OPERATIONS RESEARCH CENTER

Working Paper

PROJECTIVE PRE-CONDITIONERS FOR IMPROVING THE BEHAVIOR OF A HOMOGENEOUS CONIC LINEAR SYSTEM

by

Robert M. Freund
Alexandre Belloni

OR 375-05 **May 2005**

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Mathematical Programming manuscript No.
(will be inserted by the editor)

Alexandre Belloni · Robert M. Freund

PROJECTIVE PRE-CONDITIONERS FOR IMPROVING THE BEHAVIOR OF A HOMOGENEOUS CONIC LINEAR SYSTEM

May 31, 2005.

Abstract. In this paper we present a general theory for transforming a normalized homogeneous conic system $F:\ Ax = 0,\ \bar{s}^T x = 1,\ x \in C$ to an equivalent system via projective transformation induced by the choice of a point $\hat{v}$ in the set $H_{\bar{s}}^{\circ} = \{v : \bar{s} - A^T v \in C^*\}$. Such a projective transformation serves to pre-condition the conic system into a system that has both geometric and computational properties with certain guarantees. We characterize both the geometric behavior and the computational behavior of the transformed system as a function of the symmetry of $\hat{v}$ in $H_{\bar{s}}^{\circ}$ as well as the complexity parameter ϑ of the barrier for C. Under the assumption that F has an interior solution, $H_{\bar{s}}^{\circ}$ must contain a point v whose symmetry is at least $1/m$; if we can find a point whose symmetry is $\Omega(1/m)$ then we can projectively transform the conic system to one whose geometric properties and computational complexity will be strongly-polynomial-time in m and ϑ. We present a method for generating such a point $\hat{v}$ based on sampling and on a geometric random walk on $H_{\bar{s}}^{\circ}$ with associated complexity and probabilistic analysis. Finally, we implement this methodology on randomly generated homogeneous linear programming feasibility problems, constructed to be poorly behaved. Our computational results indicate that the projective pre-conditioning methodology holds the promise to markedly reduce the overall computation time for conic feasibility problems; for instance we observe a 46% decrease in average IPM iterations for 100 randomly generated poorly-behaved problem instances of dimension 1000×5000.

1. Introduction

Our interest lies in behavioral and computational characteristics of the following homogeneous convex feasibility problem in conic linear form:

$$F : \begin{cases} Ax = 0 \\ \ \ x \in C \backslash \{0\}\ , \end{cases} \tag{1}$$

where $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is a linear operator and C is a convex cone.

It is well known that the standard form conic feasibility problem

$$\begin{cases} \bar{A}x = \bar{b} \\ \ \ x \in K \end{cases}$$

is a special case of F under the assignments $C \leftarrow K \times \mathbb{R}_+$, $A \leftarrow \left[\bar{A}, -\bar{b}\right]$ and the qualification that we seek solutions in the interiors of the cones involved. Furthermore, this setting is general enough to encompass convex optimization as well.

Alexandre Belloni: MIT Operations Research Center, E40-149, 77 Massachusetts Ave., Cambridge, Massachusetts 02142, USA. e-mail: belloni@mit.edu

Robert M. Freund: MIT Sloan School of Management, 50 Memorial Drive, Cambridge, Massachusetts 02139-4307, USA. e-mail: rfreund@mit.edu

In the context of interior-point methods (IPMs), the system F has good computational complexity if an IPM for solving F has a good iteration bound. We also say that F has good geometric behavior if the width of the cone of feasible solutions of F is large, equivalently if F has a solution x whose relative distance from ∂C is large. Choose a point $\bar{s} \in \text{int}C^*$, and note that F is equivalent to the normalized problem $F_{\bar{s}}: \ Ax = 0, \ \bar{s}^T x = 1, \ x \in C$. We show that both the computational complexity and the geometry behavior of F can be bounded as a function of only two quantities: (i) the symmetry of the so-called *image set* $H_{\bar{s}} := \{Ax : \bar{s}^T x = 1, \ x \in C\}$ of $F_{\bar{s}}$ about the origin, denoted by $\text{sym}(0, H_{\bar{s}})$, and the complexity value ϑ of the barrier function for C. These results are shown in Section 3 after some initial definitions and analysis are developed in Section 2.

In Section 4 we present a general theory for transforming the normalized homogeneous conic system $F_{\bar{s}}$ to an equivalent system via projective transformation. Such a projective transformation serves to pre-condition the conic system into a system that has both geometric and computational properties with certain guarantees; we use the term "projective pre-conditioner" to describe such a projective transformation. Under the assumption that F has an interior solution, there must exist projective pre-conditioners that transform $F_{\bar{s}}$ into equivalent systems that are solvable in strongly-polynomial time in m and ϑ. The quality of a projective pre-conditioner depends on the ability to compute a point $\hat{v}$ that is "deep" in the set $H^\circ_{\bar{s}} = \{v : \bar{s} - A^T v \in C^*\}$. Several constructive approaches for computing such points are discussed, including a stochastic method based on a geometric random walk.

The geometric random walk approach is further developed in Section 5, with associated complexity analysis. In Section 6 we present results from computational experiments designed to assess the practical viability of the projective pre-conditioner method based on geometric random walks. We generated 300 linear programming feasibility problems (100 each in three sets of dimensions) designed to be poorly behaved. We present computational evidence that the method is very effective; for the 100 problems of dimension 1000×5000 the average IPM iterations decreased by 46% and average total running time decreased by 33%, for example.

Section 7 contains summary conclusions and next steps.

We point out that a very different pre-conditioner for F was proposed in [3] that is a linear (not projective) transformation of the range space of A and that aims to improve Renegar's condition measure $\mathcal{C}(A)$ (but does not improve the complexity of the original problem or the geometry of the feasible region).

1.1. Notation

Let $e = (1, \ldots, 1)^T \in \mathbb{R}^d$ denote the vector of ones in dimension d. Given a closed convex set $S \subset \mathbb{R}^d$ with $0 \in S$, the *polar* of S is $S^\circ := \{y \in \mathbb{R}^d : y^T x \leq 1 \text{ for all } x \in S\}$ and satisfies $S^{\circ\circ} = S$, see Rockafellar [27]. Given a closed convex cone $K \subset \mathbb{R}^d$, the (positive) *dual cone* of K is $K^* := \{y \in \mathbb{R}^d : y^T x \geq 0 \text{ for all } x \in K\}$ and satisfies $K^{**} = K$, also see [27]. For a general norm

$\|\cdot\|$, let $B(c,r)$ and $\mathrm{dist}(x,T)$ denote the ball of radius r centered at c and the distance from a point x to a set T, respectively.

2. Normalization and $\bar{s}$-norm, Behavioral Measures, and Barrier Calculus

Regarding the conic feasibility problem (1), we make the following assumptions:

Assumption 1. *C is a* regular *cone, i.e.,* $\mathrm{int}C \neq \emptyset$ *and C contains no line.*

Assumption 2. *F has a solution, i.e.,*

$$\mathcal{F} := \{x \in \mathbb{R}^n : Ax = 0,\ x \in C\backslash\{0\}\} \ \neq\ \emptyset\ .$$

Assumption 3. $\mathrm{rank}A = m$.

2.1. Normalization of F and a Class of Norms that are Linear on C

Let $\bar{s} \in \mathrm{int}C^*$ be chosen, then $x \in C\backslash\{0\}$ if and only if $x \in C$ and $\bar{s}^T x > 0$, whereby we can write F equivalently as the normalized problem:

$$F_{\bar{s}} : \left\{ \begin{array}{c} Ax = 0 \\ \bar{s}^T x = 1 \\ x \in C, \end{array} \right.$$

whose feasible region is $\mathcal{F}_{\bar{s}} := \{x \in \mathbb{R}^n : Ax = 0,\ x \in C,\ \bar{s}^T x = 1\}$.

Given the regular cone C and $\bar{s} \in \mathrm{int}C^*$, the linear functional:

$$f(x) := \bar{s}^T x$$

behaves like a norm when restricted to $x \in C$, namely $f(x)$ is (trivially) convex and positively homogeneous on C, and $f(x) > 0$ for $x \in C\backslash\{0\}$. The natural norm that agrees with $f(x) := \bar{s}^T x$ on C is:

$$\begin{array}{rl} \|x\|_{\bar{s}} \ := & \min\limits_{x^1,x^2} \ \bar{s}^T(x^1 + x^2) \\ & \text{s.t.} \ \ x^1 - x^2 = x \\ & \qquad x^1 \in C \\ & \qquad x^2 \in C\ , \end{array}$$

and note that $\|x\|_{\bar{s}} = \bar{s}^T x$ for $x \in C$. We refer to this norm as the "$\bar{s}$-norm." Indeed, $\|x\|_{\bar{s}}$ is an exact generalization of the L_1 norm in the case when $C = \mathbb{R}^n_+$ and $\bar{s} = e$:

$$\begin{array}{rl} \|x\|_1 \ := & \min\limits_{x^1,x^2} \ e^T(x^1 + x^2) \\ & \text{s.t.} \ \ x^1 - x^2 = x \\ & \qquad x^1 \in \mathbb{R}^n_+ \\ & \qquad x^2 \in \mathbb{R}^n_+\ . \end{array}$$

We will make extensive use of the family of norms $\|\cdot\|_{\bar{s}}$ herein. In the case when C is a self-scaled cone, both $\|\cdot\|_{\bar{s}}$ and its dual norm have convenient explicit formulas, for details see Section 2 of [8].

2.2. Measuring the Behavior of F: Geometry and Complexity

A natural way to think of "good" geometric behavior of F is in terms of the existence of a solution x of F that is nicely interior to the cone C. However, due to the homogeneity of F any solution $x \in \text{int}C$ can be re-scaled by a positive constant to yield a solution that is arbitrarily far from the boundary of C. Given a norm $\|\cdot\|$ on $\mathbb{R}^n$, we therefore consider the following measure of distance of x from the boundary of C that is invariant under positive scalings:

$$\text{reldist}(x, \partial C) := \frac{\text{dist}(x, \partial C)}{\|x\|} \ , \tag{2}$$

where $\text{dist}(x, S) := \inf_{y \in S} \|x - y\|$. We define the "width" or "min-width" of the feasible region $\mathcal{F}$ under the norm $\|\cdot\|$ to be the quantity τ_F defined by:

$$\tau_F = \max_{x \in \mathcal{F}} \{\text{reldist}(x, \partial C)\} = \max_{\substack{Ax = 0 \\ x \in C \setminus \{0\}}} \{\text{reldist}(x, \partial C)\} \ . \tag{3}$$

Note that τ_F is larger to the extent that F has a solution of small norm whose distance from the boundary of C is large. τ_F is a variation on the notion of the "inner measure" of Goffin [11] when the norm is Euclidean, and has also been used in similar format in [9, 7].

As is customary, we will measure the computational behavior of F using a worst-case computational complexity upper bound on the number of iterations that a suitably designed interior-point method (IPM) needs to compute a solution of F.

We will show that both the geometry measure τ_F and the computational complexity can be bounded as simple functions of the *symmetry* of the origin in the *image set* of $F_{\bar{s}}$, which we now define.

The *image set* $H = H_{\bar{s}}$ of $F_{\bar{s}}$ is defined as:

$$H = H_{\bar{s}} := \{Ax : x \in C \ , \ \bar{s}^T x = 1\} \ .$$

Note that the assumption that F has a solution implies that $0 \in H$.

We consider the *symmetry* of a point in a convex set as defined originally by Minkowski [19], see also the references and results in [1]. Let $S \subset \mathbb{R}^d$ be a convex set. Define:

$$\text{sym}(\bar{x}, S) := \max \{t \mid y \in S \Rightarrow \bar{x} - t(y - \bar{x}) \in S\} \ ,$$

which essentially measures how symmetric S is about the point x. Define

$$\text{sym}(S) := \max_{\bar{x} \in S} \ \text{sym}(\bar{x}, S) \ ,$$

and x^* is called a *symmetry point* of S if $\text{sym}(x^*, S) = \text{sym}(S)$.

2.3. Logarithmically-Homogeneous Barrier Calculus

We presume that we have a ϑ-logarithmically homogeneous (self-concordant) barrier function $f(\cdot)$ for C, see [21].

Remark 4. We will use the following properties of a ϑ-logarithmically homogeneous barrier:
(i) $\bar{u} \in \text{int}C$ if and only if $-\nabla f(\bar{u}) \in \text{int}C^*$
(ii) $f^*(s) := -\inf_{x \in \text{int}C}\{s^T x + f(x)\}$ is a ϑ-logarithmically homogeneous barrier for C^*
(iii) $\bar{s} \in \text{int}C^*$ if and only if $-\nabla f^*(\bar{s}) \in \text{int}C$
(iv) $-\nabla f(\bar{u})^T \bar{u} = \vartheta$ and $-\nabla f^*(\bar{s})^T \bar{s} = \vartheta$ for $\bar{u} \in \text{int}C$ and $\bar{s} \in \text{int}C^*$
(v) $\nabla f(\bar{u}) = -H(\bar{u})\bar{u}$ for $\bar{u} \in \text{int}C$, where $H(\cdot)$ is the Hessian of the barrier $f(\cdot)$
(vi) $\bar{u} = -\nabla f^*(\bar{s})$ if and only if $\bar{s} = -\nabla f(\bar{u})$
(vii) $-\nabla f(\bar{u})^T y \geq \sqrt{y^T H(\bar{u}) y}$ for $\bar{u} \in \text{int}C$ and $y \in C$.

Properties (i)-(vi) above are restatements of results in [21] or [26], whereas (vii) is borrowed from the proof of Lemma 5 of [23].

3. Behavioral Bounds on F

Let $\bar{s} \in \text{int}C^*$ be chosen, and let $F_{\bar{s}}$ be as defined in Section 2.1 with image set $H = H_{\bar{s}}$ as in Section 2.2. The following result shows that the width τ_F of the feasible region $\mathcal{F}$ is linearly related to the symmetry of 0 in the image set H.

Theorem 1. *Let $\bar{s} \in \text{int}C^*$ be chosen. Under the norm $\|\cdot\|_{\bar{s}}$, the width τ_F of $\mathcal{F}$ satisfies:*

$$\left(\frac{1}{\vartheta}\right) \frac{\text{sym}(0, H_{\bar{s}})}{1 + \text{sym}(0, H_{\bar{s}})} \;\leq\; \tau_F \;\leq\; \frac{\text{sym}(0, H_{\bar{s}})}{1 + \text{sym}(0, H_{\bar{s}})} \;.$$

In particular, $\frac{1}{2\vartheta}\text{sym}(0, H_{\bar{s}}) \leq \tau_F \leq \text{sym}(0, H_{\bar{s}})$.

Remark 5. The left-hand bound in the theorem depends on the complexity parameter ϑ of the barrier function $f(\cdot)$, which seems a bit unnatural since the width τ_F is a geometric object that should not directly depend on the barrier function. If we use the universal barrier of Nesterov and Nemirovskii [21], we can replace ϑ by $\text{CONST} \times n$ for the large absolute constant CONST of the universal barrier. Alternatively, we can replace ϑ by the complexity value ϑ^* of an optimal barrier for C.

Our next result shows that the computational complexity of a standard interior-point method (IPM) for computing a solution of F also depends only on $\text{sym}(0, H)$ and ϑ. In order to establish this result we first develop the model that will be solved by the IPM.

Let $\bar{s} \in C^*$ be chosen, and assign $\bar{x} \leftarrow -\frac{1}{\vartheta}\nabla f^*(\bar{s})$. It follows from Remark 4 that $\bar{x} \in \text{int}C$ and $\bar{s}^T\bar{x} = 1$. Construct the simple optimization problem:

$$\begin{array}{lll} \text{OP}: \quad \theta^* := \max\limits_{x,\theta} & \theta & \\ & Ax + (A\bar{x})\theta = 0 & \\ & \bar{s}^T x = 1 & \\ & x \in C\ , & \end{array} \tag{4}$$

and note that $(x, \theta) = (\bar{x}, -1)$ is a feasible solution of OP. Furthermore, $(\bar{x}, -1)$ is the *analytic center* associated with OP for the barrier function $f(\cdot)$, i.e., $(\bar{x}, -1)$ is the optimal solution of the problem of minimizing the barrier function $f(x)$ over the feasible region of OP. We will therefore use $(\bar{x}, -1)$ as a starting point with which to initiate a standard primal feasible interior-point method for approximately following the central path $(x(\eta), \theta(\eta))$ of the parameterized barrier problem:

$$\begin{array}{lll} \text{OP}_\eta: \quad \max\limits_{x,\theta} & -f(x) + \eta \cdot \theta & \\ & Ax + (A\bar{x})\theta = 0 & \\ & \bar{s}^T x = 1 & \\ & x \in \text{int}C & \end{array} \tag{5}$$

for an increasing sequence of values of $\eta > 0$, until we have computed a point (x, θ) that satisfies $\theta \geq 0$, whereby $(x + \theta\bar{x})$ is a feasible solution of F. The details of the algorithm scheme are presented in Algorithm A in the Appendix, where we also prove the following complexity bound for the method:

Theorem 2. *Let $\bar{s} \in \text{int}C^*$ be chosen. The standard primal-feasible interior-point Algorithm A applied to (5) will compute $\tilde{x}$ satisfying $A\tilde{x} = 0, \tilde{x} \in \text{int}C$ in at most*

$$\left\lceil 9\sqrt{\vartheta} \ln\left(11\vartheta\left(1 + \frac{1}{\text{sym}(0, H_{\bar{s}})}\right)\right)\right\rceil$$

iterations of Newton's method. Furthermore, under the norm $\|\cdot\|_{\bar{s}}$, $\tilde{x}$ will also satisfy

$$\text{reldist}_{\bar{s}}(\tilde{x}, \partial C) \geq \frac{1}{1.2\vartheta + 0.2} \cdot \tau_F\ .$$

□

(Note that the complexity bound is trivially valid even when $\text{sym}(0, H_{\bar{s}}) = 0$, using the standard convention that $1/0 = \infty$.) Taken together, Theorems 1 and 2 present bounds on both behavioral measures that are simple functions of the complexity value ϑ of the barrier function and the symmetry of 0 in the image set $H = H_{\bar{s}}$. Furthermore, Algorithm A will compute a feasible point whose relative distance from ∂C is within a factor of $O(\vartheta)$ of the maximum relative distance from ∂C over all feasible points.

Figure 1 can be used to gain some intuition on the complexity result of Theorem 2. The figure portrays the image set $H_{\bar{s}}$, which by assumption contains 0. Furthermore, $A\bar{x} \in H_{\bar{s}}$ by design of $\bar{x}$. The optimal value θ^* of (4) is the largest value of θ for which $-\theta A\bar{x} \in H_{\bar{s}}$. Also notice in Figure

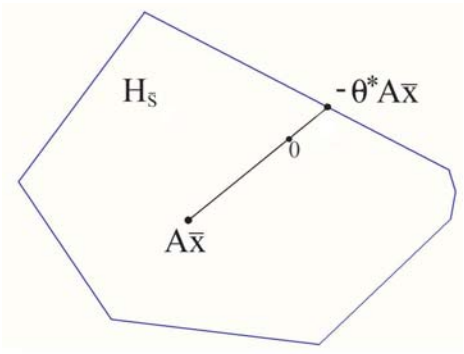

Fig. 1. The image set $H_{\bar{s}}$ and the points $A\bar{x}, 0$, and $-\theta^* A\bar{x}$.

1 that $\theta^* \geq \text{sym}(0, H_{\bar{s}})$, and so θ^* will be large if $\text{sym}(0, H_{\bar{s}})$ is large. Since the interior-point algorithm starts at the analytic center $(\bar{x}, -1)$ where $\theta = -1$ and will stop when the current iterate (x, θ) satisfies $\theta \geq 0$, it follows from the linear convergence theory of interior-point methods that the iteration bound will be proportional to the logarithm of the ratio of the initial optimality gap divided by the optimality gap at the stopping point. This ratio is simply $(1+\theta^*)/\theta^*$, which is bounded above by $1 + 1/(\text{sym}(0, H_{\bar{s}}))$.

Note that one can view $\text{sym}(0, H_{\bar{s}})$ as a condition number of sorts associated with $F_{\bar{s}}$, see [3]. In the next section, we will show how projective transformations can be used to modify $\text{sym}(0, H)$ and hence improve the behavior (both geometry and computational complexity) of $F_{\bar{s}}$.

3.1. Proof of Theorem 1

The proof of Theorem 1 is derived from the following two lemmas which we will prove in turn. For $\bar{u} \in \text{int}C$ define the ellipsoidal norm induced by $H(\bar{u})$ by $\|v\|_{\bar{u}} := \sqrt{v^T H(\bar{u}) v}$. Let $B_{\bar{s}}(c, r)$ and $B_{\bar{u}}(c, r)$ denote the balls centered at c of radius r in the norms $\|\cdot\|_{\bar{s}}$ and $\|\cdot\|_{\bar{u}}$, respectively. Note that $B_{\bar{u}}(c, r)$ is an ellipsoid whereas $B_{\bar{s}}(c, r)$ is not ellipsoidal, i.e., these two norms are not part of the same family. The following lemma shows that $\|\cdot\|_{\bar{s}}$ and $\|\cdot\|_{\bar{u}}$ are within a factor of ϑ of one another if $\bar{u} = -\nabla f^*(\bar{s})$.

Lemma 1. *Let $\bar{s} \in \text{int}C^*$ be chosen, and define $\bar{u} := -\nabla f^*(\bar{s})$. Then*

$$B_{\bar{u}}(0, 1/\vartheta) \subset B_{\bar{s}}(0, 1) \subset B_{\bar{u}}(0, 1) \quad \text{and} \quad \text{reldist}_{\bar{s}}(\bar{u}, \partial C) \geq \frac{1}{\vartheta} \ .$$

Lemma 2. *Let $\bar{s} \in \text{int}C^*$ be chosen and define $\bar{x} := -\nabla f^*(\bar{s})/\vartheta$. Then under the norm $\|\cdot\|_{\bar{s}}$,*

$$\text{reldist}_{\bar{s}}(\bar{x}, \partial C) \left(\frac{\text{sym}(0, H_{\bar{s}})}{1 + \text{sym}(0, H_{\bar{s}})} \right) \leq \text{reldist}_{\bar{s}}(\bar{x}, \partial C) \left(\frac{\theta^*}{1 + \theta^*} \right) \leq \tau_F \leq \frac{\text{sym}(0, H_{\bar{s}})}{1 + \text{sym}(0, H_{\bar{s}})} \leq \frac{\theta^*}{1 + \theta^*} \ .$$

Proof of Theorem 1: Define $\bar{x} := -\nabla f^*(\bar{s})/\vartheta$. Then $\bar{x}$ is a positive scaling of $\bar{u}$ defined in Lemma 1, and so $\text{reldist}_{\bar{s}}(\bar{x}, \partial C) = \text{reldist}_{\bar{s}}(\bar{u}, \partial C) \geq \frac{1}{\vartheta}$. Substituting this inequality into the first inequality

of Lemma 2 yields the first inequality of the theorem, and the second inequality of the theorem is simply the third inequality of Lemma 2. □

Proof of Lemma 1: Suppose that x satisfies $\|x\|_{\bar{u}} \leq 1$. Then $x^1 := \frac{1}{2}(\bar{u}+x)$ and $x^2 := \frac{1}{2}(\bar{u}-x)$ satisfy $x^1, x^2 \in C$ from the theory of self-concordance, and $x^1 - x^2 = x$, whereby from the definition of the $\bar{s}$-norm we have $\|x\|_{\bar{s}} \leq \bar{s}^T(x^1 + x^2) = \bar{s}^T\bar{u} = \vartheta$. Therefore $B_{\bar{u}}(0,1) \subset B_{\bar{s}}(0,\vartheta)$, which is equivalent to the first set inclusion of the lemma.

Let $L := \{x \in C : \bar{s}^T x = 1\}$. For any $x \in L$ we have $1 = \bar{s}^T x = -\nabla f(\bar{u})^T x \geq \sqrt{x^T H(\bar{u}) x} = \|x\|_{\bar{u}}$, where the second equality and the inequality follow from (vi) and (vii) of Remark 4, respectively. Similarly, if $x \in -L$ then $\|x\|_{\bar{u}} \leq 1$ as well. Noticing that $B_{\bar{s}}(0,1)$ is the convex hull of L and $-L$, it follows that $x \in B_{\bar{s}}(0,1)$ implies $\|x\|_{\bar{u}} \leq 1$, or equivalently, $B_{\bar{s}}(0,1) \subset B_{\bar{u}}(0,1)$.

Last of all, it follows from the theory of self-concordance that $\bar{u} + B_{\bar{u}}(0,1) \subset C$, whereby $\bar{u} + B_{\bar{s}}(0,1) \subset \bar{u} + B_{\bar{u}}(0,1) \subset C$. Therefore

$$\text{reldist}_{\bar{s}}(\bar{u}, \partial C) \geq \frac{1}{\|\bar{u}\|_{\bar{s}}} = \frac{1}{\bar{s}^T\bar{u}} = \frac{1}{\vartheta} \ ,$$

where the last equality follows from (iv) of Remark 4. □

We are grateful to Nesterov [20] for contributing to a strengthening of a previous version of Lemma 1 and its proof.

Proof of Lemma 2: Recall from Remark 4 that $\bar{x} \leftarrow -\frac{1}{\vartheta}\nabla f^*(\bar{s})$ satisfies $\bar{x} \in \text{int}C$ and $\bar{s}^T\bar{x} = 1$. Therefore $A\bar{x} \in H_{\bar{s}}$ and hence $-\text{sym}(0, H_{\bar{s}})A\bar{x} \in H_{\bar{s}}$, whereby there exists $x \in C$ satisfying $\bar{s}^T x = 1$ and $Ax = -\text{sym}(0, H_{\bar{s}})A\bar{x}$, and therefore θ^* of (4) must satisfy:

$$\theta^* \geq \text{sym}(0, H_{\bar{s}}) \ . \tag{6}$$

Therefore the second and third inequalities of the lemma imply the first and fourth inequalities of the lemma. To prove the second inequality, let (x^*, θ^*) be an optimal solution of (4), and let $x = \frac{x^* + \theta^*\bar{x}}{1+\theta^*}$. Then $x \in C$, $\|x\|_{\bar{s}} = \bar{s}^T x = 1$, and $Ax = 0$, and by construction

$$B_{\bar{s}}\left(x, \frac{\theta^*}{1+\theta^*}\text{dist}(\bar{x}, \partial C)\right) \subset C \ ,$$

whereby we have

$$\tau_F \geq \text{reldist}(x, \partial C) \geq \frac{\theta^*}{1+\theta^*}\text{dist}(\bar{x}, \partial C) = \frac{\theta^*}{1+\theta^*}\text{reldist}(\bar{x}, \partial C) \ ,$$

which demonstrates the second inequality. To prove the third inequality of the lemma, let $\hat{x} \in C$ satisfy the maximization for which τ_F is defined in (3), whereby without loss of generality $\|\hat{x}\|_{\bar{s}} = \bar{s}^T\hat{x} = 1$, $A\hat{x} = 0$, and $B_{\bar{s}}(\hat{x}, \tau_F) \subset C$. Let $y \in H_{\bar{s}}$ be given, whereby $y = Ax$ for some $x \in C$ satisfying $\|x\|_{\bar{s}} = \bar{s}^T x = 1$, and define:

$$\tilde{x} := \frac{\hat{x} - \tau_F x}{1 - \tau_F} \ ,$$

and it follows that $\tilde{x} \in C$, $\bar{s}^T\tilde{x} = 1$, and hence

$$-\frac{\tau_F}{1-\tau_F} y = A\tilde{x} \in H_{\bar{s}} \ .$$

Therefore $\mathrm{sym}(0, H_{\bar{s}}) \geq \frac{\tau_F}{1-\tau_F}$, and rearranging this last inequality yields the third inequality of the lemma. □

4. Pre-conditioning F by Projective Transformation of $F_{\bar{s}}$

Herein we present a systematic approach to transforming the problem $F_{\bar{s}}$ to an equivalent problem $F_{\hat{s}}$ for a suitably chosen vector $\hat{s} \in \mathrm{int}C^*$, with the goal of improving the symmetry of 0 in the associated image set $H_{\hat{s}}$. We first review some relevant facts about the symmetry function $\mathrm{sym}(\cdot)$:

Remark 6. Let $S \subset \mathbb{R}^m$ be a nonempty closed bounded convex set. The following properties of $\mathrm{sym}(\cdot)$ are shown in [1]:
(i) Let $\mathcal{A}(x) := Mx + g$, M nonsingular. If $\bar{x} \in S$, then $\mathrm{sym}(\mathcal{A}(\bar{x}), \mathcal{A}(S)) = \mathrm{sym}(\bar{x}, S)$.
(ii) If $0 \in S$, then $\mathrm{sym}(0, S) = \mathrm{sym}(0, S^\circ)$.
(iii) $\mathrm{sym}(S) \geq \frac{1}{m}$.

Under Assumption 2, $0 \in H_{\bar{s}}$, whereby $H_{\bar{s}}$ is a closed convex set containing the origin. Therefore $H_{\bar{s}}^\circ$, the polar of $H_{\bar{s}}$, is also a closed convex set containing the origin, and $H_{\bar{s}}^{\circ\circ} = H_{\bar{s}}$. In fact, there is a simple form for $H_{\bar{s}}^\circ$ given by the following:

Proposition 1. *Let $\bar{s} \in \mathrm{int}C^*$ be chosen. Then $H_{\bar{s}}^\circ = \{v \in \mathbb{R}^m : \bar{s} - A^T v \in C^*\}$.*

Proof. We have:

$$\begin{aligned} H_{\bar{s}}^\circ &= \{v : v^T y \leq 1 \ \text{ for all } y \in H_{\bar{s}}\} \\ &= \{v : v^T Ax \leq 1 \ \text{ for all } x \text{ that satisfy } \bar{s}^T x = 1, x \in C\} \\ &= \{v : v^T Ax \leq \bar{s}^T x \ \text{ for all } x \text{ that satisfy } \bar{s}^T x = 1, x \in C\} \\ &= \{v : (\bar{s} - A^T v)^T x \geq 0 \ \text{ for all } x \in C\} \\ &= \{v : \bar{s} - A^T v \in C^*\} \ . \end{aligned}$$

□

It is curious to note from Proposition 1 that while checking membership in $H_{\bar{s}}$ is presumably not easy (validating that $0 \in H_{\bar{s}}$ is an equivalent task to that of solving F), the set $H_{\bar{s}}^\circ$ is in fact easy to work with in at least two ways. First, $0 \in \mathrm{int}H_{\bar{s}}^\circ$, so we have a known point in the interior of $\mathrm{int}H_{\bar{s}}^\circ$. Second, checking membership in $H_{\bar{s}}^\circ$ is a relatively simple task if we have available a membership oracle for C^*.

Motivated by Theorems 1 and 2 which bound the geometric and computational behavior of F in terms of the symmetry of the origin in $H_{\bar{s}}$, we now consider replacing $\bar{s}$ by some other vector $\hat{s} \in \mathrm{int}C^*$ with the goal of improving $\mathrm{sym}(0, H_{\hat{s}})$. We proceed as follows. Taking $\bar{s} \in \mathrm{int}C^*$ as given,

suppose we choose some $\hat{v} \in \text{int}H_{\bar{s}}^{\circ} = \{v \in \mathbb{R}^m : \bar{s} - A^T v \in C^*\}$, and define $\hat{s} := \bar{s} - A^T\hat{v}$, therefore $\hat{s} \in \text{int}C^*$. We replace $\bar{s}$ by $\hat{s}$, obtaining the modified normalized feasibility problem:

$$F_{\hat{s}} : \begin{cases} Ax = 0 \\ \hat{s}^T x = 1 \\ x \in C \ , \end{cases}$$

with modified image set:

$$H_{\hat{s}} = \{Ax : x \in C \ , \ \hat{s}^T x = 1\}$$

and polar set:

$$H_{\hat{s}}^{\circ} = \{v \in \mathbb{R}^m : \hat{s} - A^T v \in C^*\} \ .$$

The following shows that $\text{sym}(0, H_{\hat{s}})$ inherits the symmetry of $\hat{v}$ in the original polar image set $H_{\bar{s}}^{\circ}$.

Theorem 3. *Let $\bar{s} \in \text{int}C^*$ be given. Let $\hat{v} \in \text{int}H_{\bar{s}}^{\circ}$ be chosen and define $\hat{s} := \bar{s} - A^T\hat{v}$. Then*

$$\text{sym}(0, H_{\hat{s}}) = \text{sym}(\hat{v}, H_{\bar{s}}^{\circ}) \ .$$

Proof. We have:

$$H_{\bar{s}}^{\circ} - \{\hat{v}\} = \{u = v - \hat{v} : \bar{s} - A^T v \in C^*\} = \{u : \bar{s} - A^T\hat{v} - A^T u \in C^*\} = \{u : \hat{s} - A^T u \in C^*\} = H_{\hat{s}}^{\circ} \ .$$

It then follows from (ii) of Remark 6 that

$$\text{sym}(0, H_{\hat{s}}) = \text{sym}(0, H_{\hat{s}}^{\circ}) = \text{sym}(0, H_{\bar{s}}^{\circ} - \{\hat{v}\}) = \text{sym}(\hat{v}, H_{\bar{s}}^{\circ}) \ ,$$

where the last equality above readily follows from (i) of Remark 6. □

Note that the following projective transformations map $F_{\bar{s}}$ and $F_{\hat{s}}$ onto one another:

$$x' \ \leftarrow \ \frac{x}{\hat{s}^T x} \ \text{ and } \ x \ \leftarrow \ \frac{x'}{\bar{s}^T x'} \ . \tag{7}$$

Furthermore, Theorem 3 has an interesting interpretation in the context of projective transformations and polarity theory which we will discuss in Subsection 4.2.

Our present goal, however, is to use Theorem 3 constructively to develop a method for transforming $F_{\bar{s}}$. Suppose we can compute a point $\hat{v} \in H_{\bar{s}}^{\circ}$ with good symmetry in $H_{\bar{s}}^{\circ}$; letting $\alpha := \text{sym}(\hat{v}, H_{\bar{s}}^{\circ})$ we seek $\hat{v}$ for which $\alpha > \text{sym}(0, H_{\bar{s}})$ and is relatively large, for example, $\alpha = \Omega(1/m)$. Then replace $\bar{s}$ by $\hat{s} := \bar{s} - A^T\hat{v}$ and work instead with with $F_{\hat{s}}$. Theorem 3 states that the transformed system will have $\text{sym}(0, H_{\hat{s}}) = \alpha$, i.e., the transformed system will take on the symmetry of $\hat{v}$ in $H_{\bar{s}}^{\circ}$. This is most important, since it then follows from Theorems 1 and 2 that the transformed system $F_{\hat{s}}$ will have geometry and complexity behavior that will depend on α as well. We formalize this method and the above conclusion as follows:

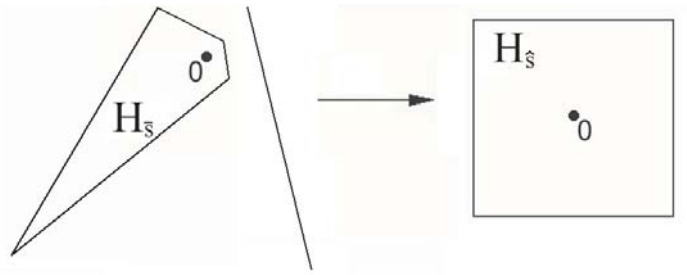

Fig. 2. Projective transformation of the image set $H_{\bar{s}}$ to improve the symmetry of 0 in the transformed image set.

Projective Pre-conditioning Method (PPM)
Step 1. Construct $H_{\bar{s}}^{\circ} := \{v \in \mathbb{R}^m : \bar{s} - A^T v \in C^*\}$
Step 2. Find a suitable point $\hat{v} \in H_{\bar{s}}^{\circ}$ (with hopefully good symmetry in $H_{\bar{s}}^{\circ}$)
Step 3. Compute $\hat{s} := \bar{s} - A^T \hat{v}$
Step 4. Construct the transformed problem:

$$F_{\hat{s}} : \begin{cases} Ax = 0 \\ \hat{s}^T x = 1 \\ x \in C \end{cases} \tag{8}$$

Step 5. The transformed image set is $H_{\hat{s}} := \{Ax \in \mathbb{R}^m : x \in C \ , \ \hat{s}^T x = 1\}$, and $\mathrm{sym}(0, H_{\hat{s}}) = \mathrm{sym}(\hat{v}, H_{\bar{s}}^{\circ})$.

Figure 2 illustrates the strategy of the Projective Pre-Conditioning Method. On the left part of the figure is the image set $H_{\bar{s}}$, and notice that $H_{\bar{s}}$ is not very symmetric about the origin, i.e., $\mathrm{sym}(0, H_{\bar{s}}) << 1$. However, under the projective transformation given by the projective plane in the slanted vertical line, $H_{\bar{s}}$ is transformed to a box that is perfectly symmetric about the origin, i.e., $\mathrm{sym}(0, H_{\hat{s}}) = 1$. (In general, of course, we can at best attain $\mathrm{sym}(0, H_{\hat{s}}) = 1/m$.)

The following corollary follows from Theorem 3 and the above discussion, using Theorems 1 and 2:

Corollary 1. *Let $\bar{s} \in \mathrm{int}C^*$ be chosen, and suppose that the Projective Pre-conditioning Method has been run, and let $\alpha := \mathrm{sym}(\hat{v}, H_{\bar{s}}^{\circ})$. Under the norm $\|\cdot\|_{\hat{s}}$, the width τ_F of $\mathcal{F}$ satisfies:*

$$\left(\frac{1}{\vartheta}\right)\frac{\alpha}{1+\alpha} \ \leq \ \tau_F \ \leq \ \frac{\alpha}{1+\alpha} \ ,$$

and the standard primal-feasible interior-point Algorithm A applied to (8) will compute $\tilde{x}$ satisfying $A\tilde{x} = 0, \tilde{x} \in \mathrm{int}C$ in at most

$$\left\lceil 9\sqrt{\vartheta}\ln\left(11\vartheta\left(1+\frac{1}{\alpha}\right)\right)\right\rceil$$

iterations of Newton's method. Furthermore, under the norm $\|\cdot\|_{\hat{s}}$, $\tilde{x}$ will also satisfy

$$\mathrm{reldist}_{\hat{s}}(\tilde{x}, \partial C) \geq \frac{1}{1.2\vartheta + 0.2} \cdot \tau_F \ .$$

Let us now presume that F has an interior solution, whereby $0 \in \text{int}H_{\bar{s}}$ and $H_{\bar{s}}^\circ$ will be bounded and $\text{sym}(0, H_{\bar{s}}) = \text{sym}(0, H_{\bar{s}}^\circ) > 0$. Furthermore, we know from (iii) of Remark 6 that there exists a point v whose symmetry value in $H_{\bar{s}}^\circ$ is at least $1/m$. Notice that if we can generate a point $\hat{v} \in H_{\bar{s}}^\circ$ with very good symmetry in $H_{\bar{s}}^\circ$ in the sense that $\alpha := \text{sym}(\hat{v}, H_{\bar{s}}^\circ) = \Omega(1/m)$, we can then compute $\tilde{x}$ of Corollary 1 using at most $O\left(\sqrt{\vartheta}\ln(\vartheta \cdot m)\right)$ Newton steps, which is strongly polynomial-time. And even if we merely satisfy $\alpha := \text{sym}(\hat{v}, H_{\bar{s}}^\circ) > \text{sym}(0, H_{\bar{s}})$, we still may improve the computational effort needed to solve F by working with $F_{\hat{s}}$ rather than $F_{\bar{s}}$.

Of course, the effectiveness of this method depends entirely on the ability to efficiently compute a point $\hat{v} \in H_{\bar{s}}^\circ$ with good symmetry. The set $H_{\bar{s}}^\circ$ has the convenient representation $H_{\bar{s}}^\circ = \{v \in \mathbb{R}^m : \bar{s} - A^T v \in C^*\}$ from Proposition 1; furthermore, we have a convenient point $0 \in \text{int}H_{\bar{s}}^\circ$ with which to start a method for finding a point with good symmetry; also, testing membership in $H_{\bar{s}}^\circ$ depends only on the capability of testing membership in C^*. Thus, the relative ease with which we can work with $H_{\bar{s}}^\circ$ suggests that excessive computation might not be necessary in order to compute a point $\hat{v}$ with good symmetry in $H_{\bar{s}}^\circ$. We explore several different approaches for computing such a point $\hat{v}$ in the following subsection.

4.1. Strategies for Computing Points with Good Symmetry in $H_{\bar{s}}^\circ$

In this subsection we presume that F has an interior solution, whereby $0 \in \text{int}H_{\bar{s}}$ and $H_{\bar{s}}^\circ$ will be bounded and $\text{sym}(0, H_{\bar{s}}^\circ) > 0$. Recall that a symmetry point of a convex set $S \subset \mathbb{R}^m$ is a point x^* whose symmetry is optimal on S. From (iii) of Remark 6 we know that $\text{sym}(x^*, S) \geq 1/m$. When $C = \mathbb{R}_+^n$, a symmetry point of $H_{\bar{s}}^\circ$ can be computed by approximately solving $n+1$ linear programs using the method developed in [1]. Thus even for the case of $C = \mathbb{R}_+^n$ the computational burden of finding a point with guaranteed good symmetry appears to be excessive. In fact, the seemingly simpler task of just evaluating $\text{sym}(x, S)$ at a particular point $x = \bar{x}$ might be hard for a general convex body S, see [1]. Therefore, one is led to investigate heuristic and/or probabilistic methods, or perhaps methods that compute other types of points that lie "deep" in a convex body. Table 1 presents the symmetry guarantee for several types of deep points in a convex body; the computational effort for $C = \mathbb{R}_+^n$ ($H_{\bar{s}}^\circ$ is the intersection of n half-spaces) is shown in the third column of the table. We now briefly discuss each of these three possible choices for such points.

4.1.1. Analytic Center Approach. Starting at $v = 0$, we could use the damped Newton method outlined in [21] to compute an approximate analytic center of $H_{\bar{s}}^\circ$ using the barrier function $b(v) := f^*(\bar{s} - A^T v)$. We know from the theory of self-concordant barriers that the analytic center v^a of $H_{\bar{s}}^\circ$ has symmetry value at least $1/\sqrt{\vartheta(\vartheta - 1)}$ (see Lemma 5 of the Appendix of Nesterov, Todd, and Ye [23]). Each iteration of the algorithm would be of comparable computational burden as an interior-point iteration for the problem OP_μ and so it would probably be wisest to perform only a few iterations and hope that the final iterate v^f would have good symmetry value in $H_{\bar{s}}^\circ$ nevertheless.

4.1.2. Löwner-John Center Approach. The Löwner-John theorem guarantees the existence of an m-rounding of $H_{\bar{s}}^{\circ}$, i.e., an ellipsoid E centered at the origin and a point v^j with the property that $\{v^j\} + E \subset H_{\bar{s}}^{\circ} \subset \{v^j\} + mE$, see [12]. Such a point v^j is called a Löwner-John center and it follows that $\mathrm{sym}(v^j, H_{\bar{s}}^{\circ}) \geq 1/m$. In the case when C is either the nonnegative orthant $\mathbb{R}_+^n$ or is the cartesian product of second-order cones, we can compute such an approximate Löwner-John center by computing the center of the maximum volume inscribed ellipsoid in $H_{\bar{s}}^{\circ}$ via semidefinite programming (see Zhang and Gao [29], for example, for the case when $C = \mathbb{R}_+^n$). The problem with this approach is that the computational effort is likely to be substantially larger than that of solving the original problem $F_{\bar{s}}$, and the approach is limited to the case when C is the cartesian product of half-lines and/or second-order cones.

4.1.3. Center of Mass Approach. The center of mass (or centroid) of a convex body $S \subset \mathbb{R}^m$ will have guaranteed symmetry of at least $1/m$, see [14]. Even when $C = \mathbb{R}_+^n$, computing the center of mass of $H_{\bar{s}}^{\circ}$ is #P-hard, see [2]. However, if we instead consider nondeterministic algorithms, then the recent work of Lovász and Vempala [16–18] points the way to computing points near the center of mass with high probability with good theoretical efficiency. This approach will be developed in more detail in Section 5.

Central Point	Symmetry Guarantee	Computational Effort when $C = \mathbb{R}_+^n$
Symmetry Point	$1/m$	$\approx$ LP $\times$ $(n+1)$
Analytic Center	$\frac{1}{\sqrt{\vartheta(\vartheta-1)}}$	$\approx$ LP
Löwner-John Center	$1/m$	$\approx$ SDP
Center of Mass	$1/m$	#P-Hard (deterministic) $\approx$ polynomial-time (stochastic)

Table 1. Summary Properties of Strategies for Computing Deep Points in $H_{\bar{s}}^{\circ}$

4.2. Polarity and Projective Transformations in Theorem 3

While it is obvious that $F_{\bar{s}}$ and $F_{\hat{s}}$ are related through the pair of projective transformations (7), it is perhaps not so obvious that the image sets $H_{\bar{s}}$ and $H_{\hat{s}}$ are related via projective transformations: $H_{\bar{s}}$ maps onto $H_{\hat{s}}$ with the following projective transformations between points $y \in H_{\bar{s}}$ and $y' \in H_{\hat{s}}$:

$$y' = T(y) := \frac{y}{1 - \hat{v}^T y} \quad \text{and} \quad y = T^{-1}(y') = \frac{y'}{1 + \hat{v}^T y'} \ . \tag{9}$$

This pair of projective transformations results from a more general theory concerning the polarity construction, translations, and projective transformations as follows, see Grünbaum [13] for example. Let S be a closed convex set containing the origin. ($S = H_{\bar{s}}$ in our context.) Then S° is a closed convex set containing the origin and $S^{\circ\circ} = S$. Let $\hat{v} \in \mathrm{int} S^{\circ}$ be given. Then $(S^{\circ} - \{\hat{v}\})$ is the translation of S

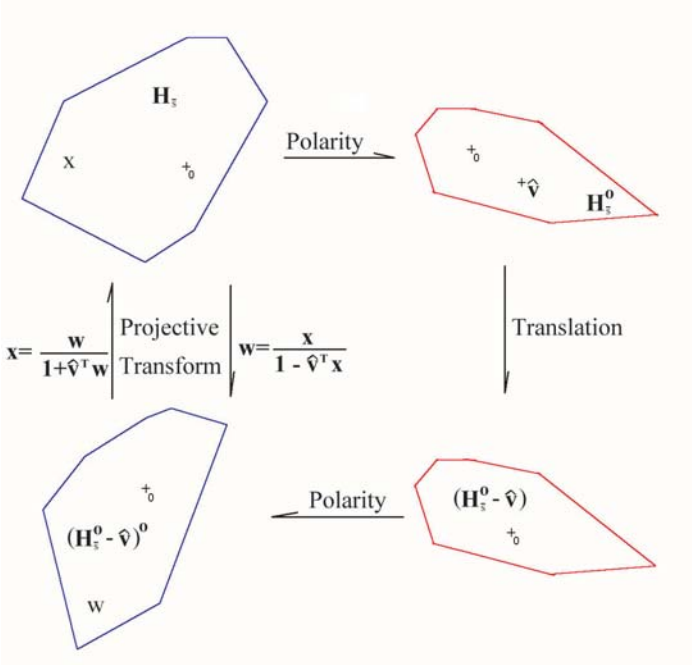

Fig. 3. Translation of the polar set corresponds to projective transformation of the original set.

by $\hat{v}$, and also is a closed convex set containing the origin, and its polar is $(S^\circ - \{\hat{v}\})^\circ$. It is elementary arithmetic to show that S and $(S^\circ - \{\hat{v}\})^\circ$ are related through the projective transformations (9), namely $(S^\circ - \{\hat{v}\})^\circ = T(S)$ and $S = T^{-1}((S^\circ - \{\hat{v}\})^\circ)$. In other words, translation of the polar set corresponds to projective transformation of the original set, see Figure 3. This correspondence was previously used in [5,6].

5. Approximate Center of Mass of $H_{\bar{s}}^\circ$ and its Symmetry

In this section we present some general results about sampling from the uniform distribution on a given convex body $S \subset \mathbb{R}^d$, which are relevant for our particular case where $S = H_{\bar{s}}^\circ$ and $d = m$. We proceed as follows. A function $f : \mathbb{R}^d \to \mathbb{R}_+$ is said to be logconcave if $\log f$ is a concave function. A random variable $Z \in \mathbb{R}^d$ is called a logconcave random variable if the probability density function of Z is a logconcave function. Note that logconcave random variables are a broad class that includes Gaussian, exponential, and uniformly distributed random variables on convex sets.

The center of mass (or centroid) and covariance matrix associated with Z are given respectively by

$$\mu_Z := E[Z] \qquad \text{and} \qquad \Sigma_Z := E[(Z - \mu_Z)(Z - \mu_Z)^T] \ .$$

The matrix Σ_Z is symmetric positive semi-definite. If Σ_Z is positive definite it can be used to define the ellipsoidal norm:

$$\|v\|_{\Sigma_Z} := \sqrt{v^T \Sigma_Z^{-1} v} \ .$$

The following are very useful properties of logconcave random variables.

Lemma 3. [15, 24, 25] *The sum of independent logconcave random variables is a logconcave random variable.*

Lemma 4. [16] *Let Z be a logconcave random variable in $\mathbb{R}^d$. Then for any $R \geq 0$:*

$$\mathsf{P}\left(\|Z - \mu_Z\|_{\Sigma_Z} \geq R\sqrt{d}\right) \leq e^{-R}.$$

Now let X be a random variable in $\mathbb{R}^d$ uniformly distributed on a convex body S, i.e., the probability density function of X is given by

$$f(x) = \frac{1}{\mathsf{Vol}(S)} 1_S(x) \ , \tag{10}$$

where $1_S(\cdot)$ is the indicator function of the set S. For simplicity, we denote its center of mass and covariance matrix respectively by μ and Σ, and note that Σ is positive definite since S has a non-empty interior. Let $B_\Sigma(x, r)$ denote the ball centered at x with radius r in the norm $\|\cdot\|_\Sigma$.

Lemma 5. [16] *Let X be a random variable uniformly distributed on a convex body $S \subset \mathbb{R}^d$. Then*

$$B_\Sigma\left(\mu \ , \sqrt{(d+2)/d}\right) \subset S \subset B_\Sigma\left(\mu \ , \sqrt{d(d+2)}\right) \ .$$

Assume that we are given M independent uniformly distributed random points $v^1, v^2, \ldots, v^M$ on the convex body S. We define the sample mean the usual way:

$$\hat{v} := \frac{1}{M} \sum_{i=1}^{M} v^i \ .$$

Lemma 6. *Let $\hat{v}$ be the sample mean of M independent uniformly distributed points on the convex body $S \subset \mathbb{R}^d$. Then*

$$\mathrm{sym}(\hat{v}, S) \geq \frac{\sqrt{(d+2)/d} - \|\hat{v} - \mu\|_\Sigma}{\sqrt{d(d+2)} + \|\hat{v} - \mu\|_\Sigma} \ .$$

Proof. Consider any chord of S that passes through $\hat{v}$. It is divided by $\hat{v}$ into two segments of length s_1 and s_2. From Lemma 5 it follows that $B_\Sigma\left(\mu \ , \sqrt{(d+2)/d}\right) \subset S \subset B_\Sigma\left(\mu \ , \sqrt{d(d+2)}\right)$. Thus, we can bound the ratio of s_1 to s_2 by

$$\frac{s_1}{s_2} \geq \frac{\sqrt{(d+2)/d} - \|\hat{v} - \mu\|_\Sigma}{\sqrt{d(d+2)} + \|\hat{v} - \mu\|_\Sigma} \ .$$

□

Theorem 4. *Let $\hat{v}$ be the sample mean of M independent uniformly distributed points on the convex body $S \subset \mathbb{R}^d$. Then for any $t \geq 0$ it holds that $\mathsf{P}(\|\hat{v} - \mu\|_\Sigma \geq t) \leq e^{-t\sqrt{\frac{M}{d}}}$.*

Proof. Let $Y = \sqrt{M}\hat{v}$. Since $v^1, v^2, \ldots, v^M$ are independent uniformly distributed random variables on S, $E[Y] = \sqrt{M}\mu$ and Σ is the covariance matrix of Y. Moreover, using Lemma 3, Y is a logconcave random variable since it is a sum of independent logconcave random variables. Applying Lemma 4 using $R = t\sqrt{\frac{M}{d}}$ we obtain:

$$\mathsf{P}\left(\|\hat{v} - \mu\|_\Sigma \geq t\right) = \mathsf{P}\left(\|\sqrt{M}\hat{v} - \sqrt{M}\mu\|_\Sigma \geq t\sqrt{M}\right) = \mathsf{P}\left(\|Y - E[Y]\|_\Sigma \geq R\sqrt{d}\right) \leq e^{-R} = e^{-t\sqrt{\frac{M}{d}}} .$$

□

Corollary 2. *For any $\delta \in (0,1)$ and setting $M = 4d\left(\ln(1/\delta)\right)^2$, we have*

$$\mathrm{sym}(\hat{v}, S) \geq \frac{1}{2d+3}$$

with probability at least $1 - \delta$.

Proof. Using Theorem 4 with $M = 4d\left(\ln(1/\delta)\right)^2$ and $t = 1/2$ we obtain $\mathsf{P}(\|\hat{v} - \mu\|_\Sigma \geq 1/2) \leq \delta$, whereby $\mathsf{P}(\|\hat{v} - \mu\|_\Sigma \leq 1/2) \geq 1 - \delta$. Finally, using Lemma 6 we obtain:

$$\mathrm{sym}(\hat{v}, S) \geq \frac{1 - 1/2}{d + 1 + 1/2} = \frac{1}{2d+3}$$

with probability at least $1 - \delta$. □

Remark 7. The proof of Corollary 2 can be extended to show that setting $M = \left(\frac{1+\frac{1-\varepsilon}{d}}{\varepsilon}\right)^2 d\left(\ln(1/\delta))^2\right)$ we obtain $\mathrm{sym}(\hat{v}, S) \geq \frac{1-\varepsilon}{d}$ with probability at least $1 - \delta$.

Keeping in mind the fact that $\mathrm{sym}(S)$ can only be guaranteed to be at most $1/d$ (and this bound is attained, for example, for a d-dimensional simplex), Corollary 2 gives an upper bound on the number of points that must be sampled to obtain a point $\hat{v}$ whose symmetry is bounded below by $\Omega(1/d)$ with high probability. Specializing to the case of $S = H_{\bar{s}}^\circ$ and $d = m$ and presuming that F has an interior solution (and hence $H_{\bar{s}}^\circ$ is a convex body), Corollary 2 provides a mechanism for achieving $\mathrm{sym}(\hat{v}, H_{\bar{s}}^\circ) = \Omega(1/m)$, and hence achieving $\mathrm{sym}(0, H_{\hat{s}}) = \Omega(1/m)$ with high probability (from Theorem 3). It follows from Corollary 1 and Corollary 2 that in the context of the Projective Pre-conditioning Method presented in Section 4, with high probability (i.e., probability at least $1-\delta$) we attain a complexity bound for solving $F_{\hat{s}}$ of

$$\left\lceil 9\sqrt{\vartheta}\ln\left(11\vartheta\left(2m+4\right)\right)\right\rceil$$

iterations of Newton's method. This iteration bound is strongly polynomial-time (with high probability). In order to make Corollary 2 constructive, we need a method for sampling on a convex body that obeys an approximate uniform distribution, which is discussed in the following subsection.

5.1. Sampling from the Uniform Distribution on a Convex Body

Herein we outline some relevant theory about uniform sampling on a convex body $S \subset \mathbb{R}^d$, see [16], [17], and [18] for recent results on this problem. We describe a standard sampling algorithm specialized to the structure of our application. To generate a random point distributed approximately uniformly on S, we will use a geometric random walk algorithm on S. The implementation of the algorithm requires only the use of a membership oracle for S and an initial point $X^0 \in S$. In the context of the Projective Pre-conditioning Method of Section 4, where $S = H_{\bar{s}}^{\circ}$ and $d = m$, the initial point is $0 \in H_{\bar{s}}^{\circ}$. The requirement of a membership oracle for $S = H_{\bar{s}}^{\circ}$ is met if we have a membership oracle for the dual cone C^* as discussed earlier.

The geometric random walk algorithm known as "Hit-and-Run" (see [17]) generates iterates $X^1, X^2, \ldots$, as follows:

Geometric Random Walk Algorithm
Step 1. Initialize with $X^0 \in S$, $k = 0$
Step 2. Choose s uniformly on the unit sphere S^{d-1} in $\mathbb{R}^d$
Step 3. Let X^{k+1} be chosen uniformly on the line segment $S \cap \{X^k + ts : t \in \mathbb{R}\}$
Step 4. Set $k \leftarrow k + 1$, goto **Step 2**

It is a simple and well known result that this random walk induces a Markov chain whose stationary distribution is the uniform distribution on S. The rate of convergence to the stationary distribution depends on the spectral gap, i.e., the difference between the two largest eigenvalues of the transition kernel. Suppose that we run the algorithm for N iterations. In [17] it is proved that to achieve an ε-approximation to the uniform distribution density function (10) in the L_1 norm, it is sufficient that

$$N = O\left(d^3 \left(\frac{R}{r}\right)^2 \ln\left(\frac{R}{\varepsilon \cdot \mathrm{dist}_2(0, \partial S)}\right)\right) ,$$

where r, R satisfy $B_2(w, r) \subset S \subset B_2(v, R)$ for some $w, v \in S$, and $B_2(c, \delta)$, $\mathrm{dist}_2(v, T)$ are the Euclidean ball centered at c with radius δ and the Euclidean distance from v to T, respectively.

Note that Step 3 of the algorithm requires that one computes the end points of the line segment in S that passes through X^k and has direction s. This can be done by binary search using a membership oracle for S. In our case $S = H_{\bar{s}}^{\circ} = \{v : \bar{s} - A^T v \in C^*\}$ and the required membership oracle for S is met if we have a membership oracle for C^*. For self-scaled cones the endpoints computation in Step 3 is a standard computation: when $C = \mathbb{R}^n_+$ the endpoints computation is a min-ratio test, when C is the cross-product of second-order cones the endpoints computation uses the quadratic formula, and when C is the positive semidefinite cone the endpoints computation is a min-ratio test of the eigenvalues of a a matrix obtained after a single Cholesky factorization.

6. Computational Results on Randomly Generated Poorly-Behaved Problems

We performed computational experiments to assess the practical viability of the projective pre-conditioning method (PPM). We tested the PPM on 300 artificially generated homogeneous linear programming feasibility problems (i.e., $C = \mathbb{R}^n_+$). These 300 problems were comprised of 100 problems each of dimensions $(m, n) = (100, 500)$, $(500, 2500)$, and $(1000, 5000)$, and were generated so as to guarantee that the resulting problems would be poorly behaved. Each problem is specified by a matrix A and the chosen value of $\bar{s}$. We first describe how A was generated. Given a pre-specified density value DENS for A, each element of A was chosen to be 0 with probability $1-$DENS, otherwise the element was generated as a standard Normal random variable. We used DENS $= 1.0$, 0.01, and 0.01 for the problem dimensions $(m, n) = (100, 500)$, $(500, 2500)$, and $(1000, 5000)$, respectively. The vector $\bar{s}$ was chosen in a manner that would guarantee that the problem would be poorly behaved as follows. Starting with $s^0 = e$, we created the polar image set $H^\circ_{s^0} = \{v : A^T v \leq e\}$. We randomly generated a non-zero vector $d \in \mathbb{R}^m$ and performed a min-ratio test to compute $\bar{t} > 0$ for which $\bar{t}A^T d \in \partial H^\circ_{s^0}$. Then $\bar{s}$ is determined by the formula:

$$\bar{s} = s^0 - (1 - 4 \times 10^{-5})\bar{t}A^T d \ .$$

This method is essentially the reverse process of the PPM, and yields $\text{sym}(0, H_{\bar{s}}) \leq 4 \times 10^{-5}$, with resulting poor geometry from Theorem 1.

We implemented the projective pre-conditioning method (PPM) using the following simplified version of the stochastic process described in Section 5: starting from $v^0 = 0 \in \text{int}H^\circ_{\bar{s}}$ we take K steps of the geometric random walk algorithm, yielding points $v^1, \ldots, v^K$, and computed $\hat{v} := \frac{1}{K}\sum_{i=1}^K v^i$, and then set $\hat{s} = \bar{s} - A^T\hat{v}$. We set $K = 30$. It is also well known that this simple method yields $\hat{v} \to \mu$ as $K \to \infty$ and that the convergence results are similar to those described in Section 5. Nonetheless, the theoretical analysis is more technical and requires additional notation and so is omitted here. See [4] and [10] for discussion and further references.

We solved the 300 original problem instances of OP (stopping as soon as $\theta \geq 0$), as well as the resulting instances after pre-conditioning, using the interior-point software SDPT3 [28]. Table 2 summarizes our computational results. Because these problems are feasibility problems the number of IPM iterations is relatively small, even for the original problems. Notice that average IPM iterations shows a marked decrease in all three dimension classes, and in particular shows a 46% decrease in average IPM iterations for the 100 problem instances of dimension 1000×5000. The total running time (which includes the time for pre-conditioning using the geometric random walk) also shows a marked decrease when using the projective pre-conditioning method, and in particular shows a 33% decrease for the 100 problem instances of dimension 1000×5000. The last two columns of Table 2 shows the average value of θ^*. Given that $(x, \theta) = (\bar{x}, -1)$ is a feasible starting point for OP (and for SDPT3), θ^* is a good measure of the computational difficulty of a problem instance – a problem is poorly behaved to the extent that θ^* is close to zero. Here we see, regardless of any IPM, that θ^* increases by a factor of roughly 400, 800, and 600 for the problem dimensions $(m, n) = (100, 500)$, $(500, 2500)$, and $(1000, 5000)$, respectively. These results all demonstrate that by taking only a small

number of steps of the geometric random walk algorithm, one can greatly improve the behavior of a poorly-behaved problem instance, and hence improve the practical performance of an IPM for solving the problem instance.

Dimensions		Average IPM Iterations		Average Total Running Time (secs.)		Average Value of θ^*	
		Original	After	Original	After	Original	After
m	n	Problem	Pre-conditioning	Problem	Pre-conditioning	Problem	Pre-conditioning
100	500	8.52	4.24	0.5786	0.2983	0.0020	0.8730
500	2500	9.30	5.17	2.4391	2.0058	0.0012	1.0218
1000	5000	9.69	5.20	22.9430	15.3579	0.0019	1.1440

Table 2. Average Performance of SDPT3 on the 300 Problem Test-bed of Linear Programming Feasibility Problems. Computation was performed on a laptop computer running Windows XP.

We also explored the sensitivity of the computational performance of the PPM to the number of steps of the random walk. Figure 4 shows the median number of IPM iterations as well as the 90% band (i.e., the band excluding the lower 5% and the upper 5%) of IPM iterations for the 100 problems of dimension 100×500 before and after pre-conditioning. Notice that only 10 steps of the random walk are needed to reduce the median and variance of the IPM iterations to a very low level. As the number of random walk steps increase, the number of IPM iterations quickly concentrates and converges to a value below the 0.05 quantile for the original problem instances.

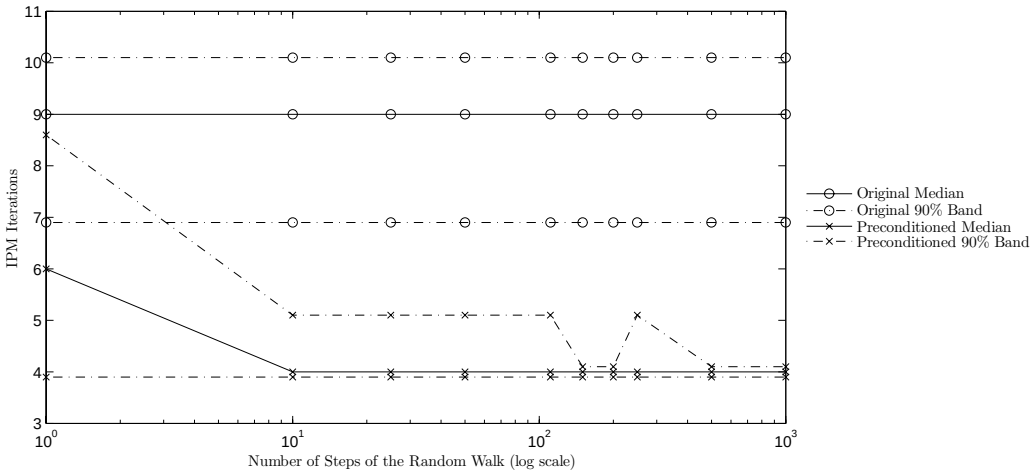

Fig. 4. IPM iterations versus number of steps of the geometric random walk for the 100 problem instances of dimension 100×500.

Figure 5 shows the median value of θ^* as well as the 90% band of θ^* values for the 100 problems of dimension 100×500 before and after pre-conditioning. As discussed earlier, θ^* is a good measure of

problem instance behavior: larger values of θ^* indicate that the problem instance is better behaved, especially for computation via an IPM. The figure indicates that there is almost no improvement in of median value of θ^* after 50 steps of the random walk.

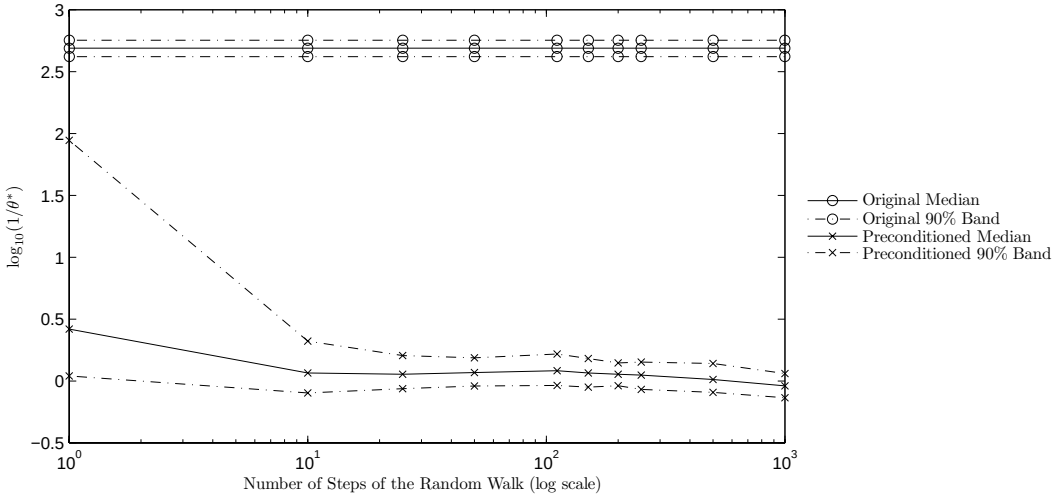

Fig. 5. $\log(\theta^*)$ versus number of steps of the geometric random walk for the 100 problem instances of dimension 100×500.

Figure 6 shows the median total running time as well as the 90% band of total running times for the 100 problems of dimension 100×500 before and after pre-conditioning. Notice that the median running time of the system with pre-conditioning rapidly decreases with a flat bottom in the range 10-100 steps of the random walk, after which the cost of the random walk steps exceeds the average benefit from computing a presumably better pre-conditioner. Also notice, however, that the variation in running time decreases with the number of random steps, which may offer some advantage in lowering the likelihood of outlier computation times by using more random walk steps.

7. Summary/Conclusions/Other Matters

In this paper we have presented a general theory for transforming a normalized homogeneous conic system $F_{\bar{s}}$ to an equivalent system $F_{\hat{s}}$ via projective transformation induced by the choice of a point $\hat{v} \in H^\circ_{\bar{s}}$. Such a projective transformation serves to pre-condition the conic system into a system that has both geometric and computational behavior with certain guarantees. We have given a characterization of both the geometric behavior and the computational behavior of the transformed system as a function of the symmetry of $\hat{v}$ in the image set $H^\circ_{\bar{s}} = \{v : \bar{s} - A^T v \in C^*\}$. Because $H^\circ_{\bar{s}}$ must contain a point v whose symmetry is at least $1/m$, if we can find a point whose symmetry is $\Omega(1/m)$ then we can projectively transform the conic system to one whose geometric and computational

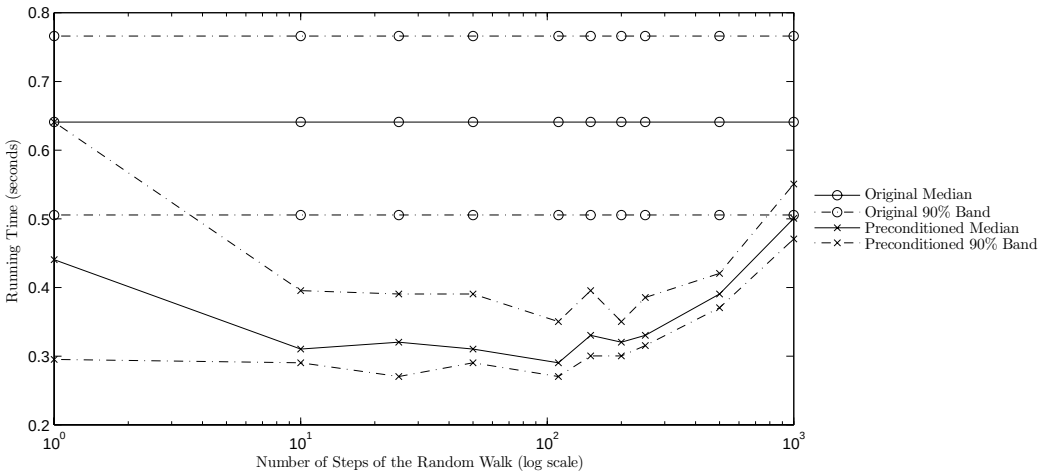

Fig. 6. Total running time versus number of steps of the geometric random walk for the 100 problem instances of dimension 100×500.

complexity behavior will be strongly-polynomial in m and the complexity value ϑ of the barrier function $f(\cdot)$ of the cone C. We have presented a method for generating such a point $\hat{v}$ using sampling on geometric random walks on $H_{\bar{s}}^{\circ}$ with associated complexity analysis. Finally, we have implemented this methodology on randomly generated homogeneous linear programming feasibility problems, constructed to be poorly behaved. Our computational results indicate that the projective pre-conditioning methodology holds the promise to markedly reduce the overall computation time for conic feasibility problems; for instance we observe a 46% improvement in average IPM iterations for the 100 problem instances of dimension 1000×5000. The next step in this line of research will be to develop a suitable adaptation of the methods developed herein to solve conic optimization problems, and to test such an adaptation on conic optimization problems that arise in practice.

7.1. The Infeasible Case

The theory presented herein is based on the assumption that F has a solution. When F does not have a solution, then one can consider the alternative/dual system:

$$F_a : \begin{cases} A^T v + s = 0 \\ \qquad s \in C^* \backslash \{0\} \ . \end{cases}$$

This system can then be re-formatted as:

$$F_a^{'} : \begin{cases} Bs = 0 \\ \quad s \in C^* \backslash \{0\} \ , \end{cases}$$

for a suitably computed matrix B whose null-space is the orthogonal complement of the null-space of A. Note that $F_a^{'}$ is of the same format as F and the results for F can be easily adapted to $F_a^{'}$.

(Actually, the computation of B is not necessary. Given $\bar{x} \in \text{int}C$, consider the analogous image sets for F_a and $F_a^{'}$ defined as $H_{\bar{x}} := \{A^T v + s : s \in C^*, \bar{x}^T s = 1, v \in \mathbb{R}^m\}$ and $H_{\bar{x}}^{'} := \{Bs : s \in C^*, \bar{x}^T s = 1\}$. Then $\text{sym}(0, H_{\bar{x}}) = \text{sym}(0, H_{\bar{x}}^{'})$ even though $H_{\bar{x}}$ is unbounded, and one can work with $H_{\bar{x}}$ and problem F_a directly.) Nevertheless, it may be more fruitful and revealing to develop a different projective pre-conditioner theory designed directly for the dual form F_a, and this is a direction for future research.

7.2. Related Complexity Matters

Nesterov [20] has suggested the following "dual approach" to solving (1): starting at $v^0 = 0$ compute an approximate analytic center v^a of $H_{\bar{s}}^{\circ}$, which is the essentially unconstrained problem $\min_v \{f^*(\bar{s} - A^T v) : \bar{s} - A^T v \in \text{int}C^*\}$. It is elementary to show from the barrier calculus that as soon as a point v is computed whose Newton step $(\Delta v, \Delta s)$ satisfies $\sqrt{(\Delta s)^T H^*(s) \Delta s} < 1$ (where $s = \bar{s} - A^T v$ and $H^*(s)$ is the Hessian of $f^*(\cdot)$ at s), then the Newton step multipliers yield an interior solution of F. Regarding the complexity of this scheme, it follows from an analysis that is almost identical to that yielding inequality (2.19) of [26] that the number of Newton steps of a short-step IPM to compute an approximate analytic center is:

$$O\left(\sqrt{\vartheta} \ln\left(\frac{\vartheta}{\text{sym}(0, H_{\bar{s}}^{\circ})}\right)\right) = O\left(\sqrt{\vartheta} \ln\left(\frac{\vartheta}{\text{sym}(0, H_{\bar{s}})}\right)\right)$$

(from (ii) of Remark 6), which is of the same order as the complexity of Algorithm A from Theorem 2. These complexity bounds depend on $\text{sym}(0, H_{\bar{s}})$ to bound the complexity of traversing the central path via a short-step method. As is shown in Nesterov and Nemirovski [22], a generically more accurate complexity bound can be found by analyzing the central path via its Riemannian geometry. However, as is demonstrated in the current paper, $\text{sym}(0, H_{\bar{s}})$ lends itself to analysis, characterization, and ultimately manipulation and reduction via the constructive projective pre-conditioning method shown herein. An interesting research challenge is to develop analogous tools/methods to work with and reduce the Riemannian distance of the central path as developed in [22].

References

1. A. Belloni and R. M. Freund. On the symmetry function of a convex set. *MIT Operations Research Center Working paper, submitted to Mathematical Programming*, OR-369-04, 2004.
2. M. E. Dyer and A. M. Frieze. On the complexity of computing the volume of a polyhedron. *SIAM Journal on Computing archive*, 17(5):967–974, 1988.
3. M. Epelman and R. M. Freund. A new condition measure, preconditioners, and relations between different measures of conditioning for conic linear systems. *SIAM Journal on Optimization*, 12(3):627–655, 2002.
4. G. S. Fishman. Choosing sample path length and number of sample paths when starting at steady state. *Operations Research Letters*, 16(4):209–220, 1994.
5. R. M. Freund. Combinatorial analogs of Brouwer's fixed-point theorem on a bounded polyhedron. *Journal of Combinatorial Theory, Series B*, 47(2):192–219, 1989.
6. R. M. Freund. Projective transformation for interior-point algorithms, and a superlinearly convergent algorithm for the w-center problem. *Mathematical Programming*, 58:203–222, 1991.

7. R. M. Freund. On the primal-dual geometry of level sets in linear and conic optimization. *SIAM Journal on Optimization*, 13(4):1004–1013, 2003.
8. R. M. Freund. On the behavior of the homogeneous self-dual model for conic convex optimization. *MIT Operations Research Center Working paper OR-372-04, submitted to Mathematical Programming*, 2004.
9. Robert M. Freund and Jorge R. Vera. Condition-based complexity of convex optimization in conic linear form via the ellipsoid algorithm. *SIAM Journal on Optimization*, 10(1):155–176, 1999.
10. C. J. Geyer. Practical markov chain monte carlo. *Statistical Science*, 7(4):473–511, 1992.
11. J. L. Goffin. The relaxation method for solving systems of linear inequalities. *Mathematics of Operations Research*, 5(3):388–414, 1980.
12. M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combiantorial Optimization*. Springer-Verlag, Berlin, second edition, 1994.
13. B. Grünbaum. *Convex Polytopes*. Wiley, New York, 1967.
14. P. C. Hammer. The centroid of a convex body. *Proc. Amer. Math. Soc.*, 5:522–525, 1951.
15. L. Leindler. On a certain converse of Hölder's inequality ii. *Acta Sci. Math. Szeged*, 33:217–223, 1972.
16. L. Lovász and S. Vempala. The geometry of logconcave functions and an $O^*(n^3)$ sampling algorithm. *Microsoft Technical Report*.
17. L. Lovász and S. Vempala. Hit-and-run is fast and fun. *Microsoft Technical Report*.
18. L. Lovász and S. Vempala. Where to start a geometric walk? *Microsoft Technical Report*.
19. H. Minkowski. Allegemeine lehzätze über konvexe polyeder. *Ges. Abh., Leipzog-Berlin*, 1:103–121, 1911.
20. Y. Nesterov. private communication. 2005.
21. Y. Nesterov and A Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1993.
22. Y. Nesterov and A Nemirovskii. Central path and Riemannian distances. Technical report, CORE Discussion Paper, CORE, Université Catholique de Louvain, Belgium, 2003.
23. Y. Nesterov, M.J. Todd, and Y. Ye. Infeasible-start primal-dual methods and infeasibility detectors. *Mathematical Programming*, 84:227–267, 1999.
24. A. Prékopa. Logarithmic concave measures and functions. *Acta Sci. Math. Szeged*, 34:335–343, 1973.
25. A. Prékopa. On logarithmic concave measures with applications to stochastic programming. *Acta Sci. Math. Szeged*, 32:301–316, 1973.
26. J. Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2001.
27. R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, New Jersey, 1970.
28. Reha H. Tütüncü, Kim Chuan Toh, and Michael J. Todd. SDPT3 — a MATLAB software package for semidefinite-quadratic-linear programming, version 3.0. Technical report, 2001. Available at http://www.math.nus.edu.sg/∼mattohkc/sdpt3.html.
29. Y. Zhang and L. Gao. On numerical solution of the maximum volume ellipsoid problem. *SIAM Journal on Optimization*, 14(1):53–76, 2003.

Appendix: A Primal-feasible Interior-Point Algorithm for $F_{\bar{s}}$ and its Complexity Analysis

Recall the parameterized barrier problem OP_η, and let $(x, \theta) = (x^k, \theta^k)$ be a feasible solution of OP_η for some $\eta > 0$. Then the locally quadratic model of OP_η at (x^k, θ^k) yields the following Newton equation system in variables (d, Δ, π, q):

$$\begin{array}{rrrl} & & \bar{x}^T A^T \pi & = \eta \\ H(x^k)d & & +A^T\pi + \bar{s} \cdot q & = -\nabla f(x^k) \\ Ad & +A\bar{x}\Delta & & = 0 \\ \bar{s}^T d & & & = 0\ . \end{array} \tag{11}$$

Here $\nabla f(x)$ and $H(x)$ denote the gradient and Hessian of $f(x)$, respectfully, (d, Δ) is the Newton step for the variables (x, θ), and (π, q) are the multipliers on the two linear equations of OP_η.

A primal-feasible interior-point algorithm for solving OP computes a sequence of approximate solutions (x^k, θ^k) of OP_{η^k} for an increasing sequence of values of η^k. Let $(\tilde{d}, \tilde{\Delta}, \tilde{\pi}, \tilde{q})$ solve the Newton system (11) using $\eta = \eta^k$, then the Newton step is $(\tilde{d}, \tilde{\Delta})$, and (x^k, θ^k) is defined to be a γ-approximate solution of OP^k_η if the norm of $\tilde{d}$ measured in the ellipsoidal norm induced by the Hessian of $f(\cdot)$ is not greater than γ, i.e., $\sqrt{(\tilde{d})^T H(x^k)\tilde{d}} \leq \gamma$. The algorithm starts by setting $(x^0, \theta^0) \leftarrow (\bar{x}, -1)$ and by computing a value η^0 for which it is guaranteed that (x^0, θ^0) is a γ-approximate solution of OP_{η^0}. Inductively, if (x^k, θ^k) is defined to be a γ-approximate solution of OP^k_η, the algorithm then increases η^k to $\eta^{k+1} = \alpha \cdot \eta^k$ for some fixed value of $\alpha > 1$, and then computes the Newton step for (x^k, θ^k) for $\text{OP}_{\eta^{k+1}}$. The algorithm continues in this manner until $\theta^k \geq 0$ at some iteration k, at which point the algorithm stops and $\hat{x} := \frac{x^k + \theta^k \bar{x}}{1+\theta^k}$ is a feasible solution of $F_{\bar{s}}$. A formal statement of the algorithm is as follows:

Algorithm A

Step 1. (Initialization) Set $\bar{x} \leftarrow -\nabla f^*(\bar{s})/\vartheta$. If $A\bar{x} = 0$, STOP. Otherwise, set $k = 0$, $(x^0, \theta^0) \leftarrow (\bar{x}, -1)$, and define the following constants:

$$\gamma = \frac{1}{9}\ , \quad \beta = \frac{1}{4}\ , \quad \alpha = \frac{\sqrt{\vartheta} + \beta}{\sqrt{\vartheta} + \gamma}\ .$$

Temporarily set $\eta = 1$ and solve (11) for $(\tilde{d}, \tilde{\Delta}, \tilde{\pi}, \tilde{q})$ and set

$$\eta^0 = \frac{\gamma}{\sqrt{\tilde{d}^T H(\bar{x})\tilde{d}}}\ .$$

Step 2. (Increase η) $\eta^{k+1} \leftarrow \alpha \cdot \eta^k$

Step 3. (Compute and Take Newton Step) Set $\eta = \eta^{k+1}$, solve (11) for $(\tilde{d}, \tilde{\Delta}, \tilde{\pi}, \tilde{q})$. Set $(x^{k+1}, \theta^{k+1}) = (x^k, \theta^k) + (\tilde{d}, \tilde{\Delta})$

Step 4. (Test Solution) If $\theta^{k+1} \geq 0$, set $\tilde{x} := \frac{x^k + \theta^k \bar{x}}{1+\theta^k}$ and STOP. Otherwise set $k \leftarrow k+1$ and go to **Step 2**.

In order to validate this algorithm, we will need to prove the following results. Let d^k denote the value of $\tilde{d}$ in (11) at (x^k, θ^k) using $\eta = \eta^k$. The norm of this Newton step in the norm induced by the Hessian $H(x^k)$ and is given by:

$$\|(d^k)\|_{x^k} := \sqrt{(d^k)^T H(x^k) d^k}\ .$$

Proposition 2. $\|(d^0)\|_{x^0} = \gamma$.

Proof. Following Step 1 of Algorithm A, let $(\tilde{d}, \tilde{\Delta}, \tilde{\pi}, \tilde{q})$ solve (11) using $x^k = x^0 = \bar{x}$ and $\eta = 1$, and set η^0 by the prescribed formula. Using the fact that $\bar{s} = -\nabla f(\bar{x})/\vartheta$ (from (vi) of Remark 4 and Theorem 2.3.9 of [26]), it follows from direct substitution that $(d^0, \Delta^0, \pi^0, q^0) := (\eta^0\tilde{d}, \eta^0\tilde{\Delta}, \eta^0\tilde{\pi}, \eta^0\tilde{q} + \vartheta - \eta^0\vartheta)$ solves (11) using $x^k = x^0 = \bar{x}$ and $\eta = \eta^0$. Therefore $\|(d^0)\|_{x^0} = \sqrt{(d^0)^T H(x^0)(d^0)} = \eta^0\sqrt{\tilde{d}^T H(\bar{x})\tilde{d}} = \gamma$. □

Proposition 3. $\|(d^k)\|_{x^k} \leq \gamma$ *for all* $k = 1, 2, \ldots$.

Proof. The proof is by induction, and for clarity we suppress the iteration counter k. Suppose that our current point is (x, θ). Let $d_{\bar{\eta}}$ denote the x-part of the Newton step for the parameter value $\eta = \bar{\eta}$. Then $d_{\bar{\eta}}$ can be decomposed as $d_{\bar{\eta}} = d^c + \bar{\eta} d^a$ where d^c is the centering direction and d^a is the affine scaling direction. It follows from the fact that $f(\cdot)$ has complexity value ϑ that $\|d^c\|_x \leq \sqrt{\vartheta}$. Furthermore, by assumption of the induction we have $\|d_{\bar{\eta}}\|_x \leq \gamma$. Then according to Step 2 of Algorithm A we increase η by the factor α. We write $d_{\alpha\bar{\eta}} = d^c + \alpha\bar{\eta} d^a = \alpha(d^c + \bar{\eta} d^a) + (1-\alpha) d^c = \alpha d_{\bar{\eta}} + (1-\alpha) d^c$, whereby:

$$\|d_{\alpha\bar{\eta}}\|_x \leq \alpha \|d_{\bar{\eta}}\|_x + (\alpha - 1)\|d^c\|_x \leq \alpha\gamma + (\alpha - 1)\sqrt{\vartheta} = \beta \ .$$

Letting $x^+ := x + d_{\alpha\bar{\eta}}$ denote the new value of x and letting $d^+_{\alpha\bar{\eta}}$ denote the Newton step at x^+ for the parameter value $\eta = \alpha\bar{\eta}$ it follows from Theorem 2.2.4 of [26] that

$$\|d^+_{\alpha\bar{\eta}}\|_{x^+} \leq \frac{\beta^2}{(1-\beta)^2} = \gamma \ ,$$

completing the inductive proof. □

Proposition 4. $\eta^0 \geq \dfrac{\gamma}{1 + \theta^*}$

Proof. Since $(x, \theta) = (\bar{x}, -1)$ is feasible for (4) and from the self-concordance of $f(\cdot)$ we have $\{\bar{x}\} + \{d : d^T H(\bar{x}) d \leq 1\} \subset C$, it follows that

$$\theta^* \geq \quad -1 + \quad \max_{d, \Delta} \quad \Delta$$

$$\begin{aligned} Ad \quad &+ (A\bar{x})\Delta = 0 \\ \bar{s}^T d \quad &\qquad\quad = 0 \\ d^T H(\bar{x}) d \quad &\qquad\quad \leq 1 \ . \end{aligned}$$

Letting (d, Δ) solve the above maximization problem, to prove the proposition it therefore suffices to show that $\Delta = \gamma/\eta^0$, which we now do. The optimality conditions for the maximization problem above can be written as:

$$\begin{aligned} \bar{x}^T A^T \pi &= 1 \\ \rho H(\bar{x}) d + A^T \pi + \bar{s} \cdot q &= 0 \\ Ad + A\bar{x}\Delta &= 0 \\ \bar{s}^T d &= 0 \\ \rho &\geq 0 \\ d^T H(\bar{x}) d &\leq 1 \\ \rho \cdot (1 - d^T H(\bar{x}) d) &= 0 \ . \end{aligned} \tag{12}$$

Following Step 1 of Algorithm A, let $(\tilde{d}, \tilde{\Delta}, \tilde{\pi}, \tilde{q})$ solve (11) using $x^k = x^0 = \bar{x}$ and $\eta = 1$, and set η^0 by the prescribed formula. One then easily checks that

$$(d, \Delta, \pi, q, \rho) := \left(\frac{\tilde{d}}{\sqrt{\tilde{d}^T H(\bar{x}) \tilde{d}}}, \frac{\tilde{\Delta}}{\sqrt{\tilde{d}^T H(\bar{x}) \tilde{d}}}, \tilde{\pi}, \tilde{q} - \vartheta, \sqrt{\tilde{d}^T H(\bar{x}) \tilde{d}} \right)$$

satisfies (12) (again using the fact that $\bar{s} = -\nabla f(\bar{x})/\vartheta$ from (vi) of Remark 4 and Theorem 2.3.9 of [26]), and so (d, Δ) is an optimal solution of the maximization problem. Straightforward manipulation of the system (12) then shows that $\Delta = \sqrt{\tilde{d}^T H(\bar{x})\tilde{d}}$, and so from the definition of η^0 we have $\Delta = \gamma/\eta^0$. □

Before proving the next proposition, we state a result which is implicit in Renegar [26], but is not stated explicitly. Rather than re-develop the notation and set-up of [26], we simply state the result and give a proof as if it appeared as part of the text of Chapter 2 of [26].

Lemma 7. (Essentially from Renegar [26]) *Under the notation and conditions of Chapter 2 of Renegar [26], suppose that y is an iterate of the barrier method for the barrier parameter η and the Newton step $n(y)$ for the function $f_\eta(x) := \eta\langle c, x\rangle + f(x)$ at y satisfies $\|n(y)\|_y \leq \gamma$ where $\gamma < 1/4$. Then*

$$c^T y \leq \text{VAL} + \frac{\vartheta}{\eta}\left(\frac{1}{1-\delta}\right)$$

where $\delta = \gamma + \frac{3\gamma^2}{(1-\gamma)^3}$.

Proof. Letting $z(\eta)$ denote the analytic center of the function $f_\eta(\cdot)$, it follows from Theorem 2.2.5 of [26] that $\|y - z(\eta)\|_y \leq \delta$, and hence from the self-concordance property that $\|y - z(\eta)\|_{z(\eta)} \leq \delta/(1-\delta)$. From inequality (2.14) of Section 2.4.1 of [26] we have:

$$c^T y \leq \text{VAL} + \frac{\vartheta}{\eta}\left(1 + \|y - z(\eta)\|_{z(\eta)}\right) \leq \text{VAL} + \frac{\vartheta}{\eta}\left(1 + \frac{\delta}{1-\delta}\right) = \text{VAL} + \frac{\vartheta}{\eta}\left(\frac{1}{1-\delta}\right) .$$

□

Proposition 5. *Suppose (x, θ) is an iterate of Algorithm A for the parameter value η and $\eta \geq \vartheta/(\theta^*(1-\delta))$ where $\delta = \gamma + \frac{3\gamma^2}{(1-\gamma)^3}$. Then $\theta \geq 0$.*

Proof. Converting Lemma 7 to the notation of problem (5) and Algorithm A, we have:

$$\theta \geq \theta^* - \frac{\vartheta}{\eta}\left(\frac{1}{1-\delta}\right) \geq \theta^* - \theta^* = 0 ,$$

where the last inequality follows from the supposition that $\eta \geq \vartheta/(\theta^*(1-\delta))$. □

Proof of Theorem 2: We first prove the iteration bound. Note that the parameter values set in Step 1 of Algorithm A imply:

$$1 - \frac{1}{\alpha} = \frac{1}{7.2\sqrt{\vartheta} + 1.8} \geq \frac{1}{9\sqrt{\vartheta}} . \tag{13}$$

Define δ as in the hypothesis of Proposition 5. Let

$$J := \left\lceil 9\sqrt{\vartheta}\ln\left(11\vartheta\left(1 + \frac{1}{\text{sym}(0, H_{\bar{s}})}\right)\right)\right\rceil .$$

From Step 2 of Algorithm A we have $\eta^0/\eta^J = (1/\alpha)^J$ and taking logarithms we obtain $\ln(\eta^0) - \ln(\eta^J) = J\ln(1/\alpha) \leq J(1/\alpha - 1)$, and rearranging we obtain:

$$\begin{aligned}
\ln(\eta^J) &\geq \ln(\eta^0) + J\left(1 - \tfrac{1}{\alpha}\right) \\
&\geq \ln\left(\tfrac{\gamma}{1+\theta^*}\right) + \ln\left(11\vartheta\left(1 + \tfrac{1}{\mathrm{sym}(0,H_{\bar{s}})}\right)\right) && \text{(from Proposition 4 and (13))} \\
&\geq \ln\left(\tfrac{\gamma}{1+\theta^*}\right) + \ln\left(11\vartheta\left(\tfrac{\theta^*+1}{\theta^*}\right)\right) && \text{(from (6))} \\
&= \ln\left(\tfrac{\vartheta}{\theta^*(1-\delta)}\right) + \ln\left(11\gamma(1-\delta)\right) \ \geq \ln\left(\tfrac{\vartheta}{\theta^*(1-\delta)}\right) \ ,
\end{aligned}$$

where the last inequality follows since $11\gamma(1-\delta) \geq 1$ for the specific values of γ and δ given. Then from Proposition 5 it follows that $\theta^J \geq 0$, and so Algorithm A will stop.

It remains to prove the bound on $\mathrm{reldist}(\tilde{x}, \partial C)$. For $x \in \mathrm{int}C$, define the norm $\|\cdot\|_x$ by $\|v\|_x := \sqrt{v^T H(x) v}$ for $x \in \mathrm{int}C$, and for $x \in C$ define:

$$B_x(c,r) := \{y : A(y-c) = 0, \bar{s}^T(y-c) = 0, \|y-c\|_x \leq r\} \ .$$

Letting z denote the analytic center of $\mathcal{F}_{\bar{s}}$ it follows from Lemma 5 of [23] that $B_z(z,\vartheta) \supset \mathcal{F}_{\bar{s}}$. Assuming for simplicity that $\theta = 0$ for the final iterate of Algorithm A, we have $\|z - \tilde{x}\|_{\tilde{x}} \leq \delta := \gamma + 3\gamma^2/(1-\gamma)^3$ from Theorem 2.2.5 of [26], whereby $B_{\tilde{x}}(\tilde{x}, \vartheta/(1-\delta) + \delta/(1-\delta)) \supset \mathcal{F}_{\bar{s}}$ follows from the self-concordance of $f(\cdot)$. Since it is also true that $B_{\tilde{x}}(\tilde{x},1) \subset \mathcal{F}_{\bar{s}}$ it follows that $\sigma := \mathrm{sym}(\tilde{x}, \mathcal{F}_{\bar{s}}) \geq (1-\delta)/(\vartheta+\delta) \geq 1/(1.2\vartheta + 0.2)$ for the specific values of γ and δ herein. Finally, noting that there exists $\hat{x} \in \mathcal{F}_{\bar{s}}$ satisfying $B_{\bar{s}}(\hat{x}, \tau_F) \subset C$, where $B_{\bar{s}}(c,r)$ is the ball centered at c of radius r in the $\bar{s}$-norm, it follows from the symmetry value of $\tilde{x}$ in $\mathcal{F}_{\bar{s}}$ that $B_{\bar{s}}(\tilde{x} - \sigma(\hat{x}-\tilde{x}), \sigma\tau_F) \subset C$, whereby taking convex combinations with $B_{\bar{s}}(\hat{x}, \tau_F)$ yields

$$B_{\bar{s}}\left(\tilde{x}, \frac{\tau_F}{1.2\vartheta + 0.2}\right) \subset B_{\bar{s}}\left(\tilde{x}, \sigma\tau_F\right) \subset B_{\bar{s}}\left(\tilde{x}, \frac{2\sigma\tau_F}{1+\sigma}\right) = \frac{1}{1+\sigma} B_{\bar{s}}(\tilde{x} - \sigma(\hat{x}-\tilde{x}), \sigma\tau_F) + \frac{\sigma}{1+\sigma} B_{\bar{s}}(\hat{x}, \tau_F) \subset C \ ,$$

from which it follows that $\mathrm{reldist}(\tilde{x}, \partial C) \geq \tau_F/(1.2\vartheta + 0.2)$. □