\} \le \frac{\tau}{\gamma^r} \right\}$$

with $Z = X \times H_0^1(\Omega) \times L^2(\Omega)$ for some fixed $\tau > 0$ and $r > 0$. In our implementation we typically choose r in accordance with our models (4.3) or (4.5). In the subsequent algorithm, for fixed γ, we stop Algorithm PDAS$_\gamma$ as soon as $(\mathbf{x}^l, p^l, \lambda^l) \in \mathcal{N}(\gamma, r)$ for the first time.

Once an approximate solution of $(\mathcal{P}_\gamma)$ is obtained we have to update γ. To this end, we introduce the feasibility measure ρ^F and the complementarity measure ρ^C as follows:

$$\begin{aligned} \rho^F(y) &:= \int_\Omega (y - \psi)^+ dx, \\ \rho^C(y) &:= \int_{I(y)} (y - \psi)^+ dx + \int_{A(y)} (y - \psi)^- dx, \end{aligned}$$

where $A(y) = \{x \in \Omega : \bar\lambda(x) + \gamma(y(x) - \psi(x)) > 0\}$, $I(y) = \Omega \setminus A(y)$, and $(\cdot)^- = \min(0, \cdot)$. Whenever $y = y_{n+1}$ and $\gamma = \gamma_n$, we write A_{n+1}, I_{n+1}, and ρ^F_{n+1}, ρ^C_{n+1}. For $\min(\rho^F_{n+1}, \rho^C_{n+1}) > 0$ we obtain a candidate γ^+_{n+1} for γ_{n+1} as

$$\gamma^+_{n+1} = \max\left(\gamma_n \max\left(\tau_1, \frac{\rho^F_{n+1}}{\rho^C_{n+1}} \right), \frac{1}{\max(\rho^F_{n+1}, \rho^C_{n+1})^q} \right) \tag{5.16}$$

with $\tau_1 > 1$ and $q \ge 1$; otherwise we set $\gamma^+_{n+1} = \tau_1 \gamma_n$. We include the quotient $\rho^F_{n+1}/\rho^C_{n+1}$ in order to significantly increase γ whenever $\rho^F_{n+1} \gg \rho^C_{n+1}$, i.e., when the iterates primarily lack feasibility rather than complementarity. The choice $q > 1$ induces certain growth rates for γ. Similar to [HK1] we also incorporate the following safeguard based on our models (4.3) respectively (4.5): Unless $\gamma_{n+1} < \tau_2 \gamma_n$, with $\tau_2 > 1$, we reduce γ^+_{n+1} until

$$|t_n(\gamma_{n+1}) - m_n(\gamma_{n+1})| \le \tau_3 |J(\mathbf{x}_{n+1}; \gamma_n) - J(\mathbf{x}_n; \gamma_{n-1})| \tag{5.17}$$

where $0 < \tau_3 < 1$, $t_n(\gamma) = J(\mathbf{x}_{n+1}; \gamma_n) + \frac{\partial J(\mathbf{x}_n; \gamma_n)}{\partial \gamma}(\gamma - \gamma_n)$, and m_n denotes one of the model functions according to (4.3) or (4.5) in iteration n. In other words, the linearization of m_n at γ_{n+1} should not be farther away from m_n than the distance of the previous two objective values of the regularized problem. As soon as (5.17) is satisfied we set $\gamma_{n+1} = \gamma^+_{n+1}$.

To determine the parameters in our model we use the actual approximate information on the value functional and its derivative as well as the value function at some reference point. In the sequel we only argue for the model (4.3). The case

(4.5) is treated similarly. Given γ_n in iteration n, for fixing D_n, $C_{1,n}$, and $C_{2,n}$ in the model $m_n(\gamma)$ we use the conditions

$$m_n(\gamma_n) = J(\mathbf{x}_n;\gamma_n), \quad \dot{m}_n(\gamma_n) = \frac{\partial J(\mathbf{x}_n;\gamma_n)}{\partial \gamma}(\mathbf{x}_n,\gamma_n), \quad m_n(\hat{\gamma}) = J(\hat{\mathbf{x}};\hat{\gamma}),$$

where $\hat{\mathbf{x}}$ denotes an approximate solution of $(\mathcal{P}_\gamma)$ at a reference value $\gamma = \hat{\gamma}$.

Now we are able to formulate our overall algorithm.

Inexact Path-Following Method (IPF).

(i) Initialized $\gamma_0 > 0$, select $r > 0$, and set $n := 0$.
(ii) Compute $(\mathbf{x}_n, p_n, \lambda_n) \in \mathcal{N}(\gamma_n, r)$.
(iii) Update γ_n by (5.16) with safeguard (5.17) to obtain γ_{n+1}.
(iv) Set $n = n + 1$, and go to (ii).

In step (ii) we use PDAS_γ for performing the inner iteration. The convergence of Algorithm IPF follows immediately from the convergence of Algorithm PDAS_γ for every fixed γ and the fact that $\gamma_{n+1} \geq \tau_1 \gamma_n$ with $\tau_1 > 1$ for all n, the property that $\frac{\tau}{\gamma_n^r} \to 0$ for $\gamma_n \to \infty$ in the definition of the neighborhood, and from Proposition 3.1.

5.2.2. *Lavrentiev regularization.* Finally we mention another regularization technique which keeps the inequality constraint explicit rather than removing it via the addition of $(2\gamma)^{-1}\|(\bar{\lambda} + \gamma(y-\psi))^+\|_{L^2}^2$ to the objective function as done by the Moreau-Yosida-based regularization. The *Lavrentiev regularization* technique, which was first introduced in [Tr], replaces $(\mathcal{P})$ by the regularized problem

$$(\mathcal{P}_\epsilon) \qquad \begin{cases} \min\ J(y,u) = \dfrac{1}{2}\displaystyle\int_\Omega (y-z)^2\ dx + \dfrac{\alpha}{2}\int_\Omega\ u^2\ dx\ , \\ -\Delta y = u \text{ in } \Omega\ ,\ y = 0\ \text{ on } \partial\Omega\ , \\ \epsilon u + y \leq \psi \text{ a.e. in } \Omega\ , \\ (y,u) \in H_0^1(\Omega) \times L^2(\Omega)\ , \end{cases}$$

with $\epsilon > 0$. Using a result in [Tr] it can be shown that the Lagrange multiplier pertinent to the mixed control-state inequality constraint in $(\mathcal{P}_\epsilon)$ exists as a function in $L^2(\Omega)$. The corresponding first order optimality condition is given by

$$-\Delta y = u, \tag{5.18a}$$

$$-\Delta p + y + \lambda = z, \tag{5.18b}$$

$$\alpha u - p + \epsilon\lambda = 0, \tag{5.18c}$$

$$\lambda \geq 0, \quad \epsilon u + y \leq \psi, \quad \lambda(\epsilon u + y - \psi) = 0 \quad \text{a.e. in } \Omega, \tag{5.18d}$$

with $y, p \in H_0^1(\Omega)$, $u, \lambda \in L^2(\Omega)$.

Setting $v = \epsilon u + y$ the problem $(\mathcal{P}_\epsilon)$ can be written as

$$(\hat{\mathcal{P}}_\epsilon) \qquad \begin{cases} \min\ \hat{J}(y,v) = \dfrac{1}{2}\displaystyle\int_\Omega (y-z)^2\ dx + \dfrac{\alpha}{2\epsilon^2}\int_\Omega\ (v-y)^2\ dx\ , \\ -\Delta y + \epsilon^{-1} y = \epsilon^{-1} v \text{ in } \Omega\ ,\ y = 0\ \text{ on } \partial\Omega\ , \\ v \leq \psi \text{ a.e. in } \Omega\ , \\ (y,u) \in H_0^1(\Omega) \times L^2(\Omega) \end{cases}$$

which resembles the control constrained optimal control problem $(\mathcal{P}_c)$ in section 2. For the numerical solution of $(\hat{\mathcal{P}}_\epsilon)$ for fixed ϵ there are two possible approaches:

- Path-following interior-point methods.
- Semismooth Newton or, equivalently, active-set methods.

Both techniques are developed and analyzed in function space below.

Short-step path-following interior-point method. This technique was pursued in [PTW] where a function space version of a short-step primal path-following interior-point method is analyzed in detail. Similar to the interior-point concept outlined in section 5.1.2 the starting point of the algorithmic development is the μ-perturbed optimality system

$$-\Delta y + \epsilon^{-1} y = \epsilon^{-1} v, \tag{5.19a}$$

$$-\Delta p + \epsilon^{-1} p + (1 + \alpha\epsilon^{-2}) y = \alpha\epsilon^{-2} v + z, \tag{5.19b}$$

$$\alpha\epsilon^{-2}(v - y) - \epsilon^{-1} p + \lambda = 0, \tag{5.19c}$$

$$\lambda(\psi - v) = \mu \quad \text{a.e. in } \Omega, \tag{5.19d}$$

for $\lambda > 0$ a.e., $\psi - v > 0$ a.e. and $\mu > 0$. This system can be reduced to

$$F(w;\mu) := D^* S^*(SDw - \hat{z}) + \alpha D^*(Dw - \hat{\psi}) - \frac{\mu}{w} = 0, \tag{5.20}$$

where $S = \imath(-\Delta)^{-1}$, with the embedding operator $\imath : H_0^1(\Omega) \to L^2(\Omega)$, and $D = (S + \epsilon I)^{-1}$. Further we use $\hat{z} = SD\psi - z$ and $\hat{\psi} = D\psi$. Note that (5.20) corresponds to the μ-perturbed first order optimality system of

$$\min \quad \frac{1}{2}|SDw - \hat{z}|_{L^2}^2 + \frac{\alpha}{2}|Dw - \hat{\psi}|_{L^2}^2 \text{ s.t. } w \geq 0 \text{ a.e. in } \Omega. \tag{5.21}$$

Given some strictly feasible estimate $w_n > 0$ a.e. in Ω of w^*, the solution of (5.21), and $\mu_n > 0$, the next iterate is obtained by performing a Newton step

$$\frac{\partial F(w_n;\mu_n)}{\partial w} \delta w_n = -F(w_n;\mu_n)$$

and setting $w_{n+1} = w_n + \delta w_n$. Here some care in the choice of μ_n is required in order to maintain strict feasibility of w_{n+1}. Then $\mu_{n+1} = \sigma\mu_n$ is set, with appropriate $\sigma \in (0,1)$, and the cycle is repeated until some stopping rule is satisfied.

Recently, it was shown in [SW] that a control reduced variant of the above method converges locally superlinearly in function space.

Semismooth Newton method. The second technique for solving $(\hat{\mathcal{P}}_\epsilon)$ is due to [H], where the functional analytic similarity between $(\mathcal{P}_c)$ and $(\hat{\mathcal{P}}_\epsilon)$ is observed and utilized in the algorithmic development. In fact, similar to the primal-dual active set strategy as outlined in section 5.1.1, given some estimate (v_{n-1}, λ_{n-1}) of $(v_\epsilon, \lambda_\epsilon)$, the solution of $(\hat{\mathcal{P}}_\epsilon)$, one may use

$$A_n = \{x \in \Omega : \lambda_{n-1}(x) + c(v_{n-1}(x) - \psi(x)) > 0\}, \tag{5.22a}$$

$$I_n = \{x \in \Omega : \lambda_{n-1}(x) + c(v_{n-1}(x) - \psi(x)) \leq 0\} \tag{5.22b}$$

as an *active* respectively *inactive* set prediction. This results in the following function space method.

Primal-dual active-set strategy (PDAS$_\epsilon$).

(1) Initialization: choose $(y_0, \lambda_0, v_0) \in H_0^1(\Omega) \times L^2(\Omega) \times L^2(\Omega)$, $c = \frac{\alpha}{\epsilon^2}$, and set $n = 1$.
(2) Determine the subset of active/inactive indices according to (5.22).
(3) If $n \geq 2$ and $A_n = A_{n-1}$, then STOP; otherwise go to step 4.

(4) Find $(y_n, v_n, p_n, \lambda_n)$ as the solution to

$$
\begin{aligned}
&-\Delta y_n + \epsilon^{-1} y_n = \epsilon^{-1} v_n, \\
&-\Delta p_n + \epsilon^{-1} p_n + (1 + \alpha\epsilon^{-2}) y_n = \alpha\epsilon^{-2} v_n + z, \\
&\alpha\epsilon^{-2}(v_n - y_n) - \epsilon^{-1} p_n + \lambda_n = 0, \\
&v_n = \psi \text{ on } A_n,\ \lambda_n = 0 \text{ on } I_n.
\end{aligned}
$$

(5) Set $n = n + 1$ and go to step 2.

Notice that in contrast to the short-step path-following interior-point method PDAS_ϵ requires no further regularization. In fact, we point out in this context that the perturbation of the complementarity system in (5.19d) acts as a second regularization.

The choice $c = \alpha/\epsilon^2$ in step (1.) of the above algorithm is due to the fact that PDAS_ϵ is equivalent to a semismooth Newton method, which, for this particular choice of $c > 0$, converges locally at a superlinear rate in function space. Moreover, it can be shown that the method is mesh-independent. For more details we refer to [H].

Approaching the limit $\epsilon \to 0$. Finally we point out that currently for the Lavrentiev-regularization technique there is no rigorous path-following respectively homotopy concept (compared to the feasible and non-interior inexact path-following method of section 5.2.1) available for $\epsilon \to 0$.

6. Numerical results

In this section we present numerical results obtained by some of the methods introduced in the previous section. Our test problems are all posed in 2D with $\Omega = (0,1)^2$. The Laplace-operator is discretized by the five-point finite difference stencil on a uniform mesh of mesh-size h. For the discretization of the integrals in the objective function we use the trapezoidal rule. As a starting point we use $y_0 \equiv \psi$, $u_0 = -\Delta y_0$, and $p_0 = \alpha u_0$ for PDAS and IPF. In addition, for PDIP we use $w_0 = \lambda_0 \equiv 1$ and $\mu_0 = 1$. The parameters in IPF had the values $\tau = 100$, $\tau_1 = 10$, $\tau_2 = 1.01$, $\tau_3 = 0.99$, $r = 0.2$, $q = 1.25$, and $\bar{\lambda} \equiv 0$. We stop the respective algorithm as soon as the primal, dual and complementarity residual norms drop below $0.1\,h^2$, respectively.

Problem 1. We use the data $\alpha = 0.1$, $z(x_1, x_2) = 10(\sin(2\pi x_1) + x_2)$ and $\psi \equiv 0.01$. Figure 1 shows the optimal solution for $h = 1/128$.

In table 1 we compare PDAS, PDIP, and IPF against each other. Recall that PDAS and PDIP are finite dimensional methods whereas IPF admits a function space analysis. For IPF next to the total number of outer iterations we denote in parenthesis the total number of inner iterations, i.e., iterations required by PDAS_γ for the selected γ-sequence. From table 1 we see that IPF is superior to the two finite dimensional methods. On the other hand, we observe that PDIP outperforms PDAS. In table 2 we further quantify the difference between the algorithms with respect to CPU-time. Rather than providing the elapsed CPU-times directly, we report on the following *CPU-ratios*:

$$\text{CPU-ratio}(\textit{method}) = \text{CPU-time}\ (\textit{method})/\text{CPU-time(IPF)}.$$

While PDIP requires approximately twice as much CPU-time as IPF, the active-set method PDAS deteriorates as h becomes small. This is due to the lack of

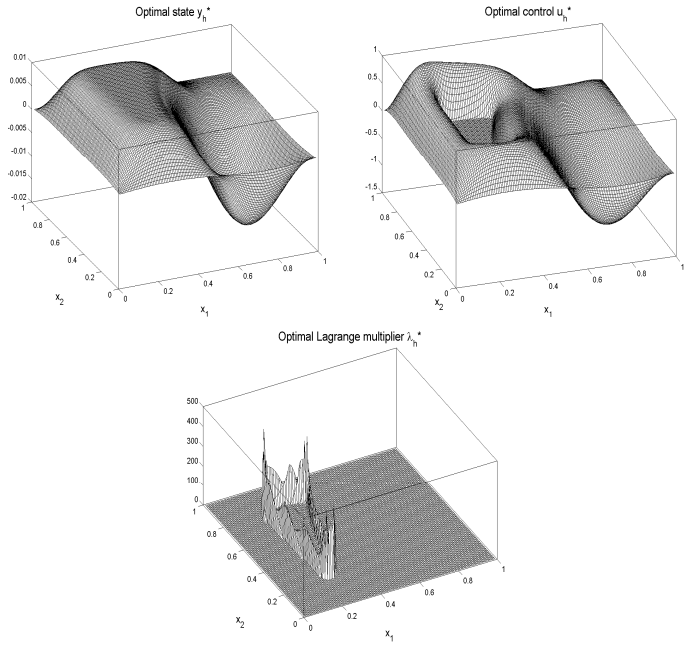

FIGURE 1. Optimal state (left), optimal control (middle), and optimal multiplier (right) for problem 1 with $h = 1/128$.

Mesh size h	1/16	1/32	1/64	1/128	1/256
PDAS	14	27	54	113	226
PDIP	13	15	17	20	19
IPF	7(11)	9(15)	9(14)	7(13)	8(15)

TABLE 1. Comparison of iteration numbers for different mesh-sizes and methods.

Mesh size h	1/16	1/32	1/64	1/128	1/256
CPU-ratio(PDAS)	0.48	0.89	1.91	4.44	5.56
CPU-ratio(PDIP)	1.08	1.11	1.57	2.02	1.97

TABLE 2. Comparison of CPU-ratios for different mesh-sizes and methods.

function space properties of PDAS. For PDIP, on the other hand, our results indicate that the method behaves essentially mesh independently. This could be an indication that a rigorous function space analysis is possible.

Finally we report on the speed-up of the respective method when utilizing a *nested iteration* concept. To be specific, we initialize the respective algorithm as described above on the coarsest grid ($h_0 = 2^{-2}$), run the respective algorithm until the residuals drop below $0.1\, h_0^2$, then we interpolate the solution to the next finer mesh with mesh-size $h_i = 0.5\, h_{i-1}$ and use the interpolated coarse grid solution as the initial guess on the fine mesh. This procedure is repeated until the finest mesh

is reached ($h = 2^{-8}$ in our case). Then we obtain the results displayed in table 3. In the nested iteration case the average CPU-ratios are

Mesh size h	1/4	1/8	1/16	1/32	1/64	1/128	1/256	total
PDAS	3	4	4	5	6	6	6	34
PDIP	3	2	4	4	5	6	7	31
IPF	4	3	3	4	5	5	5	29

TABLE 3. Problem 1. Comparison of iteration numbers for different mesh-sizes and methods based on nested iteration.

$$\text{CPU-ratio(PDAS)} \approx 0.8 \quad \text{and} \quad \text{CPU-ratio(PDIP)} \approx 2.$$

These results, which are typical in our test runs also for other problems, suggest that PDAS benefits significantly from the nested iteration concept. To appreciate the efficiency of the nested approach, we can think of the iteration levels as discrete regularization parameters. The relation between PDIP and IPF, on the other hand, stays approximately the same. We point out that although PDAS requires slightly more iterations than IPF it is still faster. This is due to the fact that in every iteration of PDAS one has to solve a linear system only on the current estimate of the inactive set at the solution. Hence, in the case where this estimate is significantly smaller than the whole computational domain, the solution times reduce remarkably.

For further numerical results comparing PDAS, PDIP and IPF we refer to [HK2]. For the Lavrentiev-regularized problem considered in section 5.2.2, in [H] one can find numerical results and comparisons obtained by a semismooth Newton technique and the short-step path-following interior-point method, both outlined in section 5.2.2 as well.

Finally, we provide a brief qualitative comparison of the Moreau-Yosida-based regularization with the Lavrentiev approach. In Figure 2 we show value functionals for both approaches. The graph on the left depicts the value functional for the Lavrentiev-regularization for a given test problem. The plot in the middle shows the Lavrentiev-based value functional for a slight modification of this problem. The right plot shows the corresponding Moreau-Yosida path for $\bar{\lambda} \equiv 0$. We point out that the qualitative behavior of the Moreau-Yosida path is independent of the underlying problem. We conclude that the Moreau-Yosida-path induces a monotonically increasing value functional, while the value functional for the Lavrentiev-regularization exhibits a problem dependent, possibly non-monotone behavior. In this respect, the Moreau-Yosida approach appears to be better suited for path-following strategies than the Lavrentiev-based technique.

Concluding we can say that some type of regularization appears to be necessary to solve state constrained problems efficiently in a nearly mesh-independent manner. The primal dual active set strategy is especially simple to implement, its path version has favorable geometric properties and combined with nested iteration it is very efficient numerically.

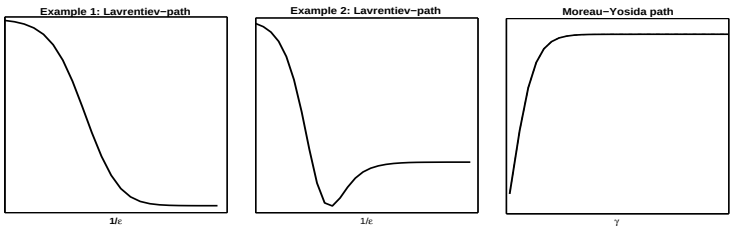

FIGURE 2. Comparison of value functionals for the Lavrentiev regularization for two different problems (left and middle) and the corresponding Moreau-Yosida regularization (in the case $\bar{\lambda} \equiv 0$; right).

APPENDIX A. NEWTON DIFFERENTIABILITY AND SEMISMOOTH NEWTON METHODS

We define the notion of a Newton differentiable function, a semismooth Newton technique for solving the operator equation $F(x) = 0$ based on this differentiability concept, and we address its local convergence properties. For details and further discussions including proofs we refer to [HIK].

Let $F : X \to Y$ denote a mapping between the Banach spaces X and Y. The associated norms are denoted by $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively.

Definition A.1. *The mapping $F\colon D \subset X \to Y$ is called* Newton differentiable *in the open subset $U \subset D$ if there exists a family of mappings $G\colon U \to \mathcal{L}(X, Y)$ such that*

$$\lim_{h\to 0} \frac{1}{\|h\|_X} \|F(x+h) - F(x) - G(x+h)h\|_Y = 0 \tag{A}$$

for every $x \in U$.

A mapping F satisfying property (A) is also called *semismooth*.

Let F satisfy (A), and assume that G is invertible in U. Then, given $x^k \in U$, the semismooth Newton step

$$x^{k+1} = x^k - G(x^k)^{-1} F(x^k) \tag{A.1}$$

is well-defined. The following local convergence result holds true.

Theorem A.1. *Suppose that x^* satisfies $F(x^*) = 0$ and that F is Newton differentiable in an open neighborhood U containing x^* with Newton derivative $G(x)$. If $G(x)$ is nonsingular for all $x \in U$ and $\{\|G(x)^{-1}\| : x \in U\}$ is bounded, then the Newton iteration* (A.1)*, with $k = 0, 1, 2 \ldots$, converges superlinearly to x^*, provided that $\|x^0 - x^*\|$ is sufficiently small.*

REFERENCES

[BHHK] M. BERGOUNIOUX, M. HADDOU, M. HINTERMÜLLER AND K. KUNISCH. A comparison of interior point methods and a Moreau-Yosida based active set strategy for constrained optimal control problems. SIAM J. Optim., 11(2):495–521, 2000.

[BIK] M. BERGOUNIOUX, K. ITO AND K. KUNISCH. Primal-dual strategy for optimal control problems. SIAM J. Control Optim., 37:1176–1194, 1999.

[BK1] M. BERGOUNIOUX AND K. KUNISCH. Augmented Lagrangian techniques for elliptic state constrained optimal control problems. SIAM J. Control Optim., 35:1524–1543, 1997.

[BK2] M. Bergounioux und K. Kunisch. On the Structure of the Lagrange Multiplier for State-Constrained Optimal Control Problems. Systems and Control Letters, 48:16–176, 2002.

[BK3] M. Bergounioux und K. Kunisch. Primal-dual strategy for state-constrained optimal control problems. Journal of Comp. Optim. Appl., 22:193–224, 2002.

[BP] V. Barbu and Th. Precupanu. *Convexity and Optimization in Banach Spaces*. Reidel Publ. Comp., Dordrecht, 1986.

[BG] J. F. Bonnans and C. C. Gonzaga, Convergence of interior point algorithms for the monotone linear complementarity problem. Math. Oper. Res., 21:1–25, 1996.

[BPR] J. F. Bonnans, C. Pola, and R. Rebaï, Perturbed path following interior point algorithms. Optim. Methods Softw., 11-12:183–210, 1999.

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

[CRZ] E. Casas, J.-P. Raymond, and H. Zidani. Pntryagin's principle for local solutions of control problems with mixed control-state constraints. SIAM. J. Control Optim., 39(4):1182–1203, 2000.

[CNQ] X. Chen, Z. Nashed, and L. Qi. Smoothing methods and semismooth methods for nondifferentiable operator equations. *SIAM J. Numer. Anal.*, 38(4):1200–1216, 2000.

[FGW] A. Forsgren, P. E. Gill, and M. H. Wright. Interior methods for nonlinear optimization. *SIAM Rev.*, 44(4):525–597, 2002.

[H] M. Hintermüller. Mesh-independence and fast local convergence of a primal-dual active-set method for mixed control-state constrained elliptic control problems. IMA-Preprint 2095, 2006.

[HH] M. Hintermüller and M. Hinze. A SQP-semismooth Newton-type algorithm applied to control of the instationary Navier-Stokes system subject to control constraints. SIAM J. Optim., 16(4):1977-2000, 2006.

[HIK] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. SIAM J. Optim., 13(3):865–888, 2003.

[HK1] M. Hintermüller and K. Kunisch. Path-following methods for a class of constrained minimization problems in function space. SIAM J. Optim., 17(1):159–187, 2006.

[HK2] M. Hintermüller and K. Kunisch. Feasible and Non-Interior Path-Following in Constrained Minimization with Low Multiplier Regularity. SIAM J. Control Optim., to appear.

[HR] M. Hintermüller and W. Ring. A level set approach for the solution of a state constrained optimal control problem. *Num. Math.*, 98(1):135–166, 2004.

[HU] M. Hintermüller and M. Ulbrich. A mesh independence result for semismooth Newton methods. Math. Prog., 101(1): 151–184, 2004.

[LMS] I. J. Lustig, R. E. Marsten, and D. F. Shanno. On implementing Mehrotra's predictor-corrector interior-point method for linear programming. SIAM J. Optim., 2:435–449, 1992.

[Meh] S. Mehrotra. On the implementation of a primal-dual interior point method. SIAM J. Optim., 2:575–601, 1992.

[MTY] S. Mizuno, M. Todd, and Y. Ye. On adaptive step primal-dual interior-point algorithms for linear programming. Math. Oper. Res., 18:964–981, 1993.

[PTW] U. Prüfert, F. Tröltzsch, and M. Weiser. The convergence of an interior point method for an elliptic control problem with mixed control-state constraints. Preprint 36-2004, TU Berlin, 2004.

[RK] J.-C. de los Reyes and K. Kunisch. A semi-smooth Newton method for control constrained optimal control of the Navier Stokes equations. *Nonlinear Analysis*. to appear.

[SW] A. Schiela, and M. Weiser. Superlinear convergence of the control reduced interior point method for PDE constrained optimization. ZIB-Report 05-15 (2005), Zuse-Institut, Berlin.

[T] G. M. Troianiello. *Elliptic Differential Equations and Obstacle Problems*. Plenum Press, New York, 1987.

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

[VY] R. J. Vanderbei and B. Yang. On the Symmetric Formulation of Interior-Point Methods. Technical report SOR 94-05, Princeton University, Princeton, NJ, 1994.

[V2] R. J. Vanderbei. *Linear Programming: Foundations and Extensions*. Kluwer Academic Publishers, Boston, MA, 1997.
[Wr] S. J. Wright. *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, 1997.
[Y] Y. Ye. *Interior Point Algorithms: Theory and Analysis*. Wiley-Intersci. Ser. Discrete Math. Optim., John Wiley, New York, 1997.
[Z2] Y. Zhang. Solving large-scale linear programs by interior-point methods under the MATLAB environment. Optim. Methods Softw., 10:1–31, 1998.

Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, A-8010 Graz, Austria.

Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, A-8010 Graz, and Radon Institute, A-4040 Linz, Austria.