A primal–dual symmetric relaxation for homogeneous conic systems

Juan Carlos Vera*, Juan Carlos Rivera*, Javier Peña†

April 27, 2002

Abstract

We address the feasibility of the pair of alternative conic systems of constraints

$$Ax = 0,\ x \in C$$

and

$$-A^T y \in C^*,$$

where $A \in \mathbb{R}^{m\times n}$, $m < n$, and $C \subseteq \mathbb{R}^n$ is a closed convex cone. We reformulate this pair of conic systems as a primal-dual pair of conic programs. Each of the conic programs corresponds to a natural relaxation of each of the two conic systems.

When C is a self-scaled cone with a known self-scaled barrier, the conic programming reformulation can be solved via interior-point methods. For a well-posed instance A, a strict solution to one of the two original conic systems can be obtained in $O(\sqrt{\nu_C}\log(\nu_C\,\mathcal{C}(A))$ interior-point iterations. Here ν_C is the complexity parameter of the self-scaled barrier of C and $\mathcal{C}(A)$ is Renegar's condition number of A, that is, the reciprocal of the relative distance from A to the set of ill-posed instances.

GSIA Working Paper 2002-06

*Department of Mathematical Sciences, Carnegie Mellon University. E-mail: jvera@andrew.cmu.edu and juanr@andrew.cmu.edu

†Graduate School of Industrial Administration, Carnegie Mellon University. E-mail: jfp@andrew.cmu.edu. Supported by NSF grant CCR-0092655.

1 Introduction

We study the conic feasibility problems

$$Ax = 0, \; x \in C \tag{1}$$

and

$$-A^T y \in C^*, \tag{2}$$

where $A \in \mathbb{R}^{m \times n}$ and $C \subseteq \mathbb{R}^n$ is a closed convex cone.

The conic systems (1), (2) are essentially *alternative* problems: (1) is well-posed feasible (see the discussion below) if and only if (2) does not have nonzero solutions.

We devise a primal-dual scheme that determines which one of these systems is feasible and generates a strictly feasible solution, provided the data A is well-posed. Our scheme does not use, nor does it assume any knowledge about the problem other than the data instance A.

Some of our ideas are inspired by previous work by the third author and Renegar [9], and by the third author and Cucker [1]. In the former, a purely *primal* relaxation scheme for solving (1) was proposed and analyzed. In the latter, the authors devised a primal-dual scheme for solving (1) or (2) for the particular case when $C = \mathbb{R}^n_+$, using a finite precision machine.

This paper combines and extends both of these previous works. On the one hand, unlike the primal approach used in [9], we reformulate the feasibility problem as a primal-dual pair of conic programs. As a nice consequence of this primal-dual approach, *both* (1) *and* (2) are treated in a unified manner, without any a priori feasibility assumption of either system. On the other hand, we extend some of the key ideas introduced in [1] to general cones. We must stress, however, that the whole treatment in [1] *cannot* be obtained as a particular case of the results in this paper. In particular, the more abstract context we address prevents us from studying the finite precision issues at the level of detail done in [1].

One of the most interesting features of our approach is that the amount of work needed to obtain a feasible solution for either (1) or (2) depends naturally on Renegar's condition number of the data instance A. This establishes a natural parallel between the feasibility problem for conic systems and the solution of linear systems of equations via iterative methods. In both cases the condition number is a key parameter in the performance of the algorithm: the amount of work necessary to solve a problem instance is proportional to the condition number of the instance. This follows a natural paradigm in numerical analysis.

Several of our results are generalizations of results previously derived in [1] and [9]. However, the proof techniques in this paper are new. They rely on an implicitly-defined function framework developed in [7]. In our opinion, the new approach is substantially more concise, insightful, and transparent than the previous approaches. To ease our presentation, we develop the main technical ideas in the last two sections of the paper.

The paper has been organized as follows. The remainder of this introduction reviews Renegar's condition number and the closely related notion of *distance to ill-posedness*. Section 2 develops our central ideas: First, we recast the pair (1), (2) as the primal-dual pair of conic programs (9), (10). Second, assuming C is a self-scaled cone, we establish a close connection between the central path of the primal-dual pair (9), (10) and the pair of alternative conic systems (1), (2) (see Propositions 2.3 and 2.4). In Section 3 we describe an interior-point algorithm that computes a strict solution to whichever of (1), (2) is feasible, provided the data A is well-posed. The algorithm finds such a solution within $O(\sqrt{\nu_C}\log(\nu_C\,\mathcal{C}(A))$ interior-point iterations, where ν_C is the complexity parameter of barrier for the cone C, and $\mathcal{C}(A)$ is the condition number of the data instance A. The proofs of the central results in Sections 2 and 3, namely Propositions 2.3, 2.4, 3.4, and 3.5 are presented in Sections 4 and 5.

Let us review the basic definitions and properties of Renegar's condition number and distance to ill-posedness (see [6, 8] for a detailed discussion on these concepts). We say that (1) is a *well-posed feasible system* if

$$\{Ax : x \in C\} = \mathbb{R}^m. \tag{3}$$

Let $\mathcal{P}$ be the set of $m \times n$ matrices A such that (3) holds. Notice that $A \in \mathcal{P}$ if and only if the alternative system (2) does not have nonzero solutions.

We say that (2) is a *well-posed feasible system* if

$$\{A^{\mathrm{T}}y : y \in \mathbb{R}^m\} + C^* = \mathbb{R}^n. \tag{4}$$

Let $\mathcal{D}$ be the set of $m \times n$ matrices A such that (4) holds. Notice that $A \in \mathcal{D}$ if and only if the alternative system (1) does not have nonzero solutions.

Endow the space $\mathbb{R}^{m\times n}$ with the operator norm. It can be shown that both $\mathcal{P}$ and $\mathcal{D}$ are open subsets of $\mathbb{R}^{m\times n}$. The set $\mathbb{R}^{m\times n} \setminus (\mathcal{P} \cup \mathcal{D})$ is the set of *ill-posed* instances. It is easy to show that this set has Lebesgue measure equal to zero. Furthermore, if $m < n$ and C is a regular cone (i.e., both C and C^* have nonempty interiors), then the closure of either $\mathcal{P}$ or $\mathcal{D}$ in $\mathbb{R}^{m\times n}$ is the complement of the other.

The *distance to infeasibility* of (1) is defined as

$$\rho_P(A) := \inf\{\|\Delta A\| : A + \Delta A \notin \mathcal{P}\}.$$

Likewise, the distance to infeasibility of (2) is defined as

$$\rho_D(A) := \inf\{\|\Delta A\| : A + \Delta A \notin \mathcal{D}\}.$$

The *distance to ill-posedness* of A is

$$\rho(A) := \inf\{\|\Delta A\| : A + \Delta A \notin \mathcal{P} \cup \mathcal{D}\} = \max\{\rho_P(A), \rho_D(A)\}.$$

The data instance A is *well-posed* if $\rho(A) > 0$, i.e., if $A \in \mathcal{P} \cup \mathcal{D}$. Renegar's condition number $\mathcal{C}(A)$ is defined as the reciprocal of the relative distance to ill-posedness, i.e.,

$$\mathcal{C}(A) := \frac{\|A\|}{\rho(A)}.$$

Our treatment will crucially rely on the following characterizations of the distance to infeasibility. For a detailed discussion on this and closely related issues see [6, 8, 10].

Proposition 1.1 (Renegar) *For any given* $A \in \mathrm{IR}^{m\times n}$,

$$\rho_P(A) = \sup\left\{\delta : \|v\| \le \delta \Rightarrow v \in \{Ax : \|x\| \le 1,\, x \in C\}\right\},$$

and

$$\rho_D(A) = \sup\left\{\delta : \|u\| \le \delta \Rightarrow u \in \{A^{\mathrm{T}}y : \|y\| \le 1\} + C^*\right\}.$$

2 Reformulation

The following reformulation scheme is a generalization of the reformulations proposed in [1] and [9]. Recast (1) as

$$\begin{array}{ll} \min & \|x''\| \\ \text{s.t.} & Ax + x'' = 0 \\ & x \in C \\ & \|x\| \le 1, \end{array} \tag{5}$$

and recast (2) as

$$\begin{array}{ll} \min & \|y'\| \\ \text{s.t.} & -A^{\mathrm{T}}y + y' \in C^* \\ & \|y\| \le 1. \end{array} \tag{6}$$

By introducing auxiliary variables, the problems (5) and (6) are equivalent to the pair

$$\begin{array}{ll} \min & \tau \\ \text{s.t.} & Ax + x'' = 0 \\ & -x + x' = 0 \\ & t_1 = 1 \\ & x \in C \\ & \|x'\| \le t_1 \\ & \|x''\| \le \tau, \end{array} \tag{7}$$

and

$$
\begin{array}{ll}
\min & \eta \\
\text{s.t.} & -A^{\mathrm{T}}y + y' \in C^* \\
& \|y\| \leq 1 \\
& \|y'\| \leq \eta.
\end{array} \tag{8}
$$

When cast appropriately, the pair (7), (8) is a primal-dual pair of conic programs in a higher dimensional space: Let $K := C \times \mathcal{K}_{n+1} \times \mathcal{K}_{m+1}$, where $\mathcal{K}_{n+1}$, $\mathcal{K}_{m+1}$ are second-order cones in $\mathbb{R}^{n+1}$, $\mathbb{R}^{m+1}$ respectively. The problem (7) can be written as

$$
\begin{array}{ll}
\min & \langle \vec{c}, \vec{x} \rangle \\
\text{s.t.} & \mathcal{A}\vec{x} = \vec{b} \\
& \vec{x} \in K,
\end{array} \tag{9}
$$

where $\vec{x} = (x, x', t_1, x'', \tau) \in \mathbb{R}^{(m+2n+2)}$, and $\mathcal{A} \in \mathbb{R}^{(m+n+1)\times(m+2n+2)}$, $\vec{c} \in \mathbb{R}^{(m+2n+2)}$, $\vec{b} \in \mathbb{R}^{(m+n+1)}$ are as follows

$$
\mathcal{A} := \begin{bmatrix} A & 0 & 0 & I_m & 0 \\ -I_n & I_n & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \; \vec{b} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \; \vec{c} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.
$$

Now the problem (8) corresponds precisely to the dual of (9), namely

$$
\begin{array}{ll}
\max & \langle \vec{b}, \vec{y} \rangle \\
\text{s.t.} & \mathcal{A}^{\mathrm{T}}\vec{y} + \vec{s} = \vec{c} \\
& \vec{s} \in K^*.
\end{array} \tag{10}
$$

where $\vec{y} = (y, y', -\eta) \in \mathbb{R}^{(m+n+1)}$.

It is obvious that at the optimal solutions to (9), (10) the variables x'' and y' are zero. It is also intuitively clear that an interior-point algorithm applied to (9), (10) would yield, in the limit, a strict solution for whichever of (1), (2) that is strictly feasible. The main goal of this paper is to formalize and make this idea more precise. Specifically, two of our key results, namely Propositions 2.3 and 2.4 below, establish a close connection between the central path of the primal-dual pair (9), (10) and the pair (1), (2) when C is self-scaled.

For the remainder of the paper we shall assume that C is self-scaled cone with self-scaled barrier f_C. (See [2, 4, 5, 11] for a detailed discussion on self-scaled cones.) It readily follows that the cone K is self-scaled as well with self-scaled barrier

$$
f_K(x, x', t_1, x'', \tau) = f_C(x) - \ln(t_1^2 - \|x'\|^2) - \ln(\tau^2 - \|x''\|^2).
$$

We also recall that as an immediate consequence of self-scaledness, both C and K are self-dual, that is, $C^* = C$ and $K^* = K$.

For a given barrier function f, we shall use g and H to denote the gradient and Hessian of f respectively. We will add the subindex C or K when we refer to the gradient and Hessian of f_C or f_K. We shall also let ν_f denote the barrier parameter of f. In such case we shall say that f is a ν_f-barrier function. We shall also let D_f denote the domain of f. For the particular functions f_C and f_K we shall abbreviate ν_{f_C} as ν_C, and ν_{f_K} as ν_K. By construction, we have $\nu_K = \nu_C + 4$.

Our development relies on the following key properties of barrier and self-scaled barrier functions. (For a detailed discussion on these properties see [3, 4, 5, 11].)

Proposition 2.1 *Let f be a ν-barrier function and $x \in D_f$. Then*

$$\{z : \langle z - x, H(x)(z - x)\rangle < 1\} \subseteq D_f, \tag{11}$$

$$\{z \in D_f : \langle z - x, g(x)\rangle \geq 0\} \subseteq \{z : \langle z - x, H(x)(z - x)\rangle \leq 4\nu + 1\}. \tag{12}$$

If f is self-scaled then $4\nu+1$ can be replaced by ν in (12), and also for all $x \in D_f$ and $t > 0$ the following identities hold

$$g(tx) = \frac{1}{t}g(x),\ H(tx) = \frac{1}{t^2}H(x), \tag{13}$$

$$-g(x) \in D_f,\ H(-g(x)) = H(x)^{-1}, \tag{14}$$

and

$$\|H(x)^{-1}\| \leq \|x\|^2. \tag{15}$$

Some of our statements are phrased in terms of the local inner product and local norm, which we now recall. Given a barrier function f and a point $x \in D_f$, the *local inner product* $\langle \cdot, \cdot \rangle_x$ induced by x is defined as

$$\langle u, v\rangle_x := \langle u, H(x)v\rangle.$$

The *local inner norm* $\|\cdot\|_x$ is defined as

$$\|v\|_x := \langle v, v\rangle_x^{1/2}.$$

Notice that the identities (11) and (12) can be rephrased in terms of the local norm induced by x.

Propositions 2.3 and 2.4 below formalize the intuitively clear fact that points on the central path of (9), (10) eventually yield solutions for whichever of (1), (2) that has strictly feasible solutions. First we recall the definition of the central path in a form that is suitable for our purposes. (For details see [4, 5, 11].)

Definition 2.2 The *central path* of (9), (10) is the set of solutions of the nonlinear system of equations

$$\begin{aligned} \mathcal{A}\vec{x} &= \vec{b} \\ \mathcal{A}^{\mathrm{T}}\vec{y} + \vec{s} &= \vec{c} \\ \vec{s} + \mu g_K(\vec{x}) &= 0, \end{aligned}$$

with $\vec{x}$, $\vec{s} \in \mathrm{int}(K)$ for all values of $\mu > 0$.

Proposition 2.3 *Let $(\vec{x}, \vec{y}, \vec{s})$ be on the central path of* (9), (10) *with $\vec{c}^{\mathrm{T}}\vec{x} = \tau$. If*

$$\tau < \frac{\rho_P(A)}{\sqrt{2}\nu_K},$$

then $\bar{x} := x + H_C(x)^{-1}A^T(AH_C(x)^{-1}A^T)^{-1}x''$ satisfies

$$A\bar{x} = 0,\ \bar{x} \in \mathrm{int}(C),\ \|\bar{x} - x\|_x \le \frac{\sqrt{2}\nu_K\tau}{\rho_P(A)}.$$

Proof. See Section 3.1. □

Proposition 2.4 *Let $(\vec{x}, \vec{y}, \vec{s})$ be on the central path of* (9), (10) *with $\vec{b}^{\mathrm{T}}\vec{y} = -\eta$. If*

$$\eta < \frac{\rho_D(A)}{\nu_K},$$

then y satisfies $-A^Ty \in \mathrm{int}(C)$.

Proof. See Section 3.1. □

3 Solving the conic pair via a primal-dual algorithm

The pair (9), (10) is amenable to the machinery of primal-dual interior-point methods for self-scaled cones (cf. [4, 5, 11, 12]). There are a number of different algorithms whose specific updates depend on the choice of a particular neighborhood of the central path.

For our purposes, we would like to ensure that results of the same kind as Propositions 2.3 and 2.4 hold for the iterates generated by the algorithm, which will only be guaranteed to lie in a neighborhood of the central path. A suitable choice of the neighborhood of the central path allows us both to obtain such kind of results and to incorporate well-studied and understood primal-dual interior-point algorithms. We shall apply a short-step primal-dual interior-point method to (9), (10). The iterates of such method are guaranteed to lie in the *local neighborhood* $\mathcal{N}_\beta$ of the central path defined next.

Definition 3.1 Let $\beta \in (0, 1/2)$ be given. The *central neighborhood* $\mathcal{N}_\beta$ of (9), (10) is defined as the set of points $(\vec{x}, \vec{y}, \vec{s})$ with $\vec{x}, \vec{s} \in K$ such that the following constraints hold

$$\begin{aligned} \mathcal{A}\vec{x} &= \vec{b} \\ \mathcal{A}^{\mathrm{T}}\vec{y} + \vec{s} &= \vec{c} \\ \|\vec{s} + \mu(\vec{x}, \vec{s})g_K(\vec{x})\|_{\vec{s}} &\leq \beta, \end{aligned}$$

where $\mu(\vec{x}, \vec{s}) := \langle \vec{x}, \vec{s} \rangle / \nu_K$.

Remark 3.2 Since K is self-scaled, by applying Proposition 2.1 it is easy to see that

$$\|\vec{s} + \mu(\vec{x}, \vec{s})g_K(\vec{x})\|_{\vec{s}} = \|\vec{x} + \mu(\vec{x}, \vec{s})g_K(\vec{s})\|_{\vec{x}}.$$

Hence either of these expressions can be used in the last inequality in Definition 3.1.

We can now state the main theorem of this paper.

Theorem 3.3 *Assume* $A \in \mathbb{R}^{m\times n}$ *with* $\mathcal{C}(A) < \infty$ *is given. A suitable short-step primal-dual interior-point algorithm applied to* (9), (10) *halts in at most* $O(\sqrt{\nu_C}\log(\nu_C\, \mathcal{C}(A))$ *interior-point iterations, yielding a strict solution to either* (1) *or* (2).

The proof of this theorem follows from the results in Sections 3.1, 3.2, and 3.3 below. These sections are interesting on their own.

3.1 Properties of the central neighborhood

The following two facts are closely related to Propositions 2.3 and 2.4 but apply to the larger set of points $\mathcal{N}_\beta$. Indeed, Propositions 2.3 and 2.4 can be obtained from Propositions 3.4 and 3.5 respectively by letting $\beta \to 0$.

Proposition 3.4 *Let* $(\vec{x}, \vec{y}, \vec{s}) \in \mathcal{N}_\beta$ *with* $\vec{c}^{\mathrm{T}}\vec{x} = \tau$. *If*

$$\tau < \frac{(1-2\beta)\rho_P(A)}{\sqrt{2}\nu_K},$$

then $\bar{x} := x + H_C(x)^{-1}A^T(AH_C(x)^{-1}A^T)^{-1}x''$ *satisfies*

$$A\bar{x} = 0,\ \bar{x} \in \mathrm{int}(C),\ \|\bar{x} - x\|_x \leq \frac{\sqrt{2}\nu_K\tau}{(1-2\beta)\rho_P(A)}.$$

Proof. See Section 5. □

Proposition 3.5 *Let* $(\vec{x}, \vec{y}, \vec{s}) \in \mathcal{N}_\beta$ *with* $\vec{b}^{\mathrm{T}}\vec{y} = -\eta$. *If*

$$\eta < \frac{(1-\beta)\rho_D(A)}{\nu_K},$$

then y *satisfies* $-A^Ty \in \mathrm{int}(C)$.

Proof. See Section 5. □

3.2 Initial point

Let e be the unique point in C such that $H_C(e) = I$. The existence of such point follows from the axioms of self-scaled barriers (see, e.g., [2, 11]). Furthermore, for the cones that most commonly arise in practice, that is, the nonnegative orthant $\mathbb{R}^n_+$, the cone of symmetric positive semidefinite matrices S^n_+, and the second-order cone $\mathcal{K}_{n+1}$, this point is readily available. In each of these cases the point e is respectively

$$\begin{bmatrix}1 & \dots & 1\end{bmatrix}^{\mathrm{T}}, \; I, \; \text{and} \; \begin{bmatrix}0 & \dots & 0 & 1\end{bmatrix}^{\mathrm{T}}.$$

We shall assume that the point e for the cone C is available. Under this reasonable assumption it is easy to construct an initial point in $\mathcal{N}_\beta$. We note that $g_C(e) = -e$ and $\|e\| = \sqrt{\nu_C}$. (See [4, 5, 11].)

Proposition 3.6 *Let*

$$\alpha := \frac{1}{\sqrt{\nu_C + 2}}, \text{ and } M := \frac{\alpha\|Ae\|}{\beta}.$$

The point $(\vec{x}, \vec{y}, \vec{s})$*, defined as follows, belongs to* $\mathcal{N}_\beta$*:*

$$\begin{aligned}\vec{x} &= (\alpha e, \alpha e, 1, -\alpha Ae, 2M)\\ \vec{y} &= (0, \tfrac{M}{\alpha}e, -\tfrac{M}{\alpha^2})\\ \vec{s} &= (\tfrac{M}{\alpha}e, -\tfrac{M}{\alpha}e, \tfrac{M}{\alpha^2}, 0, 1)\end{aligned}$$

Proof. By construction,

$$\mu(\vec{x}, \vec{s}) = \frac{\langle \vec{x}, \vec{s}\rangle}{\nu_K} = \frac{M(\frac{1}{\alpha^2} + 2)}{\nu_K} = \frac{M(\nu_C + 4)}{\nu_K} = M,$$

and

$$g_K(\vec{x}) = (-\frac{1}{\alpha}e, \frac{1}{\alpha}e, -\frac{1}{\alpha^2}, -\delta\alpha Ae, -2\delta M),$$

where $\delta := \frac{2}{4M^2 - \alpha^2\|Ae\|^2} = \frac{2}{M^2(4-\beta^2)}$.

Therefore,

$$\begin{aligned}\|\mu(\vec{x}, \vec{s})g_K(\vec{x}) + \vec{s}\|^2_{\vec{s}} &= \|(0, 0, 0, -M\delta\alpha Ae, 1 - 2\delta M^2)\|^2_{\vec{s}}\\ &= \langle \begin{bmatrix}-M\delta\alpha Ae\\ 1 - 2\delta M^2\end{bmatrix}, H_{\mathcal{K}_{m+1}}(0,1)\begin{bmatrix}-M\delta\alpha Ae\\ 1 - 2\delta M^2\end{bmatrix}\rangle\\ &= \langle \begin{bmatrix}-M\delta\alpha Ae\\ 1 - 2\delta M^2\end{bmatrix}, 2\begin{bmatrix}I & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}-M\delta\alpha Ae\\ 1 - 2\delta M^2\end{bmatrix}\rangle\\ &= 2(M^2\delta^2\alpha^2\|Ae\|^2 + (1 - 2\delta M^2)^2)\\ &= 2(\beta^2 M^4\delta^2 + (1 - 2\delta M^2)^2)\\ &= 2(\tfrac{4\beta^2}{(4-\beta^2)^2} + (1 - \tfrac{4}{4-\beta^2})^2)\\ &= \tfrac{2\beta^2(4+\beta^2)}{(4-\beta^2)^2}\\ &\le \beta^2\end{aligned}$$

(The last inequality holds because $\beta < 1/2$.)

It thus follows that $(\vec{x}, \vec{y}, \vec{s}) \in \mathcal{N}_\beta$ because by construction $\mathcal{A}\vec{x} = \vec{b}$, $\mathcal{A}^{\mathrm{T}}\vec{y} + \vec{s} = \vec{c}$, and $\vec{x},\ \vec{s} \in \mathrm{int}(K)$. □

3.3 The algorithm

We are now ready to describe our primal-dual algorithm. This is essentially a path-following short-step algorithm like those described in [5, Sec. 6], [11, Sec. 3.7], or [12, Sec. 3] enhanced with a specific starting point, and a stopping criterion.

The crucial step at each main iteration is the update of the iterate $(\vec{x}, \vec{y}, \vec{s})$. This is performed by putting

$$(\vec{x}^+, \vec{y}^+, \vec{s}^+) := (\vec{x}, \vec{y}, \vec{s}) + (\Delta\vec{x}, \Delta\vec{y}, \Delta\vec{s}), \tag{16}$$

where $(\Delta\vec{x}, \Delta\vec{y}, \Delta\vec{s})$ is the *Nesterov-Todd direction,* that is, the solution to

$$\begin{array}{rcrcl} & H_K(w)\Delta\vec{x} & + & \Delta\vec{s} & = & -(\vec{s} + \mu\, g_K(\vec{x})) \\ & \mathcal{A}\Delta\vec{x} & & & = & 0 \\ \mathcal{A}^T\Delta\vec{y} & & + & \Delta\vec{s} & = & 0, \end{array} \tag{17}$$

where w is the scaling point of $\vec{x}, \vec{s}$, namely, the unique point $w \in K$ that satisfies $H_K(w)\vec{x} = \vec{s}$.

Let $\beta, \delta \in (0, \frac{1}{2})$ be fixed constants such that

$$\frac{7(\beta^2 + \delta^2)}{1-\beta} \leq \left(1 - \frac{\delta}{\sqrt{\nu_K}}\right)\beta, \quad \frac{2\sqrt{2}\beta}{1-\beta} \leq 1.$$

Algorithm $\texttt{PD}(A)$

(i) Let

$$\alpha := \frac{1}{\sqrt{\nu_C + 2}};\ M := \frac{\alpha\|Ae\|}{\beta};$$

and

$$\begin{aligned} \vec{x} &= (\alpha e, \alpha e, 1, -\alpha A e, 2M) \\ \vec{y} &= (0, \tfrac{M}{\alpha}e, \tfrac{M}{\alpha^2}) \\ \vec{s} &= (\tfrac{M}{\alpha}e, -\tfrac{M}{\alpha}e, \tfrac{M}{\alpha^2}, 0, 1). \end{aligned}$$

(ii) If $-A^{\mathrm{T}}y \in \mathrm{int}(C)$ then `HALT` and
`return` y as a feasible solution for $A^{\mathrm{T}}y \in \mathrm{int}(C)$.

(iv) If $\bar{x} := x + H_C(x)^{-1}A^T(AH_C(x)^{-1}A^T)^{-1}x'' \in \mathrm{int}(C)$, then `HALT` and
`return` $\bar{x}$ as a feasible solution for $Ax = 0,\ x \in \mathrm{int}(C)$.

(v) Set $\bar{\mu} := \left(1 - \frac{\delta}{\sqrt{\nu_K}}\right) \mu(\vec{x}, \vec{s})$.

(vi) Update $(\vec{x}, \vec{y}, \vec{s})$ as in (16), (17) for $\mu = \bar{\mu}$.

(vii) Go to (ii).

Proof of Theorem 3.3. Arguments that are now standard in interior-point theory such as those in [11, Sec. 3.7] or [12, Sec. 3] ensure that the iterates generated by Algorithm `PD` lie in $\mathcal{N}_\beta$ and that $\mu(\vec{x}, \vec{s})$ is reduced by $(1 - \frac{\delta}{\sqrt{\nu_K}})$ at each iteration.

In addition, because $\vec{c}^{\mathrm{T}} \vec{x} - \vec{b}^{\mathrm{T}} \vec{y} = \tau + \eta = \nu_K \mu(\vec{x}, \vec{s})$, Propositions 3.4 and 3.5 ensure that the algorithm halts as soon as $\mu(\vec{x}, \vec{s})$ surpasses the threshold $\frac{(1-2\beta)\rho(A)}{\sqrt{2}\nu_K^2}$ (possibly sooner).

Since $\mu(\vec{x}, \vec{s})$ is reduced by $(1 - \frac{\delta}{\sqrt{\nu_K}})$ at each iteration, and at the initial point $\mu(\vec{x}, \vec{s}) = \frac{\alpha\|Ae\|}{\beta} = \frac{\|Ae\|}{\beta\sqrt{\nu_K - 2}}$, the threshold $\frac{(1-2\beta)\rho(A)}{\sqrt{2}\nu_K^2}$ is surpassed within

$$O(\sqrt{\nu_K}\, \log(\nu_K \sqrt{\nu_K}\, \|Ae\|/\rho(A))) = O(\sqrt{\nu_C}\, \log(\nu_C\, \mathcal{C}(A)))$$

iterations. □

3.4 Numerical considerations

Theorem 3.3 establishes a nice parallel between systems of equations and conic systems. In order to complete the parallel, we next address the computational work incurred at each main iteration.

The core of the computational work at each main iteration of Algorithm `PD` is the solution of (17). This system is typically solved via the Schur complement: First, solve the reduced system

$$(\mathcal{A}H_K(w)^{-1}\mathcal{A}^{\mathrm{T}})\Delta\vec{y} = -\mathcal{A}H_K(w)^{-1}(\vec{s} + \mu g_K(\vec{x})), \tag{18}$$

and then set

$$\Delta\vec{s} = -\mathcal{A}\Delta\vec{y}, \quad \Delta\vec{x} = H_K(w)^{-1}(\vec{s} + g_K(\vec{x}) - \Delta\vec{s}).$$

The critical step here is the solution of the reduced system (18). The conditioning of this system has been previously studied in [7] for general self-scaled programs. In particular, we have the following bound on $\kappa(\mathcal{A}H_K(w)^{-1}\mathcal{A}^{\mathrm{T}})$.

Proposition 3.7 *Assume $(\vec{x}, \vec{y}, \vec{s}) \in \mathcal{N}_\beta$. Let w be the scaling point of $\vec{x}, \vec{s}$. Then*

$$\kappa(\mathcal{A}H_K(w)^{-1}\mathcal{A}^{\mathrm{T}}) = O\left(\frac{\nu_K^4\|\mathcal{A}\|^2}{\mu(\vec{x}, \vec{s})^2}\right).$$

Proof. See [7, Cor. 1]

Corollary 3.8 *In step (vi) of every iteration of Algorithm* PD *the reduced system has condition number bounded by* $O(\nu_K^4\,\mathcal{C}(A)^2) = O(\nu_C^4\,\mathcal{C}(A)^2)$.

Theorem 3.3 and Corollary 3.8 together fully complete the parallel between systems of equations and conic systems: They state that a strictly feasible solution for either of the alternative systems (1) or (2) can be obtained (via Algorithm PD) within an amount of computational work proportional to the condition number $\mathcal{C}(A)$.

4 Proof of Propositions 2.3 and 2.4

4.1 Proof of Proposition 2.3

Elementary algebraic verifications show that the point $\bar{x}$, as defined in the statement of Proposition 2.3, satisfies $A\bar{x} = 0$ and

$$\begin{aligned} \|\bar{x} - x\|_x^2 &= {x''}^T (AH_C(x)^{-1}A^T)^{-1} x'' \\ &\leq \|(AH_C(x)^{-1}A^T)^{-1}\|\,\|x''\|^2 \\ &\leq \|(AH_C(x)^{-1}A^T)^{-1}\|\,\tau^2. \end{aligned} \tag{19}$$

By Proposition 2.1, to finish the proof it suffices to show that $\bar{x}$ satisfies

$$\|\bar{x} - x\|_x \leq \frac{\sqrt{2}\nu_K \tau}{\rho_P(A)} < 1. \tag{20}$$

But by (19), the inequality (20) follows if we show that $\|(AH_C(x)^{-1}A^T)^{-1}\| \leq \frac{2\nu_K^2}{\rho_P(A)^2}$ as long as $\tau < \frac{\rho_P(A)}{\sqrt{2}\nu_K}$. This in turn readily follows from Theorem 4.1 below, which we consider our second most important result.

Theorem 4.1 *Let* $\vec{x}$ *be a point on the central path of* (9) *with* $\vec{c}^{\mathrm{T}}\vec{x} = \tau > 0$. *Then*

$$\lambda_{\min}(AH_C(x)^{-1}A^T) \geq \left(\frac{\rho_P(A) + \tau}{\nu_K}\right)^2 - \tau^2.$$

Theorem 4.1 is closely related to results previously derived in [9]. However, the proof technique we used here, which we fully develop next, is new and in our opinion, more transparent. The ideas introduced in Section 4.2 —though a bit technical— provide deeper insight into the relaxation scheme (9), (10) and its relation with the pair of conic systems (1), (2).

4.2 Proof of Theorem 4.1

To ease our exposition we have divided the core of the proof of Theorem 4.1 into Lemmas 4.4, 4.5, and 4.6 below. We shall rely on the following *implicitly-defined-function* construction, which was introduced and studied by the third author in [7].

Let $f : D_f \to \mathbb{R}$ with $D_f \subseteq \mathbb{R}^n$ be a barrier function, and $A : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map such that $A^{-1}(Ax)$ is bounded for all $x \in D_f$. Consider the implicit function $\bar{f} : AD_f \to \mathbb{R}$ defined by

$$\begin{array}{rcll} \bar{f}(y) & = & \min & f(x) \\ & & \text{s.t.} & Ax = y. \end{array} \tag{21}$$

The following result was proven in [7].

Proposition 4.2 *Let $f : D_f \to \mathbb{R}$ with $D_f \subseteq \mathbb{R}^n$ be a barrier function, and $A : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map such that $A^{-1}(Ax)$ is bounded for all $x \in D_f$. Let $\bar{f}$ be defined in terms of A, f as in* (21)*. If f is a barrier then so is $\bar{f}$ and $\nu_{\bar{f}} \leq \nu_f$. Furthermore, for any $y \in D_{\bar{f}} = AD_f$, the Hessian $\bar{H}$ of $\bar{f}$ satisfies*

$$\bar{H}(y)^{-1} = AH(x(y))^{-1}A^T,$$

where $x(y)$ is the minimizer of (21) *for such y, and H is the Hessian of f.*

Assume $\tau > 0$ is given. Define $f_1 : D_{f_1} \subseteq \mathbb{R}^{m+n+2} \to \mathbb{R}$ implicitly as

$$\begin{array}{rcll} f_1(\vec{w}) & = & \min & f_K(\vec{x}) \\ & & \text{s.t.} & \begin{bmatrix} \mathcal{A} \\ \vec{c}^{\mathrm{T}} \end{bmatrix} \vec{x} = \begin{bmatrix} \vec{b} \\ \tau \end{bmatrix} + \vec{w} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ \tau \end{bmatrix} + \vec{w}, \end{array} \tag{22}$$

and define $f_2 : D_{f_2} \subseteq \mathbb{R}^m \to \mathbb{R}$ as

$$f_2(w) = f_1(w, 0, 0, 0). \tag{23}$$

By Proposition 4.2, f_1 is a barrier function. It thus follows that f_2 is a barrier as well. Furthermore, $\nu_{f_2} \leq \nu_{f_1} \leq \nu_K$.

Let us introduce some notation for the radius of the largest inscribed ball in a set. This notation facilitates some of our statements.

Definition 4.3 Given $S \subseteq \mathbb{R}^n$ and $x \in S$, define

$$r_S(x) := \sup\{\delta : \|w - x\| \leq \delta \Rightarrow w \in S\}.$$

For a given function $f : D_f \subseteq \mathbb{R}^n \to \mathbb{R}^m$ write r_f as a shorthand for r_{D_f}.

Notice that the inequalities (11) and (12) in Proposition 2.1 can be rephrased as

$$\frac{\|H(x)\|^{1/2}}{4\nu_f + 1} \leq \frac{1}{r_f(x)} \leq \|H(x)\|^{1/2}.$$

In particular, the functions $f_1,\ f_2$ defined above satisfy

$$\frac{\|H_i(x)\|^{1/2}}{4\nu_K+1} \le \frac{1}{r_{f_i}(x)} \le \|H_i(x)\|^{1/2},\ i=1,\ 2.$$

However, the self-scaledness of f_K yields the stronger property stated next. The proof of the following lemma is essentially the same as that of [7, Prop. 2] and hence omitted here.

Lemma 4.4 *Let f_K be a self-scaled barrier, and $f_1,\ f_2$ be defined as in* (22), (23)*. Then*

$$\frac{\|H_i(x)\|^{1/2}}{\nu_K} \le \frac{1}{r_{f_i}(x)} \le \|H_i(x)\|^{1/2},\ i=1,2.$$

Lemma 4.5 *Suppose $\tau > 0$ is given. Let f_2 be defined as in* (23)*. Then*

$$r_{f_2}(0) = \rho_P(A) + \tau.$$

Proof. The domain D_{f_2} can be written as

$$\begin{aligned}\bar{D}_{f_2} &= \{w \in \mathbb{R}^m : \exists \vec{x} \in \operatorname{int}(K) \begin{bmatrix}\mathcal{A}\\ \vec{c}^{\mathrm{T}}\end{bmatrix} \vec{x} = \begin{bmatrix} w^{\mathrm{T}} & 0 & 1 & \tau\end{bmatrix}^{\mathrm{T}}\}\\ &= \{w \in \mathbb{R}^m : \exists x \in \operatorname{int}(C),\ x'' \in \mathbb{R}^m \text{ s.t. } Ax + x'' = w,\ \|x\| < 1,\ \|x''\| < \tau\}\\ &= \{Ax : x \in \operatorname{int}(C),\ \|x\| < 1\} + \{x'' \in \mathbb{R}^m : \|x''\| < \tau\}.\end{aligned}$$

Hence

$$r_{f_2}(0) = r_S(0) + \tau,$$

where $S := \{Ax : x \in C,\ \|x\| \le 1\}$. But by Proposition 1.1, $\rho_P(A) = r_S(0)$ so

$$r_{f_2}(0) = \rho_P(A) + \tau.$$

□

In order to simplify notation, throughout the rest of the paper we let Q denote the $m \times (2m+n+2)$ projection matrix $\begin{bmatrix} I_m & 0 & 0 & 0\end{bmatrix}$, i.e., Q is the projection of the first m coordinates. We also let $\mathcal{A}_1$ denote $\begin{bmatrix} A & 0 & 0 & I_m & 0\end{bmatrix}$, i.e., the first block of rows of $\mathcal{A}$. Notice that $Q\begin{bmatrix}\mathcal{A}\\ \vec{c}^{\mathrm{T}}\end{bmatrix} = \mathcal{A}_1$.

Lemma 4.6 *Suppose $\tau > 0$ is given. Let $f_1,\ f_2$ be defined as in* (22) *and* (23)*. Then*

$$\lambda_{\min}(QH_1(0)^{-1}Q^{\mathrm{T}}) \ge \frac{1}{\|H_2(0)\|}.$$

Proof. Let $B_1 = \{\vec{w} \in \mathbb{R}^{m+n+2} : \langle \vec{w}, H_1(0)\,\vec{w}\rangle \le 1\}$, and $B_2 = \{w \in \mathbb{R}^m : \langle w, H_2(0)w\rangle \le 1\}$. Notice that $B_i = \{H_i(0)^{-1/2}u : \|u\| \le 1\}$, $i = 1, 2$ so

$$\sigma_{\min}(H_i^{-1/2}(0)) = r_{B_i}(0),\ i = 1, 2.$$

We claim $B_2 \subseteq Q(B_1)$. To see this, assume $w \in B_2$ is given. Then $w = QQ^{\mathrm{T}}w$ and $\langle w, H_2(0)w\rangle \le 1$. Hence

$$w^T Q H_1(0) Q^{\mathrm{T}} w = w^{\mathrm{T}} H_2(0) w \le 1.$$

Thus $Q^{\mathrm{T}}w \in B_1$ and in consequence $w = QQ^{\mathrm{T}}w \in Q(B_1)$, which proves the claim.

Since $B_2 \subseteq Q(B_1)$, we have

$$\begin{aligned}
\left(\lambda_{\min}(QH_1(0)^{-1}Q^{\mathrm{T}})\right)^{1/2} &= \sigma_{\min}(QH_1(0)^{-1/2}) \\
&= r_{Q(B_1)}(0) \\
&\ge r_{B_2}(0) \\
&= \sigma_{\min}(H_2(0)^{-1/2}) \\
&= \lambda_{\min}(H_2(0)^{-1})^{1/2} \\
&= \tfrac{1}{\|H_2(0)\|^{1/2}}.
\end{aligned}$$

Thus

$$\lambda_{\min}(QH_1(0)^{-1}Q^{\mathrm{T}}) \ge \frac{1}{\|H_2(0)\|}.$$

□

Proof of Theorem 4.1. Consider $\tau = \vec{c}^{\mathrm{T}}\vec{x}$ fixed and let f_1 and f_2 be defined as in (22) and (23). By Proposition 4.2

$$H_1(0)^{-1} = \begin{bmatrix} \mathcal{A} \\ \vec{c}^{\mathrm{T}} \end{bmatrix} H_K(\vec{x})^{-1} \begin{bmatrix} \mathcal{A}^T & \vec{c} \end{bmatrix},$$

and by Lemmas 4.4 and 4.5

$$\|H_2(0)\|^{1/2} \le \frac{\nu_K}{r_{f_2}(0)} = \frac{\nu_K}{\rho_P(A) + \tau}.$$

Lemma 4.6 thus yields

$$\begin{aligned}
\left(\tfrac{\rho_P(A)+\tau}{\nu_K}\right)^2 &\le \lambda_{\min}(QH_1(0)^{-1}Q^T) \\
&= \lambda_{\min}\left(Q \begin{bmatrix} \mathcal{A} \\ \vec{c}^{\mathrm{T}} \end{bmatrix} H(\vec{x})^{-1} \begin{bmatrix} \mathcal{A} & c \end{bmatrix}^T Q^{\mathrm{T}}\right) \\
&= \lambda_{\min}(\mathcal{A}_1 H(\vec{x})^{-1}\mathcal{A}_1^{\mathrm{T}}).
\end{aligned}$$

But $H^{-1} = \begin{bmatrix} H_C^{-1} & 0 & 0 \\ 0 & H_{K_{n+1}}^{-1} & 0 \\ 0 & 0 & H_{K_{m+1}}^{-1} \end{bmatrix}$ so

$$\mathcal{A}_1 H(\vec{x})^{-1}\mathcal{A}_1^{\mathrm{T}} = AH_C(x)^{-1}A^T + \begin{bmatrix} I_m & 0 \end{bmatrix} H_{K_{m+1}}(x'', \tau)^{-1} \begin{bmatrix} I_m \\ 0 \end{bmatrix}.$$

Therefore,

$$\begin{array}{rcl}\lambda_{\min}(AH_C(x)^{-1}A^{\mathrm{T}}) & \geq & \lambda_{\min}(\mathcal{A}_1 H(\vec{x})^{-1}\mathcal{A}_1^{\mathrm{T}}) - \left\| \begin{bmatrix} I_m & 0 \end{bmatrix} H_{K_{m+1}}(x'',\tau)^{-1} \begin{bmatrix} I_m \\ 0 \end{bmatrix} \right\| \\ & \geq & \left(\frac{\rho_P(A)+\tau}{\nu_K}\right)^2 - \left\| \begin{bmatrix} I_m & 0 \end{bmatrix} H_{K_{m+1}}(x'',\tau)^{-1} \begin{bmatrix} I_m \\ 0 \end{bmatrix} \right\|.\end{array}$$

To finish just notice that the Hessian $H_{K_{m+1}}$ of the second-order cone barrier satisfies

$$\begin{bmatrix} I_m & 0 \end{bmatrix} H_{K_{m+1}}(x'',\tau)^{-1} \begin{bmatrix} I_m \\ 0 \end{bmatrix} = \frac{\tau^2 - \|x''\|^2}{2} I_m + x''(x'')^{\mathrm{T}}$$

So

$$\left\| \begin{bmatrix} I_m & 0 \end{bmatrix} H_{K_{m+1}}(x'',\tau)^{-1} \begin{bmatrix} I_m \\ 0 \end{bmatrix} \right\| \leq \frac{\tau^2 + \|x''\|^2}{2} \leq \tau^2.$$

□

4.3 Proof of Proposition 2.4

The point $\vec{x} = (x, x', t_1, x'', t_2)$ satisfies

$$Ax + x'' = 0,\ 0 \neq x \in C$$

with $\|x''\| \leq \vec{c}^{\mathrm{T}}\vec{x} = \tau$. Thus

$$\left(A + \frac{x''\,x^{\mathrm{T}}}{\|x\|^2}\right) x = 0,\ 0 \neq x \in C.$$

Hence $\left(A + \frac{x''\,x^{\mathrm{T}}}{\|x\|^2}\right) \notin \mathcal{D}$, and consequently

$$\rho_D(A) \leq \left\| \frac{x''\,x^{\mathrm{T}}}{\|x\|^2} \right\| = \frac{\|x''\|}{\|x\|} \leq \frac{\tau}{\|x\|}.$$

We thus have

$$\|x\| \leq \frac{\tau}{\rho_D(A)}.$$

Since $(\vec{x}, \vec{y}, \vec{s})$ is on the central path, for some $\mu > 0$ we have $\vec{s} = -\mu g_K(\vec{x})$ and $\tau \leq \nu_K \mu$. In particular, $s = -\mu g_C(x)$. Thus, because f_C is a self-scaled barrier, Proposition 2.1 yields

$$\|H_C(s)\| = \frac{\|H_C(x)^{-1}\|}{\mu^2} \leq \frac{\|x\|^2}{\mu^2} \leq \frac{\nu_K^2}{\rho_D(A)^2}.$$

Proposition 2.1 again implies that $s + u \in C$ for all $\|u\| < \frac{\rho_D(A)}{\nu_K}$. In particular,

$$-A^{\mathrm{T}}y = s + y' \in C^*,$$

because $\|y'\| \leq -\vec{b}^{\mathrm{T}}\vec{y} = \eta < \frac{\rho_D(A)}{\nu_K}$.

5 Proof of Propositions 3.4 and 3.5

The proofs of Propositions 3.4 and 3.5 are appropriate modifications of the proofs of Propositions 2.3 and 2.4. We shall rely on the following key properties of barrier functions (cf. [3, 10]).

Proposition 5.1 *Let f be a ν_f-barrier function and $x, x' \in D_f$, with $\|x - x'\|_x < 1$. Then for every nonzero vector v*

$$1 - \|x - x'\|_x \leq \frac{\|v\|_x}{\|v\|_{x'}} \leq \frac{1}{1 - \|x - x'\|_x}. \tag{24}$$

We will use the following immediate consequences of Proposition 5.1.

Lemma 5.2 *Let f be a ν_f-barrier function and $x, x' \in D_f \subseteq \mathbb{R}^k$, with $\|x - x'\|_x \leq \beta < 1$. Then for all v*

$$v^{\mathrm{T}} H(x)^{-1} v \geq (1 - \beta)^2 v^{\mathrm{T}} H(x')^{-1} v, \tag{25}$$

and for all $Q \in \mathbb{R}^{p \times k}$

$$\lambda_{\min}(QH(x)^{-1}Q^{\mathrm{T}}) \geq (1 - \beta)^2 \lambda_{\min}(QH(x')^{-1}Q^{\mathrm{T}}). \tag{26}$$

5.1 Proof of Proposition 3.4

As in the proof of Theorem 2.3, elementary algebraic verifications show that the point $\bar{x}$, as defined in the statement of Theorem 2.3, satisfies $A\bar{x} = 0$ and

$$\|\bar{x} - x\|_x^2 \leq \|(AH_C(x)^{-1}A^T)^{-1}\| \, \tau^2. \tag{27}$$

By Proposition 2.1, to finish the proof it suffices to show that $\bar{x}$ satisfies

$$\|\bar{x} - x\|_x \leq \frac{\sqrt{2}\nu_K \tau}{(1 - 2\beta)\rho_P(A)} < 1. \tag{28}$$

But by (27), the inequality (28) follows if we show that $\|(AH_C(x)^{-1}A^T)^{-1}\| \leq \frac{2\nu_K^2}{(1-2\beta)^2\rho_P(A)^2}$ as long as $\tau < \frac{(1-2\beta)\rho_P(A)}{\sqrt{2}\nu_K}$. This in turn readily follows from Theorem 5.3 below, a natural modification of Theorem 4.1.

Theorem 5.3 *Let $(\vec{x}, \vec{y}, \vec{s}) \in \mathcal{N}_\beta$ with $\vec{c}^{\mathrm{T}} \vec{x} = \tau > 0$. Then*

$$\lambda_{\min}(AH_C(x)^{-1}A^T) \geq \left(\frac{(1 - 2\beta)(\rho_P(A) + \tau)}{\nu_K} \right)^2 - \tau^2.$$

5.2 Proof of Theorem 5.3

For a given $\tau > 0$, let f_1 and f_2 be defined as in (22) and (23). As before, f_1 and f_2 are barrier functions, and Lemmas 4.4, 4.5, and 4.6 hold.

Now given $(\vec{x}, \vec{s}, \vec{y}) \in \mathcal{N}_\beta$, let

$$\begin{aligned} \tilde{x} &:= -\mu(\vec{x}, \vec{s}) g_K(\vec{s}), \\ \tilde{b} &:= \mathcal{A}\tilde{x}, \text{ and} \\ \tilde{w} &:= \begin{bmatrix} \mathcal{A}\tilde{x} \\ \vec{c}^{\mathrm{T}}\tilde{x} \end{bmatrix} - \begin{bmatrix} \mathcal{A}\vec{x} \\ \vec{c}^{\mathrm{T}}\vec{x} \end{bmatrix} = \begin{bmatrix} \tilde{b} \\ \vec{c}^{\mathrm{T}}\tilde{x} \end{bmatrix} - \begin{bmatrix} \vec{b} \\ \tau \end{bmatrix}. \end{aligned} \tag{29}$$

It is easy to see that $(\tilde{x}, \vec{s}, \vec{y})$ is on the central path of the primal-dual pair:

$$\begin{array}{llcll} \min & \langle \vec{c}, \vec{x} \rangle & \qquad & \max & \langle \tilde{b}, \vec{y} \rangle \\ \text{s.t.} & \mathcal{A}\vec{x} = \tilde{b} & & \text{s.t.} & \mathcal{A}^{\mathrm{T}}\vec{y} + \vec{s} = \vec{c} \\ & \vec{x} \in K & & & \vec{s} \in K. \end{array} \tag{30}$$

In particular, the point $\tilde{x}$ is the minimizer of (22) for $w = \tilde{w}$. Furthermore, for $\tau = \vec{c}^{\mathrm{T}}\vec{x}$, the point $\tilde{w}$ is near the point $0 \in D_{f_1}$ as the following technical lemma states.

Lemma 5.4 *Given $(\vec{x}, \vec{y}, \vec{s}) \in \mathcal{N}_\beta$, fix $\tau = \vec{c}^{\mathrm{T}}\vec{x}$ and let $\tilde{w}$ be defined as in* (29). *Then $\tilde{w} \in D_{f_1}$ and*

$$\|\tilde{w}\|_{\tilde{w}} = (\tilde{w}^{\mathrm{T}} H_1(\tilde{w})\tilde{w})^{1/2} \leq \frac{\beta}{1-\beta}.$$

Proof. See Section 5.4. □

Proof of Theorem 5.3. Consider $\tau = \vec{c}^{\mathrm{T}}\vec{x}$ fixed and let f_1 and f_2 be defined as in (22) and (23). Let $\tilde{w}$ be defined as in (29). Since $\tilde{x}$ is the minimizer of (22) for $w = \tilde{w}$, Proposition 4.2 yields

$$H_1(\tilde{w})^{-1} = \begin{bmatrix} \mathcal{A} \\ \vec{c}^{\mathrm{T}} \end{bmatrix} H(\tilde{x})^{-1} \begin{bmatrix} \mathcal{A}^T & \vec{c} \end{bmatrix}. \tag{31}$$

On the other hand, since $(\vec{x}, \vec{y}, \vec{s}) \in \mathcal{N}_\beta$, $\|\vec{x} - \tilde{x}\|_{\vec{x}} = \|\vec{x} + \mu(\vec{x}, \vec{s}) g_K(\vec{s})\|_{\vec{x}} \leq \beta$. Therefore, by Lemma 5.2,

$$\begin{aligned} \lambda_{\min}(\mathcal{A}_1 H(\vec{x})^{-1}\mathcal{A}_1^{\mathrm{T}}) &= \lambda_{\min}(Q \begin{bmatrix} \mathcal{A} \\ \vec{c}^{\mathrm{T}} \end{bmatrix} H(\vec{x})^{-1} \begin{bmatrix} \mathcal{A}^{\mathrm{T}} & \vec{c} \end{bmatrix} Q^{\mathrm{T}}) \\ &\geq (1-\beta)^2 \lambda_{\min}(Q \begin{bmatrix} \mathcal{A} \\ \vec{c}^{\mathrm{T}} \end{bmatrix} H(\tilde{x})^{-1} \begin{bmatrix} \mathcal{A}^{\mathrm{T}} & \vec{c} \end{bmatrix} Q^{\mathrm{T}}) \\ &= (1-\beta)^2 \lambda_{\min}(Q H_1(\tilde{w})^{-1} Q^{\mathrm{T}}) \end{aligned}$$

where Q denotes the projection matrix $\begin{bmatrix} I_m & 0 & 0 & 0 \end{bmatrix}$ and $\mathcal{A}_1$ denotes the first block of rows of $\mathcal{A}$, i.e., $\begin{bmatrix} A & 0 & 0 & I_m & 0 \end{bmatrix}$.

But $\|\tilde{w}\|_{\tilde{w}} \leq \frac{\beta}{1-\beta}$ by Lemma 5.4, so applying Lemma 5.2 again we get

$$\lambda_{\min}(\mathcal{A}_1 H(\vec{x})^{-1}\mathcal{A}_1^{\mathrm{T}}) \geq (1-2\beta)^2 \lambda_{\min}(QH_1(0)^{-1}Q^{\mathrm{T}}). \tag{32}$$

Again as in the proof of Theorem 4.1, Lemmas 4.4, 4.5, and 4.6 yield

$$\lambda_{\min}(QH_1(0)^{-1}Q^{\mathrm{T}}) \geq \frac{1}{\|H_2(0)\|} \geq \left(\frac{\rho_P(A)+\tau}{\nu_K}\right)^2. \tag{33}$$

Combining (32) and (33) we get

$$\lambda_{\min}(\mathcal{A}_1 H(\vec{x})^{-1}\mathcal{A}_1^{\mathrm{T}}) \geq \left(\frac{(1-2\beta)(\rho_P(A)+\tau)}{\nu_K}\right)^2.$$

To finish, just proceed as at the end of the proof of Theorem 4.1, i.e.,

$$\begin{aligned}\lambda_{\min}(AH_C(x)^{-1}A^T) &\geq \lambda_{\min}(\mathcal{A}_1 H(\vec{x})^{-1}\mathcal{A}_1^{\mathrm{T}}) - \left\| [I_m, 0]\, H_{K_{m+1}}(x'', \tau)^{-1} \begin{bmatrix} I_m \\ 0 \end{bmatrix} \right\| \\ &\geq \left(\tfrac{(1-2\beta)(\rho_P(A)+\tau)}{\nu_K}\right)^2 - \tau^2.\end{aligned}$$

□

5.3 Proof of Proposition 3.5

Proceeding exactly as in the first part of the proof of Proposition 2.4, we get

$$\|x\| \leq \frac{\tau}{\rho_D(A)} \leq \frac{\nu_K \mu(\vec{x}, \vec{s})}{\rho_D(A)}.$$

Let $\tilde{s} := -\mu(\vec{x}, \vec{s}) g_C(x)$. Because f_C is a self-scaled barrier, Proposition 2.1 yields

$$\|H_C(\tilde{s})\| = \frac{\|H_C(x)^{-1}\|}{\mu(\vec{x}, \vec{s})^2} \leq \frac{\|x\|^2}{\mu(\vec{x}, \vec{s})^2} \leq \frac{\nu_K^2}{\rho_D(A)^2}.$$

But $\|s - \tilde{s}\|_s \leq \|\vec{s} + \mu(\vec{x}, \vec{s}) g_K(\vec{x})\|_{\vec{s}} \leq \beta$, because $(\vec{x}, \vec{y}, \vec{s}) \in \mathcal{N}_\beta$. Thus by Proposition 5.1,

$$\|H_C(s)\| \leq \frac{\|H_C(\tilde{s})\|}{(1-\beta)^2} \leq \frac{\nu_K^2}{(1-\beta)^2 \rho_D(A)^2}.$$

Proposition 2.1 implies that $s+u \in C$ for all $\|u\| < \frac{(1-\beta)\rho_D(A)}{\nu_K}$. In particular,

$$-A^{\mathrm{T}} y = s + y' \in C^*,$$

because $\|y'\| \leq -\vec{b}^{\mathrm{T}} \vec{y} = \eta < \frac{(1-\beta)\rho_D(A)}{\nu_K}$.

5.4 Proof of Lemma 5.4

Lemma 5.4 readily follows from the following more general result concerning implicitly-defined functions.

Proposition 5.5 *Let $\bar{f}$ be defined in terms of A, f as in* (21)*. Let $y \in D_{\bar{f}}$ and $x(y)$ be the minimizer of* (21) *for such y. Then every $x \in D_f$ satisfies*

$$\|Ax - y\|_y \leq \|x - x(y)\|_{x(y)}, \tag{34}$$

where $\|\cdot\|_y$ is the local norm induced by $\bar{H}(y)$.

Proof. To simplify the notation, let us abbreviate $\bar{H}(y)$ as $\bar{H}$ and $H(x(y))$ as H. By Proposition 4.2 and Schwarz inequality,

$$\begin{array}{rcl} \|Ax - y\|_y^2 & = & \|Ax - Ax(y)\|_y^2 \\ & = & \langle\, A(x - x(y))\,,\, \bar{H}A(x - x(y))\,\rangle \\ & = & \langle (x - x(y)), HH^{-1}A^{\mathrm{T}}(AH^{-1}A^{\mathrm{T}})^{-1}A(x - x(y))\rangle \\ & = & \langle (x - x(y)), H^{-1}A^{\mathrm{T}}(AH^{-1}A^{\mathrm{T}})^{-1}A(x - x(y))\rangle_{x(y)} \\ & \leq & \|x - x(y)\|_{x(y)}\|H^{-1}A^{\mathrm{T}}(AH^{-1}A^{\mathrm{T}})^{-1}A(x - x(y))\|_{x(y)}. \end{array} \tag{35}$$

But it is easy to see that $H(x(y))^{-1}A^{\mathrm{T}}(AH(x(y))^{-1}A^{\mathrm{T}})^{-1}A(x - x(y))$ is the solution to

$$\begin{array}{ll} \min & \|v\|_{x(y)} \\ \text{s.t.} & Av = A(x - x(y)). \end{array}$$

Consequently,

$$\|H^{-1}A^{\mathrm{T}}(AH^{-1}A^{\mathrm{T}})^{-1}A(x - x(y))\|_{x(y)} \leq \|x - x(y)\|_{x(y)}. \tag{36}$$

So we get (34) by putting (35) and (36) together. □

References

[1] F. Cucker and J. Peña, "A Primal-Dual Algorithm for Solving Polyhedral Conic Systems with a Finite Precision Machine," SIAM Journal on Optimization 12 (2002) pp. 522-554.

[2] R. Hauser and O. Güler, "Self-scaled barrier functions on symmetric cones and their classification," To Appear in Foundations of Computational Mathematics.

[3] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, Philadelphia, 1994.

[4] Y. Nesterov and M. Todd, "Self-scaled barriers and interior-point methods for convex programming," Mathematics of Operations Research 22 (1997) pp. 1-42.

[5] Y. Nesterov and M. Todd, “Primal-dual interior-point methods for self-scaled cones,” SIAM Journal on Optimization 8 (1998) pp. 324-364.

[6] J. Peña, “Understanding the Geometry on Infeasible Perturbations of a Conic Linear System,” SIAM Journal on Optimization 10 (2000) pp. 534-550.

[7] J. Peña, “Two Properties of Condition Numbers for Convex Programs via Implicitly Defined Functions,” To Appear in Mathematical Programming.

[8] J. Peña, “A Characterization of the Distance to Infeasibility under Structured Perturbations,” GSIA Working Paper, Carnegie Mellon University.

[9] J. Peña and J. Renegar, “Computing Approximate Solutions for Conic Systems of Constraints,” Mathematical Programming 87 (2000) pp. 351-383.

[10] J. Renegar, “Linear Programming, Complexity Theory and Elementary Functional Analysis,” Mathematical Programming, 70 (1995) pp. 279-351.

[11] J. Renegar, *A Mathematical View of Interior-Point Methods in Convex Optimization*, SIAM, Philadelphia, 2001.

[12] S. Schmieta and F. Alizadeh, “Associative and Jordan Algebras, and Polynomial Time Interior Point Algorithms for Symmetric Cones,” Mathematics of Operations Research 26 (2001) pp. 543–564.