SIAM J. OPTIM.
Vol. 12, No. 3, pp. 627–655
© 2002 Society for Industrial and Applied Mathematics

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

A NEW CONDITION MEASURE, PRECONDITIONERS, AND RELATIONS BETWEEN DIFFERENT MEASURES OF CONDITIONING FOR CONIC LINEAR SYSTEMS*

MARINA EPELMAN† AND ROBERT M. FREUND‡

Abstract. In recent years, a body of research into "condition numbers" for convex optimization has been developed, aimed at capturing the intuitive notion of problem behavior. This research has been shown to be relevant in studying the efficiency of algorithms (including interior-point algorithms) for convex optimization as well as other behavioral characteristics of these problems such as problem geometry, deformation under data perturbation, etc. This paper studies measures of conditioning for a conic linear system of the form (FP_d): $Ax = b,\ x \in C_X$, whose data is $d = (A, b)$. We present a new measure of conditioning, denoted μ_d, and we show implications of μ_d for problem geometry and algorithm complexity and demonstrate that the value of $\mu = \mu_d$ is independent of the specific data representation of (FP_d). We then prove certain relations among a variety of condition measures for (FP_d), including μ_d, σ_d, $\overline{\chi}_d$, and $\mathcal{C}(d)$. We discuss some drawbacks of using the condition number $\mathcal{C}(d)$ as the sole measure of conditioning of a conic linear system, and we introduce the notion of a "preconditioner" for (FP_d), which results in an equivalent formulation ($\mathrm{FP}_{\tilde{d}}$) of (FP_d) with a better condition number $\mathcal{C}(\tilde{d})$. We characterize the best such preconditioner and provide an algorithm and complexity analysis for constructing an equivalent data instance $\tilde{d}$ whose condition number $\mathcal{C}(\tilde{d})$ is within a known factor of the best possible.

Key words. complexity of convex programming, conditioning, preconditioners

AMS subject classifications. 90C, 90C05, 90C25, 90C60

PII. S1052623400373829

1. Introduction.

The subject of this paper is the further study and development of a new measure of conditioning for the convex feasibility problem in conic linear form:

$$(\mathrm{FP}_d) : Ax = b,\ x \in C_X, \tag{1}$$

where $A \in \mathcal{L}(X, Y)$ is a linear operator between n- and m-dimensional spaces X and Y, $b \in Y$, and $C_X \subset X$ is a closed convex cone, $C_X \neq X$. We denote the data for the problem (FP_d) by $d = (A, b)$ (the cone C_X is regarded as fixed and given) and the set of solutions of (FP_d) by

$$X_d \stackrel{\triangle}{=} \{x \in X : Ax = b,\ x \in C_X\}.$$

The problem (FP_d) is an important tool in mathematical programming. It provides a very general format for studying the feasible regions of convex optimization problems (in fact, any convex feasibility problem can be modeled as a conic linear system) and includes linear programming and semidefinite programming feasibility problems as special cases. Over the last decade many important developments in linear programming, most notably the theory of interior-point methods, have been extended

*Received by the editors June 14, 2000; accepted for publication (in revised form) July 2, 2001; published electronically January 9, 2002.
http://www.siam.org/journals/siopt/12-3/37382.html

†University of Michigan, Industrial and Operations Engineering, 1205 Beal Avenue, Ann Arbor, MI 48109-2117 (mepelman@umich.edu). This author's research has been partially supported through an NSF Graduate Research Fellowship.

‡M.I.T. Sloan School of Management, 50 Memorial Drive, Cambridge, MA 02142-1347 (rfreund @mit.edu).

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

to convex problems in this form. In recent years, largely prompted by these developments, researchers have developed new and powerful theories of condition numbers for convex optimization, aimed at capturing the intuitive notion of problem behavior; this body of research has been shown to be important in studying the efficiency of algorithms, including interior-point algorithms, for convex optimization as well as other behavioral characteristics of these problems such as problem geometry, deformation under data perturbation, etc.

In this paper, we (i) develop a new measure of conditioning μ_d for (FP$_d$) that is invariant under equivalent data representations of the problem, (ii) establish the connection of the condition numbers μ_d and $\mathcal{C}(d)$ to some of the measures of conditioning arising in recent linear programming literature, and (iii) develop a theory of "preconditioners" for improving the condition number of (FP$_d$). We begin by briefly reviewing developments in the theory of measures of conditioning in recent literature as well as by providing an overview of the issues addressed in this paper.

The study of the computational complexity of linear programming originated with the analysis of the simplex algorithm, which, while extremely efficient in practice, was shown by Klee and Minty [15] to have worst-case complexity exponential in the number of variables. Khachiyan [14] demonstrated that linear programming problems were in fact polynomially solvable via the ellipsoid algorithm. Under the assumption that the problem data is rational, the ellipsoid algorithm requires at most $O(n^2L)$ iterations, where n is the number of variables and L is the *problem size*, which is roughly equal to the number of bits required to represent the problem data. The development of interior-point methods gave rise to algorithms that are efficient in theory as well as in practice (unlike the ellipsoid algorithm). The first such algorithm, developed by Karmarkar [13], has a complexity bound of $O(nL)$ iterations, and the algorithm introduced by Renegar [23] has a complexity bound of $O(\sqrt{n}L)$ iterations, which is currently the best known bound for linear programming. Many interior-point algorithms have also proven to be extremely efficient computationally and are often superior to the simplex algorithm.

Despite the importance of the above results, there are several serious drawbacks in analyzing algorithm performance in the bit-complexity framework. One such drawback is the fact that computers use floating point arithmetic, rather than integer arithmetic, in performing computations. As a result, two problems can have data that are extremely close but have drastically different values of L. The analysis of the performance of algorithms for solving these problems will yield different performance estimates, yet actual performance of the algorithms will likely be similar due to their similar numerical properties. See Wright [39] for a detailed discussion. A second drawback is that the complexity analysis of linear programming algorithms in terms of L largely relies on the combinatorial structure of the linear program; in particular, it relies on the fact that the set of feasible solutions is a polyhedron and the solution is attained at one of the extreme points of this polyhedron.

A relevant way to measure the intuitive notion of conditioning of a convex optimization (or feasibility) problem via the so-called *distance to ill-posedness* and the closely related *condition number* was developed by Renegar in [24] in a more specific setting, but then generalized more fully in [25] and in [26] to convex optimization and feasibility problems in conic linear form. Recall that $d = (A, b)$ is the data for the problem (FP$_d$) of (1). The *condition number* $\mathcal{C}(d)$ of (FP$_d$) is essentially a scale-invariant reciprocal of the smallest data perturbation $\Delta d = (\Delta A, \Delta b)$ for which the system (FP$_{d+\Delta d}$) changes its feasibility status. The problem (FP$_d$) is well-conditioned

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

to the extent that $\mathcal{C}(d)$ is small; when the problem (FP$_d$) is "ill-posed" (i.e., arbitrarily small perturbations of the data can yield both feasible and infeasible problem instances), then $\mathcal{C}(d) = +\infty$.

One of the important issues addressed by researchers is the relationship between the condition number $\mathcal{C}(d)$ and the geometry of the feasible region of (FP$_d$). Renegar [24] demonstrated that when a feasible instance of (FP$_d$) is well-posed ($\mathcal{C}(d) < \infty$), there exists a point x feasible for (FP$_d$) which satisfies $\|x\| \leq \mathcal{C}(d)$. Furthermore, it is shown in [8] that under the above assumption the set of feasible solutions contains a so-called "reliable" solution: A solution $\hat{x}$ of (FP$_d$) is reliable if, roughly speaking, (i) the distance from $\hat{x}$ to the boundary of the cone C_X, $\operatorname{dist}(\hat{x}, \partial C_X)$, is not excessively small; (ii) the norm of the solution $\|\hat{x}\|$ is not excessively large; and (iii) the ratio $\frac{\|\hat{x}\|}{\operatorname{dist}(\hat{x}, \partial C_X)}$ is not excessively large. The importance of reliable solutions is motivated in part by considerations of finite-precision computations. The results in [8] also demonstrate that when the system (FP$_d$) is feasible, there exists a feasible point $\hat{x}$ such that

$$\frac{\|\hat{x}\|}{\operatorname{dist}(\hat{x}, \partial C_X)} \leq c_1 \mathcal{C}(d), \quad \operatorname{dist}(\hat{x}, \partial C_X) \geq c_2 \frac{1}{\mathcal{C}(d)}, \quad \|\hat{x}\| \leq c_3 \mathcal{C}(d), \tag{2}$$

where the constants c_1, c_2, and c_3 depend only on the "width" of the cone C_X (to be formally defined shortly) and are independent of the data d of the problem (FP$_d$) (but may depend on n).

The condition number $\mathcal{C}(d)$ was also shown to be crucial for analyzing the complexity of algorithms for solving (FP$_d$). Renegar [26] presented an interior-point algorithm for solving (FP$_d$) with the complexity bound of $O(\sqrt{\vartheta} \ln(\vartheta\, \mathcal{C}(d)))$ iterations, where ϑ is the complexity parameter of a self-concordant barrier for the cone C_X. In [9] it was shown that a suitably modified version of the ellipsoid algorithm will solve (FP$_d$) in $O(n^2 \ln(\mathcal{C}(d)))$ iterations. (The constants in both complexity bounds depend on the width of C_X.) In [4], a generalization of a row-action algorithm is shown to compute a reliable solution of (FP$_d$) in the sense of (2). The complexity of this algorithm is also closely tied to $\mathcal{C}(d)$.

The recent literature has explored many other important properties of the problem (FP$_d$) tied to the distance to ill-posedness and the condition number $\mathcal{C}(d)$. Renegar [24] studied the relation of $\mathcal{C}(d)$ to sensitivity of solutions of (FP$_d$) under perturbations in the problem data. (This issue was also investigated earlier by Robinson [28].) Peña and Renegar [22] discussed the role of $\mathcal{C}(d)$ in the complexity of computing approximate solutions of (FP$_d$). Freund and Vera [7] and Peña [20] addressed the theoretical complexity and practical aspects of computing the distance to ill-posedness. Vera [38] considered the numerical properties of an interior-point method for solving (FP$_d$) (and, in fact, a more general problem of optimizing a linear function over the feasible region of (FP$_d$)) in the case when (FP$_d$) is a linear programming problem. He considered the algorithm in the floating point arithmetic model, and demonstrated that the algorithm will approximately solve the optimization problem in polynomial time, while requiring roughly $O(\ln(\mathcal{C}(d)))$ significant digits of precision for computation. For additional discussion of ill-posedness and the condition number, see Filipowski [6, 5], Nunez and Freund [19], Nunez [18], Peña [21, 20], and Vera [35, 36, 37].

As we hope the above discussion conveys, the condition number $\mathcal{C}(d)$ is a relevant and important measure of conditioning of the problem (FP$_d$). Note that when (FP$_d$) is in fact a linear programming feasibility problem, $\mathcal{C}(d)$ provides a measure of

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

conditioning that, unlike L, does not rely on the assumption that the problem data is rational, and is relevant in the floating point model of computation.

Nevertheless, there are some potential drawbacks in using $\mathcal{C}(d)$ as a sole measure of conditioning of the problem (FP$_d$). To illustrate this point, note that problem (FP$_d$) of (1) can be interpreted as the problem of finding a point x in the intersection of the cone C_X with an affine subspace $\mathcal{A} \subset X$, defined as

$$\mathcal{A} \stackrel{\triangle}{=} \{x : Ax = b\} = \{x : x = x_0 + x_N,\ x_N \in \text{Null}(A)\},$$

where $x_0 \in X$ is an arbitrary point satisfying $Ax_0 = b$, and Null(A) is the null space of A. Notice that the description of the affine subspace $\mathcal{A}$ by the data instance $d = (A, b)$ is not unique. It easy to find an equivalent data instance $\tilde{d} = (\tilde{A}, \tilde{b})$ such that

$$\{x : \tilde{A}x = \tilde{b}\} = \{x : Ax = b\} = \mathcal{A}$$

(take, for example, $\tilde{b} = Bb$ and $\tilde{A} = BA$, where B is any nonsingular linear operator $B : Y \to Y$). Then the problem

$$(\text{FP}_{\tilde{d}}) : \tilde{A}x = \tilde{b},\ x \in C_X$$

is equivalent to problem (FP$_d$) in the sense that their feasible regions are identical; we can think of the systems (FP$_d$) and (FP$_{\tilde{d}}$) as different but equivalent formulations of the same feasibility problem

$$(\text{FP}): \text{find } x \in \mathcal{A} \cap C_X.$$

Since the condition number $\mathcal{C}(d)$ is, in general, different from $\mathcal{C}(\tilde{d})$, analyzing many of the properties of the problem (FP) above in terms of the condition number will lead to different results, depending on which formulation, (FP$_d$) or (FP$_{\tilde{d}}$), is being used. This observation is somewhat disconcerting, since many of these properties are of purely geometric nature. For example, the existence of a solution of small norm and the existence of a reliable solution depend only on the geometry of the feasible region, i.e., of the set $\mathcal{A} \cap C_X$, and do not depend on a specific data instance d used to "represent" the affine space $\mathcal{A}$.

An interesting research direction, therefore, is the development of relevant measures of conditioning of the problem (FP$_d$) that depend on the affine space $\mathcal{A}$ rather than on a particular data instance d used to represent it and that allow us to analyze some of the properties of the problem independently of the data used to represent the problem. The recent literature contains some results on developing such measures when (FP$_d$) is a linear programming feasibility problem. In particular, two condition measures, $\bar{\chi}_d$ and σ_d, were used in the analysis of interior-point algorithms for linear programming (Vavasis and Ye [32, 33, 34]). These measures, discussed in detail in section 4, provide a new perspective on the analysis of linear programming problems; for example, like the condition number $\mathcal{C}(d)$, they do not require the data for the problem to be rational. Also, they have the desired property that they are independent of the specific data instance d used to describe the problem and can be defined considering only the affine subspace $\mathcal{A}$. Further analysis of these measures in the setting of linear programming feasibility problems can be found in Ho [11], Todd, Tunçel, and Ye [29], and Tunçel [30].

In this paper we define a new measure of conditioning, μ_d, for feasible instances of the problem (FP$_d$) of (1), which is independent of the specific data representation

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

of the problem. We explore the relationship between μ_d and measures $\bar{\chi}_d$, σ_d, and $\mathcal{C}(d)$. (In particular, we demonstrate that the measure σ_d is directly related to μ_d in the special case of linear programming.) We show that $\mu_d \leq \mathcal{C}(d)$, i.e., μ_d is less conservative, and that for any data instance $\tilde{d}$ equivalent to d, $\mu_d \leq \mathcal{C}(\tilde{d})$. We also demonstrate that many important properties of the system (FP_d) previously analyzed in terms of $\mathcal{C}(d)$ can be analyzed through μ_d (independently of the data representation).

On the other hand, some properties of (FP_d) are not purely geometric and depend on the data d. Therefore, it might be beneficial, given a data instance d, to construct a data instance $\tilde{d}$ which is equivalent to d but is better conditioned in the sense that $\mathcal{C}(\tilde{d}) < \mathcal{C}(d)$. We develop a characterization of all equivalent data instances $\tilde{d}$ by introducing the concept of a *preconditioner* and provide an upper bound on the condition number $\mathcal{C}(\tilde{d})$ of the "best" equivalent data instance $\tilde{d}$. We also analyze the complexity of computing an equivalent data instance whose resulting condition number is within a known factor of this bound. To this end, we construct an algorithm for computing such a data instance and analyze its complexity.

An outline of the paper is as follows. Section 2 contains notation, definitions, assumptions, and preliminary results. In section 3 we introduce the new measure of conditioning μ_d for (FP_d), establish several results relating μ_d to geometric properties of the feasible region of (FP_d), and analyze the performance of several algorithms for solving (FP_d) in terms of μ_d. In section 4 we study the relationship between μ_d and other measures of conditioning, completely characterizing the relationship between $\mathcal{C}(d)$ and μ_d, as well as σ_d and $\bar{\chi}_d$, in the linear programming setting. In section 5, we develop the notion of a preconditioner for the problem (FP_d), establish an upper bound on the condition number $\mathcal{C}(\tilde{d})$ of the best equivalent data instance $\tilde{d}$, and construct and analyze an algorithm for computing an equivalent data instance whose condition number is within a known factor of this bound. Section 6 contains some final conclusions and indicates potential topics of future research.

2. Preliminaries. We work in the setup of finite-dimensional normed linear vector spaces. Both X and Y are normed linear spaces of finite dimension n and m, respectively, endowed with norms $\|x\|$ for $x \in X$ and $\|y\|$ for $y \in Y$. For $\bar{x} \in X$, let $B(\bar{x}, r)$ denote the ball centered at $\bar{x}$ with radius r, i.e., $B(\bar{x}, r) = \{x \in X : \|x - \bar{x}\| \leq r\}$, and define $B(\bar{y}, r)$ analogously for $\bar{y} \in Y$. We denote the set of real numbers by $\Re$ and the set of nonnegative real numbers by $\Re_+$. The set of real k-by-k symmetric matrices is denoted by $S^{k\times k}$. The set $S^{k\times k}$ is a closed linear space of dimension $n = \frac{k(k+1)}{2}$. We denote the set of symmetric positive semidefinite k-by-k matrices by $S^{k\times k}_+$. $S^{k\times k}_+$ is a closed convex cone in $S^{k\times k}$. The interior of the cone $S^{k\times k}_+$ is precisely the set of k-by-k positive definite matrices, and is denoted by $S^{k\times k}_{++}$.

We associate with X and Y the dual spaces X^* and Y^* of linear functionals defined on X and Y, respectively. Let $c \in X^*$. In order to maintain consistency with standard linear algebra notation in mathematical programming, we will denote the linear function $c(x)$ by $c^t x$. Similarly, for $f \in Y^*$ we denote $f(y)$ by $f^t y$. We denote $A(x)$ by Ax, and we denote the dual operator of A by $A^t : Y^* \to X^*$.

The dual norm induced on $c \in X^*$ is defined as

$$\|c\|_* \stackrel{\triangle}{=} \max\{c^t x : x \in X,\ \|x\| \leq 1\}, \tag{3}$$

and the Hölder inequality $c^t x \leq \|c\|_* \|x\|$ follows easily from this definition. The dual norm induced on $f \in Y^*$ is defined similarly.

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

We now present the development of the concepts of condition numbers and data perturbation for (FP$_d$) in detail. Recall that $d = (A, b)$ is the data for the problem (FP$_d$). Let

$$\mathcal{D} = \{d = (A, b) : A \in L(X, Y), b \in Y\}$$

denote the space of all data $d = (A, b)$ for (FP$_d$). For $d = (A, b) \in \mathcal{D}$ we define the norm on the Cartesian product $L(X, Y) \times Y$ to be

$$\|d\| = \|(A, b)\| = \max\{\|A\|, \|b\|\},$$

where $\|b\|$ is the norm specified for Y and $\|A\|$ is the operator norm, namely

$$\|A\| = \max\{\|Ax\| : \|x\| \leq 1\}.$$

We define

$$\mathcal{F} = \{(A, b) \in \mathcal{D} : \text{ there exists } x \text{ satisfying } Ax = b,\ x \in C_X\}$$

to be the set of data instances d for which (FP$_d$) is feasible. Its complement is denoted by $\mathcal{F}^{\mathcal{C}}$, the set of data instances for which (FP$_d$) is infeasible. The boundary of $\mathcal{F}$ and of $\mathcal{F}^{\mathcal{C}}$ is precisely the set $\mathcal{B} = \partial\mathcal{F} = \partial\mathcal{F}^{\mathcal{C}} = \text{cl}(\mathcal{F}) \cap \text{cl}(\mathcal{F}^{\mathcal{C}})$, where ∂S denotes the boundary and cl(S) denotes the closure of a set S. Note that if $d = (A, b) \in \mathcal{B}$, then (FP$_d$) is ill-posed in the sense that arbitrarily small changes in the data $d = (A, b)$ can yield instances of (FP$_d$) that are feasible as well as instances of (FP$_d$) that are infeasible. Also, note that $\mathcal{B} \neq \emptyset$, since $d = 0 \in \mathcal{B}$.

For a data instance $d = (A, b) \in \mathcal{D}$, the *distance to ill-posedness* is defined to be

$$\rho(d) \stackrel{\triangle}{=} \inf\{\|\Delta d\| : d + \Delta d \in \mathcal{B}\} = \begin{cases} \inf\{\|d - \bar{d}\| : \bar{d} \in \mathcal{F}^C\} & \text{if } d \in \mathcal{F}, \\ \inf\{\|d - \bar{d}\| : \bar{d} \in \mathcal{F}\} & \text{if } d \in \mathcal{F}^C; \end{cases} \tag{4}$$

see Renegar [24, 25, 26]. The *condition number* $\mathcal{C}(d)$ of the data instance d is defined to be

$$\mathcal{C}(d) = \frac{\|d\|}{\rho(d)} \tag{5}$$

when $\rho(d) > 0$, and $\mathcal{C}(d) = \infty$ when $\rho(d) = 0$. The condition number $\mathcal{C}(d)$ is a measure of the relative conditioning of the data instance d and can be viewed as a scale-invariant reciprocal of $\rho(d)$, as it is elementary to demonstrate that $\mathcal{C}(d) = \mathcal{C}(\alpha d)$ for any positive scalar α. It is easy to show that $\rho(0) = 0$, and hence $\mathcal{C}(d) \geq 1$.

If C is a convex cone in X, then the dual cone of C, denoted by C^*, is defined by

$$C^* = \{z \in X^* : z^t x \geq 0 \text{ for any } x \in C\}. \tag{6}$$

We will say that a cone C is *regular* if C is a closed convex cone, has a nonempty interior, and is pointed (i.e., contains no line). If C is a closed convex cone, then C is regular if and only if C^* is regular.

We will use the following definition of the *width* of a regular cone C.

DEFINITION 1. *If C is a regular cone in X, the* width *of C is given by*

$$\tau_C \stackrel{\triangle}{=} \max_{x,r} \left\{ \frac{r}{\|x\|} : B(x, r) \subset C \right\}.$$

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Note that $\tau_C \in (0,1]$, since C is pointed and has a nonempty interior, and τ_C is attained for some $(\bar{x},\bar{r})$ as well as along the ray $(\alpha\bar{x},\alpha\bar{r})$ for all $\alpha > 0$. By choosing the value of α appropriately, we can find $u \in C$ such that

$$\|u\| = 1 \text{ and } \tau_C \text{ is attained for } (x,r) = (u,\tau_C). \tag{7}$$

DEFINITION 2. *If C is a regular cone in X, define the norm approximation coefficient by*

$$\delta_C \stackrel{\triangle}{=} \operatorname{dist}(0,\, \partial\operatorname{conv}(C(1),-C(1)), \tag{8}$$

where $C(1) \stackrel{\triangle}{=} \{x \in C : \|x\| \le 1\}$, and $\partial\operatorname{conv}(C(1),-C(1))$ is the boundary of the convex hull of the set $C(1) \cup (-C(1))$.

The norm approximation coefficient δ_C measures the extent to which the unit ball $B(0,1) \subset X$ can be approximated by the set $\operatorname{conv}(C(1),-C(1))$. As a consequence, it measures the extent to which the norm of a linear operator can be approximated over the set $C(1)$.

PROPOSITION 3. *Suppose $A \in L(X,Y)$. Then $\|A\| \le \frac{1}{\delta_C}\max\{\|Ax\| : x \in C(1)\}$.*

LEMMA 4. *Suppose C is a regular cone with width τ_C. Then*

$$\delta_C \ge \frac{\tau_C}{1+\tau_C} \ge \frac{\tau_C}{2}. \tag{9}$$

Proof. Let $\bar{x} \in X$ be an arbitrary vector satisfying $\|\bar{x}\| \le \frac{\tau_C}{1+\tau_C}$. To establish the lemma we need to show that $\bar{x} \in \operatorname{conv}(C(1),-C(1))$.

Let $x = \frac{\bar{x}(1+\tau_C)}{\tau_C}$. If u is as in (7), then $u+\tau_C x \in C$ and $u-\tau_C x \in C$. Furthermore,

$$\frac{u+\tau_C x}{1+\tau_C} \in C(1) \quad \text{and} \quad \frac{-u+\tau_C x}{1+\tau_C} \in -C(1),$$

and so

$$\bar{x} = \frac{\tau_C}{1+\tau_C}x = \frac{1}{2}\left(\frac{u+\tau_C x}{1+\tau_C}\right) + \frac{1}{2}\left(\frac{-u+\tau_C x}{1+\tau_C}\right) \in \operatorname{conv}(C(1),-C(1)). \qquad \square$$

We will assume throughout this paper that the system (FP_d) of (1) is feasible. At this point we make no further assumptions on the cone C_X and the norms on the spaces X and Y unless stated otherwise. (We will make some additional assumptions in sections 4 and 5.)

When (FP_d) is feasible, $\rho(d)$ can be expressed via the following characterization:

$$\rho(d) = \max\{r : B(0,r) \subseteq \mathcal{H}_d\}, \tag{10}$$

where

$$\mathcal{H}_d \stackrel{\triangle}{=} \{b\theta - Ax : \theta \ge 0,\ x \in C_X,\ |\theta| + \|x\| \le 1\} \subset Y. \tag{11}$$

Note that $0 \in \mathcal{H}_d$ whenever (FP_d) is feasible, and $\rho(d) > 0$ precisely when $0 \in \operatorname{int}\mathcal{H}_d$. This interpretation, presented by Renegar in [26], will serve as an important tool in developing further understanding of the properties of the system (FP_d).

The next result follows from the definition of $\mathcal{H}_d$ and Proposition 3.

COROLLARY 5. *Suppose that $d = (A,b) \in \mathcal{D}$ and C_X is regular. Then $\|d\| \le \frac{1}{\delta_{C_X}}\max\{\|h\| : h \in \mathcal{H}_d\}$.*

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

3. The symmetry measure μ_d. In this section we define a new measure of conditioning of (FP$_d$), μ_d, which we refer to as the "symmetry measure," and we establish some of its properties relevant in the analysis of (FP$_d$). We begin by recalling the *symmetry* of a set with respect to a point, in the following definition.

DEFINITION 6. *Let $D \subset Y$ be a bounded convex set. For $y \in \operatorname{int} D$ we define* $\operatorname{sym}(D, y)$ *to be the symmetry of D about y, i.e.,*

$$\operatorname{sym}(D, y) \stackrel{\triangle}{=} \sup\{t \mid y + v \in D \Rightarrow\ y - tv \in D\}.$$

If $y \in \partial D$, we define $\operatorname{sym}(D, y) = 0$.

This definition of symmetry is equivalent to that given in [26]. Observe that $\operatorname{sym}(D, y) \in [0, 1]$, with $\operatorname{sym}(D, y) = 1$ if D is perfectly symmetric about y, and $\operatorname{sym}(D, y) = 0$ precisely when $y \in \partial D$. Moreover, the definition of $\operatorname{sym}(D, y)$ is independent of the norm on the space Y.

LEMMA 7. *Suppose that D is a compact convex set with a nonempty interior, and let $y \in \operatorname{int} D$. Then there exists an extreme point w of D such that* $\operatorname{sym}(D, y) = \operatorname{sym}_w(D, y) \stackrel{\triangle}{=} \sup\{t \mid y - t(w - y) \in D\}$.

Proof. Define $f(w) = \operatorname{sym}_w(D, y) = \sup\{t \mid y - t(w - y) \in D\}$. It follows that $f(w)$ is a quasi-concave function on D. This implies that the minimum of $f(w)$ is attained at an extreme point of D; see, for example, section 3.5.3 of [1]. ∎

To define the *symmetry measure* of the problem (FP$_d$) recall that if (FP$_d$) is feasible, then $0 \in \mathcal{H}_d$, where $\mathcal{H}_d$ is defined in (11). Hence, the following quantity is well-defined.

DEFINITION 8. *Suppose the system* (FP$_d$) *is feasible. We define*

$$\mu_d \stackrel{\triangle}{=} \frac{1}{\operatorname{sym}(\mathcal{H}_d, 0)} \tag{12}$$

when $\operatorname{sym}(\mathcal{H}_d, 0) > 0$, *and* $\mu_d = +\infty$ *when* $\operatorname{sym}(\mathcal{H}_d, 0) = 0$.

From the above definition, $\mu_d \geq 1$ and $\mu_d = +\infty$ precisely when $0 \in \partial\mathcal{H}_d$, i.e., precisely when (FP$_d$) is ill-posed.

3.1. The symmetry measure and geometric properties of solutions of (FP$_d$). We now establish two results that characterize geometric properties of the feasible region X_d of the system (FP$_d$) in terms of μ_d. Theorem 9 establishes a bound on the size of a solution of (FP$_d$) in terms of μ_d; this result is similar to the bound in terms of the condition number $\mathcal{C}(d)$ in [24]. Theorem 10 demonstrates existence of a reliable solution of (FP$_d$). This is similar to the result (2) presented in [8]; however, here the bounds on the size of the solution, its distance to the boundary of the cone C_X, and the ratio of the above quantities are established in terms of μ_d rather than $\mathcal{C}(d)$. Also, unlike for the condition number $\mathcal{C}(d)$, we can establish a converse result for μ_d; namely, if the feasible region possesses nice geometry, i.e., contains a reliable solution, then μ_d can be nicely bounded by a function of the parameters associated with the reliable solution. This result is proven in Theorem 11.

THEOREM 9. *Suppose $\mu_d < \infty$. Then there exists $x \in X_d$ such that $\|x\| \leq \mu_d$.*

Proof. By the definition of μ_d, $-\frac{1}{\mu_d} b = -\operatorname{sym}(\mathcal{H}_d, 0) b \in \mathcal{H}_d$, since $b \in \mathcal{H}_d$. Therefore there exists (θ, x) satisfying $\theta \geq 0,\ x \in C_X,\ |\theta| + \|x\| \leq 1$, and $b\theta - Ax = -\frac{1}{\mu_d} b$. Let $\hat{x} = x/(\theta + \frac{1}{\mu_d})$. Then $\hat{x} \in X_d$ and $\|\hat{x}\| = \|x\|/(\theta + \frac{1}{\mu_d}) \leq \mu_d$. ∎

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

THEOREM 10. *Suppose C_X is a regular cone with width τ, and that $\mu_d < \infty$. Then there exist $\hat{x}$ and $r > 0$ such that*

1. $\hat{x} \in X_d$,
2. $\|\hat{x}\| \leq 2\mu_d + 1$,
3. $\operatorname{dist}(\hat{x}, \partial C_X) \geq r \geq \frac{\tau}{2\mu_d+1}$,
4. $\frac{\|\hat{x}\|}{r} \leq \frac{2\mu_d+1}{\tau}$.

Proof. Let u be as in (7). Then $\frac{1}{2}b - \frac{1}{2}Au \in \mathcal{H}_d$. From the definition of μ_d we conclude that $-\frac{1}{\mu_d}\left(\frac{1}{2}b - \frac{1}{2}Au\right) \in \mathcal{H}_d$, whereby there exists $(\bar{\theta}, \bar{x}) \in \Re_+ \times C_X$, $|\bar{\theta}| + \|\bar{x}\| \leq 1$, satisfying $b\bar{\theta} - A\bar{x} = -\frac{1}{\mu_d}(\frac{1}{2}b - \frac{1}{2}Au)$.

Let $\hat{x} = \frac{2\mu_d\bar{x}+u}{2\mu_d\bar{\theta}+1}$. It is easy to verify that $\hat{x} \in X_d$, so that condition 1 of the theorem is satisfied. Moreover, $\|\hat{x}\| = \frac{\|2\mu_d\bar{x}+u\|}{2\mu_d\bar{\theta}+1} \leq 2\mu_d + 1$, establishing condition 2.

Next, let $r = \frac{\tau}{2\mu_d\bar{\theta}+1}$. Since $B(u, \tau) \subset C_X$ and $\bar{x} \in C_X$, we conclude that $B(u + 2\mu_d\bar{x}, \tau) \subset C_X$, and therefore $B(\frac{u+2\mu_d\bar{x}}{2\mu_d\bar{\theta}+1}, \frac{\tau}{2\mu_d\bar{\theta}+1}) = B(\hat{x}, r) \subset C_X$. Also, since $\bar{\theta} \leq 1$, $r \geq \frac{\tau}{2\mu_d+1}$, establishing condition 3. Finally,

$$\frac{\|\hat{x}\|}{r} = \frac{\|2\mu_d\bar{x} + u\|}{2\mu_d\bar{\theta} + 1} \cdot \frac{2\mu_d\bar{\theta} + 1}{\tau} \leq \frac{2\mu_d + 1}{\tau},$$

implying condition 4 and concluding the proof of the theorem. ∎

We conclude from Theorems 9 and 10 that, much like for the condition number $\mathcal{C}(d)$, if the symmetry measure μ_d is small, then the feasible region X_d possesses nice geometry. We now establish a converse result.

THEOREM 11. *Suppose C_X is a regular cone and there exists $\hat{x} \in X_d$ and $r > 0$ such that $\operatorname{dist}(\hat{x}, \partial C_X) \geq r$. Let $\gamma = \max\{\|\hat{x}\|, \frac{1}{r}, \frac{\|\hat{x}\|}{r}\}$. Then $\mu_d \leq 1 + 2\gamma$.*

Proof. Let $\delta = \|\hat{x}\| + 1$ and $\pi = \min\{r, 1\}$. We first show that $\operatorname{sym}(\mathcal{H}_d, 0) \geq \frac{\pi}{\delta+\pi}$.

Let $y \in \mathcal{H}_d$. From the definition of $\mathcal{H}_d$, $y = b\bar{\theta} - A\bar{x}$ for some $(\bar{\theta}, \bar{x}) \in \Re_+ \times C_X$, $|\bar{\theta}| + \|\bar{x}\| \leq 1$. Therefore

$$\frac{\pi}{\delta+\pi}(-y) = \frac{\pi}{\delta+\pi}(-b\bar{\theta} + A\bar{x}) + \frac{1}{\delta+\pi}(b - A\hat{x}) = b\left(\frac{-\pi\bar{\theta}+1}{\delta+\pi}\right) - A\left(\frac{-\pi\bar{x}+\hat{x}}{\delta+\pi}\right).$$

Let $\check{\theta} = \frac{-\pi\bar{\theta}+1}{\delta+\pi}$ and $\check{x} = \frac{-\pi\bar{x}+\hat{x}}{\delta+\pi}$. Since $\pi \leq 1$ and $\bar{\theta} \leq 1$, we have $\check{\theta} \geq 0$. Moreover, since $\pi \leq r$ and $\|\bar{x}\| \leq 1$, we have $\check{x} \in C_X$. Finally, $|\check{\theta}| + \|\check{x}\| \leq \frac{1}{\delta+\pi}(1 + \pi\|\bar{x}\| + \|\hat{x}\|) \leq 1$, and therefore $-\frac{\pi}{\delta+\pi}y \in \mathcal{H}_d$ for an arbitrary $y \in \mathcal{H}_d$, establishing that $\operatorname{sym}(\mathcal{H}_d, 0) \geq \frac{\pi}{\delta+\pi}$. Hence,

$$\mu_d = \frac{1}{\operatorname{sym}(\mathcal{H}_d, 0)} \leq \frac{\delta+\pi}{\pi} = 1 + \frac{1}{\min\{r,1\}} + \frac{\|\hat{x}\|}{\min\{r,1\}} \leq 1 + \max\{\gamma, 1\} + \gamma \leq 1 + 2\gamma.$$

The last inequality follows from the observation that $r \leq \|\hat{x}\|$ (since C_X is pointed and thus $\|\hat{x}\| \geq \operatorname{dist}(\hat{x}, \partial C_X) \geq r$) and thus $\gamma \geq \frac{\|\hat{x}\|}{r} \geq 1$. ∎

The result in Theorem 11 is quite specific to μ_d; no such result is possible for the condition number $\mathcal{C}(d)$. In fact, the example following Remark 19 in section 4 shows that $\mathcal{C}(d)$ can be arbitrarily large even when γ is fixed.

3.2. The symmetry measure and the complexity of computing a solution of ($\mathbf{FP}_d$). In this subsection we present complexity bounds for solving (FP_d) via an interior-point algorithm and via the ellipsoid algorithm, and we show that the

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

complexity of solving (FP$_d$) depends on $\ln(\mu_d)$ as well as on other naturally appearing quantities. For this subsection, we assume that the space X is an n-dimensional Euclidean space with Euclidean norm $\|x\| = \|x\|_2 = \sqrt{x^t x}$ for $x \in X$. We also assume that C_X is a regular cone with width τ and that the vector u of (7) is known.

When the cone C_X is represented as the closure of the domain of a self-concordant barrier function, a solution of (FP$_d$) can be found using the barrier method developed by Renegar, based on the theory of self-concordant functions of Nesterov and Nemirovskii [17]. Below we briefly review the barrier method as articulated in [27] and then state the main complexity result.

The version of the barrier method that we will use is designed to approximately solve a problem of the form

$$z_* = \inf\{c^t\omega : \omega \in S \cap L\}, \tag{13}$$

where S is a bounded set whose interior is convex and is the domain of a self-concordant barrier function $f(\omega)$ with complexity parameter ϑ_f (see [17] and [27] for details), and L is a closed subspace (or a translate of a closed subspace). The barrier method takes as input a point $\omega' \in \operatorname{int} S \cap L$, and proceeds by approximately following the *central path*, i.e., the sequence of solutions of the problems

$$z(\eta) = \inf_{\omega \in L} \eta \cdot c^t\omega + f(\omega),$$

where $\eta > 0$ is the *barrier parameter*. In particular, after the initialization stage, the method generates an increasing sequence of barrier parameters $\eta_k > 0$ and iterates $\omega_k \in \operatorname{int} S \cap L$ that satisfy

$$c^t\omega_k - \frac{6\vartheta_f}{5\eta_k} \le z_* \le c^t\omega_k, \qquad k = 0, 1, 2, \ldots. \tag{14}$$

It follows from the analysis in [27] that if the barrier method is initialized at the point $\omega' \in \operatorname{int} S \cap L$, then it will take at most

$$O\left(\sqrt{\vartheta_f} \ln\left(\frac{\vartheta_f(z^* - z_*)}{\operatorname{sym}(S \cap L, \omega')} \cdot \bar{\eta}\right)\right) \tag{15}$$

iterations to bring the value of the barrier parameter η above the threshold of $\bar{\eta} \ge \eta_0$ while maintaining (14). (Here, $z^* = \sup\{c^t\omega : \omega \in S \cap L\}$.) This implies the main convergence result for the barrier method, which follows.

THEOREM 12 (see [27, Theorem 2.4.10]). *Assume that S is a bounded set whose interior is convex and is the domain of a self-concordant barrier function $f(\omega)$ with complexity parameter ϑ_f, and that L is a closed subspace (or a translate of a closed subspace). Assume that the barrier method is initialized at a point $\omega' \in \operatorname{int} S \cap L$. If $0 < \epsilon < 1$, then within*

$$O\left(\sqrt{\vartheta_f} \ln\left(\frac{\vartheta_f}{\epsilon \operatorname{sym}(S \cap L, \omega')}\right)\right)$$

iterations of the method, all points ω computed thereafter satisfy $\omega \in \operatorname{int} S \cap L$ and

$$\frac{c^t\omega - z_*}{z^* - z_*} \le \epsilon.$$

In order to find a solution of (FP$_d$) we will construct a closely related problem of the form (13) and apply the barrier method to this problem. This construction was

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

carried out in [26], where the complexity of solving (FP_d) was analyzed in terms of $\mathcal{C}(d)$. The optimization problem we consider is

$$\begin{array}{rll} z_* = & \inf_{\theta,x,t} & t \\ & \text{subject to (s.t.)} & b\theta - Ax = t(\frac{1}{2}b - \frac{1}{2}Au), \\ & & x \in \operatorname{int} C_X, \\ & & \|x\| < 1, \\ & & 0 < \theta < 1, \\ & & -1 < t < 2, \end{array} \tag{16}$$

where u is chosen as in (7). We will use the barrier method to find a feasible solution $(\hat{\theta}, \hat{x}, \hat{t})$ of (16) such that $\hat{t} \le 0$, and use the transformation $x = \hat{x} - \frac{1}{2}\hat{t}u/(\hat{\theta} - \frac{1}{2}\hat{t})$ to obtain a solution of (FP_d).

Let z^* be the optimal value of the problem obtained from (16) by replacing "inf" with "sup". Let $\tilde{f}(x)$ be the self-concordant barrier function defined on $\operatorname{int} C_X$ and let $\vartheta_{\tilde{f}}$ be the complexity parameter of $\tilde{f}(x)$. Then the set $S \stackrel{\triangle}{=} \{(\theta, x, t) : x \in \operatorname{int} C_X,\ \|x\| < 1,\ 0 < \theta < 1,\ -1 < t < 2\}$ is convex and bounded, and is the domain of the self-concordant barrier function

$$f(\omega) = f(\theta, x, t) = \tilde{f}(x) - \ln(1 - \|x\|^2) - \ln\theta - \ln(1 - \theta) - \ln(t + 1) - \ln(2 - t)$$

with complexity parameter $\vartheta_f \le \vartheta_{\tilde{f}} + 5$. (See, for example, [26] or [27] for details.) If we define $L \stackrel{\triangle}{=} \{(\theta, x, t) : b\theta - Ax = t(\frac{1}{2}b - \frac{1}{2}Au)\}$, then problem (16) is of the form (13), and we can apply the barrier method initialized at the point $\omega' = (\theta', x', t') = (\frac{1}{2}, \frac{1}{2}u, 1)$. The following proposition provides bounds on all of the parameters necessary in the analysis of the complexity of the barrier method via Theorem 12.

PROPOSITION 13. $z^* \le 2,\ -1 \le z_* \le -\frac{1}{\mu_d},\ \operatorname{sym}(S \cap L, \omega') \ge \frac{1}{12\tau}$.

Proof. The upper bound on z^* and the lower bound on z_* follow from the last constraint of (16).

Let $y = \frac{1}{2}b - \frac{1}{2}Au \in \mathcal{H}_d$. From the definition of μ_d we conclude that $-\frac{y}{\mu_d} \in \mathcal{H}_d$, so there exists (θ, x) such that $\theta \ge 0,\ x \in C_X,\ |\theta| + \|x\| \le 1,\ b\theta - Ax = -\frac{1}{\mu_d}(\frac{1}{2}b - \frac{1}{2}Au)$. Therefore $(\theta, x, -1/\mu_d)$ is in the closure of the feasible set of (16), and so $z_* \le -\frac{1}{\mu_d}$.

To establish the last statement of the proposition, we appeal to Proposition 3.3 of Renegar [26], where it is shown that ω' defined above satisfies

$$\operatorname{sym}(S \cap L, \omega') \ge \frac{1}{4}\operatorname{sym}\left(C_X(1), \frac{1}{2}u\right), \text{ where } C_X(1) = \{x : x \in C_X,\ \|x\| \le 1\}.$$

Since $B\left(\frac{1}{2}u, \frac{1}{2}\tau\right) \subset C_X(1)$, it is easy to verify that $\operatorname{sym}\left(C_X(1), \frac{1}{2}u\right) \ge \frac{\tau}{3}$, establishing the proposition. ∎

THEOREM 14. *Suppose that the barrier method for problem* (16) *is initialized at the point* $(\frac{1}{2}, \frac{1}{2}u, 1)$. *Then within*

$$O\left(\sqrt{\vartheta_{\tilde{f}}} \ln\left(\frac{\vartheta_{\tilde{f}}\mu_d}{\tau}\right)\right)$$

iterations, any iterate $(\hat{\theta}, \hat{x}, \hat{t})$ *of the algorithm will satisfy* $\hat{t} \le 0$, *and therefore* $x = \hat{x} - \frac{1}{2}\hat{t}u/(\hat{\theta} - \frac{1}{2}\hat{t})$ *is a solution of (FP$_d$).*

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Proof. First note that for any iterate $(\hat{\theta}, \hat{x}, \hat{t})$ of the algorithm, $\hat{\theta} > 0$ and $\hat{x} \in \operatorname{int} C_X$. Therefore, it is easy to check that when $\hat{t} \leq 0$, x is well-defined and is a solution of (FP_d).

It remains to verify the number of iterations needed to generate an iterate such that $\hat{t} \leq 0$. Let $\epsilon = \frac{1}{3\mu_d}$. Applying Theorem 12 and substituting the bounds of Proposition 13 into the complexity bound, we conclude that after at most

$$O\left(\sqrt{\vartheta_f} \ln\left(\frac{\vartheta_f}{\epsilon \operatorname{sym}(S \cap L, \omega')}\right)\right) = O\left(\sqrt{\vartheta_{\tilde{f}}} \ln\left(\frac{\vartheta_{\tilde{f}} \mu_d}{\tau}\right)\right)$$

iterations of the barrier method, any iterate $(\hat{\theta}, \hat{x}, \hat{t})$ will satisfy

$$\hat{t} \leq \epsilon(z_* - z^*) + z^* \leq \frac{1}{3\mu_d}(2 - (-1)) - \frac{1}{\mu_d} = 0,$$

from which the theorem follows. ∎

When the cone C_X is represented via a separation oracle, a solution of (FP_d) can be found using a version of the ellipsoid algorithm. (See, for example, [2] and [10].) Below is a generic theorem for analyzing the ellipsoid algorithm for finding a point ω in a convex set $S \subset \Re^k$ given by a separation oracle.

THEOREM 15. *Suppose that a convex set $S \subset \Re^k$ given by a separation oracle contains a Euclidean ball of radius r centered at some point $\hat{\omega}$, and that an upper bound R on the quantity $(\|\hat{\omega}\|_2 + r)$ is known. Then if the ellipsoid algorithm is initiated with a Euclidean ball of radius R centered at $\omega^0 = 0$, the algorithm will compute a point in S in at most*

$$\lceil 2k(k+1) \ln(R/r) \rceil$$

iterations, where each iteration must perform a feasibility cut on S.

The main problem with trying to apply Theorem 15 directly to (FP_d) is that one needs to know the upper bound R in advance. Because such an upper bound is generically unknown in advance for (FP_d), we approach solving (FP_d) by considering finding a point in the following set:

$$S \stackrel{\triangle}{=} \{(\theta, x) : \theta > 0,\ x \in C_X,\ b\theta - Ax = 0\}, \tag{17}$$

which is a convex set in the linear subspace $T \stackrel{\triangle}{=} \{(\theta, x) : b\theta - Ax = 0\}$ of dimension $k = n + 1 - m$. Observe that it is easy to construct a separation oracle for S in the linear subspace T, provided that one has a separation oracle for C_X. We will use the ellipsoid algorithm to find a point $(\hat{\theta}, \hat{x}) \in S$ (working in the linear subspace T), and we use the obvious transformation $x = \frac{\hat{x}}{\hat{\theta}}$ to transform the output of the algorithm into a solution of (FP_d).

PROPOSITION 16. *Let S be as in* (17). *Then there exists a point $(\hat{\theta}, \hat{x}) \in S$ and $\hat{r} > 0$ such that*

$$B((\hat{\theta}, \hat{x}), \hat{r}) \cap \{(\theta, x) : b\theta - Ax = 0\} \subset S, \quad \|(\hat{\theta}, \hat{x})\| + \hat{r} \leq 3, \quad \text{and } \hat{r} \geq \frac{\tau}{2\mu_d}.$$

Proof. Let $y = \frac{1}{2}b - \frac{1}{2}Au \in \mathcal{H}_d$. From the definition of μ_d we conclude that $-\frac{y}{\mu_d} \in \mathcal{H}_d$, whereby there exists $(\bar{\theta}, \bar{x})$ such that

$$|\bar{\theta}| + \|\bar{x}\| \leq 1, \quad \bar{\theta} \geq 0, \quad \bar{x} \in C_X, \quad b\bar{\theta} - A\bar{x} = -\frac{1}{\mu_d}\left(\frac{1}{2}b - \frac{1}{2}Au\right).$$

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Let $\hat{\omega} = (\hat{\theta}, \hat{x}) \stackrel{\triangle}{=} (\bar{\theta} + \frac{1}{2\mu_d}, \bar{x} + \frac{1}{2\mu_d}u)$ and $\hat{r} = \frac{\tau}{2\mu_d}$. Then $\hat{\omega} \in S$, $B(\hat{\omega}, \hat{r}) \cap \{(\theta, x) : b\theta - Ax = 0\} \subset S$ and

$$\|\hat{\omega}\|_2 + \hat{r} = \sqrt{(\bar{\theta} + \frac{1}{2\mu_d})^2 + \|\bar{x} + \frac{1}{2\mu_d}u\|^2} + \frac{\tau}{2\mu_d} \leq |\hat{\theta}| + \|\hat{x}\| + \frac{1}{2\mu_d} + \frac{\|u\|}{2\mu_d} + \frac{\tau}{2\mu_d} \leq 3. \qquad \square$$

The following theorem is an immediate consequence of Theorem 15 and Proposition 16.

THEOREM 17. *Suppose that the ellipsoid algorithm is applied in the linear subspace T to find a point in the set S, initialized with the Euclidean ball (in the space T) of radius $R = 3$ centered at $(\theta^0, x^0) = (0, 0)$. Then the ellipsoid algorithm will find a point in S (and hence, by transformation, a solution of* (FP$_d$)*) in at most*

$$\left\lceil 2(n - m + 1)(n - m + 2) \ln\left(\frac{6\mu_d}{\tau}\right) \right\rceil$$

iterations.

4. Symmetry measure and other measures of conditioning for (FP$_d$).

4.1. Symmetry measure and the condition number. In this subsection we establish a relationship between μ_d and $\mathcal{C}(d)$. As demonstrated in Theorem 18, if an instance of (FP$_d$) is "well-conditioned" in the sense that $\mathcal{C}(d)$ is small, then μ_d is also small. This relationship, however, is one-sided, since μ_d may carry no upper-bound information about $\mathcal{C}(d)$. In particular, in Remark 19 we exhibit a sequence of instances of (FP$_d$) with $\mathcal{C}(d)$ becoming arbitrarily large while μ_d remains fixed.

THEOREM 18. $\mu_d \leq \mathcal{C}(d)$.

Proof. If $\rho(d) = 0$, then $\mathcal{C}(d) = \infty$, and the statement of the theorem holds trivially. Suppose $\rho(d) > 0$. Since $B(0, \rho(d)) \subseteq \mathcal{H}_d$, we conclude that for any $v \in \mathcal{H}_d$, $-\frac{\rho(d)}{\|v\|}v \in \mathcal{H}_d$. Therefore

$$\frac{1}{\mu_d} = \text{sym}(\mathcal{H}_d, 0) \geq \inf_{v \in \mathcal{H}_d} \frac{\rho(d)}{\|v\|} \geq \frac{\rho(d)}{\|d\|} = \frac{1}{\mathcal{C}(d)},$$

proving the theorem. $\square$

REMARK 19. *μ_d may carry no upper-bound information about $\mathcal{C}(d)$.*

To see why this is true, consider the parametric family of problems (FP$_{d_\epsilon}$), where $d_\epsilon = (A_\epsilon, b)$:

$$b = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ and } A_\epsilon = \begin{bmatrix} 1 & 1 & -1 & -1 \\ \epsilon & -\epsilon & \epsilon & -\epsilon \end{bmatrix},$$

$C_X = \Re^4_+$ and $\|x\| \stackrel{\triangle}{=} \|x\|_1$ for $x \in X$, and $\|y\| = \|y\|_2$ for $y \in Y$. Consider the values of the parameter $\epsilon \in (0, 1]$. The set $\mathcal{H}_{d_\epsilon}$ is symmetric about 0, so $\mu_{d_\epsilon} = 1$ for any value of ϵ. On the other hand, $\rho(d_\epsilon) = \epsilon$ and $\|d_\epsilon\| = \sqrt{1 + \epsilon^2}$. Therefore,

$$\mathcal{C}(d_\epsilon) = \frac{\sqrt{1 + \epsilon^2}}{\epsilon} \geq \frac{1}{\epsilon},$$

and so $\mathcal{C}(d)$ can be arbitrarily large while μ_d remains constant. Furthermore, letting $\hat{x} = (1, 1, 1, 1)$ and $r = 1$, we see that γ in Theorem 11 has fixed value $\gamma = 4$ for any $\epsilon \in (0, 1]$.

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

So far, we have made no assumptions on the norm on the space Y; in fact, it can be easily seen that μ_d is invariant under changes in the norm on Y. (This is not true for $\mathcal{C}(d)$.) We conclude this section by providing another interpretation of the relationship between the measures μ_d and $\mathcal{C}(d)$. As Theorem 20 indicates, when the space Y is endowed with the appropriate norm, μ_d and $\mathcal{C}(d)$ are within a constant factor of each other. To see this, define

$$\mathcal{T}_d \stackrel{\triangle}{=} -\mathcal{H}_d \cap \mathcal{H}_d. \tag{18}$$

Then $\mathcal{T}_d$ is a convex set that is symmetric about 0, and $0 \in \operatorname{int} \mathcal{T}_d$ when $\mu_d < \infty$. Therefore we can define the norm $\|\cdot\|$ on Y to be the norm induced by considering $\mathcal{T}_d$ to be the unit ball, namely:

$$\|y\| \stackrel{\triangle}{=} \min\{\alpha : y \in \alpha \mathcal{T}_d\}. \tag{19}$$

THEOREM 20. *Suppose C_X is regular and $\mu_d < \infty$. If the norm on Y is given by (19), then $\rho(d) = 1$ and $\mathcal{C}(d) \leq \frac{\mu_d}{\delta}$, where δ is the norm approximation coefficient of the cone C_X.*

Proof. The characterization of $\rho(d)$ in (10) easily implies that $\rho(d) = 1$. It remains to establish the bound on the condition number $\mathcal{C}(d)$. We have

$$\mathcal{C}(d) = \frac{\|d\|}{\rho(d)} = \|d\| \leq \frac{1}{\delta} \max\{\|y\| : y \in \mathcal{H}_d\} \leq \frac{\mu_d}{\delta}.$$

The first inequality above follows from Corollary 5. To verify the second inequality above, suppose that $y \in \mathcal{H}_d$. Then $\frac{1}{\mu_d} y \in \mathcal{H}_d$ because $\mu_d \geq 1$, and $-\frac{1}{\mu_d} y \in \mathcal{H}_d$ by the definition of μ_d. Therefore, $\frac{1}{\mu_d} y \in \mathcal{T}_d$, and so $\|\frac{1}{\mu_d} y\| \leq 1$, which implies that $\max\{\|y\| : y \in \mathcal{H}_d\} \leq \mu_d$. This inequality is sufficient to prove the theorem; one can, however, show that $\max\{\|y\| : y \in \mathcal{H}_d\} = \mu_d$. ∎

4.2. Relationships between the symmetry measure and other measures of conditioning for linear programming. In the special case when $C_X = \Re^n_+$, the problem (FP$_d$) becomes a linear feasibility problem and can be written as follows:

$$(\text{FP}_d): \ Ax = b, \qquad x \geq 0, \tag{20}$$

where $x \in \Re^n$, $b \in \Re^m$, and $A \in \Re^{m \times n}$. We assume in this subsection that (FP$_d$) has a strictly positive solution x^0, i.e., $Ax^0 = b$ and $x^0 > 0$, that the norm on X is $\|x\| \stackrel{\triangle}{=} \|x\|_1$, and that the norm on Y is $\|y\| \stackrel{\triangle}{=} \|y\|_2$.

Complexity analysis of linear programming sometimes relies on the complexity measures $\sigma_{(\cdot)}$ and $\bar{\chi}(\cdot)$. These measures are quite specific to the special case of linear programming, as opposed to $\mathcal{C}(d)$ and μ_d, which apply to more general conic problems. In this subsection we state both previously known as well as new results relating all of these condition measures, which in total provide a complete characterization of the relationship between these four measures of conditioning.

For simplicity of notation, we define an "expanded" matrix $\tilde{A} \stackrel{\triangle}{=} [b; -A] \in \Re^{m \times (n+1)}$. Notice that $\|\tilde{A}\| \stackrel{\triangle}{=} \max\{\|b\theta - Ax\| : \|(\theta, x)\|_1 \leq 1\} = \|d\|$.

We first review a slight variant on $\sigma_{(\cdot)}$ called σ_d, which was introduced and used in the complexity analysis of an interior-point algorithm for solving (FP$_d$) by Vavasis and Ye [32]:

$$\sigma_d \stackrel{\triangle}{=} \min_{j=1,\ldots,n+1} \max_w \{e_j^t w : \tilde{A} w = 0, \ e^t w = 1, \ w \geq 0\},$$

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where $e_j,\ j = 1, \ldots, n+1$, denotes the jth unit vector and $e \in \Re^{n+1}$ is the vector of all ones. Note that while the above does not coincide with the usual definition of σ, it does under our assumption that (FP_d) has a strictly positive solution.

We also review a slight variant on $\bar{\chi}_{(\cdot)}$ called $\bar{\chi}_d$, which has been used by Vavasis and Ye [33, 34] and Megiddo, Mizuno, and Tsuchiya [16] in the complexity analysis of another interior-point algorithm:

$$\bar{\chi}_d \stackrel{\triangle}{=} \sup\{\|\tilde{A}^t(\tilde{A}D\tilde{A}^t)^{-1}\tilde{A}D\| : D \in S_{++}^{(n+1)\times(n+1)},\ D \text{ diagonal}\}.$$

An alternative characterization of $\bar{\chi}_d$ is

$$\bar{\chi}_d = \max\{\|B^{-1}\tilde{A}\| : B \in \mathcal{B}(\tilde{A})\}, \tag{21}$$

where $\mathcal{B}(\tilde{A})$ is the set of all bases (i.e., $m \times m$ nonsingular submatrices) of $\tilde{A}$. (See [29] for the proof of the equivalence of these characterizations.)

It has been established by Vavasis and Ye [32] that σ_d and $\bar{\chi}_d$ are related by the inequality

$$\sigma_d \geq \frac{1}{\bar{\chi}_d + 1}.$$

On the other hand, Tunçel in [31] established that, in general, σ_d may carry no upper-bound information about $\bar{\chi}_d$. Specifically, he provided a family of data instances d_ϵ such that for any $\epsilon > 0$, $\sigma_{d_\epsilon} = \frac{1}{2}$, but $\bar{\chi}_{d_\epsilon} \geq \frac{1}{\epsilon}$, and so $\bar{\chi}_{d_\epsilon}$ can be arbitrarily large.

Theorem 18 and Remark 19 established a relationship between μ_d and $\mathcal{C}(d)$. Below we establish relationships between the other pairs of measures μ_d, $\mathcal{C}(d)$, $\bar{\chi}_d$, and σ_d, or provide examples that show that no such relationship exists, in the spirit of [31].

REMARK 21. *$\mathcal{C}(d)$ and $\bar{\chi}_d$ may carry no upper-bound or lower-bound information about each other.*

To establish the above result, we provide two parametric families of matrices $\tilde{A}_\epsilon$ such that by varying the value of the parameter $\epsilon > 0$ we can make one of the above measures arbitrarily bad while keeping the other measure constant or bounded.

First consider the family of matrices $\tilde{A}_\epsilon = \left[\begin{smallmatrix} \epsilon & 0 & -\epsilon \\ 0 & 1 & -1 \end{smallmatrix}\right]$. For $\epsilon > 0$ and sufficiently small, $\rho(d_\epsilon) = \frac{\epsilon}{\sqrt{\epsilon^2+4}}$. Furthermore, $\|d_\epsilon\| = \sqrt{1+\epsilon^2}$, and so

$$\mathcal{C}(d_\epsilon) = \sqrt{\frac{\epsilon^2+1}{\epsilon^2+4}} \cdot \frac{1}{\epsilon} \to +\infty \text{ as } \epsilon \to 0.$$

On the other hand, it is easy to establish using (21) that $\bar{\chi}(d_\epsilon) = \sqrt{2}$ for any $\epsilon > 0$.

To establish the second claim of the remark, consider the family $\tilde{A}_\epsilon = [1\ \epsilon\ -1]$ with $0 < \epsilon < 1$. We have $\|d_\epsilon\| = 1$, $\rho(d_\epsilon) = 1$, and so $\mathcal{C}(d_\epsilon) = 1$ for any ϵ as above. On the other hand it is easy to establish using (21) that for any $\epsilon \in (0, 1)$, $\bar{\chi}_{d_\epsilon} = \frac{1}{\epsilon} \to +\infty$ as $\epsilon \to 0$.

PROPOSITION 22. *Suppose the system (FP_d) of (20) has a positive solution. Then $\sigma_d = \frac{1}{1+\mu_d}$.*

Proof. Observe that we can redefine σ_d as follows:

$$\sigma_d = \min_{j=1,\ldots,n+1} \sigma_j, \text{ where } \sigma_j \stackrel{\triangle}{=} \max\{e_j^t w : \tilde{A}w = 0,\ e^t w = 1,\ w \geq 0\}.$$

From Lemma 7, there exists an extreme point $\bar{w}$ of

$$\mathcal{H}_d = \{b\theta - Ax : (\theta, x) \geq 0,\ \|(\theta, x)\|_1 \leq 1\} = \{\tilde{A}w : w \geq 0,\ e^t w \leq 1\}$$

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

such that $\frac{1}{\mu_d} = \text{sym}_{\bar{w}}(\mathcal{H}_d, 0) = \sup\{t : -t\bar{w} \in \mathcal{H}_d\}$. Since the set of extreme points of the set $\mathcal{H}_d$ is contained in the set $\{\tilde{A}_1, \ldots, \tilde{A}_{n+1}\}$, where $\tilde{A}_j \in \Re^m$ is the jth column of the matrix $\tilde{A}$, we can characterize μ_d as

$$\frac{1}{\mu_d} = \min_{j=1,\ldots,n+1} \frac{1}{\mu_j}, \text{ where } \frac{1}{\mu_j} \stackrel{\triangle}{=} \sup\{t : -t\tilde{A}_j \in \mathcal{H}_d\}.$$

We will now show that for any j

$$\sigma_j = \frac{1}{1+\mu_j}. \tag{22}$$

Without loss of generality we can consider $j = 1$ and the corresponding column $\tilde{A}_1$ of $\tilde{A}$. If $\tilde{A}_1 = 0$, then $\sigma_1 = 1$, $\frac{1}{\mu_1} = +\infty$, and (22) holds as a limiting relationship.

Suppose that $A_1 \neq 0$, and therefore $\mu_1 > 0$ and $\sigma_1 < 1$. By definition of μ_1, $-\frac{1}{\mu_1}\tilde{A}_1 \in \mathcal{H}_d$, i.e., there exists a point $p \geq 0,\ e^t p = 1$ such that $-\frac{1}{\mu_1}\tilde{A}_1 = \tilde{A}p$. Define $w \stackrel{\triangle}{=} \frac{\mu_1 p + e_1}{1+\mu_1}$. Then $w \geq 0$, $e^t w = 1$, and $\tilde{A}w = 0$. Therefore, $\sigma_1 \geq w_1 \geq \frac{1}{1+\mu_1}$.

Suppose now that w is a solution of the linear program defining σ_1. Then $w_1 = \sigma_1$. Let $p = \frac{1}{1-\sigma_1}(w - \sigma_1 e_1)$. Then $p \geq 0$, $e^t p = 1$, and $\tilde{A}p = \frac{-\tilde{A}_1\sigma_1}{1-\sigma_1}$. Therefore, $\frac{1}{\mu_1} \geq \frac{\sigma_1}{1-\sigma_1}$, and so $\sigma_1 \leq \frac{1}{\mu_1+1}$. Combining this with the bound in the previous paragraph, we conclude that $\sigma_1 = \frac{1}{\mu_1+1}$, and by similar argument, $\sigma_j = \frac{1}{\mu_j+1},\ j = 1, \ldots, n+1$.

Suppose now that $\sigma_d = \sigma_j$ for some j. That means that $\sigma_j \leq \sigma_i$ for any index i, or, equivalently, $\frac{1}{\mu_j+1} \leq \frac{1}{\mu_i+1}$ and hence $\mu_j \geq \mu_i$ for any index i. Therefore, $\mu_d = \mu_j$ and hence $\sigma_d = \frac{1}{1+\mu_d}$. ∎

The following two remarks, which are easy consequences of Proposition 22, establish the remaining relationships between the four measures of conditioning.

REMARK 23. $\mu_d \leq \bar{\chi}_d$. *However, μ_d may carry no upper-bound information about $\bar{\chi}_d$.*

REMARK 24. $\sigma_d \geq \frac{1}{\mathcal{C}(d)+1}$. *However, σ_d may carry no upper-bound information about $\mathcal{C}(d)$.*

In light of Proposition 22, μ_d can in fact be viewed as a generalization of the Vavasis–Ye measure σ_d to a general conic linear system. Related to this, Ho in [11] provides an argument indicating that extending $\bar{\chi}_d$ to general conic systems is not possible.

5. Preconditioners for conic linear systems. In this section we present a characterization of all data instances $\tilde{d}$ equivalent to d (in the sense that $X_d = X_{\tilde{d}}$), by introducing the concept of a *preconditioner*, and we provide an upper bound on the condition number $\mathcal{C}(\tilde{d})$ of the "best" equivalent data instance $\tilde{d}$. We conclude by analyzing the complexity of computing an equivalent data instance whose condition number is within a known factor of this bound, by constructing an algorithm for computing such an instance and analyzing its complexity.

Consider the data instance $d = (A, b) \in \mathcal{D}$ defining the system (FP_d). Let $B \in \Re^{m\times m}$ be a given nonsingular matrix and consider the data instance $Bd \stackrel{\triangle}{=} B \cdot d = (BA, Bb)$, which gives rise to the system

$$(\text{FP}_{Bd}):\ BAx = Bb, \qquad x \in C_X. \tag{23}$$

The systems (FP_d) and (FP_{Bd}) are equivalent; for this reason we say that the data instances d and Bd are equivalent as well. We can view the systems (FP_d) and (FP_{Bd})

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

as different formulations of the same feasibility problem (FP): find $x \in \mathcal{A} \cap C_X$, where $\mathcal{A}$ is the affine subspace

$$\mathcal{A} \stackrel{\triangle}{=} \{x : Ax = b\} = \{x : BAx = Bb\}. \tag{24}$$

However the condition numbers of the two systems, $\mathcal{C}(d)$ and $\mathcal{C}(Bd)$, are, in general, not equal.

On the other hand, consider the symmetry measures of the two systems, μ_d and μ_{Bd}. Observe that

$$\mathcal{H}_{Bd} \stackrel{\triangle}{=} \{Bb\theta - BAx : \theta \geq 0,\ x \in C_X,\ |\theta| + \|x\| \leq 1\} = B(\mathcal{H}_d);$$

i.e., the set $\mathcal{H}_{Bd}$ is the image of the set $\mathcal{H}_d$ under the linear transformation defined by B. Therefore, $\mathrm{sym}(\mathcal{H}_{Bd}, 0) = \mathrm{sym}(\mathcal{H}_d, 0)$, and $\mu_d = \mu_{Bd}$, since the symmetry of a set is preserved under nonsingular linear transformation, and so we can think of μ_d as depending on the affine space $\mathcal{A}$ defined in (24) but not on the specific data d. To highlight the independence of μ_d of the particular data d, we sometimes write $\mu_{\mathcal{A}}$ in place of μ_d. We record this formally in the following proposition.

PROPOSITION 25. *Let $d = (A, b) \in \mathcal{D}$, let $B \in \Re^{m \times m}$ be a nonsingular matrix, and define $\mathcal{A} \stackrel{\triangle}{=} \{x : Ax = b\}$. Then $\mu_d = \mu_{Bd} = \mu_{\mathcal{A}}$.*

We leave to the reader the proof of the next proposition.

PROPOSITION 26. *Suppose C_X is a regular cone. Let $d = (A, b) \in \mathcal{D}$ and $\tilde{d} = (\tilde{A}, \tilde{b}) \in \mathcal{D}$ be such that $X_d = X_{\tilde{d}}$. If $\mathcal{C}(d) < \infty$, then there exists a nonsingular matrix $B \in \Re^{m \times m}$ such that $\tilde{d} = Bd$.*

Suppose that a feasibility problem can be represented via two equivalent data instances d and $\tilde{d}$, and suppose that $\mathcal{C}(d) \ll \mathcal{C}(\tilde{d})$. If one were to predict, for example, the performance of the interior-point algorithm from section 3 for solving (FP_d) by analyzing its complexity in terms of the condition number, the bounds would be overly conservative if the problem were described by the data instance $\tilde{d}$. However, our analysis of the performance of the algorithm in terms of $\mu_{\mathcal{A}}$ yields a bound independent of the data instance used.

On the other hand, as detailed in the introduction, the condition number $\mathcal{C}(d)$ is a crucial parameter for analyzing properties of (FP_d) that depend on the representation of the problem ($\mathrm{FP}_{(\cdot)}$) by a specific data instance d, such as sensitivity of the feasible region to data perturbations, numerical properties of computations in algorithms for solving (FP_d), etc. Therefore, it might be beneficial to *precondition* the system (FP_d), i.e., to find another data instance $\tilde{d} = Bd$ for which $\mathcal{C}(\tilde{d}) < \mathcal{C}(d)$, and work with the corresponding system ($\mathrm{FP}_{\tilde{d}}$), which is better-behaved. In this light, we can view the matrix B above as a *preconditioner* for the system (FP_d), yielding the preconditioned system ($\mathrm{FP}_{\tilde{d}}$) with $\tilde{d} = Bd$, and Proposition 26 implies that any data instance $\tilde{d}$ for which $X_{\tilde{d}} = X_d$ can be obtained by preconditioning d with an appropriate B.

In the remainder of this section, we characterize a so-called best preconditioner, which is a preconditioner that gives rise to a condition number that is within a constant factor of the best possible, and we construct and analyze an algorithm for computing a preconditioner that yields a condition number that is within a known factor of this bound. For the remainder of this section, we assume that the space Y is the m-dimensional Euclidean space $\Re^m$ with Euclidean norm $\|y\| = \|y\|_2 = \sqrt{y^t y}$. We assume that the cone C_X is a regular cone with width τ and norm approximation coefficient δ. We also assume that $m \geq 2$. (In fact, the case $m = 1$ is trivial since in this case $\mu_{\mathcal{A}}$ and $\mathcal{C}(d)$ are within a factor of δ of each other, and thus the issue of preconditioning is essentially irrelevant.)

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

5.1. Best preconditioners and α-roundings. The main result of this subsection, Theorem 30, demonstrates the existence of a preconditioner $\bar{B}$ such that $\mathcal{C}(\bar{B}d)$ is within the factor $\frac{\sqrt{m}}{\delta}$ of $\mu_{\mathcal{A}}$. We begin by developing the tools to prove this result.

For any matrix $Q \in S^{m\times m}_{++}$ we define E_Q to be the ellipsoid $E_Q \stackrel{\triangle}{=} \{y \in Y : y^tQ^{-1}y \leq 1\}$.

DEFINITION 27. *Let $S \subset Y$ be a bounded set with a nonempty convex interior. For $\alpha \in (0,1]$, an ellipsoid E_Q is called an* α-rounding *of S if*

$$\alpha E_Q \subseteq S \subseteq E_Q.$$

We refer to the parameter α as the *tightness* of the rounding E_Q.

If the set S above satisfies $S = -S$ (i.e., is symmetric about 0), then S possesses a $\frac{1}{\sqrt{m}}$-rounding, i.e., there exists an ellipsoid E_Q such that $\frac{1}{\sqrt{m}}E_Q \subseteq S \subseteq E_Q$ (see John [12]). In particular, the ellipsoid of minimum volume containing S (often referred to as the Löwner–John ellipsoid of S) is a $\frac{1}{\sqrt{m}}$-rounding of S.

The following lemma allows us to interpret preconditioning of the system (FP_d) by B as constructing a $\frac{1}{\mathcal{C}(Bd)}$-rounding of the set $\mathcal{H}_d$.

LEMMA 28. *Let $B \in \Re^{m\times m}$ be a (nonsingular) preconditioner for the system (FP_d). Let $Q = \|Bd\|^2(B^tB)^{-1}$. Then*

$$\frac{1}{\mathcal{C}(Bd)}E_Q \subseteq \mathcal{H}_d \subseteq E_Q.$$

Proof. First, observe that $Q \in S^{m\times m}_{++}$, since B is nonsingular. To prove the first inclusion, let $h \in \frac{1}{\mathcal{C}(Bd)}E_Q$, i.e., $h^tQ^{-1}h \leq \frac{1}{\mathcal{C}(Bd)^2}$. Using the definition of Q, we have $h^t(B^tB)h \leq \frac{\|Bd\|^2}{\mathcal{C}(Bd)^2} = \rho(Bd)^2$, that is, $\|Bh\| \leq \rho(Bd)$. This implies $Bh \in \mathcal{H}_{Bd}$, and hence, $h \in \mathcal{H}_d$.

Next, suppose $h \in \mathcal{H}_d$, and so $Bh \in \mathcal{H}_{Bd}$. Then $\|Bh\| \leq \|Bd\|$, and therefore

$$h^tQ^{-1}h = h^t\left(\|Bd\|^2(B^tB)^{-1}\right)^{-1}h = \frac{\|Bh\|^2}{\|Bd\|^2} \leq 1,$$

i.e., $h \in E_Q$. ∎

LEMMA 29. *Let $Q \in S^{m\times m}_{++}$ be such that E_Q is an α-rounding of the set $\mathcal{T}_d$ of (18). Let $B = Q^{-\frac{1}{2}}$. Then B is a preconditioner for the system (FP_d) such that*

$$\mathcal{C}(Bd) \leq \frac{\mu_{\mathcal{A}}}{\alpha\delta} \leq \frac{2\mu_{\mathcal{A}}}{\alpha\tau}.$$

Proof. We establish the result by providing bounds on the distance to infeasibility $\rho(Bd)$ and the size of the data $\|Bd\|$ of the system (FP_{Bd}). First, we will show that $\rho(Bd) \geq \alpha$. Let $v \in Y$ satisfy $\|v\| \leq \alpha$. Then

$$(B^{-1}v)^tQ^{-1}B^{-1}v = (B^{-1}v)^t(B\cdot B)(B^{-1}v) = \|v\|^2 \leq \alpha^2,$$

and therefore $B^{-1}v \in \alpha E_Q \subseteq \mathcal{T}_d \subseteq \mathcal{H}_d$. Thus, $v \in \mathcal{H}_{Bd}$, and so $\rho(Bd) \geq \alpha$.

Next, recall from Corollary 5 that $\|Bd\| \leq \frac{1}{\delta}\max\{\|v\| : v \in \mathcal{H}_{Bd}\}$. Let $v \in \mathcal{H}_{Bd}$. Then $y = B^{-1}v \in \mathcal{H}_d$, and $\frac{1}{\mu_{\mathcal{A}}}y \in -\mathcal{H}_d \cap \mathcal{H}_d = \mathcal{T}_d \subseteq E_Q$. Hence $\|v\|^2 = y^tB^tBy = y^tQ^{-1}y \leq \mu^2_{\mathcal{A}}$, whereby $\|Bd\| \leq \frac{\mu_{\mathcal{A}}}{\delta}$.

Combining the obtained results, $\mathcal{C}(Bd) = \frac{\|Bd\|}{\rho(Bd)} \leq \frac{\mu_{\mathcal{A}}}{\delta\alpha} \leq \frac{2\mu_{\mathcal{A}}}{\tau\alpha}$. ∎

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

THEOREM 30. *Suppose that* (FP_d) *is feasible and* $\mathcal{C}(d) < +\infty$. *Then there exists a preconditioner* $\bar{B}$ *such that*

$$\mu_{\mathcal{A}} \leq \mathcal{C}(\bar{B}d) \leq \frac{\sqrt{m}}{\delta} \cdot \mu_{\mathcal{A}}. \tag{25}$$

Proof. By definition, $\mathcal{T}_d$ is a bounded convex set symmetric about 0. Since $\mathcal{C}(d) < \infty$, $\mathcal{T}_d$ has a nonempty interior. Therefore, there exists $Q \in S^{m\times m}_{++}$ such that E_Q is a $\frac{1}{\sqrt{m}}$-rounding of $\mathcal{T}_d$. Applying Lemma 29 with $\alpha = \frac{1}{\sqrt{m}}$, we obtain (25). ∎

REMARK 31. *In general, the upper bound in* (25) *is tight for any* m.

We verify this remark by example. Consider the system (FP_d) with $n = 2m$, $C_X = \Re^{2m}_+$, $\|x\| = \|x\|_1$ (so that $\delta = 1$), and the data $d = (A, b)$ as follows:

$$b = 0 \text{ and } A = [e_1,\ -e_1, \ldots, e_m,\ -e_m],$$

where e_i is the ith unit vector in $\Re^m$. Then $\mathcal{H}_d = \mathcal{T}_d = \mathrm{conv}\{\pm e_i,\ i = 1, \ldots, m\}$, and it can be easily verified that $\mu_{\mathcal{A}} = 1$, $\rho(d) = \frac{1}{\sqrt{m}}$, and $\|d\| = 1$, and therefore $\mathcal{C}(d) = \sqrt{m}$. Suppose B is an arbitrary preconditioner. Using Lemma 28, we can construct a $\frac{1}{\mathcal{C}(Bd)}$-rounding of the set $\mathcal{T}_d$. However, it is impossible to construct an α-rounding of the set $\mathrm{conv}\{\pm e_i,\ i = 1, \ldots, m\}$ with $\alpha > \frac{1}{\sqrt{m}}$; see, for example, [10]. Therefore, $\mathcal{C}(Bd) \geq \sqrt{m}$ for any preconditioner B.

5.2. On the complexity of computing a good preconditioner. We present an algorithm that computes a preconditioner $\tilde{B}$ for which

$$\mathcal{C}(\tilde{B}d) \leq \frac{4m\mu_{\mathcal{A}}}{\delta}. \tag{26}$$

Recall that in Lemma 29 it was shown that a tight rounding of the set $\mathcal{T}_d$ gives rise to a good preconditioner for the system (FP_d). In Theorem 30 we relied on the existence of a $\frac{1}{\sqrt{m}}$-rounding of the set $\mathcal{T}_d$ to establish the existence of a preconditioner $\bar{B}$ such that $\mu_{\mathcal{A}} \leq \mathcal{C}(\bar{B}d) \leq \frac{\sqrt{m}}{\delta}\,\mu_{\mathcal{A}}$, i.e., $\mathcal{C}(\bar{B}d)$ is within the factor of $\frac{\sqrt{m}}{\delta}$ of the lower bound. In general, we are not able to efficiently compute a $\frac{1}{\sqrt{m}}$-rounding of the set $\mathcal{T}_d$. (See [10] for commentary on the difficulty of computing an approximate $\frac{1}{\sqrt{m}}$-rounding of a set $\mathcal{S}$ that does not have an efficient half-space representation.) However, the algorithm presented in this subsection will compute an ellipsoid which is a $\frac{1}{4m}$-rounding of $\mathcal{T}_d$ (also called a weak Löwner–John ellipsoid for $\mathcal{T}_d$). In particular, the algorithm of this subsection will compute a matrix $\tilde{Q} \in S^{m\times m}_{++}$ such that

$$\frac{1}{4m} E_{\tilde{Q}} \subseteq \mathcal{T}_d \subseteq E_{\tilde{Q}}, \tag{27}$$

which can be used to obtain a preconditioner $\tilde{B}$ satisfying (26) via Lemma 29. We denote this algorithm as *Algorithm WLJ* for "Weak Löwner–John."

In order to be able to efficiently implement the algorithm described in this section, we restrict the norm $\|x\|$ for $x \in X$ to the Euclidean norm $\|x\| = \|x\|_2$ (as well as maintain the assumption that $\|y\| = \|y\|_2$ for $y \in Y$). We further assume that the interior of the cone C^*_X is the domain of a self-concordant barrier $f^*(\cdot)$ with complexity parameter ϑ^*. The width of the cone C^*_X is denoted by τ^*. We assume that we know

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

and are given the vector $u_* \in C_X^*$ for which $\|u_*\| = 1$ and $B(u_*, \tau^*) \subset C_X^*$ as in (7). Finally, we assume that an upper bound $\bar{d}$ on $\|d\|$ is known and given or is easily computable. One could, for example, take

$$\bar{d} = \sqrt{n} \max\{\|b\|_2, \|A_1\|_2, \ldots, \|A_m\|_2\},$$

where A_j is the jth column of the matrix A. Then $\bar{d}$ approximates $\|d\|$ within the factor of $\sqrt{n}$, i.e., $\frac{1}{\sqrt{n}}\bar{d} \leq \|d\| \leq \bar{d}$.

The algorithm WLJ is a version of the parallel-cut ellipsoid algorithm; see [10]. A generic iteration of this algorithm can be described as follows. At the start of each iteration, we have a matrix $Q \in S_{++}^{m \times m}$ such that $\mathcal{T}_d \subseteq E_Q$. We compute the eigenvalue decomposition of the matrix Q. In particular, we compute the eigenvalues $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m$ of the matrix Q and their corresponding (orthonormal) eigenvectors $a_1, \ldots, a_m$. Then the axes of the ellipsoid E_Q are $v_i = \sqrt{\lambda_i} a_i,\ i = 1, \ldots, m$. We denote $V \stackrel{\triangle}{=} [v_1, \ldots, v_m] \in \Re^{m \times m}$. It is elementary to verify that $Q = VV^t$.

The algorithm then checks if the scaled axes $\pm\frac{1}{4\sqrt{m}} v_i$ are elements of $\mathcal{T}_d$ for $i = 1, \ldots, m$. If so, the algorithm correctly asserts that

$$\frac{1}{4m} E_Q = \frac{1}{\sqrt{m}} \cdot \frac{1}{4\sqrt{m}} E_Q \subset \text{conv}\left\{ \pm \frac{1}{4\sqrt{m}} v_i,\ i = 1, \ldots, m \right\} \subseteq \mathcal{T}_d \subseteq E_Q, \tag{28}$$

and the algorithm terminates. On the other hand, if the algorithm finds an axis $v = \pm v_j$ for some j for which $\frac{1}{4\sqrt{m}} v_j \notin \mathcal{T}_d$, then it finds a parallel cut separating the two points $\pm\frac{1}{2\sqrt{m}} v_j$ from the set $\mathcal{T}_d$, i.e., it produces a vector s such that

$$s^t v_j = 1 \text{ for some } v_j, \text{ and } \mathcal{T}_d \subseteq \left[E_Q \cap \left\{ y : -\frac{1}{2\sqrt{m}} \leq s^t y \leq \frac{1}{2\sqrt{m}} \right\} \right]. \tag{29}$$

This cut is then used to find an ellipsoid $E_{\hat{Q}}$ which satisfies

$$E_{\hat{Q}} \supset \left[E_Q \cap \left\{ y : -\frac{1}{2\sqrt{m}} \leq s^t y \leq \frac{1}{2\sqrt{m}} \right\} \right] \supseteq \mathcal{T}_d,$$

and for which

$$\frac{\text{vol}(E_{\hat{Q}})}{\text{vol}(E_Q)} \leq \frac{1}{2} e^{\frac{3}{8}}. \tag{30}$$

The formula for $\hat{Q}$ is

$$\hat{Q} = \frac{m}{m-1} \left(1 - \frac{1}{4m\xi} \right) \left(Q - \frac{m(4\xi - 1)}{4m\xi - 1} \cdot \frac{Qss^tQ}{\xi} \right), \tag{31}$$

where

$$\xi = s^t Q s = \|V^t s\|^2 \geq s^t v_j = 1; \tag{32}$$

see formula (3.1.20) of [10], for example.

In order to implement this algorithm, it is necessary to be able to check whether the rescaled axes $\pm\frac{1}{4\sqrt{m}} v_i$ are elements of $\mathcal{T}_d$, for $i = 1, \ldots, m$, and if not, it is then necessary to produce the vector s describing the parallel cut of (29). These two tasks are accomplished in a subroutine called *Weak Check*, which is outlined as follows, and for which a more complete description is furnished in the appendix.

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Subroutine Weak Check.
Given the axes $v_1, \dots, v_m$ of an ellipsoid $E_Q \supseteq \mathcal{T}_d$, either
(i) verify that $\pm\frac{1}{4\sqrt{m}} v_i \in \mathcal{T}_d$ for all $i = 1, \dots, m$, or
(ii) find a vector s such that

$$s^t v_j = 1 \text{ for some } v_j, \text{ and } \mathcal{T}_d \subseteq \left[E_Q \cap \left\{ y : -\frac{1}{2\sqrt{m}} \le s^t y \le \frac{1}{2\sqrt{m}} \right\} \right]. \tag{33}$$

The formal description of algorithm WLJ is as follows.

ALGORITHM WLJ (Weak Löwner–John).

- *Initialization:* The algorithm is initialized with the matrix $Q^0 = \bar{d}^2 I$.
- Iteration $k \ge 1$.

  Step 1 Let $Q = Q^k$. Compute the eigenvalues $0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_m$ of Q and the corresponding (orthonormal) eigenvectors $a_1, \dots, a_m$. Define the axes of E_Q by $v_i = \sqrt{\lambda_i} a_i,\ i = 1, \dots, m$.

  Step 2 Call subroutine *Weak Check* with the input $(v_1, \dots, v_m)$. If the subroutine verifies that $\pm\frac{1}{4\sqrt{m}} v_i \in \mathcal{T}_d,\ i = 1, \dots, m$, then return $\tilde{B} = Q^{-\frac{1}{2}}$ and terminate. Otherwise, subroutine *Weak Check* returns a vector s. Define $\hat{Q}$ by (31).

  Step 3 Let $Q^{k+1} = \hat{Q}$, $k \leftarrow k + 1$; go to Step 1.

To complete the description of Algorithm WLJ, one must specify the details of subroutine *Weak Check*. The purpose of subroutine *Weak Check* is to verify whether the rescaled axes $\pm\frac{1}{4\sqrt{m}} v_i$, $i = 1, \dots, m$, are contained in $\mathcal{T}_d$, or to produce a parallel cut otherwise. This is accomplished by examining the following $2m$ optimization problems (P_{ϕ_v}), where $v = \pm v_i$, $i = 1, \dots, m$:

$$(\mathrm{P}_{\phi_v}) \quad \phi_v = \max_{\phi} \quad \phi \quad \text{s.t.} \quad \phi v \in \mathcal{H}_d. \tag{34}$$

It is easy to verify that $\pm\frac{1}{4\sqrt{m}} v_i \in \mathcal{T}_d$ for all $i = 1, \dots, m$ precisely when

$$\phi_Q \stackrel{\triangle}{=} \min_{\pm v_i} \phi_v \ge \frac{1}{4\sqrt{m}}. \tag{35}$$

(Here $\min_{\pm v_i} \phi_v$ stands for $\min\{\phi_{v_1}, -\phi_{v_1}, \dots, \phi_{v_m}, -\phi_{v_m}\}$ in order to shorten the notation.) We will therefore implement subroutine *Weak Check* by means of approximately solving the $2m$ optimization problems (34) and checking whether condition (35) is satisfied. To solve the optimization problems (34) for every value of $v = \pm v_i,\ i = 1, \dots, m$, we will apply the barrier method of [27] to a version of the Lagrangian dual of (P_{ϕ_v}). The formal description of this implementation is presented in the appendix, where the following complexity bound is proved.

LEMMA 32. *Subroutine* Weak Check *will terminate in at most*

$$O\left(m\sqrt{\vartheta^*} \ln\left(\frac{m\vartheta^*}{\tau^*} \cdot \frac{\bar{d}}{\sqrt{\lambda_1}} \cdot \sqrt{\frac{\lambda_m}{\lambda_1}} \right) \right) \tag{36}$$

iterations of the barrier method. Upon termination, it will either correctly verify that $\pm\frac{1}{4\sqrt{m}} v_i \in \mathcal{T}_d$ *for all* $i = 1, \dots, m$, *or will return a vector* s *such that*

$$s^t v_j = 1 \text{ for some } v_j, \text{ and } \mathcal{T}_d \subseteq \left[E_Q \cap \left\{ y : -\frac{1}{2\sqrt{m}} \le s^t y \le \frac{1}{2\sqrt{m}} \right\} \right]. \tag{37}$$

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Note that the skewness of the ellipsoid E_Q, which is square root of the ratio of the largest to the smallest eigenvalue of Q, comes to play in the complexity bound of subroutine *Weak Check*.

We now proceed to analyze the complexity of Algorithm WLJ. We first prove the volume reduction bound of (30) in Lemma 33. We then prove the main complexity of Algorithm WLJ in Theorem 34.

LEMMA 33. *Let Q be an iterate of Algorithm WLJ, and let $\hat{Q}$ be defined by* (31). *Then*

$$\frac{\text{vol}(E_{\hat{Q}})}{\text{vol}(E_Q)} \leq \frac{1}{2}e^{\frac{3}{8}}.$$

Proof. Let $R \in \Re^{m\times m}$ be an orthonormal matrix such that $RQ^{\frac{1}{2}}s = \|Q^{\frac{1}{2}}s\|e_1 = \sqrt{\xi}e_1$. Then $\hat{Q}$ can be expressed as

$$\hat{Q} = \frac{m}{m-1}\left(1 - \frac{1}{4m\xi}\right)Q^{\frac{1}{2}}R^t\left(I - \frac{m(4\xi-1)}{4m\xi-1}e_1e_1^t\right)RQ^{\frac{1}{2}}. \tag{38}$$

Therefore,

$$\begin{aligned}\det(\hat{Q}) &= \det\left(\frac{m}{m-1}\left(1 - \frac{1}{4m\xi}\right)Q^{\frac{1}{2}}R^t\left(I - \frac{m(4\xi-1)}{4m\xi-1}e_1e_1^t\right)RQ^{\frac{1}{2}}\right)\\ &= \left(\frac{m}{m-1}\left(1 - \frac{1}{4m\xi}\right)\right)^m\left(1 - \frac{m(4\xi-1)}{4m\xi-1}\right)\det(Q).\end{aligned}$$

We conclude that

$$\begin{aligned}\frac{\det(\hat{Q})}{\det(Q)} &= \left(\frac{m}{m-1}\left(1 - \frac{1}{4m\xi}\right)\right)^m\left(1 - \frac{m(4\xi-1)}{4m\xi-1}\right)\\ &= \frac{m^m(4m\xi-1)^{m-1}}{(m-1)^{m-1}(4m\xi)^m} = \frac{1}{4\xi}\left(\frac{4m\xi-1}{4m\xi-4\xi}\right)^{m-1}\\ &= \frac{1}{4\xi}\left(1 + \frac{4\xi-1}{4\xi(m-1)}\right)^{m-1} \leq \frac{1}{4\xi}e^{1-\frac{1}{4\xi}} \leq \frac{1}{4}e^{\frac{3}{4}}.\end{aligned}$$

The last inequality follows since the function te^{1-t} is an increasing function for $t \in [0,1]$, and from (32) we have $0 < \frac{1}{4\xi} \leq \frac{1}{4}$. Finally,

$$\frac{\text{vol}(E_{\hat{Q}})}{\text{vol}(E_Q)} = \frac{\sqrt{\det(\hat{Q})}}{\sqrt{\det(Q)}} \leq \frac{1}{2}e^{\frac{3}{8}}. \qquad \square$$

THEOREM 34. *Suppose $\mathcal{C}(d) < \infty$. Then Algorithm WLJ will terminate in at most*

$$O\left(m^2\sqrt{\vartheta^*}\ln^2\left(\frac{\bar{d}}{\rho(d)}\right)\ln\left(\frac{m\vartheta^*}{\tau^*}\right)\right) \tag{39}$$

iterations of the barrier method. It will return upon termination a preconditioner $\tilde{B}$ such that

$$\mu_{\mathcal{A}} \leq \mathcal{C}(\tilde{B}d) \leq \frac{4m\mu_{\mathcal{A}}}{\delta}.$$

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Proof. First observe that the matrix $Q^0 = \bar{d}^2 I$ used to initialize the algorithm is a valid iterate, since for any point $y \in \mathcal{T}_d$, $\|y\| \le \|d\| \le \bar{d}$, and so $\mathcal{T}_d \subseteq E_{Q^0}$.

Suppose Algorithm WLJ has performed k iterations, and let Q^k be the current iterate. Since $\mathcal{T}_d \subseteq E_{Q^k}$, we conclude that

$$\operatorname{vol}(\mathcal{T}_d) \le \operatorname{vol}(E_{Q^k}) \le \left(\frac{1}{2}e^{\frac{3}{8}}\right)^k \operatorname{vol}(E_{Q^0}) = \left(\frac{1}{2}e^{\frac{3}{8}}\right)^k \bar{d}^m \operatorname{vol}(B(0,1)).$$

On the other hand, since $B(0,\rho(d)) \subseteq \mathcal{T}_d$, we have $\operatorname{vol}(\mathcal{T}_d) \ge \operatorname{vol}(B(0,\rho(d)) = \rho(d)^m \operatorname{vol}(B(0,1))$. Therefore, $\rho(d)^m \operatorname{vol}(B(0,1)) \le \bar{d}^m (\frac{1}{2}e^{\frac{3}{8}})^k \operatorname{vol}(B(0,1))$, and Algorithm WLJ will perform at most

$$K \le m \ln\left(\frac{\bar{d}}{\rho(d)}\right) \cdot \frac{1}{\ln 2 - .375} \le \frac{10}{3} m \ln\left(\frac{\bar{d}}{\rho(d)}\right) \tag{40}$$

iterations.

To bound the skewness of the ellipsoids generated by Algorithm WLJ, note that all such ellipsoids contain the set $\mathcal{T}_d$ and therefore contain $B(0,\rho(d))$. This implies that for any ellipsoid encountered by the algorithm, $\lambda_1 \ge \rho(d)^2$.

We now estimate the change in the largest eigenvalue of the ellipsoid matrix Q^k from one iteration of the algorithm to the next. Suppose Q and $\hat{Q}$ are two consecutive iterates of the algorithm. Then from (38) we conclude that

$$\hat{\lambda}_m = \|\hat{Q}\| \le \|Q\| \frac{m}{m-1}\left(1 - \frac{1}{4m\xi}\right) = \lambda_m \frac{m}{m-1}\left(1 - \frac{1}{4m\xi}\right) \le \lambda_m \frac{m}{m-1} \le \lambda_m e^{\frac{1}{m-1}}.$$

Hence, at any iteration k,

$$\lambda_m^k \le \lambda_m^0 e^{\frac{k}{m-1}} = \bar{d}^2 e^{\frac{k}{m-1}} \le \left(\frac{\bar{d}}{\rho(d)}\right)^{\frac{10m}{3(m-1)}} \bar{d}^2,$$

the last inequality following from (40). Therefore, throughout the algorithm, the skewness of all ellipsoids generated by the algorithm is bounded above by

$$\sqrt{\frac{\lambda_m}{\lambda_1}} \le \sqrt{\left(\frac{\bar{d}}{\rho(d)}\right)^{\frac{10m}{3(m-1)}+2}} \le \left(\frac{\bar{d}}{\rho(d)}\right)^5. \tag{41}$$

Using (41) we conclude from Lemma 32 that any call to subroutine *Weak Check* will perform at most $O(m\sqrt{\vartheta^*}\ln(\frac{m\vartheta^*}{\tau^*} \cdot \frac{\bar{d}}{\rho(d)}))$ iterations of the barrier method. Combining this with (40), we can bound the total number of iterations of the barrier method performed by Algorithm WLJ by

$$O\left(m^2\sqrt{\vartheta^*}\ln^2\left(\frac{\bar{d}}{\rho(d)}\right)\ln\left(\frac{m\vartheta^*}{\tau^*}\right)\right).$$

Finally, the inequalities $\mu_{\mathcal{A}} \le \mathcal{C}(\tilde{B}d) \le \frac{4m\mu_{\mathcal{A}}}{\delta}$ follow from Theorem 18, (28), and Lemma 29. ∎

REMARK 35. *Note that the skewness of the ellipsoids does not necessarily degrade at every iteration. In fact, the last ellipsoid of the algorithm has the nice property that* $\sqrt{\frac{\lambda_m}{\lambda_1}} \le 4\sqrt{m}\mathcal{C}(d)$.

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

To see why this remark is true, notice that the axes of any ellipsoid of the algorithm will satisfy $\|v_i\| \geq \rho(d)$ for all i, and so $\sqrt{\lambda_1} \geq \rho(d)$. Also, the last ellipsoid of the algorithm satisfies $\frac{1}{4\sqrt{m}} v_i \in \mathcal{T}_d \subset B(0, \|d\|)$ for all i, and so $\|v_i\| \leq 4\sqrt{m}\|d\|$, whereby $\sqrt{\lambda_m} \leq \sqrt{m}\|d\|$.

To further interpret the complexity result of Theorem 34, suppose for simplicity that $\bar{d} = \|d\|$, i.e., the size of the data $\|d\|$ is known. Then Algorithm WLJ will perform at most

$$O\left(m^2\sqrt{\vartheta^*}\ln^2\left(\mathcal{C}(d)\right)\ln\left(\frac{m\vartheta^*}{\tau^*}\right)\right)$$

iterations. We see that the condition number $\mathcal{C}(d)$ of the initial data instance d plays a crucial role in the complexity of Algorithm WLJ, which aims to find an equivalent data instance whose condition number is within a given factor of the best possible. In particular, if the original data instance d is badly conditioned, i.e., $\mathcal{C}(d)$ is large, it might take a large number of iterations to find a "good" preconditioner as above. Another interesting observation is that the complexity of Algorithm WLJ depends on $\mathcal{C}(d)$ rather then $\mu_{\mathcal{A}}$. This result, which may seem counterintuitive at first, is actually explained by the fact that in order to obtain a preconditioner, Algorithm WLJ has to work with the set $\mathcal{T}_d$, rather then $\mathcal{H}_d$, which is symmetric about 0 regardless of the geometry of $\mathcal{H}_d$.

6. Conclusions. In this paper we have addressed several issues related to measures of conditioning for convex feasibility problems. We have discussed some potential drawbacks of using the condition number $\mathcal{C}(d)$ as the sole measure of conditioning of a conic linear system, motivating the study of data-independent measures. We have introduced the symmetry measure $\mu_{\mathcal{A}}$ for feasible conic linear systems as one such data-independent measure, and we have studied many of its implications for problem geometry, conditioning, and algorithm complexity.

One research topic that is not addressed in this paper concerns the existence of data-independent measures of conditioning for (FP_d) that are useful when (FP_d) is infeasible and/or whether any such measures can be adapted to analyze the linear optimization version of (FP_d). Such measures might or might not be an extension of the symmetry measure discussed in this paper.

Another potential topic of research stems from the importance of the inherent conditioning of the problem data for certain properties of (FP_d) such as sensitivity to data perturbations and numerical precision required for accurate computation in algorithms. The complexity bound for computing the good preconditioner in Algorithm WLJ is only reassuring in theory, as it would be unthinkable to use this algorithm in practice. Instead, much as in the case for linear optimization, it would be interesting to explore heuristic methods for preconditioning (FP_d). The notion of a heuristic preconditioning/preprocessing stage in an algorithm is well-established; most optimization software packages include some type of preprocessing options, such as variable and constraint elimination or data scaling, for improving condition numbers and other numerical measures in matrix computations. We hope that the results in this paper may inspire future research on the analysis of heuristic preconditioning techniques for solving linear and conic optimization problems.

Appendix. Implementation of subroutine *Weak Check*. In this appendix we present an implementation of the subroutine *Weak Check*. Recall that each iteration of Algorithm WLJ calls the subroutine *Weak Check* with input being the axes

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

$v_1, \ldots, v_m$ of an ellipsoid $E_Q \supseteq \mathcal{T}_d$. The purpose of *Weak Check* is to verify whether the rescaled axes $\pm\frac{1}{4\sqrt{m}}v_i$ are elements of $\mathcal{T}_d$, for $i = 1, \ldots, m$, and if not, to produce a parallel cut vector s satisfying (33).

Consider the following $2m$ optimization problems (P_{ϕ_v}), where $v = \pm v_i$, $i = 1, \ldots, m$:

$$(\mathrm{P}_{\phi_v}) \quad \begin{array}{lll} \phi_v = \max_{\phi} & \phi & \\ \text{s.t.} & \phi v \in \mathcal{H}_d & \end{array} \quad = \quad \begin{array}{ll} \max_{\theta,x,\phi} & \phi \\ \text{s.t.} & b\theta - Ax = v\phi, \\ & |\theta| + \|x\| \le 1, \\ & \theta \ge 0,\ x \in C_X. \end{array} \tag{42}$$

It is easy to verify that $\pm\frac{1}{4\sqrt{m}}v_i \in \mathcal{T}_d$ for all $i = 1, \ldots, m$ precisely when

$$\phi_Q \stackrel{\triangle}{=} \min_{\pm v_i} \phi_v \ge \frac{1}{4\sqrt{m}}. \tag{43}$$

(Here $\min_{\pm v_i} \phi_v$ stands for $\min\{\phi_{v_1}, -\phi_{v_1}, \ldots, \phi_{v_m}, -\phi_{v_m}\}$.) We will therefore implement the subroutine *Weak Check* by means of approximately solving the $2m$ optimization problems (42) and checking whether condition (43) is satisfied.

The approach we use to solve the optimization problems (42) in the subroutine *Weak Check* relies on the barrier method described in section 3. Since no obvious starting point is available for (42), we solve (42) for all $2m$ values of $v = \pm v_i$, $i = 1, \ldots, m$, by considering its dual:

$$(\mathrm{P}_{\gamma_v}) \quad \gamma_v = \begin{array}{ll} \min_{s,q,\gamma} & \gamma \\ \text{s.t.} & \|A^t s - q\| \le \gamma, \\ & b^t s \le \gamma, \\ & q \in C_X^*, \\ & v^t s = 1. \end{array} \tag{44}$$

It is straightforward to verify that strong duality holds for (P_{ϕ_v}) and (P_{γ_v}), and so

$$\phi_Q = \min_{\pm v_i} \phi_v = \min_{\pm v_i} \gamma_v.$$

In order to be able to apply the barrier method, we need the optimization problem at hand to have a bounded feasible region. To satisfy this condition, we consider the following modification of (44):

$$(\mathrm{P}_{\tilde{\gamma}_v}) \quad \tilde{\gamma}_v = \begin{array}{ll} \min_{s,q,\gamma} & \gamma \\ \text{s.t.} & \|A^t s - q\| \le \gamma, \\ & b^t s \le \gamma, \\ & \|V^t s\| \le 2\sqrt{m}, \\ & \gamma \le \frac{7\sqrt{m}\bar{d}}{\sqrt{\lambda_1}}, \\ & q \in C_X^*, \\ & v^t s = 1, \end{array} \tag{45}$$

where $\bar{d}$ is the known upper bound on the norm of the data $\|d\|$, and $V = [v_1, \ldots, v_m] \in \Re^{m\times m}$. The following two simple facts are useful in the derivation of the forthcoming results. First, for all $i = 1, \ldots, m$, we have $\sqrt{\lambda_1} \le \|v_i\| \le \sqrt{\lambda_m}$. Second, for any vector $s \in Y^*$, $\|s\| \le \frac{\|V^t s\|}{\sqrt{\lambda_1}}$. In the next proposition we show that solving $(\mathrm{P}_{\tilde{\gamma}_v})$ instead of (P_{γ_v}) still yields a valid estimate of ϕ_Q.

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

PROPOSITION 36. *For any v, $\gamma_v \leq \tilde{\gamma}_v$. Moreover,*

$$\phi_Q = \min_{\pm v_i} \gamma_v = \min_{\pm v_i} \tilde{\gamma}_v. \tag{46}$$

Proof. The first claim of the proposition is trivially true, since the feasible region of the program $(\mathrm{P}_{\tilde{\gamma}_v})$ is contained in the feasible region of the program (P_{γ_v}).

To establish the second claim, note that

$$\phi_Q = \min_{\pm v_i} \gamma_{\pm v_i} \leq \min_{\pm v_i} \tilde{\gamma}_{\pm v_i}.$$

Suppose the minimum on the left is attained for $v = v_{i_0}$, and let $(\bar{s}, \bar{q}, \gamma_v)$ be an optimal solution of the corresponding program (P_{γ_v}). Then we have

$$\gamma_v = \max\{\|A^t\bar{s} - \bar{q}\|, b^t\bar{s}\}, \quad \bar{q} \in C_X^*, \quad v^t\bar{s} = 1.$$

We can further assume without loss of generality that $\|A^t\bar{s} - \bar{q}\| \leq \|A^t\bar{s}\|$, since $\bar{q}$ can always be chosen to minimize the distance from $A^t\bar{s}$ to the cone C_X^*. If the point $(\bar{s}, \bar{q}, \gamma_v)$ is feasible for the corresponding program $(\mathrm{P}_{\tilde{\gamma}_v})$, then $\gamma_v = \tilde{\gamma}_v$ and (46) follows. Otherwise, let $\sigma = \max_i |v_i^t\bar{s}| \geq 1$. We can assume without loss of generality that $\sigma = v_j^t\bar{s}$ for some j. (If $v_j^t\bar{s} < 0$, we can redefine the jth axis of E_Q to be $-v_j$.) Define $(\tilde{s}, \tilde{q}, \tilde{\gamma}) = (\frac{1}{\sigma}\bar{s}, \frac{1}{\sigma}\bar{q}, \frac{1}{\sigma}\gamma_v)$. Note that $v_j^t\tilde{s} = 1$, $\tilde{q} \in C_X^*$, and

$$\|V^t\tilde{s}\| = \sqrt{\sum_{i=1}^{m}(v_i^t\tilde{s})^2} \leq \sqrt{m} \leq 2\sqrt{m}.$$

It remains to check whether the upper bound constraint on $\tilde{\gamma}$ is satisfied. Observe that $\|\tilde{s}\| \leq \frac{\sqrt{m}}{\sqrt{\lambda_1}}$ (since $\|V^t\tilde{s}\| \leq \sqrt{m}$). Therefore

$$\tilde{\gamma} = \max\{\|A^t\tilde{s} - \tilde{q}\|, b^t\tilde{s}\} \leq \max\{\|A^t\tilde{s}\|, b^t\tilde{s}\} \leq \bar{d} \cdot \frac{\sqrt{m}}{\sqrt{\lambda_1}} < \frac{7\sqrt{m}\bar{d}}{\sqrt{\lambda_1}}.$$

Hence the vector $(\tilde{s}, \tilde{q}, \tilde{\gamma})$ is feasible for $(\mathrm{P}_{\tilde{\gamma}_{v_j}})$, and $\tilde{\gamma}_{v_j} \leq \tilde{\gamma} \leq \gamma_v \leq \gamma_{v_j} \leq \tilde{\gamma}_{v_j}$, which implies that $\tilde{\gamma}_{v_j} = \gamma_v$, from which (46) follows. ∎

Now define

$$S \stackrel{\triangle}{=} \left\{(s, q, \gamma) : \|A^t s - q\| \leq \gamma,\ b^t s \leq \gamma,\ \|V^t s\| \leq 2\sqrt{m},\ \gamma \leq \frac{7\sqrt{m}\bar{d}}{\sqrt{\lambda_1}},\ q \in C_X^*\right\}$$

and

$$L_v \stackrel{\triangle}{=} \{(s, q, \gamma) : v^t s = 1\}.$$

Then L_v is a translate of an affine space, and S is a bounded convex set. Recall from the assumptions in section 5.2 that $f^*(\cdot)$ is a self-concordant barrier for the cone C_X^* with complexity parameter ϑ^*. Then the interior of the set S is the domain of the following self-concordant barrier $f(s, q, \gamma)$:

$$f(s, q, \gamma) \stackrel{\triangle}{=} f^*(q) - \ln(\gamma^2 - \|A^t s - q\|^2) - \ln(\gamma - b^t s) - \ln(4m - \|V^t s\|^2) - \ln\left(\frac{7\sqrt{m}\bar{d}}{\sqrt{\lambda_1}} - \gamma\right),$$

whose complexity parameter is $\vartheta_f \leq \vartheta^* + 5$.

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

In order to use the barrier method to solve $(\mathrm{P}_{\tilde{\gamma}_v})$, we need to have a point $(s', q', \gamma') \in \operatorname{int} S \cap L_v$ at which to initialize the method. The next proposition indicates that such point is readily available when the vector $u_* \in C_X^*$ of (7) is known; the second part of the proposition presents a lower bound on $\operatorname{sym}(S \cap L_v, (s', q', \gamma'))$, which is important in analyzing the complexity of the barrier method.

PROPOSITION 37.

$$(s', q', \gamma') \triangleq \left(\frac{v}{\|v\|^2}, \frac{2\bar{d}u_*}{\|v\|}, \frac{4\sqrt{m}\bar{d}}{\sqrt{\lambda_1}} \right) \in \operatorname{int} S \cap L_v,$$

and

$$\operatorname{sym}(S \cap L_v, (s', q', \gamma')) \geq \frac{\tau^*}{13\sqrt{m}} \cdot \sqrt{\frac{\lambda_1}{\lambda_m}}.$$

Proof. The first claim of the proposition is easily established by verifying directly that (s', q', γ') strictly satisfies the constraints of (45). The derivation of the bound on the symmetry in the second claim is fairly long and tedious, and is omitted. We refer the interested reader to [3] for details. ∎

We now present the formal statement of the implementation of the subroutine *Weak Check*.

Subroutine Weak Check.

- Input: Axes $v_i,\ i = 1, \ldots, m$, of an ellipsoid $E_Q \supseteq \mathcal{T}_d$.
- for $v = \pm v_i,\ i = 1, \ldots, m$,
  Step 1 Form the problem $(\mathrm{P}_{\tilde{\gamma}_v})$.
  Step 2 Run the barrier method on the problem $(\mathrm{P}_{\tilde{\gamma}_v})$ initialized at the point
  $$(s', q', \gamma') = \left(\frac{v}{\|v\|^2}, \frac{2\bar{d}u_*}{\|v\|}, \frac{4\sqrt{m}\bar{d}}{\sqrt{\lambda_1}} \right)$$
  until the value of the barrier parameter η first exceeds $\bar{\eta} = \frac{24\sqrt{m}\vartheta_f}{5}$. Let (s, q, γ) be the last iterate of the barrier method.
  Step 3 If $\gamma < \frac{1}{2\sqrt{m}}$, terminate and return s. Otherwise, continue with the next value of v.
- Assert that $\frac{1}{4\sqrt{m}} v_i \in \mathcal{T}_d$ for all $i = 1, \ldots, m$.

Proof of Lemma 32. Subroutine *Weak Check* will apply the barrier method to at most $2m$ problems of the form $(\mathrm{P}_{\tilde{\gamma}_v})$. Note that

$$\min_{(s,q,\gamma) \in S \cap L_v} \gamma \geq 0 \quad \text{and} \quad \max_{(s,q,\gamma) \in S \cap L_v} \gamma \leq \frac{7\sqrt{m}\bar{d}}{\sqrt{\lambda_1}}.$$

Therefore, applying (15) and Proposition 37, we see that each of the (at most) $2m$ applications of the barrier method will terminate in at most

$$\begin{aligned} & O\left(\sqrt{\vartheta_f} \ln \left(\frac{7\sqrt{m}\bar{d}\vartheta_f}{\sqrt{\lambda_1}} \cdot \frac{\bar{\eta}}{\operatorname{sym}(S \cap L_v, (s', q', \gamma'))} \right) \right) \\ \leq & O\left(\sqrt{\vartheta^*} \ln \left(\frac{7\sqrt{m}\bar{d}\vartheta^*}{\sqrt{\lambda_1}} \cdot \frac{24\sqrt{m}\vartheta^*}{5} \cdot \frac{13\sqrt{m}}{\tau^*} \cdot \sqrt{\frac{\lambda_m}{\lambda_1}} \right) \right) \\ = & O\left(\sqrt{\vartheta^*} \ln \left(\frac{m\vartheta^*}{\tau^*} \cdot \frac{\bar{d}}{\sqrt{\lambda_1}} \cdot \sqrt{\frac{\lambda_m}{\lambda_1}} \right) \right) \end{aligned}$$

iterations, giving (36).

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Suppose subroutine *Weak Check* has terminated in Step 3 of an iteration in which the barrier method is applied to the problem $(\mathrm{P}_{\tilde{\gamma}_{v_j}})$. (This is without loss of generality; if the termination occurs during the iteration that applies the barrier method to the problem $(\mathrm{P}_{\tilde{\gamma}_{-v_j}})$, we can redefine the jth axis of E_Q to be $-v_j$, to preserve the notation.) Then the last iterate (s, q, γ) of the barrier method satisfies

$$\begin{gathered}\|A^t s - q\| \leq \gamma < \tfrac{1}{2\sqrt{m}},\\ b^t s \leq \gamma < \tfrac{1}{2\sqrt{m}},\\ \|V^t s\| \leq 2\sqrt{m},\\ q \in C_X^*,\ v_j^t s = 1.\end{gathered}$$

The vector s above yields a parallel cut that separates $\pm\frac{v_j}{2\sqrt{m}}$ from $\mathcal{T}_d$. To see why this is true, let $h \in \mathcal{T}_d$. Then $h \in \mathcal{H}_d$, and hence $h = b\theta - Ax$ for some $(\theta, x) \in \Re_+ \times C_X$ such that $|\theta| + \|x\| \leq 1$. Therefore

$$\begin{aligned}s^t h = s^t(b\theta - Ax) &= \theta(b^t s) - x^t(A^t s) = \theta(b^t s) - x^t(A^t s - q) - x^t q\\ &\leq (|\theta| + \|x\|)\gamma \leq \gamma < \frac{1}{2\sqrt{m}} = \frac{s^t v_j}{2\sqrt{m}}.\end{aligned}$$

Applying the same argument for the point $-h \in \mathcal{H}_d$, we conclude that $s^t h > -\frac{s^t v_j}{2\sqrt{m}}$, and therefore the vector s returned by the subroutine *Weak Check* satisfies (37).

Next, suppose that the barrier method applied to $(\mathrm{P}_{\tilde{\gamma}_v})$ has not terminated in Step 3 of the subroutine *Weak Check*, i.e., we have $\gamma \geq \frac{1}{2\sqrt{m}}$. Then, using (14),

$$\tilde{\gamma}_v \geq \gamma - \frac{6\vartheta_f}{5\bar{\eta}} \geq \frac{1}{2\sqrt{m}} - \frac{6\vartheta_f}{5\bar{\eta}} \geq \frac{1}{4\sqrt{m}}.$$

Therefore, if the subroutine *Weak Check* has not terminated in Step 3 for any $v = \pm v_i,\ i = 1, \ldots, m$, we conclude that $\phi_Q = \min_{\pm v_i} \tilde{\gamma}_v \geq \frac{1}{4\sqrt{m}}$, and we correctly assert that $\pm\frac{1}{4\sqrt{m}} v_i \in \mathcal{T}_d$ for all $i = 1, \ldots, m$. ∎

Acknowledgments. The authors would like to thank James Renegar and two anonymous referees for their valuable comments and suggestions.

REFERENCES

[1] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming, Theory and Algorithms*, 2nd ed., John Wiley and Sons, New York, 1993.

[2] R. Bland, D. Goldfarb, and M. J. Todd, *The ellipsoid method: A survey*, Oper. Res., 29 (1981), pp. 1039–1091.

[3] M. Epelman, *Complexity, Condition Numbers, and Conic Linear Systems*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, 1999.

[4] M. Epelman and R. M. Freund, *Condition number complexity of an elementary algorithm for computing a reliable solution of a conic linear system*, Math. Program., 88 (2000), pp. 451–486.

[5] S. Filipowski, *On the complexity of solving sparse symmetric linear programs specified with approximate data*, Math. Oper. Res., 22 (1997), pp. 769–792.

[6] S. Filipowski, *On the complexity of solving feasible linear programs specified with approximate data*, SIAM J. Optim., 9 (1999), pp. 1010–1040.

[7] R. M. Freund and J. R. Vera, *On the complexity of computing estimates of condition measures of a conic linear system*, Technical report, Operations Research Center, MIT, Cambridge, MA, 1999.

[8] R. M. Freund and J. R. Vera, *Some characterizations and properties of the "distance to ill-posedness" and the condition measure of a conic linear system*, Math. Program., 86 (1999), pp. 225–260.

[9] R. M. Freund and J. R. Vera, *Condition-based complexity of convex optimization in conic linear form via the ellipsoid algorithm*, SIAM J. Optim., 10 (2000), pp. 155–176.
[10] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag, New York, 1994.
[11] J. C. K. Ho, *Structure of the central path and related complexity issues for linear programming*, Master's thesis, University of Waterloo, Waterloo, ON, Canada, 1999.
[12] F. John, *Extremum problems with inequalities as subsidiary conditions*, in Studies and Essays, Presented to R. Courant on His 60th Birthday, Interscience, New York, 1948, pp. 187–204.
[13] N. Karmarkar, *A new polynomial-time algorithm for linear programming*, Combinatorica, 4 (1984), pp. 373–395.
[14] L. G. Khachiyan, *A polynomial algorithm in linear programming*, Soviet Math. Dokl., 20 (1979), pp. 191–194.
[15] V. Klee and G. J. Minty, *How good is the simplex algorithm?*, in Inequalities, O. Shisha, ed., Academic Press, New York, 1972, pp. 159–175.
[16] N. Megiddo, S. Mizuno, and T. Tsuchiya, *A modified layered-step interior-point algorithm for linear programming*, Math. Programming, 82 (1998), pp. 339–355.
[17] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, Philadelphia, 1994.
[18] M. A. Nunez, *A Characterization of Ill-Posed Data Instances for Convex Programming*, informal paper, 1998.
[19] M. A. Nunez and R. M. Freund, *Condition measures and properties of the central trajectory of a linear program*, Math. Programming, 83 (1998), pp. 1–28.
[20] J. Peña, *Computing the Distance to Infeasibility: Theoretical and Practical Issues*, Technical report, Cornell University, Ithaca, NY, 1997.
[21] J. Peña, *Understanding the geometry of infeasible perturbations of a conic linear system*, SIAM J. Optim., 10 (2000), pp. 534–550.
[22] J. Peña and J. Renegar, *Computing approximate solutions for conic systems of constraints*, Math. Program., 87 (2000), pp. 351–383.
[23] J. Renegar, *A polynomial-time algorithm, based on Newnton's method, for linear programming*, Math. Programming, 40 (1988), pp. 59–93.
[24] J. Renegar, *Some perturbation theory for linear programming*, Math. Programming, 65 (1994), pp. 73–91.
[25] J. Renegar, *Incorporating condition measures into the complexity theory of linear programming*, SIAM J. Optim., 5 (1995), pp. 506–524.
[26] J. Renegar, *Linear programming, complexity theory, and elementary functional analysis*, Math. Programming, 70 (1995), pp. 279–351.
[27] J. Renegar, *A Mathematical View of Interior-Point Methods for Convex Optimization*, SIAM, Philadelphia, 2001.
[28] S. M. Robinson, *Bounds for error in the solution set of a perturbed linear program*, Linear Algebra Appl., 6 (1973), pp. 69–81.
[29] M. J. Todd, L. Tunçel, and Y. Ye, *Characterizations, bounds, and probabilistic analysis of two complexity measures for linear programming problems*, Math. Program., 90 (2001), pp. 59–69.
[30] L. Tunçel, *Approximating the complexity measure of Vavasis-Ye algorithm is NP-hard*, Math. Program., 86 (1999), pp. 219–223.
[31] L. Tunçel, *On the condition numbers for polyhedra in Karmarkar's form*, Oper. Res. Lett., 24 (1999), pp. 149–155.
[32] S. A. Vavasis and Y. Ye, *Condition numbers for polyhedra with real number data*, Oper. Res. Lett., 17 (1995), pp. 209–214.
[33] S. A. Vavasis and Y. Ye, *A primal-dual interior point method whose running time depends only on the constraint matrix*, Math. Programming, 74 (1996), pp. 79–120.
[34] S. A. Vavasis and Y. Ye, *A simplification to "a primal-dual interior point method whose running time depends only on the constraint matrix,"* in High Performance Optimization, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000, pp. 233–243.
[35] J. R. Vera, *Ill-Posedness and the Computation of Solutions to Linear Programs with Approximate Data*, Technical report, Cornell University, Ithaca, NY, 1992.
[36] J. R. Vera, *Ill-Posedness in Mathematical Programming and Problem Solving with Approximate Data*, Ph.D. thesis, Cornell University, Ithaca, NY, 1992.
[37] J. R. Vera, *Ill-posedness and the complexity of deciding existence of solutions to linear programs*, SIAM J. Optim., 6 (1996), pp. 549–569.
[38] J. R. Vera, *On the complexity of linear programming under finite precision arithmetic*, Math. Programming, 80 (1998), pp. 91–123.
[39] S. J. Wright, *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, PA, 1997.

Downloaded 11/24/15 to 18.189.82.176. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php