A Long Step Cutting Plane Algorithm That Uses the Volumetric Barrier [1]

Srinivasan Ramaswamy
ramass@rpi.edu
Dept. of Decision Sciences & Engg. Systems
Rensselaer Polytechnic Institute
Troy, NY 12180

John E. Mitchell [2]
mitchj@rpi.edu
Department of Mathematical Sciences
Rensselaer Polytechnic Institute
Troy, NY 12180

DSES Technical Report

June 20, 1995

Abstract

A cutting plane method for linear/convex programming is described. It is based on the volumetric barrier, introduced by Vaidya. The algorithm is a long step one, and has a complexity of $O(n^{1.5}L)$ Newton steps. This is better than the $O(n\sqrt{m}L)$ complexity of non-cutting plane long step methods based on the volumetric barrier, but it is however worse than Vaidya's original $O(nL)$ result (which is not a long step algorithm). Major features of our algorithm are that when adding cuts we add them right through the current point, and when seeking progress in the objective, the duality gap is reduced by half (not provably true for Vaidya's original algorithm). Further, we generate primal as well as dual iterates, making this applicable in the column generation context as well. Vaidya's algorithm has been used as a subroutine to obtain the best complexity for several combinatorial optimization problems – e.g, the Held-Karp lower bound for the Traveling Salesperson Problem. While our complexity result is weaker, this long step cutting plane algorithm is likely to be computationally more promising on such combinatorial optimization problems with an exponential number of constraints. We also discuss a multiple cuts version — where upto $p \leq n$ 'selectively orthonormalized' cuts are added through the current point. This has a complexity of $O(n^{1.5}Lp\log p)$ quasi Newton steps.

Key words: Cutting plane, volumetric center, long step, Selective Orthonormalization, linear programming, convex programming.

[1] AMS 1991 subject classification. Primary 90C05. Secondary 90C25, 65K05.

[2] Research partially supported by ONR Grant number N00014–94–1–0391.

1 Introduction

The problem of interest is to solve

$$(P)\ \ min\, c^T x$$

$$S.T\ \ Ax\ \geq\ b$$

where $x \in \Re^n,\ \ A\ \in\ \Re^{m\times n}$, and $m\ \gg\ n$.

Its dual is :

$$(D)\ \ \ max\, b^T y$$

$$S.T\ A^T y\ =\ c$$

$$y \geq 0$$

where $A^T\ \in\ \Re^{n\times m}$, $b,\ y\ \in\ \Re^m,\ c \in \Re^n$ and $m\ \gg\ n.$

Recently, there has been some work on cutting plane methods, especially in the context of the convex feasibility problem. Most of these recent methods are "central cutting" plane methods, where the test point is in the center of the polytope that approximates the convex set at the current iteration. The most commonly used test point is the analytic center of the polytope (see eg., [5, 23]), since this is fairly easy to compute. Other test points that have been used include the center of the largest inscribed sphere [10] or ellipsoid [19] and the volumetric center [20].

Since $m\ \gg\ n$, it would be preferable to use a cutting plane approach to solve (P), wherein only a promising set of constraints or hyperplanes are kept, with hyperplanes being added/deleted at each iteration. Such an approach has been provided by Vaidya [20] (based on the volumetric barrier), and Atkinson & Vaidya [5](based on the logarithmic barrier). These authors discuss a cutting plane algorithm for finding a feasible point in a convex set. At each iteration, they maintain a polytope that is guaranteed to contain the convex set, and use the analytic center of this polytope as the test point. If this point is not contained in the convex set, an oracle returns a hyperplane that separates the test point and the convex set. This hyperplane is used to update the polytope. Further, hyperplanes currently in the polytope that are deemed 'unimportant' according to some criteria are dropped. Clearly, such an algorithm can be very useful in solving problems like (P). Much like the ellipsoid algorithm (see [12]), these algorithms can be applied to linear programming problems by cutting on a violated constraint when the test point is infeasible, and cutting on the objective function when the test point is feasible but not optimal. However, from a computational point of view, this is an unwise strategy, and is likely to perform much like its worst case bound all the time (like with the ellipsoid algorithm). This is particularly so if the cuts (be it violated constraints or objective function cuts) are 'backed-off' instead of adding it through the current point, as is the case in Vaidya [20] and Atkinson & Vaidya [5]. In both these cases, the cuts are backed off and added so that the next iterate is guaranteed to be in a small ellipsoid around the current point, suggesting that objective improvement cannot be too much (or improvement towards feasibility, if we add a violated constraint). Yet another disadvantage is that in order to seek progress in the objective, we need to add a cut, increasing the number of constraints.

In [14], Mitchell and Ramaswamy give a long step cutting plane algorithm based on the logarithmic barrier. This may be viewed as casting the cutting plane algorithm of [5], which was based on the analytic center (see [18]), in a long step framework. In stages where the objective is sought to be improved, the duality gap is cut by half, which is better than in [5]. Also, seeking progress in the objective amounts to reducing the numerical value of the barrier parameter, instead of adding a constraint. The long step cutting plane algorithm had a complexity of $O(nL^2)$, same as that of Atkinson and Vaidya.

Here, our goal is to do the same with the cutting plane algorithm of Vaidya [20], which is based on the volumetric barrier. This algorithm, while theoretically the best known to date, has many computationally undesirable features. Indeed, at first sight, the algorithm and the constants associated with it seem complicated and impractical, which may explain the lack of any computational study of Vaidya's algorithm in the literature. The recent work of Anstreicher [2] achieves great improvements in the constants used by Vaidya, making the algorithm seem viable. We believe that casting it in a long step framework is a further step in this direction, even at the expense of a slightly weaker complexity result. Indeed, long step algorithms typically always have a slightly poorer global complexity, but tend to perform much better in practice.

It is important to fully appreciate the importance of Vaidya's result - recently, Bertsimas and Orlin [7] have given complexity for several combinatorial optimization subproblems using Vaidya's algorithm as a subroutine. These complexity results improve on the previously best known for several problems, including for example computing the Held-Karp lower bound for Traveling Salesperson Problems among others. We believe that our long step version, while theoretically weaker, is designed specifically to solve the kind of large linear problems encountered in [7], and is likely to show better performance.

Thus, our interest is in a long step cutting plane algorithm that uses the volumetric barrier function explicitly, partly because of the successful implementation of such algorithms (which were based on the logarithmic barrier) for linear programming. Another point is that by making use of the long step framework, we avoid having to increase the number of constraints whenever we need to drive the objective function value down - instead of cutting on the objective function, we just reduce the value of a scalar parameter. Also, we add violated cuts right through the current point, because it would not improve our complexity theoretically if we were to back off the cuts as in Vaidya [20]. We also discuss the addition of multiple cuts, which is an important aspect of any successful implementation of a cutting plane algorithm. Finally, since in such an algorithm we would maintain primal and dual variables, it allows for early termination when the sub-optimality is deemed to be within allowable limits. In our algorithm, when an iterate is infeasible in the 'overall problem', constraints are added until feasibility is regained. Additionally, constraints are dropped whenever possible according to some criterion. We prove a complexity of $O(n^{1.5}L)$ iterations for this algorithm.

2 Preliminaries

The original problem of interest is :

$$\begin{array}{lll}(\hat{P}) & min\ \hat{c}^T x & \\ & \hat{A}x & \geq \hat{b}\ , \ \ x \in \Re^{\hat{n}}\end{array}$$

where $\hat{A} \in \Re^{\hat{m}\times\hat{n}}$.

Its dual is :

$$\begin{array}{lll}(\hat{D}) & max\ \hat{b}^T y & \\ & \hat{A}^T y & = \hat{c} \\ & y & \geq 0, \ \ y \in \Re^{\hat{m}}\end{array}$$

where $\hat{A}^T \in \Re^{\hat{n}\times\hat{m}}$ and $\hat{m} \gg \hat{n}$.

We point out here that we do not actually require $\hat{m}$ to be finite — we will never work with the formulations $(\hat{P})$ or $(\hat{D})$ directly. We will just approximate them by a finite number of linear constraints. We also assume that there exists a problem dependent constant L such that
(i) The set of optimal solutions to $(\hat{P})$ is guaranteed to be contained in the n-dimensional hypercube of half-width 2^L given by $\{x \in \Re^{\hat{n}} : |x_i| < 2^L\}$, and
(ii) The set of optimal solutions to $(\hat{P})$ contains a full dimensional ball of radius 2^{-L}.
We may interpret (ii) as saying that it suffices to find a solution to $(\hat{P})$ to within an accuracy 2^{-L}.

At any given iteration, we operate with a relaxation of $(\hat{P})$. The algorithm is initialized with the relaxation

$$\begin{array}{lll} (P_0) & min\ c^T x & \\ & c^T x & \geq -\|c\|_1 2^L \\ & Ix & \geq (-2^L)e_n \\ & -Ix & \geq (-2^L)e_n \end{array}$$

where $\mathcal{F}\ (P_0)$ {ie., the feasible set of this relaxation } contains $\mathcal{X}$, which denotes the set of optimal solutions to $(\hat{P})$. The vector of dimension n whose components are all equal to one is denoted by e_n, or simply e. Indeed, e is used to denote the vector of all ones with the dimension being understood from the context. It should be noted that $\hat{n}$ is always the same as n, and we may thus drop the hats. This is also the case with $\hat{c}$.

At iteration k, we would have something like :

$$\begin{array}{lll} (P_k) & min\ c^T x & \\ & A_k x & \geq b_k \end{array}$$

with $A_k \in \Re^{m_k \times n}$.

Throughout most of the discussion, we may omit the subscripts k since that causes no ambiguity. Also, we sometimes use P to refer to the problem as well as the polytope defined by its constraints. Again, we do this since it is clear from the context which we are referring to.

We need to define the so-called logarithmic barrier function :

$$F(x) \equiv -\sum_{i=1}^{m} ln\ s_i \tag{1}$$

where s_i is equal to $a_i^T x - b_i$, and a_i is the i^{th} row of A. We use s_i without explicitly noting its functional dependence on x.

It is easily seen that the gradient of the barrier function is given by

$$\nabla F(x) = -\sum_{i=1}^{m} \frac{a_i}{s_i} = -A^T S^{-1} e \tag{2}$$

and the Hessian by

$$H(x) \equiv \nabla^2 F(x) = \sum_{i=1}^{m} \frac{a_i a_i^T}{s_i^2} = A^T S^{-2} A \tag{3}$$

Now, define the so-called Volumetric Barrier function:

$$V(x) \equiv \frac{1}{2} \log det(\nabla^2 F(x)) \tag{4}$$

Since we are interested in a long-step algorithm based on this barrier, we will instead directly work with the function

$$\psi(x, \pi, l) \equiv \frac{c^T x - l}{\pi} + V(x) \tag{5}$$

where π is of course the barrier parameter, and l is some lower bound to the optimum objective function value of $(\hat{P})$.

For a given value of π , $x(\pi)$ denotes the unique minimizer of this barrier function. This minimizer exists and is unique if and only if the polytope P is bounded — see, e.g., Anstreicher [2]. We refer to this unique point when we use the term (exact) π-center.

In our analysis towards the end, we refer to the polytope P^*. This polytope is to be understood as follows — if we have the current point x, then

$$P^* \equiv P \cap \{y \in \Re^n : c^T y \leq \beta\} \tag{6}$$

where β is equal to $max(c^T x, u)$, and u is an upper bound on the optimal value of $c^T x$. Thus, this polytope is certain to contain the optimal set to the problem of our interest and the current point. This property is important, and enables us to decide termination criteria based on the volume of this polytope. It is important to note that the polytope P^* is used only in the theoretical analysis - we do not explicitly add upper-bound constraints that restrict subsequent iterates to be in the polytope P^*, although we could.

For convenience, we sometimes use the notation

$$\|v\|_B \equiv \sqrt{v^T B^{-1} v}$$

for any $v \in \Re^n$ where B is a symmetric positive definite matrix.

The outline of this paper is as follows. In the rest of this section, we discuss the properties of the volumetric barrier function, including analyzing the convergence of the quasi Newton steps (i.e., Newton like steps based on Q instead of $\nabla^2\psi$). The reader interested in the main thrust of this paper may want to skip this, referring back only as necessary. In section 3, we present the algorithm. In the section after that, we analyze the complexity of the algorithm, considering cut addition, cut deletion and reducing the barrier parameter in separate subsections. In section 5 we consider this algorithm with the addition of multiple cuts in case 2.1. Finally, in section 6, we state our conclusions and discuss other algorithms from the literature.

2.1 Properties of the Volumetric Barrier

Here, we state some results that we need about the volumetric barrier function. When it causes no confusion to do so, we suppress the dependence of $\psi(\cdot)$ on π and l.

Proposition 1 *Let $\psi(x)$ be defined as in equation (5). Define*

$$P(s) = S^{-1}A(A^T S^{-2} A)^{-1} A^T S^{-1}$$

and let $\hat{P} = P \circ P$ where $\circ$ denotes the Schur or Hadamard product (i.e., $\hat{P}_{ij} = P_{ij}^2$). Define $\sigma_i \equiv P_{ii}$, and $\Sigma \equiv diag\{\sigma_i\}$. We then have

$$\begin{aligned} g(x) \equiv \nabla\psi(x) &= \frac{c}{\pi} - \sum_{i=1}^{m} \sigma_i \frac{a_i}{s_i} = \frac{c}{\pi} - A^T \Sigma S^{-1} e & (7) \\ \nabla^2\psi(x) &= A^T S^{-1}(3\Sigma - 2\hat{P}) S^{-1} A = \nabla^2 V(x) & (8) \end{aligned}$$

Proof :
See Anstreicher [2].

■

It is worth noting that

$$\sigma_i(x) = \frac{a_i^T H(x)^{-1} a_i}{s_i(x)^2}$$

As in Vaidya [20], we will want to deal with an approximation to $\nabla^2 V$ rather than $\nabla^2 V$ itself. We work towards that end.

Lemma 1 *For any $x \in P$, we have*
(i) $\sigma_i \leq 1 \quad \forall i$.
(ii) $\sum_{i=1}^m \sigma_i = n$.

Proof :
See Vaidya [20] or Anstreicher [2].

■

Now, let us define

$$Q(x) \equiv \sum_{i=1}^{m} \sigma_i \frac{a_i a_i^T}{s_i^2} = A^T S^{-2} \Sigma A \tag{9}$$

We want to claim that this matrix is a good approximation to the Hessian of $V(x)$.

Lemma 2 *Let $Q(x)$ be defined as in equation (9), and let $\nabla^2 V(x)$ be defined as in equation (8). We then have, $\forall x \in P$,*

$$Q(x) \preceq \nabla^2 V(x) \preceq 3Q(x)$$

where $\preceq$ denotes the partial ordering induced by the cone of symmetric positive semidefinite matrices — for any two symmetric psd matrices A and B, $A \preceq B$ if and only if $B - A$ is symmetric positive semidefinite.

Proof :
See Anstreicher [2].

■

Now that we have established that the matrix $Q(x)$ can serve as a reasonable approximation to the Hessian of $V(x)$, it is natural to want to take Newton-like steps based on $Q(x)$ rather than $\nabla^2 V(x)$. Formally speaking, we need to define the notion of proximity to the exact π-center, and show that when we are close to the exact center (measured by the proximity measure) the Newton-like steps give us good convergence. But before we can state that, we need some technical lemmas and propositions.

We first state a very well known proposition — a proof may be found in any standard text on matrices [13]. This is a result that we make heavy use of later in our proofs.

Proposition 2 *Let A and B be $n \times n$ symmetric positive definite matrices, and let t be a positive scalar. Then,*

$$A \preceq tB$$

implies that

$$A^{-1} \succeq (1/t)B^{-1}$$

Definition 1 *For any $x \in P, r \in \Re, r \geq 0$, define $\Sigma(x, r)$ as $\{y \in \Re^n : \frac{|a_i^T(y-x)|}{a_i^T x - b_i} \leq r \ \ \forall i\}$.*

We note that $\Sigma(x, r) \subseteq P$ for $r \leq 1$.

The set $\Sigma(x, r)$ was originally introduced by Vaidya [20] (see also [4, 6]), and may be thought of as a polyhedral neighbourhood, and a generalization of a small ball around the current point. Now, we can develop approximations to various quantities at points in this generalized neighborhood.

Proposition 3 *Let $x \in P$. Let $y \in \Sigma(x, r)$ for some $r < 1$. Then, for all $\xi \in \Re^n$,*

$$\frac{(1-r)^2}{(1+r)^4}\xi^T Q(x)\xi \leq \xi^T Q(y)\xi \leq \frac{(1+r)^2}{(1-r)^4}\xi^T Q(x)\xi$$

$$\frac{1}{(1+r)^2}\xi^T H(x)\xi \leq \xi^T H(y)\xi \leq \frac{1}{(1-r)^2}\xi^T H(x)\xi$$

Proof :
See Vaidya [20].

■

It will turn out that the relationship between the matrices H (the Hessian of the logarithmic barrier) and Q will be important to us. Towards quantifying that relationship, define

$$\mu(x) \equiv \inf_{\xi \in \Re^n} \frac{\xi^T Q(x)\xi}{\xi^T H(x)\xi} \tag{10}$$

It is worth noting that $\mu \leq 1$ and $\mu \geq \min_{i=1 \ to \ m} \sigma_i$. Having defined μ, the following result is quite straightforward.

Proposition 4

$$\mu H \preceq Q \leq H$$

Proof :
Follows from the definition of μ, and the fact that $\mu \leq 1$.

■

We will also want to approximate the quantity $\mu(y)$ for $y \in \Sigma(x, r)$.

Proposition 5 *Let $x \in P$. Let $y \in \Sigma(x, r), r < 1$. We then have*

$$\frac{(1-r)^4}{(1+r)^4}\mu(x) \leq \mu(y) \leq \frac{(1+r)^4}{(1-r)^4}\mu(x)$$

Proof :
See Vaidya [20].

■

We now state another technical lemma that we need for future use.

Lemma 3 *Let $x \in \Re^n$ be any point in $int(P)$. Then, $\forall i$ from 1 to m, we have*

$$\frac{a_i^T Q(x)^{-1} a_i}{s_i(x)^2} \leq \frac{1}{\sqrt{\mu(x)}} \quad \forall x \in P$$

Proof :
See Vaidya [20].

■

Lemma 4

$$E(Q(x_0), x_0, (\mu(x_0))^{\frac{1}{4}} r) \subseteq \Sigma(x_0, r)$$

Proof :
See Vaidya [20].

■

We are now ready to proceed further. Our final goal is to establish that the quantity $g(z)^T Q(z)^{-1} g(z)$ defines a measure of proximity of the point z to the π-center. Roughly speaking, we want to establish that if this measure is small, then $\psi(z) - \psi(z^*)$ is small, and vice versa, where z^* denotes the unknown minimizer of $\psi(x)$. We will do this as in [20], by integrating some quantity along a trajectory from z^* to z. To formalize this, we must first establish the existence of the trajectory.

Theorem 1 *(Vaidya) Let $v \in \Re^n$ be fixed. Define $\Phi(x,t) \equiv g(x) - tv$, where t is a scalar and $x \in P$. Then the equation $\Phi(x,t) = 0$ implicitly defines x as a function of t, and we may write $x = x(t)$ along this trajectory. Moreover, $x(t)$ is an analytic function of t, and*

$$\frac{dx(t)}{dt} = \nabla^2 \psi(x(t))^{-1} v$$

Further, if $x(t_1)$ and $x(t_2)$ are two points on this trajectory, we can write

$$\psi(x(t_2)) - \psi(x(t_1)) = \int_{t_1}^{t_2} t v^T \nabla^2 \psi(x(t))^{-1} v dt$$

Proof :
This theorem, as well as the underlying idea of using this trajectory are due to Vaidya [20]. A proof of this may be found there.

■

It may by now be clear how we plan to bound $\psi(z)-\psi(z^*)$ — by using $v = \nabla\psi(z)$ in theorem 1. We know we may approximate the quantity $v^T\nabla^2 V(x(t))^{-1}v$ inside the integrand by $v^T Q(x(t))^{-1}v$. But in order to bound $Q(x(t))$ independently of t, so that we may pull it outside the integral, we need to know that the entire trajectory lies inside $\Sigma(z,r)$ for some small r. We need some more technical lemmas before we can do this.

Lemma 5 *Let $v \in \Re^n$ be fixed. Let $\hat{x}$ be a fixed point in P, and suppose that $\nabla\psi(\hat{x}) = \hat{t}v$. Consider the trajectory defined by $\nabla\psi(x(t)) = tv$. For the portion of this trajectory that lies in $\Sigma(\hat{x},r)$ (where $r < 1$), we have for every $i = 1$ to m,*

$$|\frac{a_i^T \frac{dx(t)}{dt}}{a_i^T x(t) - b_i}| \leq \frac{(1+r)^3}{(1-r)^2}\frac{(v^T Q(\hat{x})^{-1}v)^{1/2}}{(\mu(\hat{x}))^{1/4}}$$

Proof :
Note that $\frac{dx(t)}{dt} = \nabla^2 V(x(t))^{-1}v$. Thus,

$$\begin{aligned} |a_i^T \frac{dx(t)}{dt}| &= |a_i^T \nabla^2 V(x(t))^{-1}v| \\ &\leq (a_i^T\nabla^2 V(x(t))^{-1}a_i)^{1/2}(v^T\nabla^2 V(x(t))^{-1}v)^{1/2} \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{|a_i^T \frac{dx(t)}{dt}|}{s_i(x(t))} &\leq \frac{(a_i^T\nabla^2 V(x(t))^{-1}a_i)^{1/2}}{s_i(x(t))}(v^T\nabla^2 V(x(t))^{-1}v)^{1/2} \\ &\leq \frac{(a_i^T Q(x(t))^{-1}a_i)^{1/2}}{s_i(x(t))}(v^T Q(x(t))^{-1}v)^{1/2} \\ &\leq \frac{(v^T Q(x(t))^{-1}v)^{1/2}}{(\mu(x(t)))^{1/4}} \end{aligned}$$

where the last step follows from lemma 3. Now, all we need is to approximate $Q(x(t))$ and $\mu(x(t))$ for all $x(t) \in \Sigma(\hat{x},r)$. Doing this using propositions 3 and 5, we get the desired result.

■

Now, we use the above lemma to show how far one must go along the trajectory (measured in terms of the difference in t values) before one leaves the set $\Sigma(\hat{x},r)$. This proof is exactly along the lines of a similar proof for the weighted analytic center that may be found in [4].

Lemma 6 *Let $v \in \Re^n$ be fixed. Let $\hat{x}$ be a fixed point in P, and suppose that $\nabla\psi(\hat{x}) = \hat{t}v$. Consider the trajectory defined by $\nabla\psi(x(t)) = tv$. Let $\bar{t}$ be such that $x(\bar{t})$ is not in $\Sigma(\hat{x},r)$, for some $r < 1$. Then, we have*

$$|\hat{t}-\bar{t}| \geq \frac{(r-0.5r^2)(1-r)^2(\mu(\hat{x}))^{1/4}}{(1+r)^3(v^T Q(x(t))^{-1}v)^{1/2}}$$

Proof :
WLOG, let $x(\bar{t})$ be the first point that lies outside $\Sigma(\hat{x},r)$. Then, we must have, for some k,

$$\frac{a_k^T x(\bar{t}) - b_k}{a_k^T x(\hat{t}) - b_k} = 1 + r$$

or

$$\frac{a_k^T x(\bar{t}) - b_k}{a_k^T x(\hat{t}) - b_k} = 1 - r$$

In either case, we have

$$|\ln(\frac{a_k^T x(\bar{t}) - b_k}{a_k^T x(\hat{t}) - b_k})| \geq r - \frac{r^2}{2}$$

But

$$\begin{aligned}
|\ln(\frac{a_k^T x(\bar{t}) - b_k}{a_k^T x(\hat{t}) - b_k})| &= |\ln(s_k(x(\bar{t}))) - \ln(s_k(x(\hat{t})))| \\
&= |\int_{\hat{t}}^{\bar{t}} \frac{a_k^T \frac{dx(t)}{dt}}{s_k(x(t))} dt| \\
&\leq \int_{\hat{t}}^{\bar{t}} |\frac{a_k^T \frac{dx(t)}{dt}}{s_k(x(t))}||dt| \\
&\leq |\hat{t} - \bar{t}| \frac{(1+r)^3}{(1-r)^2} \frac{(v^T Q(\hat{x})^{-1} v)^{1/2}}{(\mu(\hat{x}))^{1/4}}
\end{aligned}$$

where the last inequality follows from lemma 5. The final result of this lemma follows by rearranging terms.

■

2.2 Convergence of Quasi-Newton Steps

Now, we are finally ready to establish that $\Psi(z) \equiv g(z)^T Q(z)^{-1} g(z)$ indeed is a reasonable measure of proximity to the exact π-center.

Lemma 7 *Let $\delta = 10^{-4}$. Let $z \in P$, and suppose that $g(z)^T Q(z)^{-1} g(z) \leq \delta\sqrt{\mu(z)}$. Then,*
(i) $z^* \in \Sigma(z, 1.1\sqrt{\delta})$, where z^* denotes the unknown minimizer of $\psi(z)$.
(ii) $\mu(z) \leq 1.1\mu(z^*)$.
(iii) $\psi(z) - \psi(z^*) \leq 0.55 g(z)^T Q(z)^{-1} g(z)$.
(iv) $\psi(z) - \psi(z^*) \geq 0.156 g(z)^T Q(z)^{-1} g(z)$.

Proof :
(i) Consider the trajectory defined by $\nabla\psi(x) = t\nabla\psi(z)$. Clearly, $x(1) = z$ and $x(0) = z^*$. Now, let us look at $\Sigma(z, r)$ for $r = 1.1\sqrt{\delta}$. Let $\bar{t}$ be such that $x(\bar{t}) \notin \Sigma(z, r)$. Applying the result of lemma 6, we have

$$\begin{aligned}
|\bar{t} - 1| &\geq \frac{(r - 0.5r^2)(1-r)^2[\mu(z)]^{1/4}}{(1+r)^3[g(z)^T Q(z)^{-1} g(z)]^{1/2}} \\
&\geq 1.03
\end{aligned}$$

for $r = 1.1\sqrt{\delta}$, under the hypothesis of this lemma. As a consequence, we conclude that the trajectory $x(t)$ is contained in $\Sigma(z, 1.1\sqrt{\delta})$ for $t \in [0, 1]$. Thus, $x(0) = z^*$ lies in $\Sigma(z, 1.1\sqrt{\delta})$.
(ii) Noting that $z^* \in \Sigma(z, r)$ for $r = 1.1\sqrt{\delta}$, and applying proposition 5, this part follows directly.
(iii) We will integrate along the trajectory defined by $g(x(t)) = tg(z)$, which we know from (i) is

contained in $\Sigma(z,r)$ for $r = 1.1\sqrt{\delta}$. Using theorem 1 and lemma 2 (in the third step) and proposition 3, we can write

$$
\begin{aligned}
\psi(z) - \psi(z^*) &= \int_0^1 \frac{d\psi}{dt} dt \\
&= \int_0^1 \nabla\psi(x(t))^T \frac{dx}{dt} dt \\
&= \int_0^1 t g(z)^T \nabla^2 V(x(t))^{-1} g(z) dt \\
&\leq \int_0^1 t g(z)^T Q(x(t))^{-1} g(z) dt \\
&\leq \frac{(1+r)^4}{(1-r)^2} g(z)^T Q(z)^{-1} g(z) \int_0^1 t dt \\
&\leq 0.55 g(z)^T Q(z)^{-1} g(z)
\end{aligned}
$$

where the last two steps use $r = 1.1\sqrt{\delta}$.

(iv) As above, we can write

$$
\begin{aligned}
\psi(z) - \psi(z^*) &= \int_0^1 \frac{d\psi}{dt} dt \\
&= \int_0^1 t g(z)^T \nabla^2 V(x(t))^{-1} g(z) dt \\
&\geq \frac{1}{3} \int_0^1 t g(z)^T Q(x(t))^{-1} g(z) dt \\
&\geq \frac{1}{3} \frac{(1-r)^4}{(1+r)^2} g(z)^T Q(z)^{-1} g(z) \int_0^1 t dt \\
&\geq 0.156 g(z)^T Q(z)^{-1} g(z)
\end{aligned}
$$

where the last two steps use $r = 1.1\sqrt{\delta}$.

■

Thus, we have shown that if the measure is small, we are indeed in the proximity of the unknown minimizer. Now we address the issue of convergence to this minimizer using quasi Newton steps based on Q. As before, we first define a trajectory - here it will simply be a straight line. We will show that this trajectory is in the neighbourhood of the current point, and then we will integrate along it.

Lemma 8 *Let z be a point satisfying $g^T Q(z)^{-1} g < \delta\sqrt{\mu(z)}$. Define $\eta = -Q(z)^{-1} g$. Let $z(t) = z + t\eta$, for $t \in [0, \lambda]$, $\lambda < 1$. Then we have*

$$z(t) \in \Sigma(z, t\sqrt{\delta})$$

Proof :
Note that $z(t) - z = t\eta$. Therefore,

$$
\begin{aligned}
(z(t) - z)^T Q(z)(z(t) - z) &= t^2 \eta^T Q(z) \eta \\
&= t^2 g^T Q(z)^{-1} g \\
&\leq t^2 \delta \sqrt{\mu(z)} \\
&= (t\sqrt{\delta}(\mu(z))^{\frac{1}{4}})^2
\end{aligned}
$$

Thus, we have $z(t) \in E(Q(z), z, t\sqrt{\delta}(\mu(z))^{\frac{1}{4}})$, which, by lemma 4, implies that $z(t) \in \Sigma(z, t\sqrt{\delta})$.

■

This tells us that once we are 'close' to the weighted volumetric center, we can even take a full step in this quasi-Newton direction without violating feasibility. However, as we will soon see, we will be restricted to less than a full step in order to guarantee a certain decrease in the volumetric potential function. We will do this by integrating along the line we defined in lemma 8. But first we need some more technical lemmas.

Lemma 9 *Define $\psi(t) := \psi(z(t))$, where $z(t)$ is defined in lemma 8. Choose $\lambda = 0.2$. Under the hypothesis of lemma 8, we can say that*

$$\frac{d^2\psi(t)}{dt^2} \leq 3.037\Psi(z) \ \ \forall t \in [0, \lambda]$$

Proof :
We have $z(t) = z + t\eta$. Therefore,

$$\frac{d\psi(t)}{dt} = g(z(t))^T\eta$$

and

$$\frac{d^2\psi(t)}{dt^2} = \eta^T\nabla^2 V(z(t))\eta \ \ \forall t \in [0, \lambda]$$

From lemma 2, we have $\nabla^2 V(z(t)) \preceq 3Q(z(t))$. Thus,

$$\frac{d^2\psi(t)}{dt^2} \leq 3\eta^T Q(z(t))\eta \ \ \forall t \in [0, \lambda]$$

Also, lemma 8 and proposition 3 imply that

$$\begin{aligned}
\frac{d^2\psi(t)}{dt^2} &\leq 3\eta^T Q(z(t))\eta \\
&\leq 3\frac{(1+\lambda\sqrt{\delta})^2}{(1-\lambda\sqrt{\delta})^4}\eta^T Q(z)\eta \\
&\leq 3.037\Psi(z)
\end{aligned}$$

where the last step uses $\lambda = 0.2$ and $\delta = 10^{-4}$.

■

Now we will use the above lemma to bound the first derivative $\frac{d\psi}{dt}$ in the interval $t \in [0, \lambda]$, so that we may use that bound to guarantee a certain decrease in the volumetric potential function.

Lemma 10 *Let z satisfy $\Psi(z) < \delta\sqrt{\mu(z)}$. Let η and λ be defined as before. Let $z(t) = z + t\eta$, $t \in [0, \lambda]$. Then*

$$\frac{d\psi(t)}{dt} \leq (-1 + 3.037t)\Psi(z)$$

Proof :

$$\begin{aligned}\frac{d\psi(t)}{dt} &= \frac{d\psi(0)}{dt} + \int_0^t \frac{d^2\psi(s)}{ds^2}ds \\ &= g(z)^T\eta + \int_0^t \frac{d^2\psi(s)}{ds^2}ds \\ &\leq -\Psi(z) + 3.037t\Psi(z)\end{aligned}$$

■

We are now in a position to analyze the convergence of the quasi-Newton method.

Lemma 11 *Let z satisfy $\Psi(z) < \delta\sqrt{\mu(z)}$. Let η and λ be defined as before. Let $\bar{z} = z + \lambda\eta$, $\lambda = 0.2$. Let z^* denote the (unknown) minimizer of $\psi(\cdot, \pi)$. Then,*

$$\psi(\bar{z}) - \psi(z^*) \leq 0.75(\psi(z) - \psi(z^*))$$

Proof :
We note that $\Psi(z) < \delta\sqrt{\mu(z)}$ implies that $\psi(z) - \psi(z^*) \leq 0.55\Psi(z)$ by lemma 7, or equivalently that $\Psi(z) \geq 1.8(\psi(z) - \psi(z^*))$. Also,

$$\begin{aligned}\psi(\bar{z}) - \psi(z) &= \int_0^\lambda \frac{d\psi(t)}{dt}dt \\ &\leq \Psi(z)\int_0^{0.2}(-1 + 3.037t)dt \\ &= -0.1392\Psi(z) \\ &\leq (1.8)(-0.1392)(\psi(z) - \psi(z^*)) \\ &\leq -0.25(\psi(z) - \psi(z^*))\end{aligned}$$

Thus, we can say that

$$\begin{aligned}\psi(\bar{z}) - \psi(z^*) &= \psi(\bar{z}) - \psi(z) + \psi(z) - \psi(z^*) \\ &\leq 0.75(\psi(z) - \psi(z^*))\end{aligned}$$

■

Thus, we have shown that if a point z is reasonably close to the exact π-center, as measured by $\Psi(z)$, then taking quasi-Newton steps results in linear convergence. However, it is also useful to analyze the behaviour of the quasi-Newton steps starting from a point that is not an approximate (weighted) volumetric center. We do that now.

Lemma 12 *Let z be any point in $int(P)$ that does not satisfy $\Psi(z) < \delta\sqrt{\mu(z)}$. Let $\eta := Q(z)^{-1}g$, $\lambda = 0.2$, $\bar{\lambda} = \frac{\lambda\sqrt{\delta}(\mu(z))^{\frac{1}{4}}}{\sqrt{\Psi(z)}}$, and let $z(t) \equiv z - t\eta$ for $t \in [0, \bar{\lambda}]$. Then $z(t) \in \Sigma(z, \lambda\sqrt{\delta})$.*

Proof :
As before, note that $z(t) - z = t\eta$. Therefore,

$$\begin{aligned}(z(t)-z)^T Q(z)(z(t)-z) &= t^2 \eta^T Q(z)\eta \\ &= t^2 g^T Q(z)^{-1} g \\ &= t^2 \Psi(z) \\ &\leq \bar{\lambda}^2 \Psi(z) \\ &= \lambda^2 \delta \sqrt{\mu(z)}\end{aligned}$$

Thus, we have $z(t) \in E(Q(z), z, \lambda\sqrt{\delta}(\mu(z))^{\frac{1}{4}})$, which, by lemma 4, implies that $z(t) \in \Sigma(z, \lambda\sqrt{\delta})$.

■

Once again, as before, in order to analyze the quasi Newton step, we will need to bound the second derivative $\frac{d^2\psi(t)}{dt^2}$, and then the first derivative.

Lemma 13 *Under the hypothesis of lemma 12, we have*

$$\frac{d^2\psi(t)}{dt^2} \leq 3.037\Psi(z) \quad \forall t \in [0, \bar{\lambda}]$$

Proof :
This proof is identical to the proof of lemma 9.

■

Now we are ready to bound the first derivative along the entire portion of the step.

Lemma 14 *Under the hypothesis of lemma 12, we can say that*

$$\frac{d\psi(t)}{dt} \leq (-1 + 3.037t)\Psi(z) \quad \forall\, t \in [0, \bar{\lambda}]$$

Proof :
Again, this proof is identical to that of lemma 10.

■

We can now put together the results of the previous lemmas to show that the quasi-Newton steps still converge to the π-center even when started from a point that is not an approximate center.

Lemma 15 *Let z be a point in the interior of the polytope P that does not satisfy $\Psi(z) < \delta\sqrt{\mu(z)}$. Let η, λ and $\bar{\lambda}$ be defined as in lemma 12. Let $\bar{z} := z - \bar{\lambda}\eta$. Then,*

$$\psi(\bar{z}) - \psi(z) \leq -0.1392\delta\sqrt{\mu(z)}$$

Proof :
We have

$$\begin{aligned}\psi(\bar{z}) - \psi(z) &= \int_0^{\bar{\lambda}} \frac{d\psi(t)}{dt} dt \\ &\leq \Psi(z) \int_0^{\bar{\lambda}} (-1 + 3.037t) dt \\ &= \Psi(z)((\frac{1}{2})(3.037)\bar{\lambda}^2 - \bar{\lambda}) \\ &\leq -0.1392\delta\sqrt{\mu(z)}\end{aligned}$$

where the last step uses the definition of $\bar{\lambda}$ and the fact that $\Psi(z) > \delta\sqrt{\mu(z)}$.

■

Thus, it is clear that at least some decrease is obtainable through these quasi Newton steps. In order to bound the magnitude of this decrease below, we need to place a lower bound on $\mu(z)$ for arbitrary $z \in P$.

Lemma 16 *Let z be any point in $P := \{x \in \Re^n : Ax \geq b\}$, where $A \in \Re^{m \times n}$, and $b \in \Re^m$. Let the quantity $\mu(z)$ be defined as in equation (10). Then we can say that*

$$\mu(z) \geq \frac{1}{m}$$

Proof :
See Anstreicher [1].

■

This allows us to say that the quasi Newton steps will give us at least a $\Omega(1/\sqrt{m})$ decrease in $\psi(\cdot, \pi)$, even when started from a point that is not 'close' to the π-center. This result will be useful to us later on.

3 The Algorithm

For the purpose of initializing the algorithm, we assume that the origin is feasible in problem ($\hat{P}$). This is reasonable since if we know any feasible point in ($\hat{P}$), we can modify the problem to make the origin feasible.

The algorithm is shown in figure 1. It should be mentioned that where a lower case letter denotes a vector, the corresponding upper case letter denotes a diagonal matrix with this vector along the diagonal.

The basic idea of the algorithm is this : at any iteration, if a constraint has become unimportant (i.e., it has a small σ_j), then that hyperplane is dropped. If the constraints are deemed important, and the current point is infeasible in $(\hat{P})$, then there is a violated constraint, and the polytope is updated with this constraint. (We assume that there is an 'oracle' that can find this violated constraint.) If the current point is feasible in $(\hat{P})$, then the lower bound is updated and the barrier parameter is reduced.

Figure 1: **The Algorithm**

Step 0 : Initialization

$$\pi \;=\; 2^L \quad k \;=\; 0 \quad A = \begin{bmatrix} c^T \\ I_n \\ -I_n \end{bmatrix} \quad b = \begin{bmatrix} -2^L\|c\|_1 \\ -2^L e_n \\ -2^L e_n \end{bmatrix}$$

$$x = [0 \cdots 0]^T \; ; \quad s = Ax - b \; ;$$

Step k : The Iterative Stage

Have point $x = x^k$, which is an approximate π-center.
Calculate $\Gamma(x) = \min_{i=2\ to\ m} \sigma_i(x)$.

<u>Case 1 : $\Gamma < \epsilon$</u>

Then, for some j we have $\sigma_j(x) = \min_{2\ to\ m} \sigma_i(x) < \epsilon$
Drop the hyperplane a_j, resulting in polytope P_d.
Take $O(\sqrt{n} + \log n)$ Quasi-Newton steps to find next approximate π-center x_d.
Get dual solution $y = \pi S^{-1}\Sigma(e - S^{-1}Ad)$, where $d = Q^{-1}\nabla\psi(x^k, \pi, l)$.
If satisfactory, STOP.

<u>Case 2 : $\Gamma \geq \epsilon$</u>

<u>Subcase 2.1 : x is not feasible in ($\hat{P}$)</u>

Then $\exists\, j$ such that $\hat{a}_j^T x < \hat{b}_j$.
Add $\hat{a}_j^T x \geq \hat{a}_j^T x^k$ to the polytope, getting new polytope P_a. Set $m = m + 1$.
Find new approximate π-center x_a in $O(\sqrt{n} + \log n)$ quasi Newton steps.

<u>Subcase 2.2 : x is feasible in ($\hat{P}$)</u>

Test x for optimality. If optimal, get dual solution y and STOP.
Else,
Set $l = c^T x - 2n\pi$. Let l_{prev} denote previous lower bound.
If $l_{prev} \geq l$, the lower bound cannot be improved upon.
If $l_{prev} \leq l$, remove existing constraint $c^T x \geq l_{prev}$, add $c^T x \geq l$.
Set $\pi \equiv \pi\rho$ where $\rho \in (0.5, 1)$ is a constant independent of the problem. Take $O(n^{1.5} + \log n)$ quasi Newton steps with linesearch to find new approximate π-center.

Our convergence proof is as follows : first, we show that our dropping criterion implies that there cannot be more than $O(n)$ number of constraints at any stage of the algorithm. Then, we use this to show that if $\psi(x,\pi) > O(nL)$, then the volume of some ellipsoid that contains the polytope must have become too small, thus enabling termination. Finally, we show that this must happen in no more than $O(n^{1.5}L)$ number of Newton steps.

In much of the analysis of this algorithm, we have only assumed that the current iterate is an approximate π-center. However, for some of the proofs, we assume that we have an exact π-center. This helps in focussing attention on the key issue without having to keep track of details that would arise from having only an approximate π-center. It goes without saying that the algorithm itself uses only approximate centers.

We now formally define what we mean by an approximate center.

Definition 2 *A point $z \in int(P)$ is called an approximate π-center if it satisfies*

$$\nabla\psi(z,\pi,l)^T Q(z)^{-1} \nabla\psi(z,\pi,l) < \min(\delta\sqrt{\mu(z)}, 1/\sqrt{m})$$

The motivation for part of this definition must be clear - the other part will be considered in section 4.3

4 Complexity Analysis

In this section, we attempt to show that every stage of the algorithm may be accomplished in a reasonable number of Newton steps, and we also establish the global complexity of the algorithm. In separate subsections, we analyze the necessary lemmas pertaining to adding a cut, dropping a cut and reducing the barrier parameter. An important difference between our cutting plane algorithm and that of Vaidya [20] or Anstreicher [2] is that we never drop the lower bounding constraint $c^T x \geq l$. We need this constraint to be always present in the current system of inequalities for our complexity analysis. Therefore, no matter what the value of $\sigma_1(x) = c^T(\nabla^2 F(x))^{-1}c/(c^T x - l)^2$, we do not drop this constraint. This weakens our local convergence results - we prove a bound of $O(\sqrt{n})$ steps after adding a cut in order to find the next π-center. However, this does not affect the overall complexity of the algorithm, which would anyway be $O(n^{1.5}L)$. We do make up in some measure, because this allows us to add the new cut directly through the current point.

4.1 Adding a Cut

4.1.1 Effect of Adding a Cut on the Volumetric Potential

Here, we analyze the effect of adding a cut on the volumetric potential function $\psi(\cdot)$ - in particular, if z denotes the current π-center, $\hat{z}$ denotes the new π-center, and $\hat{\psi}(\cdot)$ denotes the potential function with the new cut included as well, then we want to show that there is some constant η_a such that

$$\hat{\psi}(\hat{z},\pi) \geq \psi(z,\pi) + \eta_a$$

We note that since this is an occurrence of case 2.1, we must have $\min_{2\ to\ m} \sigma_j > \epsilon$. However, it is possible to have $\sigma_1 < \epsilon$ at this point in the algorithm. An increase in the potential function is easy to show under the case when the lower bounding constraint satisfies $\sigma_1(z) > \epsilon$ — indeed, the lemmas of Vaidya [20] and Anstreicher [2] apply directly. When this is not the case, we are forced to use the property that all the other sigmas are bigger than ϵ. This complicates matters significantly. Since this is the harder case, we only prove our results for this case.

First we need some technical lemmas.

We will begin by considering how far the new minimizer must be from the old one. Let all quantities with tildes refer to the system of inequalities where the lower bounding constraint is *present but ignored*. What we mean is that it is not dropped (the algorithm requires that we keep it in), but we ignore it. Thus,

$$\begin{aligned} \tilde{Q} &= \sum_{2}^{m} \sigma_i/s_i^2 \; a_i a_i^T && (11) \\ \tilde{H} &= \sum_{2}^{m} 1/s_i^2 \; a_i a_i^T && (12) \\ \tilde{\mu} &\equiv \inf_{\xi\in\Re^n} \frac{\xi^T\tilde{Q}\xi}{\xi^T\tilde{H}\xi} > \epsilon && (13) \end{aligned}$$

Define $r \equiv \frac{\|\tilde{Q}(z)^{\frac{1}{2}}(\hat{z}-z)\|}{\tilde{\mu}(z)^{\frac{1}{4}}}$. It follows that

$$\hat{z} \in E(\tilde{Q}(z), z, \tilde{\mu}(z)^{\frac{1}{4}} r)$$

Thus, r is a good measure of the 'distance' of $\hat{z}$ from z.

Since we will now be working with $\tilde{Q}$ instead of with Q, we will need to relate the two.

Lemma 17 *Let $x \in int(P)$. Let $Q(x)$ and $H(x)$ be defined as before. Suppose that $\sigma_1(x) < \epsilon$. Define*

$$\tilde{Q}(x) \equiv Q(x) - \frac{\sigma_1(x)}{s_1(x)^2} cc^T$$

and

$$\tilde{H}(x) \equiv H(x) - \frac{cc^T}{s_1(x)^2}$$

(i.e., the lower bounding constraint is ignored). Then, for any $\xi \in \Re^n$, we have

$$\|\xi\|_{\tilde{Q}} \leq \|\xi\|_Q/\sqrt{1-\epsilon}$$

and

$$\|\xi\|_{\tilde{H}} \leq \|\xi\|_H/\sqrt{1-\epsilon}$$

Proof :
Consider the first part.

$$\begin{aligned} \|\xi\|_{\tilde{Q}}^2 &\equiv \xi^T\tilde{Q}^{-1}\xi \\ &= \xi^T(Q - \sigma_1 cc^T/s_1^2)^{-1}\xi \\ &\leq \xi^T Q^{-1}\xi/(1-\sigma_1\frac{c^TQ^{-1}c}{s_1^2}) \\ &\leq \xi^T Q^{-1}\xi/(1-\sigma_1\frac{c^TH^{-1}c}{\mu s_1^2}) \end{aligned}$$

$$
\begin{aligned}
&= \xi^T Q^{-1}\xi/(1-\sigma_1\frac{\sigma_1}{\mu}) \\
&\leq \xi^T Q^{-1}\xi/(1-\sigma_1) \\
&= \xi^T Q^{-1}\xi/(1-\epsilon)
\end{aligned}
$$

where in the third step, we used the fact that the smallest eigenvalue of the matrix $(I-\sigma_1 Q^{-1}cc^T/s_1^2)$ is equal to $1-(\sigma_1 c^T Q^{-1}c/s_1^2)$. In the penultimate step, we used the fact that $\mu \geq \sigma_{\min} = \sigma_1$ (see [20]), and the fact that $\sigma_1 < \epsilon$ was used in the last step. Similarly, we have

$$
\begin{aligned}
\|\xi\|_{\tilde{H}}^2 &\equiv \xi^T \tilde{H}^{-1}\xi \\
&= \xi^T (H - cc^T/s_1^2)^{-1}\xi \\
&\leq \xi^T H^{-1}\xi/(1-c^T H^{-1}c/s_1^2) \\
&= \xi^T H^{-1}\xi/(1-\sigma_1) \\
&\leq \xi^T H^{-1}\xi/(1-\epsilon)
\end{aligned}
$$

The result follows.

■

We need one other technical lemma before we can say that $\hat{z}$ is in some polyhedral neighbourhood of z.

Lemma 18 *Let $x \in int(P)$. Let $Q(x)$ and $\tilde{Q}(x)$ be defined as before. Let $\tilde{H}(x)$ and $\tilde{\mu}(x)$ be defined as in equations (12) and (13). We then have*

$$
\frac{a_i^T \tilde{Q}(x)^{-1} a_i}{s_i^2} \leq 1/\sqrt{\tilde{\mu}(x)(1-\epsilon)} \quad \forall\ i = 2\ to\ m
$$

Proof :
Let $\bar{A}$ denote the current constraint matrix A with the first row removed. Let $\bar{\Sigma}$ and $\bar{S}$ denote the bottom-right $(m-1)\times(m-1)$ submatrices of Σ and S respectively. Then consider the projection matrix

$$
M \equiv \bar{\Sigma}^{0.5}\bar{S}^{-1}\bar{A}(\bar{A}^T\bar{S}^{-2}\bar{\Sigma}\bar{A})^{-1}\bar{A}^T\bar{S}^{-1}\bar{\Sigma}^{0.5}
$$

We have

$$
M_{ii} = \sigma_{i+1}\frac{a_{i+1}^T \tilde{Q}(x)^{-1} a_{i+1}}{s_{i+1}^2} \leq 1 \quad \forall\, i = 1\ to\ m-1
$$

Therefore,

$$
\frac{a_i^T \tilde{Q}(x)^{-1} a_i}{s_i^2} \leq 1/\sigma_i \ \ \forall\, i = 2\ to\ m \tag{14}
$$

But

$$
\begin{aligned}
\sigma_i &= \frac{a_i^T H(x)^{-1} a_i}{s_i^2} \\
&= \frac{a_i^T (\tilde{H} + cc^T/s_1^2)^{-1} a_i}{s_i^2} \\
&\geq \frac{a_i^T \tilde{H}^{-1} a_i}{s_i^2}(1-\sigma_1) \\
&\geq \frac{a_i^T \tilde{Q}^{-1} a_i}{s_i^2}\tilde{\mu}(1-\sigma_1)
\end{aligned} \tag{15}
$$

Substituting (15) in (14) and using the fact that $\sigma_1 < \epsilon$, we get the desired result.

■

Lemma 19

$$\hat{z} \in E(\tilde{Q}(z), z, \tilde{\mu}(z)^{\frac{1}{4}} r) \subseteq \Sigma(z, r/\sqrt{1-\epsilon})$$

Proof :
From the very definition of r it is obvious that $\hat{z} \in E(\tilde{Q}(z), z, \tilde{\mu}(z)^{\frac{1}{4}} r)$. Only the second part needs to be shown. From lemma 18, we know that

$$\frac{a_i^T \tilde{Q}^{-1} a_i}{s_i^2} \leq 1/\sqrt{\tilde{\mu}(1-\epsilon)}$$

for $i = 2$ to m - i.e., all but the lower bounding constraint. Then, for $i = 2$ to m,

$$\begin{aligned} \frac{|a_i^T(\hat{z} - z)|}{s_i(z)} &\leq \tilde{\mu}^{\frac{1}{4}} r \sqrt{a_i^T \tilde{Q}^{-1} a_i} / s_i \\ &\leq r/\sqrt{1-\epsilon} \end{aligned}$$

Now consider the first constraint.

$$\begin{aligned} \frac{|c^T(\hat{z} - z)|}{s_1(z)} &\leq \frac{\tilde{\mu}^{\frac{1}{4}} r \sqrt{c^T \tilde{Q}^{-1} c}}{s_1(z)} \\ &\leq \frac{\tilde{\mu}^{-\frac{1}{4}} r \sqrt{c^T \tilde{H}^{-1} c}}{s_1(z)} \\ &\leq \tilde{\mu}^{-\frac{1}{4}} r \sqrt{\sigma_1/(1-\sigma_1)} \\ &\leq \tilde{\mu}^{-\frac{1}{4}} r \sqrt{\epsilon/(1-\epsilon)} \\ &\leq r\epsilon^{\frac{1}{4}}/\sqrt{1-\epsilon} \\ &\leq r/\sqrt{1-\epsilon} \end{aligned}$$

where the inequality in the last-but-one step follows from $\tilde{\mu} > \epsilon$.

The result follows.

■

Henceforth, we denote $\bar{r} \equiv r/\sqrt{1-\epsilon}$.

First, we show that if $r > \frac{1}{4}$, then there is a constant η_a such that $\hat{\psi}(\hat{z}) \geq \psi(z) + \eta_a$.

Lemma 20 *Let r be defined as above. If $r > \frac{1}{4}$, then*

$$\hat{\psi}(\hat{z}) \geq \psi(z) + \eta_a$$

where $\eta_a = 7.1 \times 10^{-7}$.

Proof :
Consider a straight line joining z and $\hat{z}$. Let $\bar{z}$ be a point on this line such that $\|\tilde{Q}(z)^{\frac{1}{2}}(\bar{z}-z)\| = 0.25\tilde{\mu}^{\frac{1}{4}}$. Let $\tilde{r} \equiv 0.25/\sqrt{1-\epsilon}$. Define $z(t) \equiv z + t(\bar{z}-z)$. Clearly we have $z(t) \in \Sigma(z,\tilde{r})$. Then we have

$$\begin{aligned} \frac{d\psi(z(t))}{dt} &= \nabla\psi(z(t))^T(\bar{z}-z) \\ \frac{d^2\psi(z(t))}{dt^2} &= (\bar{z}-z)^T\nabla^2\psi(z(t))(\bar{z}-z) \end{aligned}$$

Also,

$$\begin{aligned} \frac{d^2\psi(z(t))}{dt^2} &= (\bar{z}-z)^T\nabla^2\psi(z(t))(\bar{z}-z) \\ &\geq (\bar{z}-z)^T Q(z(t))(\bar{z}-z) \\ &\geq \frac{(1-\tilde{r})^2}{(1+\tilde{r})^4}(\bar{z}-z)^T Q(z)(\bar{z}-z) \\ &\geq \frac{(1-\tilde{r})^2}{(1+\tilde{r})^4}(\bar{z}-z)^T \tilde{Q}(z)(\bar{z}-z) \\ &= \frac{(1-\tilde{r})^2}{(1+\tilde{r})^4}\sqrt{\tilde{\mu}}/16 \\ &\geq \frac{(1-\tilde{r})^2}{(1+\tilde{r})^4}\sqrt{\epsilon}/16 \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{d\psi(z(t))}{dt} &= \int_0^t \frac{d^2\psi}{dt^2}dt \\ &\geq \frac{(1-\tilde{r})^2}{(1+\tilde{r})^4}t\sqrt{\epsilon}/16 \end{aligned}$$

Thus,

$$\begin{aligned} \hat{\psi}(\hat{z}) - \psi(z) &\geq \psi(\hat{z}) - \psi(z) \\ &\geq \psi(\bar{z}) - \psi(z) \\ &= \int_0^1 \frac{d\psi}{dt}dt \\ &\geq \int_0^1 \frac{(1-\tilde{r})^2}{(1+\tilde{r})^4}t(\sqrt{\epsilon}/16)dt \\ &= \frac{(1-\tilde{r})^2}{(1+\tilde{r})^4}\sqrt{\epsilon}/32 \\ &\geq 7.1 \times 10^{-7} \end{aligned}$$

■

Now, we will establish that we can guarantee a certain increase in the potential function even if $r \leq \frac{1}{4}$. But first, we need to prove a another technical lemma.

Lemma 21 *Let z be the current exact π-center. Let the cut $v^T x \geq v^T z$ be added, and let $\hat{z}$ denote the new exact π-center. Let r be as defined earlier. Define*

$$q(\hat{z}) \equiv \frac{v^T H(\hat{z})^{-1} v}{s_v(\hat{z})^2}$$

where $s_v(\cdot)$ denotes the slack corresponding to this new constraint. Under the hypothesis that $r < \frac{1}{4}$, we can say that

$$q(\hat{z}) \geq \frac{7}{8}\sqrt{\epsilon}(1-\epsilon)\frac{1}{r^2}$$

Proof :
First, we note that $\hat{z} \in E(\tilde{Q}(z), z, \tilde{\mu}(z)^{\frac{1}{4}} r)$. Thus,

$$\begin{aligned} s_v(\hat{z}) &\equiv v^T(\hat{z} - z) \\ &\leq \tilde{\mu}(z)^{\frac{1}{4}} r \sqrt{v^T \tilde{Q}(z)^{-1} v} \end{aligned}$$

Now,

$$\begin{aligned} q(\hat{z}) &= \frac{v^T H(\hat{z})^{-1} v}{s_v(\hat{z})^2} \\ &\geq (1-\bar{r})^2 \frac{v^T H(z)^{-1} v}{s_v(\hat{z})^2} \\ &\geq (1-\bar{r})^2 \frac{v^T H(z)^{-1} v}{\tilde{\mu}(z)^{\frac{1}{2}} r^2 v^T \tilde{Q}(z)^{-1} v} \\ &\geq (1-\bar{r})^2 (1-\epsilon) \frac{v^T \tilde{H}(z)^{-1} v}{\tilde{\mu}(z)^{\frac{1}{2}} r^2 v^T \tilde{Q}(z)^{-1} v} \\ &\geq (1-\epsilon)(1-\bar{r}^2)\sqrt{\tilde{\mu}}/r^2 \\ &\geq (1-\epsilon)(1-\bar{r}^2)\sqrt{\epsilon}/r^2 \\ &\geq \frac{7}{8}(1-\epsilon)\sqrt{\epsilon}/r^2 \end{aligned}$$

where $\bar{r} \equiv r/\sqrt{1-\epsilon}$, the fourth step uses lemma 17, and the last step uses $r < \frac{1}{4}$ and $\epsilon < \frac{1}{2}$.

∎

Now we may use this lemma to state the result we have been working towards.

Lemma 22 *Let z denote the current π-center, and let the cut $v^T x \geq v^T z$ be added. Let $\hat{z}$ denote the new π-center. Let r be defined as before. Under the hypothesis that $r < \frac{1}{4}$, we can say that*

$$\hat{\psi}(\hat{z}) \geq \psi(z) + 6.99 \times 10^{-4}$$

where ϵ is chosen to be 10^{-8}.

Proof :

$$
\begin{aligned}
\hat{\psi}(\hat{z}) &= \frac{1}{2}\log det(H(\hat{z}) + \frac{vv^T}{s_v(\hat{z})^2}) + \frac{c^T\hat{z} - l}{\pi} \\
&= \frac{1}{2}\log[(det\ H(\hat{z}))(1 + \frac{v^T H(\hat{z})^{-1} v}{s_v(\hat{z})^2})] + \frac{c^T\hat{z} - l}{\pi} \\
&= \psi(\hat{z}) + \frac{1}{2}\log(1 + q(\hat{z})) \\
&\geq \psi(z) + \frac{1}{2}\log(1 + q(\hat{z}))
\end{aligned}
$$

where the last step follows because z is the minimizer of $\psi(\cdot)$. Also, we have

$$
\begin{aligned}
q(\hat{z}) &\geq \frac{7}{8}(1-\epsilon)\sqrt{\epsilon}/r^2 \\
&\geq \frac{7}{8}(1-\epsilon)\sqrt{\epsilon}/0.25^2 \\
&\geq 1.399 \times 10^{-3}
\end{aligned}
$$

The result then follows.

■

At this point, we have shown that adding a cut always increases the potential function by at least a constant amount. We have simplified the analysis by considering the difference only between the current (exact) π-center and the (exact) π-center of the new polytope. This has been done only to minimize the additional book-keeping and focus on the main idea - similar results may easily be proven using approximate centers, by simply considering the difference in potential function between an approximate center and an exact one. Further, we have performed the analysis for the case where $\sigma_1 < \epsilon$, which is indeed the harder case. It is straightforward to repeat the above analysis for the case where this is not so. Indeed, this would be simpler since we need not consider the $\tilde{Q}$ matrix, and can do a more direct analysis. Furthermore, the lemmas in [20, 2] would be directly applicable if we had $\sigma_1 > \epsilon$, and we would not need to prove any results.

Now we will consider the affine step that we will take after adding the new cut through the current point. Again, for simplicity, we will assume that we have an exact π-center.

4.1.2 The Affine Step

Here we study the mechanics of adding a new cut right through the current (exact, for simplicity) π-center and finding the next approximate π-center. The idea of adding cuts right through the current interior point and using an affine step to regain an interior point may be found, for example, in [15, 11, 16]. In particular, if $\bar{x}$ is the point after taking an affine step, then we will show that $\hat{\psi}(\bar{x}) - \psi(z)$ is bounded above by a constant, which will then imply that the new π-center $\hat{z}$ may be approximated in $O(\sqrt{m})$ quasi-Newton steps. Again, as before, things are complicated by the fact that σ_1 may be less than ϵ, and again we need to use the fact that all other sigmas are greater than ϵ. Here we do the necessary analysis only for the case where $\sigma_1(z) < \epsilon$ – it is straightforward to see that the other case can only further strengthen our results.

Let $z \in int(P)$ be the current π-center. Let it be the case that $\sigma_1 < \epsilon$. Define $d_{aff} \equiv \frac{\beta}{r}\tilde{Q}(z)^{-1}v$, where $v^T x \geq v^T z$ is the new cut, β is a constant, and $r \equiv \|v\|_{\tilde{Q}(z)}$. Let $\bar{x} \equiv z + d_{aff}$. Our goal is to show that $\hat{\psi}(\bar{x}) - \psi(z)$ is bounded by a constant.

Lemma 23

$$\bar{x} \in \Sigma(z, \frac{\beta}{\sqrt{(1-\epsilon)\tilde{\mu}(z)}})$$

Proof :

$$\begin{aligned}
\bar{s} &= s(z) + ds \\
&= s(z) + \frac{\beta}{r} A\tilde{Q}^{-1}v \\
&= S(z)(e + \frac{\beta}{r} S^{-1}A\tilde{Q}^{-1}v)
\end{aligned}$$

Within this proof alone, denote the vector $\frac{\beta}{r}S^{-1}A\tilde{Q}^{-1}v$ by p. Then,

$$\begin{aligned}
\|p\|^2 &= (\frac{\beta}{r})^2 v^T\tilde{Q}^{-1}A^T S^{-2}A\tilde{Q}^{-1}v \\
&= (\frac{\beta}{r})^2 v^T\tilde{Q}^{-1}H(z)\tilde{Q}^{-1}v \\
&\leq (\frac{\beta}{r})^2 v^T\tilde{Q}^{-1}\tilde{H}(z)\tilde{Q}^{-1}v\frac{1}{1-\epsilon} \\
&\leq \frac{1}{\tilde{\mu}(z)}(\frac{\beta}{r})^2 v^T\tilde{Q}^{-1}\tilde{Q}\tilde{Q}^{-1}v\frac{1}{1-\epsilon} \\
&= \frac{\beta^2}{\tilde{\mu}(1-\epsilon)}
\end{aligned}$$

where we again used lemma 17 in the third step.

Thus, $\|p\| \leq \frac{\beta}{\sqrt{\tilde{\mu}(z)(1-\epsilon)}}$, which implies the result.

■

This immediately gives us a step length for the affine step - we can safely choose β to be $0.5\sqrt{\tilde{\mu}(z)} \geq 0.5 \times 10^{-4}$. The point $\bar{x}$ will then be in the interior of the new polytope, and it will also be far enough away for us to bound $\hat{\psi}(\bar{x}) - \psi(z)$ above by a constant.

Lemma 24 *Let z denote the current π-center, and let the cut $v^T x \geq v^T z$ be added. Suppose that $\sigma_1 < \epsilon = 10^{-8}$. Let $\bar{x} \equiv z + \frac{\beta}{r}\tilde{Q}(z)^{-1}v$, where $r \equiv \|v\|_{\tilde{Q}(z)}$. Let quantities with hats refer to the augmented system of inequalities. Let $\tau \equiv \beta/\sqrt{(1-\epsilon)\tilde{\mu}(z)}$. Then, we can say that*

$$\hat{\psi}(\bar{x}) \leq \psi(z) + \alpha_1$$

for some constant α_1.

Proof :
First, we observe that lemma 23 and the definition of τ imply that $\bar{x} \in \Sigma(z,\tau)$. Consider the line connecting z and $\bar{x}$. Define $z(t) \equiv z + t(\bar{x} - z)\ \ \forall t \in [0,1]$. We then have

$$\frac{d\psi(z(t))}{dt} = \nabla\psi(z(t))^T(\bar{x} - z)$$

Also,

$$\begin{aligned}
\frac{d^2\psi(z(t))}{dt^2} &= (\bar{x} - z)^T \nabla^2\psi(z(t))(\bar{x} - z) \\
&\leq 3(\bar{x} - z)^T Q(z(t))(\bar{x} - z) \\
&\leq 3\frac{(1+\tau)^2}{(1-\tau)^4}(\bar{x} - z)^T Q(z)(\bar{x} - z) \\
&\leq 3\frac{(1+\tau)^2}{(1-\epsilon)(1-\tau)^4}(\bar{x} - z)^T \tilde{Q}(z)(\bar{x} - z) \\
&= 3\frac{(1+\tau)^2}{(1-\epsilon)(1-\tau)^4}\beta^2
\end{aligned}$$

Thus, we have

$$\frac{d\psi(z(t))}{dt} \leq 3\frac{(1+\tau)^2}{(1-\epsilon)(1-\tau)^4}\beta^2 t$$

and

$$\psi(\bar{x}) - \psi(z) \leq 1.5\frac{(1+\tau)^2}{(1-\epsilon)(1-\tau)^4}\beta^2 \tag{16}$$

Also, we have

$$\begin{aligned}
q(\bar{x}) &\equiv v^T H(\bar{x})^{-1} v / s_v(\bar{x})^2 \\
&= \frac{v^T H(\bar{x})^{-1} v}{\beta^2 r^2} \\
&\leq (1+\tau)^2 \frac{v^T H(z)^{-1} v}{\beta^2 r^2} \\
&\leq (1+\tau)^2 \frac{v^T \tilde{H}(z)^{-1} v}{\beta^2 r^2} \\
&= (1+\tau)^2 \frac{v^T \tilde{H}(z)^{-1} v}{\beta^2 v^T \tilde{Q}(z)^{-1} v} \\
&\leq (1+\tau)^2/\beta^2
\end{aligned}$$

Thus,

$$\begin{aligned}
\hat{\psi}(\bar{x}) - \psi(z) &= \psi(\bar{x}) - \psi(z) + \frac{1}{2}\log(1 + q(\bar{x})) \\
&\leq 1.5\frac{(1+\tau)^2}{(1-\epsilon)(1-\tau)^4}\beta^2 + \frac{1}{2}\log(1 + \frac{(1+\tau)^2}{\beta^2})
\end{aligned}$$

Finally, choosing $\beta = 0.5 \times \sqrt{\tilde{\mu}(z)}$, we have

$$0.5 \times 10^{-4} \leq \beta \leq 0.5$$

and $\tau \leq 0.51$. Using these, it follows that $\hat{\psi}(\bar{x}) - \psi(z)$ is bounded above by a constant. Thus, the convergence lemmas 11 and 15 imply that the next approximate π-center may be found in $O(\sqrt{m})$ Newton steps.

$\blacksquare$

4.2 Dropping A Cut

In this section, we analyze what happens when case 1 of the algorithm has occurred. It is therefore the case that $\sigma_j = \sigma_{\min} < \epsilon$ for some $j \neq 1$, and the minimum is only over 2 to m. Without loss of generality, we may assume that $\sigma_m = \min_{j=2 \text{ to } m} \sigma_j$. Thus, at this stage, the m^{th} constraint is being dropped. This leads to a decrease in the potential function $\psi(\cdot)$. However, our goal is to show that this decrease is not too much. Specifically, we have to show that the maximum possible decrease when we drop a constraint is definitely less than the least possible increase when we add a cut.

As in the analysis of constraint addition, we run into trouble because of the fact that it is possible to have $\sigma_1 < \sigma_m$, and thus the results of Vaidya [20] or Anstreicher [2] do not apply here. Therefore, as before, we consider only the harder case of when $\sigma_1 < \sigma_m = \sigma_{\min}$. We prove our results for this case. It is very straightforward that the results generalize to the case where $\sigma_1 \geq \sigma_m$.

Lemma 25 *Let $P = \{x | a_i^T x \geq b_i, i = 1 \text{ to } m\}$. (Recall that the first constraint is the lower bounding constraint. Thus, $a_1 = c$.) Let it be the case that $\sigma_1 < \sigma_m = \min_{i=2 \text{ to } m} \sigma_i < \epsilon$. Let the new polytope be defined by dropping the m^{th} constraint. Let $\bar{z}$ be the new π-center, and let $\bar{\psi}(\cdot)$ denote the potential function corresponding to the new polytope. Define*

$$\begin{aligned}
\tilde{Q}(z) &= \sum_{i=2}^{m} \sigma_i(z) a_i a_i^T / s_i(z)^2 \\
\tilde{H}(z) &= \sum_{i=2}^{m} a_i a_i^T / s_i(z)^2 \\
\tilde{\mu}(z) &= \inf_{\xi \in \Re^n} \frac{\xi^T \tilde{Q}(z) \xi}{\xi^T \tilde{H}(z) \xi} \geq \sigma_m
\end{aligned}$$

Also define $r \equiv \frac{\|\tilde{Q}(z)^{\frac{1}{2}}(\bar{z} - z)\|}{\tilde{\mu}(z)^{\frac{1}{4}}}$. Then $r < \frac{1}{4}$.

Proof :
Suppose that $r > \frac{1}{4}$. Let $\bar{x}$ be a point on the line joining z and $\bar{z}$ such that

$$\frac{\|\tilde{Q}(z)^{\frac{1}{2}}(\bar{x} - z)\|}{\tilde{\mu}(z)^{\frac{1}{4}}} = \frac{1}{4}$$

Within this proof alone, define $\tilde{r} \equiv 0.25/\sqrt{1 - \epsilon}$. Exactly as in the proof of lemma 19, we can show that

$$\bar{x} \in E(\tilde{Q}(z), z, 0.25 \tilde{\mu}^{\frac{1}{4}}) \subset \Sigma(z, 0.25/\sqrt{1 - \sigma_m}) \subseteq \Sigma(z, 0.25/\sqrt{1 - \epsilon})$$

Let $z(t) \equiv z + t(\bar{x} - z)$. We then have, as before,

$$\frac{d^2 \psi(z(t))}{dt^2} = (\bar{x} - z)^T \nabla^2 \psi(z(t)) (\bar{x} - z)$$

$$
\begin{aligned}
&\geq (\bar{x}-z)^T Q(z(t))(\bar{x}-z)\\
&\geq (\bar{x}-z)^T \tilde{Q}(z(t))(\bar{x}-z)\\
&\geq \frac{(1-\tilde{r})^2}{(1+\tilde{r})^4}(\bar{x}-z)^T \tilde{Q}(z)(\bar{x}-z)\\
&= \frac{(1-\tilde{r})^2}{(1+\tilde{r})^4}\sqrt{\tilde{\mu}(z)}/16
\end{aligned}
$$

Therefore, we can say that

$$\frac{d\psi(z(t))}{dt} \geq \frac{(1-\tilde{r})^2}{(1+\tilde{r})^4} t\sqrt{\tilde{\mu}(z)}/16$$

which then leads to the result that $\psi(\bar{x}) - \psi(z) \geq (1-\tilde{r})^2(1+\tilde{r})^{-4}\sqrt{\tilde{\mu}(z)}/32$. Using $\tilde{r} = 0.25/\sqrt{1-\epsilon}$ and $\epsilon = 10^{-8}$, we get $\psi(\bar{x}) - \psi(z) \geq \frac{0.23}{32}\sqrt{\tilde{\mu}(z)} \geq \frac{0.23}{32}\sqrt{\sigma_m(z)}$. Also,

$$
\begin{aligned}
\sigma_m(\bar{x}) &= \frac{a_m^T H(\bar{x})^{-1} a_m}{s_m(\bar{x})^2}\\
&\leq \frac{(1+\tilde{r})^2}{(1-\tilde{r})^2}\sigma_m(z)\\
&\leq 2.78\sigma_m(z)
\end{aligned}
$$

Thus, we can write

$$
\begin{aligned}
\bar{\psi}(\bar{x}) &= \psi(\bar{x}) + \frac{1}{2}\log(1-\sigma_m(\bar{x}))\\
&\geq \psi(\bar{x}) + \frac{1}{2}\log(1-2.78\sigma_m(z))\\
&\geq \psi(z) + \frac{0.23}{32}\sqrt{\sigma_m(z)} + \frac{1}{2}\log(1-2.78\sigma_m(z))\\
&\geq \psi(z) + \frac{0.23}{32}\sqrt{\sigma_m} - \frac{1.39\sigma_m}{(1-2.78\sigma_m)}
\end{aligned}
$$

We also know that $\sigma_m < \epsilon = 10^{-8}$. It may be verified that the function $\frac{0.23}{32}\sqrt{x} - \frac{1.39x}{1-2.78x}$ is nonnegative and increasing in x for $x \in [0, 10^{-8}]$. Thus,

$$\bar{\psi}(\bar{x}) \geq \psi(z) \geq \bar{\psi}(z)$$

But strict convexity of $\bar{\psi}(\cdot)$ implies that it should be strictly decreasing along the line joining z and $\bar{z}$, and thus we must have $\bar{\psi}(\bar{x}) < \psi(z)$. Thus we have a contradiction. It therefore has to be the case that $r < \frac{1}{4}$.

■

Now, we are able to state the result we require.

Lemma 26 *Under the conditions of lemma 25, we have*

$$\bar{\psi}(\bar{z}) \geq \psi(z) - \eta_d$$

where $\eta_d = 2.781 \times 10^{-8}$. *Thus, we also have* $\eta_d < \eta_a$.

Proof :
Using techniques used in the previous lemma, we can write

$$\psi(\bar{z}) \geq \psi(z) + \frac{0.23}{2}\sqrt{\tilde{\mu}(z)}r^2$$

and

$$\sigma_m(\bar{z}) \leq k(r)\sigma_m(z)$$

where $k(r) \equiv \frac{(1+\bar{r})^2}{(1-\bar{r})^2}$, and $\bar{r} = r/\sqrt{1-\epsilon}$. Thus,

$$\begin{aligned}
\bar{\psi}(\bar{z}) &\geq \psi(z) + \frac{0.23}{2}\sqrt{\tilde{\mu}(z)}r^2 + \frac{1}{2}\log(1-\sigma_m(\bar{z})) \\
&\geq \psi(z) + \frac{0.23}{2}\sqrt{\sigma_m(z)}r^2 + \frac{1}{2}\log(1-k(r)\sigma_m(z)) \\
&\geq \psi(z) + \frac{1}{2}\log(1-k(r)\sigma_m(z)) \\
&\geq \psi(z) - \frac{1}{2}\frac{k(r)\sigma_m(z)}{(1-k(r)\sigma_m(z))} \\
&\geq \psi(z) - \frac{1}{2}\frac{k(0.25)\sigma_m(z)}{(1-k(0.25)\sigma_m(z))} \\
&\geq \psi(z) - \frac{1}{2}\frac{k(0.25)\epsilon}{(1-k(0.25)\epsilon)} \\
&\geq \psi(z) - 2.781 \times 10^{-8}
\end{aligned}$$

proving the result.

∎

Corollary 1 *When case 1 of the algorithm occurs and a constraint is dropped, no more than $O(\sqrt{m} + \log m)$ quasi-Newton steps are needed to find the next approximate π-center.*

Thus, we have established that adding a constraint through the current point and dropping a constraint as per the algorithm are not too expensive regarding the time needed to find the next approximate center. Now, we need to analyze subcase 2.2 of the algorithm, when we drop the barrier parameter.

4.3 Reducing the Barrier Parameter

In this subsection, we analyze subcase 2.2 of the algorithm, where we seek progress in the objective function by reducing the barrier by a constant factor $\rho \in [0.5, 1)$.

First, we have to justify the lower bound we use.

Lemma 27 *Let $x \in int(P)$ satisfy $\nabla\psi(x,\pi)^T Q(x)^{-1}\nabla\psi(x,\pi) \leq 1/\sqrt{m}$. Then $c^T x - z^* \leq 2n\pi$, where z^* denotes the optimal value of $c^T x$ over the current polytope. Also, $y \equiv \pi S^{-1}\Sigma(e - S^{-1}Ad)$, where $d \equiv Q^{-1}\nabla\psi(x,\pi,l)$ is feasible in the dual with a duality gap of smaller than $2n\pi$.*

Proof :
See Anstreicher [3], lemma 3.2.

■

This implies that in our algorithm, whenever subcase 2.2 occurs, we are justified in setting the lower bound to $c^T x - 2n\pi$ after ensuring that x satisfies the required proximity criterion. (This is true because the current polytope at that stage is guaranteed to contain x^*, the optimal solution to our original problem of interest).

Now we need to address the issue of how many quasi Newton steps may be required to find the next approximate π-center after reducing the barrier parameter. We need some technical lemmas first before we can answer this.

Here, as before, we will assume that the current point is an exact π-center for the sake of simplicity. It is not hard to extend these results for the general case.

Lemma 28 *Suppose that we have a point $z \in int(P)$, where P denotes the current polytope. Suppose that z is an exact π-center. Define $\tilde{H}(z) \equiv \sum_{i=2}^{m} a_i a_i^T / s_i(z)^2$ as before. Then we have*

$$\frac{c^T \tilde{H}(z)^{-1} c}{n^2 \pi^2} \leq 1/(1 - \sigma_1(z))$$

Proof :
Since z is an exact π-center, we have

$$c/\pi = A^T \Sigma S^{-1} e$$

Therefore,

$$\begin{aligned}
\frac{c^T H(z)^{-1} c}{\pi^2} &= e^T \Sigma S^{-1} A (A^T S^{-2} A)^{-1} A^T S^{-1} \Sigma e \\
&= \sigma S^{-1} A (A^T S^{-2} A)^{-1} A^T S^{-1} \sigma \\
&\leq \|\sigma\|^2 \\
&\leq \|\sigma\|_1^2 \\
&= n^2
\end{aligned}$$

where $\sigma \equiv [\sigma_1 \ \ \sigma_2 \ \ ...\sigma_m]^T$. Now, using the Sherman-Morrison-Woodbury formula,

$$\begin{aligned}
\frac{c^T \tilde{H}(z)^{-1} c}{\pi^2} &= \frac{c^T H(z)^{-1} c}{\pi^2 (1 - \sigma_1)} \\
&\leq n^2 / (1 - \sigma_1)
\end{aligned}$$

which proves the result.

■

Now we will bound the decrease in $\psi(\cdot)$ that results when the barrier parameter is dropped. Here, we need to be much more careful with our notation, since it must be emphasized that changing the lower bound changes the first constraint, and therefore the polytope. This is made explicit by writing $\psi(x, \pi, l)$.

Lemma 29 *Assume we have a point x_1, which is an exact π-center corresponding to a barrier parameter value of π_1. Case 2.2 of the algorithm occurs, and we reduce π_1 to $\pi_2 = \rho\pi_1$, where $\rho \in [0.5, 1)$. Let x_2 denote the new exact π-center. Let l_1 and l_2 denote the lower bounds that are active at the beginning and end of this occurrence of case 2.2.*

We then have

$$\psi(x_2, \pi_2, l_2) \geq \psi(x_1, \pi_1, l_1) - 8n$$

Proof :
Let x_0 denote the point which corresponds to the previous occurrence of case 2.2. Thus, x_0 is an approximate center for value $\pi = \pi_0 = \pi_1/\rho$. At this (previous) stage, the candidate for a new lower bound would have been $c^T x_0 - 2n\pi_0$. Thus, we can say

$$l_1 \geq c^T x_0 - 2n\pi_0$$

Therefore, we have $c^T x_1 - l_1 \leq c^T x_1 - c^T x_0 + 2n\pi_0$. But since x_0 was feasible in the overall problem, $c^T x_0 \geq l_2 \geq c^T x_1 - 2n\pi_1$. Thus,

$$(c^T x_1 - l_1)/\pi_1 \leq 2n\pi_1 + 2n\pi_0 = 2n\pi_1(1 + 1/\rho) \leq 6n \tag{17}$$

where we use the fact that $\rho \in [0.5, 1)$.

Now,

$$\begin{aligned}
\psi(x_2, \pi_2, l_2) - \psi(x_1, \pi_1, l_1) &= \frac{c^T x_2 - l_2}{\pi_2} - \frac{c^T x_1 - l_1}{\pi_1} + V(x_2, l_2) - V(x_1, l_1) \\
&\geq -\frac{c^T x_1 - l_1}{\pi_1} + V(x_2, l_2) - V(x_1, l_1) \\
&\geq -6n + V(x_2, l_2) - V(x_1, l_1) \\
&\geq -6n + V(x_2, l_1) - V(x_1, l_1) + V(x_2, l_2) - V(x_2, l_1)
\end{aligned}$$

Since $l_2 \geq l_1$, for all the points in the new polytope we must have $V(x, l_2) \geq V(x, l_1)$. Therefore,

$$\psi(x_2, \pi_2, l_2) - \psi(x_1, \pi_1, l_1) \quad \geq \quad -6n + V(x_2, l_1) - V(x_1, l_1)$$

Now we need to bound the difference $V(x_2, l_1) - V(x_1, l_1)$. Since x_1 minimizes $\psi(\cdot, \pi_1, l_1)$, we have

$$\frac{c^T x_1}{\pi_1} + V(x_1, l_1) \leq \frac{c^T x_2}{\pi_1} + V(x_2, l_1)$$

or

$$V(x_2, l_1) - V(x_1, l_1) \geq \frac{c^T x_1 - c^T x_2}{\pi_1}$$

Unfortunately, we can no longer say that $c^T x_2 \leq c^T x_1$, since by possibly changing the lower bound we may have changed the polytope. However, since x_1 is feasible, and because $c^T x_2 - 2n\pi_2$ is a valid lower bound on the optimal objective function value (irrespective of whether or not x_2 is feasible), we can say that $c^T x_2 - 2n\pi_2 \leq c^T x_1$, or

$$c^T x_1 - c^T x_2 \geq -2n\pi_2$$

Therefore,

$$(c^T x_1 - c^T x_2)/\pi_1 \geq -2n\pi_2/\pi_1 \geq -2n$$

Thus, we have

$$\psi(x_2, \pi_2, l_2) - \psi(x_1, \pi_1, l_1) \quad \geq \quad -6n - 2n = -8n$$

■

Lemma 30 *Assume that we have a point x_1 that is an exact π-center corresponding to π_1, and case 2.2 of the algorithm occurs. Therefore, the lower bound may possibly have been reset from l_1 to l_2, and $\pi_2 = \rho\pi_1$. We can then say*

$$\psi(x_1, \pi_2, l_2) \leq \psi(x_1, \pi_1, l_1) + 9n$$

Proof :
Let $l \equiv c^T x_1 - 2n\pi_1$. If $l > l_1$, then $l_2 = l$, and otherwise $l_2 = l_1$. We have

$$\begin{aligned}
\psi(x_1, \pi_2, l_2) - \psi(x_1, \pi_1, l_1) &= \frac{c^T x_1 - l_2}{\pi_2} - \frac{c^T x_1 - l_1}{\pi_1} + V(x_1, l_2) - V(x_1, l_1) \\
&\leq \frac{c^T x_1 - l_2}{\pi_2} + V(x_1, l_2) - V(x_1, l_1)
\end{aligned}$$

Now, $l_2 \geq c^T x_1 - 2n\pi_1$. Therefore,

$$\begin{aligned}
\psi(x_1, \pi_2, l_2) - \psi(x_1, \pi_1, l_1) &\leq (2n\pi_1/\pi_2) + V(x_1, l_2) - V(x_1, l_1) \\
&\leq 4n + V(x_1, l_2) - V(x_1, l_1)
\end{aligned}$$

Case 1 : $l_2 = l_1$
In this case we are done, since $V(x_1, l_2) = V(x_1, l_1)$. Therefore,

$$\psi(x_1, \pi_2, l_2) - \psi(x_1, \pi_1, l_1) \leq 4n$$

Case 2 : $l_2 = c^T x_1 - 2n\pi_1$
Here we need to bound $V(x_1, l_2) - V(x_1, l_1)$.

$$\begin{aligned}
V(x_1, l_2) - V(x_1, l_1) &= \frac{1}{2}\log\det(\tilde{H}(x_1)) + \frac{1}{2}\log(1 + \frac{c^T \tilde{H}^{-1} c}{(c^T x_1 - l_2)^2}) \\
&\quad - \frac{1}{2}\log\det(\tilde{H}(x_1)) - \frac{1}{2}\log(1 + \frac{c^T \tilde{H}^{-1} c}{(c^T x_1 - l_1)^2}) \\
&\leq \frac{1}{2}\log(1 + \frac{c^T \tilde{H}^{-1} c}{(c^T x_1 - l_2)^2}) \\
&= \frac{1}{2}\log(1 + \frac{c^T \tilde{H}^{-1} c}{4n^2\pi_1^2})
\end{aligned}$$

But by lemma 28, we have

$$\frac{c^T \tilde{H}(x_1)^{-1} c}{4n^2\pi_1^2} \leq 1/(4(1-\sigma_1)) \leq 1/(4\epsilon)$$

where the last inequality assumes that $m \geq n \geq 2$, and since case 2.2 has occurred, all other sigmas are greater than ϵ. Therefore,

$$V(x_1, l_2) - V(x_1, l_1) \leq \frac{1}{2}\log(1 + 0.25\epsilon^{-1}) \leq 8.52$$

using $\epsilon = 10^{-8}$. Thus, we have

$$\begin{aligned}
\psi(x_1, \pi_2, l_2) - \psi(x_1, \pi_1, l_1) &\leq 4n + 8.52 \\
&\leq 9n
\end{aligned}$$

where the last inequality again uses $n \geq 2$. This proves the result.

■

Thus, to find the next π-center after dropping the barrier parameter, a reduction in $\psi(\cdot)$ of no more than $17n$ is required. It follows that no more than $O(n\sqrt{m}+\log m)$ quasi Newton steps are necessary to find the next π-center.

Corollary 2 *An occurrence of case 2.2 of the algorithm will need a total of no more than $O(n\sqrt{m}+\log m)$ quasi Newton steps.*

■

4.4 Overall Complexity

Finally we are in a position to analyze the overall complexity of the algorithm.

Theorem 2 *This algorithm will terminate in at most $O(n\sqrt{n}L)$ quasi Newton steps.*

Proof :
At any stage in the algorithm, suppose that z is the current point that we have. Again, for simplicity, assume that z is an exact π-center, and is therefore the minimizer of $\psi(\cdot,\pi,l)$. Let P denote the current polytope, which is defined by m constraints. Consider the polytope

$$P^* \equiv P \cap \{x | c^T x \leq u\}$$

where $u \equiv \max(c^T z, c^T x_{prev})$. Thus, P^* is guaranteed to contain the current point as well as the optimal solution. (This is because x_{prev} is the point at which subcase 2.2 occurred most recently, which implies that it was feasible. Therefore, the optimal solution must have smaller objective function value). We are interested in bounding the volume of P^* above.

For the analysis alone, imagine that the cut $c^T x \leq u$ has been added. Thus, the polytope P^* is now defined by $m+1$ constraints. Let quantities with stars refer to this augmented polytope. Let x_a and x_v denote its analytic and volumetric centers respectively (of the augmented polytope). It is known (see, e.g., [5]) that

$$P^* \subseteq E(H^*(x_a), x_a, m+1)$$

Therefore

$$\begin{aligned}
vol(P^*) &\leq (2(m+1))^n(\det(H^*(x_a)))^{-\frac{1}{2}} \\
&\leq (2(m+1))^n(\det(H^*(x_v)))^{-\frac{1}{2}} \\
&\leq (2(m+1))^n(\det(H(x_v)))^{-\frac{1}{2}} \\
&= (2(m+1))^n e^{-V(x_v)} \\
&\leq (2(\frac{n}{\epsilon}+3))^n e^{-V(x_v)} \qquad (18)
\end{aligned}$$

where the last step follows because our dropping criterion implies that $m \leq n/\epsilon+2$. (This is because all the sigmas have to sum to n).

Now, since z is the minimizer of $\psi(\cdot,\pi,l)$, we have

$$\frac{c^T x_v - l}{\pi} + V(x_v) \geq \frac{c^T z - l}{\pi} + V(z)$$

or

$$V(x_v) \geq V(z) - \frac{c^T x_v - c^T z}{\pi}$$

If it is the case that $c^T x_v \leq c^T z$, then $V(x_v) \geq V(z)$. Otherwise, it must be the case that $u = c^T x_{prev}$. We can then say

$$\begin{aligned} c^T z &\geq l \\ &\geq c^T x_{prev} - 2n\pi_{prev} \\ &= c^T x_{prev} - 2n\pi/\rho \\ &\geq c^T x_{prev} - 4n\pi \end{aligned}$$

where the second step follows because the lower bound must have been set at the previous occurrence of case 2.2, or it must have been even better than the candidate. Also, since $x_v \in P^*$, we have $c^T x_v \leq u = c^T x_{prev}$. Therefore,

$$c^T x_v - c^T z \leq 4n\pi$$

Thus, $V(x_v) \geq V(z) - 4n$. Finally, $V(z) = \psi(z, \pi, l) - (c^T z - l)/\pi$. From an analysis exactly like the one that preceded equation (17), we can say that

$$\frac{c^T z - l}{\pi} \leq 6n$$

Therefore, $V(z) \geq \psi(z, \pi, l) - 6n$. Putting it all together, we have

$$V(x_v) \geq \psi(z, \pi, l) - 10n \tag{19}$$

Now it is clear from equations (19) and (18) that if $\psi(\cdot) > O(nL)$, then the volume of the polytope P^* would have fallen below 2^{-nL}, which is too small to contain a ball of radius 2^{-L}. Thus the algorithm must terminate if $\psi(\cdot) > O(nL)$.

Let us now look at the stages that occur in our algorithm. Each occurrence of case 2.2 can set the value of $\psi(\cdot)$ back by upto $O(n)$ — and there can be no more than $O(L)$ of these. Thus, the total setback to $\psi(\cdot)$ because of the case 2.2 stages is $O(nL)$. Also, since the number of constraint deletions cannot be more than the number of constraint additions, and since $\eta_a > \eta_d$, it follows that in $O(nL)$ non-case 2.2 stages, the value of $\psi(\cdot)$ will be sufficiently high. Thus, there can be no more than $O(nL)$ non case 2.2 stages, and $O(L)$ case 2.2 stages. Also, each occurrence of case 2.2 requires no more than $O(n\sqrt{n} + \log n)$ quasi Newton steps (since $m = O(n)$), and each non case 2.2 stage requires no more than $O(\sqrt{n} + \log n)$ quasi Newton steps. Thus, the total number of quasi Newton steps required by the algorithm is $O(n\sqrt{n}L + nL\log n) = O(n\sqrt{n}L)$.

$\blacksquare$

5 Adding Multiple Cuts

In this section we discuss a version of our long step algorithm where multiple cuts are added in case 2.1. The algorithm is exactly the same as before, except that we may be adding $p \leq n$ cuts in case 2.1. For other work in the literature, pertaining to adding multiple cutting planes, see Ye [24] and Ramaswamy and Mitchell [16].

Suppose that case 2.1 has occurred, and we have an exact π-center (for simplicity) z. Recall that again it is possible to have $\sigma_1(z) < \epsilon$, although all the other sigmas are greater than ϵ. As before, this is the harder case, and we will assume this to be the case in our analysis here. Suppose that p cuts are to be added, and these are given by $v_i^T x \geq v_i^T z$, for $i = 1$ to p. We will assume that the vectors v_i are ortho positive and normal with respective to $\tilde{Q}$ — i.e., they have the property

$$< v_i, v_j > \geq 0$$

and

$$\|v_i\|^2_{\tilde{Q}(z)} \equiv < v_i, v_i > = 1$$

where $< x, y > \equiv x^T \tilde{Q}(z)^{-1} y$ for any $x, y \in \Re^n$.

For a motivation of this, see Ramaswamy and Mitchell [16]. For completeness, we include a brief description of the Selective Orthonormalization procedure by which the v_i's may be generated from p arbitrary cuts. For further details, see [16]. Let the p arbitrary cuts be given by $u_i^T x \geq u_i^T z$.

Selective Orthonormalization Procedure

Step 0 : Define $v_1 \equiv \frac{u_1}{\|u_1\|_{\tilde{Q}}}$. Set $k = 2$.
Step k : Define $J(k) = \{i \in \{1, 2, ..., k-1\} | < u_k, v_i > \quad < 0\}$.
Set $g_k \equiv u_k - \sum_{i \in J(k)} < u_k, v_i > v_i$.
Set $v_k \equiv \frac{g_k}{\|g_k\|_{\tilde{Q}}}$
Step Check : If $k = p$, STOP. Else set $k = k + 1$, goto step k.

It is shown in [16] that the v_is define valid cuts, in addition to having other nice properties. Most importantly, a feasible direction for the affine step is now easy.

Define

$$d_{aff} \equiv \frac{\beta}{r} \tilde{Q}(z)^{-1} (\sum_{i=1}^{p} v_i)$$

where β is a constant we will soon choose, and

$$r \equiv \| \sum_{i=1}^{p} v_i \|_{\tilde{Q}} \leq p$$

Lemma 31 *Let $\bar{x}$ denote the point after an affine step — i.e., $\bar{x} \equiv z + d_{aff}$. Then,*

$$\bar{x} \in \Sigma(z, \frac{\beta}{\sqrt{\tilde{\mu}(z)(1-\epsilon)}})$$

Proof :
Exactly like the proof of lemma 23.

∎

As before, we may safely choose $\beta = 0.5\sqrt{\tilde{\mu}(z)} \geq 0.5 \times 10^{-4}$. We wish to bound $\hat{\psi}(\bar{x}) - \psi(z)$ above by a constant. But first, we state another lemma.

It will be convenient to denote $U \equiv [v_1 \ \ v_2 \ \ ... \ v_p]$. Then we can write

$$\sum_{i=1}^{p} v_i = Ue$$

Lemma 32 *Let z denote the current exact π-center. Let the p cuts defined by the v_is as above be added, and let $\bar{x}$ denote the point after an affine step. Let $\tilde{Q}$, $\tilde{H}$ and $\tilde{\mu}$ be defined as in equations (11), (12) and (13) respectively. Let $\tau \equiv \frac{\beta}{\sqrt{\tilde{\mu}(z)(1-\epsilon)}}$. Define*

$$q_i(\bar{x}) \equiv v_i^T H(\bar{x})^{-1} v_i / (v_i^T \bar{x} - v_i^T z)^2$$

for $i = 1$ to p. Then,

$$q_i(\bar{x}) \leq (1+\tau)^2 p^2/\beta^2$$

Proof :

$$\begin{aligned}
q_i(\bar{x}) &= \frac{v_i^T H(\bar{x})^{-1} v_i}{(v_i^T \bar{x} - v_i^T z)^2} \\
&= \frac{v_i^T H(\bar{x})^{-1} v_i}{(\frac{\beta}{r})^2 < v_i, Ue >^2} \\
&\leq \frac{(1+\tau)^2 r^2}{\beta^2} \frac{v_i^T H(z)^{-1} v_i}{< v_i, Ue >^2} \\
&\leq \frac{(1+\tau)^2 r^2}{\beta^2} \frac{v_i^T \tilde{H}(z)^{-1} v_i}{< v_i, Ue >^2} \\
&\leq \frac{(1+\tau)^2 r^2}{\beta^2} \frac{v_i^T \tilde{Q}(z)^{-1} v_i}{< v_i, Ue >^2} \\
&= \frac{(1+\tau)^2 r^2}{\beta^2} \frac{< v_i, v_i >}{< v_i, Ue >^2} \\
&\leq \frac{(1+\tau)^2 p^2}{\beta^2}
\end{aligned}$$

where the last step follows because $r \leq p$, $< v_i, v_i >= 1$ and $< v_i, Ue > \geq < v_i, v_i >= 1$.

■

Lemma 33 *Under the hypothesis of lemma 32, we have*

$$\psi(\bar{x}) \leq \psi(z) + 1.5\beta^2 \frac{(1+\tau)^2}{(1-\epsilon)(1-\tau)^4}$$

Proof :
Exactly like the analysis preceding equation (16).

■

It is a direct consequence of lemmas 32 and 33 that

$$\hat{\psi}(\bar{x}) - \psi(z) \leq O(p \log p)$$

(This is easily seen from the definition of $\hat{\psi}(\cdot)$ here).

Finally, if $\hat{z}$ denotes the new π-center after the addition of multiple cuts, we can say that

$$\hat{\psi}(\hat{z}) - \psi(z) \geq \eta_a$$

We do not have to prove this result — the lemmas of section 4.1.1 apply here directly since adding multiple cuts can only strengthen our result.

Thus, it follows that the next approximate π-center may be found in $O(\sqrt{n}p \log p + \log n)$ quasi Newton steps. This leads then to an overall complexity of $O(n^{1.5}Lp \log p + nL \log n)$ quasi Newton steps.

6 Conclusions

We have used Vaidya's algorithm as a framework to create a long step cutting plane algorithm based on the volumetric barrier function. Since it is a long step method, our algorithm has a slightly weaker complexity of $O(n^{1.5}L)$ Newton steps. Recently, many path following algorithms have been discussed in the literature - some are based on the logarithmic barrier function (see [17, 8, 21]), some on the volumetric barrier function (see [22, 20, 1, 3]), and some on the hybrid volumetric-logarithmic barrier function (see [22, 3]). In particular, the best known long step complexity for linear programming (without using cutting planes) is $O((nm)^{0.5}L)$, obtained using the hybrid potential function (see [22, 3]). (The same hybrid barrier yields the best known short step complexity of $O((nm)^{0.25})$ for an algorithm without cutting planes). However, all these methods have complexity depending on m, and thus when m is large enough (or infinite, as for arbitrary convex programs) these algorithms are not satisfactory. This paper is the only analysis of a long step, cutting plane algorithm based on the volumetric barrier that we are aware of.

There are two algorithms that we can compare ours with. First, Vaidya's original algorithm, which has an $O(nL)$ complexity. However, this backs off the cut instead of adding it through the current point. It is true that a version of Vaidya's algorithm which adds cuts right through the current volumetric center can be designed. An affine step followed by centering steps can then be taken to find the next volumetric center. Such an algorithm would have a complexity of $O(n^{1.5}L)$, same as ours. However, our algorithm has two advantages over this - first, we do not need to add a cut to seek progress in the objective, and secondly, we are able to guarantee cutting the duality gap by half in such iterations. Vaidya's algorithm (even with cuts added right through) does not do that. Another algorithm we could compare our algorithm with is the long step cutting plane method based on the logarithmic center, which we discuss in [14]. This algorithm has a complexity of $O(nL^2)$. However, for theoretical reasons, L is required to be $O(n\phi)$, where ϕ is the (maximum) number of bits required to store any row of the constraint matrix A. Thus, $L = O(n)$, which implies that our algorithm here improves theoretically on the algorithm in [14]. However, from a computational standpoint, methods based on the volumetric center have yet to be tested in the literature. It would be interesting to see how our algorithm here compares with, say, that of [14] in terms of performance.

We also note that our algorithm is applicable to arbitrary convex programs, provided the *separation problem* is solvable (see [12]). We also assume that the objective function is linear, but any convex optimization problem can be reformulated to have a linear objective (see [9]).

References

[1] Anstreicher, K. M. "Large Step Volumetric Potential Reduction Algorithms for Linear Programming". Department of Management Sciences, University of Iowa, Iowa City. May 5, 1994.

[2] Anstreicher, K. M. "On Vaidya's Volumetric Cutting Plane Method for Convex Programming". Department of Management Sciences, University of Iowa, Iowa City. Sep 27, 1994.

[3] Anstreicher, K. M. "Volumetric Path Following Algorithms for Linear Programming". Department of Management Sciences, University of Iowa, Iowa City. September 27, 1994.

[4] Atkinson, D. S. "Scaling and Interior Point Methods in Optimization". Ph.D Dissertation, Coordinated Science Laboratory, College of Engineering, University of Illinois at Urbana Champaign, USA, 1992.

[5] Atkinson, D.S and Vaidya, P.M. "An Analytic Center Based Cutting Plane Algorithm for Convex Programming". Technical report, Dept. of Mathematics, University of Illinois at Urbana Champaign, June 1992.

[6] Atkinson, D.S and Vaidya, P.M. "A Scaling Technique For Finding The Weighted Analytic Center Of A Polytope". Mathematical Programming, 1992 (NOV 2) Vol. 57 pp 163 - 192.

[7] Bertsimas, D. and Orlin, J. B. "A Technique for Speeding Up the Solution of the Lagrangean Dual". Mathematical Programming, vol 63, no. 1. 1994. pp23-46.

[8] Den Hertog, D. PhD Thesis. "Interior Point Approach to Linear, Quadratic and Convex Programming, Algorithms and Complexity". Faculty of Mathematics and Informatics, TU Delft, the Netherlands, Sep 1992.

[9] Den Hertog, D., Kaliski, J., Roos, C. and Terlaky, T. "A Logarithmic Barrier Cutting Plane Method for Convex Programming". Report 93-43, Delft University of Technology, 1993.

[10] Elzinga, J and Moore, T. G. "A Central Cutting Plane Algorithm for the Convex Programming Problem". Mathematical Programming, 8, 134-145, 1975.

[11] Goffin, J. L., Luo, Z. Q. and Ye, Y. "Further Complexity Analysis of a Primal-Dual Column Generation Algorithm for Convex or Quasiconvex Feasibility Problems". Faculty of Management, McGill University, Montreal, Canada. August 1993.

[12] Grotschel, M , Lovász, L and Schrijver, A. "Geometric Algorithms and Combinatorial Optimization". Springer Verlag, 1988.

[13] Horn, R. A. and Johnson, C. R. "Matrix Analysis". Cambridge University Press, 1985.

[14] Mitchell, J. E., and Ramaswamy, S. "A Long Step Cutting Plane Algorithm for Linear and Convex Programming". Technical Report number 37-93-387, Dept. of DSES, R.P.I, Troy, NY 12180. 1993.

[15] Mitchell, J. E. and Todd, M. J. "Solving Combinatorial Optimization Problems Using Karmarkar Algorithm". Mathematical Programming, 1992 (OCT 5) Vol. 56 pp 245 - 284.

[16] Ramaswamy, S. and Mitchell, J. E. "On Updating the Analytic Center After the Addition of Multiple Cuts". DSES Technical Report 37-94-423, Rensselaer Polytechnic Institute, Troy, NY 12180. October 27, 1994.

[17] Renegar, J. "A Polynomial-Time Algorithm, Based on Newton's Method, for Linear Programming". Mathematical Programming, 40, pp59-93, 1988.

[18] Sonnevend, Gy. "An 'Analytic Center' for Polyhedrons and New Classes of Global Algorithms for Linear (Smooth, Convex) Programming". Lecture Notes in Control and Information Sciences, Springer Verlag, New York, 84, 866-876, 1985.

[19] Tarasov, S. P , Khachiyan, L. G. and Erlich, I. I. "The Method of Inscribed Ellipsoids". Soviet Math Dokl., 37, 1988.

[20] Vaidya, P. M. "A New Algorithm for Minimizing Convex Functions over Convex Sets". To appear in Math. Prog.

[21] Vaidya, P. M. "An Algorithm for Linear Programming Which Requires $O(((m+n)n^2+(m+n)^{1.5}n)L)$ Arithmetic Operations". Mathematical Programming, 47, 1990. pp175-201.

[22] Vaidya, P. M. and Atkinson, D. S. "A Technique for Bounding the Number of Iterations in Path Following Algorithms", in *Complexity in Numerical Optimization*, P. M. Pardalos, editor, World Scientific, Singapore, 1993. pp462-489.

[23] Ye, Y. "A Potential Reduction Algorithm Allowing Column Generation". SIAM Journal on Optimization, 2, pp7-20, 1992.

[24] Ye, Y. "Complexity Analysis of the Analytic Center Cutting Plane Method That Uses Multiple Cuts". Dept. of Management Sciences, The University of Iowa, Iowa City, Iowa 52242. September 1994.