

Convergence of infeasible-interior-point methods for self-scaled conic programming

Bharath Kumar Rangarajan and Michael J. Todd*

October 10, 2003

Abstract

We present results on global and polynomial-time convergence of infeasible-interior-point methods for self-scaled conic programming, which includes linear and semidefinite programming. First, we establish global convergence for an algorithm using a wide neighborhood. Next, we prove polynomial complexity for the algorithm with a slightly narrower neighborhood. Both neighborhoods are related to the wide (minus infinity) neighborhood and are much larger than the 2-norm neighborhood. We also provide stopping rules giving an indication of infeasibility.

1 Introduction

The first polynomial-time interior-point algorithm for linear programming was presented by Karmarkar in [3]. Later, the interior-point framework was extended to the general class of conic programming problems by Nesterov and Nemirovskii in [7]. Conic programming can be defined as minimizing a linear objective subject to linear equality constraints and a cone membership constraint. Interior-point algorithms can be broadly classified as either feasible- or infeasible-interior-point methods. In both, the iterates remain in the interior of the cone. In the former, the iterates stay feasible to the linear equality constraints, while in the latter, they are not required to satisfy these equations.

Infeasible-interior-point algorithms are appealing in practice as it is not easy in most cases to find a starting point in the interior of the cone satisfying the linear constraints. In the case of linear programming, global convergence of an infeasible-interior-point method was first established by Kojima, Megiddo and Mizuno in [4]. Subsequently, polynomial iteration complexity for variants of this algorithm was established by Zhang [18], Mizuno [5] and Potra [11, 12]. Later, the results were extended to semidefinite programming (for instance, in Zhang [19]). For extension in the case of feasible-interior-point methods see [6, 9, 19].

Nesterov and Todd introduced self-scaled barriers and extended feasible-interior-point methods to self-scaled conic programs (see [8, 9]). Self-scaled cones are the class of cones that have an associated self-scaled barrier, and they include the non-negative orthant and the cone of positive semidefinite

*Both authors at School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York 14853, USA ([bharath|miketodd@orie.cornell.edu](mailto:{bharath|miketodd}@orie.cornell.edu)). Research supported in part by NSF through grant DMS-0209457 and ONR through grant N00014-02-1-0057.

matrices. Recently, feasible-interior-point algorithms were also extended to self-scaled cones using Jordan algebraic techniques (see [15]). In this article, we generalize an infeasible-interior-point method to self-scaled conic programs and provide a global convergence result using a wide neighborhood and establish polynomial convergence for the same algorithm using a slightly narrower neighborhood. Both neighborhoods are comparable to the wide (infinity) neighborhoods used in linear and semidefinite programming, and are much wider than the 2-norm neighborhoods used in short-step methods.

For infeasible problems, infeasible-interior-point methods can provide some information about the infeasibility. For general conic programming, using a homogeneous model, Nesterov et al. [10] introduced various measures of infeasibility and obtained complexity estimates for algorithms using them. For the type of algorithm we consider, in the context of linear programming, Todd and Ye [17] obtained complexity estimates for some reasonably strong indicators of infeasibility; here we extend these results.

In feasible-interior-point methods the search directions lie in orthogonal spaces, which simplifies the analysis very much. The main challenge in analyzing infeasible-interior-point methods lies in getting a handle on the search directions. The global convergence result in this paper is based on the arguments presented in Kojima et al. [4] for linear programming and the polynomial convergence argument for the algorithm in Section 4 closely follows the analysis by Zhang in [18]. Our proofs rely heavily on the self-scaled property of the barriers.

We start by introducing the preliminary concepts in Section 2. Section 3 presents an algorithm with a wide neighborhood and analyzes its global convergence. In Section 4, we restrict the method to a smaller neighborhood and obtain a polynomial iteration complexity result for it. We also provide results pertaining to indicators of infeasibility. The paper concludes with some remarks in Section 5. Some of the proofs are quite detailed and technical. The reader may prefer to omit some of the derivations at a first reading, for example those from (4.41) to (4.63).

2 Preliminaries

In this section, we describe the self-scaled conic programming problem and its optimality conditions. We provide an introduction to self-scaled barriers and some of their properties. Then we state the Newton system that will be used to define the search directions for the algorithms in later sections. The section concludes with some of the key properties that follow from the definition of the Newton system.

Let E, Y be given finite dimensional real vector spaces, and E^*, Y^* be their respective dual spaces. Let $\langle \cdot, \cdot \rangle$ denote the scalar product on $E^* \times E$ or $Y^* \times Y$. The primal and dual self-scaled conic programs are defined as follows :

$$(P) \quad \min\{\langle c, x \rangle : Ax = b, x \in K\}, \quad (2.1)$$

$$(D) \quad \max\{\langle b, y \rangle : A^*y + s = c, s \in K^*\}, \quad (2.2)$$

where $A : E \mapsto Y^*$ is a linear map, $A^* : Y \mapsto E^*$ is the adjoint linear map, $b \in Y^*$, $x \in E$ and $s, c \in E^*$. Here, $K \subset E$ is assumed to be a *regular* closed convex cone, i.e., it contains no lines and has a non-empty interior. It follows from standard convex analysis that its dual $K^* := \{s \in E^* : \langle s, x \rangle \geq 0 \text{ for all } x \in K\}$ is also a regular closed convex cone. We will denote the interiors of K and K^* by $\text{int } K$ and $\text{int } K^*$ respectively and \mathcal{Z} and $\text{int } \mathcal{Z}$ will denote $K \times Y \times K^*$ and $\text{int } K \times Y \times \text{int } K^*$

respectively.

In linear programming, $E = E^* = \Re^n$, $Y = Y^* = \Re^m$, and $K = K^* = \Re_+^n$, the non-negative orthant. In this case, A is a $m \times n$ matrix, $A^* = A^T$, $b \in \Re^m$, $c \in \Re^n$, and $\langle s, x \rangle = x^T s$ is the standard dot product for vectors $x, s \in \Re^n$ (or \Re^m). In semidefinite programming, $E = E^* = \mathcal{S}^n$, the space of symmetric matrices of order n , $Y = Y^* = \Re^m$, and $K = K^* = \mathcal{S}_+^n$, the cone of symmetric positive semidefinite matrices of order n . The scalar (inner) product is given by $\langle s, x \rangle = \text{trace}(xs)$ for $x, s \in \mathcal{S}^n$ and we have the standard dot product on $\Re^m \times \Re^m$. Here, $Ax = (\text{trace}(a_i x))_1^m \in \Re^m$, $A^*y = \sum_i y_i a_i$, $a_i \in \mathcal{S}^n$, $b \in \Re^m$, and $c \in \mathcal{S}^n$.

Let F be a strongly ν -self-concordant, logarithmically homogeneous, non-degenerate barrier for K (see Definitions 2.1.1, 2.3.1, 2.3.2 in [7]). The dual barrier functional, F_* , is defined by

$$F_*(s) := \sup\{-\langle s, x \rangle - F(x) : x \in \text{int } K\}.$$

Then F_* is also a strongly ν -self-concordant, logarithmically homogeneous, non-degenerate barrier for K^* (Theorem 2.4.4 in [7]). For self-concordant barriers, $\nu \geq 1$ by Corollary 2.3.3 in [7].

We say F is a self-scaled barrier for K if it is self-concordant and

- (i) for any $x, w \in \text{int } K$, $F''(w)x \in \text{int } K^*$, and
- (ii) for any $x, w \in \text{int } K$, $F_*(F''(w)x) = F(x) - 2F(w) - \nu$.

Let us further assume F to be a self-scaled barrier. By Proposition 3.1 in [8], F_* is a self-scaled barrier for K^* . By convention, we call K and K^* self-scaled cones, as there exist self-scaled barriers for them, and (P) and (D) are the primal and dual self-scaled conic programming problems. We recall Theorem 3.2 in [8]:

Lemma 2.1 *For any $(x, s) \in \text{int } K \times \text{int } K^*$, there exists a unique scaling point $w := w(x, s) \in \text{int } K$ such that $F''(w)x = s$. \square*

Hereafter, we denote the scaling point for the pair (x, s) by w .

For linear programming, $F(x) := -\sum_i \ln x_i$ and $F_*(s) = -\sum_i \ln s_i - n$ are self-scaled barriers, and the scaling point $w = \sqrt{\frac{x}{s}}$, where the square root and fraction are taken component-wise. For semidefinite programming, $F(x) := -\ln(\det(x))$ and $F_*(s) = -\ln(\det(s)) - n$ for $x, s \in \mathcal{S}_{++}^n$, the space of positive definite matrices. The scaling point is given by $w = x^{1/2}[x^{1/2}sx^{1/2}]^{-1/2}x^{1/2} = s^{-1/2}[s^{1/2}xs^{1/2}]^{1/2}s^{-1/2}$. Note that the matrix square root is uniquely defined for positive definite matrices. In both cases, the parameter ν of the barrier is n .

We collect here for later use some useful properties of the self-scaled barrier functionals. We then define local norms using the barriers and prove a useful property. For $\tau > 0$, $x \in \text{int } K$, $s \in \text{int } K^*$,

$$F'(\tau x) = \frac{1}{\tau}F'(x), \quad F''(\tau x) = \frac{1}{\tau^2}F''(x), \quad (2.3)$$

$$F''(x)x = -F'(x), \quad \langle -F'(x), x \rangle = \nu, \quad (2.4)$$

$$\langle F''(x)x, x \rangle = \nu, \quad \langle F'(x), F''(x)^{-1}F'(x) \rangle = \nu; \quad (2.5)$$

$$-F'(x) \in \text{int } K^*, \quad -F'_*(s) \in \text{int } K, \quad (2.6)$$

$$F(x) + F_*(-F'(x)) = -\nu, \quad F(-F'_*(s)) + F_*(s) = -\nu, \quad (2.7)$$

$$F'_*(-F'(x)) = -x, \quad F'(-F'_*(s)) = -s, \quad (2.8)$$

$$F''_*(-F'(x)) = F''(x)^{-1}, \quad F''(-F'_*(s)) = F''(s)^{-1}, \quad (2.9)$$

$$F(x) + F_*(s) \geq -\nu \ln \langle s, x \rangle + \nu \ln \nu - \nu. \quad (2.10)$$

Statements (2.3)-(2.5) follow from the logarithmic homogeneity of F (see Proposition 2.3.4 in [7]) and similar statements hold for F_* . Relations (2.7)-(2.10) follow from the definition of the dual barrier functional (see Theorem 2.4.2 in [7], Theorem 3.3.5 in [14]). For self-scaled barriers, by Theorem 3.1 in [8], we have

$$F''(w)x = s, \quad F''(w)F'_*(s) = F'(x), \quad \text{and} \quad F''(w)F''_*(s)F''(w) = F''(x). \quad (2.11)$$

We define the following norms induced by the barriers F and F_* : for $p \in E$, $x \in \text{int } K$, $q \in E^*$, $s \in \text{int } K^*$ and a fixed reference element $e \in \text{int } K$,

$$\begin{aligned} \|p\|_x &:= \langle F''(x)p, p \rangle^{1/2}, & \|q\|_x^* &:= \langle q, F''(x)^{-1}q \rangle^{1/2}, \\ \|q\|_s &:= \langle q, F''_*(s)q \rangle^{1/2}, & \|p\|_s^* &:= \langle F''_*(s)^{-1}p, p \rangle^{1/2}, \\ \|p\| &:= \|p\|_e, \quad \text{and} & \|q\|_* &:= \|q\|_e^*. \end{aligned}$$

Let $\|\cdot\|$ and $\|\cdot\|_*$ be dual norms on Y and Y^* . The following is an important inequality on norms:

Lemma 2.2 For $p \in K$, $x \in \text{int } K$, $q \in K^*$, $s \in \text{int } K^*$,

$$\langle s, p \rangle \geq \|p\|_s^* \quad \text{and} \quad \langle q, x \rangle \geq \|q\|_x^*.$$

Proof : From Theorem 2.1.1 (ii) in [7], $s - \frac{F''_*(s)^{-1}p}{\|F''_*(s)^{-1}p\|_s} \in K^*$ so that, $\left\langle s - \frac{F''_*(s)^{-1}p}{\|F''_*(s)^{-1}p\|_s}, p \right\rangle \geq 0$. Expanding, we get

$$\langle s, p \rangle \geq \frac{\langle F''_*(s)^{-1}p, p \rangle}{\langle F''_*(s)^{-1}p, F''_*(s)F''_*(s)^{-1}p \rangle^{1/2}} = \frac{\langle F''_*(s)^{-1}p, p \rangle}{\langle F''_*(s)^{-1}p, p \rangle^{1/2}} = \|p\|_s^*.$$

The proof of the other part proceeds similarly. □

We will make the following assumptions for the rest of the paper :

A1 A is a surjective linear map, and

A2 (P) and (D) have strictly feasible solutions, i.e., feasible solutions in the interior of K and K^* respectively.

For x feasible in (P) and (y, s) feasible in (D) , we have the following weak duality result:

$$\langle c, x \rangle - \langle b, y \rangle = \langle A^*y + s, x \rangle - \langle Ax, y \rangle = \langle s, x \rangle \geq 0.$$

Hence $\langle s, x \rangle = 0$ is sufficient for optimality. But, given our assumption of strict feasibility, it is also necessary by Theorem 4.2.1 in [7]. So, the optimality conditions for (P) and (D) are:

$$\begin{aligned} A^*y + s &= c \\ Ax &= b, \\ \langle s, x \rangle &= 0, \\ x \in K, \quad s &\in K^*. \end{aligned} \tag{2.12}$$

We will now describe the central path and the associated Newton system. We define the central path to be the set of solutions to the following system for all $\mu > 0$:

$$\begin{aligned} A^*y + s &= c \\ Ax &= b, \\ \mu F'(x) + s &= 0, \\ x \in \text{int } K, \quad s &\in \text{int } K^*. \end{aligned} \tag{2.13}$$

Given our assumptions, it is known (see Chapter 2 in [7]) that the set of equations (2.13) has a unique solution $(x(\mu), y(\mu), s(\mu))$ for each $\mu > 0$ and that $(x(\mu), y(\mu), s(\mu))$ converges to (x^*, y^*, s^*) , an optimal solution of (P) and (D), as $\mu \downarrow 0$. Following Nesterov and Todd [8], we propose the following Newton system at a given $(x, y, s) \in \text{int } \mathcal{Z}$:

Newton Equations

$$\begin{aligned} A^*\Delta y + \Delta s &= c - A^*y - s, \\ A\Delta x &= b - Ax, \\ F''(w)\Delta x + \Delta s &= h := -\beta_1\mu F'(x) - s, \\ \Delta x \in E, \quad \Delta y \in Y, \quad \Delta s \in E^*, \end{aligned} \tag{2.14}$$

where $\mu = \frac{\langle s, x \rangle}{\nu}$ and $\beta_1 \in [0, 1]$ is a given parameter.

In linear programming, the third equation of (2.14) is $S^{-1}X\Delta x + \Delta s = \beta_1\mu X^{-1}\mathbf{1} - s$, where $S = \text{diag}(s)$, $X = \text{diag}(x)$ and $\mathbf{1}$ is the vector of all ones. For semidefinite programming, Toh et. al. [16] showed that the third Newton equation in (2.14) is equivalent to $H_p(x\Delta s + \Delta xs - xs) = \sigma\mu i$, where i is the $n \times n$ identity matrix, p is any nonsingular matrix such that $p^T p = w^{-1}$, and H_p is the symmetrization operator given by

$$H_p(a) = \frac{pap^{-1} + (pap^{-1})^T}{2}.$$

Lemma 2.3 *The following hold as a consequence of the Newton equations (2.14), where $\alpha > 0$ is such that $x + \alpha\Delta x \in \text{int } K$ and $s + \alpha\Delta s \in \text{int } K^*$.*

$$(a) \quad \|\Delta x\|_w^2 + \|\Delta s\|_w^{*2} + 2\langle \Delta s, \Delta x \rangle = \|h\|_w^{*2}; \quad (2.15)$$

$$(b) \quad \langle \Delta s, x \rangle + \langle s, \Delta x \rangle = -(1 - \beta_1)\langle s, x \rangle; \quad (2.16)$$

$$(c) \quad \langle F'(x), \Delta x \rangle + \langle \Delta s, F'_*(s) \rangle = \nu - \beta_1\mu \langle F'(x), F'_*(s) \rangle; \quad (2.17)$$

$$(d) \quad \frac{\langle s + \alpha\Delta s, x + \alpha\Delta x \rangle}{\nu} = \frac{\langle s, x \rangle}{\nu}(1 - \alpha + \beta_1\alpha) + \alpha^2 \frac{\langle \Delta s, \Delta x \rangle}{\nu}; \quad (2.18)$$

$$(e) \quad \ln \left(\frac{\langle s + \alpha\Delta s, x + \alpha\Delta x \rangle}{\nu} \right) \leq \ln \left(\frac{\langle s, x \rangle}{\nu} \right) - \alpha(1 - \beta_1) + \alpha^2 \frac{\langle \Delta s, \Delta x \rangle}{\langle s, x \rangle}; \quad (2.19)$$

$$(f) \quad A(x + \alpha\Delta x) - b = (1 - \alpha)(Ax - b); \quad (2.20)$$

$$A^*(y + \alpha\Delta y) + (s + \alpha\Delta s) - c = (1 - \alpha)(A^*y + s - c).$$

Proof : The first equation is gotten by expanding the third equation of (2.14):

$$\begin{aligned} \|h\|_w^{*2} &= \langle F''(w)\Delta x + \Delta s, F''(w)^{-1}(F''(w)\Delta x + \Delta s) \rangle \\ &= \langle F''(w)\Delta x + \Delta s, \Delta x + F''(w)^{-1}\Delta s \rangle \\ &= \langle F''(w)\Delta x, \Delta x \rangle + \langle \Delta s, F''(w)^{-1}\Delta s \rangle + 2\langle \Delta s, \Delta x \rangle \\ &= \|\Delta x\|_w^2 + \|\Delta s\|_w^{*2} + 2\langle \Delta s, \Delta x \rangle. \end{aligned}$$

The second equation follows from

$$\begin{aligned} \langle \Delta s, x \rangle + \langle s, \Delta x \rangle &= \langle \Delta s, x \rangle + \langle F''(w)x, \Delta x \rangle \\ &= \langle \Delta s + F''(w)\Delta x, x \rangle \\ &= \langle -\beta_1\mu F'(x) - s, x \rangle \quad (\text{from (2.14)}) \\ &= -\langle s, x \rangle + \beta_1\mu \langle -F'(x), x \rangle \\ &= -(1 - \beta_1)\langle s, x \rangle \quad (\text{from (2.4)}). \end{aligned}$$

We obtain (2.17) from

$$\begin{aligned} \langle F'(x), \Delta x \rangle + \langle \Delta s, F'_*(s) \rangle &= \langle F''(w)F'_*(s), \Delta x \rangle + \langle \Delta s, F'_*(s) \rangle \quad (\text{from (2.11)}) \\ &= \langle F''(w)\Delta x + \Delta s, F'_*(s) \rangle \\ &= \langle -\beta_1\mu F'(x) - s, F'_*(s) \rangle \quad (\text{from (2.14)}) \\ &= \nu - \beta_1\mu \langle F'(x), F'_*(s) \rangle. \end{aligned}$$

The equation (2.18) follows from

$$\begin{aligned} \frac{\langle s + \alpha\Delta s, x + \alpha\Delta x \rangle}{\nu} &= \frac{\langle s, x \rangle}{\nu} + \alpha \frac{\langle \Delta s, x \rangle + \langle s, \Delta x \rangle}{\nu} + \alpha^2 \frac{\langle \Delta s, \Delta x \rangle}{\nu} \\ &= \frac{\langle s, x \rangle}{\nu} + \alpha \frac{-(1 - \beta_1)\langle s, x \rangle}{\nu} + \alpha^2 \frac{\langle \Delta s, \Delta x \rangle}{\nu} \quad (\text{from (2.16)}) \\ &= \frac{\langle s, x \rangle}{\nu}(1 - \alpha + \beta_1\alpha) + \alpha^2 \frac{\langle \Delta s, \Delta x \rangle}{\nu}. \end{aligned}$$

Using the result above, we can see that (as $\ln(1 + \xi) \leq \xi$ for $\xi > -1$)

$$\begin{aligned} \ln \left(\frac{\langle s + \alpha\Delta s, x + \alpha\Delta x \rangle}{\nu} \right) &= \ln \left(\frac{\langle s, x \rangle}{\nu} \right) + \ln \left(1 - \alpha(1 - \beta_1) + \alpha^2 \frac{\langle \Delta s, \Delta x \rangle}{\langle s, x \rangle} \right) \\ &\leq \ln \left(\frac{\langle s, x \rangle}{\nu} \right) - \alpha(1 - \beta_1) + \alpha^2 \frac{\langle \Delta s, \Delta x \rangle}{\langle s, x \rangle}. \end{aligned}$$

Note that the condition inside the parentheses will apply if all logarithms are defined, i.e., as long as $\langle s + \alpha \Delta s, x + \alpha \Delta x \rangle > 0$, which is ensured by our assumption on α . Finally, note that (2.20) follows directly from the first two equations of (2.14). \square

3 Global convergence in a wide neighborhood of the central path

The algorithm produces a sequence of iterates $\{(x_k, y_k, s_k)\} \subset \text{int } \mathcal{Z}$ until a termination criterion is met. At each iterate, a step direction $(\Delta x_k, \Delta y_k, \Delta s_k)$ is computed from the Newton equations (2.14), and step lengths $\alpha_p^k, \alpha_d^k \in [0, 1]$ are chosen. The next iterate is given by $(x_{k+1}, y_{k+1}, s_{k+1}) := (x_k + \alpha_p^k \Delta x_k, y_k + \alpha_d^k \Delta y_k, s_k + \alpha_d^k \Delta s_k)$. Before we describe the algorithms, we need some definitions from [9, 2]. For $x \in \text{int } K$, $s \in \text{int } K^*$, let

$$\gamma_F(x, s) := F(x) + F_*(s) + \nu \ln \left(\frac{\langle s, x \rangle}{\nu} \right) + \nu, \quad (3.1)$$

$$\begin{aligned} \gamma_G(x, s) &:= \mu \langle F'(x), F'_*(s) \rangle - \nu, & (3.2) \\ \phi_p^k &:= \prod_{i=0}^{k-1} (1 - \alpha_p^i), \text{ where } \phi_p^0 := 1, \text{ and} \\ \phi_d^k &:= \prod_{i=0}^{k-1} (1 - \alpha_d^i), \text{ where } \phi_d^0 := 1. \end{aligned}$$

The neighborhoods used in the algorithms will be defined using γ_F and γ_G . By applying (2.20) in Lemma 2.3 inductively for α_p^k (and similarly for α_d^k), it is easy to see [2, 4] that ϕ_p^k (ϕ_d^k) represents the proportion of the initial infeasibility remaining in the primal (dual) after k iterations. This we summarize as

$$Ax_k - b = \phi_p^k (Ax_0 - b), \text{ and } A^*y_k + s_k - c = \phi_d^k (A^*y_0 + s_0 - c). \quad (3.3)$$

Let

$$\mathcal{N}_F(\theta_F) := \{(x, y, s) \in \text{int } \mathcal{Z} : \gamma_F(x, s) \leq \theta_F\}.$$

In the linear programming instance, this neighborhood is

$$\left\{ (x, y, s) \in \mathfrak{R}_+^n \times \mathfrak{R}^m \times \mathfrak{R}_+^n : n \ln \left(\frac{x^T s}{n} \frac{1}{\prod_i (x_i s_i)^{1/n}} \right) \leq \theta_F \right\}.$$

Thus the arithmetic mean of the $x_i s_i$'s cannot exceed their geometric mean by more than a factor of $\exp(\theta_F/n)$. For semidefinite programming, if $(\lambda_i)_{i=1}^n$ denotes the spectrum of $x^{1/2} s x^{1/2}$, we obtain the following description for the neighborhood:

$$\left\{ (x, y, s) \in S_+^n \times \mathfrak{R}^m \times S_+^n : n \ln \left(\frac{\sum_i \lambda_i}{n} \frac{1}{\prod_i \lambda_i^{1/n}} \right) \leq \theta_F \right\},$$

with a similar interpretation in terms of the arithmetic and geometric means of the λ_i 's.

Now, we state

Algorithm 1 :

- 1 Let $1 > \beta_2 > \beta_1 > 0$, $\epsilon^* > 0$, $\Omega^* > 0$, $\theta_F > 0$, $x_0 \in \text{int } K$, $y_0 \in Y$ and $s_0 \in \text{int } K^*$ be given such that $(x_0, y_0, s_0) \in \mathcal{N}_F(\theta_F)$. Set $k = 0$.
- 2 Solve for $(\Delta x_k, \Delta y_k, \Delta s_k)$ from the Newton equations (2.14) at (x_k, y_k, s_k) .
- 3 Let $(x(\alpha), y(\alpha), s(\alpha)) := (x_k, y_k, s_k) + \alpha(\Delta x_k, \Delta y_k, \Delta s_k)$. Compute the largest step length $\bar{\alpha}_k \in (0, 1]$ such that for all $\alpha \in [0, \bar{\alpha}_k]$, $(x(\alpha), y(\alpha), s(\alpha)) \in \mathcal{N}_F(\theta_F)$, $\langle s(\alpha), x(\alpha) \rangle \geq \max(\phi_p^k, \phi_d^k)(1 - \alpha) \langle s_0, x_0 \rangle$, and $\langle s(\alpha), x(\alpha) \rangle \leq \langle s_k, x_k \rangle (1 - (1 - \beta_2)\alpha)$.
- 4 Choose a primal step length α_p^k and a dual step length α_d^k such that

$$\begin{aligned} (x_{k+1}, y_{k+1}, s_{k+1}) &:= (x_k + \alpha_p^k \Delta x_k, y_k + \alpha_d^k \Delta y_k, s_k + \alpha_d^k \Delta s_k) \in \mathcal{N}_F(\theta_F), \\ \langle s_{k+1}, x_{k+1} \rangle &\geq \max(\phi_p^k(1 - \alpha_p^k), \phi_d^k(1 - \alpha_d^k)) \langle s_0, x_0 \rangle \text{ and} \\ \langle s_{k+1}, x_{k+1} \rangle &\leq \langle s_k, x_k \rangle (1 - (1 - \beta_2)\bar{\alpha}_k). \end{aligned}$$

- 5 Increase k by 1. If $\langle s_k, x_k \rangle < \epsilon^* \langle s_0, x_0 \rangle$ or $\|x_k\| + \|s_k\|_* > \Omega^*$, then STOP. Otherwise, repeat step 2.

We would like to note that, if we choose $\alpha_p^k = \alpha_d^k = \bar{\alpha}_k$, all the conditions in Step 4 are satisfied. However, we are free to choose different step lengths as long as a comparable decrease in the complementarity is obtained, the iterate remains in the required neighborhood, and we maintain the condition

$$\langle s_k, x_k \rangle \geq \max(\phi_p^k, \phi_d^k) \langle s_0, x_0 \rangle. \quad (3.4)$$

This requirement ensures that when total complementarity, $\langle s, x \rangle$, approaches zero, the infeasibilities, from the interpretations of ϕ_p^k and ϕ_d^k , also approach zero. For simplicity, we will often write x, y, s and $\bar{\phi}$ for x_k, y_k, s_k and $\max(\phi_p^k, \phi_d^k)$ respectively. We also write γ_F for $\gamma_F(x, s)$ and γ_G for $\gamma_G(x, s)$. The arguments should be clear from the context.

Towards proving global convergence, we will assume that at the k -th iterate

$$\langle s_k, x_k \rangle \geq \epsilon^* \langle s_0, x_0 \rangle \text{ and } \|x_k\| + \|s_k\|_* \leq \Omega^*, \quad (3.5)$$

and show that there exists an $\alpha^* > 0$ (independent of k) such that all conditions in Step 3 of the algorithm are satisfied for all $\alpha \in [0, \alpha^*]$. This gives a lower bound of α^* on $\bar{\alpha}_k$, which by Step 3 of the algorithm, implies that $\langle s_k, x_k \rangle \leq (1 - \alpha^*(1 - \beta_2))^k \langle s_0, x_0 \rangle$. Hence, if $\|x_k\| + \|s_k\|_* \leq \Omega^*$ for all k , then the total complementarity goes to zero linearly and the infeasibility goes to zero at least linearly by condition (3.4). This will complete the argument for the global convergence result. At the end of Section 4, some conclusions are presented when $\|x_k\| + \|s_k\|_* > \Omega^*$ occurs. As a first step, we establish results that allow us to bound some key terms involving Δx and Δs .

Proposition 3.1 *The scaling point $w := w(x, s)$ (given by Lemma 2.1) is a continuous function of x and s .*

Proof : Let $\psi_{x,s}(v) := \langle s, v \rangle - \langle F'(v), x \rangle$ for $x, v \in \text{int } K$ and $s \in \text{int } K^*$. From the proof of Theorem 3.2 in [8] we know that $w = w(x, s)$ is the unique minimizer of $\psi_{x,s}(\cdot)$, satisfying the optimality condition $F''(w)x = s$, and that $\psi_{x,s}(v) \geq \langle s, v \rangle + F(v) - F(x) + \nu$.

Let $\{(x_n, s_n)\} \subset \text{int } K \times \text{int } K^*$ be such that $x_n \rightarrow \bar{x} \in \text{int } K$ and $s_n \rightarrow \bar{s} \in \text{int } K^*$. Let $\{w_n = w(x_n, s_n)\} \subset \text{int } K$ and $\bar{w} = w(\bar{x}, \bar{s}) \in \text{int } K$ be the scaling points given by Lemma 2.1. To establish continuity we need to show that $w_n \rightarrow \bar{w}$. By continuity of ψ in x, s we note that $\psi_{x_n, s_n}(\bar{w}) \leq \psi_{\bar{x}, \bar{s}}(\bar{w}) + 1$ for all sufficiently large n . Also, $\psi_{x_n, s_n}(w_n) \leq \psi_{x_n, s_n}(\bar{w})$ for all n .

We claim that $F(w_n) \leq M < \infty$ for some constant M . Then, from Proposition 2.1.1 in [7], as F is a strongly non-degenerate barrier, it follows that w_n is contained in $\{w \in E : F(w) \leq M\} \subset \text{int } K$ which is closed in E . To prove the claim, note that for all sufficiently large n , $F(x_n) \leq F(\bar{x}) + 1$ and hence,

$$\psi_{x_n, s_n}(w_n) \geq \langle s_n, w_n \rangle + F(w_n) - F(x_n) + \nu \geq F(w_n) - F(\bar{x}) - 1 + \nu.$$

This, along with the chain of inequalities $\psi_{x_n, s_n}(w_n) \leq \psi_{x_n, s_n}(\bar{w}) \leq \psi_{\bar{x}, \bar{s}}(\bar{w}) + 1$ for all sufficiently large n , shows that, for such n , $F(w_n) \leq \hat{M} = F(\bar{x}) + \psi_{\bar{x}, \bar{s}}(\bar{w}) + 2 - \nu$. The claim follows.

As $s_n \rightarrow \bar{s}$, for $\epsilon = \frac{1}{2} \min_{\{x \in K : \|x\|=1\}} \langle \bar{s}, x \rangle > 0$ and all n sufficiently large,

$$\langle s_n, x \rangle = \langle \bar{s}, x \rangle - \langle \bar{s} - s_n, x \rangle \geq \langle \bar{s}, x \rangle - \|x\| \|\bar{s} - s_n\|_* = \langle \bar{s}, x \rangle - \|\bar{s} - s_n\|_* \geq \langle \bar{s}, x \rangle - \epsilon \geq \epsilon$$

for every $x \in K$ such that $\|x\| = 1$. Now, for all sufficiently large n ,

$$\psi_{x_n, s_n}(w_n) = \langle s_n, w_n \rangle + \langle -F'(w_n), x_n \rangle \geq \langle s_n, w_n \rangle = \|w_n\| \left\langle s_n, \frac{w_n}{\|w_n\|} \right\rangle \geq \epsilon \|w_n\|.$$

Then $\psi_{x_n, s_n}(w_n) \leq \psi_{\bar{x}, \bar{s}}(\bar{w}) + 1$ (for all sufficiently large n) implies that $\{w_n\}$ is bounded. Hence, $\{w_n\}$ is contained in some subset of $\text{int } K$ that is compact in E .

Let \hat{w} be any limit point of $\{w_n\}$. Since $F''(w_n)x_n = s_n$ for all n , we have $F''(\hat{w})\bar{x} = \bar{s}$. As \bar{w} is the unique solution to $F''(w)\bar{x} = \bar{s}$, $\hat{w} = \bar{w}$. Therefore, \bar{w} is the the unique limit point of $\{w_n\}$. Thus w is a continuous function of (x, s) on $\text{int } K \times \text{int } K^*$. \square

Lemma 3.2 *The region*

$$R := \left\{ \begin{array}{l} (x, y, s) \in \text{int } \mathcal{Z} : \langle s, x \rangle \geq \epsilon^* \langle s_0, x_0 \rangle, \|x\| + \|s\|_* \leq \Omega^*, \gamma_F(x, s) \leq \theta_F, \\ \text{and } \exists \phi_p, \phi_d \in [0, 1] \text{ s.t. } Ax = b + \phi_p(Ax_0 - b), A^*y + s = c + \phi_d(A^*y_0 + s_0 - c) \end{array} \right\} \text{ is compact.}$$

Proof : Using the expression for γ_F in (3.1), the fact that $\langle s, x \rangle \geq \epsilon^* \langle s_0, x_0 \rangle$ and $\gamma_F(x, s) \leq \theta_F$ we get that $F(x) + F_*(s) \leq M_1$ for some constant M_1 . By Proposition 2.1.1 in [7]

$$\hat{R} := \{(x, y, s) \in \text{int } \mathcal{Z} : F(x) + F_*(s) \leq M_1\}$$

is closed in $E \times Y \times E^*$. Hence replacing $\text{int } \mathcal{Z}$ by \hat{R} in the definition of R we see that R is closed.

If A is surjective then the solution to the equation $A^*y = r$ is unique and $\|y\| \leq \|(AA^*)^{-1}\| \|A\| \|r\|_*$ with the standard operator norm on operators. Using $\|x\| + \|s\|_* \leq \Omega^*$, the remark above and the inequality $\|c + \phi_d(A^*y_0 + s_0 - c)\|_* \leq \|c\|_* + \|c - A^*y_0 - s_0\|_*$ we get $\|y\| \leq M_2$ for some constant M_2 , or $\|x\| + \|y\| + \|s\|_* \leq \Omega^* + M_2$. Hence, R is compact. \square

Note that the operator defining the Newton equations (2.14) is continuous in w and invertible for all $w \in \text{int } K$. By Lemma 3.1, w is continuous in (x, y, s) . Also note that all the iterates defined by the algorithm lie in the region R as $\phi_p^k, \phi_d^k \in [0, 1]$. We established compactness of R in Lemma 3.2. Therefore the operator is continuous and invertible over the compact set R . Hence,

Proposition 3.3 $(\Delta x, \Delta y, \Delta s)$ is a bounded continuous function of (x, y, s) on R . In particular, the sequence $\{(\Delta x_k, \Delta s_k)\}$ produced by the algorithm is uniformly bounded. \square

As a consequence of Proposition 3.3, we can choose η sufficiently large that, for all $(x, y, s) \in R$,

$$\|\Delta x\|_x^2 + \|\Delta s\|_s^2 \leq \eta/2 \text{ and } |\langle \Delta s, \Delta x \rangle| \leq \frac{\eta}{2\nu} \epsilon^* \langle s_0, x_0 \rangle. \quad (3.6)$$

Using (3.5) and (3.6) we get as a consequence

$$\frac{|\langle \Delta s, \Delta x \rangle|}{\langle s, x \rangle} \leq \eta/2\nu. \quad (3.7)$$

The reason for this choice of η will become clear later. Next, we recall the definition of the Minkowski functional and the associated norm from Section 4 of [8]. For $p \in E$, $x \in \text{int } K$, $q \in E^*$, and $s \in \text{int } K^*$,

$$\begin{aligned} \sigma_x(p) &:= \frac{1}{\sup\{\alpha \geq 0 : x - \alpha p \in K\}} = \min\{\beta \geq 0 : \beta x - p \in K\}, \text{ and} \\ |p|_x &:= \max(\sigma_x(p), \sigma_x(-p)); \end{aligned}$$

$\sigma_s(q)$ and $|q|_s$ are similarly defined, and we set

$$\sigma_x(q) := \sigma_{-F'(x)}(q), \text{ and } \sigma_s(p) := \sigma_{-F'_*(s)}(p). \quad (3.8)$$

Using (2.9) we derive the following identities: for $\hat{s} = -F'(x)$ and $\hat{x} = -F'_*(s)$

$$\|p\|_x = \langle F''(x)p, p \rangle^{1/2} = \langle F''_*(\hat{s})^{-1}p, p \rangle^{1/2} = \|p\|_{\hat{s}}^*, \text{ and} \quad (3.9)$$

$$\|q\|_s = \langle q, F''_*(s)q \rangle^{1/2} = \langle q, F''_*(\hat{x})^{-1}q \rangle^{1/2} = \|q\|_{\hat{x}}^*. \quad (3.10)$$

Let $\bar{\kappa} := \max(\|\Delta x\|_x, \|\Delta s\|_s)$. By Proposition 3.5 of [9], $\bar{\kappa} \geq \max(|\Delta x|_x, |\Delta s|_s)$. Let us define $x(\alpha) := x + \alpha\Delta x$ and $s(\alpha) := s + \alpha\Delta s$. Then we have the following result (this is a slight weakening of Theorem 4.2 in [8], but more suited for our analysis):

Lemma 3.4 For all $\alpha \in [0, 1/(2\bar{\kappa})]$

$$\begin{aligned} F(x(\alpha)) &\leq F(x) + \alpha \langle F'(x), \Delta x \rangle + \alpha^2 \|\Delta x\|_x^2 \text{ and} \\ F_*(s(\alpha)) &\leq F_*(s) + \alpha \langle \Delta s, F'_*(s) \rangle + \alpha^2 \|\Delta s\|_s^2. \end{aligned}$$

Proof : Note that $\sigma_x(-\Delta x) \leq |\Delta x|_x \leq \bar{\kappa}$. Let $x(\beta) = x + \frac{\Delta x}{\bar{\kappa}}\beta$, with $\beta \in [0, \frac{1}{2}]$. Let $\theta(\beta) := F(x(\beta))$. From the proof of Theorem 4.2 in [8], we have

$$\begin{aligned}
\theta(\beta) - \theta(0) &\leq \theta'(0)\beta + \theta''(0) \int_0^\beta \int_0^\lambda \frac{d\tau d\lambda}{(1-\tau)^2} \\
&= \theta'(0)\beta + \theta''(0) \int_0^\beta \frac{\lambda d\lambda}{1-\lambda} \\
&\leq \theta'(0)\beta + \theta''(0) \int_0^\beta 2\lambda d\lambda \quad \left(\text{as } \lambda \leq \beta \leq \frac{1}{2} \right) \\
&= \theta'(0)\beta + \theta''(0)\beta^2.
\end{aligned}$$

The first part of the lemma now follows by substituting the appropriate expressions. The second part can be proven using a similar argument for F_* . \square

Note that $\bar{\kappa}^2 \leq \|\Delta x\|_x^2 + \|\Delta s\|_s^2 \leq \eta/2$. Therefore, $\sqrt{1/(2\eta)} \leq 1/(2\bar{\kappa})$. If we define

$$\bar{\alpha}_1 := \sqrt{\frac{1}{2\eta}}, \quad (3.11)$$

then the conclusion of Lemma 3.4 and (2.19) in Lemma 2.3 hold for all $\alpha \in [0, \bar{\alpha}_1]$.

Now, using γ_F to denote $\gamma_F(x, s)$, it follows from Lemma 3.4, Lemma 2.3, and (3.2) that for $\alpha \in [0, \bar{\alpha}_1]$,

$$\begin{aligned}
\gamma_F(x(\alpha), s(\alpha)) &= F(x(\alpha)) + F_*(s(\alpha)) + \nu \ln \left(\frac{\langle s(\alpha), x(\alpha) \rangle}{\nu} \right) + \nu \\
&\leq F(x) + \alpha \langle F'(x), \Delta x \rangle + \alpha^2 \|\Delta x\|_x^2 + F_*(s) + \alpha \langle \Delta s, F'_*(s) \rangle + \alpha^2 \|\Delta s\|_s^2 \\
&\quad + \nu \left\{ \ln \left(\frac{\langle s, x \rangle}{\nu} \right) - \alpha(1 - \beta_1) + \alpha^2 \frac{\langle \Delta s, \Delta x \rangle}{\langle s, x \rangle} \right\} + \nu \\
&= \gamma_F + \alpha (\langle F'(x), \Delta x \rangle + \langle \Delta s, F'_*(s) \rangle) - \alpha\nu(1 - \beta_1) + \alpha^2 \nu \frac{\langle \Delta s, \Delta x \rangle}{\langle s, x \rangle} \\
&\quad + \alpha^2 (\|\Delta x\|_x^2 + \|\Delta s\|_s^2) \\
&= \gamma_F + \alpha (\nu - \beta_1 \mu \langle F'(x), F'_*(s) \rangle) - \alpha\nu(1 - \beta_1) + \alpha^2 \nu \frac{\langle \Delta s, \Delta x \rangle}{\langle s, x \rangle} \\
&\quad + \alpha^2 (\|\Delta x\|_x^2 + \|\Delta s\|_s^2) \\
&= \gamma_F + \alpha\beta_1 (\nu - \mu \langle F'(x), F'_*(s) \rangle) + \alpha^2 \left(\nu \frac{\langle \Delta s, \Delta x \rangle}{\langle s, x \rangle} + \|\Delta x\|_x^2 + \|\Delta s\|_s^2 \right) \\
&= \gamma_F - \alpha\beta_1 \gamma_G + \alpha^2 \left(\nu \frac{\langle \Delta s, \Delta x \rangle}{\langle s, x \rangle} + \|\Delta x\|_x^2 + \|\Delta s\|_s^2 \right) \\
&\leq \gamma_F - \alpha\beta_1 \gamma_F + \alpha^2 (\eta/2 + \eta/2) = \gamma_F - \alpha\beta_1 \gamma_F + \alpha^2 \eta.
\end{aligned}$$

The last inequality follows from $\gamma_F \leq \gamma_G$ ((4.17) in Theorem 4.2 of [9]), (3.6), and (3.7). To ensure that $\gamma_F(x(\alpha), s(\alpha)) \leq \theta_F$, it suffices to have

$$\gamma_F - \alpha\beta_1 \gamma_F + \alpha^2 \eta \leq \theta_F.$$

By considering the larger root of the quadratic $(\gamma_F - \theta_F) - \alpha\beta_1\gamma_F + \alpha^2\eta$, we see that this is guaranteed for $0 \leq \alpha \leq \bar{\alpha}_2$ with

$$\bar{\alpha}_2 = \frac{1}{2\eta} \min_{0 \leq \gamma_F \leq \theta_F} \{ \beta_1 \gamma_F + \sqrt{\beta_1^2 \gamma_F^2 + 4\eta(\theta_F - \gamma_F)} \}.$$

Setting the derivative with respect to γ_F of the minimand to zero, we obtain $\theta_F = \eta/\beta_1^2$. This implies that the function is monotone, and so it achieves its minimum at one of the extreme points. Hence

$$\bar{\alpha}_2 = \frac{1}{\eta} \min(\sqrt{\eta\theta_F}, \beta_1\theta_F). \quad (3.12)$$

So, we have $(x(\alpha), y(\alpha), s(\alpha)) \in \mathcal{N}_F$ for all $\alpha \in [0, \bar{\alpha}_2]$.

Next, we will establish a similar guarantee on step lengths for condition (3.4) at iteration $k+1$. Let $h(\alpha) := \langle s(\alpha), x(\alpha) \rangle - \bar{\phi}(1-\alpha) \langle s_0, x_0 \rangle$. Recall that $\bar{\phi} = \max(\phi_p^k, \phi_d^k)$ so that, if we took equal step lengths of α in both primal and dual at the k th iteration, the resulting ϕ^{k+1} 's would both be at most $\bar{\phi}(1-\alpha)$. We would like an $\bar{\alpha}_3 > 0$ such that $h(\alpha) \geq 0$ for all $\alpha \in [0, \bar{\alpha}_3]$. Since (3.4) holds at (x, y, s) , using (2.18) in Lemma 2.3 and (3.7) we get

$$\begin{aligned} h(\alpha) &= \langle s(\alpha), x(\alpha) \rangle - \bar{\phi}(1-\alpha) \langle s_0, x_0 \rangle \\ &= \langle s, x \rangle (1 - \alpha(1 - \beta_1)) + \alpha^2 \langle \Delta s, \Delta x \rangle - \bar{\phi}(1-\alpha) \langle s_0, x_0 \rangle \\ &= (1-\alpha)(\langle s, x \rangle - \bar{\phi} \langle s_0, x_0 \rangle) + \alpha \langle s, x \rangle \left(\beta_1 + \alpha \frac{\langle \Delta s, \Delta x \rangle}{\langle s, x \rangle} \right) \\ &\geq \alpha \langle s, x \rangle \left(\beta_1 - \alpha \frac{\eta}{2\nu} \right). \end{aligned}$$

Hence the choice of

$$\bar{\alpha}_3 := \frac{2\beta_1\nu}{\eta} \quad (3.13)$$

ensures that the second condition in Step 3 of Algorithm 1 holds for all $\alpha \in [0, \bar{\alpha}_3]$. Using (2.18) and (3.7), we find

$$\langle s(\alpha), x(\alpha) \rangle \leq \langle s, x \rangle \left(1 - \alpha(1 - \beta_1) + \alpha^2 \frac{\eta}{2\nu} \right).$$

We will obtain our desired decrease in total complementarity if α satisfies

$$\langle s, x \rangle \left(1 - \alpha(1 - \beta_1) + \alpha^2 \frac{\eta}{2\nu} \right) \leq \langle s, x \rangle (1 - \alpha(1 - \beta_2)),$$

or equivalently

$$\alpha \left[(\beta_2 - \beta_1) - \alpha \frac{\eta}{2\nu} \right] \geq 0.$$

Hence, we have $\langle s(\alpha), x(\alpha) \rangle \leq \langle s, x \rangle (1 - \alpha(1 - \beta_2))$ for all $\alpha \in [0, \bar{\alpha}_4]$, where

$$\bar{\alpha}_4 := \frac{2(\beta_2 - \beta_1)\nu}{\eta}. \quad (3.14)$$

Taking into account (3.11), (3.12), (3.13), (3.14), we obtain

$$\alpha^* := \min(1, \bar{\alpha}_1, \bar{\alpha}_2, \bar{\alpha}_3, \bar{\alpha}_4) = \min \left(1, \sqrt{\frac{1}{2\eta}}, \sqrt{\frac{\theta_F}{\eta}}, \frac{\beta_1\theta_F}{\eta}, \frac{2\beta_1\nu}{\eta}, \frac{2(\beta_2 - \beta_1)\nu}{\eta} \right) = \Omega(1/\eta). \quad (3.15)$$

We are ready to state our first main theorem.

Theorem 3.5 *Given (A, b, c, K) , $1 > \beta_2 > \beta_1 > 0$ and $\theta_F, \epsilon^*, \Omega^* > 0$, if all iterates of Algorithm 1 satisfy $\|x_k\| + \|s_k\|_* \leq \Omega^*$, then we obtain a solution (x^*, y^*, s^*) such that $\langle s^*, x^* \rangle \leq \epsilon^* \langle s_0, x_0 \rangle$, $\|Ax^* - b\| \leq \epsilon^* \|Ax_0 - b\|$, and $\|A^*y^* + s^* - c\| \leq \epsilon^* \|A^*y_0 + s_0 - c\|$ in $O\left(\eta \ln\left(\frac{1}{\epsilon^*}\right)\right)$ iterations.*

Proof : As $\|x_k\| + \|s_k\|_* \leq \Omega^*$ at each iterate, if we choose any $\alpha \in [0, \alpha^*]$, with α^* in (3.15), as the step length, all the conditions in Step 3 of Algorithm 1 are satisfied. Hence for each k , $\bar{\alpha}_k \geq \alpha^*$. Thus for $k = \left\lceil \frac{1}{(1-\beta_2)\alpha^*} \right\rceil \ln\left(\frac{1}{\epsilon^*}\right) = O\left(\eta \ln\left(\frac{1}{\epsilon^*}\right)\right)$, we have

$$\begin{aligned} \ln(\langle s_k, x_k \rangle) &\leq \ln(\langle s_{k-1}, x_{k-1} \rangle (1 - \alpha^*(1 - \beta_2))) \\ &\leq \ln\left(\langle s_0, x_0 \rangle (1 - \alpha^*(1 - \beta_2))^k\right) \\ &= \ln(\langle s_0, x_0 \rangle) + k \ln(1 - \alpha^*(1 - \beta_2)) \\ &\leq \ln(\langle s_0, x_0 \rangle) - k\alpha^*(1 - \beta_2) \\ &\leq \ln(\langle s_0, x_0 \rangle) + \ln(\epsilon^*) = \ln(\epsilon^* \langle s_0, x_0 \rangle). \end{aligned}$$

The first inequality follows from the decrease in total complementarity condition, the second from the same applied inductively, and the third inequality from the identity $\ln(1 + \xi) \leq \xi$ for all $\xi > -1$. The fourth inequality follows from our assumption on k .

From condition (3.4) it follows that $\max(\phi_p^k, \phi_d^k) \leq \frac{\langle s_k, x_k \rangle}{\langle s_0, x_0 \rangle} \leq \epsilon^*$. Then (3.3) implies that

$$\|Ax_k - b\| \leq \epsilon^* \|Ax_0 - b\|, \text{ and } \|A^*y_k + s_k - c\| \leq \epsilon^* \|A^*y_0 + s_0 - c\|.$$

□

We postpone the discussion on indicators of infeasibility to the end of Section 4.

4 Polynomial iteration complexity

We now present a variation of Algorithm 1 with a neighborhood defined using γ_G , which is narrower than one using γ_F (see (4.1) below). In this section, we obtain some key results leading to bounds on terms involving Δx and Δs , and this leads us to our analysis of polynomial complexity. Most of the results are extensions from the linear programming case proven by Zhang [18]. Though the γ_G -neighborhood is tighter than the γ_F -neighborhood, it can be related (see (4.10), (4.11)) to the γ_∞ -neighborhood used in practice. In our conclusion, we will discuss some implications of the relation between neighborhoods and complexity estimates of the algorithm. Let

$$\mathcal{N}_G(\theta_G) := \{(x, y, s) \in \text{int } \mathcal{Z} : \gamma_G(x, s) \leq \theta_G\}.$$

We note that, as $\gamma_F(x, s) \leq \gamma_G(x, s)$ ((4.17) in Theorem 4.2 of [9]), it follows that

$$\mathcal{N}_G(\theta_F) \subset \mathcal{N}_F(\theta_F), \tag{4.1}$$

and hence $\mathcal{N}_F(\theta_F)$ is a wider neighborhood than $\mathcal{N}_G(\theta_F)$.

Algorithm 2 :

- 1 Suppose given $1 > \beta_2 > \beta_1 > 0$, $\epsilon^* > 0$, $\theta_G > 0$, and $(x_0, y_0, s_0) \in \mathcal{N}_G(\theta_G)$. Set $k = 0$.
- 2 Solve for $(\Delta x_k, \Delta y_k, \Delta s_k)$ from the Newton equations (2.14) at (x_k, y_k, s_k) .
- 3 Let $(x(\alpha), y(\alpha), s(\alpha)) := (x_k, y_k, s_k) + \alpha(\Delta x_k, \Delta y_k, \Delta s_k)$. Compute the largest step length $\bar{\alpha}_k \in (0, 1]$ such that for all $\alpha \in [0, \bar{\alpha}_k]$,

$$(x(\alpha), y(\alpha), s(\alpha)) \in \mathcal{N}_G(\theta_G), \quad (4.2)$$

$$\langle s(\alpha), x(\alpha) \rangle \geq \max(\phi_p^k, \phi_d^k)(1 - \alpha) \langle s_0, x_0 \rangle, \quad (4.3)$$

$$\langle s(\alpha), x(\alpha) \rangle \leq \langle s_k, x_k \rangle (1 - (1 - \beta_2)\alpha). \quad (4.4)$$

- 4 Choose a primal step length α_p^k and a dual step length α_d^k such that

$$\begin{aligned} (x_{k+1}, y_{k+1}, s_{k+1}) &:= (x_k + \alpha_p^k \Delta x_k, y_k + \alpha_d^k \Delta y_k, s_k + \alpha_d^k \Delta s_k) \in \mathcal{N}_G(\theta_G), \\ \langle s_{k+1}, x_{k+1} \rangle &\geq \max(\phi_p^k(1 - \alpha_p^k), \phi_d^k(1 - \alpha_d^k)) \langle s_0, x_0 \rangle \text{ and} \\ \langle s_{k+1}, x_{k+1} \rangle &\leq \langle s_k, x_k \rangle (1 - (1 - \beta_2)\bar{\alpha}_k). \end{aligned}$$

- 5 Increase k by 1. If $\langle s_k, x_k \rangle < \epsilon^* \langle s_0, x_0 \rangle$, then STOP. Otherwise repeat Step 2.

In this section too, if we choose $\alpha_p^k = \alpha_d^k = \bar{\alpha}_k$, all the conditions in Step 4 are satisfied. However, we are free to choose different step lengths as long as a comparable decrease in the complementarity is obtained, the iterate remains in the required neighborhood, and we maintain the condition (3.4). In our analysis, we again consider equal step lengths α in the k -th iteration, irrespective of the choice of step lengths in previous iterations, and show that a certain minimal step length can be chosen. Then different step lengths can be chosen, but we will preserve a minimal level of decrease in the complementarity. The relations in (3.3) hold also for this algorithm. That is,

$$Ax_k - b = \phi_p^k(Ax_0 - b), \text{ and } A^*y_k + s_k - c = \phi_d^k(A^*y_0 + s_0 - c). \quad (4.5)$$

Let $(u_0, r_0, v_0) \in E \times Y \times E^*$, satisfying $Au_0 = b$, $A^*r_0 + v_0 = c$ and $(x_0 - u_0, s_0 - v_0) \in \text{int } K \times \text{int } K^*$, denote our reference point. It is feasible to the linear system but not necessarily feasible to the cone constraint. The last condition can be met by scaling our initial point by a large positive scalar.

For the given sequence of iterates $\{(x_k, y_k, s_k)\}$ we find it useful to define the following :

$$\begin{aligned} u_{k+1} &= u_k + \alpha_p^k(x_k + \Delta x_k - u_k) = (1 - \alpha_p^k)(u_k - x_k) + x_{k+1}; \\ r_{k+1} &= r_k + \alpha_d^k(y_k + \Delta y_k - r_k) = (1 - \alpha_d^k)(r_k - y_k) + y_{k+1}; \\ v_{k+1} &= v_k + \alpha_d^k(s_k + \Delta s_k - v_k) = (1 - \alpha_d^k)(v_k - s_k) + s_{k+1}. \end{aligned}$$

The properties below directly follow from the above definitions :

$$\begin{aligned} x_{k+1} - u_{k+1} &= (1 - \alpha_p^k)(x_k - u_k) = \phi_p^{k+1}(x_0 - u_0) \in \text{int } K; \\ s_{k+1} - v_{k+1} &= (1 - \alpha_d^k)(s_k - v_k) = \phi_d^{k+1}(s_0 - v_0) \in \text{int } K^*; \\ Au_k &= b \text{ and } A^*r_k + v_k = c \text{ for all } k; \\ A(x_k + \Delta x_k - u_k) &= A(x_k + \Delta x_k) - Au_k = b - b = 0; \\ A^*(y_k + \Delta y_k - r_k) + s_k + \Delta s_k - v_k &= 0. \end{aligned} \quad (4.6)$$

(The third line holds for $k = 0$ by assumption, and then holds for all k by induction using the last two lines.)

The analysis in this section is quite similar to that in the previous section. We will henceforth denote $x_k, y_k, s_k, w_k, \phi_p^k$, and ϕ_d^k by x, y, s, w, ϕ_p , and ϕ_d respectively. We also write $\bar{\phi}$ for $\max(\phi_p, \phi_d)$ and $\underline{\phi}$ for $\min(\phi_p, \phi_d)$. We will drop the subscript k unless the subscript is necessary for clarity. We first prove the following useful lemma.

Lemma 4.1 *Let (x, y, s) be any iterate generated by the algorithm and (x^*, y^*, s^*) be an optimal solution to (P) and (D). Then*

$$\frac{\langle s, x - u \rangle + \langle s - v, x \rangle}{\langle s, x \rangle} \leq 1 + \frac{\langle s^*, x_0 - u_0 \rangle + \langle s_0 - v_0, x^* \rangle + \langle s_0 - v_0, x_0 - u_0 \rangle}{\langle s_0, x_0 \rangle}.$$

If a strictly feasible point for (P) and (D) (say $(\tilde{x}, \tilde{y}, \tilde{s})$) exists, then the sequence $\{(x_k, s_k)\}$ generated by the algorithm is uniformly bounded.

Proof : We first observe the following inequality, assuming the iterate is the k th, so that $\bar{\phi} = \max(\phi_p, \phi_d)$ and $\underline{\phi} = \min(\phi_p, \phi_d)$:

$$\begin{aligned} \langle s - v, x - u \rangle &= \phi_p \phi_d \langle s_0 - v_0, x_0 - u_0 \rangle \quad (\text{see (4.6)}) \\ &= \bar{\phi} \underline{\phi} \langle s_0 - v_0, x_0 - u_0 \rangle \\ &\leq \underline{\phi} \frac{\langle s, x \rangle}{\langle s_0, x_0 \rangle} \langle s_0 - v_0, x_0 - u_0 \rangle \quad (\text{by (3.4)}). \end{aligned} \tag{4.7}$$

Next, from $\langle s^* - v, x^* - u \rangle = 0$, $\langle s^*, x^* \rangle = 0$, $x^* \in K$ and $s^* \in K^*$, we have

$$\begin{aligned} \langle s, x - u \rangle + \langle s - v, x \rangle &< \langle s, x - u \rangle + \langle s - v, x \rangle + \langle s^*, x \rangle + \langle s, x^* \rangle + \langle s^* - v, x^* - u \rangle \\ &= \langle s, x \rangle + \langle s - v, x^* \rangle + \langle s^*, x - u \rangle + \langle s - v, x - u \rangle + \langle s^*, x^* \rangle \\ &\leq \langle s, x \rangle + \bar{\phi} [\langle s_0 - v_0, x^* \rangle + \langle s^*, x_0 - u_0 \rangle] + \underline{\phi} \frac{\langle s, x \rangle}{\langle s_0, x_0 \rangle} \langle s_0 - v_0, x_0 - u_0 \rangle \\ &\leq \langle s, x \rangle + \frac{\langle s, x \rangle}{\langle s_0, x_0 \rangle} [\langle s_0 - v_0, x^* \rangle + \langle s^*, x_0 - u_0 \rangle] + \underline{\phi} \frac{\langle s, x \rangle}{\langle s_0, x_0 \rangle} \langle s_0 - v_0, x_0 - u_0 \rangle \\ &= \langle s, x \rangle \left\{ 1 + \frac{\langle s^*, x_0 - u_0 \rangle + \langle s_0 - v_0, x^* \rangle + \underline{\phi} \langle s_0 - v_0, x_0 - u_0 \rangle}{\langle s_0, x_0 \rangle} \right\}. \end{aligned}$$

The first inequality is strict because $x \in \text{int } K$ and $s \in \text{int } K^*$. We get the first equality by rearranging the terms. The second inequality above follows from (4.7) and (4.6) and the third follows from (3.4). This gives the inequality in the lemma since $\underline{\phi} \leq 1$.

For the second part of the lemma, observe that

$$\begin{aligned} \langle s, \tilde{x} \rangle + \langle \tilde{s}, x \rangle &< \langle s, \tilde{x} \rangle + \langle \tilde{s}, x \rangle + \langle s, x - u \rangle + \langle s - v, x \rangle + \langle \tilde{s} - v, \tilde{x} - u \rangle \\ &= \langle s, x \rangle + \langle s - v, \tilde{x} \rangle + \langle \tilde{s}, x - u \rangle + \langle s - v, x - u \rangle + \langle \tilde{s}, \tilde{x} \rangle \\ &\leq \langle s_0, x_0 \rangle + \langle s_0 - v_0, \tilde{x} \rangle + \langle \tilde{s}, x_0 - u_0 \rangle + \langle s_0 - v_0, x_0 - u_0 \rangle + \langle \tilde{s}, \tilde{x} \rangle \quad (\text{as } \bar{\phi} \leq 1). \end{aligned}$$

□

We define

$$\gamma_\infty(x, s) := \sigma^2 \mu - 1, \quad (4.8)$$

where $\sigma := \sigma_x(w)$ and $\mu := \frac{\langle s, x \rangle}{\nu}$. This definition follows from Lemma 3.1 and (4.10) of [9]. Then, $\mathcal{N}_\infty(\theta) := \{(x, y, s) \in \text{int } \mathcal{Z} : \gamma_\infty(x, s) \leq \theta\}$ defines the γ_∞ -neighborhood (parameterized by θ) widely used in the literature: for linear programming, this measure is the arithmetic mean of the $x_i s_i$'s divided by their minimum, minus one, and for semidefinite the same but using the eigenvalues of $x^{1/2} s x^{1/2}$.

From (4.18) in [9], we have

$$\frac{\gamma_\infty^2}{1 + \gamma_\infty} \leq \gamma_G \leq \nu \gamma_\infty. \quad (4.9)$$

From this we get $\gamma_\infty \leq \frac{\gamma_G}{2} + \sqrt{\gamma_G + \frac{\gamma_G^2}{4}}$. This leads to the following relations:

$$\frac{\gamma_G}{\nu} \leq \gamma_\infty \leq \frac{\sqrt{\gamma_G}}{2} + \sqrt{\frac{5\gamma_G}{4}} \leq 2\sqrt{\gamma_G} \leq \gamma_G + 1 \quad \text{for } \gamma_G \leq 1, \quad (4.10)$$

$$\frac{\gamma_G}{\nu} \leq \gamma_\infty \leq \frac{\gamma_G}{2} + (\frac{\gamma_G}{2} + 1) \leq \gamma_G + 1 \quad \text{for } \gamma_G \geq 1. \quad (4.11)$$

From (4.10) and (4.11) we conclude

$$\mathcal{N}_G(\theta_G) \subset \mathcal{N}_\infty(\theta_G + 1); \mathcal{N}_\infty(\theta) \subset \mathcal{N}_G(\nu\theta). \quad (4.12)$$

So $\gamma_G \leq \theta_G$ implies that $\gamma_\infty \leq \theta_\infty - 1$ for $\theta_\infty := \theta_G + 2$. This yields the following simple bound on σ :

$$\sigma \leq \sqrt{\frac{\theta_\infty}{\mu}}. \quad (4.13)$$

We will first note some useful consequences of Lemma 3.4 in [9] and Corollary 4.1 (ii) in [8]. Let $p \in E$, $x \in \text{int } K$, $q \in E^*$, $s \in \text{int } K^*$, $w \in \text{int } K$ such that $F''(w)x = s$ and $t = -F'(w)$. From Lemma 3.4 in [9] we have

$$\sigma_x(-F'_*(s)) = \sigma_s(-F'(x)) = \sigma_x(w)^2 = \sigma^2, \quad (4.14)$$

and from Corollary 4.1 (ii) in [8] we have

$$F''(x) \leq \sigma_x(-F'_*(s))F''(w) = \sigma^2 F''(w) \text{ and } F''_*(s) \leq \sigma_s(-F'(x))F''_*(t) = \sigma^2 F''_*(t). \quad (4.15)$$

Here for two self-adjoint operators A and B , $A \leq B$ means that $B - A$ is positive semidefinite.

Let us define

$$t = t_k := \sqrt{\|\Delta x\|_w^2 + \|\Delta s\|_{*w}^2}. \quad (4.16)$$

The following two quantities play a crucial role in our bound for t :

$$\chi = \chi_k := 2 \left\{ \frac{\langle s_0 - v_0, x_0 - u_0 \rangle}{\langle s_0, x_0 \rangle} \right\} + \beta_1^2 \frac{\gamma_G}{\nu} + (1 - \beta_1)^2, \text{ and} \quad (4.17)$$

$$\xi = \xi_k := \sqrt{\theta_\infty \nu} \left[\frac{\langle s, x - u \rangle + \langle s - v, x \rangle}{\langle s, x \rangle} \right]. \quad (4.18)$$

Proposition 4.2 $t_k^2 \leq \omega \langle s_k, x_k \rangle$, where ω is independent of k and

$$\left(\xi_k + \sqrt{\xi_k^2 + \chi_k} \right)^2 \leq \omega < \infty. \quad (4.19)$$

Proof : We will drop the subscripts for now. First, we note the following identity.

$$\begin{aligned} \|h\|_w^{*2} &= \langle -\beta_1 \mu F'(x) - s, -\beta_1 \mu F'_*(s) - x \rangle \\ &= \beta_1^2 \mu^2 \langle F'(x), F'_*(s) \rangle - \beta_1 \mu (\langle s, -F'_*(s) \rangle + \langle -F'(x), x \rangle) + \langle s, x \rangle \\ &= \beta_1^2 \mu^2 \langle F'(x), F'_*(s) \rangle - 2\beta_1 \mu \nu + \langle s, x \rangle \quad (\text{from (2.4)}) \\ &= \mu \nu \left[\beta_1^2 \left(\frac{\gamma G}{\nu} + 1 \right) - 2\beta_1 + 1 \right] \quad (\text{from (3.2)}) \\ &= \langle s, x \rangle \left[(1 - \beta_1)^2 + \beta_1^2 \frac{\gamma G}{\nu} \right]. \end{aligned}$$

From (2.15) in Lemma 2.3 and the expression for t in (4.16), we have

$$\|\Delta x\|_w^2 + \|\Delta s\|_w^{*2} + 2 \langle \Delta s, \Delta x \rangle = t^2 + 2 \langle \Delta s, \Delta x \rangle = \langle s, x \rangle \left[\beta_1^2 \frac{\gamma G}{\nu} + (1 - \beta_1)^2 \right]. \quad (4.20)$$

Now we will show that $0 \leq \langle s, x \rangle \frac{\langle s_0 - v_0, x_0 - u_0 \rangle}{\langle s_0, x_0 \rangle} + \langle \Delta s, \Delta x \rangle + \xi t \sqrt{\langle s, x \rangle}$.

Expanding $\langle s + \Delta s - v, x + \Delta x - u \rangle$ and using (4.6), it follows that

$$\langle s - v, x - u \rangle + \langle \Delta s, \Delta x \rangle + \langle \Delta s, x - u \rangle + \langle s - v, \Delta x \rangle = 0. \quad (4.21)$$

It follows from (4.15) that for $p \in E, q \in E^*$, $\|p\|_x \leq \sigma \|p\|_w$ and $\|q\|_s \leq \sigma \|q\|_w^*$, so that

$$\|\Delta x\|_x \leq \sigma \|\Delta x\|_w \quad \text{and} \quad \|\Delta s\|_s \leq \sigma \|\Delta s\|_w^*. \quad (4.22)$$

By letting $p = x - u, q = s - v$ in Lemma 2.2 we have

$$\|x - u\|_s^* \leq \langle s, x - u \rangle \quad \text{and} \quad \|s - v\|_x^* \leq \langle s - v, x \rangle. \quad (4.23)$$

From (4.22), (4.23), (4.13), and (4.16), we see that

$$\langle \Delta s, x - u \rangle \leq \|x - u\|_s^* \|\Delta s\|_s \leq \langle s, x - u \rangle \sigma \|\Delta s\|_w^* \leq \sqrt{\frac{\theta_\infty}{\mu}} \langle s, x - u \rangle t. \quad (4.24)$$

A similar bound holds for $\langle s - v, \Delta x \rangle$.

Substituting these bounds and (4.7) in (4.21), we get

$$\begin{aligned} 0 &\leq \frac{\langle s, x \rangle}{\langle s_0, x_0 \rangle} \langle s_0 - v_0, x_0 - u_0 \rangle + \langle \Delta s, \Delta x \rangle + \sqrt{\frac{\theta_\infty}{\mu}} \langle s, x - u \rangle t + \sqrt{\frac{\theta_\infty}{\mu}} \langle s - v, x \rangle t \\ &= \langle s, x \rangle \frac{\langle s_0 - v_0, x_0 - u_0 \rangle}{\langle s_0, x_0 \rangle} + \langle \Delta s, \Delta x \rangle + \xi t \sqrt{\langle s, x \rangle}. \end{aligned}$$

Using (4.20) to eliminate $\langle \Delta s, \Delta x \rangle$, we get

$$t^2 \leq 2 \langle s, x \rangle \frac{\langle s_0 - v_0, x_0 - u_0 \rangle}{\langle s_0, x_0 \rangle} + \langle s, x \rangle \left[\beta_1^2 \frac{\gamma G}{\nu} + (1 - \beta_1)^2 \right] + 2\xi t \sqrt{\langle s, x \rangle}.$$

So, $t^2 \leq 2\xi t \sqrt{\langle s, x \rangle} + \langle s, x \rangle \chi$ or $(t - \xi \sqrt{\langle s, x \rangle})^2 \leq \langle s, x \rangle (\chi + \xi^2)$. Therefore,

$$t_k^2 \leq \langle s_k, x_k \rangle \left(\xi_k + \sqrt{\xi_k^2 + \chi_k} \right)^2.$$

Since $\gamma_G \leq \theta_G$, χ_k is uniformly bounded by $\bar{\chi}$, and ξ_k is uniformly bounded by $\bar{\xi}$ using Lemma 4.1. We can choose $\bar{\chi}$, $\bar{\xi}$, and ω to be

$$\bar{\chi} = \beta_1^2 \frac{\theta_G}{\nu} + (1 - \beta_1)^2 + 2 \left\{ \frac{\langle s_0 - v_0, x_0 - u_0 \rangle}{\langle s_0, x_0 \rangle} \right\}, \quad (4.25)$$

$$\bar{\xi} = \sqrt{\theta_\infty \nu} \left\{ 1 + \frac{\langle s^*, x_0 - u_0 \rangle + \langle s_0 - v_0, x^* \rangle + \langle s_0 - v_0, x_0 - u_0 \rangle}{\langle s_0, x_0 \rangle} \right\}, \text{ and} \quad (4.26)$$

$$\omega \geq \left(\bar{\xi} + \sqrt{\bar{\xi}^2 + \bar{\chi}} \right)^2. \quad (4.27)$$

This completes the proof of the proposition. \square

In the following corollaries ω plays a role similar to η in the previous section. Later, we will obtain a polynomial bound for ω , given a suitable choice of starting points and an assumption on the size of optimal solutions.

Corollary 4.3 $|\langle \Delta s, \Delta x \rangle| \leq \frac{\omega}{2} \langle s, x \rangle$.

Proof : $|\langle \Delta s, \Delta x \rangle| \leq \|\Delta x\|_w \|\Delta s\|_w^* \leq \frac{\|\Delta x\|_w^2 + \|\Delta s\|_w^{*2}}{2} \leq \frac{\omega}{2} \langle s, x \rangle$. \square

Corollary 4.4 $\|\Delta x\|_x^2 + \|\Delta s\|_s^2 \leq \theta_\infty \nu \omega$ and $\|\Delta x\|_x \|\Delta s\|_s \leq \frac{1}{2} \theta_\infty \nu \omega$.

Proof : From (4.22) $\|\Delta x\|_x \leq \sigma \|\Delta x\|_w$ and $\|\Delta s\|_s \leq \sigma \|\Delta s\|_w^*$. Hence,

$$\|\Delta x\|_x^2 + \|\Delta s\|_s^2 \leq \sigma^2 (\|\Delta x\|_w^2 + \|\Delta s\|_w^{*2}) \leq \frac{\theta_\infty}{\mu} \omega \langle s, x \rangle = \theta_\infty \nu \omega.$$

The second part follows from the inequality $\|\Delta x\|_x \|\Delta s\|_s \leq \frac{\|\Delta x\|_x^2 + \|\Delta s\|_s^2}{2}$. \square

Let $x(\alpha) := x + \alpha \Delta x$ and $s(\alpha) := s + \alpha \Delta s$. We use the following lemma to guarantee that our iterates x and s remain in their respective cones.

Lemma 4.5 Let $\bar{\kappa} = \max\{\|\Delta x\|_x, \|\Delta s\|_s\}$. Then for $\alpha \in [0, 1/(2\bar{\kappa})]$, $x(\alpha) \in \text{int } K$, $\sigma_{x(\alpha)}(x) \leq 2$, $s(\alpha) \in \text{int } K^*$ and $\sigma_{s(\alpha)}(s) \leq 2$.

Proof : From the definition of $\bar{\kappa}$, we have $\|\alpha \Delta x\|_x \leq 1/2$ and $\|\alpha \Delta s\|_s \leq 1/2$. Therefore $x(\alpha) \in \text{int } K$ and $s(\alpha) \in \text{int } K^*$. Note that $2x(\alpha) - x = 2(x + \alpha \Delta x) - x = x + 2\alpha \Delta x \in K$ as $\|\alpha \Delta x\|_x \leq 1/2$. Hence, $\sigma_{x(\alpha)}(x) \leq 2$, and similarly $\sigma_{s(\alpha)}(s) \leq 2$. \square

From Corollary 4.4,

$$\frac{1}{2\bar{\kappa}} = \frac{1}{2 \max\{\|\Delta x\|_x, \|\Delta s\|_s\}} \geq \frac{1}{2\sqrt{\theta_\infty \nu \omega}} =: \bar{\alpha}_1. \quad (4.28)$$

Henceforth we will restrict our choice of α to $[0, \bar{\alpha}_1]$. Note that

$$\alpha \|\Delta x\|_x \leq 1/2, \quad \alpha \|\Delta s\|_s \leq 1/2, \quad (4.29)$$

and the conclusion of Lemma 4.5 hold for all such α .

We have so far obtained bounds on expressions involving Δx and Δs and on step lengths guaranteeing that our iterates remain in the interiors of the respective cones. Now, we proceed to guarantee the existence of a minimum positive step length such that we can satisfy (4.2)–(4.4), that is, stay inside of the $\mathcal{N}_G(\theta_G)$ neighborhood, while satisfying the condition in (3.4) and the decrease in total complementarity condition. First we will focus on the arguments that lead to bounding γ_G along the step directions. We introduce the following convenient notation.

Using Taylor series expansions, let us define R_x and R_s by

$$F'(x(\alpha)) = F'(x) + F''(x)(\alpha\Delta x) + \alpha^2 R_x, \quad F'_*(s(\alpha)) = F'_*(s) + F''_*(s)(\alpha\Delta s) + \alpha^2 R_s. \quad (4.30)$$

We have

Lemma 4.6 For $\alpha \in [0, \bar{\alpha}_1]$,

$$\|R_x\|_w^* \leq 2\sigma \|\Delta x\|_x^2, \quad \|R_s\|_w \leq 2\sigma \|\Delta s\|_s^2. \quad (4.31)$$

Proof : From Theorem 4.3 in [8], since $|\Delta x|_x \leq \|\Delta x\|_x$, we get $\|R_x\|_{x(\alpha)}^* \leq \|\Delta x\|_x^2$, and similarly $\|R_s\|_{s(\alpha)}^* \leq \|\Delta s\|_s^2$. We now just need to change the local norms to obtain our result.

The identities in (4.15) can also be written as

$$F''(w)^{-1} \leq \sigma^2 F''(x)^{-1}, \quad \text{and} \quad F''_*(t)^{-1} \leq \sigma^2 F''_*(s)^{-1}. \quad (4.32)$$

Let $p \in E$, $x, \hat{x} \in \text{int } K$, $q \in E^*$, $s, \hat{s} \in \text{int } K^*$, w such that $F''(w)x = s$ and $t = -F'(w)$. Then, using (4.32) and (4.14) we have

$$\|q\|_w^* = \langle q, F''(w)^{-1}q \rangle^{1/2} \leq \sigma \langle q, F''(x)^{-1}q \rangle^{1/2} = \sigma \|q\|_x^*, \quad (4.33)$$

$$\|p\|_w = \|p\|_t^* \leq \sigma \langle F''_*(s)^{-1}p, p \rangle^{1/2} = \sigma \|p\|_s^*; \quad (4.34)$$

$$\|q\|_x^* = \langle q, F''(x)^{-1}q \rangle^{1/2} \leq \sigma_{\hat{x}}(x) \langle q, F''(\hat{x})^{-1}q \rangle^{1/2} = \sigma_{\hat{x}}(x) \|q\|_{\hat{x}}^*, \quad \text{and} \quad (4.35)$$

$$\|p\|_s^* = \langle F''_*(s)^{-1}p, p \rangle^{1/2} \leq \sigma_{\hat{s}}(s) \langle F''_*(\hat{s})^{-1}p, p \rangle^{1/2} = \sigma_{\hat{s}}(s) \|p\|_{\hat{s}}^*. \quad (4.36)$$

In (4.35) and (4.36), \hat{x} and \hat{s} are used in the role of scaling points.

Using the above relations we get $\|R_x\|_w^* \leq \sigma \|R_x\|_x^* \leq \sigma \sigma_{x(\alpha)}(x) \|R_x\|_{x(\alpha)}^* \leq 2\sigma \|\Delta x\|_x^2$, and similarly $\|R_s\|_w \leq \sigma \|R_s\|_s^* \leq \sigma \sigma_{s(\alpha)}(s) \|R_s\|_{s(\alpha)}^* \leq 2\sigma \|\Delta s\|_s^2$. The first set of inequalities follow from (4.33) and

(4.34). The second set follows from (4.35) and (4.36). The third follows from the bounds at the beginning of the proof. \square

Now $\gamma_G(x(\alpha), s(\alpha))$ depends on $\delta := \langle F'(x(\alpha)), F'_*(s(\alpha)) \rangle$. Expanding δ into linear, quadratic, cubic and quartic terms using (4.30), we get

$$\begin{aligned} \delta &= \langle F'(x), F'_*(s) \rangle + \alpha \{ \langle F''(x)\Delta x, F'_*(s) \rangle + \langle F'(x), F''_*(s)\Delta s \rangle \} \\ &\quad + \alpha^2 \{ \langle F'(x), R_s \rangle + \langle R_x, F'_*(s) \rangle + \langle F''(x)\Delta x, F''_*(s)\Delta s \rangle \} \\ &\quad + \alpha^3 \{ \langle R_x, F''_*(s)\Delta s \rangle + \langle F''(x)\Delta x, R_s \rangle \} + \alpha^4 \langle R_x, R_s \rangle. \end{aligned} \quad (4.37)$$

We begin with a bound for the linear term.

Lemma 4.7

$$\langle F''(x)\Delta x, F'_*(s) \rangle + \langle F'(x), F''_*(s)\Delta s \rangle \leq \langle F'(x), F'_*(s) \rangle \left\{ -\frac{\beta_1}{\nu}\gamma_G + (1 - \beta_1) \right\}. \quad (4.38)$$

Proof : Using relations (2.11) and $F''(w)\Delta x + \Delta s = h = -\beta_1\mu F'(x) - s$, we get

$$\begin{aligned} &\langle F''(x)\Delta x, F'_*(s) \rangle + \langle F'(x), F''_*(s)\Delta s \rangle \\ &= \langle F''(x)\Delta x, F'_*(s) \rangle + \langle F'(x), F''(w)^{-1}F''(x)F''(w)^{-1}\Delta s \rangle \\ &= \langle F''(x)\Delta x, F'_*(s) \rangle + \langle F''(x)F''(w)^{-1}\Delta s, F'_*(s) \rangle \\ &= \langle F''(x)F''(w)^{-1}(F''(w)\Delta x + \Delta s), F'_*(s) \rangle \\ &= \langle F''(x)F''(w)^{-1}(-\beta_1\mu F'(x) - s), F'_*(s) \rangle \\ &= -\beta_1\mu \langle F''(x)F''(w)^{-1}F'(x), F'_*(s) \rangle - \langle F''(x)F''(w)^{-1}s, F'_*(s) \rangle \\ &= -\beta_1\mu \langle F''(x)F'_*(s), F'_*(s) \rangle - \langle F''(x)x, F'_*(s) \rangle \\ &= -\beta_1\mu \langle F''(x)F'_*(s), F'_*(s) \rangle + \langle F'(x), F'_*(s) \rangle. \end{aligned}$$

We note that $\langle F''(x)F'_*(s), F'_*(s) \rangle \geq \frac{\langle F'(x), F'_*(s) \rangle^2}{\nu}$. This follows from

$$\langle F'(x), F'_*(s) \rangle \leq \|F'_*(s)\|_x \|F'(x)\|_x^* = \sqrt{\nu} \|F'_*(s)\|_x.$$

Hence,

$$\begin{aligned} \langle F''(x)\Delta x, F'_*(s) \rangle + \langle F'(x), F''_*(s)\Delta s \rangle &\leq -\beta_1\mu \frac{\langle F'(x), F'_*(s) \rangle^2}{\nu} + \langle F'(x), F'_*(s) \rangle \\ &= \langle F'(x), F'_*(s) \rangle \left\{ -\frac{\beta_1}{\nu}\mu \langle F'(x), F'_*(s) \rangle + 1 \right\} \\ &= \langle F'(x), F'_*(s) \rangle \left\{ -\frac{\beta_1}{\nu}(\gamma_G + \nu) + 1 \right\} \\ &= \langle F'(x), F'_*(s) \rangle \left\{ -\frac{\beta_1}{\nu}\gamma_G + (1 - \beta_1) \right\}. \end{aligned}$$

\square

Using (2.11), we obtain

$$\|F'(x)\|_w^{*2} = \langle F'(x), F''(w)^{-1}F'(x) \rangle = \langle F'(x), F'_*(s) \rangle = \langle F''(w)F'_*(s), F'_*(s) \rangle = \|F'_*(s)\|_w^2.$$

So, we let

$$\pi := \|F'(x)\|_w^* = \|F'_*(s)\|_w. \quad (4.39)$$

For future use, we note that

$$\pi^2 = \langle F'(x), F'_*(s) \rangle = \frac{\gamma_G + \nu}{\mu} = \sigma^2 \frac{\gamma_G + \nu}{\gamma_\infty + 1} \leq \nu\sigma^2. \quad (4.40)$$

This follows from relations (3.2), (4.8), and (4.9).

We now bound the quadratic terms. Using the above and (4.31) we get

$$\begin{aligned} \langle F'(x), R_s \rangle &\leq \|R_s\|_w \|F'(x)\|_w^* \\ &\leq 2\pi\sigma \|\Delta s\|_s^2, \end{aligned} \quad (4.41)$$

and similarly

$$\langle R_x, F'_*(s) \rangle \leq 2\pi\sigma \|\Delta x\|_x^2. \quad (4.42)$$

Finally,

$$\begin{aligned} \langle F''(x)\Delta x, F''_*(s)\Delta s \rangle &\leq \|F''_*(s)\Delta s\|_w \|F''(x)\Delta x\|_w^* \\ &\leq \sigma^2 \|F''_*(s)\Delta s\|_s^* \|F''(x)\Delta x\|_x^* \quad (\text{by (4.33) and (4.34)}) \\ &= \sigma^2 \|\Delta s\|_s \|\Delta x\|_x. \end{aligned} \quad (4.43)$$

To bound the cubic and quartic terms, we have

$$\begin{aligned} \langle R_x, F''_*(s)\Delta s \rangle + \langle F''(x)\Delta x, R_s \rangle &\leq \|F''_*(s)\Delta s\|_w \|R_x\|_w^* + \|R_s\|_w \|F''(x)\Delta x\|_w^* \\ &\leq 2\sigma^2 \|F''_*(s)\Delta s\|_s^* \|\Delta x\|_x^2 + 2\sigma^2 \|F''(x)\Delta x\|_x^* \|\Delta s\|_s^2 \\ &\quad (\text{by (4.33), (4.34), and (4.31)}) \\ &= 2\sigma^2 \|\Delta s\|_s \|\Delta x\|_x^2 + 2\sigma^2 \|\Delta x\|_x \|\Delta s\|_s^2. \end{aligned} \quad (4.44)$$

Using (4.31), we have

$$\langle R_x, R_s \rangle \leq \|R_s\|_w \|R_x\|_w^* \leq 4\sigma^2 \|\Delta x\|_x^2 \|\Delta s\|_s^2. \quad (4.45)$$

Let

$$L := \langle F'(x), F'_*(s) \rangle \left\{ -\frac{\beta_1}{\nu} \gamma_G + (1 - \beta_1) \right\}, \quad Q := \sigma^2 \left\{ 2\frac{\pi}{\sigma} (\|\Delta s\|_s^2 + \|\Delta x\|_x^2) + 4\|\Delta s\|_s \|\Delta x\|_x \right\}.$$

Substituting (4.38), (4.41), (4.42), (4.43), (4.44) and (4.45) in (4.37), and using (4.29) in the first inequality we get

$$\begin{aligned} \delta &\leq \langle F'(x), F'_*(s) \rangle + \alpha L + \alpha^2 \{ 2\pi\sigma \|\Delta s\|_s^2 + 2\pi\sigma \|\Delta x\|_x^2 + \sigma^2 \|\Delta s\|_s \|\Delta x\|_x \} \\ &\quad + \alpha^3 \{ 2\sigma^2 \|\Delta s\|_s \|\Delta x\|_x^2 + 2\sigma^2 \|\Delta x\|_x \|\Delta s\|_s^2 \} + 4\alpha^4 \sigma^2 \|\Delta x\|_x^2 \|\Delta s\|_s^2 \\ &\leq \langle F'(x), F'_*(s) \rangle + \alpha L + \alpha^2 \sigma^2 \left\{ 2\frac{\pi}{\sigma} (\|\Delta s\|_s^2 + \|\Delta x\|_x^2) + \|\Delta s\|_s \|\Delta x\|_x \right\} \\ &\quad + \alpha^2 \{ 2\sigma^2 \|\Delta s\|_s \|\Delta x\|_x (1/2) + 2\sigma^2 \|\Delta x\|_x \|\Delta s\|_s (1/2) \} + 4\alpha^2 \sigma^2 \|\Delta x\|_x \|\Delta s\|_s (1/4) \\ &= \langle F'(x), F'_*(s) \rangle + \alpha L + \alpha^2 \sigma^2 \left\{ 2\frac{\pi}{\sigma} (\|\Delta s\|_s^2 + \|\Delta x\|_x^2) + 4\|\Delta s\|_s \|\Delta x\|_x \right\} \\ &= \langle F'(x), F'_*(s) \rangle + \alpha L + \alpha^2 Q. \end{aligned}$$

Note that $\mu(x(\alpha), s(\alpha)) = \frac{\langle s(\alpha), x(\alpha) \rangle}{\nu} > 0$ as $x(\alpha) \in \text{int } K$ and $s(\alpha) \in \text{int } K^*$ for all $\alpha \in [0, \bar{\alpha}_1]$. So, we have

$$\begin{aligned}
\gamma_G(x(\alpha), s(\alpha)) + \nu &= \mu(x(\alpha), s(\alpha)) \langle F'(x(\alpha)), F_*(s(\alpha)) \rangle \\
&\leq \left\{ \mu + \alpha(\beta_1 - 1)\mu + \alpha^2 \frac{\langle \Delta s, \Delta x \rangle}{\nu} \right\} \{ \langle F'(x), F'_*(s) \rangle + \alpha L + \alpha^2 Q \} \\
&= \gamma_G + \nu + \alpha \{ \mu(\beta_1 - 1) \langle F'(x), F'_*(s) \rangle + L\mu \} \\
&\quad + \alpha^2 \left\{ \mu Q + \frac{\langle \Delta s, \Delta x \rangle}{\nu} \langle F'(x), F'_*(s) \rangle + (\beta_1 - 1)\mu L \right\} \\
&\quad + \alpha^3 \left\{ \frac{\langle \Delta s, \Delta x \rangle}{\nu} L + (\beta_1 - 1)\mu Q \right\} + \alpha^4 \frac{\langle \Delta s, \Delta x \rangle}{\nu} Q.
\end{aligned} \tag{4.46}$$

We derive the following bounds on the terms of the above expansion:

$$\begin{aligned}
\mu(\beta_1 - 1) \langle F'(x), F'_*(s) \rangle + L\mu &= \mu(\beta_1 - 1) \langle F'(x), F'_*(s) \rangle + \mu \langle F'(x), F'_*(s) \rangle \left\{ -\frac{\beta_1}{\nu} \gamma_G + (1 - \beta_1) \right\} \\
&= -\frac{\beta_1}{\nu} \gamma_G \mu \langle F'(x), F'_*(s) \rangle = -\frac{\beta_1}{\nu} \gamma_G (\gamma_G + \nu).
\end{aligned} \tag{4.47}$$

We will now use our bound on π to bound Q . From (4.40), we see that

$$\begin{aligned}
\mu Q &= \mu \sigma^2 \left(2 \sqrt{\frac{\gamma_G + \nu}{\sigma^2 \mu}} (\|\Delta s\|_s^2 + \|\Delta x\|_x^2) + 4 \|\Delta s\|_s \|\Delta x\|_x \right) \\
&\leq 2\mu \sigma^2 \sqrt{\frac{\gamma_G + \nu}{\sigma^2 \mu}} (\sigma^2 \omega \langle s, x \rangle) + 4\mu \sigma^2 \frac{\theta_\infty \nu \omega}{2} \text{ (using the proof of Corollary 4.4)} \\
&= 2\sqrt{\gamma_G + \nu} (\mu \sigma^2)^{3/2} \nu \omega + 2\theta_\infty^2 \nu \omega \text{ (using } \langle s, x \rangle = \mu \nu) \\
&\leq 2\theta_\infty^{3/2} \nu \omega (\sqrt{\theta_G + \nu} + \sqrt{\theta_\infty}) \text{ (by (4.13))}.
\end{aligned} \tag{4.48}$$

Using Corollary 4.3, we get

$$\frac{\langle \Delta s, \Delta x \rangle}{\nu} \langle F'(x), F'_*(s) \rangle \leq \frac{\omega \langle s, x \rangle}{2\nu} \langle F'(x), F'_*(s) \rangle = \frac{\omega}{2} \mu \langle F'(x), F'_*(s) \rangle = \frac{\omega}{2} (\gamma_G + \nu) \leq \frac{\omega}{2} (\theta_G + \nu). \tag{4.49}$$

Next,

$$\begin{aligned}
(\beta_1 - 1)\mu L &= \mu(\beta_1 - 1) \langle F'(x), F'_*(s) \rangle \left\{ -\beta_1 \frac{\gamma_G}{\nu} + (1 - \beta_1) \right\} \\
&= (\beta_1 - 1)(\gamma_G + \nu) \left\{ -\beta_1 \frac{\gamma_G}{\nu} + (1 - \beta_1) \right\} \\
&= \beta_1(1 - \beta_1) \frac{\gamma_G(\gamma_G + \nu)}{\nu} - (1 - \beta_1)^2 (\gamma_G + \nu) \\
&\leq \beta_1(1 - \beta_1) \frac{\theta_G(\theta_G + \nu)}{\nu},
\end{aligned} \tag{4.50}$$

and

$$\begin{aligned}
\frac{\langle \Delta s, \Delta x \rangle}{\nu} L + (\beta_1 - 1)\mu Q &\leq \left| \frac{\langle \Delta s, \Delta x \rangle}{\nu} \right| |L| \\
&\leq \frac{\omega}{2} \mu \langle F'(x), F'_*(s) \rangle \left\{ \frac{\beta_1}{\nu} \gamma_G + (1 - \beta_1) \right\} \\
&\leq \frac{\omega}{2} (\gamma_G + \nu) \left\{ \frac{\beta_1}{\nu} \gamma_G + (1 - \beta_1) \right\} \\
&\leq \frac{\omega(\theta_G + \nu)}{2} (\beta_1 \frac{\theta_G}{\nu} + (1 - \beta_1)). \tag{4.51}
\end{aligned}$$

Using the expression for Q and (4.29) we have

$$\alpha^2 Q = \sigma^2 \left(2 \frac{\pi}{\sigma} (\alpha^2 \|\Delta s\|_s^2 + \alpha^2 \|\Delta x\|_x^2) + 4\alpha^2 \|\Delta s\|_s \|\Delta x\|_x \right) \leq \sigma^2 (\sqrt{\nu} + 1). \tag{4.52}$$

From (4.52), (4.13) and Corollary 4.3 we have

$$\alpha^2 Q \frac{\langle \Delta s, \Delta x \rangle}{\nu} \leq \sigma^2 (\sqrt{\nu} + 1) \frac{\omega \langle s, x \rangle}{2\nu} = \frac{\omega}{2} \theta_\infty (\sqrt{\nu} + 1). \tag{4.53}$$

To reduce the cubic to a quadratic term, we use $\alpha \leq 1$. We use (4.53) to reduce the quartic term to a quadratic. Substituting the corresponding bounds into the expansion (4.46), we get

$$\gamma_G(x(\alpha), s(\alpha)) + \nu \leq \gamma_G + \nu - \alpha \beta_1 \frac{\gamma_G(\gamma_G + \nu)}{\nu} + \alpha^2 \tau,$$

where

$$\begin{aligned}
\tau &= 2\theta_\infty^{3/2} \nu \omega (\sqrt{\theta_G + \nu} + \sqrt{\theta_\infty}) + \frac{\omega}{2} (\theta_G + \nu) + \beta_1 (1 - \beta_1) \frac{\theta_G(\theta_G + \nu)}{\nu} \\
&\quad + \frac{\omega}{2} (\theta_G + \nu) \left[\beta_1 \frac{\theta_G}{\nu} + (1 - \beta_1) \right] + \frac{\theta_\infty}{2} \omega (\sqrt{\nu} + 1) \tag{4.54}
\end{aligned}$$

and the quantities in τ are gotten from (4.47), (4.48), (4.49), (4.50), (4.51) and (4.53).

Denoting γ_G by ζ , we want an α_ζ such that $\zeta - \alpha \beta_1 \frac{\zeta(\zeta + \nu)}{\nu} + \alpha^2 \tau \leq \theta_G$ for all $\alpha \in [0, \alpha_\zeta]$. So, we have

$$\alpha_\zeta = \frac{1}{2\tau} \left\{ \beta_1 \frac{\zeta(\zeta + \nu)}{\nu} + \sqrt{\left(\beta_1 \frac{\zeta(\zeta + \nu)}{\nu} \right)^2 + 4\tau(\theta_G - \zeta)} \right\}.$$

Let $\hat{\alpha}_2 = \min_{\zeta \in [0, \theta_G]} \alpha_\zeta$. Let $f(\zeta) = \beta_1 \zeta(\zeta + \nu)$ and $g = 4\nu^2 \tau (\theta_G - \zeta)$. Then,

$$2\tau \nu \hat{\alpha}_2 = \min_{\zeta \in [0, \theta_G]} f(\zeta) + \sqrt{f^2(\zeta) + g(\zeta)} \geq \min_{\zeta \in [0, \theta_G]} f(\zeta) + \sqrt{g(\zeta)}.$$

Let the derivative of the minimand be $\ell(\zeta) = f'(\zeta) + \frac{g'(\zeta)}{2\sqrt{g(\zeta)}} = \beta_1(2\zeta + \nu) - \frac{4\nu^2 \tau}{2\sqrt{4\nu^2 \tau (\theta_G - \zeta)}} = \beta_1(2\zeta + \nu) - \frac{\nu \sqrt{\tau}}{\sqrt{\theta_G - \zeta}}$. We can check that $\ell''(\zeta)$ depends only on the second term and is negative for all $\zeta \in [0, \theta_G]$. Hence, $\ell(0) \leq 0$ and $\ell'(\theta_G) \leq 0$, implies that $\ell(\zeta) \leq 0$ for all $\zeta \in [0, \theta_G]$. This leads to the condition

that $\tau \geq \tilde{\tau} := \max(\beta_1^2 \theta_G, 16\beta_1^2 \theta_G^3 / \nu^2)$. This can be ensured by requiring that $\omega \geq 32\beta_1 \theta_G / \nu$, because then

$$\tau \geq \frac{\omega}{2}(\theta_G + \nu) + \frac{\omega}{2}(\theta_G + \nu) \left[\beta_1 \frac{\theta_G}{\nu} + (1 - \beta_1) \right] \geq \frac{\omega \nu}{2} + \frac{\omega}{2} \frac{\beta_1 \theta_G^2}{\nu} \geq \beta_1^2 \theta_G + \frac{16\beta_1^2 \theta_G^3}{\nu^2} \geq \tilde{\tau}.$$

Since the derivative is nonpositive throughout, the minimum is achieved at $\zeta = \theta_G$ and

$$\hat{\alpha}_2 = \frac{\beta_1 \theta_G (\theta_G + \nu)}{\nu \tau} \geq \frac{\beta_1 \theta_G}{\tau} =: \bar{\alpha}_2. \quad (4.55)$$

Then $(x(\alpha), y(\alpha), s(\alpha)) \in \mathcal{N}_G$ for all $\alpha \in [0, \bar{\alpha}_2]$. Taking into account the above condition on ω and the bound imposed by Proposition 4.2, we will define

$$\omega := \max \left\{ \left(\bar{\xi} + \sqrt{\bar{\xi}^2 + \bar{\chi}} \right)^2, \frac{32\beta_1 \theta_G}{\nu} \right\}. \quad (4.56)$$

Now we focus on obtaining a guarantee of a positive step length $\bar{\alpha}_3$ satisfying the condition (3.4). We want an $\bar{\alpha}_3$ such that (4.3) holds for all $0 \leq \alpha \leq \bar{\alpha}_3$. Using (2.18) and Corollary 4.3, we note that

$$\begin{aligned} \frac{\langle s(\alpha), x(\alpha) \rangle}{\langle s_0, x_0 \rangle} - \bar{\phi}(1 - \alpha) &= \frac{\langle s, x \rangle}{\langle s_0, x_0 \rangle} (1 + \alpha(\beta_1 - 1)) + \alpha^2 \frac{\langle \Delta s, \Delta x \rangle}{\langle s_0, x_0 \rangle} - \bar{\phi}(1 - \alpha) \\ &= \left(\frac{\langle s, x \rangle}{\langle s_0, x_0 \rangle} - \bar{\phi} \right) (1 - \alpha) + \alpha \beta_1 \frac{\langle s, x \rangle}{\langle s_0, x_0 \rangle} + \alpha^2 \frac{\langle \Delta s, \Delta x \rangle}{\langle s_0, x_0 \rangle} \\ &\geq \alpha \frac{\langle s, x \rangle}{\langle s_0, x_0 \rangle} \left(\beta_1 - \alpha \frac{\omega}{2} \right). \end{aligned}$$

Therefore, it suffices to have

$$\bar{\alpha}_3 := \frac{2\beta_1}{\omega} \quad (4.57)$$

in order that $\langle s(\alpha), x(\alpha) \rangle - \bar{\phi}(1 - \alpha) \langle s_0, x_0 \rangle \geq 0$ for all $\alpha \in [0, \bar{\alpha}_3]$.

Using (2.18) and Corollary 4.3, we get

$$\begin{aligned} \langle s(\alpha), x(\alpha) \rangle &= \langle s, x \rangle (1 - \alpha(1 - \beta_1)) + \alpha^2 \langle \Delta s, \Delta x \rangle \\ &\leq \langle s, x \rangle (1 - \alpha(1 - \beta_1) + \alpha^2 \omega / 2). \end{aligned}$$

Therefore, to satisfy (4.4), it suffices to ensure that

$$\langle s, x \rangle (1 - \alpha(1 - \beta_1) + \alpha^2 \omega / 2) - \langle s, x \rangle (1 - \alpha(1 - \beta_2)) = \langle s, x \rangle \alpha (-(\beta_2 - \beta_1) + \alpha \omega / 2) \leq 0.$$

Thus for all $\alpha \in [0, \bar{\alpha}_4]$ with

$$\bar{\alpha}_4 := \frac{2(\beta_2 - \beta_1)}{\omega}, \quad (4.58)$$

$$\langle s(\alpha), x(\alpha) \rangle \leq \langle s, x \rangle (1 - \alpha(1 - \beta_2)).$$

Taking into account (4.28), (4.55), (4.57), (4.58), we define

$$\alpha^* = \min \left(1, \frac{1}{2\sqrt{\theta_\infty \nu \omega}}, \frac{\beta_1 \theta_G}{\tau}, \frac{2\beta_1}{\omega}, \frac{2(\beta_2 - \beta_1)}{\omega} \right). \quad (4.59)$$

For $\alpha \in [0, \alpha^*]$, all the conditions in Step 3 of Algorithm 2 are satisfied. Observe from (4.54) that $\tau = O(\omega\nu^{1.5})$. So, $\alpha^* = \Omega(\tau^{-1}) = \Omega(\omega^{-1}\nu^{-1.5})$. We will now establish a bound on ω .

Let (u_0, r_0, v_0) be the solution to $\min\{\|u\| + \|v\|_* : Au = b, A^*r + v = c\}$. Let

$$x_0 = \rho_0 e \in \text{int } K, \quad s_0 = -\rho_0 F'(e) \in \text{int } K^*, \quad (4.60)$$

where e is the fixed reference element in $\text{int } K$ and $\rho_0 > \max(\|u_0\|, \|v_0\|_*) \geq \max(|u_0|_e, |v_0|_e)$. Then we have $\sigma_e(x_0 - u_0) = |x_0 - u_0|_e \leq 2\rho_0$ and $\sigma_e(s_0 - v_0) = |s_0 - v_0|_e \leq 2\rho_0$. Therefore,

$$2\rho_0 e - (x_0 - u_0) \in K, \quad -2\rho_0 F'(e) - (s_0 - v_0) \in K^*, \quad \text{and } \langle s_0, x_0 \rangle = \rho_0^2 \nu. \quad (4.61)$$

Let us assume that, for some constant $\Psi > 0$,

$$\rho_0 \geq \frac{1}{\Psi} \rho^* := \frac{1}{\Psi} \min\{\max(|x^*|_e, |s^*|_e) : (x^*, s^*) \text{ solves } (P) \text{ and } (D)\}. \quad (4.62)$$

(Note that we can always increase ρ_0 .) Now we can obtain a bound for ω . From (4.26) we recall that

$$\begin{aligned} \bar{\xi} &= \sqrt{\theta_\infty \nu} \left\{ 1 + \frac{\langle s^*, x_0 - u_0 \rangle + \langle s_0 - v_0, x^* \rangle + \langle s_0 - v_0, x_0 - u_0 \rangle}{\langle s_0, x_0 \rangle} \right\} \\ &\leq \sqrt{\theta_\infty \nu} \left\{ 1 + \frac{2\rho_0 \rho^* \nu + 2\rho_0 \rho^* \nu + 4\rho_0^2 \nu}{\rho_0^2 \nu} \right\} \quad (\text{using (4.61)}) \\ &= \sqrt{\theta_\infty \nu} \left\{ 1 + \frac{4(\rho^* + \rho_0)}{\rho_0} \right\} = \sqrt{\theta_\infty \nu} (4\Psi + 5) \quad (\text{using (4.62)}). \end{aligned}$$

We use (4.61) to bound $\bar{\chi}$ (see (4.25)), which is used subsequently to bound ω (see (4.56)) to get

$$\bar{\chi} = \beta_1^2 \frac{\theta_G}{\nu} + (1 - \beta_1)^2 + 2 \left\{ \frac{\langle s_0 - v_0, x_0 - u_0 \rangle}{\langle s_0, x_0 \rangle} \right\} \leq \frac{\theta_G}{\nu} + 1 + 2 \cdot \frac{4\rho_0^2 \nu}{\rho_0^2 \nu} = \frac{\theta_G}{\nu} + 9 \text{ and}$$

$$\omega = \max \left((\bar{\xi} + \sqrt{\bar{\xi}^2 + \bar{\chi}})^2, \frac{32\beta_1 \theta_G}{\nu} \right) = O(\theta_\infty \nu).$$

Hence

$$\tau = O(\omega\nu^{1.5}) = O(\nu^{2.5}).$$

Substituting the expression in (4.59), we get

$$\alpha^* = \min(1, \bar{\alpha}_1, \bar{\alpha}_2, \bar{\alpha}_3, \bar{\alpha}_4) = \min(1, \Omega(\nu^{-1}), \Omega(\nu^{-2.5}), \Omega(\nu^{-1}), \Omega(\nu^{-1})). \quad (4.63)$$

Theorem 4.8 *Given (A, b, c, K, K^*) and $\beta_1, \theta_G, \epsilon^* > 0$, let us choose x_0 and y_0 as in (4.60), where (4.62) holds. Then Algorithm 2 will produce a solution (x^*, y^*, s^*) such that $\langle s^*, x^* \rangle \leq \epsilon^* \langle s_0, x_0 \rangle$ and $\bar{\phi}^* \leq \epsilon^*$ in $O(\nu^{2.5} \ln(\frac{1}{\epsilon^*}))$ iterations.*

Proof : If we choose α^* in (4.63) as the step length at each iterate, then all the conditions in Step 3 of Algorithm 2 are satisfied. Thus $\bar{\alpha}_k \geq \alpha^*$ for each k , so that the complementarity is reduced by at

least the factor $(1 - (1 - \beta_2)\alpha^*)$ at each iteration. So for $k = \left\lceil \frac{1}{(1-\beta_2)\alpha^*} \ln\left(\frac{1}{\epsilon^*}\right) \right\rceil = O\left(\nu^{2.5} \ln\left(\frac{1}{\epsilon^*}\right)\right)$, we have

$$\begin{aligned}
\ln(\langle s_k, x_k \rangle) &\leq \ln(\langle s_{k-1}, x_{k-1} \rangle (1 - \alpha^*(1 - \beta_2))) \\
&\leq \ln\left(\langle s_0, x_0 \rangle (1 - \alpha^*(1 - \beta_2))^k\right) \\
&= \ln(\langle s_0, x_0 \rangle) + k \ln(1 - \alpha^*(1 - \beta_2)) \\
&\leq \ln(\langle s_0, x_0 \rangle) - k\alpha^*(1 - \beta_2) \\
&\leq \ln(\langle s_0, x_0 \rangle) + \ln(\epsilon^*) = \ln(\epsilon^* \langle s_0, x_0 \rangle).
\end{aligned}$$

The first inequality follows from the decrease in total complementarity condition, the second from the same applied inductively, and the third inequality from the identity $\ln(1+x) \leq x$ for all $x > -1$. The fourth inequality follows from our assumption on k .

From condition (3.4) it follows that $\phi_k \leq \frac{\langle s_k, x_k \rangle}{\langle s_0, x_0 \rangle} \leq \epsilon^*$. From (4.5), it follows that

$$\|Ax_k - b\| \leq \epsilon^* \|Ax_0 - b\|, \text{ and } \|A^*y_k + s_k - c\| \leq \epsilon^* \|A^*y_0 + s_0 - c\|.$$

□

By the second part of Lemma 4.1, we can see that strict feasibility in both primal and dual implies that all the iterates are bounded. As the result was independent of the neighborhood, this implies that there exists a Ω^* large enough that (3.5) in Section 3 will always hold. However, if such a Ω^* did not exist or is very large, then we would like some inference on the infeasibility of the primal-dual pair. We will also see the relevance of ξ and condition (4.62) in producing infeasibility indicators. Todd and Ye [17] provide some guarantees on the norms of optimal as well as feasible solutions for linear programming. We closely follow their approach and obtain analogous results in this extended setting of self-scaled conic programs. Let (x^*, y^*, s^*) denote an optimal solution to the pair (P) and (D) , if one exists.

Now, we will directly relate ξ (which contributes to the complexity estimate through ω) defined in (4.18) to indicators of infeasibility. Recall that $\bar{\phi} = \max(\phi_p, \phi_d)$ and $\underline{\phi} := \min(\phi_p, \phi_d)$. Let

$$\rho := \max(|x_0 - u_0|_e, |s_0 - v_0|_e). \quad (4.64)$$

If $\max(|x^*|_e, |s^*|_e) \leq \tilde{\rho}$, then following the proof of Lemma 4.1, we have

$$\begin{aligned}
\frac{\langle s, x - u \rangle + \langle s - v, x \rangle}{\langle s, x \rangle} &= \frac{\phi_p \langle s, x_0 - u_0 \rangle + \phi_d \langle s_0 - v_0, x \rangle}{\langle s, x \rangle} \\
&< 1 + \frac{\langle s^*, x_0 - u_0 \rangle + \langle s_0 - v_0, x^* \rangle + \underline{\phi} \langle s_0 - v_0, x_0 - u_0 \rangle}{\langle s_0, x_0 \rangle} \\
&\leq 1 + \frac{2\rho\tilde{\rho}\nu + \underline{\phi}\rho^2\nu}{\rho_0^2\nu} = 1 + \frac{\rho(2\tilde{\rho} + \underline{\phi}\rho)}{\rho_0^2}.
\end{aligned}$$

In light of our discussion above, we can use the following condition as a stopping rule.

- Stopping Rule 1. For some $\tilde{\rho}$, stop if

$$\frac{\phi_p \langle s, x_0 - u_0 \rangle + \phi_d \langle s_0 - v_0, x \rangle}{\langle s, x \rangle} \geq \left(1 + \frac{\rho(2\tilde{\rho} + \underline{\phi}\rho)}{\rho_0^2}\right).$$

Theorem 4.9 *If stopping rule 1 applies, then there is no optimal solution pair x^* and (y^*, s^*) for (P) and (D) with $|x^*|_e \leq \tilde{\rho}$ and $|s^*|_e \leq \tilde{\rho}$. \square*

We would like to show that for some Ω_* , under certain assumptions, violation of condition (3.5) does lead to some conclusions on the size of optimal solutions and the size of feasible solutions (as we shall see in Theorem 4.13). Assume that the primal is strictly infeasible, i.e., the alternative Farkas system is strictly feasible; then it can be shown that the sequence of iterates $\{x_k\}$ is bounded by some constant $M > 0$. So, suppose that $\phi_p \geq \tilde{\phi} > 0$, and $\Omega_* = M + \frac{\nu|e|(x_0-u_0)}{\tilde{\phi}} (\rho_0^2 + 2\tilde{\rho}\rho + \underline{\phi}\rho^2)$, so that $\|x\| + \|s\|_* \geq \Omega_*$ implies that $\|s\|_* \geq \frac{\nu|e|(x_0-u_0)}{\tilde{\phi}} (\rho_0^2 + 2\tilde{\rho}\rho + \underline{\phi}\rho^2)$. In this case, from

$$\phi_p \langle s, x_0 - u_0 \rangle + \phi_d \langle s_0 - v_0, x \rangle \geq \phi_p \langle s, x_0 - u_0 \rangle \geq \frac{\tilde{\phi}}{|e|(x_0-u_0)} \langle s, e \rangle \geq \frac{\tilde{\phi}}{|e|(x_0-u_0)} \|s\|_*,$$

if we substitute the lower bound on $\|s\|_*$ and use $\rho_0^2\nu \geq \langle s, x \rangle$, we get

$$\phi_p \langle s, x_0 - u_0 \rangle + \phi_d \langle s_0 - v_0, x \rangle \geq \nu (\rho_0^2 + 2\tilde{\rho}\rho + \underline{\phi}\rho^2) \geq \left(1 + \frac{\rho(2\tilde{\rho} + \underline{\phi}\rho)}{\rho_0^2}\right) \langle s, x \rangle,$$

so that stopping rule 1 applies. Note also that condition (4.62) implies a bound on $\frac{\phi_p \langle s, x_0 - u_0 \rangle + \phi_d \langle s_0 - v_0, x \rangle}{\langle s, x \rangle}$, because $\left(1 + \frac{\rho(2\tilde{\rho} + \underline{\phi}\rho)}{\rho_0^2}\right) \leq 5 + \frac{4\rho^*}{\rho_0} \leq 5 + 4\Psi$. The inequality can be seen using $\underline{\phi} \leq 1$, $\rho \leq 2\rho_0$ and condition (4.62).

Next we investigate stopping rules that provide lower bounds on the size of any feasible solutions. We use the following result:

Lemma 4.10 *Let*

$$\begin{aligned} \alpha_x &:= \min_x \{\|x\| : Ax = b, x \in K\}, \\ \alpha_y &:= \min_{y,s} \{\|y\| : A^*y + s = c, s \in K^*\}, \\ \alpha_s &:= \min_{y,s} \{\|s\|_* : A^*y + s = c, s \in K^*\}, \text{ and} \\ \alpha_{ys} &:= \min_{y,s} \{\|y\| + \|s\|_* : A^*y + s = c, s \in K^*\}. \end{aligned}$$

Let

$$\begin{aligned} \beta_c &:= \min_{\hat{c}, y} \{\|\hat{c}\|_* : \hat{c} - A^*y \in K^*, \langle b, y \rangle = 1\}, \\ \beta_b &:= \min_{\hat{b}, x} \{\|\hat{b}\|_* : Ax = \hat{b}, \langle c, x \rangle = -1, x \in K\} \\ \beta_w &:= \min_{w, x} \{\|w\| : Ax = 0, \langle c, x \rangle = -1, x + w \in K\}, \text{ and} \\ \beta_{bw} &:= \min_{\hat{b}, x, w} \{\max(\|\hat{b}\|_*, \|w\|) : Ax = \hat{b}, \langle c, x \rangle = -1, x + w \in K\}. \end{aligned}$$

Then, $\alpha_x\beta_c = \alpha_y\beta_b = \alpha_s\beta_w = \alpha_{ys}\beta_{bw} = 1$.

Proof : The proof of the first three relations follow directly by applying Lemma 3.13 in Renegar [13]. The fourth follows by using the relation between α_x and β_c with linear operators $[A^* \ I] : (y, s) \mapsto A^*y + s$ replacing A , cone $Y \times K^*$ replacing K and right-hand side c instead of b . Note that the norm $\max(\|\tilde{b}\|_*, \|\tilde{w}\|)$ on $Y^* \times E$ is dual to the norm $\|y\| + \|s\|_*$ on $Y \times E^*$. Then as a counterpart for the problem defining α_{ys} , we get

$$\beta_{bw} = \min_{\tilde{b}, \tilde{x}, \tilde{w}} \{ \max(\|\tilde{b}\|_*, \|\tilde{w}\|) : \begin{pmatrix} \tilde{b} \\ \tilde{w} \end{pmatrix} - \begin{pmatrix} A \\ I \end{pmatrix} \tilde{x} \in \begin{pmatrix} Y_d^* \\ (K^*)^* \end{pmatrix}, \langle c, \tilde{x} \rangle = 1 \},$$

where $Y_d^* = \{0\}$ is the dual cone of Y and $(K^*)^* = K$. Now changing variables to $\hat{b} = -\tilde{b}$, $x = -\tilde{x}$ and $\hat{w} = \tilde{w}$, we obtain the desired form for the problem defining β_{bw} . \square

It is straightforward to verify that

$$u_0 := F''(e)^{-1}A^*(AF''(e)^{-1}A^*)^{-1}b, \quad r_0 := (AF''(e)^{-1}A^*)^{-1}AF''(e)^{-1}c, \quad \text{and } v_0 := c - A^*r_0.$$

(We can observe here that $\langle c, u_0 \rangle = \langle b, r_0 \rangle$.) Consider the following stopping rules.

- Stopping Rule 2_p . Let $r = y - \phi_d(y_0 - r_0)$. Then, for some $\bar{\rho}_p > 0$, stop if

$$\langle b, r \rangle \geq \|c + \phi_d(s_0 - v_0)\|_* \bar{\rho}_p.$$

- Stopping Rule 2_d . Let $u = x - \phi_p(x_0 - u_0)$. Then, for some $\bar{\rho}_d > 0$, stop if

$$\langle c, u \rangle \leq -\max(\|\hat{b}\|_*, \phi_p\|x_0 - u_0\|) \bar{\rho}_d.$$

The following theorem establishes lower bounds on the norms of x in the primal space and (y, s) in the dual space.

Theorem 4.11 *If stopping rule 2_p applies, then any feasible solution to (P) has norm at least $\bar{\rho}_p$; if stopping rule 2_d applies, then any feasible solution to (D) has $\|y\| + \|s\|_*$ at least $\bar{\rho}_d$.*

Proof : If we let $\tilde{y} = \frac{r}{\langle b, r \rangle}$ and $\hat{c} = \frac{c + \phi_d(s_0 - v_0)}{\langle b, r \rangle}$ in Lemma 4.10, we will get that $\beta_c \leq \|\hat{c}\|_*$. Therefore, $\alpha_x \geq \frac{1}{\|\hat{c}\|_*} = \frac{\langle b, r \rangle}{\|c + \phi_d(s_0 - v_0)\|_*} \geq \bar{\rho}_p$, proving the first part. Next note that if we let $\tilde{x} = \frac{u}{-\langle c, u \rangle}$ and $w = \frac{x - u}{-\langle c, u \rangle}$, then $\tilde{x} + w = \frac{x}{-\langle c, u \rangle} \in K$. Now, the second part follows by using the relation between β_{bw} and α_{ys} in Lemma 4.10. \square

The following modified stopping rule is analogous to the rule proposed in [17] and this is symmetric between x and s .

- Stopping Rule $2'_p$. Let $\hat{y} = y - [r_0 + \phi_d(y_0 - r_0)]$. Then, for some $\bar{\rho}_p > 0$, stop if

$$\langle b, \hat{y} \rangle \geq \|v_0 + \phi_d(s_0 - v_0)\|_* \bar{\rho}_p.$$

- Stopping Rule $2'_d$. Let $\hat{x} = x - [u_0 + \phi_p(x_0 - u_0)]$. Then, for some $\bar{\rho}_d > 0$, stop if

$$\langle c, \hat{x} \rangle \leq -\|u_0 + \phi_p(x_0 - u_0)\| \bar{\rho}_d.$$

The proof of the following theorem follows verbatim from that of Proposition 4.2 in [17] using Lemma 4.10 in place of Lemma 2.1 in [17].

Theorem 4.12 *If stopping rule $2'_p$ applies, then any feasible solution to (P) has norm at least $\bar{\rho}_p$; if stopping rule $2'_d$ applies, then any feasible solution to (D) has $\|s\|_*$ at least $\bar{\rho}_d$. \square*

The following theorem shows that if $\tilde{\rho}$ is sufficiently large, whenever stopping rule 1 applies, so does stopping rule 2_p or 2_d (alternatively, stopping rule $2'_p$ or $2'_d$). This result provides a sharper lower bound than that provided in [17], and as should be no surprise, we can see that ϕ does not appear in the denominator of the lower bound. As a consequence, if for instance the primal is infeasible while the dual is feasible, the dual is not restricted from attaining feasibility.

Theorem 4.13 *If*

$$\tilde{\rho} \geq \frac{1}{2\rho\bar{\phi}\nu} [\|c + \phi_d(s_0 - v_0)\|_* \bar{\rho}_p + \max(\|b\|_*, \phi_p\|x_0 - u_0\|) \bar{\rho}_d],$$

then if stopping rule 1 applies, so does either 2_p or 2_d . If

$$\tilde{\rho} \geq \frac{1}{2\rho\bar{\phi}\nu} [\|v_0 + \phi_d(s_0 - v_0)\|_* \bar{\rho}_p + \|u_0 + \phi_p(x_0 - u_0)\| \bar{\rho}_d],$$

then if stopping rule 1 applies, so does either $2'_p$ or $2'_d$.

Proof : Let us first note the following identities:

$$\begin{aligned} Au &= b, & A^*r + s &= c + \phi_d(s_0 - v_0), \\ u &= x - \phi_p(x_0 - u_0), \text{ and} & r &= y - \phi_d(y_0 - r_0). \end{aligned}$$

Hence using a weak-duality like relation it follows that

$$\langle c, u \rangle = \langle c + \phi_d(s_0 - v_0), u \rangle - \langle b, r \rangle = \langle c, u \rangle - \langle b, r \rangle + \langle \phi_d(s_0 - v_0), u \rangle.$$

We first substitute for u and r in the above identity, then use the condition in stopping rule 1 and rewrite $\phi_p\phi_d$ as $\bar{\phi}\underline{\phi}$, and finally use (4.64), $\langle s_0, x_0 \rangle = \rho_0^2\nu$ and $\bar{\phi}\langle s_0, x_0 \rangle \leq \langle s, x \rangle$ to get

$$\begin{aligned} \langle c, u \rangle - \langle b, r \rangle &= \langle s, u \rangle - \langle \phi_d(s_0 - v_0), u \rangle \\ &= \langle s, x - \phi_p(x_0 - u_0) \rangle - \langle \phi_d(s_0 - v_0), x - \phi_p(x_0 - u_0) \rangle \\ &= \langle s, x \rangle - \phi_d \langle s_0 - v_0, x \rangle - \phi_p \langle s, x_0 - u_0 \rangle + \phi_p\phi_d \langle s_0 - v_0, x_0 - u_0 \rangle \\ &\leq \langle s, x \rangle \left[1 - 1 - \frac{\rho(2\tilde{\rho} + \underline{\phi}\rho)}{\rho_0^2} \right] + \underline{\phi} \frac{\langle s_0 - v_0, x_0 - u_0 \rangle}{\langle s_0, x_0 \rangle} \bar{\phi} \langle s_0, x_0 \rangle \\ &\leq \langle s, x \rangle \left[-\frac{2\rho\tilde{\rho}}{\rho_0^2} - \frac{\underline{\phi}\rho^2}{\rho_0^2} \right] + \underline{\phi} \frac{\rho^2}{\rho_0^2} \langle s, x \rangle \\ &\leq \bar{\phi}\rho_0^2\nu \left[-\frac{2\rho\tilde{\rho}}{\rho_0^2} \right] = -2\bar{\phi}\nu\rho\tilde{\rho}. \end{aligned}$$

If termination criteria of both the stopping rules 2_p and 2_d did not apply, then

$$\langle c, u \rangle - \langle b, r \rangle \geq -[\|c + \phi_d(s_0 - v_0)\|_* \bar{\rho}_p + \max(\|b\|_*, \phi_p\|x_0 - u_0\|) \bar{\rho}_d].$$

Therefore, if we choose

$$\tilde{\rho} \geq \frac{1}{2\rho\bar{\phi}\nu} [\|c + \phi_d(s_0 - v_0)\|_* \bar{\rho}_p + \max(\|b\|_*, \phi_p\|x_0 - u_0\|) \bar{\rho}_d],$$

then one of the stopping criteria must hold, otherwise it will contradict the lower bound on the difference $\langle c, u \rangle - \langle b, r \rangle$. This completes the proof of the first part.

Furthermore, note that $\hat{x} = u - u_0$, $\hat{y} = r - r_0$ and we have already observed that $\langle c, u_0 \rangle = \langle b, r_0 \rangle$. Consequently, $\langle c, \hat{x} \rangle - \langle b, \hat{y} \rangle = \langle c, u \rangle - \langle b, r \rangle$. Hence, by a similar argument, for the choice of $\tilde{\rho}$ stated in the hypothesis, it is seen that if stopping rule 1 applies, then so does either $2'_p$ or $2'_d$. \square

5 Conclusion

We have established global convergence of Algorithm 1 and polynomial iteration complexity of Algorithm 2. In Section 4, we placed restrictions on our initial points so that $x_0 - u_0 \in \text{int } K$ and $s_0 - v_0 \in \text{int } K^*$. There have been global convergence results for linear programming using arbitrary infeasible starting points [1]. It would be interesting to see if such results also extend in this setting.

Complexity estimates of algorithms for linear programming problems have normally been derived using the γ_∞ -neighborhood, because this is a reasonable approximation to the 99%-of-the-way scheme used in many practical implementations. If we are given θ and the γ_∞ -neighborhood $\mathcal{N}_\infty(\theta)$, we could relax it to a bigger neighborhood $\mathcal{N}_G(\theta_G)$ by choosing $\theta_G = \nu\theta$ and $\theta_\infty = \nu\theta + 2$. This gives us a complexity estimate of $O(\nu^4)$, larger for example than the $O(\nu^{2.5})$ bound using the γ_∞ -neighborhood for semidefinite programming in [19]. However, we must note that the neighborhood has also gotten bigger and is a better approximation to the 99%-of-the-way scheme. Moreover, our complexity analysis contains several approximations, and it is quite possible that a different analysis would yield a tighter estimate.

We can make modifications to the algorithms presented to implement them in practice without losing the convergence guarantees. The α^* in both the algorithms are hard to compute as we do not know η in Algorithm 1 or ω in Algorithm 2 beforehand. The step length at each iteration can be computed by replacing η by $\max\left(2(\|\Delta x\|_x^2 + \|\Delta s\|_s^2), \frac{2\nu|\langle \Delta s, \Delta x \rangle|}{\langle s, x \rangle}\right)$ in Algorithm 1 and ω by $\max\left(\frac{(\|\Delta x\|_x^2 + \|\Delta s\|_s^2)}{\nu\theta_\infty}, \frac{2|\langle \Delta s, \Delta x \rangle|}{\langle s, x \rangle}\right)$ in Algorithm 2. We can also obtain step lengths using local (or binary) search satisfying the conditions in Step 3, to improve the practical performance of the methods based on γ_F - and γ_G -neighborhoods. For example, as long as the step length $\tilde{\alpha}_k$ satisfies all the conditions in Step 3, but $2\tilde{\alpha}_k$ fails at least one condition, then we know that $\tilde{\alpha}_k \geq \bar{\alpha}_k/2 \geq \alpha^*/2$.

Finally, we presented results pertaining to lower bounds on optimal and feasible solutions when certain stopping rules apply and also related them to the termination condition (3.5) in Algorithm 1 and condition (4.62) in Algorithm 2.

References

- [1] S. C. Billups and M. C. Ferris. Convergence of an infeasible interior-point algorithm from arbitrary positive starting points. *SIAM Journal on Optimization*, 6:316–326, 1996.
- [2] T. J. Carpenter, I. J. Lustig, J. M. Mulvey, and D. F. Shanno. Higher order predictor–corrector interior point methods with application to quadratic objectives. *SIAM Journal on Optimization*, 3:696–725, 1993.
- [3] N. Karmarkar, A new polynomial–time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.
- [4] M. Kojima, N. Megiddo and S. Mizuno, A primal–dual infeasible–interior–point algorithm for linear programming. *Mathematical Programming*, 61:263–280, 1993.
- [5] S. Mizuno, Polynomiality of infeasible–interior–point algorithms for linear programming. *Mathematical Programming*, 67:109–119, 1994.
- [6] R. D. C. Monteiro, Primal–dual path following algorithms for semidefinite programming. *SIAM Journal on Optimization*, 7(3):663–678, 1997.
- [7] Yu. E. Nesterov, A. S. Nemirovskii, Interior-Point Polynomial Algorithms in Convex Programming. *SIAM, Philadelphia*, 1994.
- [8] Yu. E. Nesterov, M. J. Todd, Self-Scaled barriers and interior point methods for convex programming. *Mathematics of Operations Research*, 22(1):1–42, 1997.
- [9] Yu. E. Nesterov, M. J. Todd, Primal-dual interior-point methods for self-scaled cones. *SIAM Journal on Optimization*, 8(2):324–364, 1998.
- [10] Yu. E. Nesterov, M. J. Todd, Y. Ye, Infeasible-start primal-dual methods and infeasibility detectors for nonlinear programming Problems. *Mathematical Programming*, 84:227–267, 1999.
- [11] F. A. Potra, A quadratically convergent predictor-corrector method for linear programs from infeasible starting points. *Mathematical Programming*, 67(3):383–406, 1994.
- [12] F. A. Potra, An infeasible interior-point predictor-corrector method algorithm for linear programming. *SIAM Journal on Optimization*, 6(1):19–32, 1996.
- [13] J. Renegar, Some perturbation theory for linear programming. *Mathematical Programming*, 65:73–82, 1994.
- [14] J. Renegar, A Mathematical View of Interior-Point Methods in Convex Optimization. *MPS/SIAM Series on Optimization*, 2001.
- [15] S. H. Schmieta, F. Alizadeh, Extension of primal–dual interior–point algorithm to symmetric cones. *Mathematical Programming*, 96(3):409–438, 2003.
- [16] K. C. Toh, M. J. Todd, R. Tutuncü, On the Nesterov-Todd direction in semidefinite programming. *SIAM Journal on Optimization*, 8:769–796, 1998.

- [17] M. J. Todd, Y. Ye, Approximate Farkas lemmas and stopping rules for iterative infeasible-point algorithms for linear programming. *Mathematical Programming*, 81:1-21, 1998.
- [18] Y. Zhang, On the convergence of a class of infeasible interior-point methods for the horizontal linear complementarity problem. *SIAM Journal on Optimization*, 4(1):208-227, 1994.
- [19] Y. Zhang, On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming. *SIAM Journal on Optimization*, 8(1):365-386, 1998.