Improved Complexity for Maximum Volume Inscribed Ellipsoids

Kurt M. Anstreicher
Dept. of Management Sciences
University of Iowa
Iowa City, IA 52242

June 13, 2001

Abstract

Let $\mathcal{P} = \{x \mid Ax \leq b\}$, where A is an $m \times n$ matrix. We assume that $\mathcal{P}$ contains a ball of radius one centered at the origin, and is contained in a ball of radius R centered at the origin. We consider the problem of approximating the maximum volume ellipsoid inscribed in $\mathcal{P}$. Such ellipsoids have a number of interesting applications, including the inscribed ellipsoid method for convex optimization. We reduce the complexity of finding an ellipsoid whose volume is at least a factor $e^{-\epsilon}$ of the maximum possible to $O(m^{3.5} \ln(mR/\epsilon))$ operations, improving on previous results of Nesterov and Nemirovskii, and Khachiyan and Todd. A further reduction in complexity is obtained by first computing an approximation of the analytic center of $\mathcal{P}$.

Keywords: Maximum volume inscribed ellipsoid, inscribed ellisoid method.

1 Introduction

Let $\mathcal{P} = \{x \mid Ax \leq b\}$, where A is an $m \times n$ matrix. We assume that $\mathcal{P}$ is bounded, with a nonempty interior. It is then known [5] that there is a unique ellipsoid $E^* \subset \mathcal{P}$ of maximum volume. We say that an ellipsoid $E \subset \mathcal{P}$ is γ-*maximal* if $\mathrm{Vol}(E) \geq \gamma\,\mathrm{Vol}(E^*)$, where $0 < \gamma < 1$ and $\mathrm{Vol}(\cdot)$ denotes n-dimensional volume. In this paper we consider the complexity of computing a γ-maximal inscribed ellipsoid for $\mathcal{P}$. For convenience in stating complexity results we often write $\gamma = e^{-\epsilon}$ (as in [10, Section 6.5]), where $\epsilon > 0$.

There are a number of interesting applications of γ-maximal ellipsoids. For example, the computation of a γ-maximal ellipsoid, with $\gamma > 0.92$, is required on each iteration of the inscribed ellipsoid algorithm (IEM) [12] for convex programming. The IEM minimizes a convex function over an n-dimensional cube to relative accuracy ν in $O(n\ln(n/\nu))$ iterations, each requiring evaluation of the function and a subgradient. The order of this complexity, also achieved by the volumetric cutting plane algorithm [1, 13], is optimal [11].

Another application of γ-maximal ellipsoids is to provide a "rounding" of $\mathcal{P}$. It is know that for the maximum volume inscribed ellipsoid (MVIE) E^*,

$$E^* \subset \mathcal{P} \subset nE^*,$$

where for an ellipsoid E and positive scalar τ, τE denotes the dilation of E about its center by the factor τ. In the worst case the rounding factor n cannot be improved. For a γ-maximal ellipsoid E it can be shown [12] that

$$E \subset \mathcal{P} \subset n\left(\frac{1 + 3\sqrt{1-\gamma}}{\gamma}\right) E.$$

Roundings of this type are required in several contexts, including Lenstra's algorithm for integer programming in fixed dimension [9], and randomized algorithms for volume computation [6]. Alternative methodologies for obtaining $O(n)$-roundings of $\mathcal{P}$ include the shallow cut ellipsoid algorithm [3, Section 4.6] and the volumetric cutting plane algorithm [2].

Assume that $\mathcal{P}$ contains a ball of radius one centered at the origin, and is contained in a ball of radius R centered at the origin. Using the ellipsoid algorithm an $e^{-\epsilon}$–maximal inscribed ellipsoid can be computed in $O(n^6(n^2 + m)\ln(nR/\epsilon))$ operations [12]. For reasonable m this complexity was substantially improved to

$$O\left(m^{2.5}(n^2+m)\ln\left(\frac{mR}{\epsilon}\right)\right) \tag{1}$$

operations by Nesterov and Nemirovskii [10], using an interior-point algorithm with a specialized "rescaling" technique to lower the work required on each iteration. A further reduction to

$$O\left(m^{3.5}\ln\left(\frac{mR}{\epsilon}\right)\ln\left(\frac{n\ln R}{\epsilon}\right)\right) \tag{2}$$

operations was achieved by Khachiyan and Todd [8], who apply an interior-point algorithm to a sequence of problems, each of which requires less work per iteration than the original problem considered by [10]. A primal-dual algorithm for computing an approximation of the MVIE is described in [14]. In this paper we show that an $e^{-\epsilon}$–maximal inscribed ellipsoid can be computed in

$$O\left(m^{3.5}\ln\left(\frac{mR}{\epsilon}\right)\right) \tag{3}$$

operations. We also show that by first computing an approximation of the analytic center of $\mathcal{P}$ the effect of the parameter R can be further reduced, resulting in a total complexity of

$$O\left((mn^2+m^{1.5}n)\ln(R)+m^{3.5}\ln\left(\frac{m}{\epsilon}\right)\right) \tag{4}$$

operations. The difference between (3) and (4) is certainly of interest, since under standard assumptions bounds on R may be exponential in n [3, Lemma 3.1.25].

A problem related to that of computing an $e^{-\epsilon}$–maximal inscribed ellipsoid for $\mathcal{P}$ is that of computing an e^{ϵ}–minimal circumscribing ellipsoid for the convex hull of m given points in $\Re^n$. Khachiyan [7] shows that the latter problem can be solved in

$$O\left(m^{3.5}\ln\left(\frac{m}{\epsilon}\right)\right)$$

operations; note that this bound is independent of the parameter R.

Notation: If A and B are symmetric matrices, $A \preceq B$ denotes that $B-A$ is positive semidefinite, and $A \prec B$ denotes that $B-A$ is positive definite. The trace of a matrix A is denoted $\mathrm{tr}(A)$, $A \bullet B = \mathrm{tr}(AB^T)$, and $\|A\|$ denotes the Frobenius norm, $\|A\| = \sqrt{\mathrm{tr}(A^TA)}$. The Kronecker product of matrices A and B is denoted $A \otimes B$. If A is an $m \times n$ matrix, $\mathbf{vec}(A)$ is the vector in $\Re^{mn}$ formed by "stacking" the columns of A atop one another, in the natural order. We use $B(x,r)$ to denote the closed ball of radius r centered at $x \in \Re^n$.

2 Preliminaries

In this section we give definitions and basic results from [10] that will be required in the sequel.

Definition 2.1 *Let G be a closed, convex set in $\Re^N$, and let $f(\cdot) : \mathrm{Int}(G) \to \Re$ be a C^3 convex function. Then $f(\cdot)$ is said to be strongly 1-self-concordant (hereafter abbreviated strongly self-concordant) on $\mathrm{Int}(G)$ if $f(x_k) \to \infty$ for any sequence $\{x_k\}$ converging to a boundary point of G, and*

$$|D^3 f(x)[h,h,h]| \le 2\left(D^2 f(x)[h,h]\right)^{3/2}$$

for every $x \in \mathrm{Int}(G)$ and $h \in \Re^N$.

Assume that $f(\cdot)$ is strongly self-concordant on $\mathrm{Int}(G)$, and that G is bounded. It can then be shown that $\nabla^2 f(x)$ is nonsingular for every $x \in \mathrm{Int}(G)$. For $x \in \mathrm{Int}(G)$ define the *Newton direction* for $f(\cdot)$ at x to be

$$p(x) = [\nabla^2 f(x)]^{-1} \nabla^T f(x),$$

and the *Newton decrement* for $f(\cdot)$ at x to be

$$\lambda(x) = \left(\nabla f(x)[\nabla^2 f(x)]^{-1} \nabla^T f(x)\right)^{1/2}.$$

As shown in [10], for a strongly self-concordant function the Newton decrement $\lambda(x)$ provides good information regarding the difference between $f(x)$ and the minimum of $f(\cdot)$ over G. Note that if G is compact, then $f_{\min} = \min\{f(x) \,|\, x \in G\}$ is attained at an unique interior point of G.

Lemma 2.2 *Let $G \subset \Re^N$ be a compact convex set, and assume that $f(\cdot)$ is strongly self-concordant on $\mathrm{Int}(G)$. Let $x \in \mathrm{Int}(G)$, $\lambda = \lambda(x)$, $p = p(x)$.*

1. *If $\lambda \le 1/3$, then $f(x) - f_{\min} \le \lambda^2/(1 - 5.0625\lambda^2)$.*

2. *If $x^+ = x + [1/(1+\lambda)]p$, then $x^+ \in \mathrm{Int}(G)$, and $f(x^+) \le f(x) - [\lambda - \ln(1+\lambda)]$.*

Proof: Part 1 is proved in [4, Lemma 2.22]; a weaker estimate is given in [10, Theorem 2.2.2]. Part 2 is proved in [10, Proposition 2.2.2] and [4, Lemma 2.24]. □

Definition 2.3 *Let $G \subset \Re^N$ be a compact convex set, and $F(\cdot) : \mathrm{Int}(G) \to \Re$. Then $F(\cdot)$ is called a ϑ-self-concordant barrier for G if $F(\cdot)$ is strongly self-concordant on* $\mathrm{Int}(G)$*, and $\lambda^2(x) \le \vartheta$ for every $x \in \mathrm{Int}(G)$.*

It is very well known that the complexity of linear and quadratic optimization over G is characterized by the parameter ϑ. In particular, if $f(\cdot)$ is a convex quadratic function and $F(\cdot)$ is a computable ϑ-self-concordant barrier for G, then given a suitable initial interior point $x^0 \in G$ and lower bound $z^0 \le f_{\min}$, an interior-point algorithm based on $F(\cdot)$ can be used to obtain x having $f(x) - f_{\min} \le \epsilon[q(x^0) - z^0]$ in $O(\sqrt{\vartheta}|\ln \epsilon|)$ iterations, each requiring a Newton step for a linear combination of $f(\cdot)$ and $F(\cdot)$.

To analyze the complexity of optimizing more general convex functions over G, [10] uses the concept of *β-compatibility* between a convex objective $f(\cdot)$ and a barrier $F(\cdot)$ for G. The details are not important here, but we note that the complexity of approximately minimizing $f(\cdot)$ over G involves the parameters β and ϑ, as well as characteristics of the initial point.

3 The MVIE problem

As described in [10], the problem of computing the MVIE for a polyhedral set $\mathcal{P}$ can be cast as the convex programming problem

$$\begin{array}{ll} \min & -\mathrm{ldet}\, Y \\ \text{s.t.} & \|Ya_i\| \le (b_i - a_i^T y), \quad i = 1, \ldots, m \\ & Y \succeq 0, \end{array} \tag{5}$$

where a_i denotes the ith row of A. A feasible solution to (5) with objective value within ϵ of optimality provides an $e^{-\epsilon}$–maximal inscribed ellipsoid of the form $\{y + Yz \,|\, \|z\| = 1\}$. The complexity analysis in [10] uses the $2(m+n)$-self-concordant barrier

$$-2\,\mathrm{ldet}\, Y - \sum_{i=1}^{m} \ln((b_i - a_i^T y)^2 - a_i^T Y^2 a_i),$$

and the fact that $f(Y) = -\mathrm{ldet}\, Y$ is 1-compatible with this barrier. The main difficulty with this approach is that the resulting Newton equations are relatively expensive to form and solve. In [10] a "speed-up" based on rescaling the matrix Y is used to reduce this complexity to $O(m^2(n^2+m))$ operations per iteration, resulting in the overall complexity (1).

Note that by letting $X = Y^2$, (5) is equivalent to the optimization problem

$$\begin{aligned} \min \quad & -\operatorname{ldet} X \\ \text{s.t.} \quad & a_i^T X a_i \le (b_i - a_i^T x)^2, \quad i = 1, \ldots, m \\ & X \succeq 0, \end{aligned} \tag{6}$$

and a solution of (6) with objective within 2ϵ of optimimality produces an $e^{-\epsilon}$–maximal inscribed ellipsoid. Unfortunately the constraints of (6), while linear in X, are not convex in x. The approach taken in [8] is to approximately solve a sequence of problems of the form

$$P(y): \qquad \begin{aligned} \min \quad & -\operatorname{ldet} X \\ \text{s.t.} \quad & a_i^T X a_i \le (b_i - a_i^T y)(b_i - a_i^T x), \quad i = 1, \ldots, m \\ & X \succeq 0. \end{aligned} \tag{7}$$

The process is initialized using $y_0 = 0$, and if (x_k, X_k) is the approximate solution of $P(y_k)$ then $y_{k+1} = (1/2)(x_k + y_k)$. In [8] the convergence of $\{y_k\}$ is shown to be very rapid. Moreover the barrier

$$-\operatorname{ldet} X - \sum_{i=1}^{m} \ln((b_i - a_i^T y)(b_i - a_i^T x) - a_i^T X a_i) \tag{8}$$

for $P(y)$ is $(m+n)$-self-concordant, $f(X) = -\operatorname{ldet} X$ is $O(1)$-compatible with this barrier, and the Newton direction required on each iteration can be computed in only $O(m^3)$ operations. This reduces the complexity of finding an $e^{-\epsilon}$–maximal inscribed ellipsoid for $\mathcal{P}$ to (2).

The approach we take here uses the family of barriers (8) as in [8], but avoids solving the sequence of problems $P(y_k)$. In this way we reduce the computation required for each Newton step to $O(m^3)$ operations, but avoid the factor $\ln((n \ln R)/\epsilon)$ in (2). We also show that "pre-rounding" $\mathcal{P}$ by first computing an approximation of the analytic center of $\mathcal{P}$ can be used to further reduce the effect of the parameter R, resulting in the complexity (4).

4 Main stage

Let G denote the feasible region of (6), and for $y \in \text{Int}(\mathcal{P})$ let $G(y)$ denote the feasible region of $P(y)$, from (7). For $(x, X) \in \text{Int}(G(y))$ and $t \geq 1$ let

$$F_t(y; x, X) = -t \operatorname{ldet} X - \sum_{i=1}^{m} \ln \left((b_i - a_i^T y)(b_i - a_i^T x) \right) - \sum_{i=1}^{m} \ln \left((b_i - a_i^T y)(b_i - a_i^T x) - a_i^T X a_i \right).$$

It is then straightforward to show that $F_t(y; \cdot, \cdot)$ is strongly self-concordant on $\text{Int}(G(y))$ for any $y \in \text{Int}(P)$. In working with $F_t(y; x, X)$ we consider the components of y to be fixed parameters, while those of (x, X) are variables. Let $[p_t(y; x, X), P_t(y; x, X)]$ denote the Newton direction for $F_t(y; \cdot, \cdot)$ at (x, X), and let $\lambda_t(y; x, X)$ be the corresponding Newton decrement. In this section we describe and analyze the "main stage" of our barrier algorithm for obtaining an $e^{-\epsilon}$–maximal ellipsoid. The main stage is initialized with $t_0 = 1$, and a point (x_0, X_0) such that $x_0 \in \text{Int}(\mathcal{P})$, $(x_0, X_0) \in \text{Int}(G(x_0))$, and $\lambda_1(x_0; x_0, X_0) \leq 0.15$. The problem of obtaining such an initial point is considered in the next section. The main stage algorithm, described in pseudo-code below, is a variant of the standard barrier algorithm for convex optimization analyzed in [10]. The novelty of the algorithm here is that the Newton direction used on each inner iteration is obtained from a barrier function $F_t(x; \cdot, \cdot)$ that depends on the current x.

ALGORITHM (Main stage for MVIE):
Given $k = 0$, x_0, X_0, $t_0 = 1$, $t_{\max}$, $\theta > 0$.
Do Until $t_k \geq t_{\max}$ (Outer iteration)
  $t = t_{k+1} = (1 + \theta) t_k$, $x = x_k$, $X = X_k$.
  Do Until $\lambda_t(x; x, X) \leq 0.15$ (Inner iteration)
    $p = p_t(x; x, X)$, $P = P_t(x; x, X)$
    $x = x + (\alpha/2) p$, $X = X + \alpha P$.
  End
  $x_{k+1} = x, X_{k+1} = X$, $k = k + 1$
End

The steplength α on each inner iteration can be taken to be any value that produces at least the descent in $F_t(\cdot; \cdot, \cdot)$ obtained using $\alpha = 1/(1 + \lambda_t(x; x, X))$; see Lemma 4.3 below.

Our analysis of the main stage algorithm for MVIE is based on the well-known analysis of the barrier algorithm from [10] (see also [4]). The following result facilitates the use of directions based on the family of barrier functions $F_t(y; \cdot, \cdot)$.

Lemma 4.1 *Suppose that x and y are interior points of $\mathcal{P}$, and let $f(X) = -\operatorname{ldet} X$, $G(y,x) = \{X \succeq 0 \,|\, a_i^T X a_i \le (b_i - a_i^T y)(b_i - a_i^T x),\ i = 1,\ldots,m\}$. Define $\phi(y,x) = \min_{X\in G(y,x)} f(X)$, and $\phi_t(y,x) = \min_{X\in G(y,x)} F_t(y;x,X)$, $t \ge 1$. Then*

1. $\phi(y,x) \le \frac{1}{2}[\phi(y,y) + \phi(x,x)]$,
2. $\phi_t(y,x) \le \frac{1}{2}[\phi_t(y,y) + \phi_t(x,x)]$.

Proof: Part 1 is proved in [8, p.144], and we use a similar argument to prove part 2 here. Assume that $X \succ 0$ is the minimizer of $F_t(x;x,\cdot)$, and $Y \succ 0$ is the minimizer of $F_t(y;y,\cdot)$. By a change of coordinates we may assume without loss of generality that X and Y are diagonal. For each $i = 1,\ldots,m$ the inequalities $a_i^T X a_i \le (b_i - a_i^T x)^2$ and $a_i^T Y a_i \le (b_i - a_i^T y)^2$ together imply that

$$a_i^T (XY)^{1/2} a_i \le [(a_i^T X a_i)(a_i^T Y a_i)]^{1/2} \le (b_i - a_i^T x)(b_i - a_i^T y), \tag{9}$$

and $\operatorname{ldet}(XY)^{1/2} = (1/2)(\operatorname{ldet} X + \operatorname{ldet} Y)$. It is also easy to see that for any x and y in $\mathcal{P}$, and $i = 1,\ldots,m$,

$$(b_i - a_i^T x)(b_i - a_i^T y) \le \left(b_i - a_i^T \frac{(x+y)}{2}\right)^2. \tag{10}$$

To prove that $\phi_t(y,x) \le \frac{1}{2}[\phi_t(y,y) + \phi_t(x,x)]$ it then suffices to show that for each $i = 1,\ldots,m$,

$$\begin{aligned} & -\ln\left((b_i - a_i^T x)(b_i - a_i^T y) - a_i^T (XY)^{1/2} a_i\right) \\ \le \; & -\frac{1}{2}\ln\left((b_i - a_i^T x)^2 - a_i^T X a_i\right) - \frac{1}{2}\ln\left((b_i - a_i^T y)^2 - a_i^T Y a_i\right), \end{aligned}$$

which is equivalent to

$$\left((b_i - a_i^T x)(b_i - a_i^T y) - a_i^T (XY)^{1/2} a_i\right)^2 \ge \left((b_i - a_i^T x)^2 - a_i^T X a_i\right)\left((b_i - a_i^T y)^2 - a_i^T Y a_i\right). \tag{11}$$

Using the left-hand inequality in (9), to prove (11) it suffices to show that

$$\left((b_i - a_i^T x)(b_i - a_i^T y) - [(a_i^T X a_i)(a_i^T Y a_i)]^{1/2}\right)^2 \ge \left((b_i - a_i^T x)^2 - a_i^T X a_i\right)\left((b_i - a_i^T y)^2 - a_i^T Y a_i\right),$$

which easily reduces to

$$\left((b_i - a_i^T x)(a_i^T Y a_i)^{1/2} - (b_i - a_i^T y)(a_i^T X a_i)^{1/2}\right)^2 \ge 0.$$

□

We now use Lemma 4.1, and standard results on self-concordant functions, to bound the possible reduction in $F_t(\cdot;\cdot,\cdot)$ and $f(\cdot)$ when the Newton decrement is sufficiently small.

Lemma 4.2 *Suppose that* $t > 1$, $x \in \mathrm{Int}(\mathcal{P})$, $(x, X) \in \mathrm{Int}(G(x))$, *and* $\lambda_t(x; x, X) \leq \lambda \leq 1/3$. *Let* $\delta = \delta(\lambda) = \lambda^2/(1 - 5.0625\lambda^2)$. *Then*

1. $F_t(y; y, Y) \geq F_t(x; x, X) - 2\delta$ *for all* $(y, Y) \in G$.

2. $f(X) \leq f_{\min} + [12m + 2\delta]/(t - 1)$.

Proof: From part 1 of Lemma 2.2 we have

$$F_t(x; y, Y) \geq F_t(x; x, X) - \delta \tag{12}$$

for any $(y, Y) \in G(x)$. Part 2 of Lemma 4.1 then implies that for any $y \in \mathrm{Int}(\mathcal{P})$,

$$\begin{aligned} \phi_t(y, y) &\geq 2\phi_t(y, x) - \phi_t(x, x) \\ &\geq 2(F_t(x; x, X) - \delta) - F_t(x; x, X) \\ &= F_t(x; x, X) - 2\delta, \end{aligned}$$

which proves part 1. Next, note that $F_t(x; x, X) = (t-1)f(X) + F_1(x; x, X)$, where $F_1(x; \cdot, \cdot)$ is a $(2m + n)$-self-concordant barrier for $G(x)$. From (12) and a standard argument (see for example [10, p.75]) we conclude that for all $(y, Y) \in G(x)$,

$$f(Y) \geq f(X) - \frac{2(2m + n) + \delta}{t - 1},$$

and therefore $\phi(x, y) \geq f(X) - (6m + \delta)/(t - 1)$ for all $y \in \mathrm{Int}(\mathcal{P})$, since $m > n$. Part 1 of Lemma 4.1 then implies

$$\begin{aligned} \phi(y, y) &\geq 2\phi(y, x) - \phi(x, x) \\ &\geq 2\left(f(X) - \frac{6m + \delta}{t - 1}\right) - f(X) \\ &= f(X) - \frac{12m + 2\delta}{t - 1}, \end{aligned}$$

which proves part 2. □

The next lemma considers the effect of increasing t on the Newton decrement for $F_t(x; x, X)$, and the reduction in $F_t(\cdot; \cdot, \cdot)$ that can be assured if the Newton decrement is not sufficiently small.

Lemma 4.3 *For $t \geq 1$ let $\lambda_t(y; x, X)$ be the Newton decrement for $F_t(y; \cdot, \cdot)$ at (x, X).*

1. *Suppose that $\lambda_t(x; x, X) \leq .15$, and let $t^+ = (1 + \theta)t$, $\theta \geq 0$. Then $\lambda_{t^+}(x; x, X) \leq .15(1 + \theta) + \sqrt{2m}\theta$.*

2. *Suppose that $\lambda_t(x; x, X) = \lambda > .15$. Let $\alpha = 1/(1 + \lambda)$, $x^+ = x + (\alpha/2)p_t(x; x, X)$, $X^+ = X + \alpha X_t(x; x, X)$. Then $(x^+, X^+) \in \text{Int}(G)$, and $F_t(x^+; x^+, X^+) \leq F_t(x; x, X) - 0.01$.*

Proof: Part 1 is proved in [4, Lemma 2.25]. Since $F_t(x; \cdot, \cdot)$ is strongly-self-concordant for $t \geq 1$, part 2 of Lemma 2.2 implies that if $x^{++} = \alpha p_t(x; x, X)$, then $(x^{++}, X^+) \in \text{Int}(G(x))$, and

$$F_t(x; x^{++}, X^+) \leq F_t(x; x, X) - (\lambda - \ln(1 + \lambda)) \leq F_t(x; x, X) - 0.01. \tag{13}$$

However (10) implies that if $(y, Y) \in G(x)$, then

$$F_t\left(\frac{x+y}{2}; \frac{x+y}{2}, Y\right) \leq F_t(x; y, Y). \tag{14}$$

The proof of part 2 is completed by combining (13) and (14), with $(y, Y) = (x^{++}, X^+)$. $\square$

We can now combine Lemmas 4.2 and 4.3 to obtain the final complexity result for the main stage.

Theorem 4.4 *Let $\theta = 0.07/\sqrt{m}$. Then the main stage algorithm requires $O(1)$ inner iterations per outer iteration. Moreover, for $t_{\max} = O(m/\epsilon)$ the algorithm terminates with an $e^{-\epsilon}$–maximal ellipsoid in $O(m^{.5}\ln(m/\epsilon))$ outer iterations, requiring a total of $O(m^{3.5}\ln(m/\epsilon))$ operations.*

Proof: From part 1 of Lemma 4.3, at the start of each sequence of inner iterations we have

$$\lambda_{t_{k+1}}(x_k; x_k, X_k) \leq .15\left(1 + \frac{.07}{\sqrt{m}}\right) + .1 < 0.26.$$

Part 1 of Lemma 4.2 then implies that for any $(x, X) \in G$,

$$F_{t_{k+1}}(x; x, X) > F_{t_{k+1}}(x_k; x_k, X_k) - .21,$$

and from Part 2 of Lemma 4.3 there can be at most 20 inner iterations on each outer iteration.

From part 2 of Lemma 4.2, to obtain an $e^{-\epsilon}$–maximal inscribed ellipsoid it suffices to terminate the algorithm using $t_{\max} = O(m/\epsilon)$, which requires $O(m^{.5}\ln(m/\epsilon))$ outer iterations using $\theta =$

$.07/\sqrt{m}$. Finally, it can be shown using a small modification of the argument used in [8, Section 6] that each inner iteration can be executed in $O(m^3)$ operations. $\square$

It is worth noting that the analysis of this section does not require the term

$$-\sum_{i=1}^{m} \ln\left((b_i - a_i^T y)(b_i - a_i^T x)\right), \tag{15}$$

in the definition of $F_t(y; x, X)$, and in fact this term slightly degrades the theoretical performance of the main stage algorithm. However (15) is very helpful for the analysis of the preliminary stage, in the next section.

5 Preliminary stage

In this section we consider the preliminary stage of our barrier algorithm for the MVIE problem. The goal of the preliminary stage is to produce the initial point (x_0, X_0) required by the main stage algorithm of the previous section. Our preliminary stage is based on the general preliminary stage described in [10, Section 3.2.3], except that we work with directions based on $F_1(y; x, X)$ so as to keep the work per iteration $O(m^3)$.

The preliminary stage is initialized at $x_0 = 0$, and a suitable X_0 to be described below. Let

$$c = -\nabla_x F_1(0; 0, X_0)^T = -\sum_{i=1}^{m} \left(\frac{b_i}{\Delta_i^0} + \frac{1}{b_i}\right) a_i, \tag{16a}$$

$$C = -\nabla_X F_1(0; 0, X_0) = X_0^{-1} - \sum_{i=1}^{m} \frac{1}{\Delta_{i,0}} a_i a_i^T, \tag{16b}$$

where $\Delta_{i,0} = b_i^2 - a_i^T X_0 a_i$. For $t \leq 1$ define the preliminary stage barrier function

$$F_t^0(y; x, X) = t(c^T(x + y) + C \bullet X) + F_1(y; x, X).$$

Let $[p_t^0(y; x, X), P_t^0(y; x, X)]$ denote the Newton direction for $F_t^0(y; \cdot, \cdot)$ at (x, X), and $\lambda_t^0(y; x, X)$ the corresponding Newton decrement. Note that by construction $\lambda_1^0(0; 0, X_0) = 0$. The preliminary stage algorithm, given below, is very similar to the main stage, except that the preliminary stage uses $F_t^0(\cdot; \cdot, \cdot)$ and t is decreased rather than increased on each outer iteration.

ALGORITHM (Preliminary stage for MVIE):
Given $k = 0$, $x_0 = 0$, X_0, $t_0 = 1$, $t_{\min}$, $\theta > 0$.

Do Until $t_k \le t_{\min}$ (Outer iteration)
$t = t_{k+1} = (1-\theta)t_k,\ x = x_k,\ X = X_k.$
Do Until $\lambda_t^0(x; x, X) \le 0.15$ (Inner iteration)
$p = p_t^0(x; x, X),\ P = P_t^0(x; x, X)$
$x = x + (\alpha/2)p,\ X = X + \alpha P.$
End
$x_{k+1} = x, X_{k+1} = X,\ k = k+1$
End

To analyze the preliminary stage we require an extension of part 2 of Lemma 4.1 that applies to $F_t^0(y; x, X)$. This turns out to be straightforward under the assumption that $C \succeq 0$.

Lemma 5.1 *For interior points x and y of $\mathcal{P}$, let $\phi_t^0(y, x) = \min_{X \in G(y,x)} F_t^0(y; x, X)$, $0 < t \le 1$. Assume that $C \succeq 0$. Then $\phi_t^0(y, x) \le \frac{1}{2}[\phi_t^0(y, y) + \phi_t^0(x, x)]$.*

Proof: The proof is identical to the proof of part 2 of 4.1. Note that the change of coordinates that simultaneously diagonalizes X and Y preserves the semidefiniteness of C. After this change of coordinates we have

$$C \bullet X = \sum_{i=1}^{m} C_{ii} X_{ii}, \quad C \bullet Y = \sum_{i=1}^{m} C_{ii} Y_{ii}, \quad C \bullet (XY)^{1/2} = \sum_{i=1}^{m} C_{ii} \sqrt{X_{ii} Y_{ii}},$$

where $C_{ii} \ge 0$, $i = 1, \ldots, m$. It immediately follows that $C \bullet (XY)^{1/2} \le (1/2)(C \bullet X + C \bullet Y)$.
□

For $C \succeq 0$ and a given value of $t_{\min}$, the analysis of the preliminary stage is very similar to the analysis of the main stage given in the previous section, and is omitted here. The final complexity result has the following form.

Theorem 5.2 *Assume that $C \succeq 0$. Then for $\theta = \eta/\sqrt{m}$, where $\eta > 0$ is an appropriately chosen positive constant, the preliminary stage requires $O(m^{.5} \ln(1/t_{\min}))$ outer iterations, $O(1)$ inner iterations per outer iteration, and a total of $O(m^{3.5} \ln(1/t_{\min}))$ operations.*

To complete the analysis of the preliminary stage we must show that X_0 can be chosen so that $C \succeq 0$, and characterize the value $t_{\min}$ so that termination of the preliminary stage produces a suitable initial point for the main stage.

Lemma 5.3 *Assume that $B(0,1) \subset \mathcal{P}$, and let $X_0 = \frac{1}{m+1} I$. Then $C \succeq 0$.*

Proof: By the assumption that $B(0,1) \in \mathcal{P}$ we must have $b_i \geq \|a_i\|$, $i = 1, \ldots, m$. Then for each i,

$$\Delta_{i,0} = b_i^2 - a_i^T X_0 a_i = b_i^2 - \frac{1}{m+1}\|a_i\|^2 \geq b_i^2 \frac{m}{m+1} \ . \tag{17}$$

From (16b) we then have

$$C \succeq (m+1)I - \frac{m+1}{m}\sum_{i=1}^{m} \frac{1}{b_i^2} a_i a_i^T \succeq (m+1)\left(I - \frac{1}{m}\sum_{i=1}^{m}\frac{1}{\|a_i\|^2} a_i a_i^T\right) \succeq 0.$$

□

It remains to obtain a lower bound on the required value of $t_{\min}$. To this end, note that for any strongly self-concordant function $F(\cdot)$ defined on the interior of a compact, convex set $G \in \Re^N$, the Newton decrement $\lambda(x)$ at $x \in \mathrm{Int}(G)$ is equal to the solution value in the optimization problem

$$\begin{array}{ll} \max & DF(x)[h] \\ \text{s.t.} & D^2F(x)[h,h] \leq 1. \end{array}$$

It follows from this characterization and the fact that $F_t^0(y; x, X) = t(c^T x + c^T y + C \bullet X) + F_1(y; x, X)$ that for any $(x, X) \in \mathrm{Int}(G)$,

$$\lambda_1(x; x, X) \leq \lambda_t^0(x; x, X) + t v(x, X), \tag{18}$$

where

$$\begin{array}{rll} v(x,X) = & \max & c^T h + C \bullet H \\ & \text{s.t.} & D^2 F_1(x; x, X)[(h,H),(h,H)] \leq 1. \end{array} \tag{19}$$

Thus to obtain a lower bound on the required value $t_{\min}$ we require an upper bound for $v(\cdot,\cdot)$.

Lemma 5.4 *Assume that $B(0,1) \subset \mathcal{P} \subset B(0,R)$, and let $X_0 = \frac{1}{m+1} I$. Then at any $(x, X) \in \mathrm{Int}(G)$, $v(x,X) \leq (4+\sqrt{n})(m+1)R^2$.*

Proof: It is straightforward to compute (see for example [8]) that

$$\begin{aligned} D^2F_1(x; x, X)[(h,H),(h,H)] \;=\; & \mathbf{vec}(H)^T (X^{-1} \otimes X^{-1}) \mathbf{vec}(H) + \sum_{i=1}^{m} \frac{1}{s_i^2}(a_i^T h)^2 \\ & + \sum_{i=1}^{m} \frac{1}{\Delta_i^2}(a_i^T H a_i + s_i a_i^T h)^2, \end{aligned} \tag{20}$$

where $s_i = b_i - a_i^T x$, and $\Delta_i = (b_i - a_i^T x)^2 - a_i^T X a_i$. From the assumption that $\mathcal{P} \in B(0, R)$ and the relationship between (5) and (6), we must have $X \preceq R^2 I$, so $X^{-1} \succeq (1/R^2)I$ and $X^{-1} \otimes X^{-1} \succeq (1/R^4)I$. Therefore

$$\mathbf{vec}(H)^T (X^{-1} \otimes X^{-1})\, \mathbf{vec}(H) \geq (1/R^4)\|H\|^2,$$

for any H. From (20), if (h, H) is feasible in (19) we clearly have

$$C \bullet H \leq \|C\| \|H\| \leq (m+1)\sqrt{n} R^2, \tag{21}$$

since $0 \preceq C \preceq X_0^{-1}$. In addition, the fact that $\mathcal{P} \in B(0, R)$ implies that for any $h \neq 0$ there is an index $i = i(h)$ so that $a_i^T h / \|h\| \geq b_i / R$, implying that $b_i \leq R\|a_i\|$, and $a_i^T h \geq b_i \|h\|/R \geq \|a_i\|\,\|h\|/R$. For this same index it must also be that $s_i = b_i - a_i^T x \leq b_i + R\|a_i\| \leq 2R\|a_i\|$. Combining these facts we conclude that for any h,

$$\sum_{i=1}^m \frac{1}{s_i^2}(a_i^T h)^2 \geq \left(\frac{1}{2R\|a_{i(h)}\|}\right)^2 \left(\frac{\|a_{i(h)}\|\,\|h\|}{R}\right)^2 = \frac{\|h\|^2}{4R^4}.$$

From (17) we have

$$\frac{b_i}{\Delta_{i,0}} + \frac{1}{b_i} \leq \frac{2m+1}{mb_i} < \frac{2(m+1)}{m\|a_i\|}$$

for each i, so if (h, H) is feasible in (19),

$$c^T h \leq \|c\|\,\|h\| \leq 4(m+1)R^2. \tag{22}$$

The proof is completed by combining (21) and (22). □

Theorem 5.5 *Assume that $B(0,1) \subset \mathcal{P} \subset B(0,R)$, and let $X_0 = \frac{1}{m+1}I$. Then for $1/t_{\min} = O(\sqrt{n} m R^2)$ the preliminary stage terminates with an (x, X) having $\lambda_1(x; x, X) \leq .26$, using a total of $O(m^{3.5}\ln(mR))$ operations.*

Proof: This follows immediately from Theorem 5.2, (18), and Lemma 5.4. □

Given (x, X) with $\lambda_1(x; x, X) \leq .26$, using at most 20 inner iterations of the main stage algorithm (see the proof of Theorem 4.4) we can obtain (x, X) having $\lambda_1(x; x, X) \leq .15$, as required to initialize the main stage.

Note that the second term on the right-hand side of (20) arises from the presence of (15) in the definition of $F_t(\cdot;\cdot,\cdot)$, and this term is responsible for the bound (22) used in the proof of Theorem 5.2. Without (15) we would be forced to rely on the third term of (20) to bound $c^T h$. Such an analysis may be possible, but appears to require that s_i be bounded away from zero for each i. It is interesting to note that a similar issue appears in the analysis of [8]; see the proof of [8, Theorem 3].

Combining Theorems 4.4 and 5.5 we immediately obtain the overall complexity bound (3). Note that the parameter R only appears in the complexity bound for the preliminary stage. We next show that the effect of R can be reduced by first computing an approximation of the ordinary analytic center of $\mathcal{P}$.

The analytic center of $\mathcal{P}$ is the minimizer of the logarithmic barrier function

$$F(x) = -\sum_{i=1}^{m} \ln(b_i - a_i^T x). \tag{23}$$

Let $\lambda(x)$ be the Newton decrement for $F(\cdot)$ at x. It is then well known [10, Section 3.2.3] that under the assumption that $B(0,1) \in \mathcal{P} \in B(0,R)$, a point x with $\lambda(x) < .2$ can be computed using $O(m^{.5}\ln(mR))$ damped Newton steps. Moreover, it is straightforward to show (for example using a small modification of the proof of [2, Lemma 3.1]) that for such an x,

$$E(x, \nabla^2 F(x), 1) \subset \mathcal{P} \subset E(x, \nabla^2 F(x), 1.25m),$$

where for a positive definite matrix H, $E(x,H,r) = \{y \mid (y-x)^T H (y-x) \le r^2\}$. Using a change of coordinates (that scales volume by a constant factor) we can then move x to the origin, and obtain $R = 1.25m$. Finally, by using "partial updating" of the Newton equations required on each iteration, the total complexity of obtaining the approximate analytic center x is

$$O((mn^2 + m^{1.5}n)\ln(mR))$$

operations; this complexity is comprised of a total of $O(m\ln(mR))$ updating steps, each requiring $O(n^2)$ work, and $O(mn)$ other operations per iteration (see [10, Chapter 8] or [4, Chapter 4]). It follows that by first computing an approximation of the analytic center of $\mathcal{P}$ the overall complexity of computing an $e^{-\epsilon}$–maximal ellipsoid is reduced to (4).

References

[1] K.M. Anstreicher, "On Vaidya's volumetric cutting plane method for convex programming," *Math. Oper. Res.* 22 (1997) 63-89.

[2] K.M. Anstreicher, "Ellipsoidal approximations of convex sets based on the volumetric barrier." *Math. Oper. Res.* 24 (1999) 193-203.

[3] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Springer–Verlag, Berlin, 1988.

[4] D. den Hertog, *Interior Point Approach to Linear, Quadratic, and Convex Programming - Algorithms and Complexity*, Kluwer, Dordrecht, 1994.

[5] F. John, "Extremum problems with inequalities as subsidiary conditions," in *Studies and Essays, Presented to R. Courant on his 60th Birthday, January 8, 1948*, Wiley Interscience, New York, 187-204.

[6] R. Kannan, L. Lovász, and M. Simonovits, "Random walks and an $O^*(n^5)$ volume algorithm for convex bodies," *Random Structures and Algorithms* 11 (1997) 1-50.

[7] L.G. Khachiyan, "Rounding of polytopes in the real number model of computation," *Math. Oper. Res.* 21 (1996) 307-320.

[8] L.G. Khachiyan and M.J. Todd, "On the complexity of approximating the maximal volume inscribed ellipsoid for a polytope," *Math. Programming* 61 (1993) 137-159.

[9] H.W. Lenstra, Jr., "Integer programming with a fixed number of variables," *Math. Oper. Res.* 8 (1983) 538-548.

[10] N. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, Philadelphia, 1994.

[11] A.S. Nemirovskii and D.B. Yudin, *Problem complexity and method efficiency in optimization,* John Wiley, Chichester, 1983.

[12] S.P. Tarasov, L.G. Khachiyan, and I.I. Erlich, "The method of inscribed ellipsoids," *Soviet Math. Doklady* 27 (1988) 226-230.

[13] P.M. Vaidya, "A new algorithm for minimizing convex functions over convex sets," *Math. Programming* 73 (1996) 291-341.

[14] Y. Zhang, "An interior-point algorithm for the maximum-volume ellipsoid problem," Technical Report TR98-15, Center for Computational and Applied Math, Rice University, Houston, 1998.