Math. Program., Ser. A (2013) 138:83–114
DOI 10.1007/s10107-012-0595-y

FULL LENGTH PAPER

An interior point method in function space for the efficient solution of state constrained optimal control problems

Anton Schiela

Received: 7 July 2008 / Accepted: 8 September 2012 / Published online: 6 October 2012

Abstract We propose and analyze an interior point path-following method in function space for state constrained optimal control. Our emphasis is on proving convergence in function space and on constructing a practical path-following algorithm. In particular, the introduction of a pointwise damping step leads to a very efficient method, as verified by numerical experiments.

Keywords Interior point methods in function space · Optimal control · State constraints

Mathematics Subject Classification (2000) 90C51 · 49M05

1 Introduction

The construction and analysis of efficient algorithms for state constrained optimal control problems is still a considerable challenge. Presently, most popular methods that admit a (partial) analysis in function space are path-following methods, such as exterior penalty methods [8], Lavrentiev regularization [10,11] and interior point methods [14–16]. Except for [14] and partially [11] (for a fixed Lavrentiev parameter) the available results are restricted to properties of the *homotopy path*, such as its existence, convergence and continuity. Except for these two works, not much is known about convergence of the associated *path-following algorithms*. This includes the important question whether it is at all possible to follow the homotopy path by a

This study was supported by the DFG Research Center MATHEON "Mathematics for key technologies".

A. Schiela (✉)
Zuse Institute Berlin, Takustraße 7, 14195 Dahlem, Berlin, Germany
e-mail: schiela@zib.de

practical algorithm, or whether the sequence of iterates may stagnate far away from the desired solution. Closely connected and even more relevant from a practical point of view is the question how to choose homotopy parameters to obtain a fast and robust algorithm. These questions can certainly not be answered by an analysis of the path alone.

The aim of this paper is to propose and analyze an interior point method in function space that is capable of solving state constrained optimal control problems efficiently. The corresponding homotopy path has been analyzed in [15,16], so our emphasis here is on the Newton path-following method and on giving positive answers to the above questions. We establish qualitative convergence results in the following sense. Under suitable conditions there is a sequence of homotopy parameters that converges to zero and a sequence of corresponding iterates produced by a Newton corrector scheme that converges to the solution of the original problem. The quantities used in the analysis, which yields convergence of the scheme as an a-priori result, can be modelled and estimated inside a numerical algorithm to yield a criterion for controlling the path-following algorithm efficiently. This is done in the spirit of [4, Chapter 5], but modified in a way that fits into our particular setting in function space.

To establish a rigorous analysis we essentially need estimates for two quantities. The first one, which reflects the most basic analytic properties of the homotopy path, is its local Lipschitz constant $\eta(\mu)$. The second captures the nonlinearity of the equations that define the homotopy path. This quantity, which governs the local convergence behavior of Newton's method and in particular its radius of convergence, is an affine covariant Lipschitz constant for the Jacobian, denoted by $\omega(\mu)$. Since good a-posteriori estimates are available for η and ω, their role is not a purely analytic one, but they establish a close connection between a-priori theory and algorithmic implementation. In some sense, the algorithm is driven by an a-posteriori counterpart of the convergence theory established in this work.

Compared to [14] we introduce, as an algorithmic modification, a pointwise damping step, which prevents Newton's method from leaving the feasible domain and enhances the efficiency of the path-following scheme significantly. It is motivated by the idea to exploit the pointwise structure of the problem and has several useful interpretations. In our numerical experiments we observe that this modification allows the solution of state constrained optimal control problems in a few Newton steps.

2 A class of state constrained optimal control problems

Let Ω be an open and bounded domain in $\mathbb{R}^d$ and $\overline{\Omega}$ its closure. Let $Y = (C(\overline{\Omega}), \|\cdot\|_\infty)$ denote the space of states and $U = (L_2(Q), \|\cdot\|_{L_2(Q)})$ the space of controls, where Q is a subset of $\mathbb{R}^n$, equipped with a positive, regular and finite measure.

Consider a convex minimization problem of the following form, the details of which are fixed in the remaining section.

$$\min_{(y,u)\in Y\times U} J(y,u) := \frac{1}{2}\,\|y-y_d\|^2_{L_2(\Omega)} + \frac{\alpha}{2}\,\|u\|^2_{L_2(Q)} \quad \text{s.t. } Ay - Bu = 0$$

$$\underline{y} \leq y. \quad (1)$$

We will now specify our theoretical framework, which holds throughout this work and collect a couple of basic results about this class of problems. Further we will give an example in Sect. 2.3 below.

2.1 Equality constraints

In the following we will describe our abstract framework for the equality constraint $Ay - Bu = 0$ which models a partial differential equation.

Recall that a linear operator $T : V \supset \operatorname{dom} T \to W$ on a domain of definition $\operatorname{dom} T$ is called closed, if for any sequence v_k in $\operatorname{dom} T$ we have the relation

$$(v_k \to v \text{ and } Tv_k \to w) \Rightarrow (v \in \operatorname{dom} T \text{ and } Tv = w).$$

Assumption 1 Let $Y = C(\overline{\Omega})$, $U = L_2(Q)$, as defined above and let R be a reflexive Banach space. Assume that the linear operator

$$A : Y \supset \operatorname{dom} A \to R$$

is *densely defined, closed*, and maps $\operatorname{dom} A$ to R *bijectively*. Let

$$B : U \to R$$

be a continuous linear operator.

We consider A as a model of a differential operator, which, due to the choice $Y = C(\overline{\Omega})$ is usually unbounded. Closed, densely defined operators between Banach spaces are a classical concept of functional analysis. They generalize the concept of continuous operators and retain much of their structure. In particular, there is an open-mapping theorem, a closed range theorem, and adjoint operators are well defined. In this work and in [15,16] only these basic properties of A are needed for a successful analysis. A classical introduction to unbounded operators is [5], but most elementary facts can also be found in standard textbooks on functional analysis.

There is a simple correspondence between a bijective closed operator and its inverse.

Lemma 1 *Let V and W be Banach spaces and $T : V \supset \operatorname{dom} T \to W$ be a linear operator. T is closed and bijective if and only if T possesses a continuous inverse $T^{-1} : W \to \operatorname{dom} T \subset V$.*

Let $S : V \to W$ be continuous and $T : V \supset \operatorname{dom} T \to W$ closed. Then

$$S + T : V \supset \operatorname{dom} T \to W$$

is closed.

Proof For the first assertion, see [5, IV.1.1]. The second assertion is an easy calculation. By continuity of S and closedness of T we compute

$$\begin{aligned} & v_k \to v \text{ and } (S+T)v_k \to w \\ \Rightarrow\quad & v_k \to v \text{ and } Tv_k \to w - Sv \\ \Rightarrow\quad & v \in \operatorname{dom} T \text{ and } (T+S)v = w \end{aligned}$$

Hence $S + T$ is closed. □

Because $\operatorname{dom} A$ is dense in Y one can define an adjoint operator (cf. [5, Definition II.2.2])

$$A^* : R^* \supset \operatorname{dom} A^* \to Y^*.$$

for which

$$\langle l, Ay\rangle = \langle A^* l, y\rangle \qquad \forall\, l \in \operatorname{dom} A^*,\, y \in \operatorname{dom} A \tag{2}$$

holds. Its domain $\operatorname{dom} A^*$ is defined canonically and A^* is automatically closed (cf. [5, Theorem II.2.6]). Because R is reflexive and A is closed, $\operatorname{dom} A^*$ is dense in R^* (cf. [5, Theorem II.2.14]) and A^* is continuously invertible, because A is closed and surjective (cf. [5, Theorem II.4.4]).

2.2 Inequality constraints

For the inequality constraints $\underline{y} \le y$ we assume:

Assumption 2 Let $\underline{y} \in C(\overline{\Omega})$. The inequality constraints in (1) are interpreted in the pointwise sense:

$$\underline{y}(t) \le y(t)\ \forall t \in \overline{\Omega}.$$

We call y strictly feasible, if $\underline{y}(t) < y(t)$ for all $t \in \overline{\Omega}$. By compactness, this implies

$$0 < \min_{t\in\overline{\Omega}} \{y(t) - \underline{y}(t)\}. \tag{3}$$

We assume that there is $(\breve{y}, \breve{u})$ that satisfies $A\breve{y} - B\breve{u} = 0$ and $\breve{y}$ is strictly feasible.

The point $(\breve{y}, \breve{u})$ is called a Slater point. This condition together with the topology of Y defined by $\|\cdot\|_\infty$ is needed in the analysis of dual variables and subdifferentials and in the derivation of first order optimality conditions.

2.3 An optimal control problem with an elliptic partial differential equation

As a simple example we consider the operators A and B associated with a elliptic distributed control problem with Neumann boundary conditions on a smoothly bounded domain $\Omega \subset \mathbb{R}^d$ with $d \in \{1, 2, 3\}$. More general classes of problems are discussed in [15, Section 3].

For some $\infty > p > d$ and $1/p + 1/p' = 1$ let

$$\begin{aligned} A : W^{1,p}(\Omega) &\to W^{1,p'}(\Omega)^* \\ y &\mapsto Ay : \langle Ay, v\rangle := \int_\Omega \langle \nabla y, \nabla v\rangle + yv\, dt. \end{aligned}$$

It follows from regularity theory (cf. e.g. [2, Thm. 9.2]) that A is an isomorphism. Moreover, by $p > d$ there exists a continuous Sobolev embedding $W^{1,p}(\Omega) \hookrightarrow C(\overline{\Omega})$, which is dense. These two results allow us to define A as a closed, densely defined, bijective operator via Lemma 1:

$$A : C(\overline{\Omega}) \supset W^{1,p}(\Omega) \to W^{1,p'}(\Omega)^*,$$

i.e. $R = (W^{1,p'}(\Omega))^*$, dom $A = W^{1,p}(\Omega)$. An example for the operator B is given by

$$\begin{aligned} B : L_2(\Omega) &\to W^{1,p'}(\Omega)^* \\ u &\mapsto Bu : \langle Bu, v\rangle := \int_\Omega uv\, dt. \end{aligned}$$

B is continuous by the Sobolev embedding theorems, if $W^{1,p'}(\Omega) \hookrightarrow L_2(\Omega)$ is continuous, i.e. $p' \geq 2d/(d+2)$. Since $p > d$ implies $p' < d/(d-1)$ (or $p' < \infty$ for $d = 1$), there is a range of possible choices of p and p' for $d \in \{1, 2, 3\}$.

Regularity theory yields that for $u \in L_2(\Omega)$, we even obtain $y \in H^2(\Omega) \hookrightarrow C^\beta(\overline{\Omega})$, where $\beta > 0$ depends on the spatial dimension d. Hölder continuity of states can also be established in a much more general framework (cf. [7]).

Remark 1 Neumann boundary conditions can, of course, be replaced by Dirichlet or mixed boundary conditions. In this case $Y = C(\overline{\Omega})$ has to be replaced by a suitable space of continuous functions with prescribed boundary values.

3 Barrier regularizations for state constraints

Our theoretical framework, described above is contained in the framework of [15], where barrier regularizations for a class of state constraints were analyzed. In this section we recapitulate basic results of [15], needed later.

Definition 1 For all $q \geq 1$ and $\mu > 0$ the functions $l(z; \mu; q) : [0; \infty[\rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$ defined by

$$l(z; \mu; q) := \begin{cases} -\mu \ln(z) & : q = 1 \\ \dfrac{\mu^q}{(q-1)z^{q-1}} & : q > 1 \end{cases}$$

are called *barrier functions of order* q. We extend their domain of definition to $\mathbb{R}$ by setting $l(z; \mu; q) = \infty$ for $z \leq 0$. We include finite sums of these barrier functions, and define their order to be the maximum order of the summands. We denote their derivatives, which are defined for $z > 0$ by l' and l''.

Usually we do not have to consider special values of q or μ. In these cases we may abbreviate the notation $l(z; \mu; q)$ by $l(z; \mu)$ or even $l(z)$.

Using these barrier *functions* and a given lower bound $\underline{y}$ we construct barrier *functionals* $b(y; \mu; q)$ to implement constraints of the form $\underline{y} \leq y$ on $\overline{\Omega}$ by computing the integral over l:

$$\begin{aligned} b(\cdot; \mu; q) &: C(\overline{\Omega}) \rightarrow \overline{\mathbb{R}} \\ y &\mapsto \int_{\overline{\Omega}} l(y(t) - \underline{y}; \mu; q)\, dt. \end{aligned}$$

It is easy to see that b is a well defined, extended real valued functional on $C(\overline{\Omega})$. With these definitions we may perform a barrier regularization of problem (1):

$$\min_{(y,u) \in Y \times U} J(y, u) + b(y; \mu; q) \quad \text{s.t. } Ay - Bu = 0. \tag{4}$$

As usual, the inequality constraints have been removed and a barrier functional has been added. If y is infeasible, then $b(y; \mu; q) = \infty$, and thus minimizers of (4) have to be feasible.

We denote by b' and b'' the *formal* derivatives of b. Here,

$$\langle b'(y; \mu; q), \delta y \rangle = \int_{\overline{\Omega}} l'(y(t) - \underline{y}; \mu; q) \delta y(t)\, dt, \tag{5}$$

if the right hand side is well defined. An analogous definition holds for b''. We call these quantities *formal derivatives*, because in general they may not have the properties of a derivative, and it is not even clear, a-priori whether (5) is well defined, because for given y, δy the function $l'(y; \mu; q)\delta y$ may not be integrable on $\overline{\Omega}$. If, however, y is strictly feasible for all $t \in \overline{\Omega}$, then the functions $t \rightarrow l'(y(t) - \underline{y}(t); \mu; q)$ and $t \rightarrow l''(y(t) - \underline{y}(t); \mu; q)$ are continuous. With slight abuse of notation we will identify $b'(y; \mu; q)$ and $b''(y; \mu; q)$ with these functions on $\overline{\Omega}$, respectively.

Our next assertion captures existence and analytic properties of our solutions.

Theorem 1 *For all $\mu > 0$, problem (4) admits a unique solution $(y(\mu), u(\mu))$.*

The set of solutions of (4) forms a path that converges to the solution (y_*, u_*) of (1) with the error estimates

$$J(y(\mu), u(\mu)) - J(y_*, u_*) \leq C\mu \tag{6}$$

$$\|y(\mu) - y_*\|_\infty + \|u(\mu) - u_*\|_{L_2(Q)} \leq c\sqrt{\mu}. \tag{7}$$

This path is locally Lipschitz continuous for each $\mu > 0$, and satisfies

$$\|y(\mu) - y(\nu)\|_\infty + \|u(\mu) - u(\nu)\|_{L_2(Q)} \leq c\mu^{-1/2}|\mu - \nu|. \tag{8}$$

If $\mu \geq \nu \geq \mu/2$, then

$$\left\| \sqrt{b''(y(\mu); \mu; q)}(y(\mu) - y(\nu)) \right\|_{L_2(\Omega)} \leq c\mu^{-1/2}|\mu - \nu|. \tag{9}$$

Proof Existence of minimizers follows from [15, Theorem 5.2]. Equation (7) follows from [15, Theorem 6.3], (8) and (9) follow from [15, Theorem 6.5]. □

As usual in interior point methods we will call this homotopy path of solutions the *central path*.

Next we consider first order optimality conditions for barrier problems, which are necessary and sufficient by convexity of the problem.

Theorem 2 *Let $(y(\mu), u(\mu))$ be a barrier minimizer for $\mu > 0$. If $y(\mu)$ is strictly feasible, then there exists a unique adjoint state $p(\mu) \in R^*$ such that the following equations are satisfied:*

$$0 = y(\mu) - y_d + A^* p(\mu) + b'(y(\mu); \mu; q) \tag{10}$$

$$0 = \alpha u(\mu) - B^* p(\mu), \tag{11}$$

$$0 = Ay(\mu) - Bu(\mu). \tag{12}$$

If, in turn, (10)–(12) are satisfied by some $(y(\mu), u(\mu), p(\mu))$, then $(y(\mu), u(\mu))$ are barrier minimizers.

Moreover, for every $\mu_0 > 0$ $\|b'(y(\mu); \mu; q)\|_{L_1(\Omega)}$ is bounded above on the interval $]0, \mu_0]$.

Proof First order optimality conditions follow from [15, Theorem 5.4] together with [15, Proposition 4.6]. Uniform boundedness of b' in L_1 is a consequence of [15, Proposition 5.6]. □

Remark 2 First order optimality conditions can be extended to the case where $y(\mu)$ is not strictly feasible. Then additional measure terms appear at regions, where $y(\mu)$ touches the bounds. Details can be found in [15]. In this work we will, however, consider the strictly feasible case only, which holds for an appropriate choice of q (cf. Sect. 3.1, below).

We may use (11) to compute $u = \alpha^{-1} B^* p$, and eliminate u from the state equation $Ay - Bu = 0$. This, together with (10) yields the following system of equations

$$\begin{aligned} y - y_d + b'(y; \mu; q) + A^* p &= 0 \\ Ay - BB^* \alpha^{-1} p &= 0, \end{aligned} \tag{13}$$

which characterizes strictly feasible barrier minimizers.

Proposition 1 *Let (y, p) be a solution of (13) for some $\mu > 0$. Then*

$$(y, \alpha^{-1} B^* p) = (y(\mu), u(\mu))$$

is the barrier minimizer for μ and hence (y, p) is the unique solution of (13).

If $(y(\mu), u(\mu))$ is the barrier minimizer for μ and $y(\mu)$ is strictly feasible, then (13) has the unique solution $(y, p) = (y(\mu), p(\mu))$.

Proof It is easy to see that (13) and (10)–(12) are equivalent. Now Theorem 2 yields the result. □

The system of Eq. (13) will be in the center of our considerations. Our path-following method is based on solving this system approximately by Newton's method. For an analysis in function space it is therefore necessary to guarantee strict feasibility of $y(\mu)$.

3.1 Strict feasibility

Strict feasibility of $y(\mu)$, i.e. $\underline{y}(t) < y(\mu)(t)$ for all $t \in \overline{\Omega}$ is not guaranteed a-priori under Assumptions 1 and 2. However, under slightly stronger assumptions on the regularity of y and $\underline{y}$ strict feasibility of $y(\mu)$ was studied in [15, Section 7]. In this section we will recapitulate the most important results.

We recall from [1, Definition IV.4.3] that $\Omega \subset \mathbb{R}^d$ satisfies a so-called *cone property* if there is a "cone" K (defined as the convex hull of a ball and a point in $\mathbb{R}^d$) such that each point $t \in \Omega$ is the vertex of a cone $K_t \subset \Omega$, which is the image of a rigid motion of K. Further, we denote for $0 < \beta < 1$ by $C^\beta(\overline{\Omega})$ the space of Hölder continuous functions. Regularity theory of partial differential equations asserts Hölder continuity of solutions under very mild assumptions (cf. e.g. [7]). Moreover, since optimal controls $u(\mu)$ are uniformly bounded in U, the following assumption can be verified for a large class of problems:

Assumption 3 Let $\Omega \subset \mathbb{R}^d$ satisfy the cone property. For some $0 < \beta \leq 1$ assume that $y(\mu) - \underline{y}$ is uniformly bounded in $C^\beta(\overline{\Omega})$ on some positive interval $]0, \mu_0]$.

Proposition 2 *Suppose that Assumptions 1–3 hold and choose the order of the barrier functional $q \geq d/\beta$. Define for each $\mu \in]0, \mu_0]$ the quantity*

$$\psi(\mu) := \min_{t \in \overline{\Omega}} (y(\mu)(t) - \underline{y}(t)). \tag{14}$$

Then the function $\mu \to \psi(\mu)$ is continuous and $\psi(\mu) > 0$ on $]0, \mu_0]$.

In particular, $\psi(\mu)$ *is uniformly continuous and bounded away from* 0 *on every compact subinterval* $[\underline{\mu}, \mu_0]$.

Proof Continuity of ψ follows from the continuity of the mapping $\mu \to y(\mu)$ with respect to $\|\cdot\|_\infty$. Positivity of $\psi(\mu)$ from [15, Lemma 7.1]. The proof uses Hölder continuity of $y(\mu)$ and the cone property of Ω to show that $y(\mu)(t) = \underline{y}(t)$ implies $b'(y(\mu); \mu; q) \notin L_1(\Omega)$ for $q \geq d/\beta$. Our final assertion follows from a simple compactness argument. □

Hence, by an appropriate choice of q we can force the central path to be strictly feasible, approaching the bounds for $\mu \to 0$ in a controlled fashion. For a problem at hand, regularity theory usually yields information on β, and thus allows an a-priori choice of q.

Remark 3 If we know that $y(\mu)(t) > \underline{y}(t)$ on the boundary $\partial\Omega$ of Ω (e.g. by Dirichlet boundary conditions), then the above proposition can be extended to the condition $q \geq d/(1+\beta)$, if $y \in C^{1,\beta}(\Omega)$ (cf. [15, Lemma 7.1]).

Remark 4 It would be possible to impose inequality constraints on a compact subdomain of $\overline{\Omega}$, as long as this subdomain satisfies a cone condition. In a similar fashion, piecewise Hölder continuous functions $\underline{y}$ could be treated as well.

4 A simple Newton path-following method

In the following we are going to study a path-following algorithm, which is based on the solution of the system (13) for a sequence of μ by Newton's method. In the whole section we suppose that Assumptions 1–3 hold and that $q \geq d/\beta$ has been chosen so that strict feasibility of barrier minimizers is guaranteed.

Let $X = Y \times R^*$ be the space of states and adjoint states, i.e. $x = (y, p)$, equipped with the norm

$$\|\cdot\|_X := \|\cdot\|_\infty + \|\cdot\|_{R^*}.$$

Then by reflexivity of R we have $X^* = Y^* \times R$. Let us denote by $X_{\mathrm{sf}} \subset X$ those x with strictly feasible states

$$X_{\mathrm{sf}} := \{(y, p) \in X : \underline{y}(t) < y(t)\ \forall t \in \overline{\Omega}\}.$$

It follows from compactness of $\overline{\Omega}$ and the choice $(Y, \|\cdot\|_\infty)$ that X_{sf} is an open subset of X. Further, let

$$D_{\mathrm{sf}} := X_{\mathrm{sf}} \cap (\mathrm{dom}\, A \times \mathrm{dom}\, A^*).$$

Then for $\mu > 0$ we define (cf. 13)

$$\begin{aligned} &F(\cdot;\mu): X \supset D_{\mathrm{sf}} \to X^* \\ &x \mapsto F(x;\mu) := \begin{pmatrix} y - y_d + b'(y;\mu) + A^* p \\ Ay - BB^*\alpha^{-1} p \end{pmatrix}. \end{aligned}$$

Our aim in this section is to prove a qualitative convergence result for a simple Newton path-following algorithm in function space, based on the homotopy of equations $F(x;\mu) = 0$.

4.1 The linearization of F and its inverse

For $\mu > 0$ and $x \in X_{\mathrm{sf}}$ we define the linear operator

$$\begin{aligned} &F'(x;\mu): X \supset \mathrm{dom}\, A \times \mathrm{dom}\, A^* \to X^* \\ &\delta x \mapsto F'(x;\mu)\delta x := \begin{pmatrix} I + b''(y;\mu) & A^* \\ A & -\alpha^{-1} BB^* \end{pmatrix} \begin{pmatrix} \delta y \\ \delta p \end{pmatrix}. \end{aligned} \tag{15}$$

This is a *formal* linearization of F, because we do not specify in which sense $F'(x;\mu)$ is a derivative of $F(x;\mu)$. Observe from (15) that $F'(x;\mu)$ is defined *for all* $x \in X_{\mathrm{sf}}$, in contrast to $F(x;\mu)$ which is only defined on D_{sf}. As the sum of a continuous operator (the diagonal blocks) and a closed operator (the off-diagonal blocks) $F'(x;\mu)$ is closed by Lemma 1.

Let us establish the solvability of the linear system $F'(x;\mu)\delta x = r$, which reads in detail:

$$\begin{pmatrix} I + b''(y;\mu) & A^* \\ A & -\alpha^{-1} BB^* \end{pmatrix} \begin{pmatrix} \delta y \\ \delta p \end{pmatrix} = \begin{pmatrix} r_a \\ r_s \end{pmatrix}. \tag{16}$$

To formulate a precise result we introduce for $x \in X_{\mathrm{sf}}$ and $\mu > 0$ the following local scaled norm on Y

$$\|\delta y\|_{x,\mu} := \|\delta y\|_\infty + \left\| \sqrt{1 + b''(y;\mu)}\,\delta y \right\|_{L_2(\Omega)}, \tag{17}$$

which depends on y and μ.

Theorem 3 *Let $x \in X_{\mathrm{sf}}$. Then $F'(x;\mu)$ is continuously invertible and in particular $F'(x;\mu)^{-1} r \in \mathrm{dom}\, A \times \mathrm{dom}\, A^*$ for each $r \in X^*$.*

If $r_a \in L_1(\Omega)$ and $r_s = 0$ in (16), then the following estimate holds:

$$\|\delta y\|_{x,\mu} + \left\| \alpha^{-1/2} B^* \delta p \right\|_{L_2(\Omega)} \le C \sup_{v \in Y} \frac{\langle r_a, v \rangle}{\|v\|_{x,\mu}} \le C \|r_a\|_{L_1(\Omega)}. \tag{18}$$

Here C is independent of x and μ.

Proof Consider the quadratic minimization problem:

$$\min_{(y,u)\in Y\times U} \varphi(y,u) := \frac{1}{2}\left\|\sqrt{1+b''}\,y\right\|^2_{L_2} + \frac{\alpha}{2}\|u\|^2_{L_2(Q)} - \langle r_a, y\rangle \quad \text{s.t. } Ay - Bu = r_s. \tag{19}$$

By our assumptions this problem has a unique solution $(\delta y, \delta u) \in Y \times U$. First order optimality conditions for (19) yield existence of a unique $\delta p \in \operatorname{dom} A^*$, such that the equations

$$\begin{aligned}(I + b'')\delta y + A^*\delta p &= r_a\\ \alpha\delta u - B^*\delta p &= 0\\ A\delta y - B\delta u &= r_s\end{aligned}$$

are satisfied. The first row is identical to the first row of (16). The second row yields, $\delta u = \alpha^{-1}B^*\delta p$. Inserting this into the equation $A\delta y - B\delta u = r_s$ yields the second row of (16). Hence $(\delta y, \delta p)$ is the solution of (16) and $F'(x;\mu)$ is bijective and thus continuously invertible by Lemma 1.

Now let $r_s = 0$. Because $\varphi(0) = 0$, we have $\varphi(\delta y, \delta u) \le 0$. Hence, we conclude

$$\left\|\sqrt{1+b''}\,\delta y\right\|^2_{L_2(\Omega)} + \alpha\|\delta u\|^2_{L_2(Q)} \le 2|\langle r_a, \delta y\rangle|$$

and thus, dividing by the square-root of the left hand side, using $\|\delta y\|_\infty \le C\|\delta u\|_{L_2(Q)}$ (which holds, because $A\delta y - B\delta u = 0$ and $A^{-1} : R \to Y$ is continuous by Lemma 1) we obtain

$$\|\delta y\|_{x,\mu} + \alpha^{1/2}\|\delta u\|_{L_2(Q)} \le C\frac{\langle r_a, \delta y\rangle}{\|\delta y\|_{x,\mu}} \le C\sup_{v\in Y}\frac{\langle r_a, v\rangle}{\|v\|_{x,\mu}}. \tag{20}$$

Inserting $\delta p = \alpha^{-1}B^*\delta u$ into (20) yields the first inequality (18). The second inequality follows from $\|\cdot\|_\infty \le \|\cdot\|_{x,\mu}$. □

Remark 5 Observe that $F'(x;\mu)^{-1}$ possesses a strong smoothing property. In particular, $\|\delta y\|_\infty \le C\|r_a\|_{L_1}$. Such a smoothing property is important for the robustness of function space oriented methods. The most popular interior point methods in finite dimensions are primal-dual methods, which introduce additional *algebraic* equations. However, the presence of purely algebraic equations spoils the smoothing property of the inverse Jacobian. In Sect. 5 we propose an algorithmic variant that retains the smoothing property of $F'(x;\mu)^{-1}$, but similarly to primal-dual methods alleviates the nonlinearity of the barrier terms.

Having analyzed invertibility of $F'(x;\mu)$ for fixed x, we now study the dependence of $F'(x;\mu)^{-1}$ on x. Because most parts of $F'(x;\mu)$ are constant we compute

$$F'(x;\mu) - F'(\tilde{x};\mu) = \begin{pmatrix} b''(y) - b''(\tilde{y}) & 0\\ 0 & 0\end{pmatrix} \tag{21}$$

and thus we have the Lipschitz property

$$\begin{aligned}\|F'(x;\mu)-F'(\tilde{x};\mu))\delta x\|_{Y^*\times R} &= \|(b''(y)-b''(\tilde{y}))\delta y\|_{Y^*}\\ &\le \left\|(b''(y)-b''(\tilde{y}))\right\|_{Y^*}\|\delta y\|_\infty\,. \end{aligned} \tag{22}$$

This is the key to the following operator perturbation lemma. We denote the space of continuous mappings $X^* \to X$ by $L(X^*, X)$.

Lemma 2 *The mapping*

$$\begin{aligned} F'(\cdot;\mu)^{-1} : X_{\mathrm{sf}} &\to L(X^*, X)\\ x &\mapsto F'(x;\mu)^{-1}\end{aligned}$$

is continuous.

Proof Let $x \in X_{\mathrm{sf}}$ and consider a sequence $x_k \to x$ in X_{sf}. By continuity of $F'(x;\mu)$ with respect to x due to (22) and by Theorem 3 we conclude

$$T(x;x_k)(\cdot) := F'(x;\mu)^{-1}(F'(x;\mu)-F'(x_k;\mu))(\cdot)$$

is a continuous linear mapping $X \supset \operatorname{dom} A \times \operatorname{dom} A^* \to X$ and can thus be extended continuously and uniquely to a mapping in $L(X, X)$. We even obtain $\|T(x;x_k)\| \to 0$ for $x_k \to x$. This implies via construction of the Neumann series that $(I_X - T(x;x_k))^{-1}$ is well defined for sufficiently large k and that

$$\lim_{k\to\infty}(I_X - T(x;x_k))^{-1} = I_X. \tag{23}$$

Finally, we verify the identity

$$\begin{aligned} F'(x_k;\mu)^{-1} &= \big(I - F'(x;\mu)^{-1}(F'(x;\mu)-F'(x_k;\mu))\big)^{-1}F'(x;\mu)^{-1}\\ &= (I - T(x;x_k))^{-1}F'(x;\mu)^{-1}\end{aligned}$$

by left-multiplication with $F'(x_k;\mu)$. By (23) we obtain

$$\lim_{k\to\infty} F'(x_k;\mu)^{-1} = F'(x;\mu)^{-1}$$

and thus the desired result. □

4.2 Local convergence of Newton's method

We consider a Newton step, which maps $x \in D_{\mathrm{sf}}$ to

$$x_+ := x - F'(x;\mu)^{-1}F(x;\mu) \in X.$$

We thus obtain a mapping

$$\begin{aligned} &N(\cdot;\mu) : X \supset D_{\text{sf}} \to X \\ &x \mapsto x_+ := x - F'(x;\mu)^{-1}F(x;\mu). \end{aligned} \tag{24}$$

The starting point of our consideration is that we can use $F(x(\mu);\mu) = 0$ to write

$$x_+ - x(\mu) = F'(x;\mu)^{-1}\underbrace{\big(F'(x;\mu)(x - x(\mu)) - (F(x;\mu) - F(x(\mu);\mu))\big)}_{r(x)} \tag{25}$$

and compute

$$r(x) = \begin{pmatrix} r_a(y) \\ 0 \end{pmatrix} = \begin{pmatrix} b''(y;\mu)(y - y(\mu)) - (b'(y;\mu) - b'(y(\mu);\mu)) \\ 0 \end{pmatrix}. \tag{26}$$

We observe that in this expression the operators A and A^* have disappeared. This allows us to establish a surprising result, namely that the mapping N has a unique continuous extension from D_{sf} to X_{sf}.

Proposition 3 *The mapping $N(\cdot;\mu)$, defined in (24) has a unique continuous extension to a mapping $\overline{N}(\cdot;\mu) : X_{\text{sf}} \to X$. For all $x \in X_{\text{sf}}$, $x_+ = \overline{N}(x;\mu) \in$* dom $A \times$ dom A^*.

Proof First of all we note that D_{sf} is dense in X_{sf}, because dom $A \times$ dom A^* is dense in $Y \times R^*$ and X_{sf} is open.

By the discussion above we have seen that $F'(x;\mu)^{-1}r(x)$ is not only well defined for $x \in D_{\text{sf}}$, but for all $x \in X_{\text{sf}}$. Hence, the extended mapping

$$\overline{N}(\cdot;\mu) : x \to x_+ := F'(x;\mu)^{-1}r(x) + x(\mu)$$

is well defined for all $x \in X_{\text{sf}}$ and coincides with $N(\cdot;\mu)$ on D_{sf}. Moreover, by Theorem 3 and since $x(\mu) \in$ dom $A \times$ dom A^* we have $x_+ \in$ dom $A \times$ dom A^*.

It remains to show that $\overline{N}(\cdot;\mu)$ is continuous at x for all $x \in X_{\text{sf}}$, which implies that it is the unique continuous extension of $N(\cdot;\mu)$. To show this, chose a sequence $x_k \to x$ in X_{sf}. By continuity of r we have $r(x_k) \to r(x)$. We compute

$$\overline{N}(x_k;\mu) = (F'(x_k;\mu)^{-1} - F'(x;\mu)^{-1})r(x_k) + F'(x;\mu)^{-1}r(x_k) + x(\mu).$$

By Lemma 2 and boundedness of $r(x_k)$, $(F'(x_k;\mu)^{-1} - F'(x;\mu)^{-1})r(x_k) \to 0$, as $x_k \to x$. Moreover, $F'(x;\mu)^{-1}r(x_k) \to F'(x;\mu)^{-1}r(x)$ by Theorem 3 and $r(x_k) \to r(x)$. This shows continuity of $\overline{N}(\cdot;\mu)$ at x. □

Remark 6 In applications, where A and A^* are differential operators and dom $A \times$ dom A^* is a set of weakly differentiable functions, Proposition 3 states that Newton's method does not rely on smoothness of the iterates. This reflects the fact that pointwise state constraints lead to pointwise nonlinearities. We will use this additional freedom in Sect. 5 below to introduce a pointwise modification of Newton's method.

In the next step we relate convergence of Newton's method to an analytic quantity $\Theta(x;\mu)$, for which we are going to derive a-priori estimates in this section and a-posteriori estimates in Sect. 6. For a simple characterization of local convergence we will use the following scaled norm

$$\|\delta x\|_\mu := \|\delta y\|_{x(\mu),\mu} + \left\|\alpha^{-1/2} B^*\delta p\right\|_{L_2(Q)}, \tag{27}$$

which only depends on μ. Observe that $\delta u := \alpha^{-1}B^*\delta p$ corresponds to an element of the control space U, so that this norm indirectly reflects also convergence in u.

Theorem 4 *For all $x \in D_{\mathrm{sf}}$ the quantity*

$$\Theta(x;\mu) := \frac{\left\|F'(x;\mu)^{-1}\big(F'(x;\mu)(x-x(\mu)) - (F(x;\mu) - F(x(\mu);\mu))\big)\right\|_\mu}{\|x-x(\mu)\|_\mu} \tag{28}$$

is well defined and extends continuously to X_{sf}. The following contraction estimate holds

$$\|x_+ - x(\mu)\|_\mu = \Theta(x;\mu)\,\|x-x(\mu)\|_\mu\,. \tag{29}$$

Proof Using (25) and (28) we have

$$\begin{aligned}\|x_+ - x(\mu)\|_\mu &= \left\|F'(x;\mu)^{-1}\big(F'(x;\mu)(x-x(\mu)) - (F(x;\mu) - F(x(\mu);\mu))\big)\right\|_\mu\\ &= \|F'(x;\mu)^{-1}r(x)\|_\mu = \Theta(x;\mu)\,\|x-x(\mu)\|_\mu\,,\end{aligned}$$

which yields (29) for all $x \in D_{\mathrm{sf}}$ and by extension also for all $x \in X_{\mathrm{sf}}$. □

Equation (29) gives us the interpretation of

$$\Theta(x;\mu) = \frac{\|x_+ - x(\mu)\|_\mu}{\|x-x(\mu)\|_\mu}$$

as a local Newton contraction. If $\Theta(x;\mu) \le \tilde{\Theta} < 1$ in a neighborhood of $x(\mu)$, then Newton's method converges locally to $x(\mu)$. We will show now $\Theta(x;\mu) = O(\|x-x(\mu)\|_\mu)$ in a suitable neighborhood of $x(\mu)$, which implies local quadratic convergence of Newton's method for each $\mu > 0$. Proving local quadratic convergence alone, however, would not be sufficient in the context of path-following. In addition we need a more quantitative result that relates $\Theta(x;\mu)$ to μ and yields bounds from below on the *radius of convergence* to conclude convergence of the overall path-following method.

Lemma 3 *Let $z_1, z_2 > 0$, and $\tilde{z} := \min\{z_1; z_2\}$. For the barrier function $l(z) = l(z; \mu; q)$ the following pointwise estimate holds:*

$$|l''(z_1)(z_1 - z_2) - (l'(z_1) - l'(z_2))| \leq \frac{c}{\tilde{z}} |l''(\tilde{z})(z_1 - z_2)^2|. \tag{30}$$

The constant c depends only on q.

Proof Since l' is a sum of functions of the form $\mu^q y^{-q}$, it is twice differentiable for positive z and all derivatives are monotonically decreasing in absolute value. Hence, application of the fundamental theorem of calculus twice yields (30), taking into account the rules of differentiation. □

Let now $x(\mu) = (y(\mu), p(\mu))$ be a solution of $F(x; \mu) = 0$ for some fixed $\mu > 0$, which corresponds to a barrier minimizer by Proposition 1. By Proposition 2 there is $\psi(\mu) > 0$ such that

$$y(\mu)(t) - \underline{y}(t) \geq \psi(\mu) \quad \forall t \in \overline{\Omega}$$

We introduce a strictly feasible neighborhood X_μ of the central path by

$$X_\mu = \{x \in X : \|y - y(\mu)\|_\infty \leq \rho\psi(\mu)\}, \tag{31}$$

for some fixed $0 < \rho < 1$. By the triangle inequality the equivalence relation

$$(1 - \rho)(y(\mu) - \underline{y})(t) \leq (y - \underline{y})(t) \leq (1 + \rho)(y(\mu) - \underline{y})(t) \quad \forall t \in \overline{\Omega} \tag{32}$$

holds for all $x \in X_\mu$, which we will use often in the following. This helps us to derive *a-priori* estimates for $\Theta(x; \mu)$. In particular, the norms $\|\cdot\|_{x;\mu}$ and $\|\cdot\|_{x(\mu);\mu}$ are equivalent for $x \in X_\mu$.

For our analysis we will choose the sequence μ_k such that all x_k and x_{k+1} remain in X_{μ_k}. In a practical algorithm, where *a-posteriori* estimates are available, this neighborhood can be dropped.

Proposition 4 *For each $\mu > 0$ Newton's method converges locally quadratically to the solution $x(\mu)$. More precisely, there is a positive function $\omega(\mu)$, which is bounded on every compact positive interval, such that for $x \in X_\mu$*

$$\Theta(x; \mu) \leq \frac{1}{2}\omega(\mu)\|x - x(\mu)\|_\mu, \tag{33}$$

together with the bound $\omega(\mu) \leq c_\omega \psi^{-1}(\mu)$.

Proof In the following c denotes a generic constant that may vary from step to step. By definition of Θ and by (26)

$$\Theta(x; \mu) = \frac{\|x_+ - x(\mu)\|_\mu}{\|x - x(\mu)\|_\mu} = \frac{\left\|F'(x; \mu)^{-1} r(x)\right\|_\mu}{\|x - x(\mu)\|_\mu},$$

with first component $r_a(y) = b''(y)(y - y(\mu)) - (b'(y) - b'(y(\mu)))$. Lemma 3 gives us a pointwise estimate for $\tilde{y} = \min\{y(\mu), y\}$:

$$|r_a(y)|(t) \leq \frac{c}{\tilde{y} - \underline{y}} \left| b''(\tilde{y})(y - y(\mu))^2 \right| (t),$$

and by the equivalence relation (32)

$$|r_a(y)|(t) \leq \frac{c}{y(\mu) - \underline{y}} \left| b''(y(\mu))(y - y(\mu))^2 \right| (t).$$

Then

$$\begin{aligned} \|r_a(y)\|_{L_1(\Omega)} &\leq \int_\Omega \frac{c}{y(\mu) - \underline{y}} |b''(y(\mu))(y - y(\mu))^2| \, dt \\ &\leq \left\| \frac{c}{y(\mu) - \underline{y}} \right\|_\infty \|y - y(\mu)\|_{x,\mu}^2 \leq c\psi(\mu)^{-1} \|y - y(\mu)\|_{x,\mu}^2 . \end{aligned}$$

Hence, Theorem 3 and the equivalence of norms yields

$$\begin{aligned} \|\delta x\|_\mu &\leq c \|\delta y\|_{x,\mu} + \left\| \alpha^{-1/2} B^* \delta p \right\|_{L_2(Q)} \leq c \|r_a(y)\|_{L_1(\Omega)} \\ &\leq c\psi(\mu)^{-1} \|y - y(\mu)\|_{x,\mu}^2 \leq c\psi(\mu)^{-1} \|x - x(\mu)\|_\mu^2 . \end{aligned}$$

This implies (33), which in turn yields local quadratic convergence by Theorem 4. □

Remark 7 Inserting (33) into Theorem 4 yields an L_∞-radius of the region of convergence proportional to $\psi(\mu)$ – the distance of $y(\mu)$ to its lower bound $\underline{y}$. This is a result as good as one can expect in an L_∞-framework, because the radius of convergence is restricted anyway by the distance of $y(\mu)$ to the bound. Unfortunately, $\psi(\mu)$ may shrink rapidly for $\mu \to 0$, which makes our estimated radius of convergence shrink rapidly, too.

An attempt to overcome this difficulty would be to introduce a scaled L_∞-norm of the form $\|b''(y)^{1/2} \cdot \|_\infty$. In finite dimensional settings this norm can be bounded in terms of $\| \cdot \|_{x,\mu}$ using the general result that all norms are equivalent on $\mathbb{R}^n$ up to a factor of $O(\sqrt{n})$. This is a method of proving linear convergence for interior point methods in $\mathbb{R}^n$, but of course, such an argument is not applicable in our infinite dimensional setting. Here the link between L_2 and L_∞ has to be established via PDE regularity results. However, no such results are available in terms of scaled L_∞-norms.

4.3 Convergence of a simple path-following algorithm

In this section we consider convergence of Algorithm 1. We will show that a choice μ_k is possible, such that Algorithm 1 is well defined and converges to the optimal solution of the problem. In Sect. 6 we describe how to choose the sequence μ_k in practice, based on a-posteriori quantities.

Algorithm 1

```
select $\mu_0 > 0$, and $x_0 \in X_{\mathrm{sf}}$
for $k = 0, \ldots$
    $x_{k+1} := \overline{N}(x_k; \mu_k)$    (extended Newton step, cf. Proposition 3)
    select $\mu_{k+1}$
```

Let us recapitulate the information on the non-linearity of our problem gathered so far. The first piece of information is an estimate of the local Lipschitz constant of the homotopy path $\mu \to x(\mu)$. From (8) and (9) we obtain

$$\|x(\mu) - x(\nu)\|_\mu \leq \eta(\mu)(\mu - \nu) \text{ with } \eta(\mu) \leq c_\eta \mu^{-1/2} \tag{34}$$

if $\mu \geq \nu \geq \mu/2$. The second piece of information (33) characterizes local convergence of Newton's method and can be written as

$$\|x_+ - x(\mu)\|_\mu \leq \frac{\omega(\mu)}{2}\|x - x(\mu)\|_\mu^2 \text{ with } \omega(\mu) \leq c_\omega \psi(\mu)^{-1}. \tag{35}$$

For each μ this estimate holds, as long as $x \in X_\mu$. This is satisfied in particular, if

$$\|x - x(\mu)\|_\mu \leq r(\mu) \text{ with } r(\mu) := \min\{\rho\psi(\mu), c_\omega^{-1}\psi(\mu), \sqrt{\mu}\}.$$

Note that $r(\mu)$ is a continuous function in μ on $]0, \mu_0]$ for all $\mu_0 > 0$, since ψ is continuous in μ. Moreover, by Proposition 4 we observe that

$$\|x - x(\mu)\|_\mu \leq r(\mu) \;\Rightarrow\; \|x_+ - x(\mu)\|_\mu \leq r(\mu)/2$$

and thus $x_+ \in X_\mu$ again. The inequality $r(\mu) \leq \sqrt{\mu}$ forces convergence of the iterates in the (trivial) case that $\psi(\mu)$ is bounded away from zero.

Finally, we need information on the dependence of $\|\cdot\|_\mu$, defined in (27), on μ.

Lemma 4 *For each $\mu, \nu > 0$ there is a constant $\gamma(\mu, \nu)$ such that the following norm estimate holds:*

$$\|\delta x\|_\nu \leq \gamma(\mu, \nu)\, \|\delta x\|_\mu \, \forall \delta x \in X. \tag{36}$$

The constant $\gamma(\mu, \nu)$ depends continuously on (μ, ν), and $\gamma(\mu, \mu) = 1$.

Proof For arbitrary $\delta x \in X$ we can write

$$\|\sqrt{1 + b''(y(\nu); \nu)}\delta y\|_{L_2(\Omega)} \leq \left\| \frac{\sqrt{1 + b''(y(\nu); \nu)}}{\sqrt{1 + b''(y(\mu); \mu)}} \right\|_\infty \|\sqrt{1 + b''(y(\mu); \mu)}\delta y\|_{L_2(\Omega)}.$$

and we set

$$\gamma(\mu, \nu) := \max\left\{1, \left\| \frac{\sqrt{1 + b''(y(\nu); \nu)}}{\sqrt{1 + b''(y(\mu); \mu)}} \right\|_\infty \right\}.$$

Obviously, $\gamma(\mu,\mu)=1$ and γ depends continuously on (μ,ν) because $\mu\to y(\mu)$ is continuous w.r.t. $\|\cdot\|_\infty$ and $y(\mu)$ is strictly feasible. By definition of $\|\cdot\|_\mu$ it is now easy to see that (36) holds. □

If we connect these assertions we can show that there is a sequence μ_k such that Algorithm 1 produces iterates x_k that remain close to the central path and converge to the solution of the original problem. With this method it is possible in principle but rather technical to compute a rate of convergence. In the context of pure control constraints this has been done in [13].

Theorem 5 *Suppose Assumptions 1–3 hold, and choose $q\ge d/\beta$ as the order of the barrier function. Assume further that initial guesses $x_0\in X_{\mathrm{sf}}$ and $\mu_0>0$ are given such that*

$$\|x_0-x(\mu_0)\|_\mu\le r(\mu_0).$$

Then there is a sequence $\mu_k\to 0$, such that Algorithm 1 is well defined and produces a converging sequence x_k. Setting $u_k:=\alpha^{-1}B^*p_k$, the sequence (y_k,u_k) converges to the solution (y_*,u_*) of (1) and there is a constant C such that:

$$\|y_k-y_*\|_\infty+\|u_k-u_*\|_U\le C\sqrt{\mu_k}. \tag{37}$$

Proof Let $\mu>0$, $\|x-x(\mu)\|_\mu\le r(\mu)$, and denote by x_+ the result of one Newton step at $(x;\mu)$. We study an update of μ of the form $\sigma\mu$ for $\sigma\in[1/2,1[$.

Due to Proposition 4 and by construction of $r(\mu)$ we have $\|x_+-x(\mu)\|_\mu\le 1/2r(\mu)$. Moreover, using the definition of γ from Lemma 4 we have:

$$\begin{aligned}\|x_+-x(\sigma\mu)\|_{\sigma\mu}&\le\gamma(\mu,\sigma\mu)\,\|x_+-x(\sigma\mu)\|_\mu\\&\le\gamma(\mu,\sigma\mu)(\|x_+-x(\mu)\|_\mu+\|x(\mu)-x(\sigma\mu)\|_\mu)\\&\le\gamma(\mu,\sigma\mu)\left(\frac{1}{2}r(\mu)+c_\eta\mu^{-1/2}(\mu-\sigma\mu)\right).\end{aligned}$$

We want to choose σ such that $\|x_+-x(\sigma\mu)\|_{\sigma\mu}\le r(\sigma\mu)$. Defining

$$\phi_\mu(\sigma):=\gamma(\mu,\sigma\mu)\left(\frac{1}{2}r(\mu)+c_\eta\mu^{-1/2}(\mu-\sigma\mu)\right)-r(\sigma\mu).$$

this is equivalent to the requirement $\phi_\mu(\sigma)\le 0$.

Since r and γ are continuous, $\phi_\mu(\sigma)$ is a continuous function in σ, and we can compute $\phi_\mu(1)=-\frac{1}{2}r(\mu)<0$. Hence, by continuity of ϕ_μ there is a smallest $\sigma\in[1/2,1[$, such that $\phi_\mu(\sigma)\le 0$ and it is easy to see that either $\sigma=1/2$, or $\phi_\mu(\sigma)=0$.

Starting with some $\mu_0>0$ and $\|x_0-x(\mu_0)\|_\mu\le r(\mu_0)$, successive choice of σ_k according to the above rule and application of one Newton step yields a well defined algorithm with $\|x_k-x(\mu_k)\|_{\mu_k}\le r(\mu_k)$ by induction. Since $\mu_{k+1}\le\mu_k$, the sequence μ_k is decreasing, and thus convergent. We will show by contradiction that $\mu_k\to 0$.

Otherwise there would be some $\mu_* > 0$ with $\mu_k \to \mu_*$. Then $\lim_{k\to\infty} \sigma_k = 1$, and thus $\phi_{\mu_k}(\sigma_k) = 0$ for all sufficiently large k. However, by continuity $r(\sigma_k \mu_k) \to r(\mu_*)$, and by Lemma 4 $\gamma(\mu_k, \sigma_k \mu_k) \to 1$. Thus

$$\phi_{\mu_k}(\sigma_k) = \gamma(\mu_k, \sigma_k \mu_k) \left(\frac{1}{2} r(\mu_k) + c_\eta \mu_k^{-1/2} (\mu_k - \sigma_k \mu_k) \right) \to \frac{1}{2} r(\mu_*) - r(\mu_*),$$

This yields a contradiction, since $r(\mu_*) \neq 0$.

To show (37), we compute

$$\sqrt{\mu} \geq r(\mu_k) \geq \|x(\mu_k) - x_k\|_\mu \geq \|y(\mu_k) - y_k\|_\infty + \|\alpha^{-1/2} B^*(p(\mu_k) - p_k)\|_{L_2(Q)}$$
$$= \|y(\mu_k) - y_k\|_\infty + \|\alpha^{1/2}(u(\mu_k) - u_k)\|_{L_2(Q)}.$$

Now (37) follows from (7) via the triangle inequality. □

5 A pointwise modification

Barrier methods rely on iterates that are feasible with respect to the inequality constraints. Since in the barrier context Newton's method approximates a rational function by a linear one, Newton steps tend to violate the constraints. This issue should be addressed algorithmically. Otherwise, this may restrict the speed of convergence of practical algorithms.

In the following we propose a modification of Newton's method, which may be considered as a pointwise damping strategy. The idea exploits the pointwise structure of the problem and guarantees feasibility of the iterates. In the whole discussion $\mu > 0$ is fixed.

Consider the first row of the Newton equation $F(x; \mu) + F'(x; \mu)(x_+ - x) = 0$. It is posed in Y^* and reads

$$y - y_d + b'(y) + A^* p + (I + b''(y))(y_+ - y) + A^*(p_+ - p) = 0. \qquad (38)$$

Our principle idea is to construct a modified feasible iterate $y_C > 0$ that satisfies

$$y_C - y_d + b'(y_C) + A^* p_+ = 0. \qquad (39)$$

Unfortunately, this equation is not well defined, because $A^* p_+$ is not necessarily a function. In particular, in the context of finite elements and weak formulations (39) cannot be interpreted as a pointwise equation.

However, subtraction of (38) and (39) yields a pointwise equation for y_C that depends on y and y_+:

$$y_C - y + b'(y_C) - b'(y) = (1 + b''(y))(y_+ - y). \qquad (40)$$

If p_+ is sufficiently smooth, then (39) and (40) are equivalent. Hence, (40) extends (39) for general p_+. The idea is now to solve this equation pointwise, but to use only those y_C, for which $|y_C - y| \leq |y_+ - y|$. We obtain a pointwise damping step, which enables us to compute a *strictly* feasible corrected iterate y_C from a possibly infeasible iterate y_+ in a *natural way*.

In the case of a logarithmic barrier function (40) is a quadratic equation in y_C and can be solved explicitly as such. For rational barrier functions we may use iterative techniques, of which bisection is the simplest. Because this computation is a pointwise operation to be performed at each node of the discretization, its contribution to the overall computational effort is marginal.

5.1 Interpretation as a pointwise damped primal correction

In the following lemma we gather the basic properties of our pointwise modification. In view of (40) we consider for given $z \in]0, \infty[$ and $z_+ \in \mathbb{R}$ existence and properties of a solution z_C of the equation

$$z_C - z + l'(z_C) - l'(z) = (1 + l''(z))(z_+ - z). \tag{41}$$

Lemma 5 *Let $z \in]0, \infty[$ and $z_+ \in \mathbb{R}$. Then (41) admits a unique solution $z_C > 0$.*

If $z_+ \geq z$, then $z \leq z_+ \leq z_C$.
If $z_+ \leq z$, then $z_+ \leq z_C \leq z$ and

$$|z_C - z_+| \leq c\frac{z^{q+1}}{z_C^{q+2}}|z - z_+|^2. \tag{42}$$

Proof On $]0, \infty[$ the function $f(z_C) := z_C + l'(z_C)$ is well defined, monotonically increasing, continuous, $\lim_{z_C \to 0} = -\infty$, and $\lim_{z_C \to \infty} = +\infty$. By the mean value theorem this implies unique solvability of the equation $f(z_C) = r$ for any $r \in \mathbb{R}$ with $z_C \in]0, \infty[$. By (40) and the fundamental theorem of calculus:

$$(1 + l''(z))(z_+ - z) = f(z_C) - f(z) = \int_z^{z_C} (1 + l''(\zeta))d\zeta \ (z_C - z). \tag{43}$$

Since l'' is monotonically decreasing in z, this relation yields $z_+ \leq z_C \leq z$ for $z_+ \leq z$ and $z \leq z_+ \leq z_C$ otherwise.

Let now $z_+ \leq z$. We may rewrite (43) as

$$z_C - z_+ = l''(z)(z_+ - z) - (l'(z_C) - l'(z)),$$

and hence

$$(z_C - z_+)(1 + l''(z)) = l''(z)(z_C - z) - (l'(z_C) - l'(z)),$$

and thus (42) follows from $z_C = \min\{z, z_C\}$ and Lemma 3

$$|z_C - z_+| \le \frac{1}{1+l''(z)}|l''(z)(z_C - z) - (l'(z_C) - l'(z))| \le c\frac{|l''(z_C)|}{(1+l''(z))z_C}|z_C - z|^2$$
$$\le c\frac{z^{q+1}}{z_C^{q+2}}|z_C - z|^2 \le c\frac{z^{q+1}}{z_C^{q+2}}|z_+ - z|^2.$$

□

In order to obtain a *damping* strategy we use (40) only for those $t \in \overline{\Omega}$ for which $y_+(t) \le y(t)$ and define

$$y_D(t) := \begin{cases} y_+(t) : y_+(t) \ge y(t) \\ y_C(t) : y_+(t) < y(t) \end{cases}$$

Further, we define $x_D := (y_D, p_+)$. In this case Lemma 5 asserts that $y_+ \mapsto y_D$ is indeed a pointwise *damping*, i.e.

$$|y_D - y|(t) \le |y_+ - y|(t) \quad \forall t \in \overline{\Omega}.$$

Moreover, (42) implies

$$\lim_{\|y - y(\mu)\|_\infty \to 0} \frac{\|y_D - y_+\|_\infty}{\|y - y_+\|_\infty} = 0.$$

This means that undamped steps are recovered asymptotically close to a (strictly feasible) solution.

Theorem 6 (Convergence Theorem for Damping) *The conclusions of Theorem 5 remain valid, if a pointwise damping step (40) is used for all $t \in \Omega$ for which $y_+(t) \le y(t)$.*

Proof By Proposition 3 extended Newton steps are well defined for all $x \in X_{\mathrm{sf}}$. Hence, a pointwise modification never leads to undefined Newton steps. If $x, x_+ \in X_\mu$ as defined in (31), then (42) yields

$$\|x_D - x_+\|_\infty \le \min\{1; c\rho\}\|x - x_+\|_\infty$$

and thus for small ρ the damping step is only a small perturbation of the undamped Newton step. Thus, the proof of Theorem 5 carries over. □

We will see in our numerical experiments that in practice the pointwise damped variant is far more efficient than the undamped version.

5.2 Interpretation as a blended primal-dual correction

The pointwise modification (39) has another useful interpretation in terms of a dual method. Let us introduce the variable v, defined by

$$v(t) = y(t) + l'(y(t); \mu).$$

We can solve this equation for y to obtain a nonlinear function $y(v)$ with derivative

$$y_v(v) = \big(I + b''(y(v);\mu)\big)^{-1},$$

and a nonlinear system of equations in the variables p and v:

$$\begin{aligned} v - y_d + A^*p &= 0 \\ Ay(v) - \alpha^{-1}BB^*p &= 0, \end{aligned}$$

which we will abbreviate by $\tilde{F}(\tilde{x};\mu)$, setting $\tilde{x} := (v, p)$. In contrast to F, which is only defined for $\underline{y} \le y \le \overline{y}$, $\tilde{F}$ is well defined for all sufficiently smooth $\tilde{x}$. Its (again formal) linearization is given by

$$\begin{aligned} \tilde{F}'(\tilde{x};\mu) &:= \begin{pmatrix} I & A^* \\ Ay_v(v) & -\alpha^{-1}BB^* \end{pmatrix} \\ &= \begin{pmatrix} I + b'' & A^* \\ A & -\alpha^{-1}BB^* \end{pmatrix} \begin{pmatrix} \big(I + b''\big)^{-1} & 0 \\ 0 & I \end{pmatrix}. \end{aligned} \tag{44}$$

Although this formulation has the advantage of guaranteed feasibility, it suffers from its poor regularity properties. The presence of the nonlinearity $Ay(v)$ in $\tilde{F}$ makes a pure dual method rather unstable.

If we consider the factorization of $\tilde{F}'(\tilde{x};\mu)^{-1}$ in (44), we see that performing one Newton step for $F(x;\mu)$ and one correction of the form (40) is equivalent to performing one Newton step for $\tilde{F}(\tilde{x};\mu)$ and computing $y(v)$.

Hence, by our damping strategy we implicitly compute both the primal and the dual Newton step and take the pointwise minimum in absolute value to obtain a blended correction that has the favorable properties of both methods and avoids their problems.

6 An adaptive path-following scheme

In this section we consider the construction of a practical path-following algorithm that adaptively chooses the sequence μ_k. Our starting point is again the system (13). Hence, our algorithm uses the state y and the adjoint state p as iteration variables.

For efficient path-following several extensions of Algorithm 1 are useful. Most importantly, we have to provide a practical criterion for the choice of the sequence μ_k. If this choice is made, then performing one single Newton step is a too rigid concept in practice. Rather, one should aim for some (loose) convergence criterion. If the choice of μ_k was too aggressive, it may be useful to reject the path-following step and select μ_k more conservatively, based on more accurate information. This leads to Algorithm 2.

Our considerations are based on the ideas of [4, Chapter 5]. The main idea is to introduce parameterized models for the quantities $\eta(\mu)$ and $\Theta(x;\mu)$ (and closely related $\omega(\mu)$) used in the a-priori analysis and supply computational estimates for the parameters. This strategy guarantees a close connection between a-priori results and the algorithmic realization and helps to take into account the special structure of the function space problem. In the following we will introduce these models. For additional details we refer to [13, Section 8.2].

Algorithm 2

```
select $\mu_0 > 0$, and $x_0$ with $\|x_0 - x(\mu_0)\| \le r(\mu_0)$, $k := 0$
do (Homotopy Method)
    $\tilde{x}_0 = x_k$, $j = 0$
    do (Newton Corrector)
        $\delta\tilde{x}_j \leftarrow$ compute pointwise damped Newton step
        (failure,converged) $\leftarrow$ estimate Newton contraction
        if(not failure) $\tilde{x}_{j+1} := \tilde{x}_j + \delta\tilde{x}_j$
        j := j+1
    while not(converged or failure)
    if(converged)
        $x_{k+1} := \tilde{x}_{j+1}$
        $\mu_{k+1} \leftarrow$ predict new step size
        k := k+1
    if(failure)
        $\mu_k \leftarrow$ reduce step size
while(termination criterion not reached)
```

For the evaluation of our algorithmic quantities we have to choose a norm. Our convergence theory and numerical experience suggest to use a scaled local norm, similar to $\|\cdot\|_{x,\mu}$ defined in (17). Experience shows that it is favorable in practice to drop the $\|\cdot\|_\infty$-part and use the following L_2-type norm:

$$\|\delta x\|^2 := \|\delta x\|^2_{x,\mu,2} := \left\|\sqrt{1 + b''(x;\mu)}\delta y\right\|^2_{L_2(\Omega)} + \alpha^{-1/2}\left\|B^*\delta p\right\|^2_{L_2(Q)}$$

It is of practical importance that this norm can be evaluated easily and accurately. In particular, we do not rely on the evaluation of norms of residuals. Corresponding residual norms would be dual norms, which are cumbersome and expensive to evaluate.

Following the ideas of [18] it is possible and useful to take into account algorithmically that our scaled norm depends on μ and x. We will, however, neglect this issue here for simplicity.

6.1 A model of the central path

A direct way to estimate the Lipschitz constant η would be using finite differences:

$$[\eta]_{Exact}(\mu_k) := \frac{\|x(\mu_k) - x(\mu_{k+1})\|}{|\mu_k - \mu_{k+1}|}.$$

However, because exact barrier minimizers are not available, we replace them by finite differences of our computational values:

$$[\eta](\mu_k) := \frac{\|x_k - x_{k+1}\|}{|\mu_k - \mu_{k+1}|}. \tag{45}$$

The closer x_k and $x(\mu_k)$, the more accurate the estimate $[\eta]$. With this estimate (34) suggests to model $\eta(\mu)$ by

$$\eta_M(\mu) := [\eta](\mu_k)\left(\frac{\mu}{\mu_k}\right)^{-1/2}. \tag{46}$$

Once $[\eta](\mu)$ is computed we may also use it to estimate the remaining length of the central path via integration of $\eta_M(\mu)$ on $[0, \mu_k]$:

$$\|x_{\mu_k} - x_*\| \approx 2[\eta](\mu_k)\cdot\mu_k. \tag{47}$$

6.2 Estimating the Newton contraction

The main step of Algorithm 2 is the evaluation of a Newton step that defines the corrector: $\tilde{x}_{j+1} \leftarrow \tilde{x}_j$. To capture the behavior of Newton's method we derive a model for the contraction $\Theta(x;\mu)$. This is done, in a way that uses only quantities, defined on the domain space of F, and this model can thus be called affine covariant. The role of affine covariance for the control of algorithms based on Newton's method has been pointed out in [4]. In the context of partial differential equations affine covariance has the important advantage that the hard to compute residual norm $\|\cdot\|_{X^*}$ is not needed.

To derive a model for $\Theta(x;\mu)$ we consider (28), which is of course computationally unavailable, because $x(\mu)$ is unknown. However, after one Newton step $x \to x_+$ we may replace $x(\mu)$ by x_+ in (28) and obtain

$$\begin{aligned}[\Theta]^N(x;\mu) :&= \frac{\left\|F'(x;\mu)^{-1}\big(F'(x;\mu)(x-x_+) - (F(x;\mu) - F(x_+;\mu))\big)\right\|}{\|x-x_+\|}\\ &= \frac{\left\|(x - F'(x;\mu)^{-1}F(x;\mu)) - (x_+ - F'(x;\mu)^{-1}F(x_+;\mu))\right\|}{\|x-x_+\|}\\ &= \frac{\|x_+ - \overline{x}_+\|}{\|x-x_+\|},\end{aligned}$$

where $\overline{x}_+$ is the result of a simplified Newton step $\overline{x}_+ := x_+ - F'(x;\mu)^{-1}F(x_+;\mu)$.

If we use our pointwise damping strategy we have to design a modification of $[\Theta](x;\mu)$ in terms of x_D, because x_+ may be infeasible. The most obvious modification is to replace the result of a Newton step $x \to x_+$ by a pointwise damped Newton step $x \to x_D$, and the simplified Newton step at $x_+ \to \overline{x}_+$ by a pointwise damped simplified Newton step $x_D \to \overline{x}_D$:

$$[\Theta]^{PD}(x;\mu) := \frac{\|x_D - \overline{x}_D\|}{\|x - x_D\|}.$$

By (42) the pointwise damped Newton steps merge into ordinary Newton steps close to the solution, and the same holds for the simplified versions. So $[\Theta]^{PD}(x;\mu)$ is asymptotically equivalent to $[\Theta]^N(x;\mu)$ close to the solution.

If $[\Theta]$ is small enough, then the simplified Newton step needed for this evaluation can be used to improve the quality of the solution. If direct sparse solvers are used for the linear equations, then the additional computational effort needed is rather small. We only have to assemble another right hand side and perform one forward-backward substitution. In our implementation we even perform a second simplified Newton step in case of very small $[\Theta]$, because this improves the efficiency slightly.

If iterative solvers are used, then the relative additional effort for a simplified Newton step depends on the relation between assembly of the stiffness matrix, construction of a preconditioner and iterative solution. Also in this case the obtained simplified Newton step can be used to improve the solution, or as a starting value for the next iterative solution process.

Let us now concentrate on $[\Theta]^{PD}$. Since Newton's method is a *locally* convergent method it may happen that for bad initial values the iteration diverges. This happens, if $\Theta(x;\mu) > 1$. So it is natural to terminate the Newton corrector with failure, if $[\Theta]^{PD}(x;\mu) > 1$ in order to start a correction step with a more conservative choice of μ_k.

Next we derive a convergence criterion for the Newton corrector. Motivated by the triangle inequality $\|x - x(\mu)\| \le \|x - x_D\| + \|x_D - x(\mu)\|$ we may estimate the error $e(x;\mu) := \|x - x(\mu)\|$ by setting

$$[e](x;\mu) := \left(1 + [\Theta]^{PD}\right)\|x - x_D\|, \quad [e](x_D;\mu) := [\Theta]^{PD}\left(1 + [\Theta]^{PD}\right)\|x - x_D\|.$$

If $[e](x_C;\mu)$ is sufficiently small, then the Newton corrector is terminated successfully. A useful convergence criterion for a Newton correction method is to require

$$[e](\tilde{x}_{j+1};\mu) \le \Lambda \left\|\tilde{x}_0 - \tilde{x}_{j+1}\right\|, \tag{48}$$

which means that the corrector has reduced the error about a factor Λ. Useful choices are in a range of 0.01–0.5.

6.3 Step size selection

Proposition 4 suggests to model the Newton contraction $\Theta(x;\mu)$ by

$$\Theta_M(x;\mu) := \omega_M(\mu)\|x - x(\mu)\|, \tag{49}$$

where we—similarly to η_M—define

$$\omega_M(\mu) := [\omega](\mu_k)\left(\frac{\mu}{\mu_k}\right)^{-1/2} \tag{50}$$

and, again replacing $x(\mu)$ by x_C:

$$[\omega](\mu_k) := \frac{[\Theta]^{PD}(x;\mu)}{\|x - x_D\|}. \tag{51}$$

Assume first that the corrector for μ_k has terminated successfully. By the triangle inequality we have

$$\|x_{k+1} - x(\mu_{k+1})\| \le \|x_{k+1} - x(\mu_k)\| + \|x(\mu_{k+1}) - x(\mu_k)\|$$

Recalling that $x_{k+1} = \tilde{x}_{j+1}$ the first summand is estimated by

$$\|x_{k+1} - x(\mu_k)\| \approx [e](x_{k+1};\mu_k),$$

while the second summand is estimated via (46) as

$$\|x(\mu_{k+1}) - x(\mu_k)\| \approx \eta_M(\mu_{k+1})|\mu_k - \mu_{k+1}|.$$

To compute μ_{k+1} it is sensible to aim for a certain contraction Θ_d, supplied by the user, which should be achieved by the next Newton correction step. The requirement $\Theta_M(x_{k+1};\mu_{k+1}) = \Theta_d$ yields

$$\|x_{k+1} - x(\mu_{k+1})\| = \frac{\Theta_d}{\omega_M(\mu_{k+1})}.$$

Inserting our estimates into these equations we obtain, setting $\sigma := \mu_{k+1}/\mu_k$:

$$[e](x_{k+1};\mu_k) + [\eta](\mu_k)\mu_k(1-\sigma)\sigma^{-1/2} = \frac{\Theta_d}{[\omega](\mu_k)\sigma^{-1/2}}.$$

Since $[e](x_{k+1};\mu_k)$, $[\eta](\mu_k)$, $[\omega](\mu_k)$ are computationally available and Θ_d is given we may compute σ, and thus $\mu_{k+1} = \sigma\mu_k$ from this equation.

If the corrector has terminated with a failure, we still have $[e](x_k;\mu_{k-1})$, $[\eta](\mu_{k-1})$, and $[\omega](\mu_k)$ at hand. Note that $[\omega](\mu_k)$ is the result of the evaluation of the failed Newton step and thus gives rise to a step size reduction. Hence we can compute σ analogously to the successful case with $\mu_k < \sigma\mu_{k-1} < \mu_{k-1}$, via

$$[e](x_k;\mu_{k-1}) + [\eta](\mu_{k-1})\mu_{k-1}(1-\sigma)\sigma^{-1/2} = \frac{\Theta_d}{[\omega](\mu_k)\sigma^{-1/2}}.$$

which serves as a step size reduction.

6.4 Termination criteria

Depending on the application there are several termination criteria conceivable. For example we may use (47) or a modification to obtain a stopping criterion in terms of a norm of interest.

As an alternative we may stop if the estimated error in the functional is below a certain bound. Theorem 1 provides us with a linear convergence result via (6). By evaluation of the function values during the iteration we may estimate the missing constant and arrive at a good estimate for the error in the functional. Because the function values converge linearly in μ the difference $J(x(\mu)) - J(x_*)$ becomes small very quickly.

It is worth pointing out that interior point methods terminate at a *feasible* suboptimal solution. This feature, not shared by exterior penalty methods, reliefs users from the difficulty to decide how much infeasibility they are willing to accept. Rather, users can balance between optimality and computational effort.

If a quantitative discretization error estimate is available, then the error bounds can be matched with the these estimates. For a-priori error estimates for interior point methods we refer to [9], while quantitative a-posteriori error estimates together with an adaptive grid refinement strategy are subject to recent research [17].

7 Numerical examples

In our numerical examples we will investigate the performance of the proposed variants of interior point methods for some model problems. In particular, we are interested in the qualitative convergence behavior of the path-following scheme and a-posteriori estimates for the quantities which govern the path-following method, namely $[\eta]$ (cf. 45) and $[\omega]$ (cf. 51). Further we are interested in the convergence behavior of the function values at the iterates. Finally we are interested in the efficiency of the proposed method.

For a simple numerical example we consider an elliptic distributed control problem as described in our example in Sect. 2.3 on the unit square $\Omega\ =]0;1[\times]0;1[$. As for state constraints we choose $\overline{y} = 0.5$ as an upper bound, for the definition of the functional J we choose $y_d = 2 \cdot t_1 \cdot t_2, \alpha = 10^{-3}$. As boundary conditions we choose homogeneous Neumann conditions. The optimal state has a relatively large active set, and the Lagrange multiplier apparently consists of a regular part and a line measure, concentrated at the boundary of the active set. The control reveals an edge at the boundary of the active set.

As a second numerical example we change the boundary conditions to homogeneous Dirichlet conditions and choose $\overline{y} = 0.55$. Inspection of the numerical solution yields that the active constraint set seems to be concentrated on an single point with a point measure as Lagrange multiplier. The adjoint state (and thus the control) has a sharp peak at the active point.

The discretization of y and p is performed by linear finite elements as described in [9] on a uniform triangular grid. The implementation has been performed with the library KASKADE7 [6] based on the DUNE library [3]. For the evaluation of the barrier integrals we use the trapezoidal rule, as analyzed in [9]. The resulting linear systems of equations are solved by the direct sparse solver PARDISO [12].

Let us first have a look at the algorithmic quantities $[\eta]$ and $[\omega]$. It will turn out below that our method is able to perform very large reductions of μ per step. To obtain smooth plots we deliberately set the algorithmic parameters to very conservative values

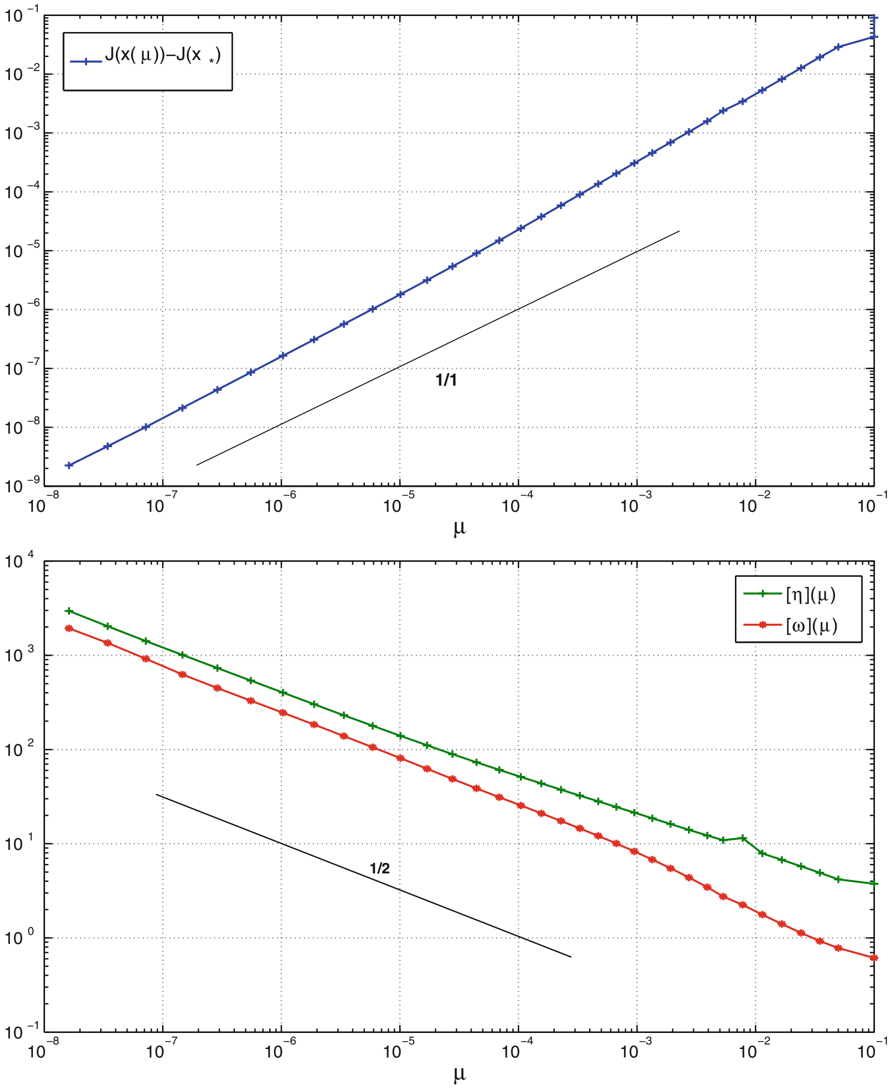

Fig. 1 Iteration history for first problem, using conservative parameters. *Top* Error in functional values. *Bottom* Algorithmic quantities

(in particular $\Theta_d = 0.05$) in the following. We also choose the stopping value for $\mu_{end} = 10^{-8}$ much smaller than appropriate for a practical application.

Comparing Fig. 1 to our theoretic results we conclude that the theoretic predictions $J(x(\mu)) - J(x_*) = O(\mu)$ and $\eta(\mu) = O(\mu^{-1/2})$ made in Theorem 1 correspond rather well to the computational estimates. A very close look at the graphs suggests that the convergence is slightly faster in this particular problem. In contrast the a-posteriori estimate $[\omega](\mu) = O(\mu^{-1/2})$ is much better than the predicted bound from Proposition 4 for this particular problem.

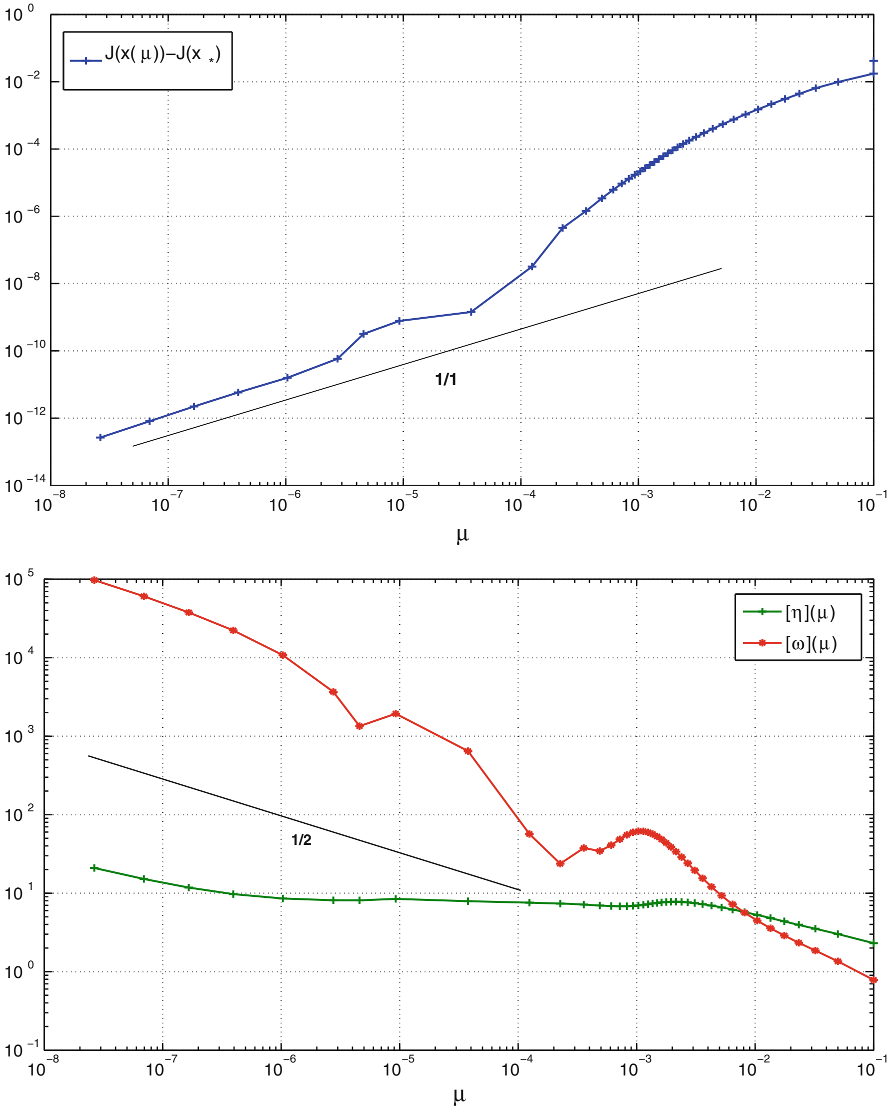

Fig. 2 Iteration history for first problem, using conservative parameters. *Top* Error in functional values. *Bottom* Algorithmic quantities

Figure 2, which corresponds to the case with a point functional shows a quite different behavior. While $J(x(\mu)) - J(x_*) = O(\mu)$ seems to hold asymptotically, $\eta(\mu)$ grows much more slowly than in our first example, while $\omega(\mu)$ grows faster and less regularly with local maximum near $\mu = 10^{-3}$. Observe how our algorithm reduces the stepsize in this difficult region to be able to comply to our (very restrictive) contraction demands. Summarizing, the second problem seems to be more nonlinear than the first, while having a shorter central path. This underlines the necessity of modeling the Newton nonlinearity as well as the properties of the central path.

$j \setminus i$	0	1	2	3
0	13	-	-	-
1	12	19	-	-
2	17	21	21	-
3	17	22	23	23

$j \setminus i$	0	1	2	3
0	14	-	-	-
1	12	12	-	-
2	13	15	12	-
3	15	15	13	13

Fig. 3 Number of Newton steps used by the various barrier functions $l_{i,j}(y;\mu)$, using aggressive parameters. *Left* First problem. *Right* Second problem

Problem $\setminus k$	4	5	6	7	8	9
#1	13	14	12	12	13	13
#2	9	11	11	12	14	13

Fig. 4 Number of Newton steps depending on the mesh size $h = 2^{-k}$, using aggressive parameters

Let us turn to the efficiency of our algorithm. The results of [9] suggest that for mesh sizes up to $h = 2^{-8}$ the discretization error (at least for the first problem) is above 2.5×10^{-3}. Hence, the choice of 10^{-3} as an accuracy requirement seems appropriate. For the first problem our algorithm detects this accuracy around $\mu \approx 5 \times 10^{-7}$, for the second problem this criterion is reached around $\mu \approx 10^{-5}$. The error in the functional is around 10^{-7} in the first problem and around 10^{-9} in the second problem.

For the desired contraction Θ_d we now choose a more aggressive value $\Theta_d = 0.8$ and a relative accuracy $\Lambda = 0.5$ (cf. 48) for the corrector. To assess the influence of the order parameter q on the computational performance we compute the solutions of our problem with the help barrier functions of the form

$$l_{i,j}(y;\mu) = \sum_{k=i}^{j} l(y;\mu;q = 1 + k/2)$$

with $0 \leq i \leq j \leq 3$ (cf. Fig. 3). The table indicates that for the first problem and for $h = 2^{-8}$ low order barrier functions seem to be the more efficient than their higher order counterparts. In particular, efficiency degrades, if low order terms are dropped. For the second problem there seems to be no clear advantage for any type of barrier function.

To inspect mesh dependence of our algorithm (using the pure logarithmic barrier function from now on) test runs were performed for $h = 2^{-k}$ for $k = 4 \ldots 9$ (cf. Fig. 4). While the number of Newton steps used for the first problem appears to be constant, for the second problem iteration counts increase slightly for finer discretizations. This reflects the fact that the structure of the solution of the first problem is already well resolved on coarse grids, while for the second problem the peak in the control is resolved only gradually when the grid is refined.

Finally, to give a qualitative result on the impact of pointwise damping, we performed a comparative run of our algorithm with pointwise damping and without pointwise damping. In Fig. 5 the choice of μ-reduction factors $\sigma_k = \mu_{k+1}/\mu_k$ is plotted. First, compare the variant without pointwise damping to a damped variant, using the same (conservative) algorithmic parameters. Strikingly, both methods behave identically in the starting phase but very differently afterwards. While the undamped variant

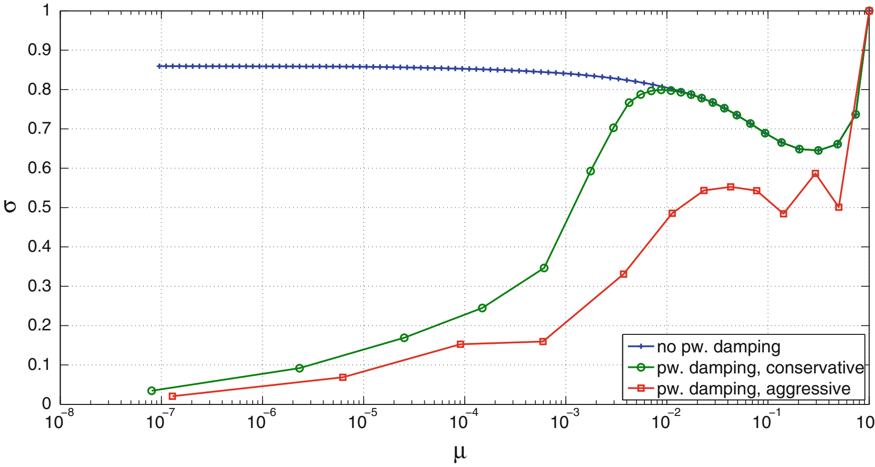

Fig. 5 Comparison of μ reduction factors with and without pointwise damping

seems to converge linearly and slowly (i.e. with an asymptotic choice $\sigma_\infty \approx 0.86$), the damped variant takes larger and larger steps (seemingly $\sigma_k \to 0$) asymptotically. Moreover, pointwise damping allows a more aggressive choice of parameters, leading to even larger steps.

8 Conclusion

This work provides analytic and algorithmic results for interior point methods in function space in the context of state constraints. On the analytic side, a proof of convergence for a simple path-following scheme has been given. Although it does not yield a sharp result on the rate of convergence (which seems to be linear from numerical experiments) it provides mathematical evidence that Newton path-following can be performed successfully. Similar results are not available yet for competing methods, such as Moreau-Yosida regularization [8] or Lavrentiev regularization [10]. There the homotopy path and local convergence of each homotopy step are analysed, but not the overall path-following scheme.

Major algorithmic aspects of our work are the pointwise modification of Newton steps, which solves the problem of generating feasible iterates, and an adaptive step size control, which is based on the analytic insights, gained in the first part of the paper. The proposed pointwise modification has a remarkable impact on the efficiency of our method, and numerical experiments indicate that the corresponding path-following scheme converges locally superlinearly. It would thus be interesting to gain a deeper understanding of this modification, and explore the application of similar techniques also in other contexts.

Acknowledgments The author wishes to thank Dr. Martin Weiser for helpful discussions and the close cooperation during the development of the computational framework.

References

1. Adams, R.A.: Sobolev Spaces. Academic, New York (1975)
2. Amann, H.: Nonhomogeneous linear and quasilinear elliptic and parabolic boundary value problems. In: Schmeisser, H.J., Triebel, H. (eds.) Function spaces, differential operators and nonlinear analysis, pp. 9–126. Teubner, Stuttgart, Leipzig (1993)
3. Bastian, P., Blatt, M., Engwer, C., Dedner, A., Klöfkorn, R., Kuttanikkad, S., Ohlberger, M., Sander, O.: The distributed and unified numerics environment (DUNE). In: Proceedings of the 19th Symposium on Simulation Technique in Hannover, Sep 12–14 (2006)
4. Deuflhard, P.: Newton methods for nonlinear problems, 2nd edition. Affine Invariance and Adaptive Algorithms, Vol. 35 of Series Computational Mathematics. Springer (2006)
5. Goldberg, S.: Unbounded Linear Operators. Dover Inc., New York (1966)
6. Götschel, S., Weiser, M., Schiela, A.: Solving Optimal Control Problems with the Kaskade 7 Finite Element Toolbox. In: Dedner, A., Flemisch, B., Klöfkorn, R. (eds.) Advances in DUNE, pp. 101–112. Springer, Berlin (2012)
7. Haller-Dintelmann, R., Rehberg, J., Meyer, C., Schiela, A.: Hölder continuity and optimal control for nonsmooth elliptic problems. Appl. Math. Optim. **60**(3), 397–428 (2009)
8. Hintermüller, M., Kunisch, K.: Feasible and non-interior path-following in constrained minimization with low multiplier regularity. SIAM J. Control Optim. **45**(4), 1198–1221 (2006)
9. Hinze, M., Schiela, A.: Discretization of interior point methods for state constrained elliptic optimal control problems: optimal error estimates and parameter adjustment. Comput. Optim. Appl. **48**(3), 581–600 (2011)
10. Meyer, C., Tröltzsch, F., Rösch, A.: Optimal control problems of PDEs with regularized pointwise state constraints. Comput. Optim. Appl. **33**, 206–228 (2006)
11. Prüfert, U., Tröltzsch, F., Weiser, M.: The convergence of an interior point method for an elliptic control problem with mixed control-state constraints. Comput. Optim. Appl. **39**(2), 183–218 (2008)
12. Schenk, O., Gärtner, K.: On fast factorization pivoting methods for sparse symmetric indefinite systems. Electron. Trans. Numer. Anal. **23**, 158–179 (2006)
13. Schiela, A.: The control reduced interior point method—a function space oriented algorithmic approach. PhD thesis, Department of Mathematics and Computer Science, Free University of Berlin (2006)
14. Schiela, A.: Convergence of the Control Reduced Interior Point Method for PDE Constrained Optimal Control with State Constraints. ZIB Report 06–16, Zuse Institute Berlin (2006)
15. Schiela, A.: Barrier methods for optimal control problems with state constraints. SIAM J. Optim. **20**(2), 1002–1031 (2009)
16. Schiela, A.: An extended mathematical framework for barrier methods in function space. In: Domain Decomposition Methods in Science and Engineering XVIII, Vol. 70 of Lecture Notes in Computational Science and Engineering, pp. 201–208. Springer, Berlin (2009)
17. Schiela, A., Günther, A.: An interior point algorithm with inexact step computation in function space for state constrained optimal control. Numerische Mathematik **119**(2), 373–407 (2011)
18. Weiser, M.: Function space complementarity methods for optimal control problems. PhD thesis, Department of Mathematics and Computer Science, Free University of Berlin (2001)