Superlinear and Quadratic Convergence of Affine-Scaling Interior-Point Newton Methods for Problems with Simple Bounds without Strct Complementarity Assumption

Matthias Heinkenschloss, Michael Ulbrich and Stefan Ulbrich

CRPC-TR97759
Revised September 1998

Center for Research on Parallel Computation
Rice University
6100 South Main Street
CRPC - MS 41
Houston, TX 77005

Submitted August 1998; Also available as Rice CAAM TR97-30 and submitted to "Mathematical Programming"

Superlinear and quadratic convergence of affine-scaling interior-point Newton methods for problems with simple bounds without strict complementarity assumption

Matthias Heinkenschloss [a,1], Michael Ulbrich [b,2] and Stefan Ulbrich [b,2]

[a] *Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77251–1892, USA.*

[b] *Lehrstuhl für Angewandte Mathematik und Mathematische Statistik, Zentrum Mathematik, Technische Universität München, 80290 München, Germany.*

Abstract

A class of affine-scaling interior-point methods for bound constrained optimization problems is introduced which are locally q–superlinear or q–quadratic convergent. It is assumed that the strong second order sufficient optimality conditions at the solution are satisfied, but strict complementarity is not required. The methods are modifications of the affine-scaling interior-point Newton methods introduced by T. F. Coleman and Y. Li (*Math. Programming*, 67:189–224, 1994). There are two modifications. One is a modification of the scaling matrix, the other one is the use of a projection of the step to maintain strict feasibility rather than a simple scaling of the step. A comprehensive local convergence analysis is given. A simple example is presented to illustrate the pitfalls of the original approach by Coleman and Li in the degenerate case and to demonstrate the performance of the fast converging modifications developed in this paper.

Key words: Bound constraints, affine scaling, interior-point algorithms, superlinear convergence, nonlinear programming, degeneracy, optimality conditions.

[1] Email: heinken@caam.rice.edu. Supported by the North Atlantic Treaty Organization under Grant CRG 960945, the DoE under Grant DE-FG03-95ER25257, and the AFSOR under Grant F49620–96–1–0329.
[2] Email: mulbrich/sulbrich@mathematik.tu-muenchen.de. Supported by the North Atlantic Treaty Organization under Grant CRG 960945.

1 Introduction

In this paper we present and analyze an extension of an affine-scaling interior-point method for the solution of simply constrained minimization problems of the form

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to } x \in \mathcal{B} \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n \;:\; a_i \le x_i \le b_i\}\,. \end{aligned} \tag{1}$$

Throughout the paper we assume that the interior $\mathcal{B}^\circ \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n \;:\; a_i < x_i < b_i\}$ of the feasible set is non–empty. We allow the strict complementarity condition at a solution $\bar{x}$ to be violated, i.e., we allow that $\bar{x}_i \in \{a_i, b_i\}$ and $\nabla f(\bar{x})_i = 0$ for some i. We call a solution $\bar{x}$ of (1) degenerate if the strict complementarity conditions are violated at $\bar{x}$.

The basis of our method is the affine-scaling interior-point method introduced in [3,4]. Several extensions to this interior-point method have been made subsequently [5,8,9,25,24,23]. Many of these works are motivated by discretized or infinite dimensional optimal control problems with bound constraints on the controls. In particular the research [24,23] on affine–scaling algorithms for infinite dimensional problems with pointwise simply constrained controls in L^p have partly motivated this work.

All papers listed above assume strict complementarity at the solution to establish local convergence results. The main purpose of this paper is to weaken this assumption and to develop an affine-scaling interior-point algorithm that converges locally q–quadratic. Strict complementarity is also required in local convergence proofs for several other interior-point approaches. We will discuss these results after a brief review of the affine-scaling interior-point method and the motivation of this paper. In this review we will make a few slight modifications to the original algorithm in [3,4] to match it with our presentation in the subsequent sections and we will concentrate on the local convergence.

Assumptions.

(A1) $f : \mathcal{D} \longrightarrow \mathbb{R}$ is twice continuously differentiable on an open neighborhood $\mathcal{D} \supset \mathcal{B}$.
(A2) The Hessian $\nabla^2 f$ is locally Lipschitz continuous on $\mathcal{B}$.

Let $\bar{x}$ be a solution of (1) and let (A1), (A2) hold. The fundamental observation in [3,4] is that the necessary optimality conditions for (1) can be written as $\bar{x} \in \mathcal{B}$ and

$$D(\bar{x})\nabla f(\bar{x}) = 0.$$

where $D(x) \stackrel{\text{def}}{=} \operatorname{diag}(d_1(x), \ldots, d_n(x))$ with

$$d_i(x) = d_i^{\mathrm{CL}}(x) \stackrel{\text{def}}{=} \begin{cases} x_i - a_i & \text{if } \nabla f(x)_i > 0, \\ b_i - x_i & \text{if } \nabla f(x)_i < 0, \\ \min\{x_i - a_i, b_i - x_i\} & \text{else.} \end{cases} \tag{2}$$

The nonlinear equation $D(x)\nabla f(x) = 0$ is solved using a modification of Newton's method. The iterates are constructed to remain in $\mathcal{B}^\circ$. At a point $x^c \in \mathcal{B}^\circ$ a formal application of the product rule yields a Newton-like step of the form

$$M(x^c)s = -D(x^c)\nabla f(x^c)$$

with $M(x) \stackrel{\text{def}}{=} D(x)\nabla^2 f(x) + E(x)$, where $E(x) \stackrel{\text{def}}{=} \text{diag}(e_1(x), \ldots, e_n(x))$ with $e_i(x) = |\nabla f(x)_i|$. The step is scaled by σt so that $x^+ = x^c + \sigma t s$ is in the interior. The step size $t \in [0,1]$ is the largest step size giving $x^c + ts \in [a,b]$ and $\sigma \in (0,1)$ is a scalar close to one. If $\bar{x}$ is a point at which the second order sufficient optimality conditions are satisfied and which is nondegenerate, i.e., at which $|\nabla f(\bar{x})_i| > 0$ whenever $\bar{x}_i \in \{a_i, b_i\}$, then one can show that $M(x)^{-1}$ exists and is uniformly bounded for all $x \in \mathcal{B}$ in a neighborhood of $\bar{x}$. Moreover, it is shown in [3] that $1 - \sigma t$ is of the order $\|x^c - \bar{x}\|$ and does not interfere with the local convergence of Newton's method if $1 - \sigma$ is of the order $\|s\|$. The affine-scaling interior-point method is locally q–quadratic convergent.

If strict complementarity is violated, the previous statements about the Coleman-Li affine-scaling interior-point method do not hold. In the degenerate case, if $x^c \to \bar{x}$, there exist indices i such that $x_i^c \to a_i$ or $x_i^c \to b_i$ and $\nabla f(x^c)_i \to 0$ at the same time. In this case, the matrix $M(x^c)$ asymptotically becomes singular. As a result, the step sizes t may become very small. In general, this will prohibit fast local convergence, cf. the second example in §6. It is worth noting that even if the unit stepsize can be used, the convergence still may be only linear. This is illustrated by the first example in §6.

The purpose of this paper is to modify the affine-scaling interior-point method so that fast local convergence in the degenerate case can be maintained. This will be accomplished by two modifications. The first one concerns the scaling matrix $D(x)$ and/or the matrix $E(x)$. Roughly speaking, we will modify them so that scaling will be switched off for components in which degeneracy is detected. This detection will be done automatically. Secondly, the scaling of the step by σt to keep the new iterate in the interior will be replaced by a projection. Since the scaling t is determined by the components s_i that are the most infeasible compared to their size, 'bad components' in s may result in a very small t, which is then also applied to the 'good components' of s, i.e. the ones which are sufficiently interior. We will show in this paper, cf. (46), that the step sizes leading to small scalings t are small relative to the 'good components' of s. The projection cuts off 'bad components', but leaves 'good components' unchanged or changes them only slightly and therefore preserves fast convergence. We will prove local convergence with q–order $\min\{2, p\}$, where $p > 1$ is a fixed parameter used in the modification of scaling matrices $D(x)$ and $E(x)$. We emphasize that the use of a projection does not destroy the descent property of the steps, see §5.

Our local convergence analysis differs from those in, e.g., [3,4] even in the nondegenerate case. Our modifications for the degenerate case are partly motivated by our analyses for the infinite dimensional case in [23,24]. For example, using a projection instead of simple scaling to make x^+ an interior point is already featured in [23]. However, the analysis in the present paper is different and, in some parts, uses proof techniques that do not carry over to the infinite dimensional framework.

Of course, the affine-scaling interior-point method is only one interior-point approach that can be used to solve problems like (1). Interior-point methods for nonlinear programming problems are a very active research area. In the following we will only discuss a few papers on this topic that discuss local convergence results. Strict complementarity is assumed in [10] and [28] to establish local convergence proofs of primal-dual Newton and quasi-Newton interior-point methods. See also [25]. In [2] a trust-region Dikin-Karmarkar scaling interior-point method is introduced for problems of the type (1) with additional linear equality constraints. Local convergence is proven without strict complementarity requirement but assuming strict second order sufficient optimality. However, since a Dikin-Karmarkar scaling is used, local convergence is slow. The algorithm can not be expected to converge superlinearly [2,22]. Strict complementarity is used to prove convergence of the barrier method with logarithmic barrier function [26, Thm. 3.6]. For linear and quadratic programs the structure of the problems can be used to make stronger statements about the convergence of interior-point methods. For example, in the case of linear programs there always exists at least one solution which satisfies strict complementarity, e.g., [27, Thm. 2.4], and there are many results which prove convergence of interior-point methods towards one such solution – the analytic center of the solution set. See, e.g., [17] for a discussion. For quadratic programs and linear complementarity problems (LCPs) strict complementarity may not be satisfied at any solution. In this case, interior-point methods may convergence significantly slower. For example, [11,20] consider a class of Newton interior-point methods for monotone LCPs and show that if there exists no solution that satisfies strict complementarity, then the duality gap converges only q–linearly with factor $1/4$ instead of q–superlinearly.

In addition to interior-point approaches, other methods can be used to solve problems like (1). In particular trust-region modifications of the projected gradient and projected Newton method [1] are often used. In [18] a modification of the trust region algorithm [6,7] is proposed that converges locally q–quadratic under assumptions comparable to ours. [3]

Finally, we remark that there is a class of optimization methods for nonlinear programs which for (1) requires conditions comparable to the ones used in this paper. These methods are based on equivalently rewriting the necessary Karush-Kuhn-Tucker optimality conditions as a nonlinear system. This system is then solved by a Newton type method [15]. Local q–quadratic convergence of this Newton-type method can be proven without strict complementarity assumptions, but with the strong second order sufficient optimality condition stated in Theorem 2 below. See, e.g, [12,16] for recent developments in this direction.

This paper is organized as follows. Section 2 introduces the basics of the algorithm. In it the affine-scaling framework is stated with minimal requirements on the matrices $D(x)$ and $E(x)$. Section 3 contains several convergence estimates. These will form the basis for the final convergence analysis of our algorithm. The algorithm, in particular our precise

[3] The report [19], published after the submission of this paper, extends these results. For problems (1) the assumptions in [19] for local convergence are comparable to ours.

choices of the matrices $D(x)$ and $E(x)$ will be given in Section 4. This section also contains the main convergence result. In Section 5 we show that our convergence result remains valid if the Newton system is solved inexactly with appropriately controlled accuracy. Moreover, we sketch a trust region globalization. An illustrative example is presented in Section 6.

Notations. We use $\bar{x}$, x^c, x^+ to denote a solution of (1), a current iterate of the algorithm, and the new iterate of the algorithm respectively. The matrices $D(x)$ and $E(x)$ are diagonal matrices generated by vectors $d(x)$ and $e(x)$, respectively. Analogously, we will generate a diagonal matrix $W(x)$ from a vector $w(x)$. We recall that $\mathcal{B}^\circ \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n \ : \ a_i < x_i < b_i\}$.

2 Optimality conditions and basic formulation of the algorithm

We first review optimality conditions for (1).

Theorem 1 (First order necessary conditions)
If the assumption (A1) *holds and $\bar{x}$ is a local solution of* (1), *then $\bar{x} \in \mathcal{B}$, and*

$$\text{(ON)} \qquad \nabla f(\bar{x})_i \begin{cases} = 0 & \text{if } a_i < \bar{x}_i < b_i, \\ \geq 0 & \text{if } \bar{x}_i = a_i, \\ \leq 0 & \text{if } \bar{x}_i = b_i. \end{cases}$$

For $\bar{x} \in \mathcal{B}$ we decompose $\{1, \ldots, n\}$ into the active index set with strict complementarity $\bar{A}$, the inactive set $\bar{I}$, and the active set without strict complementarity $\bar{N}$, i.e.,

$$\begin{aligned} \bar{A} &\stackrel{\text{def}}{=} \{i \in \{1, \ldots, n\} \ : \ \bar{x}_i \in \{a_i, b_i\}, \ \nabla f(\bar{x})_i \neq 0\}, \\ \bar{I} &\stackrel{\text{def}}{=} \{i \in \{1, \ldots, n\} \ : \ a_i < \bar{x}_i < b_i\}, \\ \bar{N} &\stackrel{\text{def}}{=} \{i \in \{1, \ldots, n\} \ : \ \bar{x}_i \in \{a_i, b_i\}, \ \nabla f(\bar{x})_i = 0\}. \end{aligned}$$

Theorem 2 (Strong second order sufficient optimality conditions)
Let the assumption (A1) *hold. If the conditions in Theorem* 1 *are satisfied and if*

(OS) *there exists $\alpha > 0$ such that $s^T \nabla^2 f(\bar{x}) s \geq \alpha \|s\|_2^2$ for all $s \in T(\bar{x})$,*

where $T(\bar{x}) \stackrel{\text{def}}{=} \left\{ s \in \mathbb{R}^n \ : \ s_i = 0 \ \forall \ i \in \bar{A} \right\}$, then $\bar{x}$ is a local solution of (1).

The name strong second order sufficient optimality condition is, to our knowledge, due to [21, p. 55]. It is important for sensitivity analysis if the strict complementarity condition is violated, see, e.g., [13, p. 43]. It is also used in the convergence analyses in, e.g, [2,15,18,19]. If strict complementarity holds at $\bar{x}$, i.e., if $\bar{N} = \emptyset$, then the second order sufficient optimality condition in Theorem 2 is the standard one. If $\bar{N} \neq \emptyset$, then the second order sufficient optimality condition, e.g., [14, Thm. 4], only requires that $s^T \nabla^2 f(\bar{x}) s \geq \alpha \|s\|_2^2$ for all s in the cone $\left\{ s \in T(\bar{x}) \ : \ s_i \geq 0 \text{ if } i \in \bar{N} \text{ and } \bar{x}_i = a_i, \ s_i \leq 0 \text{ if } i \in \bar{N} \text{ and } \bar{x}_i = b_i \right\}$.

By introducing a scaling-matrix $D(x) \stackrel{\text{def}}{=} \operatorname{diag}(d_1(x), \ldots, d_n(x))$ with

$$d_i(x) \begin{cases} = 0 & \text{if } x_i = a_i \text{ and } \nabla f(x)_i > 0, \\ = 0 & \text{if } x_i = b_i \text{ and } \nabla f(x)_i < 0, \\ \geq 0 & \text{if } x_i \in \{a_i, b_i\} \text{ and } \nabla f(x)_i = 0, \\ > 0 & \text{else,} \end{cases} \tag{3}$$

the condition (ON) can be written in the form

$$D(\bar{x}) \nabla f(\bar{x}) = 0. \tag{4}$$

The affine-scaling interior-point method is a Newton-like method applied to solve (4) while maintaining $x \in \mathcal{B}^\circ$ for the iterates. A formal linearization of $D(x)\nabla f(x)$ around x^c leads to the Newton-like system

$$M(x^c)(x^+ - x^c) = -D(x^c)\nabla f(x^c). \tag{5}$$

Here

$$M(x) \stackrel{\text{def}}{=} D(x)\nabla^2 f(x) + E(x) \tag{6}$$

is obtained by formal application of the product rule and $E(x)$ is a substitute for the in general not existing derivative $\frac{\mathrm{d}}{\mathrm{d}\tilde{x}}(D(\tilde{x})\nabla f(x))|_{\tilde{x}=x}$. We choose $E(x)$ as a diagonal matrix $E(x) \stackrel{\text{def}}{=} \operatorname{diag}(e_1(x), \ldots, e_n(x))$ with

$$e_i(x) = d_i'(x)\nabla f(x)_i, \quad i = 1, \ldots, n. \tag{7}$$

In (7) the real valued functions $d_i'(x)$ are substitutes for the not everywhere existing derivatives $\frac{\partial}{\partial x_i} d_i(x)$. We assume there exists $c_{d'} > 0$ such that

$$0 \leq e_i(x), \quad |d_i'(x)| \leq c_{d'}, \quad i = 1, \ldots, n, \tag{8}$$

for all $x \in \mathcal{B}$. If we choose the Coleman and Li scaling $d(x) = d^{\mathrm{CL}}(x)$ defined in (2), then we set

$$d_i'(x) = \operatorname{sgn}(\nabla f(x)_i), \quad i = 1, \ldots, n. \tag{9}$$

If $x \in \mathcal{B}^\circ$ and $\nabla f(x)_i \neq 0$, then (9) is the actual derivative of (2) with respect to x_i. With the choice (9) the entries (7) of $E(x)$ are given by

$$e_i(x) = |\nabla f(x)_i|, \quad i = 1, \ldots, n. \tag{10}$$

The solution x^+ of (5) in general does not satisfy $x^+ \in \mathcal{B}^\circ$. To complete the description of the local algorithm one has to describe how the solution x^+ of (5) is modified to

ensure that $x^+ \in \mathcal{B}^\circ$. We will discuss this issue in Section 4. First, we will analyze the iteration (5).

3 Basic convergence estimates

Let the necessary optimality conditions in Theorem 1 hold at $\bar{x}$. With (4) we obtain the trivial equality $M(x^c)(\bar{x} - \bar{x}) = -D(\bar{x})\nabla f(\bar{x})$. Subtracting this equality from (5) yields

$$M(x^c)(x^+ - \bar{x}) = r(x^c) \tag{11}$$

with remainder term

$$r(x) = D(\bar{x})\nabla f(\bar{x}) - D(x)\nabla f(x) - M(x)(\bar{x} - x). \tag{12}$$

The following lemma provides important estimates for r.

Lemma 3 i. *Let the assumption* (A1) *hold and let $\bar{x}$ satisfy the first order necessary conditions in Theorem* 1. *If D, M, E are matrices satisfying* (3), (6), (7) *and if r is given by* (12), *then for all $x \in \mathcal{B}$ and $1 \le i \le n$*

$$\begin{aligned} |r_i(x)| \le\ & d_i(x)\,|(\nabla f(\bar{x}) - \nabla f(x) - \nabla^2 f(x)(\bar{x} - x))_i| \\ & + |\nabla f(\bar{x})_i(d_i(\bar{x}) - d_i(x)) - e_i(x)(\bar{x}_i - x_i)|\,. \end{aligned} \tag{13}$$

ii. *If, in addition to the assumptions in* i., (A2) *holds, if* (8) *is satisfied, and if $d_i(x)$, $e_i(x)$ are defined by* (2), (10), *respectively, for all $i \in \bar{A}$, then there exist $\rho > 0$ and $L_2 > 0$ such that for all $x \in \mathcal{B}$ with $\|x - \bar{x}\|_2 < \rho$ the following estimate is satisfied:*

$$|r_i(x)| \le L_2\, d_i(x)\|x - \bar{x}\|_2^2 + \max\{1, c_{d'}\}\, |\nabla f(x)_i - \nabla f(\bar{x})_i|\, |x_i - \bar{x}_i| \tag{14}$$

for $i = 1, \ldots, n$, where $c_{d'}$ is given by (8).
iii. *Under the assumptions in* ii. *the following inequality is valid:*

$$|r_i(x)| \le L_2\, d_i(x)\|x - \bar{x}\|_2^2 \quad \forall i \in \bar{N} \text{ with } e_i(x) = 0. \tag{15}$$

Proof. i. Using (6) and (12) we obtain

$$\begin{aligned} r_i(x) =\ & d_i(x)(\nabla f(\bar{x}) - \nabla f(x) - \nabla^2 f(x)(\bar{x} - x))_i \\ & + \nabla f(\bar{x})_i\,(d_i(\bar{x}) - d_i(x)) - e_i(x)(\bar{x}_i - x_i), \quad i = 1, \ldots, n. \end{aligned}$$

This implies (13).

ii. By continuity there exists $\rho > 0$ such that for all $x \in \mathcal{B}$ with $\|x - \bar{x}\|_2 < \rho$ the implication

$$\nabla f(\bar{x})_i \neq 0 \implies \nabla f(\bar{x})_i \nabla f(x)_i > 0$$

holds. Since $d_i(x)$ is defined by (2) and $\nabla f(\bar{x})_i \neq 0$, $d_i(\bar{x}) = 0$ if $i \in \bar{A}$, we obtain for all $x \in \mathcal{B}$ with $\|x - \bar{x}\|_2 < \rho$ and all $i \in \bar{A}$ that $d_i(\bar{x}) - d_i(x) = \operatorname{sgn}(\nabla f(\bar{x})_i)(\bar{x}_i - x_i)$. Since $\nabla f(\bar{x})_i = 0$ for all $i \in \{1, \dots, n\} \setminus \bar{A}$, we have for all $i \in \{1, \dots, n\}$

$$\nabla f(\bar{x})_i (d_i(\bar{x}) - d_i(x)) = \begin{cases} |\nabla f(\bar{x})_i|(\bar{x}_i - x_i) & \text{if } i \in \bar{A}, \\ 0 & \text{else.} \end{cases} \tag{16}$$

Similarly, if (8) is satisfied and if for all $i \in \bar{A}$ the scalars $e_i(x)$ are defined by (10), then

$$e_i(x)(\bar{x}_i - x_i) = |\nabla f(x)_i|(\bar{x}_i - x_i) \tag{17}$$

for all $i \in \bar{A}$ and

$$|e_i(x)(\bar{x}_i - x_i)| \leq c_{d'} |\nabla f(x)_i| \, |\bar{x}_i - x_i| = c_{d'} |\nabla f(x)_i - \nabla f(\bar{x})_i| \, |\bar{x}_i - x_i| \tag{18}$$

for all $i \in \{1, \dots, n\} \setminus \bar{A}$. Using (13), (16), and (17), we find that for $i \in \bar{A}$

$$\begin{aligned} |r_i(x)| &\leq d_i(x) \left| (\nabla f(\bar{x}) - \nabla f(x) - \nabla^2 f(x)(\bar{x} - x))_i \right| \\ &\quad + \left| |\nabla f(\bar{x})_i| - |\nabla f(x)_i| \right| |\bar{x}_i - x_i| \\ &\leq d_i(x) \left| (\nabla f(\bar{x}) - \nabla f(x) - \nabla^2 f(x)(\bar{x} - x))_i \right| \\ &\quad + |\nabla f(x)_i - \nabla f(\bar{x})_i| \, |\bar{x}_i - x_i| \, . \end{aligned}$$

From (13), (16), and (18) we deduce that for $i \in \{1, \dots, n\} \setminus \bar{A}$

$$\begin{aligned} |r_i(x)| &\leq d_i(x) \, |(\nabla f(\bar{x}) - \nabla f(x) - \nabla^2 f(x)(\bar{x} - x))_i| \\ &\quad + c_{d'} \, |\nabla f(x)_i - \nabla f(\bar{x})_i| \, |\bar{x}_i - x_i| \, . \end{aligned}$$

For $\rho > 0$ sufficiently small, assumption (A2) provides us with a Lipschitz constant L_2 of $\nabla^2 f$ on $\{x \in \mathcal{B} \,:\, \|x - \bar{x}\|_2 < \rho\}$. Thus

$$\left| (\nabla f(\bar{x}) - \nabla f(x) - \nabla^2 f(x)(\bar{x} - x))_i \right| \leq L_2 \|x - \bar{x}\|_2^2, \tag{19}$$

and the assertion (14) is proven.

iii. Let $i \in \bar{N}$ and $e_i(x) = 0$. The assertion (15) follows from (13), (16), and (19), since $\nabla f(\bar{x})_i = 0$ for $i \in \bar{N}$. $\square$

If a second order sufficiency condition with strict complementarity holds at $\bar{x}$ then Lemma 3 indicates the second order convergence of the Coleman and Li algorithm in which $d_i(x)$, $e_i(x)$ are defined by (2), (10). As we have stated earlier, in this case $M(x)^{-1}$ exists and is uniformly bounded in a neighborhood of $\bar{x}$. The estimate (14) implies that x^+ defined by (5) obeys $\|x^+ - \bar{x}\|_2 \leq C\|x^c - \bar{x}\|_2^2$, provided that $x^c \in \mathcal{B}^\circ$ is sufficiently close to $\bar{x}$. Moreover, the scalar σt in $x^c + \sigma t(x^+ - x^c)$ converges sufficiently fast to one so that it does not interfere with the q–quadratic convergence.

If strict complementarity is violated at $\bar{x}$, then the set $\bar{N}$ is nonempty and for all $i \in \bar{N}$ we have $d_i(\bar{x}) = e_i(\bar{x}) = 0$. The uniform boundedness of $M(x)^{-1}$ is no longer ensured in any neighborhood of $\bar{x}$. Assume, for example, that $\bar{x}_i = a_i$, $\nabla f(\bar{x})_i = 0$ for some i and that there is a sequence $\{x^k\}$ with $x^k \in \mathcal{B}^\circ$, $\lim_{k\to\infty} x^k = \bar{x}$, and $\nabla f(x^k)_i > 0$, $d_i(x^k) = d_i^{CL}(x^k)$. Then $d_i(x^k) = x_i^k - a_i = x_i^k - \bar{x}_i$ and $\lim_{k\to\infty} d_i(x^k) = 0$, $\lim_{k\to\infty} e_i(x^k) = \lim_{k\to\infty} d_i'(x^k)\nabla f_i(x^k) = 0$. Consequently, the definition (6) of $M(x^k)$ shows that the ith row of $M(x^k)$ tends to zero and thus

$$\lim_{k\to\infty} \|M(x^k)^{-1}\| = \infty.$$

To overcome this difficulty we premultiply equation (5) by the diagonal matrix $W(x^c)$ with $W : \mathcal{B}^\circ \longrightarrow \mathbb{R}^{n\times n}$, $W(x) = \mathrm{diag}(w_1(x), \ldots, w_n(x))$,

$$w_i(x) \stackrel{\text{def}}{=} \frac{1}{d_i(x) + e_i(x)}.$$

Note that $W(x)$ is well-defined on $\mathcal{B}^\circ$, since $d_i(x) > 0$ and $e_i(x) \geq 0$, $i = 1, \ldots, n$, for $x \in \mathcal{B}^\circ$. We introduce the matrix

$$H(x) \stackrel{\text{def}}{=} W(x)M(x) = \mathrm{diag}\left(\frac{d_i(x)}{d_i(x) + e_i(x)}\right)\nabla^2 f(x) + \mathrm{diag}\left(\frac{e_i(x)}{d_i(x) + e_i(x)}\right), \quad (20)$$

and we use it to equivalently rewrite (11) as

$$H(x^c)(x^+ - \bar{x}) = W(x^c)r(x^c). \quad (21)$$

We first show that $H(x)^{-1}$ exists and is uniformly bounded for all $x \in \mathcal{B}^\circ$, $\|x - \bar{x}\|_2 < \rho$, if $\bar{x}$ satisfies (OS) and if $\rho > 0$ is small enough.

Lemma 4 *Let the assumption* (A1) *hold and let $\bar{x} \in \mathcal{B}$ satisfy the strong second order sufficient optimality conditions stated in Theorem 2. Moreover, let D, M, E satisfy* (3), (6), (7), *and* (8) *and let H be defined by* (20). *If there exists $\bar{\rho} > 0$ such that for all $x \in \mathcal{B}^\circ$ with $\|x - \bar{x}\|_2 < \bar{\rho}$ and all $i \in \bar{A}$ the diagonals $d_i(x)$ are defined by* (2) *and the diagonals $e_i(x)$ are defined by* (10), *then for $\rho \in (0, \bar{\rho})$ sufficiently small and all $x \in \mathcal{B}^\circ$ with $\|x - \bar{x}\|_2 < \rho$ the matrix $H(x)$ is nonsingular with*

$$\|H(x)^{-1}\|_2 \leq \frac{1 + C_H + \alpha/2}{\alpha/2} \stackrel{\text{def}}{=} C_{H^{-1}}, \quad (22)$$

where α is the constant in (OS) *and* $C_H = \sup_{\|x-\bar{x}\|_2<\rho} \|\nabla^2 f(x)\|_2$.

Proof. For $s \in \mathbb{R}^n$ we set

$$s_d \stackrel{\text{def}}{=} W(x)D(x)s, \quad s_e \stackrel{\text{def}}{=} W(x)E(x)s.$$

i. First we show that for all $\kappa > 0$ there is $\rho > 0$ such that for all $x \in \mathcal{B}^\circ$ with $\|x-\bar{x}\|_2 < \rho$

$$s_d^T \nabla^2 f(x) s_d \geq \frac{\alpha}{2} \|s_d\|_2^2 \quad \forall s \in \mathbb{R}^n \text{ with } \|s_d\|_2 \geq \kappa \|s\|_2. \tag{23}$$

To prove (23), we choose $\rho \in (0, \bar{\rho})$ so small that

$$\max_{i\in\bar{A}} |\nabla f(x)_i - \nabla f(\bar{x})_i| \leq \frac{1}{2} \min_{i\in\bar{A}} |\nabla f(\bar{x})_i| \stackrel{\text{def}}{=} \gamma \tag{24}$$

for all $x \in \mathcal{B}^\circ$ with $\|x-\bar{x}\|_2 < \rho$. Let $x \in \mathcal{B}^\circ$ with $\|x-\bar{x}\|_2 < \rho$ be arbitrary. From (24) we obtain

$$\min_{i\in\bar{A}} |\nabla f(x)_i| \geq \gamma. \tag{25}$$

As a consequence, $\nabla f(x)_i \nabla f(\bar{x})_i > 0$ for all $i \in \bar{A}$, and, since $d_i(x)$, $i \in \bar{A}$, is defined by (2), it follows that

$$|d_i(x)| = |x_i - \bar{x}_i| < \rho \quad \forall\, i \in \bar{A}. \tag{26}$$

The vector s_T with

$$(s_T)_i = \begin{cases} (s_d)_i & \text{if } i \notin \bar{A}, \\ 0 & \text{else} \end{cases}$$

is an element of the linear space $T(\bar{x})$ defined in Theorem 2. Using (25), (26), and the fact that $e_i(x) = |\nabla f(x)_i|$ for $i \in \bar{A}$, we obtain an estimate for $s_{T^c} \stackrel{\text{def}}{=} s_d - s_T$:

$$\|s_{T^c}\|_2 \leq \max_{i\in\bar{A}} \left| \frac{d_i(x)}{d_i(x) + |\nabla f(x)_i|} \right| \|s\|_2 \leq \frac{\rho}{\gamma} \|s\|_2 \leq \frac{\rho}{\kappa\gamma} \|s_d\|_2. \tag{27}$$

Since $s_T^T s_{T^c} = 0$, we have by (27)

$$\|s_T\|_2^2 = \|s_d\|_2^2 - \|s_{T^c}\|_2^2 \geq \left(1 - \frac{\rho^2}{\kappa^2\gamma^2}\right) \|s_d\|_2^2.$$

Moreover, (A1) and (OS) guarantee that

$$s_T^T \nabla^2 f(x) s_T \geq \frac{3}{4}\alpha \|s_T\|_2^2 \tag{28}$$

for all $x \in \mathcal{B}$ with $\|x - \bar{x}\|_2 < \rho$, provided ρ is sufficiently small. Using the previous two estimates, (27), and (28), we get

$$\begin{aligned}
s_d^T \nabla^2 f(x) s_d &= s_T^T \nabla^2 f(x) s_T + s_{T^c}^T \nabla^2 f(x)(s_T + s_d) \\
&\geq \frac{3}{4}\alpha \|s_T\|_2^2 - \frac{2\rho}{\kappa\gamma}\|\nabla^2 f(x)\|_2 \|s_d\|_2^2 \\
&\geq \frac{3}{4}\alpha \left(1 - \frac{\rho^2}{\kappa^2\gamma^2}\right)\|s_d\|_2^2 - \frac{2\rho}{\kappa\gamma}\|\nabla^2 f(x)\|_2 \|s_d\|_2^2.
\end{aligned}$$

Together with (A1), the previous estimate implies (23) for sufficiently small $\rho > 0$.

ii. By (A1) there is a constant $C_H > 0$ such that for sufficiently small $\rho > 0$,

$$\|\nabla^2 f(x)\|_2 \leq C_H \quad \forall\, x \in \mathcal{B} \text{ with } \|x - \bar{x}\|_2 < \rho.$$

We set $\kappa = 1/(1 + C_H + \alpha/2)$. We have shown in part i. that – possibly after reducing $\rho > 0$ – (23) is satisfied for all $x \in \mathcal{B}^\circ$ with $\|x - \bar{x}\|_2 < \rho$. We now show that for all these x

$$\|H(x)^T s\|_2 \geq \frac{\alpha/2}{1 + C_H + \alpha/2}\|s\|_2 \quad \forall\, s \in \mathbb{R}^n. \tag{29}$$

This inequality proves that

$$\|H(x)^{-1}\|_2 = \|H(x)^{-T}\|_2 \leq \frac{1 + C_H + \alpha/2}{\alpha/2}$$

as asserted. To prove (29), let $x \in \mathcal{B}^\circ$ with $\|x - \bar{x}\|_2 < \rho$ be arbitrary. For $s \in \mathbb{R}^n$ we consider two cases.

1. Let $\|s_d\|_2 \geq \kappa\|s\|_2$. From the nonnegativity of $d_i(x)$, $e_i(x)$, and $w_i(x)$ we find that $s_d^T s_e = s^T D(x) W^2(x) E(x) s \geq 0$. This inequality and (23) yield

$$\|s_d\|_2 \|H(x)^T s\|_2 \geq s_d^T H(x)^T s = s_d^T \nabla^2 f(x) s_d + s_d^T s_e \geq s_d^T \nabla^2 f(x) s_d \geq \frac{\alpha}{2}\|s_d\|_2^2.$$

Hence,

$$\|H(x)^T s\|_2 \geq \frac{\alpha}{2}\|s_d\|_2 \geq \frac{\alpha}{2}\kappa\|s\|_2 = \frac{\alpha/2}{1 + C_H + \alpha/2}\|s\|_2.$$

2. Let $\|s_d\|_2 \leq \kappa\|s\|_2$. Then

$$\|s_e\|_2 \|H(x)^T s\|_2 \geq s_e^T H(x)^T s = s_e^T \nabla^2 f(x) s_d + s_e^T s_e \geq \|s_e\|_2^2 - C_H \|s_d\|_2 \|s_e\|_2.$$

Since $s = s_d + s_e$ we have $\|s_e\|_2 \geq \|s\|_2 - \|s_d\|_2$ and thus

$$\begin{aligned}\|H(x)^T s\|_2 &\geq \|s_e\|_2 - C_H\|s_d\|_2 \geq \|s\|_2 - (1+C_H)\|s_d\|_2 \\ &\geq (1-\kappa(1+C_H))\|s\|_2 = \frac{\alpha/2}{1+C_H+\alpha/2}\|s\|_2.\end{aligned}$$

The proof is complete. $\square$

As a first result we get the following error bound.

Theorem 5 *Let* (A1), (A2) *hold and let the strong second order sufficient optimality conditions stated in Theorem 2 be satisfied at $\bar{x}$. Furthermore, let D, M, E satisfy* (3), (6) (7), *and* (8). *If there exists $\bar{\rho} > 0$ such that for all $x \in \mathcal{B}^\circ$ with $\|x - \bar{x}\|_2 < \rho$ and all $i \in \bar{A}$ the diagonals $d_i(x)$ are defined by* (2) *and the diagonals $e_i(x)$ are defined by* (10), *then there exists $C > 0$ such that for $\rho \in (0, \bar{\rho})$ small enough* (21) *admits a unique solution x^+ and the estimate*

$$\|x^+ - \bar{x}\|_2 \leq C\left(\|x^c - \bar{x}\|_2^2 + \|\tilde{r}(x^c)\|_2\right), \tag{30}$$

where

$$\tilde{r}_i(x) = \begin{cases} 0 & \text{if } i \in \bar{N} \text{ and } e_i(x) = 0, \\ \dfrac{|\nabla f(x)_i - \nabla f(\bar{x})_i|\,|x_i - \bar{x}_i|}{d_i(x) + e_i(x)} & \text{otherwise,} \end{cases} \tag{31}$$

is valid for all $x^c \in \mathcal{B}^\circ$ with $\|x^c - \bar{x}\|_2 < \rho$.

Proof. Since $d_i(x^c) > 0$ for all $i \in \{1, \ldots, n\}$, the matrix $W(x^c)$ is well defined and the equations (5), (11), and (21) are equivalent. For sufficiently small $\rho > 0$, Lemma 3 and 4 are applicable. The estimates (14), (15) imply that

$$\begin{aligned}|(W(x^c)r(x^c))_i| &\leq L_2\,\frac{d_i(x^c)}{d_i(x^c)+e_i(x^c)}\|x^c - \bar{x}\|_2^2 + \max\{1, c_{d'}\}|\tilde{r}_i(x^c)| \\ &\leq L_2\,\|x^c - \bar{x}\|_2^2 + \max\{1, c_{d'}\}|\tilde{r}_i(x^c)|.\end{aligned} \tag{32}$$

Now (21) and (22) yield (30) with $C = C_{H^{-1}} \max\{\sqrt{n}L_2, 1, c_{d'}\}$. $\square$

Note that if $e_i(x)$ is given by (10), then

$$\tilde{r}_i(x) = \frac{|\nabla f(x)_i - \nabla f(\bar{x})_i|\,|x_i - \bar{x}_i|}{d_i(x) + |\nabla f(x)_i|}. \tag{33}$$

4 Formulation of the algorithm and local convergence analysis

The results in Lemmas 3, 4 and Theorem 5 indicate that the Coleman-Li scaling and the corresponding E given by (10) are good choices for $i \in \bar{A}$. We will see in a moment that they are also good choices for indices $i \in \bar{I}$. However, if we use the Coleman-Li scaling matrix (2) and the corresponding matrix E given by (10) for all indices, then the estimate (30) with $\tilde{r}$ given by (33) indicates that the term $\|\tilde{r}\|_2$ in (30) may only be of first order in $\|x^c - \bar{x}\|_2$ if $|\nabla f(x)_i|$ and $d_i(x)$ are of the same order of magnitude. But for $i \in \bar{N}$ this may happen if $d_i(x)$ and $e_i(x)$ are defined by (2), (10), respectively. On the other hand, for $i \in \bar{A} \cup \bar{I}$, and $d_i(x)$ according to (2) there is $\varepsilon > 0$ such that $d_i(x^c) + |\nabla f(x^c)_i| \geq \varepsilon$ for $\|x^c - \bar{x}\|_2$ small, yielding a quadratic order of these components of $\tilde{r}(x^c)$.

These are the key observations for the development of a fast convergent method in the case without strict complementarity. We must control the components of $\tilde{r}(x)$ with indices $i \in \bar{N}$ by suitable choice of $d_i(x)$ and/or $e_i(x)$.

This can be done in various ways. We propose two choices.

Scaling matrices I. For $p > 1$ set

$$d_i(x) = \begin{cases} d_i^{\mathrm{CL}}(x) & \text{if } |\nabla f(x)_i| < \min\{x_i - a_i, b_i - x_i\}^p \text{ or} \\ & \text{if } \min\{x_i - a_i, b_i - x_i\} < |\nabla f(x)_i|^p \\ 1 & \text{else.} \end{cases} \tag{34}$$

The matrix $E(x)$ is either defined by (10) or by

$$e_i(x) = \begin{cases} |\nabla f(x)_i| & \text{if } |\nabla f(x)_i| < \min\{x_i - a_i, b_i - x_i\}^p \text{ or} \\ & \text{if } \min\{x_i - a_i, b_i - x_i\} < |\nabla f(x)_i|^p \\ 0 & \text{else.} \end{cases} \tag{35}$$

Scaling matrices II. The matrix $D(x)$ is defined by (2) and E is defined by (35) with some fixed $p > 1$.

If $i \in \bar{N}$ and $|\nabla f(x)_i| \sim |x_i - \bar{x}_i| \ll 1$ then $|\nabla f(x)_i| \geq |x_i - \bar{x}_i|^p = \min\{x_i - a_i, b_i - x_i\}^p$ and $\min\{x_i - a_i, b_i - x_i\} = |x_i - \bar{x}_i| \geq |\nabla f(x)_i|^p$. Hence, the choice (34) of D ensures that for these components, which turn out to be the pathological ones, see the proof of Theorem 8, the Coleman-Li scaling is eventually switched off. On the other hand, we will show in Lemma 7 below that for all indices $i \in \bar{A} \cup \bar{I}$ the diagonal entries $d_i(x)$ and $e_i(x)$ are chosen according to (2) and (7), respectively, if x is sufficiently close to $\bar{x}$. The choice (10) of E is the standard choice in the Coleman and Li setting, the choice (35) of E corresponds to the 'derivative' of D in (34).

First, we remark that the definition (34) of D is admissible, i.e. satisfies (3). This can

be seen readily from (34).

Lemma 6 *The definition* (34) *of* d_i *satisfies the requirement* (3).

Lemma 7 *Let* (A1) *hold, let* $\bar{x} \in \mathcal{B}$ *and let* $D(x)$, $E(x)$ *be defined by* (34) *and* (35), *respectively. Then there exists* $\rho > 0$ *such that*

$$d_i(x) = d_i^{CL}(x), \quad e_i(x) = |\nabla f(x)|$$

for all $i \in \bar{A} \cup \bar{I}$ *and all* $x \in \mathcal{B}$ *with* $\|x - \bar{x}\|_2 < \rho$.

Proof. Consider an index $i \in \bar{A}$. Since $0 = \min\{\bar{x}_i - a_i, b_i - \bar{x}_i\} < |\nabla f(\bar{x})_i|^p$, the continuity of the expression on both sides of the previous inequality ensures the existence of $\rho > 0$ with $\min\{x_i - a_i, b_i - x_i\} < |\nabla f(x)_i|^p$ for all $i \in \bar{A}$ and all $x \in \mathcal{B}$ with $\|x - \bar{x}\|_2 < \rho$. This proves the assertion for $i \in \bar{A}$.

Now, consider an index $i \in \bar{I}$. We have that $0 = |\nabla f(\bar{x})_i| < \min\{\bar{x}_i - a_i, b_i - \bar{x}_i\}^p$. As before, there exists $\rho > 0$ such that $|\nabla f(x)_i| < \min\{x_i - a_i, b_i - x_i\}^p$ for all $i \in \bar{I}$ and all $x \in \mathcal{B}$ with $\|x - \bar{x}\|_2 < \rho$. $\square$

We now state the main result:

Theorem 8 *Let* (A1), (A2) *hold and let the strong second order sufficient conditions of Theorem* 2 *be satisfied at* $\bar{x}$. *If* M *is defined by* (6) *and if* D *and* E *are computed according to choice* I *or choice* II, *then there are* $C > 0$ *and* $\rho > 0$ *such that for all* $x^c \in \mathcal{B}^\circ$, $\|x^c - \bar{x}\|_2 < \rho$, *the equation* (21) *has a unique solution* x^+ *and*

$$\|x^+ - \bar{x}\|_2 \le C\,\|x^c - \bar{x}\|_2^{\min\{p,2\}}, \tag{36}$$

where $p > 1$ *is the scalar in the definition of* D *or* E, *respectively.*

Proof. By Lemma 7 there exists $\bar{\rho} > 0$ such that if $d_i(x)$ is defined by (2) or (34) and if $e_i(x)$ is defined by (10) or (35), then

$$d_i(x) = d_i^{\mathrm{CL}}(x), \quad e_i(x) = |\nabla f(x)_i| \quad \text{for all } i \in \bar{A} \cup \bar{I} \tag{37}$$

and for all $x \in \mathcal{B}$ with $\|x - \bar{x}\|_2 < \bar{\rho}$. In particular, both choices I and II of the scaling matrix D and of E satisfy the assumptions in Theorem 5 and the estimate (30) is valid for sufficiently small $\rho > 0$.

We complete the proof by estimating the quantities $\tilde{r}_i(x)$ defined in (31).

By the definition of $\bar{A}$ and $\bar{I}$ there exists $\gamma > 0$ with

$$\min_{i \in \bar{A} \cup \bar{I}} \left(|\nabla f(\bar{x})_i| + \min\{\bar{x}_i - a_i, b_i - \bar{x}_i\}\right) = 2\gamma.$$

Since $d_i(x) \geq \min\{1, x_i - a_i, b_i - x_i\}$ we have, after a possible reduction of $\rho > 0$,

$$\min_{i \in \bar{A} \cup \bar{I}} (d_i(x) + |\nabla f(x)_i|) \geq \gamma > 0$$

for all $x \in \mathcal{B}$ with $\|x - \bar{x}\|_2 < \rho$.

Since (A1) implies the local Lipschitz continuity of ∇f with a constant L_1, (31), (37), and the previous inequality imply that

$$\tilde{r}_i(x) = \frac{|\nabla f(x)_i - \nabla f(\bar{x})_i|\,|x_i - \bar{x}_i|}{d_i(x) + |\nabla f(x)_i|} \leq \frac{L_1 \|x - \bar{x}\|_2\,|x_i - \bar{x}_i|}{\gamma} \quad \forall i \in \bar{A} \cup \bar{I}. \tag{38}$$

For the remaining indices $i \in \bar{N}$ we have $\nabla f(\bar{x})_i = 0$ and $\bar{x}_i \in \{a_i, b_i\}$. We reduce $\rho > 0$ such that $x \in \mathcal{B}$, $\|x - \bar{x}\|_2 < \rho$, implies

$$|x_i - \bar{x}_i| = \min\{x_i - a_i, b_i - x_i\} = \min\{1, x_i - a_i, b_i - x_i\} \leq d_i(x) \quad \forall i \in \bar{N}. \tag{39}$$

We consider three cases. In each of the following cases, we consider an arbitrary $i \in \bar{N}$.

1. Let $|\nabla f(x)_i| < \min\{x_i - a_i, b_i - x_i\}^p$. In this case $d_i(x) = d_i^{\mathrm{CL}}(x)$ and for all choices of $E(x)$ in I and II we have that $e_i(x) = |\nabla f(x)_i|$. Moreover,

$$|\nabla f(x)_i - \nabla f(\bar{x})_i| = |\nabla f(x)_i| < \min\{x_i - a_i, b_i - x_i\}^p = |x_i - \bar{x}_i|^p.$$

Hence, using (39),

$$|\tilde{r}_i(x)| = \frac{|\nabla f(x)_i - \nabla f(\bar{x})_i||x_i - \bar{x}_i|}{d_i(x) + |\nabla f(x)_i|} \leq \frac{|x_i - \bar{x}_i|^{p+1}}{d_i(x)} \leq \frac{|x_i - \bar{x}_i|^{p+1}}{|x_i - \bar{x}_i|} = |x_i - \bar{x}_i|^p. \tag{40}$$

2. Let $\min\{x_i - a_i, b_i - x_i\} < |\nabla f(x)_i|^p$. As in the previous case, $d_i(x) = d_i^{\mathrm{CL}}(x)$ and for all choices of $E(x)$ in I and II we have that $e_i(x) = |\nabla f(x)_i|$. Moreover, $\nabla f(\bar{x})_i = 0$,

$$|x_i - \bar{x}_i| = \min\{x_i - a_i, b_i - x_i\} < |\nabla f(x)_i|^p = |\nabla f(x)_i - \nabla f(\bar{x})_i|^p,$$

and

$$\begin{aligned} |\tilde{r}_i(x)| &= \frac{|\nabla f(x)_i - \nabla f(\bar{x})_i||x_i - \bar{x}_i|}{d_i(x) + |\nabla f(x)_i|} \leq \frac{|\nabla f(x)_i - \nabla f(\bar{x})_i|^{p+1}}{d_i(x) + |\nabla f(x)_i - \nabla f(\bar{x})_i|} \\ &\leq |\nabla f(x)_i - \nabla f(\bar{x})_i|^p \leq L_1^p \|x - \bar{x}\|_2^p. \end{aligned} \tag{41}$$

3. Finally, assume that neither case 1 nor case 2 apply. If D and E are chosen according to I, then $d_i(x) = 1$ and $e_i(x) \geq 0$. In this case we obtain

$$\begin{aligned} |\tilde{r}_i(x)| &= \frac{|\nabla f(x)_i - \nabla f(\bar{x})_i||x_i - \bar{x}_i|}{d_i(x) + e_i(x)} \\ &\leq |\nabla f(x)_i - \nabla f(\bar{x})_i|\,|x_i - \bar{x}_i| \leq L_1 \|x - \bar{x}\|_2\,|x_i - \bar{x}_i|. \end{aligned} \tag{42}$$

If D and E are chosen according to II, then $e_i(x) = 0$ and $\tilde{r}_i(x) = 0$.

From the estimates (38) to (42) we obtain the existence of $C_1 > 0$ with

$$|\tilde{r}_i(x)| \le C_1 \, \|x - \bar{x}\|_2^{\min\{2,p\}} \tag{43}$$

for all $i \in \{1, \dots, n\}$. Using this estimate in (30) gives the desired result. $\square$

The solution x^+ of (5) will in general not satisfy $x^+ \in \mathcal{B}^\circ$. Therefore, we need to modify x^+ to ensure that the new iterate is in the interior. Of course, the modification should be so that the convergence result in Theorem 8 remain valid. In most cases the new iterate is chosen as $x^c + \sigma t s$, $s = x^+ - x^c$, where the step size

$$t = \min\left\{1, \min\left\{(b_i - x_i^c)/s_i \mid s_i > 0\right\}, \min\left\{(a_i - x_i^c)/s_i \mid s_i < 0\right\}\right\} \in [0,1], \tag{44}$$

is the largest step size giving $x^c + t^c s \in [a,b]$ and $\sigma \in (0,1)$ is scalar close to one. This is used, e.g, in [3,4]. Under the conditions in [3, Lemma 12], which include strict complementarity, it is shown that this step size satisfies $\sigma t = 1 - O(\|x^c - \bar{x}\|)$ if σ is chosen of the order $1 - O(\|s\|)$. Together with an estimate like (36) for $p \ge 2$, this implies q–quadratic convergence of the iteration.

In the degenerate or near degenerate case this scaling of $s = x^+ - x^c$ by σt is problematic. It can not be guaranteed that t is close to one if x^c is close to $\bar{x}$. In fact, in the next section we will give an example in which the step size even converges towards zero. From (44) we see that $t \ll 1$ can be caused by only one component. In such a case all other components of $x^+ - x^c$ will be scaled by this small t, even though many of the other components of x^+ may be safely inside $\mathcal{B}^\circ$. To see how badly the scaling by t can fail and to find a remedy, we make two observations. First, we note that if (36) holds and if x_i^+ is infeasible, then

$$\operatorname{dist}(x_i^+, [a_i, b_i]) \le |\bar{x}_i - x_i^+| = O(\|\bar{x} - x^c\|^{\min\{2,p\}}) = O(\|x^+ - x^c\|^{\min\{2,p\}}). \tag{45}$$

The first equality in (45) follows from (36) and the second equality from (36) and

$$\begin{aligned} &\|\bar{x} - x^c\|(1 - C\|\bar{x} - x^c\|^{(\min\{2,p\}-1)}) \le \|\bar{x} - x^c\| - \|x^+ - \bar{x}\| \\ &\le \|x^+ - x^c\| \le \|x^+ - \bar{x}\| + \|\bar{x} - x^c\| \le \|\bar{x} - x^c\|(1 + C\|\bar{x} - x^c\|^{(\min\{2,p\}-1)}). \end{aligned}$$

Equation (45) shows that the infeasibility of the unscaled step is always small relative to the step length. For our second observation suppose that a small scaling $t \ll 1$ occurs. Let $t = (b_i - x_i^c)/s_i$, $s_i > 0$, or $t = (a_i - x_i^c)/s_i$, $s_i < 0$. Simple calculations show that $\operatorname{dist}(x_i^+, [a_i, b_i]) = (1-t)|s_i|$. Since $t \ll 1$, $\operatorname{dist}(x_i^+, [a_i, b_i]) \approx |s_i|$. With (45) we see that components of $s = x^+ - x^c$ that lead to a small scaling t are always small, more precisely,

$$|s_i| = O(\|s\|^{\min\{2,p\}}) \ll \|s\|. \tag{46}$$

Therefore, if the scaling is small, this is caused by the 'unimportant' small components of s. The example in Section 6 will show that these components can in fact lead to very small step sizes t and, consequently, to a deterioration of convergence. Thus, a modification of the scaling matrix D and the corresponding matrix E alone is not enough to guarantee fast convergence.

Our first observation points to a remedy. Equation (45) shows that only a modification of the order $O(\|s\|^p)$ is necessary to obtain a feasible point from x^+. This can be easily accomplished. Instead of the scaling by t, we use a projection P onto $\mathcal{B}$. In the case of box constraints the projection can be easily computed and is given by $P(x) \stackrel{\text{def}}{=} \max\{a, \min\{b, x\}\}$. Instead of $x^c + \sigma t(x^+ - x^c)$ we use $x^c + \sigma(P(x^+) - x^c)$ as the next iterate, where σ is chosen close to one, e.g. $0.9995 \le \sigma < 1$. This choice was first proposed in [23]. The advantage of the projection over the simple scaling is that the projection only cuts off components x_i^+ which are larger than b_i or smaller than a_i. Other components are unchanged. Moreover, since $P(x^+)$ cuts off components x_i^+ outside $[a_i, b_i]$, the point $P(x^+)$ will actually be closer to $\bar{x} \in \mathcal{B}$ than x^+ if $P(x^+) \ne x^+$. We define

$$P[x^c](x) \stackrel{\text{def}}{=} x^c + \max\{\sigma, 1 - \|P(x) - x^c\|_2\}\ (P(x) - x^c)$$

and obtain the following method:

Algorithm 1 (Projected Affine-Scaling Interior-Point Newton Method)
Let $x^0 \in \mathcal{B}^\circ$ be given. Select the way the scaling matrix D and E is computed from choice I *and* II.

For $k = 0, 1, 2, \dots$:

1. *Compute $D(x^k)$ and $E(x^k)$ according to choice* I *or choice* II.
2. *If $D(x^k)\nabla f(x^k) = 0$ STOP with solution x^k.*
3. *Solve $H(x^k)(x^{k+1/2} - x^k) = -W(x^k)D(x^k)\nabla f(x^k)$.*
4. *Set $x^{k+1} = P[x^k](x^{k+1/2}) = x^k + \max\left\{\sigma, 1 - \|P(x^{k+1/2}) - x^k\|_2\right\}\ (P(x^{k+1/2}) - x^k)$.*

This algorithm is locally convergent with q–order $\min\{p, 2\}$. More precisely:

Theorem 9 *Let the assumptions of Theorem* 8 *be valid and let p be the scalar in the definition* (34), (35) *of the modified scaling matrices D or E. The iterates generated by Algorithm* 1 *converge locally with q–order* $\min\{p, 2\}$ *towards a point $\bar{x}$ satisfying the second order sufficient optimality conditions in Theorem* 2.

Proof. Let x^0 be sufficiently close to $\bar{x}$ so that (36) is valid, i.e.,

$$\|x^{1/2} - \bar{x}\|_2 \le C\|x^0 - \bar{x}\|_2^{\min\{p,2\}}.$$

The nonexpansion property of the projection and $P(\bar{x}) = \bar{x}$, $P(x^0) = x^0$ imply

$$\|P(x^{1/2}) - \bar{x}\|_2 \le \|x^{1/2} - \bar{x}\|_2, \quad \|P(x^{1/2}) - x^0\|_2 \le \|x^{1/2} - x^0\|_2.$$

Step 4 of the Algorithm 1 implies that

$$x^1 - \bar{x} = \left(1 - \max\left\{\sigma, 1 - \|P(x^{1/2}) - x^0\|_2\right\}\right)\left(x^0 - P(x^{1/2})\right) + P(x^{1/2}) - \bar{x}.$$

If $\max\left\{\sigma, 1 - \|P(x^{1/2}) - x^0\|_2\right\} = 1 - \|P(x^{1/2}) - x^0\|_2$, the two inequalities from above yield that

$$\begin{aligned}\|x^1 - \bar{x}\|_2 &\le \|P(x^{1/2}) - x^0\|_2^2 + \|P(x^{1/2}) - \bar{x}\|_2 \le \|x^{1/2} - x^0\|_2^2 + \|x^{1/2} - \bar{x}\|_2 \\ &\le 2\|x^0 - \bar{x}\|_2^2 + 2\|x^{1/2} - \bar{x}\|_2^2 + \|x^{1/2} - \bar{x}\|_2 \\ &\le 2\left(C^2\|x^0 - \bar{x}\|_2^{2\min\{p,2\}-2} + 1\right)\|x^0 - \bar{x}\|_2^2 + C\|x^0 - \bar{x}\|_2^{\min\{p,2\}}.\end{aligned}$$

The proof can be completed using standard induction arguments. □

5 Inexact solution of linear systems and globalization

For large-scale problems, the exact solution of the linear system in step 3 of Algorithm 1 may be impossible or at least very expensive. In this case iterative methods are used to compute approximate solutions $x^{k+1/2}$ satisfying

$$H(x^k)(x^{k+1/2} - x^k) = -W(x^k)D(x^k)\nabla f(x^k) + v^k. \tag{47}$$

The control of the residual error v_k can be done similarly to the control of inexactness in Newton's method for unconstrained optimization. The following theorem states a typical convergence result.

Theorem 10 *Let the assumptions of Theorem 9 hold. Assume that in step 3 of Algorithm 1 the iterates $x^{k+1/2}$ are computed only inexactly such that they satisfy the accuracy requirement* (47) *with*

$$\|v^k\|_2 \le \eta_k \|W(x^k)D(x^k)\nabla f(x^k)\|_2, \tag{48}$$

where $\eta_k \ge 0$. If $\lim_{k\to\infty}\eta_k = 0$ and $\eta_k \le \eta$, $\eta > 0$ sufficiently small, or even $\eta_k \le K\|W(x^k)D(x^k)\nabla f(x^k)\|_2^{(\min\{2,p\}-1)}$ for some $K > 0$, then the iterates x^k converge locally q–superlinearly, or locally with q–order $\min\{p, 2\}$, *respectively, towards a point $\bar{x}$ satisfying the second order sufficient optimality conditions in Theorem* 2.

Proof. In the following, let $C > 0$ denote a generic constant. For $x^0 \in \mathcal{B}^\circ$ sufficiently close to $\bar{x}$, Theorems 5 and 8 are applicable. In particular, (32) and (43) hold, proving that $\|W(x^0)r(x^0)\|_2 \le C\|x^0 - \bar{x}\|_2^{\min\{2,p\}}$. From (A1) it is obvious that $H(x)$, $x \in \mathcal{B}^\circ$, is bounded in a neighborhood of $\bar{x}$. Hence, the definition (12) of r yields

$$\|W(x^0)D(x^0)\nabla f(x^0)\|_2 = \|H(x^0)(x^0 - \bar{x}) - W(x^0)r(x^0)\|_2 \le C\|x^0 - \bar{x}\|_2.$$

This, assumption (48), and Lemma 4 imply

$$\|x^{1/2} - \bar{x}\|_2 = \|H(x^0)^{-1}(W(x^0)r(x^0) + v^0)\|_2 \le C(\|x^0 - \bar{x}\|_2^{(\min\{2,p\}-1)} + \eta_k)\|x^0 - \bar{x}\|_2.$$

See also (21). Now we can proceed as in the proof of Theorem 9. $\square$

Algorithm 1 can be globalized using a trust-region method analogously to the one in [3,4,23,24]. We only sketch the results and leave the details to a forthcoming paper. The main observation for the transition from global convergence to fast local convergence in the trust-region method is that Algorithm 1 produces a descent directions if the iterates are close enough to $\bar{x}$. Given $x^k \in \mathcal{B}^\circ$, $x^{k+1/2}$ solves the equation in step 3 of Algorithm 1 if and only if $x^{k+1/2} - x^k$ is a stationary point of the quadratic function

$$\psi_k(s) \stackrel{\text{def}}{=} \nabla f(x^k)^T s + \frac{1}{2} s^T Q(x^k) s$$

with $Q(x^k) \stackrel{\text{def}}{=} D(x^k)^{-1} M(x^k) = \nabla^2 f(x^k) + D(x^k)^{-1} E(x^k)$. One can show that $Q(x^k)$ is positive definite for x^k close enough to $\bar{x}$ satisfying the strong second order sufficiency condition of Theorem 2. Hence, ψ_k may be used for a trust region globalization where the trial steps s^k are approximate solutions of the subproblems

$$\text{minimize} \quad \psi_k(s) \quad \text{subject to} \quad x^k + s \in \mathcal{B}, \ \|s\|_2 \le \Delta_k.$$

Denoting by ψ_k^o the optimal value of ψ_k for this subproblem one can show that $s^k = x^{k+1} - x^k$ with x^{k+1} from step 4 in Algorithm 1 satisfies for any fixed $\beta \in (0,1)$ the *fraction of optimal decrease condition* $\psi_k(s^k) \le \beta\psi_k^o$ if the sufficiency condition of Theorem 2 holds for $\bar{x}$ and $\|x^k - \bar{x}\|_2 \le \rho$, $\rho > 0$ small enough. If the acceptance of steps and the update of the trust region radius Δ_k are performed in a standard way by comparing actual reduction $ared = f(x^k + s^k) - f(x^k)$ and predicted reduction $pred = \nabla f(x^k)^T s^k + \frac{1}{2} {s^k}^T \nabla^2 f(x^k) s^k$ it can be shown that for $\Delta_k \ge \Delta_{\min} > 0$ the step s^k is always accepted for sufficiently small ρ. Hence, close to $\bar{x}$ Algorithm 1 generates trial steps that are accepted by the trust region method outlined above, yielding local convergence with q-order $\min\{2, p\}$.

6 Two simple examples

The first example is $\min_{x \ge 0} f(x)$. We assume that $\bar{x} = 0$ is a minimizer and that $f'(0) = 0$, $f''(0) > 0$. Since $f'(x), f''(x) > 0$ for sufficiently small $x > 0$, it is not difficult to show that the Coleman and Li affine-scaling interior-point method is equivalent to Newton's method applied to the root finding problem $g(x) \stackrel{\text{def}}{=} x f'(x)$. In particular, the Coleman and Li steps are always shorter than the Newton steps for $\min f(x)$ and they are always strictly feasible. Since $g'(0) = 0$ and $g''(0) = 2f''(0) > 0$, the Coleman and Li affine-scaling interior-point method is linearly convergent. Depending on the choice

of the scaling matrix, our modification of the affine-scaling interior-point method will either take steps $s = (f''(x) + f'(x))^{-1} f'(x)$ or $s = f''(x)^{-1} f'(x)$. If $x + s < 0$, then the projection in Step 4 of the algorithm will move the new iterate even closer to the solution. Our algorithm converges q–quadratically, if the parameter in the choice of the scaling matrices is $p = 2$.

The previous example shows that for one-dimensional problems only the choice of the scaling matrices D and E in the Coleman and Li affine-scaling interior-point method cause a deterioration of the local convergence in the degenerate case. The next example shows that in the degenerate case the original Coleman and Li affine-scaling method can produce step sizes that do not converge towards one. Our previous discussion shows that this can not happen in one dimension. We consider

$$\min_{0 \le x, y \le 1} f(x, y) \stackrel{\text{def}}{=} -\frac{1}{2}x^2 + \frac{1}{2}y^2 - x^2 y + x. \tag{49}$$

It is easy to check that $(\bar{x}, \bar{y}) = (0, 0)$ is a minimizer of f at which the second order sufficient optimality conditions of Theorem 2 are satisfied. The strict complementarity is violated in the component y. For small $x, y > 0$, the step in the original Coleman and Li affine-scaling method with D and E given by (2), (10), respectively, is the solution of

$$\begin{pmatrix} d_1(-1 - 2y) + (1 - x - 2xy) & -2xd_1 \\ -2xd_2 & d_2 + |y - x^2| \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = - \begin{pmatrix} d_1(1 - x - 2xy) \\ d_2(y - x^2) \end{pmatrix}, \tag{50}$$

where $d_1 = x$ and $d_2 = y$ if $y - x^2 \ge 0$ and $d_2 = 1 - y$ if $y - x^2 < 0$.

We fix $\mu \in (0, 1)$ and look at the special case $y = \mu x^2$, $x > 0$ small. Then $d_2 = 1 - \mu x^2$ and the solution of (50) is given by $s_1 = -x + O(x^2)$, $s_2 = -(1 + \mu)x^2 + O(x^3)$. Hence, if a simple scaling is used to maintain strict feasibility, then $s_2 \approx -((1 + \mu)/\mu)y$ leads to $t \approx \mu/(1 + \mu) < \mu$. This example illustrates our statement (46) that the components s_i which lead to small scalings t are small relative to the step length.

If the modified scaling is used, then we obtain the following. Let again $\mu \in (0, 1)$ be fixed, $y = \mu x^2$, $x > 0$ sufficiently small. If E is computed from (35), then the entry $|y - x^2| = (1 - \mu)x^2$ will be replaced by zero and $d_2 > 0$, which is equal to $d_2 = 1$ for Choice I and $d_2 = 1 - \mu x^2$ for Choice II, can be cancelled. The solution of the resulting system (50) is given by $s_1 = -x + O(x^2)$, $s_2 = -(1 + \mu)x^2 + O(x^3)$. Hence, as before, a simple scaling would lead to a small $t \approx \mu/(1 + \mu)$ and it would prevent fast convergence. The same is true for all other variants of Choice I and II, respectively, since they all generate a step with $s_1 = -x + O(x^2)$, $s_2 = -(1 + \mu)x^2 + O(x^3)$.

Using a projection instead remedies the situation. In our numerical results below we find that using the projection instead of the simple scaling leads to convergence for the original Coleman-Li scaling as well as for its modifications. However, to obtain fast local convergence, one has to use the modifications of the scaling D and/or of E combined with the projection.

We complete our illustration with a few Matlab computations for the example (49). In all cases the step is computed using equation (5) multiplied by $W(x^k)$. The iteration was stopped if $\|D_k \nabla f(x^k)\|_2 < 10^{-11}$ and $\|x^k - x^{k-1}\|_2 < 10^{-11}$ or if 20 iterations were reached. We used $\sigma = 0.9995$. Our starting point was $x_1^0 = 10^{-3}$, $x_2^0 = (x_1^0)^2/2$. The left part of Table 1 shows the performance of the original Coleman and Li affine-scaling method discussed in the introduction with D and E given by (2), (10), respectively, and simple scaling of the steps. It can be clearly seen that convergence is prohibited by small simple scalings t. In particular, we see that $t_0 \approx 1/3 = \frac{1}{2}/(1+\frac{1}{2})$, as expected.

Table 1
Coleman and Li affine-scaling interior-point method and its modifications ($\hat{g}_k = D_k \nabla f(x^k)$). a. Coleman and Li affine-scaling interior-point method with simple scaling of the step. b. Coleman and Li affine-scaling interior-point method with projection of the step. c. Modified Coleman and Li affine-scaling interior-point method, Algorithm 1 ($p = 2$, Choice I/II).

	a.				b.			c.		
k	t_k	$\|\hat{g}_k\|_2$	x_1^k	x_2^k	$\|\hat{g}_k\|_2$	x_1^k	x_2^k	$\|\hat{g}_k\|_2$	x_1^k	x_2^k
0	3.3e-01	9.9e-04	1.0e-03	5.0e-07	9.9e-04	1.0e-03	5.0e-07	9.9e-04	1.0e-03	5.0e-07
1	5.6e-04	6.6e-04	6.6e-04	2.5e-10	5.0e-07	5.0e-07	2.5e-10	5.0e-07	5.0e-07	2.5e-10
2	2.8e-07	6.6e-04	6.6e-04	1.2e-13	2.5e-13	2.5e-13	1.2e-10	2.5e-13	2.5e-13	1.2e-16
3	1.4e-10	6.6e-04	6.6e-04	6.2e-17	3.8e-21	1.5e-23	6.2e-11	6.2e-26	6.2e-26	3.1e-29
4	7.0e-14	6.6e-04	6.6e-04	3.1e-20	9.7e-22	4.8e-34	3.1e-11			
5	3.5e-17	6.6e-04	6.6e-04	1.5e-23	2.4e-22	7.5e-45	1.5e-11			
6	1.7e-20	6.6e-04	6.6e-04	7.8e-27	6.0e-23	5.9e-56	7.8e-12			
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$						
19	2.1e-63	6.6e-04	6.6e-04	9.5e-70						

The situation is improved considerably if the simple scaling is replace by a projection identical to the one in Step 4 of Algorithm 1. However, a closer look at the middle part of Table 1 shows that the iterates do not converge q–quadratically towards the solution, but show a linear convergence behavior with factor $\approx \frac{1}{2}$.

The right part of Table 1 shows the performance of the modified Coleman and Li affine-scaling method, Algorithm 1. We used $p = 2$ and $\sigma = 0.9995$. With Choice I, D is computed by (34) and E is computed by (35). We observed that $d_2(x^k) = 1$ and $e_2(x^k) = 0$ for $k \geq 1$. The iterates converge almost q–quadratically. The numerical results with Choice II for D and E are indistinguishable from those for Choice I if the result are printed to 5 digits. Of course, these illustrations do not replace numerical tests of the modifications of the affine-scalings and their comparisons with existing ones. Such tests will be part of a forthcoming report. We believe, however, that they make a strong case for the projection instead of the simple scaling. This is also supported by

the examples in [23].

7 Conclusions

Affine-scaling interior-point Newton methods for simply constrained nonlinear problems were introduced by Coleman and Li [3,4] and have proven to be very robust and efficient. However, in the original design of Coleman and Li, they are q–quadratically convergent only if strict complementarity holds at the solution. In this paper we have introduced and analyzed two modifications that overcome the two main problems arising from degeneracy. The first modification addresses the fact that the affine scaling of degenerate components may destroy fast local convergence. This effect is remedied by switching off the affine scaling for components which are identified to be degenerate. This is done completely automatically and involves a parameter $p > 1$ which can be chosen arbitrarily. The solution of the affine-scaling Newton equation might result in an iterate that is outside of $\mathcal{B}^\circ$ and, therefore, has to be transported back into the interior $\mathcal{B}^\circ$. This is usually accomplished by a simple scaling with an appropriate step size. In the presence of degeneracy, however, this may lead to almost vanishing step sizes even arbitrarily close to the solution and a loss of fast convergence. Therefore, the second modification is an implementation of the back-transport into $\mathcal{B}^\circ$ by an interior-point modification of the projection onto $\mathcal{B}$. We have proved that the modified affine-scaling interior-point Newton method converges locally with q–order $\min\{p, 2\}$. This rate of convergence is preserved if the Newton equations are solved only inexactly with sufficient accuracy. Moreover, it was pointed out that the algorithm generates locally descent directions and can be globalized by trust region techniques.

The failures of the original Coleman-Li algorithm in the degenerate case and effects of the two modifications introduced in this paper have been illustrated on an example.

Finally, we would like to point out that the convergence results were proven using some non-standard techniques, which partly were developed in the analysis of affine-scaling interior-point Newton methods for infinite dimensional problems [23]. As an example we mention the preconditioning by the diagonal matrix W. Also the importance of the projection as back-transport was already discovered in [23], since in the infinite dimensional context the simple scaling fails even in the nondegenerate case.

Acknowledgement

Most of the research in this paper was done during a visit of M. Heinkenschloss at the TU München and a subsequent visit of M. and S. Ulbrich at Rice University. We would like to thank our respective hosts for their hospitality and support. M. Heinkenschloss is grateful for interesting discussions with Amr El-Bakry on interior-point methods for degenerate problems Finally, we would like to thank the referees and the associate editor for their careful reading of this paper and their constructive comments.

References

[1] D. P. Bertsekas. Projected Newton methods for optimization problems with simple constraints. *SIAM Journal on Control and Optimization*, 20:221–246, 1982.

[2] J. F. Bonnans and C. Pola. A trust region interior point algorithm for linearly constrained optimization. *SIAM Journal on Optimization*, 7:717–731, 1997.

[3] T. F. Coleman and Y. Li. On the convergence of interior–reflective Newton methods for nonlinear minimization subject to bounds. *Mathematical Programming*, 67:189–224, 1994.

[4] T. F. Coleman and Y. Li. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6:418–445, 1996.

[5] T. F. Coleman and Y. Li. A trust region and affine scaling interior point method for nonconvex minimization with linear inequality constraints. CORNELLCS TR97–1642, Department of Computer Science, Cornell University, Ithaca, NY 14853–3801, 1997.

[6] A. R. Conn, N. I. M. Gould, and P. L. Toint. Global convergence of a class of trust region algorithms for optimization with simple bounds. *SIAM Journal on Numerical Analysis*, 25:433–460, 1988. See [7].

[7] A. R. Conn, N. I. M. Gould, and P. L. Toint. Correction to the paper on global convergence of a class of trust region algorithms for optimization with simple bounds. *SIAM Journal on Numerical Analysis*, 26:764–767, 1989.

[8] J. E. Dennis, M. Heinkenschloss, and L. N. Vicente. Trust–region interior–point algorithms for a class of nonlinear programming problems. *SIAM J. Control and Optimization*, 36:1750–1794, 1998.

[9] J. E. Dennis and L. N. Vicente. Trust-region interior-point algorithms for minimization methods with simple bounds. In H. Fischer, B. Riedmüller, and S. Schäffler, editors, *Applied Mathematics and Parallel Computing, Festschrift for Klaus Ritter*, pages 97–107, Heidelberg, 1996. Physica-Verlag.

[10] A. S. El-Bakry, R. A. Tapia, T. Tsuchiya, and Y. Zhang. On the formulation and theory of the primal–dual Newton interior–point method for nonlinear programming. *Journal of Optimization Theory and Applications*, 89:507–541, 1996.

[11] A. S. El-Bakry, R. A. Tapia, and Y. Zhang. On the convergence rate of Newton interior–point methods in the absence of strict complementarity. *Computational Optimization and Applications*, 6:157–167, 1996.

[12] F. Facchinei and C. Kanzow. A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems. *Mathematical Programming*, 76:493–512, 1997.

[13] A. V. Fiacco. *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming.* Academic Press, New York, 1983.

[14] A. V. Fiacco and G. P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques.* John Wiley & Sons, Chichester, New York, 1968. Republished by SIAM, Philadelphia, 1990.

[15] A. Fischer. A special Newton-type optimization method. *Optimization*, 24:269–284, 1992.

[16] A. Fischer. Solution of monotone complementarity problems with locally Lipschitzian functions. *Mathematical Programming*, 76:513–532, 1997.

[17] C. C. Gonzaga and R. A. Tapia. On the convergence of the Mizuno-Todd-Ye algorithm to the analytic center of the solution set. *SIAM Journal on Optimization*, 7:47–65, 1997.

[18] M. Lescrenier. Convergence of trust region algorithms for optimization with bounds when strict complementarity does not hold. *SIAM Journal on Numerical Analysis*, 28:476–495, 1991.

[19] C.-J. Lin and J. J. Moré. Newton's method for large bound-constrained optimization problems. ANL/MCS-P724-0898, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois 60439, 1998.

[20] R. D. C. Monteiro and S. J. Wright. Local convergence of interior point algorithms for degenerate monotone LCP. *Computational Optimization and Applications*, 3:131–155, 1994.

[21] S. M. Robinson Strongly regular generalized equations. *Math. of Operations Research*, 5:43–62, 1980.

[22] J. Sun. A convergence proof for an affine-scaling algorithm for convex quadratic programming without nondegeneracy assumptions. *Mathematical Programming*, 60:69–79, 1993.

[23] M. Ulbrich and S. Ulbrich. Superlinear convergence of affine-scaling interior-point Newton methods for infinite-dimensional nonlinear problems with pointwise bounds. TR97–05, Department of Computational and Applied Mathematics, Rice University, Houston, Texas, 1997.

[24] M. Ulbrich, S. Ulbrich, and M. Heinkenschloss. Global convergence of affine-scaling interior-point Newton methods for infinite-dimensional nonlinear problems with pointwise bounds. To appear in *SIAM Journal on Control and Optimization.*

[25] L. N. Vicente. On interior–point Newton algorithms for discretized optimal control problems with state constraints. *Optimization Methods and Software*, 8:249–275, 1998.

[26] M. H. Wright. Interior point methods for constrained optimization. In A. Iserles, editor, *Acta Numerica 1992*, pages 341–407. Cambridge University Press, Cambridge, London, New York, 1992.

[27] S. J. Wright. *Primal–Dual Interior–Point Methods.* SIAM, Philadelphia, PA, 1996.

[28] H. Yamashita and H. Yabe. Superlinear and quadratic convergence of some primal–dual interior point methods for constrained optimization. *Mathematical Programming*, 75:377–397, 1996.