

MARTIN WEISER

**Linear convergence of an interior point
method for linear control constrained
optimal control problems**

Linear convergence of an interior point method for linear control constrained optimal control problems

Martin Weiser

March 21, 2002

Abstract

The paper provides a detailed analysis of a short step interior point algorithm applied to linear control constrained optimal control problems. Using an affine invariant local norm and an inexact Newton corrector, the well-known convergence results from finite dimensional linear programming can be extended to the infinite dimensional setting of optimal control.

The present work complements a recent paper of Weiser and Deuffhard on a similar multilevel interior point algorithm applied to more general optimal control problems, where convergence rates have not been derived.

The choice of free parameters, i.e. the corrector accuracy and the number of corrector steps, is discussed.

Keywords: linear optimal control, infinite dimensional, interior point methods, inexact Newton methods

AMS classification (2000): 90C51, 49N05, 65J10, 65K05

1 Introduction

Primal-dual interior point methods are a powerful tool to solve linear optimization problems. In the optimal control area, however, they are most often used for solving a priori discretized problems.

Recently, WEISER AND DEUFLHARD [17, 18] have presented a novel multilevel algorithm combining primal-dual complementarity methods in function space and inexact Newton pathfollowing. This algorithm internally adapts the mesh refinement to the progress on the central path, and has been shown to solve a rather intricate real-world problem up to high accuracy, which had previously been achieved only by indirect methods.

There remained, however, a gap in the accompanying theory. Convergence of the central path could only be proved for purely control constrained optimal control problems, and no lower bound on the pathfollowing method's convergence rate had been derived. The present paper is a step towards filling this gap.

We analyze the convergence behavior of a very similar algorithm applied to *linear control constrained* optimal control problems, which are more accessible to theory. The main result is that under reasonable assumptions the *duality gap parameter* μ can be reduced by a constant factor in each step of the inexact Newton pathfollowing method, taking exactly one corrector step.

The setting is inspired by recent work of POTRA [13] on finite dimensional linear complementarity problems.

This is an extension of similar results for finite dimensional linear programming problems to the infinite dimensional optimal control setting. This extension relies on two important properties of linear control constrained optimal control problems: (a) they decompose pointwise into finite dimensional linear programming problems and (b) solutions of the interior point Kuhn-Tucker systems are continuous. Property (a) allows to adapt techniques from a recent work by POTRA [13] to the infinite dimensional setting.

The consequence for the application of primal-dual interior point methods to *a priori discretized* optimal control problems is, that the convergence rate does not deteriorate as the discretization becomes finer, as could be expected from the $1 - \text{const}/\sqrt{n}$ convergence factor predicted by finite dimensional theory (see e.g. WRIGHT [19]). However, the convergence result does not imply $\mathcal{O}(\sqrt{n}L)$ computational complexity for solving the whole problem, since the effort spent per step increases with the size of the discretization.

Remarkably few publications are concerned with applying interior point methods to infinite dimensional optimal control problems. ITO, KELLEY, AND SACHS [10] analyze inexact primal-dual interior point methods for infinite dimensional linear programs in standard form. The inexactness considered is, however, in general not sufficient to account for discretization errors, but only for truncation errors of iterative solvers.

We refer to ULBRICH, ULBRICH, AND HEINKENSCHLOSS [15] and ULBRICH AND ULBRICH [14] for affine scaling interior point methods. FAYBUSOVICH AND MOORE [7] extend the Nesterov-Nemirovskii framework to infinite dimensional problems with, however, finitely many inequality constraints.

More papers treat inexact interior point methods applied to finite dimensional linear programs, predominantly employing iterative linear solvers in order to solve the occurring linear equations. We refer to FREUND, JARRE, AND

MIZUNO [8], Bellavia [2], KORZAK [11], and BARYAMUREEBA AND STEIHAUG [1].

The paper is organized as follows. In Section 2 we state the class of problems considered here and formulate the primal-dual interior point algorithm. Section 3 contains the convergence analysis of the inexact Newton corrector towards the central path. In Section 4, a lower bound on the continuation stepsize is derived.

Notation. The algorithm is formulated in terms of the compound variable $v = (u, y, \lambda, \eta)^T$, which we will write interchangeably, i.e. $\bar{v} = (\bar{u}, \bar{y}, \bar{\lambda}, \bar{\eta})^T$ etc. With $\text{co } A$ we denote the convex hull of the set A . The Hadamard (or pointwise, or componentwise) product of two vectors or functions a, b is written as $a * b$. With a bold symbol \mathbf{a} we denote the Nemyckii (or diagonal) operator associated to a , i.e. $\mathbf{a}b = a * b$ for all b . $\mathbf{1}$ stands for the vector (or function) with all components equal to 1 (almost everywhere).

2 Primal-dual interior point method

Setting. We consider the optimization problem

$$\text{minimize } \int_{t=0}^T (J_u(t)^T u(t) + J_y(t)^T y(t)) dt$$

subject to

$$\dot{y}(t) + C_y(t)y(t) + C_u(t)u(t) + c(t) = 0 \quad \text{a. e.} \quad (1)$$

$$R_0 y(0) + R_T y(T) + r = 0 \quad (2)$$

$$G(t)u(t) + b(t) \geq 0 \quad \text{a. e.} \quad (3)$$

where $u \in L_\infty(0, T)^{n_u}$, and $y \in W_\infty^1(0, T)^{n_y}$. The constituting functions $J_u : [0, T] \rightarrow \mathbb{R}^{n_u}$, $J_y : [0, T] \rightarrow \mathbb{R}^{n_y}$, $C_y : [0, T] \rightarrow \mathbb{R}^{n_y, n_y}$, $C_u : [0, T] \rightarrow \mathbb{R}^{n_u, n_y}$, $c : [0, T] \rightarrow \mathbb{R}^{n_y}$, $G : [0, T] \rightarrow \mathbb{R}^{n_g, n_u}$, and $b : [0, T] \rightarrow \mathbb{R}^{n_g}$ are assumed to be continuous. Together with the boundary condition (2) defined in terms of $R_0, R_T \in \mathbb{R}^{n_y, n_y}$ and $r \in \mathbb{R}^{n_y}$, the differential equation (1) is assumed to form a well-posed boundary value problem, such that for every $u \in L_\infty(0, T)^{n_u}$ there is a unique solution $y \in W_\infty^1(0, T)^{n_y}$ with

$$\|y\|_{W_\infty^1(0, T)^{n_y}} \leq \alpha_C \|u\|_{L_\infty(0, T)^{n_u}}.$$

G is assumed to satisfy the following uniformity condition:

$$\|G(t)\xi\|_2 \geq \alpha_G \|\xi\|_2 \quad \text{for all } t \in [0, 1]. \quad (4)$$

In order to ease the notation, we will write the boundary value problem given by (1) and (2) in the more compact form

$$C_y y + C_u u + c = 0.$$

Primal-dual interior point method. Under reasonable conditions, any solution (u, y) of the optimization problem above satisfies the following first-order

necessary condition (cf. MAURER AND ZOWE [12]):

$$\begin{aligned} \exists \lambda \in (L_\infty^{n_y} \times \mathbb{R}^{n_y})^*, \eta \in (L_\infty^{n_g})^* : & -J_u + C_u^* \lambda + G^* \eta = 0 \\ & -J_y + C_y^* \lambda = 0 \\ & C_y y + C_u u + c = 0 \\ & \langle \eta, Gu + b \rangle = 0 \\ & \eta, Gu + b \geq 0 \end{aligned}$$

A well-known method to compute solution candidates for linear optimization problems is to regularize the complementarity condition $\langle \eta, Gu + b \rangle = 0, \eta, Gu + b \geq 0$ by introducing a duality gap parameter $\mu > 0$ and restricting the iterates to the interior of the feasible region:

$$\begin{aligned} \eta * (Gu + b) &= \mu \mathbf{1} \\ \eta, Gu + b &> 0 \end{aligned}$$

The homotopy in μ defines a *central path* that can in general be followed towards the solution at $\mu = 0$. The interior-point Karush-Kuhn-Tucker system to which a continuation method must then be applied is

$$F(u, y, \lambda, \eta; \mu) = \begin{bmatrix} -J_u + C_u^* \lambda + G^* \eta \\ -J_y + C_y^* \lambda \\ C_y y + C_u u + c \\ \eta * (Gu + b) - \mu \mathbf{1} \end{bmatrix}. \quad (5)$$

For F to be continuously differentiable we have to make the regularity assumption $\eta \in L_\infty^{n_g}$. Additionally we assume $\lambda \in L_\infty^{n_y} \times \mathbb{R}^{n_y}$. In fact, both assumptions are justified by the regularity conditions imposed on J_u, J_y, C_u, C_y, c , and G . F is defined on the interior of the feasible region:

$$\text{dom}(F) := \{u \in L_\infty^{n_u} : Gu + b > 0\} \times (W_\infty^1)^{n_y} \times L_\infty^{n_y} \times \{\eta \in L_\infty^{n_g} : \eta > 0\}.$$

For actually computing an approximate solution we employ a predictor-corrector continuation method with classic predictor and Newton type corrector. However, since the pathfollowing algorithm is applied to a non-discretized, infinite dimensional problem in function space, in practice we will not be able to compute Newton corrections δv exactly. Therefore, we have to employ an *inexact* Newton method, where an *inner residual* $r^{k,i}$ remains.

Algorithm 1.

```

initialize  $\mu^0 > 0$  and  $v^0 \in \text{dom}(F)$ 
for  $k = 0, \dots$ :
  perform  $i_*^k$  corrector steps:
     $v^{k,0} := v^k$ 
    for  $i = 0, \dots, i_*^k$ :
      compute inexact Newton correction:
         $F'(v^{k,i}; \mu^k) \delta v^{k,i} = -F(v^{k,i}; \mu^k) + r^{k,i}$ 
         $v^{k,i+1} := v^{k,i} + \delta v^{k,i}$ 
     $v^{k+1} := v^{k,i_*^k}$ 
  decrease continuation parameter:
     $\mu^{k+1} := (1 - \sigma / \sqrt{n_\eta}) \mu^k$ 

```

Remark 1. For the problem type considered here, this approach is not the most efficient and is in fact not advocated. One single solve of the adjoint equation decouples the problem temporally, such that only independent low-dimensional linear programs for the control u need to be solved. In contrast to this specialized approach, however, the analyzed algorithm extends in a natural way to nonlinear problems. The presented analysis is to be seen as a step towards understanding the behavior of the algorithm for the more general case.

Local norm. First we define a local norm which is suitable for the analysis of Newton's method. The construction of the norm is inspired by POTRA [13].

For every pair (u, η) of continuous functions from $L_\infty^{n_u} \times L_\infty^{n_\eta}$ with $Gu + b > 0$ and $\eta > 0$, we define

$$\|(r_u, r_y, r_\lambda, r_\eta)^T\|_{(u, \eta)}^2 := \|r_y\|_{L_\infty^{n_y} \times \mathbb{R}^{n_y}}^2 + \|(r_u, r_\lambda, r_\eta)^T\|_{(u, \eta)}^2$$

with

$$\|(r_u, r_\lambda, r_\eta)^T\|_{(u, \eta)}^2 := \|\bar{\lambda}\|_{L_\infty^{n_y} \times \mathbb{R}^{n_y}}^2 + \|(\bar{u}, \bar{\eta})^t\|_{(u, \eta)}^2,$$

where

$$\begin{bmatrix} C_u^* & G^* \\ C_y^* & \mathbf{w} \\ \eta G & \end{bmatrix} \begin{bmatrix} \bar{u} \\ \bar{\lambda} \\ \bar{\eta} \end{bmatrix} = \begin{bmatrix} r_u \\ r_\lambda \\ r_\eta \end{bmatrix}, \quad w := Gu + b, \quad (6)$$

$$\|(\bar{u}, \bar{\eta})^T\|_{(u, \eta)} := \sup_{t \in [0, T]} \|(\bar{u}(t), \bar{\eta}(t))^T\|_{(u(t), \eta(t))},$$

$$\|(\bar{u}(t), \bar{\eta}(t))^T\|_{(u(t), \eta(t))}^2 := \|D(t)G(t)\bar{u}(t)\|_2^2 + \|D(t)^{-1}\bar{\eta}(t)\|_2^2,$$

and

$$D := \mathbf{w}^{-1/2} \eta^{1/2}.$$

3 Inexact Newton Corrector

In this section, we will analyze the convergence of Newton's method towards the central path. In order to simplify the notation, we will drop indices corresponding to the continuation iteration and also omit the (arbitrary but fixed) continuation parameter μ from the parameter list of F .

We would like to emphasize the fact that the Newton convergence result below is *affine invariant* as long as the local norm in terms of which it is formulated, is itself affine invariant. As the ordinary Newton method itself, the convergence result is not affected by invertible affine transformations of the domain or image space, under which the norm is invariant. Note that the local norm defined in Section 2 is invariant under transformations of the domain space of (u, y) . For detailed treatment of affine invariance we refer to DEUFLHARD AND HEINDL [4], HOHMANN [9], DEUFLHARD AND WEISER [5, 6], VOLKWEIN AND WEISER [16], and, in particular, to DEUFLHARD [3].

Theorem 1. *Let V and Z be Banach spaces, $\text{dom}(F) \subset V$ open, and $F : \text{dom}(F) \rightarrow Z$ a twice Fréchet-differentiable mapping. For $v \in \text{dom}(F)$ let $\|\cdot\|_{(v)}$ be a local norm on Z , and let $\mathcal{L}_\gamma(v)$ denote the connected component of the level set $\{\bar{v} \in \text{dom}(F) : \|F(\bar{v})\|_{(\bar{v})} \leq (1 + \gamma/2) \|F(v)\|_{(v)}\}$ that contains v . Let $v_0 \in \text{dom}(F)$ be a given starting point. Assume there exist constants $0 < \gamma < \infty$, $\omega < \infty$, $\bar{\delta} < 1$, and $\Theta < 1$, such that*

1. $\mathcal{L}_\gamma(v_0)$ is closed,
2. F' satisfies the Lipschitz condition

$$\|(F'(v + \Delta v) - F'(v))\Delta v\|_{(v)} \leq \omega \|F'(u, \eta)\Delta v\|_{(v)}^2, \quad (7)$$

for $v, \Delta v$ such that $v \in \mathcal{L}_0(v^0)$ and $\text{co}\{v, v + \Delta v\} \subset \mathcal{L}_\gamma(v^0)$,

3. the local norm $\|\cdot\|_{(v)}$ satisfies the Lipschitz condition

$$\|z\|_{(v+\Delta v)} \leq (1 + \gamma\omega \|F'(v)\Delta v\|_{(v)}) \|z\|_{(v)} \quad (8)$$

for all v and Δv with $\|F'(v)\Delta v\|_{(v)} \leq (1 + \bar{\delta}) \|F(v^0)\|_{(v^0)}$ and $\text{co}\{v, v + \Delta v\} \subset \mathcal{L}_\gamma(v^0)$, and

4. the relative error

$$\delta^k = \frac{\|r^k\|_{(v^k)}}{\|F'(v^k)\|_{(v^k)}} \leq \bar{\delta} \quad (9)$$

of the inexact Newton iteration

$$\begin{aligned} F'(v^k)\delta v^k &= -F(v^k) + r^k \\ v^{k+1} &= v^k + \delta v^k \end{aligned}$$

satisfies the accuracy matching condition

$$(1 + \gamma(1 + \delta^k)h^k) \left(\delta^k + \frac{h^k}{2}(1 + \delta^k)^2 \right) \leq \Theta \quad (10)$$

for all $k \geq 0$, where $h^k := \omega \|F(v^k)\|_{(v^k)}$.

Then the inexact Newton iteration converges linearly to the solution:

$$\|F(v^{k+1})\|_{(v^{k+1})} \leq \Theta \|F(v^k)\|_{(v^k)} \quad (11)$$

Proof. By induction, assume $\mathcal{L}_\gamma(v^k)$ is closed. For all $s \in [0, 1]$ such that $\text{co}\{v^k, v^k + \delta v^k\} \subset \mathcal{L}(\|F(v_0)\|_{(v_0)})$, we have

$$\begin{aligned} F(v^k + s\delta v^k) &= F(v^k) + \int_{\sigma=0}^s F'(v^k + \sigma\delta v^k)\delta v^k d\sigma \\ &= (1 - s)F(v^k) + sr^k + \int_{\sigma=0}^s (F'(v^k + \sigma\delta v^k) - F'(v^k))\delta v^k d\sigma \end{aligned}$$

and consequently, with (7) and (9),

$$\begin{aligned}
\|F(v^k + s\delta v^k)\|_{(v^k)} &\leq (1-s) \|F(v^k)\|_{(v^k)} + s\delta^k \|F(v^k)\|_{(v^k)} \\
&\quad + \int_{\sigma=0}^s \|(F'(v^k + \sigma\delta v^k) - F'(v^k))\delta v^k\|_{(v^k)} d\sigma \\
&\leq (1-s + \delta^k s) \|F(v^k)\|_{(v^k)} + \int_{\sigma=0}^s \omega\sigma \|F'(v^k)\delta v^k\|_{(v^k)}^2 d\sigma \\
&\leq (1-s + \delta^k s) \|F(v^k)\|_{(v^k)} + \frac{s^2}{2}\omega(1+\delta^k)^2 \|F(v^k)\|_{(v^k)}^2 \\
&= \left(1 - (1-\delta^k)s + \frac{s^2}{2}h^k(1+\delta^k)^2\right) \|F(v^k)\|_{(v^k)}.
\end{aligned}$$

Since by (9), (10) and by induction

$$\|F'(v^k)s\delta v^k\|_{(v^k)} \leq (1+\delta^k) \|F(v^k)\|_{(v^k)} \leq (1+\bar{\delta}) \|F(v^0)\|_{(v^0)},$$

we can apply (8) and (9) in order to obtain

$$\begin{aligned}
&\frac{\|F(v^k + s\delta v^k)\|_{(v^k+s\delta v^k)}}{\|F(v^k)\|_{(v^k)}} \\
&\leq \left(1 + s\gamma\omega \|F'(v^k)\delta v^k\|_{(v^k)}\right) \left(1 - (1-\delta^k)s + \frac{s^2}{2}(1+\delta^k)^2h^k\right) \\
&\leq (1 + s\gamma(1+\delta^k)h^k) \left(1 - (1-\delta^k)s + \frac{s}{2}(1+\delta^k)^2h^k\right). \tag{12}
\end{aligned}$$

Using Lemma 1 below with $\chi_1 = \gamma(1+\delta^k)h^k$ and $\chi_2 = 1 - \delta^k - (1+\delta^k)^2h^k/2$, and observing that the accuracy condition (10) implies $(1+\delta^k)h^k < 2$, yields

$$\frac{\|F(v^k + s\delta v^k)\|_{(v^k+s\delta v^k)}}{\|F(v^k)\|_{(v^k)}} \leq 1 + \frac{\gamma(1+\delta^k)h^k}{4} < 1 + \frac{\gamma}{2}. \tag{13}$$

Thus, for all $s \in [0, 1]$ such that $\text{co}\{v^k, v^k + s\delta v^k\} \subset \mathcal{L}_\gamma(v^k)$, (13) holds. Now assume $\text{co}\{v^k, v^k + \delta v^k\} \not\subset \mathcal{L}_\gamma(v^k)$. Because of $\gamma > 0$, v^k is in the interior of $\mathcal{L}_\gamma(v^k)$, thus there is a minimal $s^* > 0$, such that $\text{co}\{v^k, v^k + s^*\delta v^k\} \subset \mathcal{L}_\gamma(v^k)$ and $v^k + s^*\delta v^k \in \partial\mathcal{L}_\gamma(v^k)$. By continuity,

$$\|F(v^k + s^*\delta v^k)\|_{(v^k+s^*\delta v^k)} = (1 + \gamma/2) \|F(v^k)\|_{(v^k)},$$

which contradicts (13). Therefore, $\text{co}\{v^k, v^k + \delta v^k\} \subset \mathcal{L}_\gamma(v^k)$.

Inserting $s = 1$ into (12) and using (10) leads to

$$\begin{aligned}
\|F(v^{k+1})\|_{(v^{k+1})} &\leq (1 + \gamma(1+\delta^k)h^k) \left(\delta^k + \frac{h^k}{2}(1+\delta^k)^2\right) \|F(v^k)\|_{(v^k)} \\
&\leq \Theta \|F(v^k)\|_{(v^k)}
\end{aligned}$$

and verifies the contraction result (11). Consequently, $\mathcal{L}_\gamma(v^{k+1}) \subset \mathcal{L}_\gamma(v^k)$ is closed, which completes the induction step. \square

Lemma 1. Let $\chi_1, \chi_2 \geq 0$ and $f(s) = (1 + \chi_1 s)(1 - \chi_2 s)$. If $f(1) < 1$, then

$$f(s) < 1 + \frac{\chi_1}{4} \quad \text{for all } s \in [0, 1].$$

Proof. From $1 > f(1) = 1 + (\chi_1 - \chi_2) - \chi_1 \chi_2$ we infer $\chi_1 - \chi_2 < \chi_1 \chi_2$. Then

$$\begin{aligned} f(s) &= 1 + (\chi_1 - \chi_2)s - \chi_1 \chi_2 s^2 < 1 + (\chi_1 - \chi_2)(s - s^2) \\ &\leq 1 + \frac{\chi_1 - \chi_2}{4} \leq 1 + \frac{\chi_1}{4} \end{aligned}$$

for all $s \in [0, 1]$. \square

In order to apply Theorem 1 to the interior point formulation (5), all the assumptions have to be verified. This is done in the remaining part of this section.

Lipschitz continuity of F' . First we establish local Lipschitz continuity of the derivative F' . Lemmas 2 and 3 are derived from POTRA [13].

Lemma 2. Let $G \in \mathbb{R}^{n_g, n_u}$ be of full rank, and $\eta, b \in \mathbb{R}^{n_g}$, $u \in \mathbb{R}^{n_u}$. Assume that both η and $w := Gu + b$ are positive. Any solution $(\bar{u}, \bar{\eta})^T$ of

$$\begin{bmatrix} G^T \\ \eta G \quad \mathbf{w} \end{bmatrix} \begin{bmatrix} \bar{u} \\ \bar{\eta} \end{bmatrix} = \begin{bmatrix} 0 \\ a \end{bmatrix} \quad (14)$$

satisfies

$$\|(\bar{u}, \bar{\eta})^T\|_{(u, \eta)} = \left\| (\eta \mathbf{w})^{-1/2} a \right\|_2.$$

Proof. Multiplication of the bottom equation in (14) with $(\eta \mathbf{w})^{-1/2}$ yields

$$DG\bar{u} + D^{-1}\bar{\eta} = (\eta \mathbf{w})^{-1/2} a,$$

where the diagonal matrix D is given by $D = \mathbf{w}^{-1/2} \eta^{1/2}$. Taking the norm on both sides leads to

$$\|DG\bar{u}\|_2^2 + 2\langle \bar{\eta}, G\bar{u} \rangle + \|D^{-1}\bar{\eta}\|_2^2 = \left\| (\eta \mathbf{w})^{-1/2} a \right\|_2^2.$$

The top equation in (14) implies $\langle \bar{\eta}, G\bar{u} \rangle = 0$, which completes the proof. \square

Lemma 3. Let $G \in \mathbb{R}^{n_\eta, n_u}$ be of full rank, and $\eta, \Delta\eta, \bar{\eta}, b \in \mathbb{R}^{n_\eta}$, $u, \Delta u, \bar{u} \in \mathbb{R}^{n_u}$, such that both η and $w := Gu + b$ are positive. Then

$$\|\Delta\eta * G\bar{u} + \bar{\eta} * G\Delta u\|_2 \leq \|(\bar{u}, \bar{\eta})^T\|_{(u, \eta)} \|(\Delta u, \Delta\eta)^T\|_{(u, \eta)}.$$

Proof. Defining $D := \mathbf{w}^{-1/2} \eta^{1/2}$ and using the Cauchy-Schwarz inequality in \mathbb{R}^2 we estimate

$$\begin{aligned} \|\Delta\eta * G\bar{u} + \bar{\eta} * G\Delta u\|_2 &\leq \|\Delta\eta * G\bar{u}\|_2 + \|\bar{\eta} * G\Delta u\|_2 \\ &\leq \|DG\bar{u}\|_2 \|D^{-1}\Delta\eta\|_2 + \|D^{-1}\bar{\eta}\|_2 \|DG\Delta u\|_2 \\ &\leq \sqrt{\|DG\bar{u}\|_2^2 + \|D^{-1}\bar{\eta}\|_2^2} \sqrt{\|DG\Delta u\|_2^2 + \|D^{-1}\Delta\eta\|_2^2} \\ &= \|(\bar{u}, \bar{\eta})^T\|_{(u, \eta)} \|(\Delta u, \Delta\eta)^T\|_{(u, \eta)}, \end{aligned}$$

which completes the proof. \square

Theorem 2. F' satisfies the Lipschitz condition

$$\|(F'(v_1 + \Delta v) - F'(v_1))\Delta v\|_{(u,\eta)} \leq \omega_{(u,\eta)} \|F'(v)\Delta v\|_{(u,\eta)}^2 \quad (15)$$

with a local Lipschitz constant

$$\omega_{(u,\eta)} = \left(\min_{t \in [0, T]} \min \eta(t) * (G(t)u(t) + b(t)) \right)^{-1/2}.$$

Proof. First we consider the left hand side of (15), which originates from

$$(F'(u_1 + \Delta u, \eta_1 + \Delta \eta) - F'(u_1, \eta_1)) \begin{bmatrix} u_a \\ y_a \\ \lambda_a \\ \eta_a \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ a \end{bmatrix}$$

with $a = \Delta \eta * Gu_a + (G\Delta u) * \eta_a$. Let $w := Gu + b$. From the reduced system (6) we eliminate $\bar{\lambda} = 0$ and obtain

$$\begin{bmatrix} \boldsymbol{\eta}G & G^* \\ \boldsymbol{w} & \boldsymbol{w} \end{bmatrix} \begin{bmatrix} \bar{u} \\ \bar{\eta} \end{bmatrix} = \begin{bmatrix} 0 \\ a \end{bmatrix}.$$

This system is completely decoupled in time and can be written as

$$\begin{bmatrix} \boldsymbol{\eta}(t)G(t) & G(t)^* \\ \boldsymbol{w}(t) & \boldsymbol{w}(t) \end{bmatrix} \begin{bmatrix} \bar{u}(t) \\ \bar{\eta}(t) \end{bmatrix} = \begin{bmatrix} 0 \\ a(t) \end{bmatrix} \quad \text{for all } t \in [0, T].$$

Applying Lemmas 2 and 3 we obtain

$$\begin{aligned} & \|(\bar{u}(t), \bar{\eta}(t))^T\|_{(u(t), \eta(t))}^2 \\ & \leq (\min \eta(t) * w(t))^{-1} \|a(t)\|_2^2 \\ & \leq \omega_{(u,\eta)}^2 \|a(t)\|_2^2 \\ & \leq \omega_{(u,\eta)}^2 \|(u_a(t), \eta_a(t))^T\|_{(u(t), \eta(t))}^2 \|(\Delta u(t), \Delta \eta(t))^T\|_{(u(t), \eta(t))}^2 \end{aligned}$$

for all $t \in [0, T]$. Taking the supremum finally yields

$$\begin{aligned} \|(F'(v_1 + \Delta v) - F'(v_1))\Delta v\|_{(u,\eta)} & \leq \|(\bar{u}, \bar{\eta})^T\|_{(u,\eta)} \\ & \leq \omega_{(u,\eta)} \|(u_a, \eta_a)^T\|_{(u,\eta)} \|(\Delta u, \Delta \eta)^T\|_{(u,\eta)}. \end{aligned} \quad (16)$$

Concerning the right hand side of (15), by definition the local norm satisfies

$$\|(\bar{u}, \bar{\eta})^T\|_{(u,\eta)} \leq \|F'(u, \eta)(\bar{u}, \bar{y}, \bar{\lambda}, \bar{\eta})^T\|_{(u,\eta)} \quad (17)$$

for all $(\bar{u}, \bar{y}, \bar{\lambda}, \bar{\eta})$. Combining (16) and (17) yields (15) and completes the proof. \square

Lipschitz continuity of the local norm. Second we establish the Lipschitz continuity of the local norm.

Lemma 4. Let $|a, b|_{x,y}^2 := \frac{y}{x}a^2 + \frac{x}{y}b^2$. For real numbers $a, b, x, y, \Delta x, \Delta y$ such that $x, y, x + \Delta x$, and $y + \Delta y$ are positive, the following inequality holds:

$$|a, b|_{x+\Delta x, y+\Delta y} \leq \sqrt{\frac{1+h}{1-h}} |a, b|_{x,y}, \quad (18)$$

where $h = \omega |\Delta x, \Delta y|_{x,y}$ and $\omega = 1/\sqrt{xy}$.

Proof. To begin with, we have

$$\begin{aligned} |a, b|_{x+\Delta x, y+\Delta y}^2 &= \frac{y + \Delta y}{x + \Delta x} a^2 + \frac{x + \Delta x}{y + \Delta y} b^2 \\ &= \frac{y + \Delta y}{x + \Delta x} \frac{x}{y} a^2 + \frac{x + \Delta x}{y + \Delta y} \frac{y}{x} b^2 \\ &\leq \max\left(\frac{y}{y + \Delta y} \frac{x + \Delta x}{x}, \frac{y + \Delta y}{y} \frac{x}{x + \Delta x}\right) \left(\frac{y}{x} a^2 + \frac{x}{y} b^2\right) \\ &= \max\left(\frac{1 + \frac{\Delta x}{x}}{1 + \frac{\Delta y}{y}}, \frac{1 + \frac{\Delta y}{y}}{1 + \frac{\Delta x}{x}}\right) |a, b|_{x,y}^2. \end{aligned}$$

With

$$\frac{|\Delta x|}{x} = \omega \sqrt{\frac{y}{x}} |\Delta x| = \omega |\Delta x, 0|_{x,y} \leq h \quad \text{and similarly} \quad \frac{|\Delta y|}{y} \leq h \quad (19)$$

we obtain

$$|a, b|_{x+\Delta x, y+\Delta y}^2 \leq \max\left(\frac{1+h}{1-h}, \frac{1+h}{1-h}\right) |a, b|_{x,y}^2 = \frac{1+h}{1-h} |a, b|_{x,y}^2$$

and immediately (18). \square

Componentwise application of Lemma 4 results in

Corollary 1. Let $G \in \mathbb{R}^{n_\eta, n_u}$ be of full rank, and $\eta, \Delta\eta, \bar{\eta}, b \in \mathbb{R}^{n_\eta}$, $u, \Delta u, \bar{u} \in \mathbb{R}^{n_u}$, such that $\eta, \eta + \Delta\eta, Gu + b$, and $G(u + \Delta u) + b$ are all positive. Then

$$\|(\bar{u}, \bar{\eta})^T\|_{(u+\Delta u, \eta+\Delta\eta)} \leq \sqrt{\frac{1+h}{1-h}} \|(\bar{u}, \bar{\eta})^T\|_{(u, \eta)},$$

where $h = \omega \|(\Delta u, \Delta\eta)^T\|_{(u, \eta)}$ and $\omega = 1/\sqrt{\min \eta * (Gu + b)}$.

Lemma 5. Denoting $\omega_{x,y} := 1/\sqrt{xy}$ for positive real numbers x and y , the following inequality holds for increments $|\Delta x| < x$ and $|\Delta y| < y$:

$$\frac{\omega_{x,y}}{1+h} \leq \omega_{x+\Delta x, y+\Delta y} \leq \frac{\omega_{x,y}}{1-h},$$

where $h := \omega_{x,y} |\Delta x, \Delta y|_{x,y}$.

Proof. First we have

$$\omega_{x+\Delta x, y+\Delta y}^{-2} = (x + \Delta x)(y + \Delta y) = \omega_{x,y}^{-2} \left(1 + \frac{\Delta x}{x}\right) \left(1 + \frac{\Delta y}{y}\right)$$

and therefore

$$\omega_{x,y}^{-2} \left(1 + \frac{|\Delta x|}{x}\right) \left(1 + \frac{|\Delta y|}{y}\right) \geq \omega_{x+\Delta x, y+\Delta y}^{-2} \geq \omega_{x,y}^{-2} \left(1 - \frac{|\Delta x|}{x}\right) \left(1 - \frac{|\Delta y|}{y}\right).$$

From (19) we then infer

$$\omega^{-2}(1+h)^2 \geq \omega_{x+\Delta x, y+\Delta y}^{-2} \geq \omega^{-2}(1-h)^2,$$

which completes the proof. \square

Again, componentwise application of Lemma 5 gives

Corollary 2. *For increments $(\Delta u, \Delta \eta)^T$ with $\|(\Delta u, \Delta \eta)^T\|_{(u,\eta)} \leq \omega_{(u,\eta)}^{-1}$, the following inequality holds:*

$$\frac{\omega_{(u,\eta)}}{1+h} \leq \omega_{(u+\Delta u, \eta+\Delta \eta)} \leq \frac{\omega_{(u,\eta)}}{1-h},$$

where $h := \omega_{(u,\eta)} \|(\Delta u, \Delta \eta)^T\|_{(u,\eta)}$.

Theorem 3. *The local norm $\|\cdot\|_{(u,\eta)}$ satisfies the Lipschitz condition*

$$\|r\|_{(u+\Delta u, \eta+\Delta \eta)} \leq \left(\sqrt{\frac{1+h}{1-h}} + \frac{h}{1-h} \right) \|r\|_{(u,\eta)}, \quad (20)$$

where $h := \omega_{(u,\eta)} \|(u, \eta)^T\|_{(u,\eta)}$.

Proof. Let $w := Gu + b$ and $\Delta w := G\Delta u$. For any $r = (r_u, r_y, r_\lambda, r_\eta)$ we have

$$\begin{aligned} \left| \|r\|_{(u+\Delta u, \eta+\Delta \eta)} - \|r\|_{(u,\eta)} \right| &\leq \left| \|\bar{\lambda} + \Delta \bar{\lambda}\|_{L_\infty} + \|(\bar{u} + \Delta \bar{u}, \bar{\eta} + \Delta \bar{\eta})^T\|_{(u+\Delta u, \eta+\Delta \eta)} \right. \\ &\quad \left. - \|\bar{\lambda}\|_{L_\infty} - \|(\bar{u}, \bar{\eta})^T\|_{(u,\eta)} \right| \end{aligned} \quad (21)$$

where

$$\begin{bmatrix} C_u^* & G^* \\ C_y^* & \mathbf{w} \\ \eta G & \mathbf{w} \end{bmatrix} \begin{bmatrix} \bar{u} \\ \bar{\lambda} \\ \bar{\eta} \end{bmatrix} = \begin{bmatrix} r_u \\ r_y \\ r_\eta \end{bmatrix} = \begin{bmatrix} C_u^* & G^* \\ C_y^* & \mathbf{w} + \Delta \mathbf{w} \\ (\eta + \Delta \eta)G & \mathbf{w} + \Delta \mathbf{w} \end{bmatrix} \begin{bmatrix} \bar{u} + \Delta \bar{u} \\ \bar{\lambda} + \Delta \bar{\lambda} \\ \eta + \Delta \bar{\eta} \end{bmatrix}.$$

Subtraction yields

$$\begin{bmatrix} C_u^* & G^* \\ C_y^* & \mathbf{w} + \Delta \mathbf{w} \\ (\eta + \Delta \eta)G & \mathbf{w} + \Delta \mathbf{w} \end{bmatrix} \begin{bmatrix} \Delta \bar{u} \\ \Delta \bar{\lambda} \\ \Delta \bar{\eta} \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ \Delta \eta * G\bar{u} + \bar{\eta} * G\Delta u \end{bmatrix}.$$

Obviously, $\Delta \bar{\lambda} = 0$. Applying Lemmas 2 and 3 and taking the supremum over $t \in [0, T]$, we get

$$\begin{aligned} \|(\Delta \bar{u}, \Delta \bar{\eta})^T\|_{(u+\Delta u, \eta+\Delta \eta)} &\leq \omega_{(u+\Delta u, \eta+\Delta \eta)} \|\Delta \eta * G\bar{u} + \bar{\eta} * G\Delta u\|_{L_\infty} \\ &\leq \omega_{(u+\Delta u, \eta+\Delta \eta)} \|(\bar{u}, \bar{\eta})^T\|_{(u,\eta)} \|(\Delta u, \Delta \eta)^T\|_{(u,\eta)}. \end{aligned} \quad (22)$$

Combining (21) and (22) and applying Corollaries 1 and 2 yields

$$\begin{aligned}
& \|r\|_{(u+\Delta u, \eta+\Delta \eta)} - \|r\|_{(u, \eta)} \\
& \leq \left| \left\| (\bar{u} + \Delta \bar{u}, \bar{\eta} + \Delta \bar{\eta})^T \right\|_{(u+\Delta u, \eta+\Delta \eta)} - \left\| (\bar{u}, \bar{\eta})^T \right\|_{(u, \eta)} \right| \\
& \leq \left| \left\| (\bar{u}, \bar{\eta})^T \right\|_{(u+\Delta u, \eta+\Delta \eta)} - \left\| (\bar{u}, \bar{\eta})^T \right\|_{(u, \eta)} \right| \\
& \quad + \omega_{(u+\Delta u, \eta+\Delta \eta)} \left\| (\bar{u}, \bar{\eta})^T \right\|_{(u, \eta)} \left\| (\Delta u, \Delta \eta)^T \right\|_{(u, \eta)} \\
& \leq \left(\sqrt{\frac{1+h}{1-h}} - 1 \right) \left\| (\bar{u}, \bar{\eta})^T \right\|_{(u, \eta)} \\
& \quad + \frac{\omega_{(u, \eta)}}{1-h} \left\| (\bar{u}, \bar{\eta})^T \right\|_{(u, \eta)} \left\| (\Delta u, \Delta \eta)^T \right\|_{(u, \eta)} \\
& \leq \left(\sqrt{\frac{1+h}{1-h}} - 1 + \frac{h}{1-h} \right) \|r\|_{(u, \eta)}.
\end{aligned}$$

Adding $\|r\|_{(u, \eta)}$ on both sides proves the result. \square

Closedness of the level sets. Having Lipschitz constants for F' and the local norm at hand, we can now establish closedness of sufficiently small level sets. This also leads to global Lipschitz constants on the level sets.

Lemma 6. *Assume $F(v^*; \mu) = 0$ and let*

$$L(\phi; \mu) := \left\{ \bar{v} \in \text{dom}(F) : \begin{array}{l} 2\omega_{(\bar{u}, \bar{\eta})} \|F(\bar{v}; \mu)\|_{(\bar{u}, \bar{\eta})} \leq \phi \\ \wedge \omega_{(\bar{u}, \bar{\eta})} \|F'(\bar{v}; \mu)(v^* - \bar{v})\|_{(\bar{u}, \bar{\eta})} \leq 1 \end{array} \right\}.$$

For $\phi < 1$ and $v \in L(\phi; \mu)$,

$$\omega_{(u, \eta)} \leq \omega(\phi, \mu) := \frac{2 - \sqrt{1 - \phi}}{\sqrt{\mu}} \quad (23)$$

holds.

Proof. In order to simplify the notation, we define $v := v^* + \Delta v$ and omit the (constant) parameter μ . Then

$$\begin{aligned}
F(v) &= F(v^*) + \int_{s=0}^1 F'(v^* + s\Delta v)\Delta v \, ds \\
&= 0 + \int_{s=0}^1 (F'(v^* + s\Delta v) - F'(v))\Delta v \, ds + F'(v)\Delta v
\end{aligned}$$

and thus

$$\begin{aligned}
\|F(v)\|_{(u, \eta)} &\geq \|F'(v)\Delta v\|_{(u, \eta)} - \int_{s=0}^1 \left\| (F'(v^* + s\Delta v) - F'(v))\Delta v \right\|_{(u, \eta)} \, ds \\
&\geq \|F'(v)\Delta v\|_{(u, \eta)} - \frac{\omega_{(u, \eta)}}{2} \|F'(v)\Delta v\|_{(u, \eta)}^2.
\end{aligned}$$

Solving this quadratic inequality yields

$$\|F'(v)\Delta v\|_{(u,\eta)} \leq \frac{1 - \sqrt{1 - \phi}}{\omega(u,\eta)} \quad \text{or} \quad \|F'(v)\Delta v\|_{(u,\eta)} \geq \frac{1 + \sqrt{1 - \phi}}{\omega(u,\eta)}.$$

By definition of $L(\phi; \mu)$, only the left hand inequality can be true, so that we have

$$\|(\Delta u, \Delta \eta)^T\|_{(u,\eta)} \leq \|F'(v)\Delta v\|_{(u,\eta)} \leq \frac{1 - \sqrt{1 - \phi}}{\omega(u,\eta)},$$

and by pointwise application of Corollary 2, left hand inequality,

$$\omega(u,\eta) \leq (2 - \sqrt{1 - \phi})\omega(u^*, \eta^*).$$

Observing that $\omega(u^*, \eta^*) = 1/\sqrt{\mu}$ completes the proof. \square

Theorem 4. Assume $F(v^*; \mu) = 0$ and let $\mathcal{L}(\alpha; \mu)$ denote the connected component around v^* of the level set $\{\bar{v} \in \text{dom}(F) : \|F(\bar{v}; \mu)\|_{(\bar{u}, \bar{\eta})} \leq \alpha\}$.

Then the level set

$$\mathcal{L}\left(\frac{\phi}{2\omega(\phi, \mu)}; \mu\right)$$

is closed and F' satisfies thereon the Lipschitz condition (7) with the Lipschitz constant $\omega(\phi, \mu)$.

Proof. First we note that by continuity of both F and the local norm, and the boundedness of $\omega(u,\eta)$ on $L(\phi; \mu)$, the level set $L(\phi; \mu)$ is closed.

For $0 < \epsilon < \phi$ assume $v \in L(\phi; \mu) \setminus L(\phi - \epsilon; \mu)$. By Lemma 6,

$$\|F(v; \mu)\|_{(u,\eta)} > \frac{\phi - \epsilon}{2\omega(u,\eta)} \geq \frac{\phi - \epsilon}{2\omega(\phi, \mu)},$$

and therefore

$$\|F(v; \mu)\|_{(u,\eta)} \leq \frac{\phi - \epsilon}{2\omega(\phi, \mu)} \quad \Rightarrow \quad v \in L(\phi - \epsilon; \mu) \vee v \notin L(\phi; \mu).$$

Since $\mathcal{L}(\cdot)$ is connected by definition, we have

$$\mathcal{L}\left(\frac{\phi - \epsilon}{2\omega(\phi, \mu)}; \mu\right) \subset L(\phi; \mu)$$

for all $0 < \epsilon < \phi$. Since $L(\rho; \mu)$ is closed and the local norm is continuous, we also have

$$\mathcal{L}\left(\frac{\phi}{2\omega(\phi, \mu)}; \mu\right) \subset L(\phi; \mu),$$

which is then closed, too.

By Lemma 6, $\omega(u,\eta) \leq \omega(\phi, \mu)$ for all $v \in \mathcal{L}(\phi/(2\omega(\phi, \mu)); \mu)$, and by Theorem 2, the Lipschitz condition (7) is then satisfied on $\mathcal{L}(\phi/(2\omega(\phi, \mu)); \mu)$ with a Lipschitz constant of $\omega(\phi, \mu)$. \square

Accuracy requirement. For the inexact Newton corrector to be applicable, we must be able to satisfy the accuracy requirement (10) in the local norm $\|\cdot\|_{(u,\eta)}$.

Theorem 5. *Using piecewise polynomial functions on a sufficiently fine grid for approximating the variables u , y , λ , and η , the accuracy requirement (9) of the inexact Newton corrector can be satisfied for arbitrary $\delta^k > 0$.*

Proof. Since $F'(u, \eta)^{-1}$ is bounded, it suffices to show that the exact Newton correction $\Delta v = -F'(u, \eta)^{-1}F(u, y, \lambda, \eta)$ is continuous and thus can be approximated to arbitrary precision by piecewise polynomials.

With $w := Gu + b$, the exact Newton correction Δv is given by

$$\begin{bmatrix} & C_u^* & G^* \\ & C_y^* & \\ C_u & C_y & \\ \eta G & & w \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta y \\ \Delta \lambda \\ \Delta \eta \end{bmatrix} = - \begin{bmatrix} -J_u + C_u^* \lambda + G^* \eta \\ -J_y + C_y^* \lambda \\ C_y y + C_u u + c \\ \eta * (Gu + b) - \mu \mathbf{1} \end{bmatrix}.$$

First we eliminate $\Delta \lambda = -\lambda + C_y^{-*} J_y$. The sum $(\xi, \xi_0) = \lambda + \Delta \lambda$ satisfies the adjoint differential equation

$$\dot{\xi}(t) = C(t)^T \xi(t) - J_y(t)$$

with the transversality boundary conditions

$$\xi(0) = R_0^T \xi_0 \quad \text{and} \quad \xi(T) = -R_T^T \xi_0.$$

From this it is immediately clear that the function component of $\Delta \lambda$ is continuous if J_y , C , and the function component of λ are continuously differentiable.

Second we eliminate $\Delta y = -y - C_y^{-1}(C_u(u + \Delta u) + c)$, which is obviously continuously differentiable if u , Δu , and c are continuous and y is continuously differentiable. Continuity of Δu will be shown below.

The reduced system now reads

$$\begin{bmatrix} & G^* \\ \eta G & w \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta \eta \end{bmatrix} = - \begin{bmatrix} -J_u + C_u^*(\lambda + \Delta \lambda) + G^* \eta \\ \eta * (Gu + b) - \mu \mathbf{1} \end{bmatrix}.$$

Elimination of $\Delta \eta = -\eta + w^{-1}(\mu \mathbf{1} + \eta * G \Delta u)$ leads to

$$\Delta u = -(G^* \eta G)^{-1} (J_u + C_u^*(\lambda + \Delta \lambda) + G^* w^{-1} \mu),$$

which is evidently continuous. \square

Newton convergence. Now that all building blocks are available, they merely need to be assembled in order to prove linear convergence of the inexact Newton iteration.

Theorem 6. *For $\phi < 1$ and $\bar{\delta} < 1$ define*

$$\gamma(\phi, \bar{\delta}) := \frac{1}{h} \left(\sqrt{\frac{1+h}{1-h}} + \frac{h}{1-h} - 1 \right) \quad (24)$$

with $h := (1 + \bar{\delta})\phi/4$. If

$$v^0 \in \mathcal{L} \left(\frac{\phi}{2\omega(\phi, \mu)(1 + \gamma(\phi, \bar{\delta})/2)}; \mu \right),$$

and the inexact Newton corrections δv^k are computed up to a relative accuracy $\delta^k \leq \bar{\delta} < 1$, then the local norm satisfies the Lipschitz condition (8) with a Lipschitz constant of $\gamma(\phi, \bar{\delta})$, and the inexact Newton corrector converges towards the central path with a contraction factor of

$$\Theta(\phi, \bar{\delta}) := \left(1 + \frac{\gamma(\phi, \bar{\delta})(1 + \bar{\delta})\kappa}{2(1 + \gamma(\phi, \bar{\delta})/2)} \right) \left(\bar{\delta} + \frac{(1 + \bar{\delta})^2\kappa}{4(1 + \gamma(\phi, \bar{\delta})/2)} \right), \quad (25)$$

where

$$\kappa := \phi \frac{\omega(\phi/(1 + \gamma(\phi, \bar{\delta})), \mu)}{\omega(\phi, \mu)} = \phi \frac{2 - \sqrt{1 - \phi/(1 + \gamma(\phi, \bar{\delta}))}}{2 - \sqrt{1 - \phi}}.$$

Proof. By Theorem 4, $\mathcal{L}_\gamma(v^0) = \mathcal{L}(\phi/(2\omega(\phi, \mu)); \mu)$ is closed, and since

$$\mathcal{L}_0(v^0) = \mathcal{L} \left(\frac{\phi/(1 + \gamma/2)}{2\omega(\phi, \mu)}; \mu \right),$$

the Lipschitz condition (7) holds with $\omega := \omega(\phi/(1 + \gamma/2), \mu) \leq \omega(\phi, \mu)$.

Next observe that $\gamma := \gamma(\phi, \bar{\delta}) \geq 2$. The Lipschitz condition (8) must be satisfied for all Δv with

$$\omega \|F'(v)\Delta v\|_{(v)} \leq \omega(\phi, \mu)(1 + \bar{\delta}) \|F(v^0)\|_{(v^0)} \leq \frac{(1 + \bar{\delta})\phi}{2(1 + \gamma/2)} \leq \frac{(1 + \bar{\delta})\phi}{4},$$

which means, by Theorem 3,

$$\sqrt{\frac{1+h}{1-h}} + \frac{h}{1-h} \leq 1 + \gamma h \quad \text{for } h \leq (1 + \bar{\delta})\phi/4.$$

Solving for γ yields (24) and verifies (8).

By Theorem 5 it is clear that the relative accuracy $\bar{\delta}$ can always be obtained. Applying Theorem 1, the contraction factor Θ is then given by

$$\begin{aligned} \Theta &\leq \left(1 + \gamma(1 + \bar{\delta})\omega \|F(v^0)\|_{(v^0)} \right) \left(\bar{\delta} + (1 + \bar{\delta})^2 \frac{\omega}{2} \|F(v^0)\|_{(v^0)} \right) \\ &\leq \left(1 + \frac{\gamma(1 + \bar{\delta})\omega\phi}{2\omega(\phi, \mu)(1 + \gamma/2)} \right) \left(\bar{\delta} + \frac{(1 + \bar{\delta})^2\omega\phi}{4\omega(\phi, \mu)(1 + \gamma/2)} \right), \end{aligned}$$

which proves (25). \square

4 Classic Continuation

In this section, we will analyze a particularly simple instantiation of Algorithm 1, which performs exactly $i_*^k \equiv j$ Newton steps with an upper bound $\bar{\delta}$ on the relative accuracy of the computation. We will derive a lower bound on the continuation step size σ such that the iterates remain in a neighborhood of the central path which is sufficiently small for the corrector to converge. We begin with an estimate for the slope of the central path.

Lemma 7. For $0 \leq \sigma < 1$ we have

$$\|F(v; (1 - \sigma/\sqrt{n_\eta})\mu)\|_{(v)} \leq \|F(v, \mu)\|_{(v)} + \beta\sigma,$$

where $\beta = \mu\omega_{(u, \eta)}$.

Proof. From $F(v; (1 - \sigma/\sqrt{n_\eta})\mu) - F(v; \mu) = (0, 0, 0, \sigma\mu\mathbf{1}/\sqrt{n_\eta})^T$ Lemma 2 yields

$$\begin{aligned} \|F(v; (1 - \sigma/\sqrt{n_\eta})\mu) - F(v; \mu)\|_{(v)} &= (\min \eta * (Gu + b))^{-1/2} \|\sigma\mu\mathbf{1}/\sqrt{n_\eta}\|_{L^\infty} \\ &= \sigma\mu\omega_{(u, \eta)} \sup_{t \in [0, T]} \|\mathbf{1}\|_2 / \sqrt{n_\eta} \\ &= \sigma\mu\omega_{(u, \eta)}, \end{aligned}$$

such that we end up with $\beta = \mu\omega_{(u, \eta)}$. \square

Theorem 7. Assume there is $\mu^0 > 0$ and v^0 such that

$$v^0 \in \mathcal{L}\left(\frac{\phi/(1 + \gamma(\phi, \bar{\delta})/2)}{2\omega(\phi, \mu^0)}; \mu^0\right)$$

for some $\phi < 1$. The continuation can be performed for $k \in \mathbb{N}$ with all continuation stepsizes $\sigma > 0$ satisfying

$$\Theta(\phi, \bar{\delta})^j \phi + 2\kappa(1 + \gamma(\phi, \bar{\delta})/2)\sigma \leq \phi\sqrt{1 - \sigma}, \quad (26)$$

where

$$\kappa := \left(2 - \sqrt{1 - \phi}\right) \left(2 - \sqrt{1 - \Theta(\phi, \bar{\delta})^j \phi / (1 + \gamma(\phi, \bar{\delta})/2)}\right).$$

Proof. In order to simplify notation, let $\gamma := \gamma(\phi, \bar{\delta})$ and $\Theta := \Theta(\phi, \bar{\delta})$. By induction, assume

$$v^k \in \mathcal{L}(\phi, \mu^k) := \mathcal{L}\left(\frac{\phi/(1 + \gamma/2)}{2\omega(\phi, \mu^k)}; \mu^k\right).$$

Theorem 6 then guarantees

$$\|F(v^{k+1}, \mu^k)\|_{(v^{k+1})} \leq \Theta^j \|F(v^k; \mu^k)\|_{(v^k)} \leq \frac{\Theta^j \phi / (1 + \gamma/2)}{2\omega(\phi, \mu^k)}$$

and thus

$$v^{k+1} \in \mathcal{L}\left(\frac{\Theta^j \phi / (1 + \gamma/2)}{2\omega(\phi, \mu^k)}; \mu^k\right).$$

Hence, by Theorem 4 and Lemma 7,

$$\omega_{(u^{k+1}, \eta^{k+1})} \leq \omega(\Theta^j \phi / (1 + \gamma/2), \mu^k)$$

and

$$\beta \leq \sqrt{\mu^k} \left(2 - \sqrt{1 - \Theta^j \phi / (1 + \gamma/2)}\right)$$

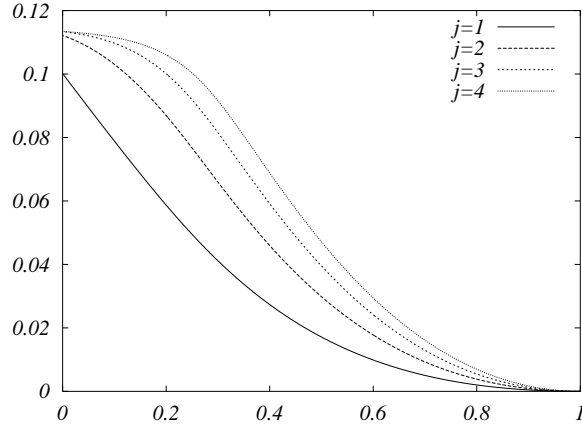


Figure 1: Continuation stepsize $\sigma(\bar{\delta})$ versus $\bar{\delta}$ for different numbers $j = 1, \dots, 4$ of Newton corrector steps.

hold. With $\mu^{k+1} = (1 - \sigma/\sqrt{n_\eta})\mu^k$, we have

$$\begin{aligned}
& 2\omega(\phi, \mu^{k+1})(1 + \gamma/2) \|F(v^{k+1}; \mu^{k+1})\|_{(v^{k+1})} \\
& \leq 2 \frac{\omega(\phi, \mu^k)}{\sqrt{1 - \sigma/\sqrt{n_\eta}}} (1 + \gamma/2) \left(\Theta^j \|F(v^k; \mu^k)\|_{(v^k)} + \beta\sigma \right) \\
& \leq 2 \frac{\omega(\phi, \mu^k)}{\sqrt{1 - \sigma}} (1 + \gamma/2) \left(\frac{\Theta^j \phi}{2\omega(\phi, \mu^k)(1 + \gamma/2)} + \mu^k \omega(\phi, \mu^k) \sigma \right) \\
& = \frac{1}{\sqrt{1 - \sigma}} (\Theta^j \phi + 2\kappa(1 + \gamma/2)\sigma) \\
& \leq \phi.
\end{aligned}$$

Since this estimate holds for arbitrary $v^k \in \mathcal{L}(\phi, \mu^k)$, v^{k+1} is in the connected component $\mathcal{L}(\phi, \mu^{k+1})$, which completes the proof. \square

Using the definitions of $\gamma(\phi, \bar{\delta})$ and $\Theta(\phi, \bar{\delta})$ in equations (24) and (25), a simple numerical computation of

$$\sigma(\bar{\delta}) = \max\{\sigma > 0 : \exists \phi < 0 \text{ such that } \sigma \text{ satisfies (26)}\}$$

reveals the results shown in Figure 1. Even for infinite dimensional problems and reasonably inexact solution of the Newton systems, reduction factors comparable to those derived for finite dimensional linear programs (see [13]) can be obtained. Taking the number of Newton corrector steps as a measure of the complexity (which is, however, not really appropriate in the adaptive grid refinement setting considered here), it becomes clear from Figure 2, that performing only one corrector step is optimal. By means of a more detailed model for effort and information gain, but optimistically assuming $\gamma = 0$, the same observation was made in [17]. Additionally, a Newton corrector accuracy of $\bar{\delta} \approx 0.1$ and a predictor target of $\omega \|F(v^k; \mu^k)\|_{(v^k)} \approx 0.9$ have been shown to be theoretically

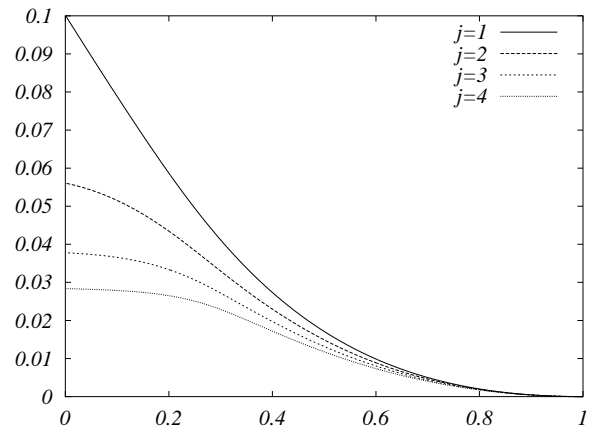


Figure 2: Averaged continuation stepsize per Newton corrector step $\sigma(\bar{\delta})/j$ versus $\bar{\delta}$ for different numbers $j = 1, \dots, 4$ of Newton corrector steps.

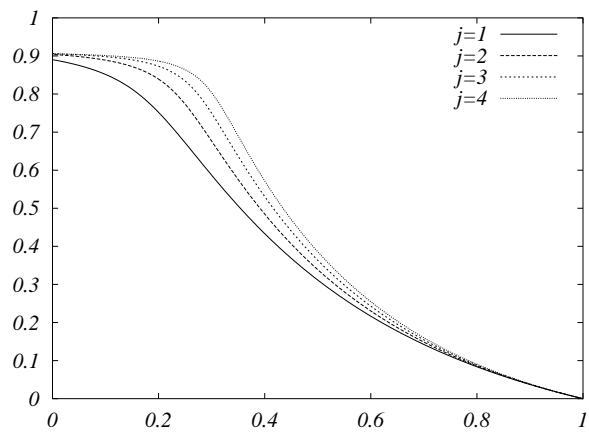


Figure 3: Upper bound on the predictor residuum.

nearly optimal. Taking the value of $\gamma \approx 2.2$ into account, this is also reflected by the present analysis, as depicted in Figure 3.

For the default accuracy of $\bar{\delta} = 0.1$, a guaranteed reduction factor of $1 - 0.079/\sqrt{n_\eta}$ is obtained.

Conclusion

The convergence rate of a primal dual interior point method of the short step pathfollowing type applied to linear control constrained optimal control problems has been analyzed. Using an affine invariant inexact Newton method and an affine invariant local norm, linear convergence with a convergence rate comparable to the ones obtained for finite dimensional linear programs was obtained. In particular, performing only one corrector step is optimal for all tolerances of the inexact Newton method. Quite large relative errors of up to 0.2 can be accepted without severely deteriorating the interior point convergence rate.

It is interesting to note that for discretized optimal control problems the reduction factor $1 - \text{const}/\sqrt{n_\eta}$ is not determined by the size n of the discrete linear program, but by the much smaller number n_η of inequality constraint components in the continuous problem.

References

- [1] V. Baryamureeba and T. Steihaug. On the convergence of an inexact primal-dual interior point method for linear programming. Report 188, University of Bergen, 2000.
- [2] S. Bellavia. Inexact interior-point method. *J. Optim. Theory Appl.*, 96(1):109–121, 1988.
- [3] P. Deuffhard. *Newton Methods for Nonlinear Problems. Affine Invariance and Adaptive Algorithms*. Springer, to be published (2002).
- [4] P. Deuffhard and G. Heindl. Affine invariant convergence theorems for Newton's method and extensions to related methods. *SIAM J. Numer. Anal.*, 16:1–10, 1979.
- [5] P. Deuffhard and M. Weiser. Local inexact Newton multilevel FEM for nonlinear elliptic problems. In M.-O. Bristeau, G. Etgen, W. Fitzgibbon, J.-L. Lions, J. Periaux, and M. Wheeler, editors, *Computational science for the 21st century*, pages 129–138. Wiley, 1997.
- [6] P. Deuffhard and M. Weiser. Global inexact Newton multilevel FEM for nonlinear elliptic problems. In W. Hackbusch and G. Wittum, editors, *Multigrid Methods V*, Lecture Notes in Computational Science and Engineering, pages 71–89. Springer, 1998.
- [7] L. Faybusovich and J. B. Moore. Infinite-dimensional quadratic optimization: Interior-point methods and control applications. *Appl. Math. Optim.*, 36(1):43–66, 1997.

- [8] R. W. Freund, F. Jarre, and S. Mizuno. Convergence of a class of inexact interior-point algorithms for linear programs. *Math. Oper. Res.*, 24(1):50–71, 1999.
- [9] A. Hohmann. *Inexact Gauss Newton Methods for Parameter Dependent Nonlinear Problems*. PhD thesis, Free University of Berlin, Dept. Math. and Comp. Sci., 1994.
- [10] S. Ito, C. T. Kelley, and E. W. Sachs. Inexact primal-dual interior point iteration for linear programs in function spaces. *Comput. Optim. Appl.*, 4:189–201, 1995.
- [11] J. Korzák. Convergence analysis of inexact infeasible-interior-point algorithms for solving linear programming problems. *SIAM J. Optim.*, 11(1):133–148, 2000.
- [12] H. Maurer and J. Zowe. First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems. *Math. Programming*, 16:98–110, 1979.
- [13] F. Potra. A path-following method for linear complementarity problems based on the affine invariant Kantorovich theorem. ZIB-Report 00-30, Zuse Institute Berlin, 2000.
- [14] M. Ulbrich and S. Ulbrich. Superlinear convergence of affine-scaling interior-point Newton methods for infinite-dimensional nonlinear problems with pointwise bounds. *SIAM J. Control Optim.*, 38(6):1938–1984, 2000.
- [15] M. Ulbrich, S. Ulbrich, and M. Heinkenschloss. Global convergence of trust-region interior-point algorithms for infinite-dimensional nonconvex minimization subject to pointwise bounds. *SIAM J. Control Optim.*, 37(3):731–764, 1999.
- [16] S. Volkwein and M. Weiser. Affine invariant convergence analysis for inexact augmented Lagrangian-SQP methods. *SIAM J. Control Optim.*, (accepted), 2002.
- [17] M. Weiser. *Function Space Complementarity Methods for Optimal Control Problems*. PhD thesis, Free University of Berlin, Dept. Math. and Comp. Sci., 2001. (<http://www.diss.fu-berlin.de/2001/189/>).
- [18] M. Weiser and P. Deuffhard. The central path towards the numerical solution of optimal control problems. ZIB-Report 01-12, Zuse Institute Berlin, 2001.
- [19] S. J. Wright. *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, 1997.