A Globally and Superlinearly Convergent Potential Reduction Interior Point Method for Convex Programming*

R. D. C. Monteiro[†] S. J. Wright[‡]

July 15, 1992

Abstract

We consider an interior point algorithm for convex programming in which the steps are generated by using a primal-dual affine scaling technique. A "local" variant of the algorithm is shown to have superlinear convergence with q-order up to (but not including) 2. The technique is embedded in a potential reduction algorithm and the resulting method is shown to be globally and superlinearly convergent. An important feature of the convergence analysis is its use of a novel strict interiority condition, which generalizes the usual conical neighborhood of the central path.

1 Introduction

In the past two years, we have seen the appearance of several papers dealing with the construction of primal-dual interior-point algorithms for linear programs (LP) and linear complementarity problems (LCP) which are superlinearly or quadratically convergent. For LP, these works include McShane [7], Mehrotra [8], Tsuchiya [10], Ye [11], Ye et al. [13] and Zhang et al. [14, 16]. For LCP, we cite Ji et al. [3, 4], Kojima et al. [5, 6], Ye and Anstreicher [12] and Zhang et al. [17]. For more details on the historic development of superlinearly or quadratically convergent primal-dual interior-point algorithms, we refer the reader to the introduction section of Ye and Anstreicher [12].

In this paper, we discuss superlinearly convergent primal-dual affine scaling methods for solving the convex programming problem

$$\min_x \ f(x), \qquad Ax = b, \qquad x \geq 0, \tag{1}$$

*This research was supported by the National Science Foundation (NSF) under Grant No. DDM-9109404 and the Applied Mathematical Sciences subprogram of the Office of Energy Research, U. S. Department of Energy, under Contract W-31-109-Eng-38.

[†]Department of Systems and Industrial Engineering, University of Arizona, Tucson, Arizona 85721.

[‡]Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439.

where $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ and $f : \mathbb{R}^n \to \mathbb{R}$ is convex and smooth in a sense to be defined below. We assume that the feasible set $\{x \mid Ax = b, \;\; x \geq 0\}$ is nonempty and that $m < n$. The Wolfe dual problem for (1) can be stated as

$$\max_{x,y} \; f(x) - x^T \nabla f(x) + b^T y, \qquad \nabla f(x) - A^T y \geq 0. \tag{2}$$

It is well known that if $\bar{x}$ solves (1), then there is $\bar{y} \in \mathbb{R}^m$ such that $(\bar{x}, \bar{y})$ solves (2) and, moreover, the optimal values of the primal and dual objective functions are identical. Introducing a slack vector s by

$$s = \nabla f(x) - A^T y,$$

we deduce that the following conditions must be satisfied by the primal and dual solutions:

$$\begin{aligned}
\bar{s} &= \nabla f(\bar{x}) - A^T \bar{y}, && \text{(3a)}\\
A\bar{x} &= b, && \text{(3b)}\\
\bar{x} &\geq 0, && \text{(3c)}\\
\bar{s} &\geq 0, && \text{(3d)}\\
\bar{x}^T \bar{s} &= 0. && \text{(3e)}
\end{aligned}$$

In the subsequent discussion, we say that a point (x, y, s) is "feasible" if it satisfies the equations (3a)–(3d), and "strictly feasible" if (3a)–(3d) are satisfied with $x > 0$ and $s > 0$.

Interior point algorithms for the linear and quadratic versions of (1), and the related complementarity problems, have usually required all iterates to belong to a neighborhood of the central path defined by either

$$\mathcal{N}_1(\beta) = \left\{ (x, y, s) \;\; \text{feasible} \;\mid\; \|XSe - (x^T s/n)e\|_2 \leq \beta \right\}$$

or

$$\mathcal{N}_2(\beta) = \left\{ (x, y, s) \;\; \text{feasible} \;\mid\; x_i s_i \geq \beta(x^T s/n), \;\; \forall i = 1, \ldots, n \right\},$$

where

$$X = \operatorname{diag}(x_1, x_2, \ldots, x_n), \qquad S = \operatorname{diag}(s_1, s_2, \ldots, s_n), \qquad e = (1, 1, \ldots, 1)^T,$$

and β is a constant in $(0, 1)$. For example, the predictor-corrector algorithms of Ye et al. [13] for linear programming and Ji, Potra and Huang [3] and Ye and Anstreicher [12] for linear complementarity problems use neighborhoods of the form $\mathcal{N}_1$, while the linear programming algorithm of Zhang and Tapia [15] uses $\mathcal{N}_2$. In this paper, we use a different neighborhood defined with respect to two parameters $\delta > 0$ and $\eta > 0$ in the following way:

$$\mathcal{N}(\delta, \eta) = \left\{ (x, y, s) \;\; \text{feasible} \;\mid x^T s \leq 1, \quad x_i s_i \geq \eta (x^T s)^{1+\delta}, \;\; \forall i = 1, \ldots, n \right\}. \tag{4}$$

For analytical purposes, this neighborhood has the advantage over $\mathcal{N}_1$ and $\mathcal{N}_2$ that local convergence of q-order arbitrarily close to 2 can be obtained *without* expanding $\mathcal{N}$ as the solution is approached. In the papers by Ye [11] and Ye and Anstreicher [12], it is necessary to expand the neighborhood $\mathcal{N}_1(\beta)$ during the final stages in order to obtain rapid local convergence.

The basic algorithm we describe in this paper calculates search directions by using a primal-dual affine scaling technique. This technique is equivalent to finding a Newton direction for the system of nonlinear equations formed by (3a), (3b) and the complementarity condition $XSe = 0$. Because of this connection to Newton's method, it is immediately clear that such an algorithm would be quadratically convergent if it is started sufficiently close to a nondegenerate solution and if it is allowed to take full steps. In our local analysis, we show that we can ensure that each iterate lies in the neighborhood $\mathcal{N}(\delta, \eta)$ without jeopardizing fast local convergence properties. Moreover, we also show that superlinear convergence can be obtained under certain assumptions which allow the possibility of problem (1) to have multiple solutions! Membership of $\mathcal{N}(\delta, \eta)$ is enforced by taking a step length α less than one along the affine scaling direction and by "bending" the search path for s in a way to be described in the next section. We show that it is easy to choose an α to ensure that both continued membership of $\mathcal{N}(\delta, \eta)$ and fast local convergence apply. By a suitable choice of parameters, the q-order of the local convergence can lie anywhere in the range $(1, 2)$.

To make the algorithm converge globally from any strictly feasible starting point, we embed the affine scaling technique into the potential reduction algorithm discussed in Monteiro [9]. The search direction is calculated as before, but now we choose the step size so that the Tanabe-Todd-Ye potential function

$$\phi_q(x, s) = q \log x^T s - \sum_{i=1}^{n} \log x_i s_i. \tag{5}$$

(where $q > n$) is reduced at every iteration according to the Armijo rule, a requirement that is often used in unconstrained optimization algorithms. The actual amount of reduction must be at least a small multiple of the "predicted" reduction obtained from a first order model of ϕ_q around the current point. Rules for making an initial guess of the step size and for reducing the step size when it fails the Armijo test are also given.

Finally, we also show that if q is chosen from the range $(n, n + 1/4)$, we can define an algorithm for which both our global and local convergence theories hold. That is, the method is both globally and superlinearly convergent.

The remainder of the paper is laid out as follows: In Section 2, we start by stating our main assumptions. These are, for the most part, standard. The exception is Assumption 5, and the bulk of the section is taken up with proving that this assumption holds in a variety of familiar circumstances, including the cases of nondegeneracy, weak sharp minima, and quadratic objective functions.

In Section 3, we state the linear system of equations that must be solved to obtain the search directions. We then derive bounds on the components of these search directions, in terms of the duality gap $x^T s$ and the parameters δ and η that define the neighborhood $\mathcal{N}$. The local convergence theory for our primal-dual affine scaling algorithm appears in Section 4. We show that if the algorithm is started from a point at which $x^T s$ is sufficiently small, and which is not too close to the boundary, superlinear convergence of the duality gap to zero can be obtained. Section 5 contains the global convergence theory. We outline the globally convergent algorithm presented in Monteiro [9] and show that if some iterate K satisfies the initial point conditions for the locally convergent algorithm, superlinear convergence also occurs. Finally, we prove that for appropriate choices of q in ϕ_q, namely for every $q \in (n, n+1/4)$, these conditions for superlinear convergence are guaranteed to be satisfied by some iterate. As a consequence, we obtain superlinearly convergent algorithms based on the potential function (5). Since the step sizes of the algorithm are selected by means of Armijo rule applied to the potential function (5), there is no reason to believe that the iterates remain in a neighborhood of the form $\mathcal{N}_1(\beta)$ or $\mathcal{N}_2(\beta)$. The neighborhood $\mathcal{N}(\delta,\eta)$ plays a crucial role: we are able to show that the iterates (generated via Armijo rule) remain within a certain $\mathcal{N}(\delta,\eta)$ and that one-step superlinear convergence can be obtained.

The following notational conventions are used in the remainder of the paper: Unless otherwise specified, $\|.\|$ denotes the the Euclidean norm. We denote the open ball with center $x \in \mathbb{R}^n$ and radius $\epsilon > 0$ by

$$B(x,\epsilon) \equiv \{u \in \mathbb{R}^n \mid \|u - x\| < \epsilon\}.$$

When x and y are two vectors in $\mathbb{R}^n$, the notation $[x, y]$ denotes the set of vectors on the line joining x and y.

For a general vector $z \in \mathbb{R}^n$ and index set $J \subseteq \{1,\ldots,n\}$, z_J denotes the vector made up of components z_i for $i \in J$. If $M \in \mathbb{R}^{n\times n}$ and $I, J \subseteq \{1,\ldots,n\}$ then M_{IJ} denotes the submatrix consisting of elements M_{ij} for which $i \in I$ and $j \in J$. The matrix $M_{I.}$ refers to the submatrix M_{IJ} for which $J = \{1,\ldots,n\}$. Similarly, $M_{.J}$ denotes the submatrix M_{IJ} in which $I = \{1,\ldots,m\}$ If $D \in \mathbb{R}^{n\times n}$ is diagonal, then D_B denotes the diagonal matrix whose diagonal entries are D_{ii} for $i \in B$.

We say that (J, L) is a partition of $\{1,\ldots,n\}$ if $J \cup L = \{1,\ldots,n\}$ and $J \cap L = \emptyset$.

Finally, we define some problem-dependent notation: For $\epsilon \in [0,\infty]$, let

$$\mathcal{F}_\epsilon \triangleq \{(x, y, s) \ \text{ feasible } \ \mid x^T s \le \epsilon\}.$$

Clearly, $\mathcal{F}_0$ is just the set of all primal-dual solutions of (1),(2) and $\mathcal{F}_\infty$ is the set of all primal-dual feasible points. We also define a projection operator Π by $\Pi(x, y, z) \triangleq x$ and, with a slight abuse of notation,

$$\mathcal{X}_\epsilon \triangleq \Pi(\mathcal{F}_\epsilon)$$

for $\epsilon \in [0,\infty]$. Note that $\mathcal{X}_0$ is the solution set for (1). The operator $P : \mathbb{R}^n \to \mathbb{R}^n$ defines projection into $\mathcal{X}_0$, that is

$$P(x) \triangleq \arg\min_{y\in\mathcal{X}_0} \|y - x\|.$$

2 Assumptions

In this section, we state all the assumptions on problem (1) that will be needed in our development. We then show that certain well-studied cases are special cases of these assumptions.

We start by stating all our major assumptions. In this work, we adopt the convention of explicitly stating the assumptions needed in the statement of each result. The main results of the paper will require all the assumptions below to hold.

Assumption 1 (Differentiability and Convexity of f.) *The function f is convex over the set $\{x \mid Ax = b,\ x \geq 0\}$ and twice continuously differentiable in a neighborhood of this set.*

Assumption 2 (Local Lipschitz continuity of the Hessian.) *For every $x \in \{x \mid Ax = b,\ x \geq 0\}$, there exists a neighborhood $\mathcal{J}(x)$ of x and a positive constant C_x such that for all x^1, $x^2 \in \mathcal{J}(x)$ we have*

$$\|\nabla^2 f(x^1) - \nabla^2 f(x^2)\| \leq C_x \|x^1 - x^2\|.$$

Assumption 3 *The set of strictly feasible solutions of (3a)–(3d) is non-empty and a strictly feasible solution (x^0, y^0, s^0) is given.*

Assumption 4 (Existence of a strictly complementary solution.) *There exist a partition (B, N) of $\{1, \ldots, n\}$ and a solution (x^*, y^*, s^*) of (3) such that $x^*_B > 0$ and $s^*_N > 0$.*

Assumption 5 *There exists a constant $L > 0$ such that*

$$\left\|\nabla f(P(x)) - \nabla f(x) - \nabla^2 f(x)(P(x) - x)\right\| \leq L(x^T s)^2, \qquad \forall (x, y, s) \in \mathcal{F}_1. \tag{6}$$

Since the left hand side of (6) is uniformly zero when f is quadratic, we can immediately note the following result.

Lemma 2.1 *If f is linear or quadratic, then Assumption 5 holds.*

In Lemmas 2.6 and 2.7 below, we show that Assumption 5 also holds in other well-studied situations.

Boundedness of the set $\mathcal{F}_1$, and therefore of the neighborhood $\mathcal{N}(\delta, \eta)$, is an immediate consequence of the following lemma:

Lemma 2.2 *Suppose that Assumptions 1 and 3 hold. Then there exists a constant $C_1 > 0$ such that if $(x, y, s) \in \mathcal{F}_1$ then*

$$\begin{aligned} x_i &\leq C_1 \qquad \text{for } i = 1, \ldots, n, \\ s_i &\leq C_1 \qquad \text{for } i = 1, \ldots, n. \end{aligned}$$

Proof. Consider the strictly feasible point (x^0, y^0, s^0) as in Assumption 3. By the convexity of f, we can easily show that $(x - x^0)^T (s - s^0) \geq 0$. Hence

$$x^T s^0 + {x^0}^T s \leq x^T s + {x^0}^T s^0 \leq 1 + {x^0}^T s^0.$$

This implies that

$$x_i \leq (1 + {x^0}^T s^0)/s^0_i, \tag{7}$$

$$s_i \leq (1 + {x^0}^T s^0)/x^0_i, \tag{8}$$

for every $i = 1, \ldots, n$. The result now follows by setting

$$C_1 = \frac{1 + x^{0^T} s^0}{r^0},$$

where r^0 is the smallest component of (x^0, s^0). ■

Lemma 2.3 *Suppose that Assumptions 1 and 3 hold and assume that there exist constants $\epsilon > 0$ and $L_\epsilon > 0$ such that*

$$\left\| \nabla f(P(x)) - \nabla f(x) - \nabla^2 f(x)(P(x) - x) \right\| \le L_\epsilon (x^T s)^2, \qquad \forall (x, y, s) \in \mathcal{F}_\epsilon. \tag{9}$$

Then Assumption 5 holds.

Proof. If $\epsilon \ge 1$ then Assumption 5 obviously holds with $L = L_\epsilon$. Assume then that $\epsilon < 1$ and consider the following set

$$\mathcal{U} \stackrel{\triangle}{=} \{(x, y, s) \text{ feasible} \mid x^T s \le 1 \, \text{and} \, x^T s \ge \epsilon\}.$$

Using Lemma 2.2, it is easy to see that $\mathcal{U}$ is compact. Hence,

$$a \equiv \sup_{(x,y,s) \in \mathcal{U}} \left\| \nabla f(P(x)) - \nabla f(x) - \nabla^2 f(x)(P(x) - x) \right\|$$

is finite. It is now easy to see that Assumption 5 is satisfied with $L \stackrel{\triangle}{=} \max(L_\epsilon, a/\epsilon)$. ■

Lemma 2.4 *Suppose that Assumptions 1 and 2 hold. Then for every compact subset $\mathcal{K} \subseteq \{x \mid Ax = b, x \ge 0\}$, there exists a constant $C = C(\mathcal{K}) > 0$ such that*

$$\|\nabla^2 f(x^1) - \nabla^2 f(x^2)\| \le C \|x^1 - x^2\|, \qquad \forall x^1, x^2 \in \mathcal{K}. \tag{10}$$

Proof. For $x \in \mathcal{K}$, consider the constant C_x and the neighborhood $\mathcal{J}(x)$ of x as in Assumption 2. Since $\mathcal{J}(x)$ is a neighborhood of x, there exists $\epsilon(x) > 0$ such that $B(x, 2\epsilon(x)) \subseteq \mathcal{J}(x)$. By the compactness of $\mathcal{K}$, there exist $\bar{x}^1, \ldots, \bar{x}^p \in \mathcal{K}$ such that

$$\mathcal{K} \subseteq B(\bar{x}^1, \epsilon(\bar{x}^1)) \cup \ldots \cup B(\bar{x}^p, \epsilon(\bar{x}^p)). \tag{11}$$

Let $\omega \equiv \sup\{\|\nabla^2 f(x)\| \mid x \in \mathcal{K}\} < \infty$, $\epsilon \equiv \min\{\epsilon(\bar{x}^1), \ldots, \epsilon(\bar{x}^p)\}$ and $C_{\max} \equiv \max\{C_{\bar{x}^1}, \ldots, C_{\bar{x}^p}\}$. We will show that $C \equiv \max\{C_{\max}, 2w/\epsilon\}$ satisfies the requirement of the lemma. Indeed, let $x^1, x^2 \in \mathcal{K}$ be given. There are two cases to consider. In the first case, assume that $\|x^1 - x^2\| < \epsilon$. By (11), there exists an $l \in \{1, \ldots, p\}$ such that $x^1 \in B(\bar{x}^l, \epsilon(\bar{x}^l))$. Hence, $x^2 \in B(\bar{x}^l, 2\epsilon(\bar{x}^l))$ since $\|x^1 - x^2\| < \epsilon \le \epsilon(\bar{x}^l)$. Hence, it follows that $x^1, x^2 \in \mathcal{J}(\bar{x}^l)$. By Assumption 2, we have

$$\|\nabla^2 f(x^1) - \nabla^2 f(x^2)\| \le C_{\bar{x}^l} \|x^1 - x^2\| \le C \|x^1 - x^2\|.$$

For the second case, assume that $\|x^1 - x^2\| \ge \epsilon$. Then we have

$$\|\nabla^2 f(x^1) - \nabla^2 f(x^2)\| \le 2w \le C\epsilon \le C \|x^1 - x^2\|.$$

Hence, the lemma follows. ■

Using Lemmas 2.3 and 2.4, we can now give a general condition which guarantees the validity of Assumption 5.

Lemma 2.5 *Suppose that Assumptions 1, 2 and 3 hold. Assume that there exists constants* $\epsilon > 0$ *and* $v > 0$ *such that the following implication holds:*

$$(x, y, s) \in \mathcal{F}_\epsilon \Longrightarrow \|x - P(x)\| \leq v(x^T s). \tag{12}$$

Then Assumption 5 holds.

Proof. We may assume without loss of generality that $\epsilon \leq 1$. Lemma 2.2 implies that $\mathcal{X}_1$ is bounded, and hence that the sets $\mathcal{X}_0$ and $\mathcal{X}_\epsilon$ are also bounded. Since these two sets are also closed, they are compact. Hence, the set defined as

$$\mathcal{K} \equiv \bigcup \{[x, \bar{x}] \mid x \in \mathcal{X}_\epsilon \text{ and } \bar{x} \in \mathcal{X}_0\}$$

is also compact. By lemma 2.4, there exists $C > 0$ such that (10) holds. Hence, for every $(x, y, s) \in \mathcal{F}_\epsilon$, we obtain

$$\begin{aligned}
& \left\| \nabla f(P(x)) - \nabla f(x) - \nabla^2 f(x)(P(x) - x) \right\| \\
\leq \ & \max_{\tilde{x} \in [x, P(x)]} \left\| \left[\nabla^2 f(\tilde{x}) - \nabla^2 f(x) \right] [P(x) - x] \right\| \\
\leq \ & C \|\tilde{x} - x\| \, \|P(x) - x\| \\
\leq \ & C \|P(x) - x\|^2 \\
\leq \ & C v^2 (x^T s)^2,
\end{aligned}$$

where the second inequality is due to the fact that both x and $\tilde{x}$ are in $\mathcal{K}$ and the last inequality is due to implication (12). Therefore, the assumptions of Lemma 2.3 are satisfied and, as a consequence, Assumption 5 holds. ∎

Two important situations in which the implication (12) holds are stated in the next two lemmas.

Lemma 2.6 *Suppose that Assumptions 1, 2 and 3 hold. Then Assumption 5 holds if problem (1) has a weak sharp minimum in the sense of Burke and Ferris [1], that is, if there exists a constant* $\lambda > 0$ *such that*

$$\|x - P(x)\| \leq \lambda \left[f(x) - f(P(x)) \right],$$

for all $x \in \{x \mid Ax = b, \ x \geq 0\}$.

Proof. For every feasible (x, y, s), we have

$$\|x - P(x)\| \leq \lambda \left[f(x) - f(P(x)) \right] \leq \lambda (x^T s)$$

which shows that the implication (12) holds with $\epsilon = \infty$ and $v = \lambda$. ∎

The other important case which guarantees the validity of implication (12), and hence of Assumption 5, requires that the solution of problem (1) be nondegenerate. We now make precise the notion of nondegeneracy of a solution of problem (1).

Definition 1 *A solution* $(\bar{x}, \bar{y}, \bar{s})$ *of problem (1) is called* nondegenerate *if the following conditions are satisfied:*

(a) $(\bar{x}, \bar{y}, \bar{s})$ *is a strictly complementary solution, that is, there exists a partition* (B, N) *of* $\{1, \dots, n\}$ *such that* $\bar{x}_B > 0$ *and* $\bar{s}_N > 0$*;*

(b) The matrix $\bar{A}$ *defined as*

$$\bar{A} \equiv \begin{bmatrix} A_B & A_N \\ 0 & I \end{bmatrix}$$

has linearly independent columns;

(c) $\nabla^2 f(\bar{x})$ *is nonsingular in the null space of* $\bar{A}$*.*

We observe that the rows of $\bar{A}$ are the gradients of the constraints of problem (1) which are active at the solution $\bar{x}$. Clearly, a nondegenerate solution must be unique. We are now ready to state the second case under which Assumption 5 holds.

Lemma 2.7 *Suppose that Assumptions 1, 2 and 3 hold and assume that problem (1) has the unique nondegenerate solution* $(\bar{x}, \bar{y}, \bar{s})$*. Then Assumption 5 holds.*

Proof. Consider the function $F : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ defined by

$$F(x, y, s) = \begin{bmatrix} XSe \\ Ax - b \\ -\nabla f(x) + A^T y + s \end{bmatrix}. \tag{13}$$

Using the assumption that $(\bar{x}, \bar{y}, \bar{s})$ is a nondegenerate solution, one can easily show that $\nabla F(\bar{x}, \bar{y}, \bar{s})$ is nonsingular. To simplify notation, let $w = (x, y, s)$ and $\bar{w} = (\bar{x}, \bar{y}, \bar{s})$. The differentiability of F at $\bar{w}$ implies that we can write

$$F(w) = F(\bar{w}) + \nabla F(\bar{w})(w - \bar{w}) + \|w - \bar{w}\| r(w - \bar{w}), \tag{14}$$

where $\lim_{\|h\| \to 0} r(h) = 0$. Define $M \stackrel{\triangle}{=} (\|\nabla F(\bar{w})\|)^{-1}$. Then there exists $\xi > 0$ such that $\|r(w - \bar{w})\| \leq M/2$ for every $w \in B(\bar{w}, \xi)$. Hence, for every $w \in B(\bar{w}, \xi)$ we obtain using (14) that

$$\begin{aligned} \|F(w) - F(\bar{w})\| &\geq \|\nabla F(\bar{w})(w - \bar{w})\| - \|r(w - \bar{w})\| \, \|w - \bar{w}\| \\ &\geq M\|w - \bar{w}\| - \|r(w - \bar{w})\| \, \|w - \bar{w}\| \\ &\geq \frac{M}{2}\|w - \bar{w}\|, \end{aligned} \tag{15}$$

where the last inequality is due to the fact that $\|r(w - \bar{w})\| \leq M/2$. Using the definition (13) of F and choosing $w = (x, y, s)$ to be feasible, we obtain

$$\|F(w) - F(\bar{w})\| = \|F(x, y, s) - F(\bar{x}, \bar{y}, \bar{s})\| = \|XSe\| \leq x^T s. \tag{16}$$

Using (15) and (16), we obtain

$$\|x - \bar{x}\| \leq \|w - \bar{w}\| \leq \frac{2}{M}\|F(w) - F(\bar{w})\| \leq \frac{2}{M} x^T s, \tag{17}$$

for every $w \in B(\bar{w}, \xi)$. If $\mathcal{F}_1 \subseteq B(\bar{w}, \xi)$, then the assumptions of Lemma 2.5 hold with $\epsilon = 1$ and $v = 2/M$, so we are done. Otherwise, consider the set $\mathcal{F}_1 \backslash B(\bar{w}, \xi)$. Clearly, $x^T s > 0$ for all $(x, y, s) \in \mathcal{F}_1 \backslash B(\bar{w}, \xi)$. Because of Lemma 2.2, the set $\mathcal{F}_1 \backslash B(\bar{w}, \xi)$ is compact, and hence there exists $\epsilon > 0$ such that

$$x^T s > \epsilon, \qquad \forall (x, y, s) \in \mathcal{F}_1 \backslash B(\bar{w}, \xi).$$

This implies that $\mathcal{F}_\epsilon \subseteq B(\bar{w}, \xi)$, and hence that implication (12) holds with $v = 2/M$. By Lemma 2.5, Assumption 5 must hold. ∎

3 Technical Results

In this section, we briefly review how the primal-dual affine scaling search direction is computed from a given strictly feasible point (x, y, s). The remainder of the section is taken up with finding bounds on the components of the search direction vectors for points (x, y, s) lying in the neighborhood $\mathcal{N}(\delta, \eta)$, where δ and η are fixed *a priori*.

We start by describing how the primal-dual affine scaling search direction is computed. Given a strictly feasible point (x, y, s), we obtain the primal-dual affine scaling search direction $(\Delta x, \Delta y, \Delta s)$ by applying one step of Newton's method to the nonlinear system defined by (3a), (3b) and the equation $XSe = 0$. Hence $(\Delta x, \Delta y, \Delta s)$ can be obtained by solving the linear system

$$\begin{aligned} S\Delta x + X\Delta s &= -SXe, && \text{(18a)} \\ A\Delta x &= 0, && \text{(18b)} \\ -\nabla^2 f(x)\Delta x + A^T \Delta y + \Delta s &= 0. && \text{(18c)} \end{aligned}$$

The remaining of the section is devoted to finding bounds on the components of the search direction $(\Delta x, \Delta y, \Delta s)$. The main result of this section is stated in Theorem 3.8 which in turn is obtained by combining the results of Lemmas 3.3 and 3.7. We start by stating a simple result that is used in a number of places in this section.

Lemma 3.1 *Suppose Assumption 1 holds. Then if $(\Delta x, \Delta y, \Delta s)$ satisfies relations (18b) and (18c) then $\Delta x^T \Delta s \geq 0$.*

Proof. The proof follows immediately from the convexity of $f(x)$. ∎

We can now state some simple results concerning bounds on certain components of x, Δx, s, and Δs.

Lemma 3.2 *Under Assumptions 1 and 4, if (x, y, s) is feasible then there exists a constant $r > 0$ such that*

$$
\begin{aligned}
x_i &\leq (x^T s)/r, \qquad \forall i \in N, & (19)\\
s_i &\leq (x^T s)/r, \qquad \forall i \in B. & (20)
\end{aligned}
$$

Proof. Consider the strictly complementary solution (x^*, y^*, s^*) as in Assumption 4. Using convexity of f, we obtain

$$0 \leq (x - x^*)^T (s - s^*) = x^T s - {x^*}^T s - x^T s^*$$

where the last equality is due to the fact that ${x^*}^T s^* = 0$. Hence, ${x^*}^T s \leq x^T s$ which implies that

$$s_i \leq (x^T s)/x_i^*, \qquad \forall i \in B.$$

Similarly, we can show that

$$x_i \leq (x^T s)/s_i^*, \qquad \forall i \in N.$$

The result now follows by setting r to be the smallest component of (x_B^*, s_N^*). ■

Lemma 3.3 *Suppose that Assumptions 1 and 4 hold and, for some $\delta > 0$ and $\eta \in (0, 1]$, let $(x, y, s) \in \mathcal{N}(\delta, \eta)$ be given. Let $(\Delta x, \Delta y, \Delta s)$ denote the solution of (18). Then there exists $r > 0$ such that*

$$
\begin{aligned}
|\Delta x_i| &\leq \frac{(x^T s)^{1-(1/2)\delta}}{r \eta^{1/2}}, \qquad \forall i \in N, & (21)\\
|\Delta s_i| &\leq \frac{(x^T s)^{1-(1/2)\delta}}{r \eta^{1/2}}, \qquad \forall i \in B, & (22)\\
s_i &\geq r\eta (x^T s)^{\delta}, \qquad \forall i \in N, & (23)\\
x_i &\geq r\eta (x^T s)^{\delta}, \qquad \forall i \in B. & (24)
\end{aligned}
$$

Proof. Defining $D = X^{1/2} S^{-1/2}$ and using relation (18a), we obtain

$$D^{-1} \Delta x + D \Delta s = -(SX)^{1/2} e.$$

Taking the square of the norm of both sides of the above relation, we obtain

$$\|D^{-1} \Delta x\|^2 + \|D \Delta s\|^2 + 2 \Delta x^T \Delta s = x^T s,$$

and since by Lemma 3.1 $\Delta x^T \Delta s \geq 0$, this relation implies

$$
\begin{aligned}
\|D^{-1} \Delta x\| &\leq (x^T s)^{1/2}, & (25)\\
\|D \Delta s\| &\leq (x^T s)^{1/2}. & (26)
\end{aligned}
$$

Since $(x, y, s) \in \mathcal{N}(\delta, \eta)$, we have

$$x_i s_i \geq \eta (x^T s)^{1+\delta}, \qquad \forall i = 1, \ldots, n. \tag{27}$$

Using relations (25), (19) and (27), we obtain

$$|\Delta x_i| \leq x_i \left(\frac{x^T s}{x_i s_i}\right)^{1/2} \leq \left(\frac{x^T s}{r}\right) \left(\frac{1}{\eta (x^T s)^\delta}\right)^{1/2} = \frac{(x^T s)^{1-(\delta/2)}}{r \eta^{1/2}},$$

for every $i \in N$. This yields (21). Relation (22) follows similarly with the aid of (26), (20) and (27). To show (23), observe that relations (27) and (19) imply

$$s_i \geq \frac{\eta (x^T s)^{1+\delta}}{x_i} \geq r\eta (x^T s)^\delta,$$

for every $i \in N$. Hence (23) follows. Similarly, (24) follows with the aid of (27) and (20). ∎

Providing upper bounds for the remaining components of the search directions, namely Δx_B and Δs_N, is more difficult. This part of the development is based on the approach which appears in the paper of Ye and Anstreicher [12].

We start with the following lemma of Ye and Anstreicher [12, Lemma 3.4].

Lemma 3.4 *Let $M \in \mathbb{R}^{p \times p}$ be a positive semi-definite matrix and assume that (J, L) form a partition of $\{1, \ldots, p\}$. Then,*

$$\mathcal{R} \begin{pmatrix} M_{JJ} & M_{JL} \\ 0 & I \end{pmatrix} = \mathcal{R} \begin{pmatrix} M_{JJ}^T & M_{LJ}^T \\ 0 & -I \end{pmatrix}. \tag{28}$$

As a consequence of the above result, we obtain

Lemma 3.5 *Let $Q \in \mathbb{R}^{n \times n}$ be a positive semi-definite matrix and let $A \in \mathbb{R}^{m \times n}$ be arbitrary. Let (B, N) form an arbitrary partition of $\{1, \ldots, n\}$. Define*

$$U_1 \equiv \begin{pmatrix} Q_{BB} & -A_{.B}^T & Q_{BN} \\ A_{.B} & 0 & A_{.N} \\ 0 & 0 & I \end{pmatrix}, \qquad U_2 \equiv \begin{pmatrix} Q_{BB}^T & A_{.B}^T & Q_{NB}^T \\ -A_{.B} & 0 & -A_{.N} \\ 0 & 0 & -I \end{pmatrix}. \tag{29}$$

Then, $\mathcal{R}(U_1) = \mathcal{R}(U_2)$.

Proof. First observe that the matrix

$$M \equiv \begin{pmatrix} Q_{BB} & -A_{.B}^T & Q_{BN} \\ A_{.B} & 0 & A_{.N} \\ Q_{NB} & -A_{.N}^T & Q_{NN} \end{pmatrix} \tag{30}$$

is positive semi-definite since it can be obtained by a symmetric permutation of the positive semi-definite matrix

$$\begin{pmatrix} Q & -A^T \\ A & 0 \end{pmatrix}. \tag{31}$$

The lemma now follows from Lemma 3.4 applied to the matrix M and the partition (J, L) of $\{1, \ldots, m+n\}$ for which

$$M_{JJ} = \begin{pmatrix} Q_{BB} & -A_{.B}^T \\ A_{.B} & 0 \end{pmatrix}. \tag{32}$$

■

Using the above result, we next prove a lemma that is similar to Lemma 3.5 of Ye and Anstreicher [12]. In fact, the two results are identical if the function $f(x)$ is quadratic.

Lemma 3.6 *Suppose that Assumption 4 holds. Let (x, y, s) be a strictly feasible point and let $D \equiv X^{1/2}S^{-1/2}$. Let $(\Delta x, \Delta y, \Delta s)$ denote the solution of (18). Then $(u, v, w) = (\Delta x_B, \Delta s_N, \Delta y)$ solves the problem*

$$\min_{(u,v,w)} \tfrac{1}{2}\|D_B^{-1}u\|^2 + \tfrac{1}{2}\|D_N v\|^2 + g_B^T u, \tag{33}$$

subject to

$$\begin{aligned} A_{.B}u &= -A_{.N}\Delta x_N, \\ -Q_{BB}u + A_{.B}^T w &= Q_{BN}\Delta x_N - \Delta s_B, \\ -Q_{NB}u + A_{.N}^T w + v &= Q_{NN}\Delta x_N, \end{aligned} \tag{34}$$

where

$$Q = \nabla^2 f(x), \qquad g = -\nabla f(\bar{x}) + \nabla f(x) + \nabla^2 f(x)(\bar{x} - x),$$

and $\bar{x}$ is any primal solution of (1).

Proof. ¿From relations (18b) and (18c) it follows immediately that $(u, v, w) = (\Delta x_B, \Delta s_N, \Delta y)$ is feasible with respect to the constraints (34). The result follows once we verify that $(u, v, w) = (\Delta x_B, \Delta s_N, \Delta y)$ satisfies the KKT conditions for the above problem, namely:

$$\begin{pmatrix} g_B + D_B^{-2}\Delta x_B \\ 0 \\ D_N^2 \Delta s_N \end{pmatrix} \in \mathcal{R}\begin{pmatrix} A_{.B}^T & -Q_{BB}^T & -Q_{NB}^T \\ 0 & A_{.B} & A_{.N} \\ 0 & 0 & I \end{pmatrix} = \mathcal{R}(U_2), \tag{35}$$

where U_2 is the matrix defined in Lemma 3.5. Indeed, using equation (18a), it is easily seen that

$$D^{-2}\Delta x = -(s + \Delta s), \tag{36}$$

$$D^2 \Delta s = -(x + \Delta x). \tag{37}$$

Let $(\bar{y}, \bar{z})$ denote a pair which together with the solution $\bar{x}$ satisfies (3). Using the fact that both (x, y, s) and $(\bar{x}, \bar{y}, \bar{z})$ satisfy (3a)–(3b) and that $(\Delta x, \Delta y, \Delta s)$ satisfy relations (18b) and (18c), one can easily show that

$$A(x + \Delta x - \bar{x}) = 0, \tag{38}$$

$$-Q(x + \Delta x - \bar{x}) + A^T(y + \Delta y - \bar{y}) + (s + \Delta s - \bar{s}) = g. \tag{39}$$

Since $\bar{x}_N = 0$, $\bar{s}_B = 0$ and using relations (36) and (39), we obtain

$$
\begin{aligned}
g_B + D_B^{-2}\Delta x_B &= g_B - (s_B + \Delta s_B) \\
&= g_B - (s_B + \Delta s_B - \bar{s}_B) \\
&= -Q_{B.}(x + \Delta x - \bar{x}) + A_{.B}^T(y + \Delta y - \bar{y}) \\
&= -Q_{BB}(x_B + \Delta x_B - \bar{x}_B) - Q_{BN}(x_N + \Delta x_N) \\
&\quad + A_{.B}^T(y + \Delta y - \bar{y}).
\end{aligned}
\tag{40}
$$

Also, from (37) and (38) it follows that

$$
D_N^2 \Delta s_N = -(x_N + \Delta x_N), \tag{41}
$$

$$
0 = -A_{.B}(x_B + \Delta x_B - \bar{x}_B) - A_{.N}(x_N + \Delta x_N). \tag{42}
$$

Using relations (40), (41) and (42), we obtain

$$
\begin{pmatrix} g_B + D_B^{-2}\Delta x_B \\ 0 \\ D_N^2 \Delta s_N \end{pmatrix} \in \mathcal{R}(U_1). \tag{43}
$$

where U_1 is the matrix defined in Lemma 3.5. Relation (35) now follows from the fact that $\mathcal{R}(U_1) = \mathcal{R}(U_2)$ in view of Lemma 3.5. ∎

Lemma 3.7 *Suppose that Assumptions 1, 3, 4, and 5 hold, and that $(x, y, s) \in \mathcal{N}(\delta, \eta)$ for some $\delta > 0$ and $\eta \in (0, 1]$. Then there exists a constant $C_2 > 0$ independent of δ and η such that*

$$
\|\Delta x_B\| \leq \frac{C_2}{\eta^{3/2}}(x^T s)^{1-(3/2)\delta}, \tag{44}
$$

$$
\|\Delta s_N\| \leq \frac{C_2}{\eta^{3/2}}(x^T s)^{1-(3/2)\delta}. \tag{45}
$$

Proof. It is well known that since (34) is a consistent set of equalities, there is a triple $(\bar{u}, \bar{v}, \bar{w})$ which satisfies (34) and is bounded above by a constant times the norm of the right hand side of (34). Therefore,

$$
\|(\bar{u}, \bar{v}, \bar{w})\| \leq C_3 \|(\Delta x_N, \Delta s_B)\|, \tag{46}
$$

for some constant $C_3 > 0$ independent of δ and η. Define $K_1 \equiv \max(\|D_B\|, \|D_N^{-1}\|)$. Then,

$$
\begin{aligned}
&\|\Delta x_B + D_B^2 g_B\|^2 + \|\Delta s_N\|^2 \\
&\quad \leq \|D_B\|^2 \, \|D_B^{-1}\Delta x_B + D_B g_B\|^2 + \|D_N^{-1}\|^2 \, \|D_N \Delta s_N\|^2 \\
&\quad \leq K_1^2 \left[\|D_B^{-1}\Delta x_B + D_B g_B\|^2 + \|D_N \Delta s_N\|^2 \right] \\
&\quad \leq K_1^2 \left[\|D_B^{-1}\bar{u} + D_B g_B\|^2 + \|D_N \bar{v}\|^2 \right],
\end{aligned}
\tag{47}
$$

where the third inequality follows from the fact that $(\Delta x_B, \Delta s_N, \Delta y)$ is an optimal solution for problem (33)-(34). Relation (47) can be rewritten as

$$\left\| \begin{pmatrix} \Delta x_B + D_B^2 g_B \\ \Delta s_N \end{pmatrix} \right\| \le K_1 \left\| \begin{pmatrix} D_B^{-1}\bar{u} + D_B g_B \\ D_N \bar{v} \end{pmatrix} \right\|.$$

Define $K_2 \equiv \max(\|D_B^{-1}\|, \|D_N\|)$. Then, using the triangle inequality twice in the above relation and relations (46), (21) and (22), we obtain

$$\begin{aligned} \left\| \begin{pmatrix} \Delta x_B \\ \Delta s_N \end{pmatrix} \right\| &\le K_1 \left\{ \left\| \begin{bmatrix} D_B^{-1} & 0 \\ 0 & D_N \end{bmatrix} \right\| \left\| \begin{pmatrix} \bar{u} \\ \bar{v} \end{pmatrix} \right\| + \left\| \begin{pmatrix} D_B g_B \\ 0 \end{pmatrix} \right\| \right\} + \left\| \begin{pmatrix} D_B^2 g_B \\ 0 \end{pmatrix} \right\| \\ &\le K_1 \left\{ K_2 C_3 \left\| \begin{pmatrix} \Delta x_N \\ \Delta s_B \end{pmatrix} \right\| + \|D_B\| \, \|g_B\| \right\} + \|D_B\|^2 \|g_B\| \qquad (48) \\ &\le K_1 K_2 C_3 \sqrt{n} \, \frac{(x^T s)^{1-(\delta/2)}}{r\eta^{1/2}} + 2K_1^2 \|g_B\|. \end{aligned}$$

It remains to estimate K_1, K_2 and $\|g_B\|$ in the above relation. First, we estimate K_1. Observe that since $(x, y, s) \in \mathcal{N}(\delta, \eta)$, relation (27) holds. Using (27), the definition of D and Lemma 2.2, we obtain

$$\|D_B^2\| = \|X_B S_B^{-1}\| = \max_{i \in B} \left\{ \frac{x_i^2}{x_i s_i} \right\} \le \frac{C_1^2}{\eta} (x^T s)^{-(1+\delta)}.$$

Similarly, we have

$$\|D_N^{-2}\| = \|X_N^{-1} S_N\| = \max_{i \in N} \left\{ \frac{s_i^2}{x_i s_i} \right\} \le \frac{C_1^2}{\eta} (x^T s)^{-(1+\delta)}.$$

Hence,

$$K_1^2 \le \frac{C_1^2}{\eta} (x^T s)^{-(1+\delta)}. \qquad (49)$$

We now estimate K_2. Using (27), the definition of D and Lemma 3.2, we obtain

$$\|D_B^{-2}\| = \|X_B^{-1} S_B\| = \max_{i \in B} \left\{ \frac{s_i^2}{x_i s_i} \right\} \le \frac{(x^T s/r)^2}{\eta (x^T s)^{1+\delta}} = \frac{(x^T s)^{1-\delta}}{\eta r^2}.$$

Similarly, we have

$$\|D_N^2\| = \|X_N S_N^{-1}\| = \max_{i \in N} \left\{ \frac{x_i^2}{x_i s_i} \right\} \le \frac{(x^T s/r)^2}{\eta (x^T s)^{1+\delta}} = \frac{(x^T s)^{1-\delta}}{\eta r^2}.$$

Hence,

$$K_2^2 \le \frac{(x^T s)^{1-\delta}}{\eta r^2}. \qquad (50)$$

Next, we estimate $\|g_B\|$. We can assume that the primal solution $\bar{x}$ of Lemma 3.6 is equal to $P(x)$. Using $\bar{x} = P(x)$ in the definition of g and Assumption 5, we obtain

$$\|g_B\| \leq \|g\| = \| - \nabla f(P(x)) + \nabla f(x) + \nabla^2 f(x)(P(x) - x)\| \leq L(x^T s)^2. \tag{51}$$

Substituting the estimates (49), (50) and (51) into expression (48), we obtain

$$\begin{aligned} \left\| \begin{pmatrix} \Delta x_B \\ \Delta s_N \end{pmatrix} \right\| &\leq \left(\frac{C_1^2}{\eta} (x^T s)^{-(1+\delta)} \right)^{1/2} \left(\frac{(x^T s)^{1-\delta}}{\eta r^2} \right)^{1/2} \sqrt{n}\, C_3 \frac{(x^T s)^{1-(1/2)\delta}}{r \eta^{1/2}} \\ &\quad + 2 \left(\frac{C_1^2}{\eta} (x^T s)^{-(1+\delta)} \right) \left(L (x^T s)^2 \right) \\ &= \frac{(x^T s)^{1-(3/2)\delta}}{\eta^{3/2}} \left[\frac{\sqrt{n}\, C_1 C_3}{r^2} + 2 C_1^2 L \eta^{1/2} (x^T s)^{\delta/2} \right]. \end{aligned}$$

Since $x^T s \leq 1$ and $\eta \in (0, 1]$, we can bound the bracketed term by a constant C_2 independent of η and δ. We thus obtain

$$\left\| \begin{pmatrix} \Delta x_B \\ \Delta s_N \end{pmatrix} \right\| \leq \frac{C_2}{\eta^{3/2}} (x^T s)^{1-(3/2)\delta},$$

which is the desired result. ∎

We can now merge the results of Lemmas 3.3 and 3.7 to obtain the following theorem.

Theorem 3.8 *Suppose that Assumptions 1, 3, 4, and 5 hold, and that $(x, y, s) \in \mathcal{N}(\delta, \eta)$ for some $\delta > 0$ and $\eta \in (0, 1]$. Then there exists a constant C_4 independent of δ and η such that*

$$\begin{aligned} \|\Delta x\| &\leq \frac{C_4}{\eta^{3/2}} (x^T s)^{1-(3/2)\delta}, \\ \|\Delta s\| &\leq \frac{C_4}{\eta^{3/2}} (x^T s)^{1-(3/2)\delta}. \end{aligned}$$

4 A locally convergent algorithm

In this section, we discuss a basic primal-dual affine scaling algorithm based on the step calculation (18) and show that rapid local convergence can be obtained for certain choices of the step lengths. In the next section, this algorithm is embedded in the globally convergent potential reduction algorithm of Monteiro [9].

In describing the algorithm, we first show how the step calculated in (18) can be used to move from one iterate to the next. In general, we cannot simply move along the direction $(\Delta x, \Delta y, \Delta s)$ from the current point (x, y, s), since this may lead to a new point that violates the condition (3a). Instead, we take a "curved" step in the s component and make the transition

$$(x, y, s) \to (x + \alpha\Delta x, y + \alpha\Delta y, s(\alpha)),$$

where

$$s(\alpha) = \nabla f(x + \alpha\Delta x) - A^T(y + \alpha\Delta y). \tag{52}$$

The algorithm proceeds as follows:

Algorithm PDA

initially: choose (x^0, y^0, s^0) strictly feasible

for $k = 1, 2, \ldots$

Find $(\Delta x^k, \Delta y^k, \Delta s^k)$ by solving (18) with $(x, y, s) = (x^k, y^k, s^k)$;

Choose $\alpha_k > 0$ such that $(x^k + \alpha_k\Delta x^k, y^k + \alpha_k\Delta y^k, s^k(\alpha_k))$
is strictly feasible;

Set $(x^{k+1}, y^{k+1}, s^{k+1}) = (x^k + \alpha_k\Delta x^k, y^k + \alpha_k\Delta y^k, s^k(\alpha_k))$;

end for

The following neighborhood will play an important role in the subsequent development. Let $p > 0$, $\delta > 0$, $\eta \in (0, 1]$ and $\gamma > 0$ be given and define

$$\mathcal{N}_p(\delta, \eta, \gamma) \equiv \{(x, y, s) \in \mathcal{N}(\delta, \eta) \mid (x^T s)^{1-p\delta}/(\delta\eta^p) \leq 1/\gamma\}. \tag{53}$$

The proof of the following lemma is straightforward and is left to the reader.

Lemma 4.1 *Assume that $p > q > 0$, $\delta > 0$ and $\eta \in (0, 1]$. Then*

$$(x^T s)^{1-q\delta}/\eta^q \leq (x^T s)^{1-p\delta}/\eta^p, \qquad \forall x \in \mathcal{N}(\delta, \eta), \tag{54}$$

and

$$\mathcal{N}_p(\delta, \eta, \gamma) \subseteq \mathcal{N}_q(\delta, \eta, \gamma), \qquad \forall \gamma > 0. \tag{55}$$

Lemma 4.2 *Suppose that all Assumptions 1–5 hold and let $(x, y, s) \in \mathcal{N}_4(\delta, \eta, \gamma)$ be given, where $\delta > 0$, $\eta \in (0, 1]$, and $\gamma \geq \delta$. Then, there exist a constant C_5 independent of δ, η and γ such that for all $\alpha \in [0, 1]$,*

$$|(x_i + \alpha\Delta x_i)s_i(\alpha) - (1 - \alpha)x_i s_i| \leq C_5\alpha^2 \frac{(x^T s)^{2-3\delta}}{\eta^3}, \qquad \forall i = 1, \ldots, n, \tag{56}$$

and

$$\left|(x+\alpha\Delta x)^T s(\alpha) - (1-\alpha)x^T s\right| \le nC_5\alpha^2 \frac{(x^Ts)^{2-3\delta}}{\eta^3}. \tag{57}$$

Proof. We first note that (57) follows immediately from (56). Using relations (52), (3a) and (18c), we obtain

$$\begin{aligned} s(\alpha) &= \nabla f(x+\alpha\Delta x) - A^T y - \alpha A^T \Delta y \\ &= \nabla f(x+\alpha\Delta x) - (\nabla f(x) - s) - \alpha(\nabla^2 f(x)\Delta x - \Delta s) \\ &= s + \alpha\Delta s + [\nabla f(x+\alpha\Delta x) - \nabla f(x) - \alpha\nabla^2 f(x)\Delta x] \\ &= s + \alpha\Delta s + \alpha\int_0^1 [\nabla^2 f(x+t\alpha\Delta x) - \nabla^2 f(x)]\Delta x\, dt \\ &= s + \alpha\Delta s + \alpha R(x,\alpha\Delta x)\Delta x \end{aligned} \tag{58}$$

where

$$R(x,\Delta x) \stackrel{\triangle}{=} \int_0^1 [\nabla^2 f(x+t\Delta x) - \nabla^2 f(x)]\, dt. \tag{59}$$

Relations (58) and (18a) imply

$$\begin{aligned} (x_i+\alpha\Delta x_i)s_i(\alpha) &= (x_i+\alpha\Delta x_i)\left[s_i + \alpha\Delta s_i + \alpha\left\{R(x,\alpha\Delta x)\Delta x\right\}_i\right] \\ &= (1-\alpha)x_i s_i + \alpha^2\Delta x_i\Delta s_i + \alpha(x_i+\alpha\Delta x_i)\left\{R(x,\alpha\Delta x)\Delta x\right\}_i, \end{aligned}$$

so that

$$\begin{aligned} &|(x_i+\alpha\Delta x_i)s_i(\alpha) - (1-\alpha)x_i s_i| \\ &\qquad \le \alpha^2|\Delta x_i\Delta s_i| + \alpha|(x_i+\alpha\Delta x_i)|\,|\left\{R(x,\alpha\Delta x)\Delta x\right\}_i|. \end{aligned} \tag{60}$$

We will now estimate each of the quantities which appears in the right hand side of (60). Regardless of whether $i \in B$ or $i \in N$, we have from Lemmas 3.3 and 3.7 that

$$|\Delta x_i \Delta s_i| \le \frac{(x^Ts)^{1-(1/2)\delta}}{r\eta^{1/2}} \frac{C_2}{\eta^{3/2}}(x^Ts)^{1-(3/2)\delta} \le \frac{C_6}{\eta^2}(x^Ts)^{2-2\delta}, \tag{61}$$

where

$$C_6 \equiv C_2/r. \tag{62}$$

Relation (55) with $p=4$ and $q=3/2$ and the assumption that $(x,y,s)\in\mathcal{N}_4(\delta,\eta,\gamma)$ imply that $(x,y,s)\in\mathcal{N}_{3/2}(\delta,\eta,\gamma)$. Hence, $(x^Ts)^{1-(3\delta/2)}/\eta^{3/2} \le \delta/\gamma \le 1$ in view of (53) and the assumption that $\gamma \ge \delta$. Using this inequality, Lemma 2.2 and Theorem 3.8, we obtain

$$|x_i + \alpha\Delta x_i| \le |x_i| + |\Delta x_i| \le C_1 + C_4\frac{(x^Ts)^{1-(3\delta/2)}}{\eta^{3/2}} \le C_1 + C_4. \tag{63}$$

Consider now the set

$$\mathcal{K} \stackrel{\triangle}{=} \{x \in \mathbb{R}^n \mid Ax = b,\ 0 \le x_i \le C_1 + C_4,\ \ \forall i = 1,\ldots,n\}.$$

Clearly, $\mathcal{K}$ is a compact subset of $\{x \mid Ax = b,\ x \geq 0\}$. Hence, by Lemma 2.4 there exists a constant $C > 0$ such that

$$\|\nabla^2 f(x^1) - \nabla^2 f(x^2)\| \leq C\|x^1 - x^2\|, \qquad \forall x^1, x^2 \in \mathcal{K}.$$

Using this last relation and Theorem 3.8, we obtain

$$\begin{aligned} |\{R(x,\alpha\Delta x)\Delta x\}_i| &\leq \|R(x,\alpha\Delta x)\|\,\|\Delta x\| \\ &\leq \left(\int_0^1 \|\nabla^2 f(x+t\alpha\Delta x) - \nabla^2 f(x)\|\,dt\right)\|\Delta x\| \\ &\leq \alpha\tfrac{C}{2}\|\Delta x\|^2 \\ &\leq \alpha\tfrac{CC_4^2}{2}\tfrac{(x^Ts)^{2-3\delta}}{\eta^3}. \end{aligned} \tag{64}$$

Substituting relations (61), (63) and (64) into (60), we obtain

$$\begin{aligned} &|(x_i+\alpha\Delta x_i)s_i(\alpha) - (1-\alpha)x_is_i| \\ &\leq \alpha^2\tfrac{C_6}{\eta^2}(x^Ts)^{2-2\delta} + \alpha\,(C_1+C_4)\left(\alpha\tfrac{CC_4^2}{2}\tfrac{(x^Ts)^{2-3\delta}}{\eta^3}\right) \\ &\leq \alpha^2\tfrac{(x^Ts)^{2-3\delta}}{\eta^3}\left\{C_6\eta(x^Ts)^{\delta} + (C_1+C_4)(CC_4^2/2)\right\} \\ &\leq \alpha^2\tfrac{(x^Ts)^{2-3\delta}}{\eta^3}\left\{C_6 + (C_1+C_4)(CC_4^2/2)\right\}. \end{aligned} \tag{65}$$

Relation (56) now follows if we let

$$C_5 \equiv C_6 + (C_1+C_4)(CC_4^2/2). \tag{66}$$

■

In Lemma 4.3, we show that the property of belonging to the set $\mathcal{N}_4(\delta,\eta,\gamma)$ is "inheritable" when γ is sufficiently large; that is, we can take a substantial step with parameter α_k from the current iterate (x^k, y^k, s^k) and remain within $\mathcal{N}_4(\delta,\eta,\gamma)$.

Lemma 4.3 *Suppose that $\delta \in (0, 1/4)$ and $\eta \in (0,1]$ are given and that Assumptions 1–5 hold. Then there exists a constant $\bar\gamma \geq 1$ independent of δ and η that satisfies the following implication. If $\gamma \geq \bar\gamma$ and $(x,y,s) \in \mathcal{N}_4(\delta,\eta,\gamma)$ then $(x+\alpha\Delta x, y+\alpha\Delta y, s(\alpha)) \in \mathcal{N}_4(\delta,\eta,\gamma)$ for all*

$$\alpha \in \left[0, 1 - \frac{\gamma}{\delta\eta^4}(x^Ts)^{1-4\delta}\right]. \tag{67}$$

Moreover, we have that

$$(x+\alpha\Delta x)^T s(\alpha) \leq \left(1 - \alpha\left(1 - \frac{\delta}{2}\right)\right)x^Ts, \qquad \forall\alpha \in [0,1]. \tag{68}$$

Proof. Consider the constant C_5 as in the statement of Lemma 4.2 and define

$$\bar{\gamma} \equiv \max\{1, (2n+1)C_5)\}. \tag{69}$$

We will show that $\bar{\gamma}$ fulfills the requirements of the lemma. Indeed, let $\gamma \geq \bar{\gamma}$ and $(x, y, s) \in \mathcal{N}_4(\delta, \eta, \gamma)$ be given. Note that (69) implies that $\gamma \geq 1$. Our first task is to show that the range of α values such that

$$(x_i + \alpha \Delta x_i) s_i(\alpha) \geq \eta \left[(x + \alpha \Delta x)^T s(\alpha) \right]^{1+\delta}, \tag{70}$$

for all $i = 1, 2, \ldots, n$, includes the interval in (67). By relations (56) and (27), we have

$$\begin{aligned}
(x_i + \alpha \Delta x_i) s_i(\alpha) &\geq (1-\alpha) x_i s_i - C_5 \alpha^2 \frac{(x^T s)^{2-3\delta}}{\eta^3} \\
&\geq (1-\alpha) \eta (x^T s)^{1+\delta} - C_5 \alpha^2 \frac{(x^T s)^{2-3\delta}}{\eta^3} \\
&\geq \eta (x^T s)^{1+\delta} \left[1 - \alpha \left(1 + C_5 \frac{(x^T s)^{1-4\delta}}{\eta^4} \right) \right].
\end{aligned} \tag{71}$$

for all $\alpha \in [0, 1]$. Similarly, from relation (57), we obtain

$$(x + \alpha \Delta x)^T s(\alpha) \leq x^T s \left[1 - \alpha \left(1 - n C_5 \frac{(x^T s)^{1-3\delta}}{\eta^3} \right) \right], \qquad \forall \alpha \in [0, 1]. \tag{72}$$

Relation (55) with $p = 4$ and $q = 3$ and the assumption that $(x, y, s) \in \mathcal{N}_4(\delta, \eta, \gamma)$ imply that $(x, y, s) \in \mathcal{N}_3(\delta, \eta, \gamma)$. Hence, using (53) we obtain

$$(x^T s)^{1-3\delta} / \eta^3 \leq \delta / \gamma. \tag{73}$$

We now use the fact that for all $\delta \in (0, 1)$ and $u \geq -1$, $(1+u)^\delta \leq 1 + \delta u$. Using this inequality, (72), (73) and the fact that $n C_5 \delta / \gamma \leq 1$, we find that

$$\begin{aligned}
&\eta \left[(x + \alpha \Delta x)^T s(\alpha) \right]^{1+\delta} \\
&\leq \eta (x^T s)^{1+\delta} \left[1 - \alpha \left(1 - n C_5 \tfrac{(x^T s)^{1-3\delta}}{\eta^3} \right) \right] \left[1 - \delta \alpha \left(1 - n C_5 \tfrac{(x^T s)^{1-3\delta}}{\eta^3} \right) \right] \\
&\leq \eta (x^T s)^{1+\delta} \left[1 - \alpha (1+\delta) \left(1 - n C_5 \tfrac{(x^T s)^{1-3\delta}}{\eta^3} \right) + \delta \alpha^2 \left(1 - n C_5 \tfrac{(x^T s)^{1-3\delta}}{\eta^3} \right)^2 \right] \\
&\leq \eta (x^T s)^{1+\delta} \left[1 - \alpha (1+\delta) \left(1 - n C_5 \tfrac{(x^T s)^{1-3\delta}}{\eta^3} \right) + \delta \alpha^2 \right].
\end{aligned} \tag{74}$$

The inequality (70) will hold if the right-hand side of (74) is smaller than or equal to the right-hand side of (71). In other words, we want α to be chosen so that

$$\begin{aligned}
&\eta (x^T s)^{1+\delta} \left[1 - \alpha (1+\delta) \left(1 - n C_5 \tfrac{(x^T s)^{1-3\delta}}{\eta^3} \right) + \delta \alpha^2 \right] \\
&\quad \leq \eta (x^T s)^{1+\delta} \left[1 - \alpha \left(1 + C_5 \tfrac{(x^T s)^{1-4\delta}}{\eta^4} \right) \right].
\end{aligned}$$

This is equivalent to

$$(1+\delta)\left(1 - nC_5 \frac{(x^Ts)^{1-3\delta}}{\eta^3}\right) - \delta\alpha \geq \left(1 + C_5 \frac{(x^Ts)^{1-4\delta}}{\eta^4}\right),$$

that is,

$$\alpha \leq \frac{1}{\delta}\left\{(1+\delta)\left[1 - nC_5 \frac{(x^Ts)^{1-3\delta}}{\eta^3}\right] - \left[1 + C_5 \frac{(x^Ts)^{1-4\delta}}{\eta^4}\right]\right\} \tag{75}$$

To show that $(x + \alpha\Delta x, y + \alpha\Delta y, s(\alpha)) \in \mathcal{N}(\delta, \eta)$ for every α satisfying (67), it is sufficient to show that the right hand side of (75) is greater than or equal to $1 - (\gamma/\delta\eta^4)(x^Ts)^{1-4\delta}$. Indeed, we have

$$\begin{aligned}
&\frac{1}{\delta}\left\{(1+\delta)\left[1 - nC_5 \frac{(x^Ts)^{1-3\delta}}{\eta^3}\right] - \left[1 + C_5 \frac{(x^Ts)^{1-4\delta}}{\eta^4}\right]\right\} \\
&\quad \geq \frac{1}{\delta}\left\{\delta - (1+\delta)nC_5 \frac{(x^Ts)^{1-3\delta}}{\eta^3} - C_5 \frac{(x^Ts)^{1-4\delta}}{\eta^4}\right\} \\
&\quad \geq \frac{1}{\delta}\left\{\delta - \frac{(x^Ts)^{1-4\delta}}{\eta^4}\left[C_5 + 2nC_5\eta(x^Ts)^{\delta}\right]\right\} \\
&\quad \geq 1 - \frac{\gamma}{\delta\eta^4}(x^Ts)^{1-4\delta},
\end{aligned}$$

where the last inequality follows from the fact that $\eta(x^Ts)^{\delta} \leq 1$, $\gamma \geq \bar{\gamma}$ and relation (69). We have thus shown that (70) holds for every α satisfying (67). The validity of relation (68) guarantees that $(x + \alpha\Delta x, y + \alpha\Delta y, s(\alpha)) \in \mathcal{N}_4(\delta, \eta, \gamma)$ for all α satisfying (67). Therefore, to complete the proof of the lemma, it is sufficient to show that (68) holds. Indeed, using (72), (73) and (69), we obtain

$$\begin{aligned}
(x + \alpha\Delta x)^T s(\alpha) &\leq x^Ts\left[1 - \alpha + \alpha nC_5 \frac{(x^Ts)^{1-3\delta}}{\eta^3}\right] \\
&\leq x^Ts\left[1 - \alpha + \alpha\frac{nC_5\delta}{\gamma}\right] \\
&\leq x^Ts\left(1 - \alpha\left(1 - \frac{\delta}{2}\right)\right),
\end{aligned}$$

for every $\alpha \in [0,1]$. ■

The following lemma will be useful later in the selection of the step sizes α_k for algorithm PDA.

Lemma 4.4 *Suppose that Assumptions 1–5 hold and let $\delta \in (0, 1/4)$, $\eta \in (0, 1]$ and $\gamma \geq \bar{\gamma}$ be given, where $\bar{\gamma}$ is the constant specified in Lemma 4.3. Let $(x, y, s) \in \mathcal{N}_4(\delta, \eta, \gamma)$ be given. Consider the solution $(\Delta x, \Delta y, \Delta s)$ of (18) and define $\hat{\alpha}$ as*

$$\hat{\alpha} = \sup\{\alpha \geq 0 \mid x + \alpha\Delta x \geq 0, s + \alpha\Delta s \geq 0\}. \tag{76}$$

Then,

$$|\hat{\alpha} - 1| \leq \frac{\gamma}{\delta\eta^4}(x^T s)^{1-4\delta}.$$

Proof. ¿From the definition of $\hat{\alpha}$, it is easy to show that

$$\hat{\alpha}^{-1} = \max(-X^{-1}\Delta x, -S^{-1}\Delta s). \tag{77}$$

It follows from relations (23), (24), (62) and (54) with $p = 4$ and $q = 5/2$, and Lemma 3.7 that

$$\begin{aligned} \max\left\{\max_{i\in B}\left(\frac{|\Delta x_i|}{x_i}\right), \max_{i\in N}\left(\frac{|\Delta s_i|}{s_i}\right)\right\} &\leq \frac{C_2(x^T s)^{1-(3/2)\delta}/\eta^{3/2}}{r\eta(x^T s)^\delta} \\ &\leq \frac{C_6(x^T s)^{1-(5/2)\delta}}{\eta^{5/2}} \\ &\leq \frac{C_6(x^T s)^{1-4\delta}}{\eta^4}. \end{aligned} \tag{78}$$

By (69) and (66), we have $C_6 \leq C_5 \leq \bar{\gamma}/(2n+1) \leq \gamma/2$. Using this fact in (78), together with $\delta < 1$ and $(x, y, s) \in \mathcal{N}_4(\delta, \eta, \gamma)$, we find that

$$\begin{aligned} \max\left\{\max_{i\in B}\left(\frac{|\Delta x_i|}{x_i}\right), \max_{i\in N}\left(\frac{|\Delta s_i|}{s_i}\right)\right\} &\leq \frac{C_6(x^T s)^{1-4\delta}}{\eta^4} \\ &\leq \frac{\gamma}{2\delta\eta^4}(x^T s)^{1-4\delta} \leq \frac{1}{2}. \end{aligned} \tag{79}$$

¿From (18a), we have that

$$\frac{\Delta x_i}{x_i} + \frac{\Delta s_i}{s_i} = -1, \qquad \forall i = 1, \ldots, n.$$

Using this relation and (79), we obtain

$$\begin{aligned} \left|\tfrac{\Delta x_i}{x_i} + 1\right| &\leq \tfrac{|\Delta s_i|}{s_i} \leq \tfrac{\gamma}{2\delta\eta^4}(x^T s)^{1-4\delta} \leq \tfrac{1}{2}, \qquad \forall i \in N, \\ \left|\tfrac{\Delta s_i}{s_i} + 1\right| &\leq \tfrac{|\Delta x_i|}{x_i} \leq \tfrac{\gamma}{2\delta\eta^4}(x^T s)^{1-4\delta} \leq \tfrac{1}{2}, \qquad \forall i \in B. \end{aligned} \tag{80}$$

¿From (79) and (80) it follows that

$$\max(-X^{-1}\Delta x, -S^{-1}\Delta s) = \max(-X_N^{-1}\Delta x_N, -S_B^{-1}\Delta s_B). \tag{81}$$

Using (80) again and this last relation, we obtain

$$|\hat{\alpha}^{-1} - 1| \leq \frac{\gamma}{2\delta\eta^4}(x^T s)^{1-4\delta} \leq \frac{1}{2}. \tag{82}$$

Hence,

$$|\hat{\alpha}-1| \leq \hat{\alpha}\frac{\gamma}{2\delta\eta^4}(x^Ts)^{1-4\delta} \leq \frac{\gamma}{\delta\eta^4}(x^Ts)^{1-4\delta},$$

since $\hat{\alpha} \leq 2$ due to (82). ∎

Finally we show that, given (strong) assumptions on the starting point for Algorithm PDA, a sequence of iterates for which the duality gap ${x^k}^T s^k$ converges superlinearly to zero can be generated.

Theorem 4.5 *Suppose that the Assumptions 1–5 hold and let $\delta \in (0, 1/4)$ be given. Let $\bar{\gamma}$ be the constant as in the statement of Lemma 4.3. Assume that a constant $\gamma \geq \bar{\gamma}$ and a strictly feasible solution (x^0, y^0, s^0) satisfy*

$$\min_i x_i^0 s_i^0 \geq \left(\frac{2\gamma}{\delta}\right)^{1/4} \left({x^0}^T s^0\right)^{5/4}, \tag{83}$$

and

$$\eta \triangleq \left(\frac{2\gamma}{\delta}\right)^{1/4} \left({x^0}^T s^0\right)^{(1-4\delta)/4} \leq 1. \tag{84}$$

Then the following statements hold.

(i) If (x^0, y^0, s^0) is used as the starting point for Algorithm PDA and the sequence of step sizes $\{\alpha_k\}$ satisfies

$$\alpha_k \in J_k \triangleq \left[0, 1 - \frac{\gamma}{\delta\eta^4}\left({x^k}^T s^k\right)^{1-4\delta}\right], \qquad \forall k = 1, 2, \ldots \tag{85}$$

then

$$(x^k, y^k, s^k) \in \mathcal{N}_4(\delta, \eta, 2\gamma), \qquad \forall k \geq 0.$$

(ii) Let $\bar{\delta}$ be an arbitrary constant satisfying $\bar{\delta} \in (0, 1-4\delta)$. If the sequence of step sizes $\{\alpha_k\}$ is defined as

$$\alpha_k = \max\left(\frac{1}{3}, 1 - \sigma\left({x^k}^T s^k\right)^{\bar{\delta}}\right)\hat{\alpha}_k, \qquad \forall k \geq 0, \tag{86}$$

where

$$\sigma = \left({x^0}^T s^0\right)^{-(1-4\delta)} = \frac{2\gamma}{\delta\eta^4}, \tag{87}$$

and $\hat{\alpha}_k$ is any number satisfying

$$|\hat{\alpha}_k - 1| \leq \frac{\gamma}{\delta\eta^4}\left({x^k}^T s^k\right)^{1-4\delta}, \tag{88}$$

then, $\alpha_k \in J_k$ for all $k \geq 0$ and the sequence of duality gaps corresponding to $\{x_k, y_k, s_k\}$ eventually converges q-superlinearly to zero according to

$$ {x^{k+1}}^T s^{k+1} \leq \frac{31}{9}\sigma\left({x^k}^T s^k\right)^{(1+\bar{\delta})}. \tag{89}$$

Proof. We first show (i). In view of Lemma 4.3, it is sufficient to show that $\{x_0, y_0, s_0\} \in \mathcal{N}_4(\delta, \eta, 2\gamma)$. Indeed, relation (84) implies that ${x^0}^T s^0 \leq 1$ and

$$({x^0}^T s^0)^{1-4\delta}/(\delta\eta^4) = 1/(2\gamma).$$

Moreover, using (83) and (84), we obtain

$$\begin{aligned}
\min_i x_i^0 s_i^0 &\geq \left(\frac{2\gamma}{\delta}\right)^{1/4} ({x^0}^T s^0)^{5/4} \\
&= \left(\frac{2\gamma}{\delta}\right)^{1/4} ({x^0}^T s^0)^{(1-4\delta)/4} ({x^0}^T s^0)^{1+\delta} \\
&= \eta ({x^0}^T s^0)^{1+\delta}.
\end{aligned}$$

Hence, $\{x_0, y_0, s_0\} \in \mathcal{N}_4(\delta, \eta, 2\gamma)$.

We now show (ii), starting with a proof that $\alpha_k \in J_k$ for all $k \geq 0$. Indeed, using relation (88) and the fact that $(x^k, y^k, s^k) \in \mathcal{N}_4(\delta, \eta, 2\gamma)$, and in particular that

$$({x^k}^T s^k)^{1-4\delta}/(\delta\eta^4) \leq 1/(2\gamma), \tag{90}$$

we obtain

$$\begin{aligned}
\frac{1}{3}\hat{\alpha}_k &\leq \frac{1}{3}\left(1 + \frac{\gamma}{\delta\eta^4}\left({x^k}^T s^k\right)^{1-4\delta}\right) \\
&\leq \frac{1}{3}\left(1 + \frac{1}{2}\right) = \frac{1}{2} \leq 1 - \frac{\gamma}{\delta\eta^4}\left({x^k}^T s^k\right)^{1-4\delta}.
\end{aligned} \tag{91}$$

Moreover, from (88), (87) and the fact that $\bar{\delta} \in (0, 1-4\delta)$, it follows that

$$\begin{aligned}
&\left(1 - \sigma\left({x^k}^T s^k\right)^{\bar{\delta}}\right)\hat{\alpha}_k \\
&\quad\leq \left(1 - \frac{2\gamma}{\delta\eta^4}\left({x^k}^T s^k\right)^{\bar{\delta}}\right)\left(1 + \frac{\gamma}{\delta\eta^4}\left({x^k}^T s^k\right)^{1-4\delta}\right) \\
&\quad\leq 1 - \frac{2\gamma}{\delta\eta^4}\left({x^k}^T s^k\right)^{\bar{\delta}} + \frac{\gamma}{\delta\eta^4}\left({x^k}^T s^k\right)^{1-4\delta} \\
&\quad\leq 1 - \frac{\gamma}{\delta\eta^4}\left({x^k}^T s^k\right)^{1-4\delta}.
\end{aligned} \tag{92}$$

Together, (91) and (92) imply that $\alpha_k \in J_k$. Before showing that (89) holds, we show that ${x^k}^T s^k$ converges at least q-linearly to zero. Indeed, by (90) and (88), we have

$$\alpha_k \geq \frac{1}{3}\hat{\alpha}_k \geq \frac{1}{3}\left[1 - \frac{\gamma}{\delta\eta^4}\left({x^k}^T s^k\right)^{1-4\delta}\right] \geq \frac{1}{6}.$$

Using this inequality together with (68) and $\delta \in (0, 1/4)$, we obtain

$$x^{k+1^T} s^{k+1} \leq (1 - \alpha_k(1 - \delta/2)) \, x^{k^T} s^k \leq \left(1 - \frac{1}{6}(1 - \delta/2)\right) x^{k^T} s^k < \frac{41}{48} x^{k^T} s^k, \qquad (93)$$

as required. We now show that (89) holds. Because of the q-superlinear convergence property, we can choose an integer K such that

$$\sigma (x^{k^T} s^k)^{\bar{\delta}} \leq 2/3, \qquad \forall k \geq K. \qquad (94)$$

In view of (86), this implies that

$$\alpha_k = \left(1 - \sigma (x^{k^T} s^k)^{\bar{\delta}}\right) \hat{\alpha}_k, \qquad \forall k \geq K. \qquad (95)$$

¿From relations (88), (95) and (72), it follows that

$$x^{k+1^T} s^{k+1} \leq x^{k^T} s^k \left[1 - \alpha_k \left(1 - nC_5 \frac{(x^{k^T} s^k)^{1-3\delta}}{\eta^3}\right)\right] \leq$$
$$x^{k^T} s^k \left[1 - \left(1 - \sigma (x^{k^T} s^k)^{\bar{\delta}}\right) \left(1 - \frac{\gamma}{\delta \eta^4} (x^{k^T} s^k)^{1-4\delta}\right) \left(1 - nC_5 \frac{(x^{k^T} s^k)^{1-3\delta}}{\eta^3}\right)\right],$$

for all $k \geq K$. Using (69), (87), relation (54) with $p = 4$ and $q = 3$, and (94), we have that

$$nC_5 \frac{(x^{k^T} s^k)^{1-3\delta}}{\eta^3} \leq \frac{\gamma}{\delta} \frac{(x^{k^T} s^k)^{1-4\delta}}{\eta^4} \leq \frac{\sigma}{2} (x^{k^T} s^k)^{\bar{\delta}} \leq \frac{1}{3}.$$

Hence,

$$\begin{aligned} x^{k+1^T} s^{k+1} &\leq (x^{k^T} s^k) \left[1 - \left(1 - \sigma (x^{k^T} s^k)^{\bar{\delta}}\right)^3\right] \\ &\leq (x^{k^T} s^k) \left[3\sigma (x^{k^T} s^k)^{\bar{\delta}} + \sigma^3 (x^{k^T} s^k)^{3\bar{\delta}}\right] \\ &\leq \left(3 + \frac{4}{9}\right) \sigma (x^{k^T} s^k)^{1+\bar{\delta}}, \end{aligned}$$

for every $k \geq K$. Hence, (89) holds. ∎

A number of choices for $\hat{\alpha}_k$ that insures the validity of (88) are possible. The simplest is to set $\hat{\alpha}_k \equiv 1$ but, from Lemma 4.4, we see that choosing $\hat{\alpha}_k$ according to (76) is another possibility. A third possibility is to choose

$$\hat{\alpha}_k = \sup \left\{ \alpha \mid (x^k + \theta \Delta x^k, s^k(\theta)) \geq 0 \text{ for all } \theta \in [0, \alpha] \right\},$$

which can be shown, by a technique similar to that used in Lemma 4.4, to satisfy (88). However, this last choice is not attractive since $\hat{\alpha}_k$ is not easily computable.

5 A globally and superlinearly convergent potential reduction algorithm

In this section we embed the locally convergent algorithm of the previous section in the potential reduction algorithm of Monteiro [9].

We start by outlining a version of the algorithm in [9] in which the centering parameter (namely, the parameter σ_k in the notation of [9]) is chosen to be zero. The directions so obtained are referred to as *primal-dual affine scaling directions*. For simplicity of exposition, we fix certain parameters in his algorithm to numerical values. The algorithm makes use of the Tanabe-Todd-Ye potential function:

$$\phi_q(x, s) = q \log x^T s - \sum_{i=1}^{n} \log x_i s_i \tag{96}$$

where $q > n$.

Algorithm PDPR

initially: choose fixed constants $\beta > 1$ and $\mu \in (0, 1)$ and let (x^0, y^0, s^0) be a strictly feasible point;

for $k = 1, 2, \ldots$

Find $(\Delta x^k, \Delta y^k, \Delta s^k)$ by solving (18) with $(x, y, s) = (x^k, y^k, s^k)$;

Set $\tau_k = \theta_k \hat{\alpha}_k$ where $\theta_k \geq 1/3$ and

$$\hat{\alpha}_k \equiv \max\{\alpha \mid x_k + \alpha \Delta x_k \geq 0, s_k + \alpha \Delta s_k \geq 0\}; \tag{97}$$

Define $\alpha_k = \beta^{-m_k}\tau_k$, where m_k is the smallest non-negative integer for which

$$x^k + \beta^{-m_k}\tau_k \Delta x^k > 0, \qquad s^k(\beta^{-m_k}\tau_k) > 0,$$

and

$$\begin{aligned} &\phi_q(x^k + \beta^{-m_k}\tau_k \Delta x^k, s^k(\beta^{-m_k}\tau_k)) \\ \leq\ \ &\phi_q(x^k, s^k) + \mu(\beta^{-m_k}\tau_k)\nabla\phi_q(x^k, s^k)^T \begin{bmatrix} \Delta x^k \\ \Delta s^k \end{bmatrix}; \end{aligned} \tag{98}$$

Set $(x^{k+1}, y^{k+1}, s^{k+1}) = (x^k + \alpha_k \Delta x^k, y^k + \alpha_k \Delta y^k, s^k(\alpha_k))$;

end for

Some comments are in order. First, there is nothing special with the constant 1/3 which gives a lower bound for θ_k. In fact, any positive constant could be used in place of 1/3. Second, α_k defined in (97) is well-defined in the sense that the maximum of the set exists. This can be shown as follows. By relation (18a), it follows that ${x^k}^T \Delta s^k + {s^k}^T \Delta x^k = -{x^k}^T s^k < 0$. Hence, some component of either Δx^k or Δs^k must be negative. Thus, the maximum in (97) is finite.

The following global convergence theorem can be proved as a consequence of the results presented in Monteiro [9].

Theorem 5.1 *In Algorithm PDPR,*

$$\lim_{k \to \infty} {x^k}^T s^k = 0,$$

and the sequence $\{(x^k, y^k, s^k)\}$ *has at least one limit point. If* $(\bar{x}, \bar{y}, \bar{s})$ *is one such limit point then* $\bar{x}$ *solves the primal problem (1), while* $(\bar{x}, \bar{y})$ *solves the corresponding dual problem (2).*

The theorem above guarantees that Algorithm PDPR converges globally from any feasible starting point. To prove local superlinear convergence of Algorithm PDPR, we need the assumption stated below that some iterate (x^K, y^K, s^K) of the algorithm satisfies a condition similar to (83) and (84). In the final result of this section, we show that this assumption can be dispensed with for certain values of q.

Assumption 6 *For fixed constants* $\gamma \geq \bar{\gamma}$ *and* $\delta \in (0, 1/4)$, *assume that an integer* $K > 0$ *is* known *for which the iterate* (x^K, y^K, s^K) *of Algorithm PDPR satisfies*

$$\min_i x_i^K s_i^K \geq \left(\frac{2\gamma}{\delta} \right)^{1/4} \left({x^K}^T s^K \right)^{5/4}, \ \textit{and} \tag{99}$$

$$\eta = \left[\frac{2\gamma}{\delta} \right]^{1/4} \left({x^K}^T s^K \right)^{(1-4\delta)/4} \leq 1. \tag{100}$$

If Assumption 6 holds then the following rule for selecting the sequence of initial step sizes can be employed:

$$\tau_k \triangleq \max \left(\frac{1}{3}, 1 - \sigma_k \left({x^k}^T s^k \right)^{\bar{\delta}} \right) \hat{\alpha}_k, \tag{101}$$

where $\bar{\delta} \in (0, 1 - 4\delta)$ and

$$\sigma_k = \begin{cases} \text{any value}, & \text{if } k < K; \\ \sigma & \text{if } k \geq K, \end{cases} \tag{102}$$

with

$$\sigma \triangleq \frac{2\gamma}{\delta \eta^4} = ({x^K}^T s^K)^{-(1-4\delta)}.$$

Superlinear convergence follows from the results of Section 4 if we can show that $\alpha_k = \tau_k$ for all k sufficiently large. That is, the relation (98) must be satisfied with $m_k = 0$ for all sufficiently large k. We have the following result:

Theorem 5.2 *Suppose that Assumptions 1–6 hold. Assume that the sequence of initial trial step sizes is chosen as described in the previous paragraph. Then Algorithm PDPR converges superlinearly with q-order at least $1+\bar{\delta}$.*

Proof. First, we observe that if $(x^k, y^k, s^k) \in \mathcal{N}(\delta, \eta, 2\gamma)$ then Lemma 4.4 and Theorem 4.5(ii) imply that the initial k-th step size satisfies $\tau_k \in J_k$. Since $\alpha_k \le \tau_k$, we also have that $\alpha_k \in J_k$. Lemma 4.3 then implies that $(x^{k+1}, y^{k+1}, s^{k+1}) \in \mathcal{N}(\delta, \eta, 2\gamma)$. Since $(x^K, y^K, s^K) \in \mathcal{N}(\delta, \eta, 2\gamma)$, the above argument implies that $(x^k, y^k, s^k) \in \mathcal{N}(\delta, \eta, 2\gamma)$, for all $k \ge K$. As mentioned in the discussion above, superlinear convergence with q-order at least $1+\bar{\delta}$ follows as a consequence of Theorem 4.5(ii) once we show that τ_k satisfies the following relation

$$\phi_q(x^k + \tau_k \Delta x^k, s^k(\tau_k)) - \phi_q(x^k, s^k) \le \mu \tau_k \nabla \phi_q(x^k, s^k)^T \left[\begin{array}{c} \Delta x^k \\ \Delta s^k \end{array} \right], \tag{103}$$

for every k sufficiently large. Indeed, from Theorem 5.1, we know that $x^{k^T} s^k \to 0$ as k tends to ∞, that is Algorithm PDPR converges globally. Therefore, we can choose $K_1 \ge K$ such that

$$\sigma \left(x^{k^T} s^k \right)^{\bar{\delta}} \le \frac{2}{3}, \qquad \forall k \ge K_1,$$

and so

$$\tau_k = \max\left(\frac{1}{3}, 1 - \sigma \left(x^{k^T} s^k \right)^{\bar{\delta}} \right) \hat{\alpha}_k = \left(1 - \sigma (x^{k^T} s^k)^{\bar{\delta}} \right) \hat{\alpha}_k, \qquad \forall k \ge K_1. \tag{104}$$

Moreover, it follows from (104) and Lemma 4.4 that

$$\begin{aligned} \left| \tau_k - \left(1 - \sigma (x^{k^T} s^k)^{\bar{\delta}} \right) \right| &= \left(1 - \sigma (x^{k^T} s^k)^{\bar{\delta}} \right) \, |\hat{\alpha}_k - 1| \\ &\le |\hat{\alpha}_k - 1| \\ &\le \frac{\gamma}{\delta \eta^4} (x^{k^T} s^k)^{1-4\delta} \\ &\le \frac{\sigma}{2} (x^{k^T} s^k)^{\bar{\delta}}. \end{aligned}$$

Hence, we obtain that

$$\left(1 - \frac{3}{2} \sigma (x^{k^T} s^k)^{\bar{\delta}} \right) \le \tau_k \le \left(1 - \frac{1}{2} \sigma (x^{k^T} s^k)^{\bar{\delta}} \right). \tag{105}$$

In the remainder of the proof, we drop the superscript and subscript k for the sake of clarity. The right hand side of (103) is easy to evaluate. Indeed, from (18a) and the fact that by (105), $\tau \le 1$, we have

$$\begin{aligned} & \mu \tau \nabla \phi_q(x, s)^T \left[\begin{array}{c} \Delta x \\ \Delta s \end{array} \right] \\ = \; & \mu \tau \left\{ \left[\frac{q}{x^T s} s - X^{-1} e \right]^T \Delta x + \left[\frac{q}{x^T s} x - S^{-1} e \right]^T \Delta s \right\} \end{aligned}$$

$$
\begin{aligned}
&= \mu\tau \left\{ \frac{q}{x^T s}(s^T \Delta x + x^T \Delta s) - \sum_{i=1}^{n} \left[\frac{\Delta x_i}{x_i} + \frac{\Delta s_i}{s_i} \right] \right\} \\
&= -\mu\tau(q-n) \\
&\geq -\mu(q-n).
\end{aligned}
$$

We now evaluate the left hand side of (103). ¿From (57) it follows that

$$
\frac{(x+\alpha\Delta x)^T s(\alpha)}{x^T s} \leq 1 - \alpha + nC_5 \frac{(x^T s)^{1-3\delta}}{\eta^3}, \qquad \forall \alpha \in [0,1]. \tag{106}
$$

Moreover, it follows from (56) that for $i = 1, \ldots, n$

$$
\begin{aligned}
\frac{(x_i+\alpha\Delta x_i)s_i(\alpha)}{x_i s_i} &\geq 1 - \alpha - C_5 \frac{(x^T s)^{2-3\delta}}{\eta^3 (x_i s_i)} \\
&\geq 1 - \alpha - C_5 \frac{(x^T s)^{1-4\delta}}{\eta^4}, \qquad \forall \alpha \in [0,1],
\end{aligned} \tag{107}
$$

where the last inequality is due to the fact that $x_i s_i \geq \eta (x^T s)^{1+\delta}$. Using (106), (107) and the definition of ϕ_q given in (96), we obtain

$$
\begin{aligned}
&\phi_q(x+\alpha\Delta x, s(\alpha)) - \phi_q(x,s) \\
&\quad = q \log \left[\frac{(x+\alpha\Delta x)^T s(\alpha)}{x^T s} \right] - \sum_{i=1}^{n} \log \left[\frac{(x_i + \alpha\Delta x_i) s_i(\alpha)}{x_i s_i} \right] \\
&\quad \leq q \log \left[1 - \alpha + nC_5 \frac{(x^T s)^{1-3\delta}}{\eta^3} \right] - \sum_{i=1}^{n} \log \left[1 - \alpha - C_5 \frac{(x^T s)^{1-4\delta}}{\eta^4} \right] \\
&\quad = (q-n)\log(1-\alpha) + q \log \left[1 + \frac{nC_5 (x^T s)^{1-3\delta}}{\eta^3 (1-\alpha)} \right] \\
&\qquad - n \log \left[1 - \frac{C_5 (x^T s)^{1-4\delta}}{\eta^4 (1-\alpha)} \right],
\end{aligned}
$$

for all $\alpha \in [0,1]$. Setting $\alpha = \tau_k$ in the above relation and using (105), we obtain

$$
\begin{aligned}
&\phi_q(x+\tau\Delta x, s(\tau)) - \phi_q(x,s) \\
&\quad \leq (q-n)\log\left[\frac{3}{2}\sigma (x^T s)^{\bar{\delta}} \right] + q \log \left[1 + \frac{nC_5 (x^T s)^{1-3\delta-\bar{\delta}}}{\eta^3 (\sigma/2)} \right] \\
&\qquad - n \log \left[1 - \frac{C_5 (x^T s)^{1-4\delta-\bar{\delta}}}{\eta^4 (\sigma/2)} \right].
\end{aligned}
$$

Consider each of the three terms in this expression. Since ${x^k}^T s^k \to 0$ and $\bar{\delta} \in (0, 1-4\delta)$, we have for k sufficiently large that

$$
(q-n)\log\left[\frac{3\sigma}{2} (x^T s)^{\bar{\delta}} \right] \;\; \leq \;\; -2(q-n)\mu,
$$

$$
\begin{aligned}
q \log\left[1+\frac{nC_5(x^Ts)^{1-3\delta-\bar{\delta}}}{\eta^3(\sigma/2)}\right] &\le \frac{q-n}{2}\mu, \\
-n \log\left[1-\frac{C_5(x^Ts)^{1-4\delta-\bar{\delta}}}{\eta^4(\sigma/2)}\right] &\le \frac{q-n}{2}\mu.
\end{aligned}
$$

Hence, for k sufficiently large,

$$
\phi_q(x^k+\tau_k\Delta x^k, s^k(\tau_k)) - \phi_q(x^k, s^k) \le -(q-n)\mu \le \mu\tau_k \nabla\phi_q(x^k,s^k)^T \left[\begin{array}{c} \Delta x^k \\ \Delta s^k \end{array}\right].
$$

■

We now address two important issues in the implementation of Algorithm PDPR, namely, the choice of γ greater than the (unknown) constant $\bar{\gamma}$, and the availability of an index K such that the conditions (99) and (100) are satisfied. We address the latter issue first. Given any $\gamma > 0$, we show that a K satisfying the two conditions of Assumption 6 is always available provided that the parameter q that defines the potential function ϕ_q lies in the range $(n, n+1/4)$.

Theorem 5.3 *Let $\gamma > 0$ and $\delta \in (0, 1/4)$ be given and assume that $q \in (n, n+1/4)$. Then there exists an integer K such that both (99) and (100) hold.*

Proof. For convenience, we define the quantity $\nu = q - n$ so that $\nu \in (0, 1/4)$. Since the potential function ϕ_q decreases at each iteration, we have

$$
\phi_q(x^k, s^k) \le \phi_q(x^0, s^0) \stackrel{\triangle}{=} M.
$$

Now, for any $i \in \{1, 2, \ldots, n\}$, we have

$$
\begin{aligned}
&(1+\nu)\log {x^k}^T s^k - \log x_i^k s_i^k \\
&\quad \le \; M - (n-1)\log {x^k}^T s^k + \sum_{j\ne i} \log x_j^k s_j^k \\
&\quad \le \; M - (n-1)\log\left({x^k}^T s^k - x_i^k s_i^k\right) + \sum_{j\ne i} \log x_j^k s_j^k \\
&\quad \le \; M - (n-1)\log(n-1) \\
&\quad \stackrel{\triangle}{=} \; \bar{M}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
\log \frac{x_i^k s_i^k}{\left({x^k}^T s^k\right)^{1+\nu}} &\ge -\bar{M} \\
\Longrightarrow \qquad x_i^k s_i^k &\ge e^{-\bar{M}} \left({x^k}^T s^k\right)^{1+\nu} \\
\Longrightarrow \qquad x_i^k s_i^k &\ge \left[\frac{2\gamma}{\delta}\right]^{1/4} \left({x^k}^T s^k\right)^{5/4} \left\{ e^{-\bar{M}} \left[\frac{\delta}{2\gamma}\right]^{1/4} \left({x^k}^T s^k\right)^{\nu-1/4} \right\}.
\end{aligned}
$$

Hence (99) will be satisfied for $k = K$ if we have

$$e^{-\bar{M}} \left[\frac{\delta}{2\gamma} \right]^{1/4} \left(x^{K^T} s^K \right)^{\nu - 1/4} \geq 1.$$

Since $x^{k^T} s^k \to 0$ by Theorem 5.1, this last condition, and hence (99), is guaranteed to hold for any sufficiently large choice of K. By increasing K if necessary, we can also ensure that (100) is also satisfied. Hence, the result follows. ∎

In the remainder of the paper, we assume that $q \in (n, n + 1/4)$.

The remaining issue is the choice of a constant γ. In order for the theory of sections 4 and 5 to hold, we require that $\gamma \geq \bar{\gamma}$, where the constant $\bar{\gamma}$, which depends on the problem data but not on δ, is not known. Our strategy is similar to one that has been suggested for the choice of penalty parameter in the nonsmooth penalty function approach for constrained optimization (see Fletcher [2]); namely, to embed Algorithm PDPR in an outer loop in which γ is successively increased until a suitable value is found. We formalize this procedure as follows:

Algorithm G-PDPR

initially: Choose fixed constants $\beta > 1$, $\bar{\delta} \in (0, 1)$, $\delta \in (0, (1 - \bar{\delta})/4)$, $\mu \in (0, 1)$ and $\gamma_0 \geq \delta$;
Let (x^0, y^0, s^0) be a strictly feasible point;

for $\ell = 0, 1, \ldots$

Phase 1:

Apply Algorithm PDPR with $\gamma = \gamma_\ell$, defining the initial steps τ_k by (101), with σ_k arbitrary;

terminate Phase 1 at the first iterate $k = K_\ell$ for which (99) and (100) hold with $K = K_\ell$.

Phase 2:

Set $\sigma = (x^{K_\ell T} s^{K_\ell})^{-(1-4\delta)}$;
Set $\eta = (2\gamma/\delta)^{1/4} (x^{K_\ell T} s^{K_\ell})^{(1-4\delta)/4}$;
Continue Algorithm PDPR with $\sigma_k \equiv \sigma$;

terminate Phase 2 if any of the following three conditions is violated by the current iterate (x^k, y^k, s^k):

$$(x^k + \tau_k \Delta x^k, y^k + \tau_k \Delta y^k, s^k(\tau_k)) \in \mathcal{N}(\delta, \eta, 2\gamma); \tag{108}$$

$$|\hat{\alpha}^k - 1| \leq \frac{\gamma}{\delta \eta^4} (x^{kT} s^k)^{1-4\delta}; \tag{109}$$

$$(x^k + \tau_k \Delta x^k)^T s^k(\tau_k) \leq \frac{31}{9} \sigma \left(x^{kT} s^k \right)^{(1+\bar{\delta})}. \tag{110}$$

Set $\gamma_{\ell+1} \leftarrow 10\gamma_\ell$ and $\ell \leftarrow \ell + 1$;

end for

The increase factor of 10 for γ can of course be replaced by any constant that exceeds 1. Note that Algorithm G-PDPR is a special case of Algorithm PDPR, in which we define the choice of initial step size in a somewhat elaborate way. Note also that the choices of $\bar{\delta}$ and δ guarantee that $\delta \in (0, 1/4)$ and that $\bar{\delta} \in (0, 1 - 4\delta)$.

The following result guarantees global and superlinear convergence of Algorithm G-PDPR.

Theorem 5.4 *Suppose that $q \in (n, n+1/4)$ and that Assumptions 1–5 hold. Then Algorithm G-PDPR converges converges globally and superlinearly with q-order at least $1 + \bar{\delta}$*

Proof. As noted above, Algorithm G-PDPR is simply a special case of Algorithm PDPR, and so global convergence follows as a consequence of Theorem 5.1.

To prove local convergence properties, we show first that the number of "outer loops" in Algorithm G-PDPR is finite. Let us assume for contradiction that this is not true. Then there must be a loop index ℓ for which $\gamma_\ell > \bar{\gamma}$. Theorem 5.3 implies that Phase 1 of loop ℓ terminates, while Lemma 4.4 and Theorem 4.5 ensure that conditions (108), (109), and (110) are *not* violated on any subsequent iteration. This means that Phase 2 of loop ℓ does not terminate, giving a contradiction.

Now consider the final loop of Algorithm G-PDPR. Since all the conditions (108), (109) and (110) are satisfied for every k sufficiently large, it is clear from (110) that superlinear convergence with Q-order at least $(1 + \bar{\delta})$ is obtained if we can show that $\alpha_k = \tau_k$, for all k sufficiently large. The proof that this holds is identical to that of Theorem 5.2. We need to observe that (105) can be obtained by using relation (109), while the relations (106) and (107) follow from Lemma 4.2 and (108).

■

6 Conclusions and discussion

We have described an algorithm for general convex programming problems that converges globally and superlinearly under fairly weak assumptions. An interesting feature of the global convergence framework is that the choice of q for the potential function does not depend on δ. Since smaller values of $q - n$ are typically (and, for some algorithms, rigorously) associated with slower convergence, the choices of q we have described in Section 5 are preferable to the more obvious choice $q = n + \delta$, which would have made for simpler analysis.

The analysis can be generalized in at least two ways. First, a centering parameter can be introduced into the step calculation (18). Equation (18a) can be replaced by

$$S^k \Delta x^k + X^k \Delta s^k = -S^k X^k e + \sigma_k ({x^k}^T s^k)/n,$$

where $\sigma_k \in [0, \sigma]$, $\sigma \in [0, 1)$. The second term on the right hand side serves to bias the step towards the central path, and is a common feature of most interior point algorithms. The global convergence analysis in Monteiro [9] still holds when σ_k is chosen in this way, while our local convergence analysis still goes through provided that σ_k approaches zero sufficiently rapidly (that is, like some power of $({x^k}^T s^k)$) during the latter stages of the algorithm.

A second possible generalization is to functions whose Hessian satisfies a Hölder continuity condition

$$\|\nabla^2 f(x^1) - \nabla^2 f(x^2)\| \leq C_x \|x^1 - x^2\|^\gamma$$

(where $\gamma \in (0, 1]$) in place of the Lipschitz continuity condition in Assumption 2. In this case, it is still possible to design an algorithm with global and superlinear convergence. The q-order for the convergence of the duality gap to zero may lie in the range $[1, 1 + \gamma)$.

Acknowledgment

Part of this work was done while the first author was visiting the Mathematics and Computer Science Division of Argonne National Laboratory. He thanks Argonne for its financial support and for the congenial scientific atmosphere that it provided.

References

[1] J. V. Burke and M. C. Ferris, *Weak sharp minima in mathematical programming*, Research Report 1050, iComputer Sciences Department, University of Wisconsin, Wisconsin, Madison, 1991. SIAM Journal on Control and Optimization, to appear.

[2] R. Fletcher, *Practical Methods of Optimization*, John Wiley & Sons, New York, second ed., 1987.

[3] J. Ji, F. Potra, and S. Huang, *A predictor-corrector method for linear complementarity problems with polynomial complexity and superlinear convergence*, Technical Report, Department of Mathematics, The University of Iowa, Iowa City, Iowa, 52240, USA, August 1991.

[4] J. Ji, F. Potra, R. A. Tapia, and Y. Zhang, *An interior–point method with polynomial complexity and superlinear convergence for linear complementarity problems*, Technical Report TR–91–23, Dept. of Mathematical Sciences, Rice University, Houston, TX 77251, USA, 1991.

[5] M. Kojima, Y. Kurita, and S. Mizuno, *Large–step interior point algorithms for linear complementarity problems*, Research Reports on Information Sciences, Ser. B : Operations Research B–243, Dept. of Information Sciences, Tokyo Institute of Technology, Oh–Okayama, Meguro–ku, Tokyo 152, Japan, 1991.

[6] M. Kojima, N. Megiddo, and T. Noma, *Homotopy continuation methods for nonlinear complementarity problems*, Mathematics of Operations Research, 16 (1991), pp. 754–774.

[7] K. A. McShane, *A superlinearly convergent* $O(\sqrt{n}L)$ *iteration primal–dual linear programming algorithm*, manuscript, U. S. Government, 2537 Villanova Drive, Vienna, Virginia, USA, 1991.

[8] S. Mehrotra, *Quadratic convergence in a primal-dual method*, Technical Report 91–15, Dept. of Industrial Engineering and Management Science, Northwestern University, Evanston, IL 60208, USA, September 1991.

[9] R. D. C. Monteiro, *A globally convergent interior point algorithm for convex programming*, Technical Report 91–021, Dept. of Systems and Industrial Engineering, University of Arizona, Tucson, AZ 85721, USA, July 1991.

[10] T. Tsuchiya, *Quadratic convergence of Iri and Imai's method for degenerate linear programming problems*, Research Memorandum 412, The Institute of Statistical Mathematics, 4–6–7 Minami–Azabu, Minato–ku, Tokyo 106, Japan, 1991.

[11] Y. Ye, *On the q-order of convergence of interior-point algorithms for linear programming*, Working Paper 91–17, The College of Business Administration, The University of Iowa, Iowa City, IA 52242, USA, 1991.

[12] Y. Ye and K. Anstreicher, *On quadratic and* $O(\sqrt{n}L)$ *convergence of a predictor-corrector algorithm for LCP*, Technical Report, Department of Management Sciences, The University of Iowa, Iowa City, IA 52242, USA, November 1991.

[13] Y. Ye, O. Güler, R. A. Tapia, and Y. Zhang, *A quadratically convergent* $O(\sqrt{n}L)$*–iteration algorithm for linear programming*, Technical Report TR–91–26, Dept. of Mathematical Sciences, Rice University, Houston, TX 77251, USA, August 1991.

[14] Y. Zhang and R. A. Tapia, *A quadratically convergent polynomial primal–dual interior–point algorithm for linear programming*, Technical Report TR 90–40, Dept. of Mathematical Sciences, Rice University, Houston, TX 77251, USA, 1990.

[15] ——, *Superlinear and quadratic convergence of primal–dual interior–point methods for linear programming revisited*, Journal of Optimization Theory and Applications, 73 (1992), pp. 229–242.

[16] Y. Zhang, R. A. Tapia, and J. E. Dennis, *On the superlinear and quadratic convergence of primal–dual interior point linear programming algorithms*, Technical Report TR 90–6, Dept. of Mathematical Sciences, Rice University, Houston, TX 77251, USA, January 1990.

[17] Y. Zhang, R. A. Tapia, and F. Potra, *On the superlinear convergence of interior point algorithms for a general class of problems*, Technical Report TR 90–9, Dept. of Mathematical Sciences, Rice University, Houston, TX 77251, USA, March 1990.