http://doi.org/10.1007/s10107-003-0435-1
"Complexity of Convex Optimization using Geometry-Based Measures and a Reference Point." Freund, Robert M. *Mathematical Programming* Vol. 99, No. 2 (2004): 197-221.

COMPLEXITY OF CONVEX OPTIMIZATION USING GEOMETRY-BASED MEASURES AND A REFERENCE POINT [1]

Robert M. Freund[2]
M.I.T.

September, 2001
revised August, 2002

Abstract

Our concern lies in solving the following convex optimization problem:

$$\begin{array}{lll} G_P: & \mathrm{minimize}_x & c^T x \\ & \mathrm{s.t.} & Ax = b \\ & & x \in P, \end{array}$$

where P is a closed convex subset of the n-dimensional vector space X. We bound the complexity of computing an almost-optimal solution of G_P in terms of natural geometry-based measures of the feasible region and the level-set of almost-optimal solutions, relative to a given *reference point* x^r that might be close to the feasible region and/or the almost-optimal level set. This contrasts with other complexity bounds for convex optimization that rely on data-based condition numbers or algebraic measures, and that do not take into account any *a priori* reference point information.

AMS Subject Classification: 90C, 90C05, 90C60
Keywords: Convex Optimization, Complexity, Interior-Point Method, Barrier Method

[1]This research has been partially supported through the Singapore-MIT Alliance. Portions of this research were undertaken when the author was a Visiting Scientist at Delft University of Technology.

[2]MIT Sloan School of Management, 50 Memorial Drive, Cambridge, MA 02142-1347, USA. email: rfreund@mit.edu

1 Introduction, Motivation, and Main Result

Consider the following convex optimization problem:

$$\begin{array}{llll} G_P: & z^* := \text{minimum}_x & c^T x \\ & \text{s.t.} & Ax = b \\ & & x \in P, \end{array}$$

where P is a closed convex set in the (finite) n-dimensional linear vector space X, and b lies in the (finite) m-dimensional vector space Y. We call this problem linear optimization with a ground-set, and we call P the ground-set. In practical applications, P could be the solution of box constraints of the form $l \le x \le u$, a convex cone C, or perhaps the solution to network flow constraints of the form $Nx = b, x \ge 0$. However, for ease of presentation, we will make the following assumption:

Assumption A: P has an interior, and $\{x \mid Ax = b\} \cap \text{int}P \neq \emptyset$.

For $\epsilon > 0$, we call x an ϵ-optimal solution of G_P if x is a feasible solution of G_P that satisfies $c^T x \le z^* + \epsilon$. The chief concern in this paper is an algorithm and associated complexity bound for computing an ϵ-optimal solution of G_P.

Let $\|\cdot\|$ be any norm on X, and let $B(x, r)$ denote the ball of radius r centered at x:

$$B(x, r) := \{w \in X \mid \|w - x\| \le r\} \ .$$

The norm $\|\cdot\|$ might be a problem-appropriate norm for the actual problem context at hand. However, in Section 5.1, we will examine in detail two norms on X that arise "naturally" in association with the ground-set P.

The computational engine that we will use to solve G_P is the barrier method based on the theory of self-concordant barriers, and we presume that the reader has a general familiarity with this topic as developed in [5] and/or [7], for example. We therefore assume that we have a ϑ_P-self-concordant barrier $F_P(\cdot)$ for P. We also assume that we have a $\vartheta_{\|\ \|}$-self-concordant barrier $F_{\|\ \|}(\cdot)$ for the unit ball:

$$B(0, 1) = \{x \mid \|x\| \le 1\} \ .$$

The work here is motivated by a desire to generalize and improve several aspects of the general complexity theory for conic convex optimization developed by Renegar in [7], key elements of which we now attempt to summarize in a brief and somewhat simplified manner. In [7], the general convex optimization problem G_P is assumed to be conic, that is, P is assumed to be closed convex cone C, and the data for the problem is given by the array

$d = (A, b, c)$. One complexity result that can be gleaned from [7] is as follows: assuming that G_P has a feasible solution, there is an algorithm based on interior-point methods that will compute an ϵ-optimal solution of G_P in

$$O\left(\sqrt{\vartheta_C}\ln\left(\mathcal{C}(d) + \vartheta_C + \frac{\|d\|}{\epsilon} + \frac{\|\check{x}\|}{\operatorname{dist}(\check{x}, \partial C)} + \frac{\max\{\bar{s}, \|d\|\}}{\min\{\bar{s}, \|d\|\}}\right)\right) \tag{1}$$

iterations of Newton's method (see Theorem 3.1 and Corollary 7.3 of [7]), where we use the notation ϑ_C to denote the complexity value of the barrier for the cone C. Here $\check{x} \in C$ is a given interior point of the cone C that is specified as part of the input to the algorithm, $\operatorname{dist}(\check{x}, \partial C)$ is the distance from $\check{x}$ to the boundary of C, $\|d\|$ is the (suitably defined) norm of the data d, and $\bar{s}$ is a positive scalar that must be specified as input to the algorithm. The quantity $\mathcal{C}(d)$ is the condition number of the data d, defined as:

$$\mathcal{C}(d) := \frac{\|d\|}{\min\{\rho_P(d), \rho_D(d)\}}\ , \tag{2}$$

where $\rho_P(d), \rho_D(d)$ are the primal and dual distances to ill-posedness, see [7] for details and motivating discussion. ($\mathcal{C}(d)$ naturally extends the concept of condition number of a system of equations to the far broader problem of conic convex optimization.) The complexity result (1) is remarkable for its breadth and generality, as well as for its reliance on natural data-dependent concepts imbedded in condition-number theory. In order to keep the presentation brief, we have shown a simplified and slightly weaker complexity result in (1) than the verbatim complexity bound in [7]. Furthermore, [7] has many other complexity results related to conic convex optimization in finite as well as infinite-dimensional settings.

While the significance of (1) and the many related results in [7] cannot be overstated, there are certain issues with this type of complexity bound that are not very satisfactory. One issue has to do with undue data dependence. Given a data instance $d = (A, b, c)$ for G_P and a nonsingular matrix B and a vector π of multipliers, we can create an equivalent representation of the problem G_P using the different data $\bar{d} = (\bar{A}, \bar{b}, \bar{c}) := (B^{-1}A, B^{-1}b, c - A^T\pi)$. The two data instances d and $\bar{d}$ will generally give rise to different complexity bounds using (1) since in general $\mathcal{C}(d) \neq \mathcal{C}(\bar{d})$, etc., yet both data instances represent the same underlying optimization problem.

Another issue with the condition number approach is that the problem must be in conic form. While any convex optimization problem can be transformed to conic form, such a transformation might not be natural (such as converting a quadratic objective to a linear objective using a second-order cone constraint, etc.) or unique (and so might further introduce arbitrarily different data for the same original problem). Yet a third issue with the condition number approach has to do with the fact that the theory assumes that the data is arising only in the linear equation system and the objective function, and that the cone C is fixed independent of any data. In this format, data used to defined the cone is not accounted for in the theory.

A fourth issue has to do with the role of the starting point $\check{x}$. The bound (1) is not dependent on or sensitive to the extent to which $\check{x}$ might be nearly feasible and/or nearly optimal. It would be nice to have a complexity bound that accounted for the proximity of $\check{x}$ to the feasible and/or optimal solution set.

The algorithm and analysis presented in this paper represent an attempt to overcome the above-mentioned issues. In Sections 3 and 4, we develop and analyze interior-point algorithms FEAS and OPT for finding a feasible solution $\bar{x}$ of G_P and an ϵ-optimal solution $\hat{x}$ of G_P, respectively. This pair of algorithms and their complexity analysis depend on certain geometry-based measures for analyzing convex optimization problems and the concept of a *reference point* x^r, which we now discuss.

1.1 Reference Point and Interior Point

The phase-I algorithm FEAS requires that the user specify two points as part of the input of the algorithm, the *reference point* x^r, and an interior point $x^0 \in \text{int}P$. The reference point x^r might be chosen to be an initial guess of a feasible and/or optimal solution, the solution to a previous version of the problem (such as in warm-start methodologies), or the origin 0 of the space X, etc. If P is the box defined by the constraints $l \le x \le u$, then x^r might be chosen as a given corner of the box such as $x^r = l$; if P is a convex cone C, x^r might be chosen to be the origin $x^r = 0$, or a known point on the boundary or the interior of C, etc. Certain properties of x^r will enter into the complexity bounds derived herein, particularly related to the distance from x^r to the feasible region, to the set of ϵ-optimal solutions, and to the set of nicely-interior feasible solutions. There is no assumption concerning whether or not x^r is in the ground-set P or satisfies the linear equations $Ax = b$.

Algorithm FEAS also requires an interior point $x^0 \in \text{int}P$. This interior point will be used in many ways to measure how interior other points in P are. (By analogy, in linear optimization $e := (1, \ldots, 1)^T$ is used to measure the positivity of other vectors v by computing the largest α for which $v \ge \alpha e$.) It will be desirable for x^0 to be nicely interior to P, as the quantity $\text{dist}(x^0, \partial P)$ will enter into our complexity bounds for solving G_P,; however, we do not require that we know the value of $\text{dist}(x^0, \partial P)$ in advance.

1.2 Phase I Geometry Measure g

The complexity of the phase-I algorithm FEAS will be bounded by the following geometric measure which we denote by g :

$$\begin{array}{rll} g := \mathrm{minimum}_{x,\rho} & \dfrac{\max\{\|x-x^r\|,1\}}{\min\{\rho,1\}} & \\ \text{s.t.} & Ax = b & \\ & B(x,\rho) \subset P \ . & \end{array} \tag{3}$$

If we ignore for the moment the "1"s in the numerator and denominator of the ratio defining g, (3) could be re-written as:

$$\begin{array}{rll} \tilde{g} := \mathrm{minimum}_{x} & \dfrac{\|x-x^r\|}{\mathrm{dist}(x,\partial P)} & \\ \text{s.t.} & Ax = b & \\ & x \in P \ . & \end{array} \tag{4}$$

and so g (or $\tilde{g}$) measures the extent to which x^r is close to an interior feasible solution x that is itself not too close to the boundary of P, and g (or $\tilde{g}$) is smaller to the extent that x^r is close to feasible solutions $x \in \mathrm{int}P$ that are themselves far from the boundary of P.

The ratios defining g and $\tilde{g}$ arise naturally in the complexity of the ellipsoid algorithm applied to the problem of finding a feasible solution of G_P. If one were to initiate the ellipsoid algorithm at the ball centered at the reference point x^r with a radius given by $\|x^r - \tilde{x}\| + \tilde{\rho}$ (where $\tilde{x}$ is an optimal solution of (4) and $\tilde{\rho} = \mathrm{dist}(\tilde{x}, \partial P)$), then it is easy to see that a suitably designed version of the ellipsoid method would compute a feasible solution of G_P in $O\left(n^2 \ln(\tilde{g})\right)$ iterations, under the presumption that the norm $\|\cdot\|$ is ellipsoidal. (See [4] for an expository treatment of the ellipsoid algorithm.) In the more typical context in continuous optimization where we do not have an *a priori* bound on the distance from the feasible region to the reference point, there is a natural projective transformation of the problem for which the ellipsoid algorithm will compute a feasible solution of G_P in $O\left(n^2 \ln(g)\right)$ iterations, see Lemma 4.1 of [2]. Therefore g is a very relevant geometric measure for the Phase-I problem in the context of the ellipsoid algorithm. Herein, we will see that g is also relevant for the complexity of the Phase-I problem for a suitably constructed interior-point algorithm.

The geometric measure g is also a generalization of the condition measure σ of Ye [9] for linear programming, which has been used in linear programming complexity analysis. The linear programming feasibility problem in standard form is to find a solution x of the system:

$$LF: \quad \begin{array}{l} Ax = b \\ x \in \Re^n_+ \ , \end{array} \tag{5}$$

which is an instance of the feasibility problem of G_P with $X = \Re^n$ and $P = \Re^n_+ := \{x \in \Re^n \mid x \geq 0\}$. The measure σ is associated with the homogenized version of LF, and is defined as:

$$\sigma = \min_{j=1,\ldots,n+1} \begin{array}{ll} \max & x_j \\ \text{s.t.} & Ax - bx_{n+1} = 0 \\ & \sum_{k=1}^{n+1} x_k = 1 \\ & x_1, \ldots, x_{n+1} \geq 0 \ . \end{array} \tag{6}$$

This definition of σ coincides with the usual definition of σ under Assumption A. The following proposition, whose proof is presented in the appendix, shows essentially that g is a generalization of the reciprocal of σ.

Proposition 1.1 *Under Assumption A and the L_1-norm on $X = \Re^n$, we have:*

$$\frac{1}{2}\left(\frac{1}{\sigma}\right) \leq g \leq (n+1)\left(\frac{1}{\sigma}\right) \ .$$

1.3 Phase II Geometry Measure D_ϵ

Our complexity analysis of the phase-II algorithm OPT will rely on the maximum distance from the reference point x^r to the set of ϵ-optimal solutions:

$$D_\epsilon := \max_x \left\{ \|x - x^r\| \mid Ax = b, x \in P, c^T x \leq z^* + \epsilon \right\} \ . \tag{7}$$

At first glance it may seem odd to maximize rather than minimize in defining D_ϵ. However, consider the ill-posed case when z^* is finite but the set of optimal solutions is unbounded, which can arise, for example, in semidefinite optimization. Then the dual feasible region has no interior, and so we would not expect to have an efficient complexity bound for solving G_P. In this context, the more relevant complexity measure is the maximum distance to the ϵ-optimal solution set (which would be infinite in this case) rather than the minimum distance (which would be finite in this case). Also, in [1] in the case of conic optimization with $x^r = 0$, it is shown that D_ϵ defined using (7) is inversely proportional to the size of the largest ball contained in the level sets of the dual problem, and so D_ϵ is very relevant in studying the behavior of primal-dual and/or dual interior-point algorithms for conic problems.

1.4 Main Result

Theorems 3.1 and 4.1 contain complexity bounds on the phase-I and phase-II algorithms FEAS and OPT, respectively. Taken as a pair, the combined complexity bound for the algorithms to compute an ϵ-optimal solution of G_P using the reference point x^r and the interior-point x^0 is:

$$O\left(\sqrt{\vartheta_P + \vartheta_{\|\ \|}}\ln\left(\begin{array}{c}\vartheta_P + \vartheta_{\|\ \|} + \frac{1}{\min\{\mathrm{dist}(x^0,\partial P),1\}} + \|x^0 - x^r\| \\ \\ +g + D_\epsilon + \max\left\{\frac{\tilde{s}}{\epsilon}, 1\right\}\end{array}\right)\right) \tag{8}$$

iterations of Newton's method, where

$$\tilde{s} := \max_w \left\{c^T w \mid \|w\| \le 1, Aw = 0\right\} \le \|c\|_* \ .$$

Note that (8) depends logarithmically on the phase-I and phase-II geometry measures g and D_ϵ, the inverse of the distance from x^0 to the boundary of P, as well as the distance from x^0 to the reference point x^r.

In Section 5.1, we present two choices of norms $\|\cdot\|$ on X that arise naturally and for which the complexity bound (8) simplifies to:

$$O\left(\sqrt{\vartheta_P}\ln\left(g + D_\epsilon + \vartheta_P + \max\left\{\frac{\tilde{s}}{\epsilon}, 1\right\} + \|x^0 - x^r\|\right)\right) \tag{9}$$

iterations of Newton's method. In Section 1.5.3, we show how the condition-number based complexity bound (1) can be derived as a special case of Theorems 3.1 and 4.1.

1.5 Comments and Discussion

In this subsection we make some comments and discuss several issues regarding the complexity bounds (8) and (9).

1.5.1 General Comments

The complexity bounds (8) or (9) do not explicitly depend on the data representation of the problem G_P in the sense that the bounds are invariant if we replace $d = (A, b, c)$ by $\bar{d} = (\bar{A}, \bar{b}, \bar{c}) := (B^{-1}A, B^{-1}b, c - A^T\pi)$ for any nonsingular matrix B and any multipliers π. The direct impact of the ground-set P on the complexity bounds is through the complexity value ϑ_P of the barrier function $F_P(\cdot)$ on $\text{int}P$. The indirect impact of P on the complexity bounds lies in the geometry of the feasible region of G_P, namely $P \cap \{x \in X \mid Ax = b\}$ through the geometry measure g, and in the geometry of the ϵ-optimal level set $P \cap \{x \in X \mid Ax = b, c^Tx \leq z^* + \epsilon\}$ through the measure D_ϵ; recall that g measures the extent to which there are feasible points close to x^r and far from the boundary of P, and D_ϵ measures how far the farthest point in the ϵ-optimal level set is from x^r.

Incidentally, in the case when P is a cone and $x^r = 0$, it is shown in [1] that D_ϵ is large if and only if the objective function level sets of the conic dual problem are close to the boundary of P^*, and so one can then interpret the presence of D_ϵ in the complexity bounds (8) or (9) as conveying information about the behavior of the dual problem.

1.5.2 On the Reference Point and Warm Starts

One of the goals of the research described herein has been to explore the extent to which a good and/or bad starting point can impact the complexity of finding an approximate solution of G_P. The complexity bound (8) or its simpler version (9) show how the quality of the reference point x^r influences the complexity bound through g, D_ϵ, and $\|x^0 - x^r\|$; and the number of Newton steps grows at most linearly in the logarithm of these quantities. However, we point out that even when x^r is already an ϵ-optimal solution and $x^0 = x^r$, then $g = 1$ and D_ϵ could nevertheless be large, and the iteration bounds (8) or (9) could still be quite large due also to the typically very large constants hidden in the big-O notation. It is a worthwhile goal to try to construct a truly viable warm-start algorithm for G_P whose iteration bound would be small to the extent that the starting point is close to the set of ϵ-optimal solutions. The recent work of Yildirim and Wright [10] is a promising step in this direction for the case of linear programming.

1.5.3 Relation to Condition-Number based Complexity Bounds

In this subsection we indicate how the condition-number based complexity bound for conic convex optimization presented in (1) can be obtained as a special case of the complexity bound (8). To do so, assume that P is a closed convex cone C, and for convenience we will assume that C is pointed and has an interior. We assume that we have a ϑ_C-self-concordant barrier $F_C(\cdot)$ for C, and we assume as in [7] that the norm $\|\cdot\|$ on X is an inner-product norm $\|v\| := \sqrt{v^T v}$, and so the barrier function

$$F_{\|\ \|}(v) := -\ln\left(1 - v^T v\right)$$

is a $\vartheta_{\|\ \|} = 1$-self-concordant barrier for the unit ball.

Let us set $x^r := 0$ and let $x^0 \in \text{int}C$ be given, and assume that we have rescaled x^0 so that $\|x^0\| = 1$. Then from Theorem 17 of [3], it follows that g will satisfy:

$$g \le 3\mathcal{C}(d)\frac{\|x^0\|}{\text{dist}(x^0, \partial C)}$$

and from Theorem 1.1 and Lemma 3.2 of [6] it follows that

$$D_\epsilon \le \mathcal{C}(d)^2 + \mathcal{C}(d)\frac{\epsilon}{\|c\|_*}$$

where $\mathcal{C}(d)$ is defined here using (2). Then under the hypothesis that $\epsilon \le \|c\|_*$, the complexity bound (8) simplifies to:

$$O\left(\sqrt{\vartheta_C}\ln\left(\vartheta_C + \frac{\|x^0\|}{\text{dist}(x^0, \partial P)} + \mathcal{C}(d) + \frac{\|c\|_*}{\epsilon}\right)\right) \tag{10}$$

iterations of Newton's method. This bound is almost identical to the bound (1) except that the term involving the user-specified scalar $\bar{s}$ in (1) is now not needed. As will be seen in Section 4, our methodology requires an explicit computation of $\bar{s}$ beforehand, and so is not as general as the methodology in [7]; however our method is more advantageous if the computational cost of accurately estimating $\|d\|$ in (1) is high.

2 Summary of Interior-Point Methodology

We employ the basic theoretical machinery of interior-point methods in our analysis using the theory of self-concordant barrier functions as articulated in Renegar [7] and [8], based on the theory of self-concordant functions of Nesterov and Nemirovskii [5]. The barrier method is essentially designed to approximately solve a problem of the form

$$OP: \quad \hat{z} = \min\{\bar{c}^T w \mid w \in S\},$$

where S is a compact convex subset of the n-dimensional space X, and $\bar{c} \in X^*$. The method requires the existence of a ϑ-self-concordant barrier function $F(w)$ for the relative interior of the set S, see [7] and [5] for details, and proceeds by approximately solving a sequence of problems of the form

$$OP_\mu: \quad \min\{\bar{c}^T w + \mu F(w) \mid w \in \mathrm{relint} S\},$$

for a decreasing sequence of values of the barrier parameter μ. We base our complexity analysis on the general convergence results for the barrier method presented in Renegar [7], which are similar to (but are more accessible for our purposes than) related results found in [5]. The barrier method starts at a given point $w^0 \in \mathrm{relint} S$. The method performs two stages. In stage I, the method starts from w^0 and computes iterates based on Newton's method, ending when it has computed a point $\hat{w}$ that is an approximate solution of $OP_{\hat{\mu}}$ for some barrier parameter $\hat{\mu}$ that is generated internally in stage I. In stage II, the barrier method computes a sequence of approximate solutions w^k of OP_{μ_k}, again using Newton's method, for a decreasing sequence of barrier parameters μ_k converging to zero. The goal of the barrier method is to find an ϵ-optimal solution of OP, which is a feasible solution w of OP for which $\bar{c}^T w \le \hat{z} + \epsilon$. One description of the complexity of the barrier method is as follows:

- *Assume that S is a bounded set, and that $w^0 \in \mathrm{relint} S$ is given. The barrier method requires*

$$O\left(\sqrt{\vartheta} \ln\left(\vartheta + \frac{1}{\mathrm{sym}(w^0, S)} + \frac{R}{\epsilon}\right)\right) \tag{11}$$

iterations of Newton's method to compute an ϵ-optimal solution of OP.

In the above expression, R is the range of the objective function $\bar{c}^T w$ over the set S, that is, $R = z^u - z^l$ where

$$z^l = \min\{\bar{c}^T w \mid w \in S\} \quad \text{and} \quad z^u = \max\{\bar{c}^T w \mid w \in S\},$$

and $\mathrm{sym}(w, S)$ is a measure of the symmetry of the point w with respect to the set S, and is defined as

$$\mathrm{sym}(w, S) := \max\{t \mid y \in S \Rightarrow w - t(y - w) \in S\}.$$

This term in the complexity of the barrier method arises since the closer the starting point is to the boundary, the larger is the value of the barrier function at this point, and so more effort is generally required to proceed from such a point.

The barrier method can also be used in equation-solving mode, to solve the system:

$$\begin{array}{c} w \in S \\ \bar{c}^T w = \delta \end{array} \tag{12}$$

for some given value of δ. A description of the complexity of the barrier method for equation-solving mode is as follows:

- *Assume that S is a bounded set and that $w^0 \in \mathrm{relint} S$. If $\delta \in \left(z^l, z^u\right)$, the barrier method requires*

$$O\left(\sqrt{\vartheta}\ln\left(\vartheta + \frac{1}{\mathrm{sym}(w^0, S)} + \frac{R}{\min\{z^u - \delta, \delta - z^l\}}\right)\right) \tag{13}$$

iterations to compute a point $\hat{w}$ that satisfies $\hat{w} \in S, \bar{c}^T\hat{w} = \delta$. Furthermore, $\hat{w}$ will also satisfy:

$$sym(\hat{w}, T) \geq \frac{1}{3.5\vartheta + 1.25} \tag{14}$$

where T is the level set:

$$T := \left\{ w \mid w \in S, \bar{c}^T w = \delta \right\} \ .$$

Because it will play a prominent role in our analysis, we present a derivation of (14) based on [5] and [8], under the assumption that S has an interior. Let w^c denote the analytic center of T, namely

$$w^c := \mathrm{argmin}_w \left\{ F(w) \mid w \in T \right\} \ .$$

For $w \in S$, let $\|\cdot\|_w$ denote the norm induced by the Hessian $H(w)$ of the barrier function $F(\cdot)$ at w, namely $\|v\|_w := \sqrt{v^T H(w) v}$, and let $\eta(w)$ denote the Newton direction for $F(\cdot)$ at w, namely $\eta(w) := -H(w)^{-1}\nabla F(w)$. Recall from Proposition 2.3.2 of [5] that all $w \in T$ satisfy $\|w - w^c\|_{w^c} \leq 3\vartheta + 1$. There is a fixed constant $\gamma < 1$, the value of γ being dependent on the specific implementation of the barrier method, such that the final iterate $\hat{w} \in T$ of the barrier method will satisfy $\|\eta(\hat{w}) - \hat{\lambda}\bar{c}\|_{\hat{w}} \leq \gamma$ for some multiplier $\hat{\lambda}$. By taking a fixed extra number of Newton steps if necessary, we can assume that $\gamma := \frac{1}{12}$. Then from Theorem 2.2.5 of [8] we have $\|\hat{w} - w^c\|_{\hat{w}} \le \gamma + \frac{3\gamma^2}{(1-\gamma)^3} \le \frac{1}{9}$, and so for all $w \in T$ we have

$$\begin{aligned} \|w - \hat{w}\|_{\hat{w}} &\le \|w - w^c\|_{\hat{w}} + \|w^c - \hat{w}\|_{\hat{w}} \\ &\le \left(\frac{1}{1-\frac{1}{9}}\right) \|w - w^c\|_{w^c} + \frac{1}{9} \\ &\le \frac{9}{8}(3\vartheta + 1) + \frac{1}{9} \\ &\le 3.5\vartheta + 1.25 \ . \end{aligned} \tag{15}$$

(The second inequality above follows from Theorem 2.1.1 of [5].) Furthermore, if w satisfies $\bar{c}^T w = \delta$ and $\|w - \hat{w}\|_{\hat{w}} \le 1$, then $w \in T$ (also from Theorem 2.1.1 of [5]), and together with (15) this then implies that $\mathrm{sym}(\hat{w}, T) \ge \frac{1}{3.5\vartheta + 1.25}$.

Remark 2.1 *Inequality (14) implies that for any objective function vector $s \in X^*$:*

$$\max_{w \in T} s^T w - s^T \hat{w} \le (3.5\vartheta + 1.25)\left(s^T \hat{w} - \min_{w \in T} s^T w\right) \tag{16}$$

and

$$\max_{w \in T} s^T w - s^T \hat{w} \ge \left(\frac{1}{3.5\vartheta + 1.25}\right)\left(s^T \hat{w} - \min_{w \in T} s^T w\right) \ . \tag{17}$$

3 Complexity of Computing a Feasible Solution of G_P

In this section we present and analyze algorithm FEAS for computing a feasible solution of G_P using the barrier method in equation-solving mode. The output of algorithm FEAS will be a point $\bar{x}$ that will satisfy $A\bar{x} = b, \bar{x} \in P$, as well as several other important properties that will be described in this section. The computed point $\bar{x}$ will also be used to initiate algorithm OPT, to be presented in Section 4, that will start from $\bar{x}$ and will then compute an ϵ-optimal solution of G_P.

Algorithm FEAS will employ the barrier method in equation-solving mode to solve the following optimization problem denoted by P_1:

$$
\begin{array}{lll}
P_1: & t^* := \text{maximum}_{z,t,\theta} & t \\
& \text{s.t.} & Az = (b - Ax^r)\theta \\
& & z + \theta x^r - t x^0 \in (\theta - t)P \\
& & \theta \leq 1 \\
& & \theta \geq t \\
& & t \geq -2 \\
& & \|z\| \leq 1
\end{array}
\tag{18}
$$

where x^r and x^0 are the pre-specified reference point and interior-point, respectively, and where we use the notation αP as follows:

$$
\alpha P := \begin{cases} \{x \in X \mid x = \alpha w \text{ for some } w \in P\} & \text{if } \alpha > 0 \\ \text{rec}P & \text{if } \alpha = 0 \\ \emptyset & \text{if } \alpha < 0 \end{cases}
\tag{19}
$$

and where $\text{rec}P$ denotes the recession cone of P.

Note that P_1 is an instance of the optimization problem OP of Section 2, with $w = (z, t, \theta)$, $\bar{c} = (0, 1, 0)$, etc. We will employ the barrier method in equation-solving mode to solve P_1 for a feasible solution $(\hat{z}, \hat{t}, \hat{\theta})$ with objective value $\bar{c}^T w = \delta := 0$, i.e., for a feasible solution $(\hat{z}, \hat{t}, \hat{\theta})$ of P_1 for which $\hat{t} = 0$. We will then convert this solution to a feasible solution of G_P via the elementary transformation:

$$
\bar{x} := \frac{\hat{z}}{\hat{\theta}} + x^r
\tag{20}
$$

(where the algorithm will ensure that $\hat{\theta} > 0$ and so (20) will be legal). Note that if $(\hat{z}, 0, \hat{\theta})$ is feasible for P_1 and $\hat{\theta} > 0$, then it is straightforward to verify that $\bar{x}$ from (20) satisfies $A\bar{x} = b, \bar{x} \in P$.

In order to solve P_1 for a solution $(z, t, \theta) = (\hat{z}, 0, \hat{\theta})$, we first must construct a suitable barrier function for P_1. Define:

$$
S_0 := \left\{(z, t, \theta) \mid z + \theta x^r - t x^0 \in (\theta - t)P, \theta \leq 1, \theta \geq t, t \geq -2, \|z\| \leq 1\right\}
$$

and let S_1 denote the feasible region of P_1, namely:

$$
S_1 := S_0 \cap \{(z, t, \theta) \mid Az = (b - Ax^r)\theta\} \ ,
\tag{21}
$$

and consider the barrier function for $\text{int}S_0$:

$$F(z,t,\theta) := -\ln(t+2) - \ln(1-\theta) + F_{\|\ \|}(z) + 400\left[F_P\left(\frac{z+\theta x^r - tx^0}{\theta - t}\right) - 2\vartheta_P \ln(\theta - t)\right] . \tag{22}$$

Define:

$$\vartheta := 2 + \vartheta_{\|\ \|} + 800\vartheta_P \ .$$

Then from the barrier calculus, and in particular from Proposition 5.1.4 of [5], we have:

Proposition 3.1 *$F(z,t,\theta)$ is a ϑ-self-concordant barrier for intS_0, and its restriction to S_1 is a ϑ-self-concordant barrier for relintS_1.* ▮

Note that $\vartheta = O\left(\vartheta_P + \vartheta_{\|\ \|}\right)$.

We will initiate the barrier method at the point $(z,t,\theta)^0 := (0,-1,0)$. Thus our algorithm for finding a feasible point of G_P is as follows:

Algorithm FEAS: Construct problem P_1 and the barrier function (22). Using the starting point $(z,t,\theta)^0 := (0,-1,0)$, apply the barrier method, in equation-solving mode, to compute a feasible solution $(\hat{z},\hat{t},\hat{\theta})$ of P_1 that satisfies $\hat{t} = \delta = 0$. If such a solution is computed, then compute $\bar{x}$ using (20).

We now examine the complexity of algorithm FEAS. To do so, we first bound the symmetry of S_1 at the point $(z,t,\theta)^0$:

Proposition 3.2 *$(z,t,\theta)^0 := (0,-1,0)$ is a feasible solution of P_1, and*

$$\mathrm{sym}\left((z,t,\theta)^0, S_1\right) \ge \frac{\min\{\mathrm{dist}(x^0,\partial P),1\}}{3+2\|x^0-x^r\|} \ .$$

The proof of Proposition 3.2 is deferred to the end of the section.

We will also need the following relationship between the optimal value of P_1 and g :

Proposition 3.3 *Let t^* denote the optimal value of P_1. Then*

$$\left(\min\left\{\mathrm{dist}(x^0,\partial P),1\right\}\right)\cdot g \le \frac{1}{t^*} \le g\left(g+1+\|x^0-x^r\|\right) \ .$$

The proof of Proposition 3.3 is also deferred to the end of the section.

We next examine the range of the objective function value of P_1. Because

$$z^u = \max\{t \mid (z, t, \theta) \in S_1\} = t^* \le 1$$

(from the constraints $t \le \theta \le 1$ of P_1) and

$$z^l = \min\{t \mid (z, t, \theta) \in S_1\} \ge -2 \ ,$$

we have

$$R := \max\{t \mid (z, t, \theta) \in S_1\} - \min\{t \mid (z, t, \theta) \in S_1\} \le 3 \ . \tag{23}$$

Finally, observe that $z^l \le -1$ (from Proposition 3.2), and so with $\delta := 0$ we have:

$$\min\{z^u - \delta, \delta - z^l\} \ge \min\{t^*, 1\} = t^* \ge \frac{1}{g\left(g + 1 + \|x^0 - x^r\|\right)} \ , \tag{24}$$

where the last inequality is from Proposition 3.3. Combining Propositions 3.2 and 3.1 as well as (23) and (24), and using (13) and Proposition A.1 of the Appendix, we obtain the following:

Theorem 3.1 *Under Assumption A, algorithm FEAS will compute a feasible solution* $(\hat{z}, \hat{t}, \hat{\theta})$ *of* P_1 *and by transformation a feasible solution* $\bar{x}$ *of* G_P*, in at most:*

$$O\left(\sqrt{\vartheta_P + \vartheta_{\|\ \|}} \ln\left(\vartheta_P + \vartheta_{\|\ \|} + \frac{1}{\min\{\mathrm{dist}(x^0, \partial P), 1\}} + \|x^0 - x^r\| + g\right)\right)$$

iterations of Newton's method. ∎

Note that the complexity bound in Theorem 3.1 is vacuously true even if Assumption A does not hold; in that case the optimal value t^* of P_1 will satisfy $t^* = 0$ (assuming G_P has a feasible solution), and algorithm FEAS will not halt. However, $g = +\infty$ in this case, and so the complexity bound of the theorem remains valid but is meaningless in this case.

Given the output $(\hat{z}, \hat{t}, \hat{\theta})$ and the transformed point $\bar{x}$ given in (20) from algorithm FEAS, define the following set:

$$S_2 := \left\{ x \in X \mid Ax = b, x \in P, \|x - x^r\| \le \frac{1}{\hat{\theta}} \right\} \ . \tag{25}$$

The following characterizes important properties of $(\hat{z}, \hat{t}, \hat{\theta})$ and $\bar{x}$ that will be used in the analysis in Section 4:

Lemma 3.1 *Suppose that Assumption A is satisfied, and let $(\hat{z}, \hat{t}, \hat{\theta})$ and $\bar{x}$ be the output of algorithm FEAS. Then*

(i) $\bar{x} \in S_2$, *and* $\mathrm{sym}(\bar{x}, S_2) \ge \frac{1}{3.5\vartheta + 1.25}$

(ii) $\frac{1}{\hat{\theta}} \le (3.5\vartheta + 2.25)\, g$

(iii) $\|\bar{x} - x^r\| \le (3.5\vartheta + 2.25)\, g$

(iv) $1 - \hat{\theta} \ge \frac{1}{3.5\vartheta + 2.25}$

(v) $1 - \|\hat{z}\| \ge \frac{1}{3.5\vartheta + 2.25}$

(vi) $B\left(\bar{x}, \frac{1}{(3.5\vartheta + 3.25)^2 \cdot g}\right) \subset P$.

Proof of Lemma 3.1: It follows from Proposition 3.3 that $t^* > 0$, and so from the barrier method the point $(\hat{z}, \hat{t}, \hat{\theta}) = (\hat{z}, 0, \hat{\theta})$ will satisfy $A\hat{z} = (b - Ax^r)\hat{\theta}$, $\hat{z} + \hat{\theta}x^r \in \mathrm{int}(\hat{\theta}P)$, $\hat{\theta} < 1$, $\hat{\theta} > 0$, and $\|\hat{z}\| < 1$. Then $\hat{\theta} > 0$ validates (20), and also $A\bar{x} = b$, $\bar{x} \in P$, and $\|\bar{x} - x^r\| = \frac{1}{\hat{\theta}}\|\hat{z}\| < \frac{1}{\hat{\theta}}$, whereby we see that $\bar{x} \in S_2$. This proves the first assertion of (i).

Let $T_1 := S_1 \cap \{(z, t, \theta) |\ t = 0\}$ where recall that S_1 is the feasible region of P_1, see (21), and let $T_2 := S_1 \cap \{(z, t, \theta) |\ t = 0, \theta = \hat{\theta}\}$. Then from (14), $(\hat{z}, \hat{t}, \hat{\theta}) = (\hat{z}, 0, \hat{\theta})$ will satisfy

$$\mathrm{sym}((\hat{z}, \hat{t}, \hat{\theta}), T_1) \ge \frac{1}{3.5\vartheta + 1.25} \ .$$

Furthermore, since T_2 is the intersection of T_1 with an affine space passing through $(\hat{z}, \hat{t}, \hat{\theta})$, then it also follows that

$$\mathrm{sym}((\hat{z}, \hat{t}, \hat{\theta}), T_2) \ge \frac{1}{3.5\vartheta + 1.25} \ .$$

Also, the affine transformation $(z, t, \theta) \mapsto (\frac{z}{\hat{\theta}} + x^r)$ maps T_2 onto S_2 (see (25)) and maps $(\hat{z}, \hat{t}, \hat{\theta})$ to $\bar{x}$, and since symmetry is preserved under affine transformations, it follows that

$$\mathrm{sym}(\bar{x}, S_2) \ge \frac{1}{3.5\vartheta + 1.25} \ , \tag{26}$$

completing the proof of (i).

Let $\delta := \min\{\|x - x^r\| \ |Ax = b, x \in P\}$, and observe that under the transformation $x = \frac{z}{\theta} + x^r$ that

$$\begin{aligned}\max_{(z,t,\theta)\in T_1} \theta &= \max\{\theta \mid Ax = b, x \in P, \theta \le 1, \theta\|x - x^r\| \le 1\}\\ &= \max\{\theta \mid Ax = b, x \in P, \theta \le \tfrac{1}{\max\{1,\|x-x^r\|\}}\}\\ &= \tfrac{1}{\max\{\delta,1\}}\end{aligned}$$

and so

$$\max_{(z,t,\theta)\in T_1} \theta = \frac{1}{\max\{\delta, 1\}} \ge \frac{1}{\max\{g, 1\}} = \frac{1}{g} \quad ,$$

since $g \ge \delta$ and $g \ge 1$. Noting as well that $\min_{(z,t,\theta)\in T_1} \theta = 0$, it follows from (16) that

$$\begin{aligned}(3.5\vartheta + 1.25)\hat{\theta} &= (3.5\vartheta + 1.25)(\hat{\theta} - \min_{(z,t,\theta)\in T_1} \theta)\\ &\ge \max_{(z,t,\theta)\in T_1} \theta - \hat{\theta}\\ &\ge \tfrac{1}{g} - \hat{\theta} \quad ,\end{aligned}$$

and rearranging yields $\frac{1}{\hat{\theta}} \le (3.5\vartheta + 2.25)g$. This proves (ii). (iii) then follows since $\|\bar{x} - x^r\| = \frac{\|\hat{z}\|}{\hat{\theta}} \le \frac{1}{\hat{\theta}} \le (3.5\vartheta + 2.25)g$. Noting that $\max_{(z,t,\theta)\in T_1} \theta \le 1$ and $\min_{(z,t,\theta)\in T_1} \theta = 0$, it follows from (17) that

$$\begin{aligned}1 - \hat{\theta} \ge \max_{(z,t,\theta)\in T_1} \theta - \hat{\theta} &\ge (\tfrac{1}{3.5\vartheta+1.25})(\hat{\theta} - \min_{(z,t,\theta)\in T_1} \theta)\\ &= \tfrac{\hat{\theta}}{3.5\vartheta+1.25} \quad ,\end{aligned}$$

and rearranging yields $1 - \hat{\theta} \ge \frac{1}{3.5\vartheta+2.25}$, which proves (iv).

We now prove (v). Given $\hat{z}$, there exists $\bar{z} \in X^*$ satisfying $\|\bar{z}\|_* = 1$ and $\bar{z}^T\hat{z} = \|\hat{z}\|$, see Proposition A.3 of the Appendix. Then

$$\begin{aligned}1 - \|\hat{z}\| &\ge \max_{(z,t,\theta)\in T_1} \bar{z}^T z - \bar{z}^T\hat{z}\\ &\ge (\tfrac{1}{3.5\vartheta+1.25})(\bar{z}^T\hat{z} - \min_{(z,t,\theta)\in T_1} \bar{z}^T z)\\ &\ge (\tfrac{1}{3.5\vartheta+1.25})(\|\hat{z}\| - 0) \quad ,\end{aligned}$$

where the second inequality above is from (17), and rearranging yields $1 - \|\hat{z}\| \le \frac{1}{3.5\vartheta+2.25}$, proving (v).

In order to prove (vi), we will use the following claim:

$$\text{there exist } (\tilde{x},\tilde{r}) \text{ satisfying } A\tilde{x}=b,\ B(\tilde{x},\tilde{r})\subset P,\ \|\tilde{x}-x^r\|+\tilde{r}\le \frac{1}{\hat{\theta}} \quad , \tag{27}$$

and

$$\tilde{r} \ge \frac{1}{(3.5\vartheta+3.25)g} \quad . \tag{28}$$

Before proving (27) and (28), we use them to prove part (vi) of the Lemma. From (27) and (28) we have $\tilde{x}\in S_2$ and so from (26) $x^1 := \bar{x} - (\frac{1}{3.5\vartheta+1.25})(\tilde{x}-\bar{x}) \in S_2$. Rearranging this, we obtain

$$\bar{x} = \frac{3.5\vartheta+1.25}{3.5\vartheta+2.25}x^1 + \frac{1}{3.5\vartheta+2.25}\tilde{x} \tag{29}$$

and so $\bar{x}$ is a convex combination of $x^1\in S_2$ and $\tilde{x}\in S_2$. It then follows from (29) that $B\left(\bar{x}, \frac{\tilde{r}}{3.5\vartheta+2.25}\right)\subset P$, which combined with (28) proves (vi).

It remains to prove (27) and (28). Let $(\breve{x},\breve{\rho})$ be an optimal solution of (3). We assume for the rest of the proof without loss of generality that $\breve{\rho}\le 1$. We consider two cases:

Case 1: $(\|\breve{x}-x^r\|+\breve{\rho})\hat{\theta}\le 1$. Let $\tilde{x}=\breve{x}$ and $\tilde{r}=\breve{\rho}$. Then $A\tilde{x}=b$, $B(\tilde{x},\tilde{r})\subset P$, and $\|\tilde{x}-x^r\|+\tilde{r} = \|\breve{x}-x^r\|+\breve{\rho}\le\frac{1}{\hat{\theta}}$. Furthermore, $\tilde{r}=\breve{\rho}\ge\frac{1}{g}\ge\frac{1}{(3.5\vartheta+3.25)g}$, and so (27) and (28) are proved.

Case 2: $(\|\breve{x}-x^r\|+\breve{\rho})\hat{\theta}\ge 1$. Let $\tilde{x}=(1-\beta)\bar{x}+\beta\breve{x}$, where

$$\beta = \frac{1}{1+(3.5\vartheta+2.25)(\hat{\theta}(\|\breve{x}-x^r\|+\breve{\rho})-1)} \quad , \tag{30}$$

and let $\tilde{r}=\beta\breve{\rho}$. Then $\beta\in[0,1]$, and so $A\tilde{x}=b$, and $B(\tilde{x},\tilde{r})\subset P$. Also,

$$\begin{aligned}
\|\tilde{x}-x^r\|+\tilde{r} &= \|(1-\beta)(\bar{x}-x^r)+\beta(\breve{x}-x^r)\|+\beta\breve{\rho} \\
&\le (1-\beta)\|\bar{x}-x^r\|+\beta\|\breve{x}-x^r\|+\beta\breve{\rho} \\
&= (1-\beta)\tfrac{\|\hat{z}\|}{\hat{\theta}}+\beta\|\breve{x}-x^r\|+\beta\breve{\rho} \\
&\le (1-\beta)(\tfrac{1}{\hat{\theta}})(\tfrac{3.5\vartheta+1.25}{3.5\vartheta+2.25})+\beta\|\breve{x}-x^r\|+\beta\breve{\rho} \\
&= \tfrac{1}{\hat{\theta}}
\end{aligned}$$

where the last inequality follows from part (v) of the lemma, and the last equality follows directly from (30). Also,

$$\begin{aligned}
\tilde{r}=\beta\breve{\rho} &= \tfrac{\breve{\rho}}{1+(3.5\vartheta+2.25)(\hat{\theta}(\|\breve{x}-x^r\|+\breve{\rho})-1)} \\
&\ge \tfrac{\breve{\rho}}{1+\|\breve{x}-x^r\|(3.5\vartheta+2.25)} \\
&\ge \tfrac{1}{(3.5\vartheta+3.25)g} \ ,
\end{aligned} \tag{31}$$

where the first inequality arises since $\hat{\theta} \le 1$ and $\breve{\rho} \le 1$, and the second inequality is valid since $g \ge \frac{1}{\breve{\rho}}$ and $g \ge \frac{\|\breve{x}-x^r\|}{\breve{\rho}}$. This then proves (27) and (28) in this case.∎

Proof of Proposition 3.2: Note first that $(z,t,\theta)^0 = (0,-1,0)$ is feasible for P_1. It suffices to show that if $(0+d,\ -1+\alpha,\ 0+\delta)$ is feasible for P_1, then $(0-\beta d,\ -1-\alpha\beta,\ 0-\beta\delta)$ is feasible for P_1, where $\beta = \frac{\tau}{3+2\|x^0-x^r\|}$ and $\tau := \min\{\mathrm{dist}(x^0,\partial P), 1\}$. Note that $\beta \le \frac{1}{3} < \frac{1}{2}$. Since $(0+d,\ -1+\alpha,\ 0+\delta)$ is by presumption feasible for P_1, then $Ad = (b-Ax^r)\delta$, $d+\delta x^r + (1-\alpha)x^0 \in (\delta+1-\alpha)P$, $\delta \le 1$, $\delta \ge -1+\alpha$, $\|d\| \le 1$, $-1+\alpha \ge -2$, and it follows that

$$-2 \le \delta \le 1 \quad \text{and} \quad -1 \le \alpha \le 2\ . \tag{32}$$

Let $(z,t,\theta) := (-\beta d, -1-\alpha\beta, -\beta\delta)$. Then $Az = (b-Ax^r)\theta$, and $\|z\| \le 1$ since $\beta \le 1$. Also $\theta = -\beta\delta \le 2\beta \le 1$ from (32) and $\beta \le \frac{1}{2}$. Next, notice that $\theta - t = -\beta\delta + 1 + \alpha\beta = 1 - \beta(\delta-\alpha) \ge 1 - \frac{1}{2}(1+1) \ge 0$ from (32) and $\beta \le \frac{1}{2}$. Also, $t = -1-\alpha\beta \ge -2$ since $\alpha \le 2$ and $\beta \le \frac{1}{2}$. It remains to prove that $z + \theta x^r - t x^0 \in (\theta - t)P$. To see this, note first that

$$\begin{array}{rcll} \|-\beta d - \delta\beta(x^r-x^0)\| - \alpha\beta + \delta\beta & \le & \beta(\|d\| + 2\|x^r-x^0\|) + 2\beta & (\text{from}(32)) \\ & \le & \beta(3+2\|x^r-x^0\|) & \\ & = & \tau\ . & \end{array}$$

Then

$$\begin{array}{rcll} \frac{\beta}{1+\alpha\beta-\delta\beta}\|-d-\delta(x^r-x^0)\| & \le & \frac{\beta}{1-2\beta}(1+2\|x^r-x^0\|) & (\text{from}(32)) \\ & = & \tau\left(\frac{1+2\|x^r-x^0\|}{3-2\tau+2\|x^r-x^0\|}\right) & \\ & \le & \tau & (\text{since } \tau \le 1). \end{array}$$

Therefore $x^0 + \frac{\beta}{1+\alpha\beta-\delta\beta}(-d-\delta(x^r-x^0)) \in P$ from the definition of τ, and rearranging yields

$$-\beta d - \beta\delta x^r + (1+\alpha\beta)x^0 \ \in\ (1+\alpha\beta-\delta\beta)P,$$

which is the same as $z + \theta x^r - t x^0 \in (\theta - t)P$. This shows that (z,t,θ) is feasible for P_1, and so $\mathrm{sym}((z,t,\theta)^0,\ S_1) \ge \beta$ as desired.∎

Proof of Proposition 3.3: Let $(\breve{x}, \breve{\rho})$ be an optimal solution of (3), and note from (3) that we can assume that $\breve{\rho} \le 1$. From Assumption A, $\breve{\rho} > 0$. Define the following:

$$z = \frac{\breve{x} - x^r}{\max\{\|\breve{x}-x^r\|, 1\}},\quad t = \frac{\beta\breve{\rho}}{\max\{\|\breve{x}-x^r\|, 1\}},\quad \theta = \frac{1}{\max\{\|\breve{x}-x^r\|, 1\}} \tag{33}$$

where

$$\beta = \frac{1}{g+1+\|x^r-x^0\|}. \tag{34}$$

Then $\beta < 1$, since in particular $g \geq 1$, and from (33) we have $Az = (b - Ax^r)\theta$, $\theta \leq 1$, $-2 \leq t \leq \theta$, and $\|z\| \leq 1$. If (z, t, θ) also satisfies

$$z + \theta x^r - tx^0 \in (\theta - t)P, \tag{35}$$

then (z, t, θ) is feasible for P_1, whereby

$$t^* \geq t = \frac{\beta\breve{\rho}}{\max\{\|\breve{x} - x^r\|, 1\}} \geq \frac{\beta}{g} = \frac{1}{g(g + 1 + \|x^r - x^0\|)} \ , \tag{36}$$

proving the second inequality of the proposition. Therefore, to prove the second inequality of the proposition, we must show (35). Note first that

$$\frac{1}{\beta} = g + 1 + \|x^r - x^0\| \geq \|\breve{x} - x^r\| + \breve{\rho} + \|x^r - x^0\| \geq \breve{\rho} + \|\breve{x} - x^0\|, \tag{37}$$

and

$$\frac{t}{\theta - t}\|\breve{x} - x^0\| = \frac{\beta\breve{\rho}}{1 - \beta\breve{\rho}}\|\breve{x} - x^0\| = \frac{\breve{\rho}}{\frac{1}{\beta} - \breve{\rho}}\|\breve{x} - x^0\| \leq \breve{\rho}, \tag{38}$$

where the last inequality above follows from (37), and $\theta > t$ (since $\beta < 1$ and $\breve{\rho} \leq 1$), and so

$$\frac{z + \theta x^r - tx^0}{\theta - t} = \breve{x} + \frac{t}{\theta - t}(\breve{x} - x^0) \in P,$$

since from (38) we have $\frac{t}{\theta - t}\|\breve{x} - x^0\| \leq \breve{\rho}$ and $B(\breve{x}, \breve{\rho}) \subset P$. Therefore $z + \theta x^r - tx^0 \in (\theta - t)P$, and we have proven the second inequality of the proposition.

To prove the first inequality of the proposition, let (z^*, t^*, θ^*) be any optimal solution of P_1, and note from (36) that $t^* > 0$. Therefore $\theta^* > 0$, and define

$$x = \frac{z^*}{\theta^*} + x^r, \quad r = \frac{t^*\tau}{\theta^*} \ , \tag{39}$$

where $\tau := \min\{\mathrm{dist}(x^0, \partial P), 1\}$. Then $Ax = b$, and for any d satisfying $\|d\| \leq 1$ we have $x^0 + \tau d \in P$ from the definition of τ. Also, $z^* + \theta^* x^r - t^* x^0 \in (\theta^* - t^*)P$. If $\theta^* > t^*$, then

$$x + rd = \left(\frac{\theta^* - t^*}{\theta^*}\right)\frac{z^* + \theta^* x^r - t^* x^0}{\theta^* - t^*} + \left(\frac{t^*}{\theta^*}\right)(x^0 + \tau d) \in P \ .$$

If $\theta^* = t^*$, then $z^* + \theta^* x^r - t^* x^0 \in \mathrm{rec}P$, and so

$$x + rd = \frac{z^* + \theta^* x^r - t^* x^0}{\theta^*} + (x^0 + \tau d) \in P.$$

In either case, $x + rd \in P$, and so $B(x, r) \subset P$. Therefore

$$g \leq \frac{\max\{\|x^r - x\|, 1\}}{\min\{r, 1\}} = \max\left\{\frac{\|x^r - x\|}{r}, \frac{1}{r}\right\},$$

since $r \le 1$. Now

$$\frac{1}{r} = \frac{\theta^*}{t^*\tau} \le \frac{1}{t^*\tau}$$

and $\frac{\|x-x^r\|}{r} = \frac{\|z^*\|}{t^*\tau} \le \frac{1}{t^*\tau}$, so $g \le \frac{1}{t^*\tau}$, proving the left inequality of the proposition.∎

Certain ideas and constructs used in the results of this section arose from or were inspired directly from Section 3 of Renegar [7], including the idea of solving phase-I by using the barrier method in equation-solving mode, transforming to the original problem via an elementary projective transformation, and establishing key properties of the output of the algorithm (upper bounds on norms and lower bounds on distances from constraints) using symmetry properties of the output of the barrier method.

4 Complexity of Computing an ϵ-optimal Solution of G_P

In this section we present algorithm OPT for computing an ϵ-optimal solution of G_P initiated at the point $\bar{x}$, using the barrier method in optimization-mode, where $\bar{x}$ is the output of algorithm FEAS.

Using $\bar{x}$ as a starting point, we will modify G_P slightly by adding a level-set constraint of the form "$c^T x \le c^T \bar{x} + \bar{s}$" to G_P for some suitably chosen positive scalar offset $\bar{s}$ which will then render the point $\bar{x}$ in the interior of the half-space generated by the constraint $c^T x \le c^T \bar{x} + \bar{s}$. The question then arises as to how to choose the offset $\bar{s}$. One would think that $\bar{s}$ should be chosen proportional to the norm of c :

$$\|c\|_* := \max_{w} \left\{ c^T w \mid \|w\| \le 1 \right\} \ . \tag{40}$$

However, because the objective function $c^T x$ of G_P differs only by a constant from the modified objective function $(c - A^T\pi)^T x$ over the feasible region of G_P for any given value of π, we must be mindful of the equations $Ax = b$. From this perspective, it is natural to choose $\bar{s}$ proportional to:

$$\tilde{s} := \max_{w} \left\{ c^T w \mid \|w\| \le 1, Aw = 0 \right\} \ , \tag{41}$$

and note $\tilde{s}$ is the maximum objective value over the unit ball, "reduced" by the subspace constraints $Aw = 0$, and so $\tilde{s}$ is the norm of the linear functional $c^T x$ over the vector subspace of solutions to $Aw = 0$. However, even for otherwise computationally tractable norms such as the L_∞ norm in $\Re^n$, the computation of $\tilde{s}$ is not trivial; in fact its computation is a linear program for the L_∞ norm. We therefore will instead use the information inherent in the barrier function $F_{\|\ \|}(\cdot)$ for the unit ball as a proxy for the $\|\cdot\|$ in constructing the offset $\bar{s}$. Let $H(0)$ denote the Hessian matrix of $F_{\|\ \|}(\cdot)$ at $x = 0$, and define:

$$s_2 := \max_w \left\{ c^T w \mid Aw = 0, w^T H(0) w \leq 1 \right\} \ , \tag{42}$$

and note that s_2 admits a closed form solution when $\text{rank}(A) = m$:

$$s_2 = \sqrt{c^T H(0)^{-1} c - c^T H(0)^{-1} A^T \left(A H(0)^{-1} A^T \right)^{-1} A H(0)^{-1} c}$$

It will be convenient for our purposes to determine $\bar{s}$ proportional to s_2 as follows:

$$\bar{s} := \left(\frac{6\vartheta_{\|\ \|} + 1}{\sqrt{2}} \right) s_2 \ , \tag{43}$$

and we consider the following amended version of G_P:

$$\begin{array}{lrll} P_{\bar{s}}: & z^* := \text{minimum}_x & c^T x & \\ & \text{s.t.} & Ax = b & \\ & & x \in P & \\ & & c^T x \leq c^T \bar{x} + \bar{s} \ . & \end{array} \tag{44}$$

Note that since $\bar{x}$ is feasible for G_P, then $\bar{x}$ is also feasible for $P_{\bar{s}}$, and $P_{\bar{s}}$ and G_P have the same optimal objective function value and the same set of optimal solutions. (The idea of solving phase-II by adding a level set constraint of the objective function was used by Renegar [7], but without an explicit construction for computing the offset $\bar{s}$.)

In order to apply the barrier method (in optimization mode) to compute an ϵ-optimal solution of $P_{\bar{s}}$, we need to specify the barrier function to be used. The obvious choice is:

$$F(x) := F_P(x) - \ln\left(c^T \bar{x} + \bar{s} - c^T x \right) \ , \tag{45}$$

whose complexity value is at most $\vartheta_P + 1$.

It is easily seen that $\bar{s} \geq 0$, and that $\bar{s} > 0$ except when the objective function $c^T x$ is constant over the entire feasible region of G_P, in which case $\bar{x}$ is then an optimal solution of G_P. In light of this observation, the algorithm for computing an ϵ-optimal solution of G_P is as follows:

Algorithm OPT: Compute $\bar{s}$ and construct problem $P_{\bar{s}}$ and the barrier function (45), using (42) and (43). If $\bar{s} > 0$, then using the starting point $\bar{x}$ (where $\bar{x}$ is the output of algorithm FEAS), apply the barrier method, in optimization mode, to compute an ϵ-optimal solution $\hat{x}$ of $P_{\bar{s}}$. Otherwise, $\bar{s} = 0$, and $\bar{x}$ is an optimal solution of G_P and no further computation is required.

We now examine the complexity of algorithm OPT. We first relate the offset $\bar{s}$ to the theoretical quantity $\tilde{s}$ defined in (41) in the following proposition, whose proof is deferred to the end of this section:

Proposition 4.1 $\left(\frac{1}{6\vartheta_{\|\ \|}+1}\right)\bar{s} \le \tilde{s} \le \bar{s}$. $\blacksquare$

Define S_3 to be the feasible region of $P_{\bar{s}}$, namely

$$S_3 := \left\{x \mid Ax = b, x \in P, c^T x \le c^T \bar{x} + \bar{s}\right\} \ . \tag{46}$$

The following result, whose proof is also deferred to the end of the section, presents a bound on the symmetry of S_3 at the point $\bar{x}$:

Lemma 4.1

$$\frac{1}{\mathrm{sym}\left(\bar{x}, S_3\right)} \le (3.5\vartheta + 2.25)^2\, \bar{h}$$

where $\bar{h}$ is given by:

$$\bar{h} := 3D_\epsilon + (3.5\vartheta + 2.25)\, g + 4gD_\epsilon\,(g + D_\epsilon)\left[(3.5\vartheta + 2.25)\, g + D_\epsilon + 6\vartheta_{\|\ \|} + 1\right]\max\left\{\frac{\tilde{s}}{\epsilon}, 1\right\} \ . \tag{47}$$

$\blacksquare$

Remark 4.1 $\frac{1}{\mathrm{sym}(\bar{x},S_3)}$ *is bounded from above by a polynomial in $g, \vartheta_P, \vartheta_{\|\ \|}, D_\epsilon$, and* $\max\left\{\frac{\tilde{s}}{\epsilon}, 1\right\}$.

We now state and prove a complexity bound on algorithm OPT:

Theorem 4.1 *Under Assumption A, and starting from the point $\bar{x}$ computed by algorithm FEAS, algorithm OPT will compute an ϵ-optimal solution of G_P in at most:*

$$O\left(\sqrt{\vartheta_P}\ln\left(g + D_\epsilon + \vartheta_P + \vartheta_{\|\ \|} + \max\left\{\frac{\tilde{s}}{\epsilon}, 1\right\}\right)\right)$$

iterations of Newton's method. ▮

Remark 4.2 *Note that we could replace $\tilde{s}$ by $\|c\|_*$ in the iteration bound of Theorem 4.1, since $\tilde{s} \leq \|c\|_*$.*

Note also that the complexity bound in Theorem 4.1 is vacuously true even if $D_\epsilon = +\infty$; in this case algorithm OPT will not halt but the complexity bound of the theorem remains trivially valid.

Proof of Theorem 4.1: To prove the theorem, we invoke the complexity bound for the barrier method in optimization mode stated in (11). The barrier $F(x)$ for $P_{\bar{s}}$ defined in (45) has complexity value $\vartheta \leq \vartheta_P + 1 = O(\vartheta_P)$. The starting point $\bar{x}$ has symmetry bounded by Lemma 4.1:

$$\frac{1}{\mathrm{sym}(\bar{x}, S_3)} \leq (3.5\vartheta + 2.25)^2 \bar{h} \quad ,$$

this bound being a polynomial in g, ϑ_P, $\vartheta_{\|\ \|}$, D_ϵ, and $\max\{\frac{\tilde{s}}{\epsilon}, 1\}$, see Remark 4.1. The range R of the objective function of $P_{\bar{s}}$ is bounded as follows:

$$\begin{array}{rcll} R & \leq & c^T\bar{x} + \bar{s} - z^* & \\ & = & c^T\bar{x} - c^Tx^* + \bar{s} & \text{(where } x^* \text{ solves } G_p) \\ & \leq & \tilde{s}(\|\bar{x} - x^r\| + \|x^* - x^r\|) + \bar{s} & \text{(from Proposition A.2)} \\ & \leq & \tilde{s}(\|\bar{x} - x^r\| + \|x^* - x^r\|) + \tilde{s}(6\vartheta_{\|\ \|} + 1) & \text{(from Proposition 4.1)} \\ & \leq & \tilde{s}\left[(3.5\vartheta + 2.25)g + D_\epsilon + 6\vartheta_{\|\ \|} + 1\right] & \end{array} \tag{48}$$

where the last inequality uses (7) and Lemma 3.1. Therefore $\frac{R}{\epsilon}$ is bounded above by a polynomial in $\max\{\frac{\tilde{s}}{\epsilon}, 1\}$, ϑ_P, $\vartheta_{\|\ \|}$, g, and D_ϵ. Combining all of these terms and using Proposition A.1 of the Appendix, we obtain the complexity bound of

$$O\left(\sqrt{\vartheta_P}\ln\left(g + D_\epsilon + \vartheta_p + \vartheta_{\|\ \|} + \max\left\{\frac{\tilde{s}}{\epsilon}, 1\right\}\right)\right)$$

iterations of Newton's method.∎

We now prove Proposition 4.1 and Lemma 4.1. The proof of Proposition 4.1 is a straightforward application of barrier calculus:

Proof of Proposition 4.1: Since $F_{\|\ \|}(x)$ is a $\vartheta_{\|\ \|}$-self-concordant barrier for $B(0,1) = \{x \mid \|x\| \le 1\}$, then $\bar{F}(x) := F_{\|\ \|}(x) + F_{\|\ \|}(-x)$ is a $2\vartheta_{\|\ \|}$-self-concordant barrier for $B(0,1)$, whose analytic center is $x^c = 0$, and note that the Hessian of $\bar{F}(x)$ at $x = 0$ is $2H(0)$ where $H(x)$ is the Hessian of $F_{\|\ \|}(\cdot)$ at x. Then from Proposition 2.3.2 of [5] it follows that

$$\left\{x \mid \sqrt{x^T(2H(0))x} \le 1\right\} \subset B(0,1) \subset \left\{x \mid \sqrt{x^T(2H(0))x} \le 3(2\vartheta_{\|\ \|}) + 1\right\}. \tag{49}$$

From this it then follows that $\frac{1}{\sqrt{2}}s_2 \le \tilde{s} \le \frac{(6\vartheta_{\|\ \|}+1)}{\sqrt{2}}s_2$, and therefore the result follows from (43).∎

The proof of the symmetry bound in Lemma 4.1 relies primarily on the following "level set growth" lemma, which bounds the sizes of the objective function level sets of G_P. Let L_α denote the level set of the objective function of G_P at value α, namely:

$$L_\alpha := \{x \in X \mid Ax = b, x \in P, c^Tx \le \alpha\} \ .$$

The following lemma bounds the sizes of $L_{\alpha+t}$ for $t \ge 0$ in terms of the size of L_α and other geometric quantities:

Lemma 4.2 *Suppose that $\tilde{x} \in L_\alpha$ and $B(\tilde{x}, \tilde{r}) \subset P$ for some $\tilde{r} > 0$, and let $Q \ge \max\{\|x - \tilde{x}\| \mid x \in L_\alpha\}$. Then for all $t \ge 0$, we have*

$$x \in L_{\alpha+t} \quad \Rightarrow \quad \|x - \tilde{x}\| \le Q\left(1 + \frac{2t}{\tilde{s}\cdot\tilde{r}}\right) \ ,$$

where $\tilde{s}$ is defined in (41).

Proof: Given the hypotheses of the lemma, suppose that $x \in L_{\alpha+t}$, whereby $Ax = b$, $x \in P$, and $c^Tx \le \alpha + t$. Define $s^1 := c^Tx - \alpha$. If $s^1 \le 0$, then $c^Tx \le \alpha$, and therefore $\|x - \tilde{x}\| \le Q \le Q(1 + \frac{2t}{\tilde{s}\tilde{r}})$, proving the result. Suppose instead that $s^1 > 0$, and define

$$w := \left(\frac{s^1}{\alpha - c^T\tilde{x} + \tilde{r}\tilde{s} + s^1}\right)(\tilde{x} - \tilde{r}\tilde{c}) + \left(\frac{\alpha - c^T\tilde{x} + \tilde{r}\tilde{s}}{\alpha - c^T\tilde{x} + \tilde{r}\tilde{s} + s^1}\right)x, \tag{50}$$

where $\tilde{c} \in \arg\max_v\{c^T v \mid Av = 0, \ \|v\| \leq 1\}$. Then $\|\tilde{c}\| \leq 1$, and $c^T\tilde{c} = \tilde{s}$. Also $\tilde{x} - \tilde{r}\tilde{c} \in P$, $A(\tilde{x} - \tilde{r}\tilde{c}) = b$, $Ax = b$, $x \in P$, and so it follows from (50) that $Aw = b$, $w \in P$, and $c^T w = \alpha$, since $c^T\tilde{c} = \tilde{s}$. Therefore $w \in L_\alpha$, and so

$$\begin{aligned} Q &\geq \|w - \tilde{x}\| = \| \left(\tfrac{-\tilde{r}s^1}{\alpha - c^T\tilde{x} + \tilde{r}\tilde{s} + s^1}\right)\tilde{c} + \left(\tfrac{\alpha - c^T\tilde{x} + \tilde{r}\tilde{s}}{\alpha - c^T\tilde{x} + \tilde{r}\tilde{s} + s^1}\right)(x - \tilde{x})\| \\ &\geq \|x - \tilde{x}\|\left(\tfrac{\alpha - c^T\tilde{x} + \tilde{r}\tilde{s}}{\alpha - c^T\tilde{x} + \tilde{r}\tilde{s} + s^1}\right) - \tfrac{\tilde{r}s^1}{\alpha - c^T\tilde{x} + \tilde{r}\tilde{s} + s^1} \end{aligned}$$

Therefore

$$\begin{aligned} \|x - \tilde{x}\| &\leq Q\left(1 + \tfrac{s^1}{\alpha - c^T\tilde{x} + \tilde{r}\tilde{s}}\right) + \tfrac{\tilde{r}s^1}{\alpha - c^T\tilde{x} + \tilde{r}\tilde{s}} \\ &\leq Q\left(1 + \tfrac{s^1}{\tilde{r}\tilde{s}}\right) + \tfrac{s^1}{\tilde{s}}. \end{aligned} \tag{51}$$

However, as noted above, $\tilde{x} - \tilde{r}\tilde{c} \in P$ and $A(\tilde{x} - \tilde{r}\tilde{c}) = b$, and $c^T(\tilde{x} - \tilde{r}\tilde{c}) \leq c^T\tilde{x} \leq \alpha$, and so

$$Q \geq \|\tilde{x} - (\tilde{x} - \tilde{r}\tilde{c})\| = \tilde{r}.$$

Therefore from (51) we have

$$\begin{aligned} \|x - \tilde{x}\| &\leq Q\left(1 + \tfrac{s^1}{\tilde{r}\tilde{s}}\right) + \tfrac{s^1}{\tilde{s}}\left(\tfrac{Q}{\tilde{r}}\right) \\ &= Q\left(1 + \tfrac{2s^1}{\tilde{r}\tilde{s}}\right) \\ &\leq Q\left(1 + \tfrac{2t}{\tilde{r}\tilde{s}}\right), \end{aligned}$$

since $s^1 = c^T x - \alpha \leq t$.∎

We will use Lemma 4.2 to bound the size of the feasible region S_3 of the phase-II problem $P_{\bar{s}}$, since $S_3 = L_{c^T\bar{x}+\bar{s}}$ from (46). In order to do so, we need to first show that the ϵ-optimal level set $L_{z^*+\epsilon}$ contains a point $\tilde{x}$ sufficiently far from the boundary of P, which we do now:

Lemma 4.3 *Suppose that D_ϵ is finite. Then there exists $(\tilde{x}, \tilde{r})$ satisfying*

$$\tilde{x} \in L_{z^*+\epsilon} \quad \textit{and} \quad B(\tilde{x}, \tilde{r}) \subset P$$

where

$$\tilde{r} \geq \breve{\rho}\min\left\{1, \frac{\epsilon}{c^T\breve{x} - z^*}\right\}\ , \tag{52}$$

and $(\breve{x}, \breve{\rho})$ is an optimal solution of (3).

Proof: Suppose first that $c^T\breve{x} \le z^* + \epsilon$. Then setting $\tilde{x} = \breve{x}$ and $\tilde{r} = \breve{\rho}$, we have $c^T\tilde{x} \le z^* + \epsilon$, $A\tilde{x} = b$, $B(\tilde{x}, \tilde{r}) \subset P$, and $\tilde{r} = \breve{\rho} = \breve{\rho}\min\{1, \frac{\epsilon}{c^t\breve{x} - z^*}\}$, proving the result. Suppose instead that $c^T\breve{x} > z^* + \epsilon$, and let

$$\tilde{x} = \lambda\breve{x} + (1-\lambda)x^* \quad , \quad \tilde{r} = \lambda\breve{\rho} \quad ,$$

where

$$\lambda = \frac{\epsilon}{c^T\breve{x} - z^*} \quad ,$$

and x^* is an optimal solution of G_P (x^* is guaranteed to exist since D_ϵ is finite by hypothesis). Then $\lambda \in [0,1]$, and so $\tilde{x}$ satisfies $A\tilde{x} = b$, $\tilde{x} \in P$, and by construction of λ we have $c^T\tilde{x} = z^* + \epsilon$. Furthermore $B(\tilde{x}, \tilde{r}) \subset P$, and $\tilde{r} = \lambda\breve{\rho} = \frac{\breve{\rho}\epsilon}{c^T\breve{x} - z^*} = \breve{\rho}\min\left\{1, \frac{\epsilon}{c^T\breve{x} - z^*}\right\}$, proving the result.∎

We now combine Lemma 4.2 and Lemma 4.3 to bound the size of feasible region S_3:

Lemma 4.4 *Suppose that $x \in S_3$. Then*

$$\|x - \bar{x}\| \le \bar{h}$$

where $\bar{h}$ is defined in (47).

Proof of Lemma 4.4: Suppose that $x \in S_3$, and let $(\tilde{x}, \tilde{r})$ be as described in Lemma 4.3. Then

$$\begin{aligned} \|x - \bar{x}\| &\le \|x - \tilde{x}\| + \|\tilde{x} - x^r\| + \|x^r - \bar{x}\| \\ &\le \|x - \tilde{x}\| + D_\epsilon + (3.5\vartheta + 2.25)g \end{aligned} \tag{53}$$

where the last inequality uses (7) and Lemma 3.1. It thus remains to bound $\|x - \tilde{x}\|$. To this end, we will invoke Lemma 4.2 with $\alpha = z^* + \epsilon$. Then $Q := 2D_\epsilon$ satisfies

$$\begin{aligned} & \max_x\{\|x - \tilde{x}\| \mid Ax = b,\ x \in P,\ c^Tx \le \alpha\} \\ \le\ & \max_x\{\|x - x^r\| + \|x^r - \tilde{x}\| \mid Ax = b,\ x \in P,\ c^Tx \le z^* + \epsilon\} \\ \le\ & 2D_\epsilon = Q\ , \end{aligned}$$

and so $\tilde{x}$, $\tilde{r}$, α, and Q satisfy the hypotheses of Lemma 4.2. Now let $t := [c^T\bar{x} + \bar{s} - \alpha]^+$, and so $\alpha + t \ge c^T\bar{x} + \bar{s}$. Then if $x \in S_3$, x also satisfies $Ax = b$, $x \in P$, $c^Tx \le \alpha + t$, and so from Lemma 4.2 we have:

$$\begin{array}{rcl}\|x-\tilde{x}\| & \leq & Q(1+\frac{2t}{\tilde{s}\tilde{r}}) \\ & = & 2D_\epsilon + \frac{4D_\epsilon t}{\tilde{s}\tilde{r}} \\ & \leq & 2D_\epsilon + \frac{4D_\epsilon t}{\tilde{s}\min\{\breve{\rho},1\}}\max\{1,\frac{c^T\breve{x}-z^*}{\epsilon}\} \qquad \text{(from Lemma 4.3)} \\ & \leq & 2D_\epsilon + \frac{4D_\epsilon(c^T\bar{x}-z^*+\bar{s})}{\tilde{s}\min\{\breve{\rho},1\}}\max\{1,\frac{c^T\breve{x}-z^*}{\epsilon}\}\ .\end{array} \tag{54}$$

But now recall from (48) that

$$c^T\bar{x}-z^*+\bar{s} \ \ \leq \ \ \tilde{s}\left[(3.5\vartheta+2.25)g+D_\epsilon+6\vartheta_{\|\ \|}+1\right]\ . \tag{55}$$

We also bound $c^T\breve{x}-z^*$ using Proposition A.2:

$$c^T\breve{x}-z^* = c^T\breve{x}-c^Tx^* \leq \tilde{s}(\|\breve{x}-x^r\|+\|x^*-x^r\|) \leq \tilde{s}(g+D_\epsilon)\ . \tag{56}$$

Substituting (48) and (56) into (54) we obtain

$$\begin{array}{rcl}\|x-\tilde{x}\| & \leq & 2D_\epsilon + \frac{4D_\epsilon[(3.5\vartheta+2.25)g+D_\epsilon+6\vartheta_{\|\ \|}+1]\max\left\{1,\frac{\tilde{s}(g+D_\epsilon)}{\epsilon}\right\}}{\min\{\breve{\rho},1\}} \\ & \leq & 2D_\epsilon+4D_\epsilon g[(3.5\vartheta+2.25)g+D_\epsilon+6\vartheta_{\|\ \|}+1](g+D_\epsilon)\max\left\{1,\frac{\tilde{s}}{\epsilon}\right\}\ .\end{array} \tag{57}$$

Finally, substituting (57) into (53) we obtain the desired bound.∎

Last of all, we use the bound in Lemma 4.4 to prove the bound on the symmetry of S_3 of Lemma 4.1:

Proof of Lemma 4.1: Let $\beta := \frac{1}{(3.5\vartheta+2.25)^2\bar{h}}$. For any v satisfying $\bar{x}+v\in S_3$, we must show that $\bar{x}-\beta v\in S_3$. To do so, we must show that $A(\bar{x}-\beta v)=b$, $\bar{x}-\beta v\in P$, and $c^T(\bar{x}-\beta v)\leq c^T\bar{x}+\bar{s}$. Note that $\bar{x}\in S_3$ and $\bar{x}+v\in S_3$ imply that $Av=0$, and so $A(\bar{x}-\beta v)=b$. Also, from Lemma 4.4, we have $\|v\|\leq\bar{h}$. And with $\alpha := \min\left\{1,\frac{1-\|\hat{z}\|}{\hat{\theta}\bar{h}}\right\}$ where $\hat{z},\hat{\theta}$ are part of the output of algorithm FEAS, we have $\bar{x}+\alpha v\in P$ (since $\bar{x}+v\in P$ and $\alpha\in(0,1]$). Observe that

$$\|\bar{x}+\alpha v-x^r\| \leq \|\bar{x}-x^r\|+\alpha\bar{h} \leq \frac{\|\hat{z}\|}{\hat{\theta}}+\frac{1-\|\hat{z}\|}{\hat{\theta}} = \frac{1}{\hat{\theta}}\ ,$$

and so $\bar{x}+\alpha v\in S_2$ (see(25)). Then from Lemma 3.1 we have $\bar{x}-\frac{\alpha}{3.5\vartheta+1.25}v\in S_2$, and note that

$$\begin{array}{rcll} \alpha = \min\left\{1, \frac{1-\|\hat{z}\|}{\hat{\theta}\bar{h}}\right\} & \geq & \min\left\{1, \frac{1}{(3.5\vartheta+2.25)\hat{\theta}\bar{h}}\right\} & \text{(from Lemma 3.1)} \\ & \geq & \min\left\{1, \frac{1}{\bar{h}(3.5\vartheta+2.25)}\right\} & (\text{since } \hat{\theta} \leq 1) \\ & = & \frac{1}{\bar{h}(3.5\vartheta+2.25)} \ . & \end{array} \tag{58}$$

Therefore $\frac{\alpha}{3.5\vartheta+1.25} \geq \frac{1}{\bar{h}(3.5\vartheta+2.25)^2} = \beta$, and so $\bar{x}-\beta v \in S_2$, whereby $\bar{x}-\beta v \in P$. Finally, note that

$$\begin{array}{rcll} c^T(\bar{x}-\beta v) & \leq & c^T\bar{x} + \beta\tilde{s}\|v\| & \\ & \leq & c^T\bar{x} + \tilde{s}\beta\bar{h} & \\ & \leq & c^T\bar{x} + \bar{s}\beta\bar{h} & \text{(from Proposition 4.1)} \\ & \leq & c^T\bar{x} + \bar{s} \quad . & (\text{since } \beta\bar{h} < 1) \end{array}$$

Therefore $\bar{x}-\beta v \in S_3$, and the result is proved.▮

5 On Natural Norms, and Condition-Number Complexity

5.1 Two Natural Norms on X

In this subsection we briefly discuss two norms on X that arise naturally based on x^0 and P.

The First Norm. For the given point $x^0 \in \text{int}P$, define the set B_{x^0}:

$$B_{x^0} := \left[P - x^0\right] \cap \left[x^0 - P\right] = \left\{v \mid x^0 + v \in P, x^0 - v \in P\right\} .$$

Then B_{x^0} is the largest symmetric set B for which $x^0 + B \subset P$. Under the assumption that P contains no line, B_{x^0} will be compact, convex, and contain the origin in its interior, and so can be used as the unit ball of a norm. Indeed this norm is constructed as follows:

$$\begin{array}{rll} \|v\|_{x^0} := & \min_\alpha & \alpha \\ & \text{s.t.} & x^0 + \frac{1}{\alpha}v \in P \\ & & \\ & & x^0 - \frac{1}{\alpha}v \in P \ . \end{array}$$

(The norm $\|\cdot\|_{x^0}$, in either explicit or implicit form, appears throughout much of the analysis in [5].) Under $\|\cdot\|_{x^0}$, it is easily shown that $\tau(x^0) = \operatorname{dist}(x^0, \partial P) = 1$, and so the explicit dependence of the complexity bounds in Theorems 3.1 and 4.1 on $\operatorname{dist}(x^0, \partial P)$ disappears. Also, we can construct a barrier function for the unit ball B_{x^0} using the barrier $F_P(\cdot)$ for P:

$$F_{\|\ \|}(v) := F_P(x^0 + v) + F_P(x^0 - v) \ ,$$

whose complexity parameter $\vartheta_{\|\ \|}$ is bounded above as follows:

$$\vartheta_{\|\ \|} \le 2\vartheta_P \ .$$

Therefore the explicit dependence of the complexity bounds in Theorems 3.1 and 4.1 on $\vartheta_{\|\ \|}$ disappears as well. With this choice of norm, then, the combined complexity bound of the algorithms FEAS and OPT becomes:

$$O\left(\sqrt{\vartheta_P}\ln\left(g + D_\epsilon + \vartheta_P + \max\left\{\frac{\tilde{s}}{\epsilon}, 1\right\} + \|x^0 - x^r\|\right)\right)$$

iterations of Newton's method.

(The norm $\|\cdot\|_{x^0}$ is referred to as a generalization of the L_∞-norm because in the case when $P = \Re^n_+$ and $x^0 = e$, we recover the L_∞-norm as $\|v\|_{x^0}$ for $v \in \Re^n$.)

The Second Norm. The second norm we consider is constructed using the barrier function $F_P(\cdot)$ for P. For the given point $x^0 \in \operatorname{int}P$, define the norm

$$\|v\|_{F,x^0} := \sqrt{v^T H_P(x^0) v} \ ,$$

where $H_P(x^0)$ is the Hessian of $F_P(\cdot)$ evaluated at $x = x^0$. It then follows from Theorem 2.1.1 of [5] that $B(x^0, 1) \subset P$ and so $\operatorname{dist}(x^0, \partial P) \ge 1$. Also,

$$F_{\|\ \|}(v) := -\ln\left(1 - v^T H_P(x^0) v\right)$$

is a $\vartheta_{\|\ \|} = 1$-self-concordant barrier function for the unit ball of this norm. Therefore the explicit dependence of the complexity bounds in Theorems 3.1 and 4.1 on $\operatorname{dist}(x^0, \partial P)$ and $\vartheta_{\|\ \|}$ disappear, and like the previous norm, the combined complexity bound of the algorithms FEAS and OPT becomes:

$$O\left(\sqrt{\vartheta_P}\ln\left(g + D_\epsilon + \vartheta_P + \max\left\{\frac{\tilde{s}}{\epsilon}, 1\right\} + \|x^0 - x^r\|\right)\right)$$

iterations of Newton's method.

APPENDIX

Proposition A.1: *If* $a, b > 0$ *then* $\frac{1}{2}\ln 2 + \frac{1}{2}(\ln a + \ln b) \leq \ln(a+b)$. *If in addition* $a, b \geq 1$, *then* $\ln(a+b) \leq \ln 2 + (\ln a + \ln b)$.

Proof: We have $\sqrt{2ab} \leq \sqrt{a^2+b^2+2ab} = a+b$. If also $a, b \geq 1$, then $a+b \leq 2\max\{a,b\} \leq 2\max\{a,b\}\min\{a,b\} = 2ab$. The results then follow by taking logarithms.∎

Proposition A.2: *Let* $\tilde{s} := \max_w\{c^T w \mid Aw = 0, \|w\| \leq 1\}$. *If* x^1, x^2 *satisfy* $Ax^1 = Ax^2 = b$, *then*

$$|c^T x^1 - c^T x^2| \leq \tilde{s}\|x^1 - x^2\| \leq \tilde{s}\left(\|x^1 - x^r\| + \|x^2 - x^r\|\right) \ .$$

Proof: From the definition of $\tilde{s}$, we have

$$|c^T x^1 - c^T x^2| = |c^T(x^1 - x^2)| \leq \tilde{s}\|x^1 - x^2\| \leq \tilde{s}\left(\|x^1 - x^r\| + \|x^2 - x^r\|\right) \ .∎$$

The following proposition is a special case of the Hahn-Banach Theorem; for a short proof of this proposition based on the subdifferential operator, see Proposition 2 of [3].

Proposition A.3: *For every* $z \in X$, *there exists* $\bar{z} \in X^*$ *with the property that* $\|\bar{z}\|_* = 1$ *and* $\|z\| = \bar{z}^T z$. ∎

We end the appendix with a proof of Proposition 1.1.

Proof of Proposition 1.1: Let $(\breve{x}, \breve{\rho})$ solve (3), whereby $\|\breve{x}\|_1 \leq g, \breve{x}_j \geq \frac{1}{g}$, and $\frac{\|\breve{x}\|_1}{\breve{x}_j} \leq g$ for $j = 1, \ldots, n$. Let $(x, x_{n+1}) = \frac{(\breve{x}, 1)}{\|\breve{x}\|_1 + 1}$. Then (x, x_{n+1}) is feasible for each of the $n+1$ maximization linear programs in the definition of σ. For $j = 1, \ldots, n$ we have $x_j = \frac{\breve{x}_j}{\|\breve{x}\|_1 + 1} \geq \frac{1}{2g}$ and $x_{n+1} = \frac{1}{\|\breve{x}\|_1 + 1} \geq \frac{1}{2g}$, and therefore $\sigma \geq \frac{1}{2g}$.

Now let $x^{(j)}, j = 1, \ldots, n+1$, be optimal solutions of the maximization linear programs in the definition of σ, whereby $x_j^{(j)} \geq \sigma$ for $j = 1, \ldots, n+1$, and let $\tilde{x} := \frac{1}{n+1}\sum_{j=1}^{n+1} x^{(j)}$ and $\bar{x} := \frac{(\tilde{x}_1, \ldots, \tilde{x}_n)}{\tilde{x}_{n+1}}$. Then $\tilde{x}_j \geq \frac{\sigma}{n+1}$ for $j = 1, \ldots, n+1$, and $\sum_{j=1}^{n+1} \tilde{x}_j = 1$, and $\bar{x}$ is feasible for (5). We have $\|\bar{x}\|_1 = \sum_{j=1}^{n} \bar{x}_j = \frac{1 - \tilde{x}_{n+1}}{\tilde{x}_{n+1}} \leq \frac{n+1}{\sigma} - 1 \leq \frac{n+1}{\sigma}$, and $\bar{x}_j = \frac{\tilde{x}_j}{\tilde{x}_{n+1}} \geq \frac{\frac{\sigma}{n+1}}{1} = \frac{\sigma}{n+1}$ for $j = 1, \ldots, n$, whereby $\text{dist}(\bar{x}, \partial\Re^n_+) \geq \frac{\sigma}{n+1}$. Also $\frac{\|\bar{x}\|_1}{\text{dist}(\bar{x}, \partial\Re^n_+)} = \frac{\sum_{j=1}^n \bar{x}_j}{\min_{j=1,\ldots,n}\bar{x}_j} = \frac{\sum_{j=1}^n \tilde{x}_j}{\min_{j=1,\ldots,n}\tilde{x}_j} \leq \frac{1}{\min_{j=1,\ldots,n}\tilde{x}_j} \leq \frac{n+1}{\sigma}$. Therefore $g \leq \max\left\{\|\bar{x}\|_1, \frac{1}{\text{dist}(\bar{x}, \partial\Re^n_+)}, \frac{\|\bar{x}\|_1}{\text{dist}(\bar{x}, \partial\Re^n_+)}, 1\right\} \leq \frac{n+1}{\sigma}$.∎

References

[1] R.M. Freund. On the primal-dual geometry of level sets in linear and conic optimization. Operations Research Center Working Paper 999-01, Massachusetts Institute of Technology, July 2001.

[2] R.M. Freund and J.R. Vera. Condition-based complexity of convex optimization in conic linear form via the ellipsoid algorithm. *SIAM Journal on Optimization*, 10(1):155–176, 1999.

[3] 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. *Mathematical Programming*, 86:225–260, 1999.

[4] M. Grötschel, L. Lovasz, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Springer-Verlag, Berlin, 1988.

[5] Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1994.

[6] J. Renegar. Some perturbation theory for linear programming. *Mathematical Programming*, 65(1):73–91, 1994.

[7] J. Renegar. Linear programming, complexity theory, and elementary functional analysis. *Mathematical Programming*, 70(3):279–351, 1995.

[8] J. Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization*. Society for Industrial and Applied Mathematics, Philadelphia, 2001.

[9] Y. Ye. Toward probabilistic analysis of interior-point algorithms for linear programming. *Mathematics of Operations Research*, 19:38–52, 1994.

[10] E.A. Yildirim and S.J. Wright. Warm-start strategies in interior-point methods for linear programming. *SIAM Journal on Optimization*, 12(3):782–810, 2002.