Alexandre Belloni · Robert M. Freund

PROJECTIVE RE-NORMALIZATION FOR IMPROVING THE BEHAVIOR OF A HOMOGENEOUS CONIC LINEAR SYSTEM⋆

REVISED AUGUST 3, 2007

Abstract. In this paper we study the homogeneous conic system $F : Ax = 0, \ x \in C \setminus \{0\}$. We choose a point $\bar{s} \in \text{int} C^*$ that serves as a *normalizer* and consider computational properties of the normalized system $F_{\bar{s}} : Ax = 0, \ \bar{s}^T x = 1, \ x \in C$. We show that the computational complexity of solving F via an interior-point method depends only on the complexity value ϑ of the barrier for C and on the symmetry of the origin in the *image set* $H_{\bar{s}} := \{Ax : \bar{s}^T x = 1, \ x \in C\}$, where the symmetry of 0 in $H_{\bar{s}}$ is

$$\text{sym}(0, H_{\bar{s}}) := \max\{\alpha : y \in H_{\bar{s}} \Rightarrow -\alpha y \in H_{\bar{s}}\} \ .$$

We show that a solution of F can be computed in $O(\sqrt{\vartheta} \ln(\vartheta/\text{sym}(0, H_{\bar{s}})))$ interior-point iterations. In order to improve the theoretical and practical computation of a solution of F, we next present a general theory for projective re-normalization of the feasible region $F_{\bar{s}}$ and the image set $H_{\bar{s}}$ and prove the existence of a normalizer $\bar{s}$ such that $\text{sym}(0, H_{\bar{s}}) \geq 1/m$ provided that F has an interior solution. We develop a methodology for constructing a normalizer $\bar{s}$ such that $\text{sym}(0, H_{\bar{s}}) \geq 1/m$ with high probability, based on sampling on a geometric random walk with associated probabilistic complexity analysis. While such a normalizer is not itself computable in strongly-polynomial-time, the normalizer will yield a conic system that is solvable in $O(\sqrt{\vartheta} \ln(m\vartheta))$ iterations, which is strongly-polynomial-time. 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 re-normalization 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 computational characteristics of the following homogeneous convex feasibility problem in conic linear form:

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

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

Alexandre Belloni: IBM T. J. Watson Research Center and MIT, 32-221, 1101 Kitchawan Road, Yorktown Heights, New York 10598. 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`

It is well known that the standard form conic feasibility problem

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

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.

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. Choose a point $\bar{s} \in \text{int}C^*$ (the interior of the dual cone of 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 the computational complexity of solving F can be bounded as a function of only two quantities: (i) the symmetry of the *image set* $H_{\bar{s}} := \{Ax : \bar{s}^T x = 1, \ x \in C\}$ 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 re-normalizing the conic system $F_{\bar{s}}$ to an equivalent system $F_{\hat{s}}$ via projective transformation of the image set (that also projectively transforms the feasible regions of the normalized problem). The re-normalization transforms the conic system into one with computational properties with certain guarantees; we use the term "projective re-normalization" to describe such a projective transformation. Under the assumption that F has an interior solution, there must exist projective re-normalizations that transform $F_{\bar{s}}$ into equivalent systems that are solvable in strongly-polynomial time in m and ϑ. The quality of the projective re-normalization depends on the ability to compute a point $\hat{v}$ that is "deep" in the set $H_{\bar{s}}^{\circ} = \{v : \bar{s} - A^T v \in C^*\}$. 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 probabilistic complexity analysis. In Section 6 we present results from computational experiments designed to assess the practical viability of the projective re-normalization 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 methodology for modifying F was proposed in [5] 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).

1.1. Notation

Let $e = (1, \dots, 1)^T \in \mathbb{R}^d$ 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 [28]. 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 [28]. For a general norm $\|\cdot\|$, let $B(c, r)$ and $\text{dist}(x, T)$ denote the ball of radius r centered at c and the distance from a point x to a set T, respectively.

2. Normalization, Complexity, and Barrier Calculus

2.1. Assumptions

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

Assumption 1. *C is a regular closed convex cone, i.e.,* $\text{int} C \neq \emptyset$ *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. $\text{rank}(A) = m.$

It follows immediately from Assumption 1 that C^* is also a regular cone, see Rockafellar [28]. Assumptions 2 and 3 together assure that F has a solution and A contains no redundant equations.

2.2. Normalization, Image Set, and Symmetry

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

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

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

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 the complexity of computing a solution of F can be bounded quite simply using 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 Assumptions 1, 2 and 3 imply that $0 \in H$ and that H has dimension m.

We consider the *symmetry* of a point in a convex set as defined originally by Minkowski [20], see also the references and results in [2]. 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 [22].

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 [22] or [27], whereas (vii) is borrowed from the proof of Lemma 5 of [24]. We presume that function values, gradients, and Hessians of $f(\cdot)$ and $f^*(\cdot)$ are computable.

3. Complexity Bound on F

Let $\bar{s} \in \text{int}C^*$ be chosen, and let $F_{\bar{s}}$ be as defined in Section 2.2 with image set $H = H_{\bar{s}}$ as in Section 2.2. In this section we show that the computational complexity of a standard interior-point method (IPM) for computing a solution of F depends only on $\text{sym}(0, H)$ and the complexity value ϑ of the barrier function $f(\cdot)$ for C. In order to establish this result we first present the model that will be solved by the IPM.

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

$$
\begin{aligned}
\text{OP}: \quad t^* := \max_{x,t} \quad & t \\
\text{s.t.} \quad Ax + (A\bar{x})t &= 0 \\
\bar{s}^T x \quad\quad\quad &= 1 \\
x \quad\quad\quad &\in C \, .
\end{aligned}
\tag{2}
$$

It follows from Remark 4 that $\bar{x} \in \text{int}C$ and $\bar{s}^T\bar{x} = 1$, and that $(x,t) = (\bar{x},-1)$ 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 use $(\bar{x},-1)$ as a starting point with which to initiate a standard short-step primal-feasible interior-point method (IPM) for improving the objective function value of (2) until we have computed a point (x,t) that satisfies $t \geq 0$, whereby $\tilde{x} := (x + t\bar{x})$ is a feasible solution of F.

Theorem 1. *Let $\bar{s} \in \text{int}C^*$ be chosen. The standard short-step primal-feasible interior-point algorithm applied to (2) will compute $\tilde{x}$ satisfying $A\tilde{x} = 0, \tilde{x} \in \text{int}C$ in at most*

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

iterations of Newton's method.

The proof of this theorem uses the following proposition.

Proposition 1. *Let t^* be defined in (2). Then $t^* \geq \text{sym}(0,H_{\bar{s}})$.*

Proof. 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 $t^* \geq \text{sym}(0,H_{\bar{s}})$. $\square$

Proof of Theorem 1: Let OP_η denote the optimization problem (2) with the objective function replaced by $\eta f(x)+t$ where $f(\cdot)$ is the barrier function of Remark 4. Because the point $(\bar{x},-1)$, whose objective value in (2) is -1, is the analytic center of the feasible region of OP, it follows from the arguments on pages 47-48 of [27] that the barrier method can be initiated at $(\bar{x},-1)$ using the initial barrier parameter $\eta = \eta^0 \geq c_1/(1+t^*)$ for some absolute constant c_1. Furthermore, since the optimal value of OP is t^*, and the IPM will stop when the current feasible point (x,t) satisfies $t \geq 0$, the final optimality gap is t^* and will be attained when the barrier parameter η reaches $\eta = \eta^f = c_2\vartheta/t^*$ for some absolute constant c_2. Therefore, the ratio of the final to the initial barrier parameter is $O(\vartheta(1+t^*)/t^*) \leq O(\vartheta/\text{sym}(0,H_{\bar{s}}))$ from Proposition 1, and it follows from the linear convergence theory of interior-point methods that the iteration bound will be $O(\sqrt{\vartheta}\ln(\vartheta/\text{sym}(0,H_{\bar{s}})))$. $\square$

Note that the complexity bound is trivially valid even when $\text{sym}(0,H_{\bar{s}}) = 0$, using the standard convention that $1/0 = \infty$. An alternate proof of Theorem 1 that does not use big-O notation can be found in the appendix of [1].
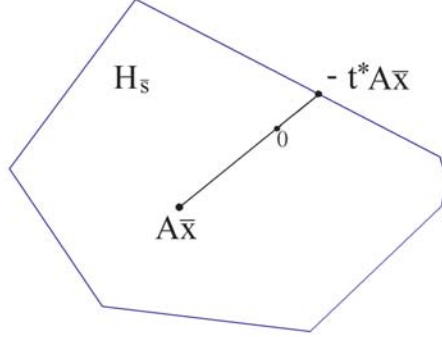
Fig. 1. The image set $H_{\bar{s}}$ and the points $A\bar{x}, 0$, and $-t^* A\bar{x}$.

Figure 1 can be used to gain some intuition into both Proposition 1 and Theorem 1. The figure portrays the image set $H_{\bar{s}}$, which contains 0 as a consequence of Assumption 2. Furthermore, $A\bar{x} \in H_{\bar{s}}$ by design of $\bar{x}$. The optimal value t^* of (2) is the largest value of t for which $-tA\bar{x} \in H_{\bar{s}}$. Figure 1 and Proposition 1 show that $t^* \geq \text{sym}(0, H_{\bar{s}})$, and so t^* will be large if $\text{sym}(0, H_{\bar{s}})$ is large. Since the interior-point algorithm starts at the analytic center $(\bar{x}, -1)$ with objective value $t = -1$ and will stop when the current iterate (x, t) satisfies $t \geq 0$, the ratio of the initial to the final optimality gap is $(1 + t^*)/t^* \leq 1 + 1/(\text{sym}(0, H_{\bar{s}}))$.

In Section 4, we will show how projective transformations can be used to modify $\text{sym}(0, H)$ and hence improve the computational complexity of the system $F_{\bar{s}}$.

4. Re-Normalizing F by Projective Transformation

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 $\hat{s} \in \text{int}C^*$, 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 $\text{sym}(\cdot)$:

Remark 5. Let $S \subset \mathbb{R}^m$ be a nonempty closed bounded convex set. The following properties of $\text{sym}(\cdot, S)$ are shown in [2]:
(i) Let $\mathcal{A}(x) := Mx + g$, M nonsingular. If $\bar{x} \in S$, then $\text{sym}(\mathcal{A}(\bar{x}), \mathcal{A}(S)) = \text{sym}(\bar{x}, S)$.
(ii) If $0 \in S$, then $\text{sym}(0, S) = \text{sym}(0, S^\circ)$.
(iii) $\text{sym}(S) \geq \frac{1}{m}$.

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 $H_{\bar{s}}^\circ$, as the following proposition shows.

Proposition 2. *Let $\bar{s} \in \text{int}C^*$ 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 2 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, $0 \in \text{int} H_{\bar{s}}^{\circ}$, so we have a known point in the interior of $\text{int} H_{\bar{s}}^{\circ}$. Second, checking membership in $H_{\bar{s}}^{\circ}$ is a relatively simple task if we have available a membership oracle for C^*.

Motivated by Theorem 1 which bounds the 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 $\hat{s} \in \text{int} C^*$ with the goal of improving $\text{sym}(0, H_{\hat{s}})$. We proceed as follows. Taking $\bar{s} \in \text{int} C^*$ 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:

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

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 2. *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 5 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 5. □

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

$$x' \leftarrow \frac{x}{\hat{s}^T x} \quad \text{and} \quad x \leftarrow \frac{x'}{\bar{s}^T x'} \ . \tag{3}$$

Furthermore, Theorem 2 has an interesting interpretation in the context of projective transformations and polarity theory which we will discuss shortly.

Our present goal, however, is to use Theorem 2 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 2 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 Theorem 1 that the transformed system $F_{\hat{s}}$ will have complexity that will depend on α as well. We formalize this method and the above conclusion as follows:

Projective Re-Normalization Method (PRM)

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:

$$F_{\hat{s}} : \begin{cases} Ax = 0 \\ \hat{s}^T x = 1 \\ x \in C \end{cases} \tag{4}$$

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

Figure 2 illustrates the strategy of the Projective Re-Normalization 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., $\text{sym}(0, H_{\bar{s}}) << 1$. 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., $\text{sym}(0, H_{\hat{s}}) = 1$. (In general, of course, we can at best attain $\text{sym}(0, H_{\hat{s}}) = 1/m$.)

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

Corollary 1. *Let $\bar{s} \in \text{int}C^*$ be chosen, and suppose that the Projective Re-Normalization Method has been run, and let $\alpha := \text{sym}(\hat{v}, H_{\bar{s}}^{\circ})$. The standard primal-feasible interior-point algorithm applied to the re-normalized problem (4) will compute $\tilde{x}$ satisfying $A\tilde{x} = 0, \tilde{x} \in \text{int}C$ in at most*

$$O \left(\sqrt{\vartheta} \ln \left(\frac{\vartheta}{\alpha} \right) \right)$$

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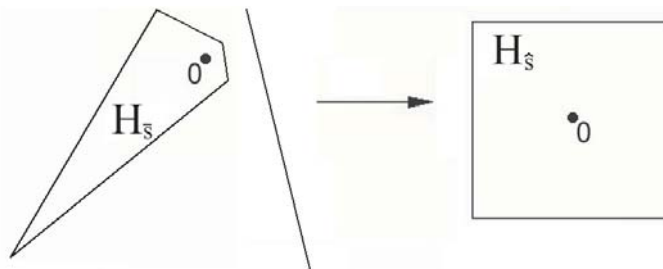
Fig. 2. Projective transformation of the image set $H_{\bar{s}}$ to improve the symmetry of 0 in the transformed image set.

While it is obvious that $F_{\bar{s}}$ and $F_{\hat{s}}$ are related through the pair of projective transformations (3), 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 $y' \in H_{\hat{s}}$:

$$y' = T(y) := \frac{y}{1 - \hat{v}^T y} \quad \text{and} \quad y = T^{-1}(y') = \frac{y'}{1 + \hat{v}^T y'} \; . \tag{5}$$

This pair of projective transformations results from a more general theory concerning the polarity construction, translations, and projective transformations as follows, see Grünbaum [11] 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 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 (5), 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 [7,8].

If F has an interior solution, it follows from Assumptions 1, 2, and 3 that $0 \in \mathrm{int}H_{\bar{s}}$, whereby $H_{\bar{s}}^\circ$ will be bounded and $\mathrm{sym}(0, H_{\bar{s}}) = \mathrm{sym}(0, H_{\bar{s}}^\circ) > 0$. Furthermore, we know from (iii) of Remark 5 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 $\alpha := \mathrm{sym}(\hat{v}, H_{\bar{s}}^\circ) = \Omega(1/m)$, 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 $\alpha := \mathrm{sym}(\hat{v}, H_{\bar{s}}^\circ) > \mathrm{sym}(0, H_{\bar{s}})$, 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 2; furthermore, we have a convenient point $0 \in \mathrm{int}H_{\bar{s}}^\circ$ 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 convenient description of $H_{\bar{s}}^\circ$ 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.
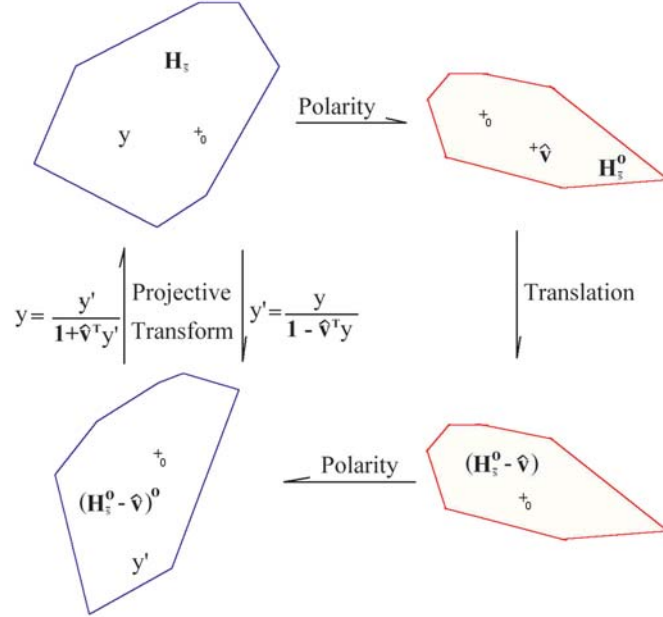
Fig. 3. Translation of the polar set corresponds to projective transformation of the original set.

Remark 6. One implication of Corollary 1 is that the efficiency of computing only a "barely-deep" (in the symmetry measure) point in a convex body starting from a given (interior) point implies that F can be solved with similar efficiency. To see why this is true, recall that $H_{\bar{s}}^\circ := \{v : \bar{s} - A^T v \in C^*\}$ has a convenient representation, and in fact $0 \in \operatorname{int} H_{\bar{s}}^\circ$. From Corollary 1, if we could use $v = 0$ as a starting point to find a point $\hat{v} \in H_{\bar{s}}^\circ$ satisfying $\operatorname{sym}(\hat{v}, H_{\bar{s}}^\circ) \geq 2^{-m}$ in strongly-polynomial-time, we could then compute a solution of F in strongly-polynomial-time. More formally, we have:

> Suppose that an interior solution of F exists, and that there is an absolute constant $c > 0$ and a polynomial $T(n, m, \vartheta)$ such that a point $\hat{v} \in H_{\bar{s}}^\circ := \{v : \bar{s} - A^T v \in C^*\}$ satisfying $\operatorname{sym}(\hat{v}, H_{\bar{s}}^\circ) \geq e^{-m^c}$ can be computed in $T(n, m, \vartheta)$ operations. Then F can be solved in $O(T(n, m, \vartheta) + \sqrt{\vartheta} \ln(\vartheta) n^3 m^c)$ operations.

While this result follows immediately from Corollary 1, it also is directly implied by the discussion on pages 47-48 of Renegar [27]. In the context of linear inequalities, this suggests that unless linear programming is strongly-polynomial-time, the problem of computing a barely-deep point in $H_e^\circ := \{v : A^T v \leq e\}$ is not solvable in strongly-polynomial-time. This result is comparable with that of Khachiyan for the rounding problem on H_e° [15].

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

If F has an interior solution, it follows from Assumptions 1, 2, and 3 that $0 \in \text{int} H_{\bar{s}}$, whereby $H_{\bar{s}}^{\circ}$ 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 5 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 [2]. 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 [2]. 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 $C = \mathbb{R}_+^n$ (wherein $H_{\bar{s}}^{\circ}$ 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 [22] 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 [24]). Each iteration of the algorithm would be of comparable computational burden as an interior-point iteration for the problem (2) 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 [10]. Such a point v^j is called a Löwner-John center and it follows that $\text{sym}(v^j, H_{\bar{s}}^{\circ}) \geq 1/m$. 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 [30], 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 Hammer [12]. Even when $C = \mathbb{R}_+^n$, computing the center of mass of $H_{\bar{s}}^{\circ}$ is #P-hard, see Dyer and Frieze [4]. However, if we instead consider nondeterministic algorithms, then the recent work of Lovász and Vempala [17–19] 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	$1/m$	$\approx \text{LP} \times (n+1)$
Analytic Center	$\frac{1}{\sqrt{\vartheta(\vartheta-1)}}$	$\approx \text{LP}$
Löwner-John Center	$1/m$	$\approx \text{SDP}$
Center of Mass	$1/m$	#P-Hard (deterministic)
		$\approx$ polynomial-time (stochastic)

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

5. Approximate Center of Mass of $H^\circ_{\bar{s}}$ 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 $S = H^\circ_{\bar{s}}$ 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 1. [16, 25, 26] *The sum of independent logconcave random variables is a logconcave random variable.*

Lemma 2. [17] *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

$$f(x) = \frac{1}{\mathsf{Vol}(S)} 1_S(x) \ , \tag{6}$$

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 3. [17] *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 4. *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 3 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

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

□

Corollary 2. *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 1, Y is a logconcave random variable since it is a sum of independent logconcave random variables. Applying Lemma 2 using $R = t\sqrt{\frac{M}{d}}$ we obtain:

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

□

Theorem 3. *For any $\delta \in (0,1)$ and setting $M = \lceil 4d\left(\ln(1/\delta)\right)^2 \rceil$, let $\hat{v}$ be the sample mean of M independent uniformly distributed points on the convex body $S \subset \mathbb{R}^d$. Then*

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

with probability at least $1 - \delta$.

Proof. Using Corollary 2 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 4 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 Theorem 3 can be extended to show that setting $M = \left\lceil \left(\frac{1 + \frac{1-\varepsilon}{d}}{\varepsilon} \right)^2 d \left(\ln(1/\delta) \right)^2 \right\rceil$ 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), we see that Theorem 3 provides 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 $d = m$, under Assumptions 1, 2, and 3 and the additional assumption that F has an interior solution x^o (i.e., $Ax^o = 0$, $x^o \in \mathrm{int}C$), it follows that $0 \in \mathrm{int}H_{\bar{s}}$ and hence $H_{\bar{s}}^\circ$ is a convex body. In this case Theorem 3 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 2). It follows from Corollary 1 and Theorem 3 that in the context of the Projective Re-Normalization 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 $O(\sqrt{\vartheta} \ln(m \cdot \vartheta))$ iterations of Newton's method.

In order to make Theorem 3 constructive, we need a method for sampling on $S = H_{\bar{s}}^\circ$ that (approximately) obeys a uniform distribution. There are many recent results on this problem for general convex bodies S, see, e.g., [17], [18], and [19]. Herein we focus on the "Hit-and-Run" geometric random walk (see [18]) that 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**

In [18] it is proved that to achieve an ε-approximation to the uniform distribution density function (6) 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 requires the largest and smallest eigenvalues of a matrix.

6. Computational Results on Randomly Generated Poorly-Behaved Problems

We performed computational experiments to assess the practical viability of the projective re-normalization method (PRM). We tested the PRN 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 $H^{\circ}_{s^0} = \{v : A^T v \le e\}$. 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 PRN, and yields $\mathrm{sym}(0, H_{\bar{s}}) \le 4 \times 10^{-5}$.

We implemented the projective re-normalization method (PRN) using the following simplified version of the stochastic process described in Section 5: starting from $v^0 = 0 \in \mathrm{int} H^{\circ}_{\bar{s}}$ 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 [6] and [9] for discussion and further references.

We solved the 300 original problem instances of OP (stopping as soon as $t \ge 0$), as well as the resulting instances after re-normalization, using the interior-point software SDPT3 [29]. 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 re-normalization using the geometric random walk) also shows a marked decrease when using the projective re-normalization 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 t^*. Given that $(x, t) = (\bar{x}, -1)$ is a feasible starting point for OP (and for SDPT3), t^* is a good measure of the computational difficulty of a problem instance – a problem is poorly behaved to the extent that t^* is close to zero. Here we see, regardless of any IPM, that t^* 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 t^*	
		Original	After	Original	After	Original	After
m	n	Problem	Re-Normalization	Problem	Re-Normalization	Problem	Re-Normalization
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 PRN 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 re-normalization. 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.

Figure 5 shows the median value of t^* as well as the 90% band of t^* values for the 100 problems of dimension 100×500 before and after re-normalization. As discussed earlier, t^* is a good measure of problem instance behavior: larger values of t^* indicate that the problem instance is better behaved, especially for computation via an IPM. The figure indicates that there is almost no improvement in the median value of t^* after 50 steps of the random walk. Note that Figure 5 is almost a mirror image of Figure 4, verifying intuition that IPM iterations should be inversely proportional to $\log(t^*)$.

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 re-normalization. Notice that the median running time of the system with re-normalization 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 re-normalization. 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.
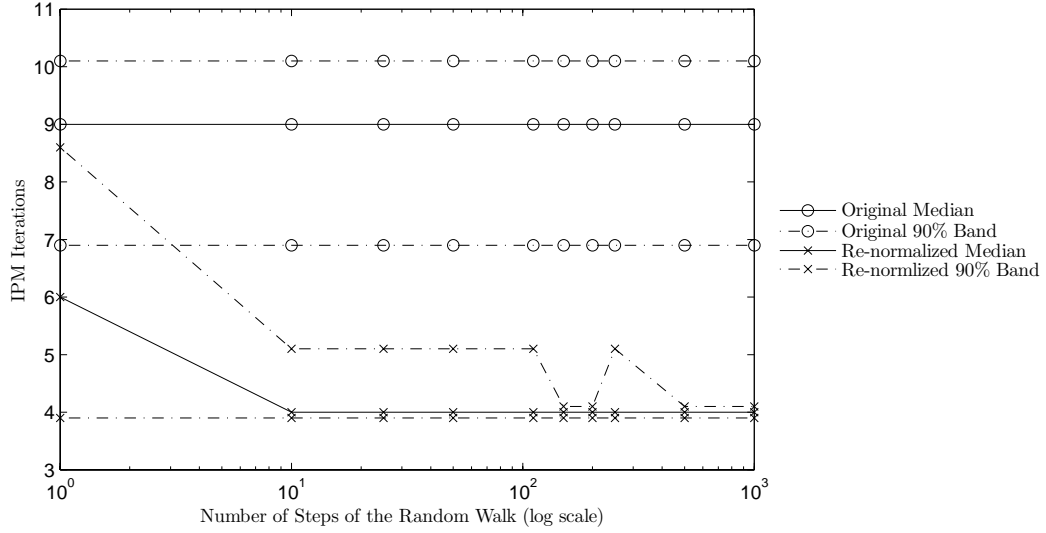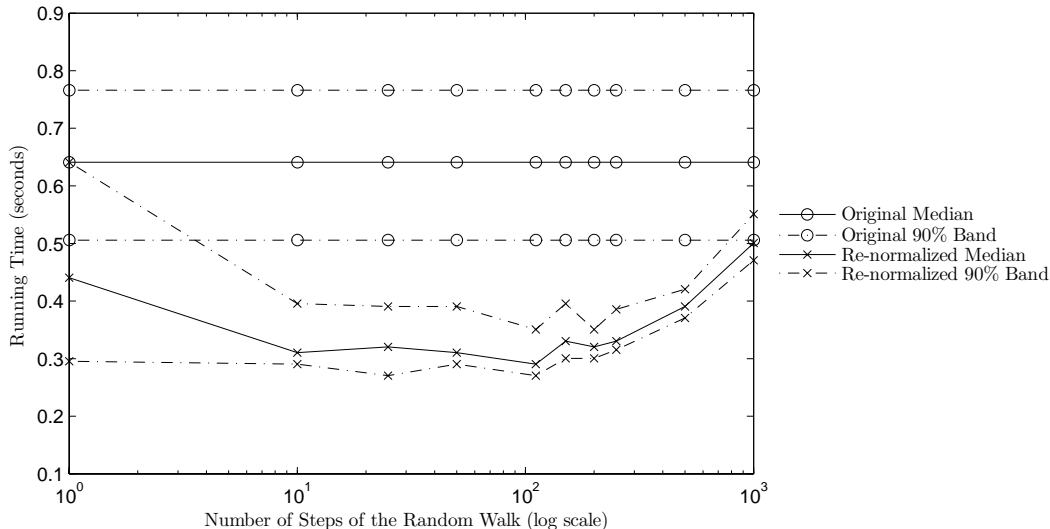
Fig. 4. IPM iterations versus number of steps of the geometric random walk for the 100 problem instances of dimension 100×500.
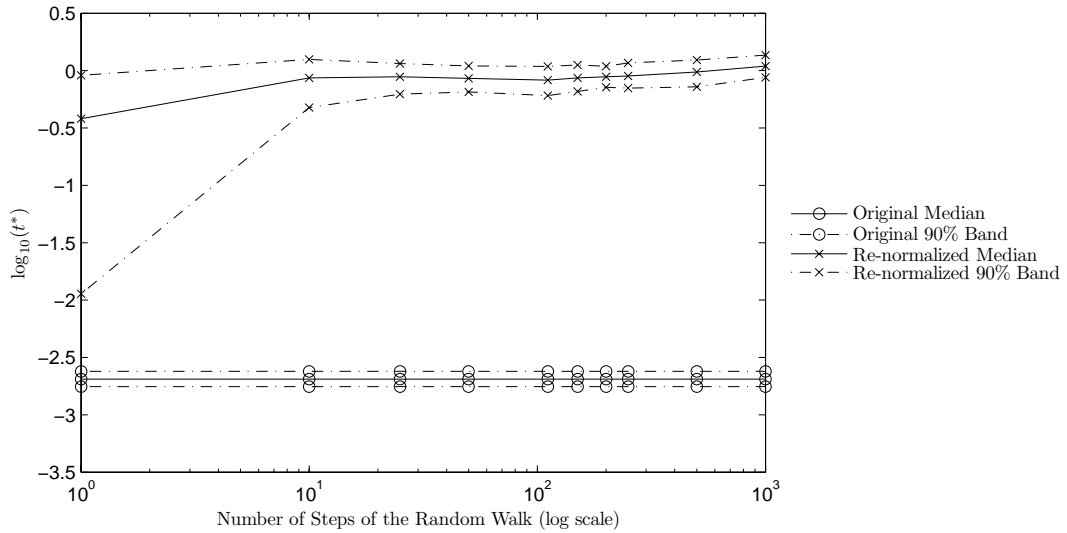
Fig. 5. $\log(t^*)$ versus number of steps of the geometric random walk for the 100 problem instances of dimension 100×500.

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

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

$\hat{v} \in H^{\circ}_{\bar{s}}$. Such a projective transformation serves to re-normalize the conic system into a system that has computational behavior with certain guarantees. We have given a characterization of 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 re-normalize the conic system to one whose computational complexity 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 $H^{\circ}_{\bar{s}}$ 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 re-normalization 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 $\bar{x} \in \mathrm{int}C$, 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 $\mathrm{sym}(0, H_{\bar{x}}) = \mathrm{sym}(0, H_{\bar{x}}')$ 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 re-normalization theory designed directly for the dual form F_a, and this is a direction for future research.

7.2. Conic Optimization

Given the structural equivalence of a homogeneous conic system (1) and conic optimization of the form $\max_x \{c^T x : Ax = b, \ x \in C\}$ under a regularity condition, one is led to the question of how to adapt/apply the projective re-normalization methodology herein to the optimization setting. The naive approach of simply homogenizing the primal and dual strong duality conditions will not work, as the resulting system will not have an interior solution, and the origin will lie on the boundary of the resulting image set. A more appropriate approach is to solve for an ε-optimal solution; this approach still has numerical challenges and is the subject of ongoing research.

7.3. Related Complexity Matters

When the projective re-normalization methodology is implemented using a random walk to find a deep point in $H_{\bar{s}}^\circ$, then the combined re-normalization and interior-point solution method can be viewed as a probabilistic algorithm with associated polynomial-time complexity with high probability. Viewed this way, the methodology proposed and tested herein is an addition to a growing number of polynomial-time probabilistic algorithms for convex feasibility and/or optimization, see Bertsimas and Vempala [3], Kalai and Vempala [13], and Kelner and Spielman [14].

Nesterov [21] 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 $\min_v \{f^*(\bar{s} - A^T v) : \bar{s} - A^T v \in \operatorname{int} C^*\}$. 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 [27] that the number of Newton steps of a short-step IPM to compute an approximate analytic center is: $O(\sqrt{\vartheta}\ln(\vartheta/\operatorname{sym}(0, H_{\bar{s}}^\circ))) = O(\sqrt{\vartheta}\ln(\vartheta/\operatorname{sym}(0, H_{\bar{s}})))$ (from (ii) of Remark 5), which is of the same order as the complexity for solving F developed herein in Theorem 1. These complexity bounds depend on $\operatorname{sym}(0, H_{\bar{s}})$ to bound the complexity of traversing the central path via a short-step method. As is shown in Nesterov and Nemirovski [23], 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, $\operatorname{sym}(0, H_{\bar{s}})$ lends itself to analysis, characterization, and ultimately manipulation and reduction via the constructive projective re-normalization 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 [23].

References

1. A. Belloni and R. M. Freund. Proejctive pre-conditioners for improving the behavior of a homogeneous conic linear system. Working Paper OR375-05, MIT, Operations Research Center, 2005.
2. A. Belloni and R. M. Freund. On the symmetry function of a convex set. *Mathematical Programming*, to appear, 2006.
3. D. Bertsimas and S. Vempala. Solving convex programs by random walks. *Journal of the ACM*, 51(4):540–556, 2004.
4. 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.
5. 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.
6. 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.
7. 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.
8. 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.
9. C. J. Geyer. Practical markov chain monte carlo. *Statistical Science*, 7(4):473–511, 1992.
10. M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combiantorial Optimization*. Springer-Verlag, Berlin, second edition, 1994.
11. B. Grünbaum. *Convex Polytopes*. Wiley, New York, 1967.
12. P. C. Hammer. The centroid of a convex body. *Proc. Amer. Math. Soc.*, 5:522–525, 1951.
13. A. Kalai and S. Vempala. Simulating annealing for convex optimization. *Mathematic of Operations Research*, to appear, 2006.
14. J.A. Kelner and D.A. Spielman. A randomized polynomial-time simplex algorithm for linear programming. Technical report, Proceedings of the 38th annual ACM symposium on theory of computing, 2006.
15. L. Khachiyan. Rounding of polytopes in the real number model of computation. *Mathematics of Operations Research*, 21(2):307–320, 1996.

16. L. Leindler. On a certain converse of Hölder's inequality ii. *Acta Sci. Math. Szeged*, 33:217–223, 1972.

17. L. Lovász and S. Vempala. The geometry of logconcave functions and an $O^*(n^3)$ sampling algorithm. *Microsoft Technical Report*.

18. L. Lovász and S. Vempala. Hit-and-run is fast and fun. *Microsoft Technical Report*.

19. L. Lovász and S. Vempala. Where to start a geometric walk? *Microsoft Technical Report*.

20. H. Minkowski. Allegemeine lehzätze über konvexe polyeder. *Ges. Abh., Leipzog-Berlin*, 1:103–121, 1911.

21. Y. Nesterov. private communication. 2005.

22. Y. Nesterov and A Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1993.

23. Y. Nesterov and A Nemirovskii. Central path and Riemannian distances. Technical report, CORE Discussion Paper, CORE, Université Catholique de Louvain, Belgium, 2003.

24. Y. Nesterov, M.J. Todd, and Y. Ye. Infeasible-start primal-dual methods and infeasibility detectors. *Mathematical Programming*, 84:227–267, 1999.

25. A. Prékopa. Logarithmic concave measures and functions. *Acta Sci. Math. Szeged*, 34:335–343, 1973.

26. A. Prékopa. On logarithmic concave measures with applications to stochastic programming. *Acta Sci. Math. Szeged*, 32:301–316, 1973.

27. J. Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2001.

28. R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, New Jersey, 1970.

29. 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.

30. Y. Zhang and L. Gao. On numerical solution of the maximum volume ellipsoid problem. *SIAM Journal on Optimization*, 14(1):53–76, 2003.