An analytic center cutting plane approach for conic programming

Vasile L Basescu
email: vbasescu@comcast.net

John E Mitchell
Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180 USA
email: mitchj@rpi.edu `http://www.rpi.edu/~mitchj`

We analyze the problem of finding a point strictly interior to a bounded, fully dimensional set from a finite dimensional Hilbert space. We generalize the results obtained for the LP, SDP and SOCP cases. The cuts added by our algorithm are central and conic. In our analysis, we find an upper bound for the number of Newton steps required to compute an approximate analytic center. Also, we provide an upper bound for the total number of cuts added to solve the problem. This bound depends on the quality of the cuts, the dimensionality of the problem and the thickness of the set we are considering.

Key words: cutting plane, cutting surface, analytic center, conic programming, feasibility problem

MSC2000 Subject Classification: Primary: 90C51, 90C25; Secondary: 90C05, 90C22

OR/MS subject classification: Primary: theory and algorithms of nonlinear programming; Secondary: convexity

---

1. Introduction In this paper we will analyze the feasibility problem:

"Given an m-dimensional Hilbert space $(Y, \langle \cdot, \cdot \rangle_Y)$, find a point y in the convex bounded set $\Gamma \subset Y$."

Feasibility problems can be as hard to solve as optimization problems. In fact, once we have an algorithm for solving the feasibility problem, we can use it for solving optimization problems by using binary search.

Because the set Γ is convex, the problem we analyze is of interest in the larger context of non-differentiable convex optimization.

The first assumption made in any feasibility problem is that at least part of the domain Γ is strictly included in a larger set Ω_0. This larger set can be described using a set of so called "box-constraints". These "box-constraints" have different forms, depending on the nature of the Hilbert space $(Y, \langle \cdot, \cdot \rangle_Y)$. In the most general setting, the set Ω_0 is given by

$$\Omega_0 := \{y \in Y : c_1 \leq y \leq c_2\}.$$

The inequality sign "$\leq$" used in describing Ω_0 is a partial order defined on Y. This partial order generates a cone of "positive" vectors K (hence the name of conic programming),

$$K := \{x \in Y : x \geq 0\}.$$

Note here that $u \geq v \Leftrightarrow u - v \geq 0$. This partial order is what distinguishes different classes of feasibility problems.

The most basic class of such problems is linear programming. Linear programming deals with problems that have a linear objective and linear constraints. One of the multiple equivalent forms a linear programming problem can have is:

$$\begin{array}{rl} \max & b^T y, \\ \text{subject to} & A^T y \leq c. \end{array}$$

In this setting, the inequality between two vectors is to be understood componentwise,

$$u \geq v \text{ iff } u_i \geq v_i \text{ for all } i.$$

This vector inequality "$\geq$" introduces a partial ordering on the vector space $\mathbb{R}^n$. The first orthant is the corresponding cone of positive vectors.

More general than linear programming is second order cone programming. The partial order involved in this case is given by

$$u \geq 0,\, u \in I\!R^n \;\Leftrightarrow u_1 \geq \sqrt{\sum_{i=2}^{n} u_i^2}.$$

The induced cone is called the second order cone or the Lorentz cone or the ice-cream cone. Linear programming can be considered a special case for second order cone programming. To see this it is enough to observe that if $n = 1$ the second order cone is $I\!R_+$. Then the first orthant $I\!R_+^n$ can be represented as a cartesian product of n lines $I\!R_+$ or of n one dimensional second order cones.

Even more general is semidefinite programming. In this case the cone K is the cone of positive semidefinite matrices $\mathcal{S}_n$. The partial order, denoted $\succeq$ is given by

$$A \succeq B \Leftrightarrow A - B \in \mathcal{S}_n.$$

To see that second order cone programming is a subcase of semidefinite programming it is enough to notice that the second order cone can be embedded in the cone of positive semidefinite matrices because

$$u_n \geq \sqrt{\sum_{i=1}^{n-1} u_i^2} \Leftrightarrow \begin{pmatrix} u_n I & v \\ v^T & u_n \end{pmatrix} \succeq 0,$$

where v is a $n-1$ - dimensional vector with $v_i = u_i$ for $i = 1, \ldots, n-1$.

All these cases are part of the conic programming family of problems. In this general case, the cone considered is a so called self-scaled cone (it will be defined later). The second order cone, the cone of positive semidefinite matrices and their cartesian products are examples of such cones.

This is the general context in which we intend to analyze the feasibility problem. We assume that this problem has a solution. One way of insuring that is to require that Γ contains a small ball of radius ε. This assumption insures that the set is not too flat. Any point from the interior of Γ is called a feasible point.

Next we will describe the main idea of our approach (most of the terms encountered here will be defined later on, in the second section).

Going back to our problem we assume that

$$\Omega_0 := \{y \in Y : -\tilde{c}_0 \leq y \leq \tilde{c}_0 \text{ with } \tilde{c}_0 \in int(\tilde{K}_0)\}$$

Here $\tilde{K}_0$ is a full-dimensional self-scaled cone in the Hilbert space $(\tilde{X}_0, \langle\cdot,\cdot\rangle_0)$ with $\dim(\tilde{X}_0) = m$. We assume the existence of an oracle which, given a point $\hat{y}$ either recognizes that the point is in Γ or returns a p-dimensional Hilbert space $(X, \langle\cdot,\cdot\rangle_X)$ together with an injective linear operator $A : X \to Y$ such that:

$$\Gamma \subseteq \{y \in Y : A^*(\hat{y} - y) \in K\}.$$

Here K is a full-dimensional self-scaled cone in the Hilbert space $(X, \langle\cdot,\cdot\rangle_X)$. We will say that the operator A defines p central cuts.

In solving the problem we will generate a sequence of closed, bounded sets Ω_i such that $\Gamma \subseteq \Omega_i \subset \Omega_{i-1}$ for any $i \geq 1$. Each set Ω_i is obtained from the previous set Ω_{i-1} by introducing p_i central cuts through a special point $\hat{y}_{i-1} \in \Omega_{i-1}$:

$$\Omega_i := \Omega_{i-1} \cap \{y \in Y : A_i^*(\hat{y}_{i-1} - y) \in K_i\}. \tag{1}$$

The operator $A_i : (X_i, \langle\cdot,\cdot\rangle_i) \to Y$ is injective and linear, X_i is a p_i-dimensional Hilbert space and K_i is a full-dimensional self-scaled cone in X_i.

The special chosen points $\hat{y}_i$ are θ - analytic centers of the corresponding domains Ω_i with respect to an intrinsically self-conjugate functional $f_i : \bigoplus_{j=0}^{i} K_j \to I\!R$.

We will prove that if the total number of cuts added is big enough then the θ - analytic center of the last generated set Ω_i is guaranteed to be in Γ. We will get an estimate on the number of cuts that are added in order to solve the problem. Also we will study the complexity of obtaining one θ - analytic center $\hat{y}_i$ from the previous one $\hat{y}_{i-1}$.

We will prove that the algorithm will stop with a solution after no more than $O^*(\frac{mP^3\Theta^3}{\varepsilon^2\Lambda^2})$ (O^* means that terms of low order are ignored) cuts are added. Here P is the maximum number of cuts added at any of the iterations, Θ is a parameter characterizing the self-concordant functionals and Λ is the minimum eigenvalue of all $A_i^*A_i$ (A_i is the injective operator describing the cuts added at step i). The complexity result we obtain is comparable with the results obtained for less general cases.

Our presentation starts by introducing some previous results corresponding to some particular cases (LP, SDP, SOCP). We will start our analysis by setting up the theoretical background. We will present the notion of self-concordant functional and some related results in *Section 3*. The entire analysis is based on using local inner-products. This notion together with some properties are introduced in *Section 4*. After setting up the theoretical structure we will define in *Section 5* the notion of analytic center. Because computationally it is impossible to work with exact analytic centers, the notion of an approximate analytic center will be introduced. We will analyze then some its properties. In *Section 6* we will introduce more carefully all the assumptions we make about the problem.

After describing the algorithm in *Section 7*, we will analyze in *Section 8* how feasibility can be recovered after the cuts are added centrally, right through the analytic center. In order to keep track of changes in the potentials (another name for the self-concordant functionals used to define the analytic centers) some scaled recovery steps need to be taken.

Section 9 is dedicated to analysis of potentials. The main result will characterize how the potentials at two consecutive analytic centers are related. Using a primal-dual potential and two types of Newton steps we will prove in *Section 10* that a new analytic center can be easily recovered

We will derive an upper bound for the potentials evaluated at the corresponding analytic centers in *Section 11*. This upper bound will be the one that will be used to prove that the algorithm eventually stops with a solution. As expected, this bound depends on the radius ε of the ball we assumed that Γ contains, on the characteristics of the potentials introduced and also on the condition number of the operators describing the cuts.

Finally, in *Section 12* we prove that the algorithm will arrive at a solution in a certain number of steps. We will use the approach employed by Ye in [17] in deriving the bound for the total number of constraints that can be introduced before the algorithm stops with a solution.

2. Previous Work The notion of analytic center was introduced for the first time by Sonnevend in [15] in the context of LP feasibility problems. Atkinson and Vaidya are the ones to introduce for the first time in [1] a complete analysis of a cutting plane method using analytic centers. In their approach the cuts are introduced one by one and "short-steps" are used. Dropping cuts is also allowed. The set Γ is included in a cube of side 2^{L+1} and contains a ball of radius 2^{-L}. The complexity obtained is $O(mL^2)$ iterations. Mitchell and Ramaswamy extended this result in [6] to "long-steps". The complexity was the same but the "long-steps" method is more promising from the computational point of view.

The first analysis of the complexity of the analytic center cutting plane method with multiple cuts was done by Ye in [17]. He proved that by adding multiple cuts, the solution to the feasibility problem can be obtained in no more than $O^*(\frac{m^2P^2}{\varepsilon^2})$ iterations. The same complexity was obtained by Goffin and Vial in [2]. They proved that the recovery of a new analytic center can be done in $O(p\ln(p+1))$ damped Newton steps. This number of steps is the same regardless of the scaling matrix that is used (primal, dual or primal-dual). In our approach we will use a primal-dual approach.

The SOCP case is treated by Oskoorouchi and Goffin in [11]. They analyze the case when one SOCP cut is added at each call of the oracle. They prove that the analytic center of the new domain can be recovered in one Newton step and the total number of analytic centers generated before getting a feasible point is fully polynomial. This was generalized to multiple SOCP cuts by Oskoorouchi and Mitchell [12].

The semidefinite programming case is treated by Toh et. al. [16]. They consider the case of adding multiple central linear cuts. In this case the cuts are added centrally through the analytic center $\hat{Y}$. The form of these cuts is given by $\{Y \in \mathcal{S}_+^m : A_i \bullet Y \leq A_i \bullet \hat{Y}, i = 1, \ldots, p\}$. If P is the maximum of all p, the complexity they obtain is $O(\frac{m^3P}{\varepsilon^2})$. Oskoorouchi and Goffin proved in [10] that when cuts corresponding to SDP cones are added centrally, the analytic center can be recovered in $O(p\ln(p+1))$ damped Newton steps and the total number of steps required to obtain the solution is $O(\frac{m^3P^2}{\varepsilon^2})$.

Peton and Vial extend the analytic center cutting plane method to the general case of convex programming. In [13] they study the introduction of multiple central cuts in a conic formulation of the analytic center cutting plane method. They prove that the new analytic center can be recovered in $O(p \ln wp)$ damped Newton iterations, where w is a parameter depending of the data.

A general survey of non-differentiable optimization problems and methods with a special focus on the analytic center cutting plane method is presented by Goffin and Vial in [3]. This paper presents also the case of multiple cuts and the case of deep cuts.

The analytic center cutting plane class of methods is a member of the larger class of interior point cutting plane methods. Mitchell in [5] gives an overview of polynomial interior point cutting plane methods, including methods based on the volumetric center.

3. Preliminaries on self-concordant functionals Self-concordant functionals are of the utmost importance for optimization theory. In this section we will define this notion and will give some results regarding them that are relevant for our analysis. Most of the definitions/theorems presented in this section are taken from or inspired by [14] and [9].

Let $(X, \langle \cdot, \cdot \rangle_X)$ be a finite dimensional Hilbert space and let $f : X \to I\!R$ be a strictly convex functional with the following properties: D_f, the domain of f is open and convex, $f \in C^2$ and its hessian $H(x)$ is positive definite for all $x \in D_f$. Using the functional f we introduce for each $x \in D_f$ the local (intrinsic) inner product (at x):

$$\langle u, v \rangle_x := \langle u, v \rangle_{H(x)} = \langle u, H(x)v \rangle_X$$

with the corresponding induced norm:

$$\|u\|_x^2 = \langle u, u \rangle_x = \langle u, H(x)u \rangle_X.$$

More generally, for any positive definite operator S we can define a new inner product given by

$$\langle u, v \rangle_S = \langle u, Sv \rangle. \tag{2}$$

Let $B_x(y, r)$ be the open ball of radius r centered at y given by:

$$B_x(y, r) = \{z : \|z - y\|_x \leq r\}. \tag{3}$$

Definition 3.1 *A functional f is said to be* (strongly nondegenerate) self-concordant *if for all $x \in D_f$ we have $B_x(x, 1) \subseteq D_f$, and if whenever $y \in B_x(x, 1)$ we have:*

$$1 - \|y - x\|_x \leq \frac{\|v\|_y}{\|v\|_x} \leq \frac{1}{1 - \|y - x\|_x}, \text{ for all } v \neq 0.$$

Let SC be the family of such functionals.

Let $g(y)$ be the gradient of the functional f defined using the original inner product $\langle \cdot, \cdot \rangle$. In the local intrinsic inner product $\langle \cdot, \cdot \rangle_x$, the corresponding gradient $g_x(y)$ and hessian $H_x(y)$ are given by:

$$g_x(y) := H(x)^{-1} g(y), \tag{4}$$

$$H_x(y) := H(x)^{-1} H(y). \tag{5}$$

Definition 3.2 *A functional is a* (strongly nondegenerate self-concordant) barrier functional *if $f \in SC$ and*

$$\vartheta_f := \sup_{x \in D_f} \|g_x(x)\|_x^2 < \infty. \tag{6}$$

Let SCB be the family of such functionals.

Definition 3.3 *Let K be a closed convex cone and $f \in SCB$, $f : int(K) \to I\!R$. f is* logarithmically homogeneous *if for all $x \in int(K)$ and $t > 0$:*

$$f(tx) = f(x) - \vartheta_f \ln(t). \tag{7}$$

Equivalently, f is logarithmically homogeneous *if, for all $x \in int(K)$ and all $t > 0$:*

$$g_x(tx) = \frac{1}{t} g_x(x). \tag{8}$$

THEOREM 3.1 *(Theorem 2.3.9. [14].) If f is a self-concordant logarithmically homogeneous barrier functional then:*

$$H(tx) = \frac{1}{t^2}H(x),\ g_x(x) = -x \text{ and } \|g_x(x)\|_x = \sqrt{\vartheta_f}.$$

In linear programming such a logarithmically homogeneous self-concordant barrier functional is: $f : \mathbb{R}^n_+ \to \mathbb{R}$ with $f(x) := -\sum_{i=1}^{n} \ln(x_i)$. In this case $\vartheta_f = n$. For the SOCP case, the functional is given by $f(x) := -\ln(x_1^2 - \sum_{i=2}^{n} x_i^2)$, with $\vartheta_f = 2$. In the case of semidefinite programming such a functional is given by $f(X) := -\ln\det(X)$, with X a positive semidefinite matrix, $X \in \mathcal{S}_n$. The corresponding value for ϑ_f is $\vartheta_f = n$.

Most of the following results (taken from [14]) are technical in nature. They are needed in our analysis of the algorithm.

THEOREM 3.2 *(Theorem 2.2.2. [14].) If $f \in SC$, $x \in D_f$ and $y \in B_x(x,1)$, then*

$$|f(y) - f(x) - \langle g(x), y-x\rangle_X - \frac{1}{2}\|y-x\|_x^2| \leq \frac{\|y-x\|_x^3}{3(1-\|y-x\|_x)}.$$

If we take $y = x + d$ with $\|d\|_x < 1$ then

$$f(x+d) - f(x) \leq \langle g(x), d\rangle_X + \frac{1}{2}\|d\|_x^2 + \frac{\|d\|_x^3}{3(1-\|d\|_x)}. \tag{9}$$

THEOREM 3.3 *(Theorem 2.3.8. [14].) Assume $f \in SCB$ and $x \in D_f$. If $y \in \bar{D}_f$, then for all $0 < t \leq 1$,*

$$f(y + t(x-y)) \leq f(x) - \vartheta_f \ln t. \tag{10}$$

If the functional f is also logarithmically homogeneous, then a direct consequence of *Theorem 3.3* is the next lemma.

LEMMA 3.1 *Let $f \in SCB$ be a logarithmically homogeneous functional. If $x \in D_f$, $y \in \bar{D}_f$ and for all $t \geq 0$ then*

$$f(x+ty) \leq f(x). \tag{11}$$

If the domain of f is a cone K then the geometrical interpretation of *Lemma 3.1* is that x maximizes f over the cone $x + K$.

DEFINITION 3.4 *Let K be a cone and $z \in int(K)$. The* dual cone *of K is*

$$K^* = \{s \in X : \langle x, s\rangle_X \geq 0 \text{ for all } x \in K\}. \tag{12}$$

The dual cone of K with respect to the local inner product $\langle\cdot,\cdot\rangle_z$ *is given by*

$$K^*_z := \{s \in X : \langle x, s\rangle_z \geq 0, \text{ for all } x \in K\}. \tag{13}$$

The cone K is intrinsically self-dual *if $K^*_z = K$ for all $z \in$ int (K).*

DEFINITION 3.5 *The* conjugate *of $f \in SCB$ with respect to $\langle\cdot,\cdot\rangle$ is*

$$f^*(s) := -\inf_{x \in int(K)} (\langle x, s\rangle + f(x)) \text{ with } s \in int(K^*_z).$$

In particular, the conjugate *of $f \in SCB$ with respect to $\langle\cdot,\cdot\rangle_z$ is*

$$f^*_z(s) := -\inf_{x \in int(K)} (\langle x, s\rangle_z + f(x)) \text{ with } s \in int(K^*_z).$$

A final definition:

Definition 3.6 *A functional $f \in SCB$ is* intrinsically self-conjugate *if f is logarithmically homogeneous, if K is intrinsically self-dual, and for each $z \in int(K)$ there exists a constant C_z such that $f_z^*(s) = f(s) + C_z$ for all $s \in int(K)$.*

A cone K is self-scaled *or* symmetric *if $int(K)$ is the domain of an intrinsically self-conjugate barrier functional.*

Lemma 3.2 *Let K be a self-scaled cone. Then*

$$K = K_z^* = H(z)^{-1}K^* = H(z)^{-1}K. \tag{14}$$

Hence, for any $z \in int(K)$, $H(z)$ is a linear automorphism of K.

Lemma 3.3 *(Proposition 3.5.1. [14].) If $f : int(K) \to I\!R$ is an intrinsically self-conjugate barrier functional, then for all $z \in int(K)$,*

$$f_z^*(s) = f(s) - (\vartheta_f + 2f(z)).$$

As a direct consequence:

$$g^* \equiv g \text{ and } H^* \equiv H.$$

Theorem 3.4 *(Theorem 3.3.4. [14].) Assume f is self-concordant. Then $f^* \in \mathrm{C}^2$. Moreover, if x and s satisfy $s = -g(x)$, then*

$$-g^*(s) = x \text{ and } H^*(s) = H(x)^{-1}.$$

Starting now, all the functionals we will deal with will be intrinsically self-conjugate barrier functionals.

For each cone K we will consider a fixed vector $e \in \mathrm{int}(K)$ and we will take all the inner products to be scaled by e.

Starting now, unless explicitly stated otherwise, each time we deal with an intrinsic self-conjugate functional f defined on a Hilbert space $(X, \langle\cdot,\cdot\rangle_X)$, the inner product will be thought to be the one induced by e (i.e. $\langle u, v\rangle = \langle u, H(e)v\rangle_X$ where $\langle\cdot,\cdot\rangle_X$ is the original inner-product on X). Accordingly, we will denote $K^* := K_e^*$, $g(x) := g_e(x)$ to be the gradient of f, $H(x) := H_e(x)$ to be the hessian and so on.

Also if A_X^* is the adjoint operator of A in the original inner product, then $H(e)^{-1}A_X^*$ is the adjoint operator of A in the local inner product induced by e. We will denote

$$A^* := H(e)^{-1}A_X^*. \tag{15}$$

With this notation in mind, the vector e has some immediate and useful properties:

$$\|e\| = \sqrt{\vartheta_f}, g(e) = -e, H(e) = I. \tag{16}$$

Renegar proved in [14] the following result

Theorem 3.5 *Let f be an intrinsically self-conjugate barrier functional. Then, for any $x \in int(K)$:*

$$H(x)^{\frac{1}{2}}e = -g(x). \tag{17}$$

with H and g being the hessian and gradient of f considered in the local inner product induced by e.

4. On scaled inner products Let $(X, \langle\cdot,\cdot\rangle_X)$ be a finite dimensional Hilbert space, with K a self-scaled cone and $f : X \to I\!R$ the corresponding self-conjugate functional. Let $e \in \mathrm{int}(K)$ be a fixed point chosen arbitrarily.

Define the inner product $\langle\cdot,\cdot\rangle := \langle\cdot,\cdot\rangle_{H(e)}$ to be the local inner product induced by e, i.e.:

$$\langle u, v\rangle = \langle u, H(e)v\rangle_X.$$

For this point e define the set $\mathcal{B} := \{v \in X : e \pm v \in \mathrm{int}(K)\}$. Using this set define a new norm on X:

$$|v| := \inf\{t \geq 0 : \frac{1}{t}v \in \mathcal{B}\}.$$

Lemma 4.1 *(Theorem 3.5.7. [14].) Assume K is self-scaled. If $x \in K$ satisfies $|x-e|<1$, then for all $v \neq 0$:*

$$\frac{1}{1+|x-e|} \leq \frac{\|v\|_x}{\|v\|} \tag{18}$$

and

$$\frac{\|v\|_{-g(x)}}{\|v\|} \leq 1+|x-e|. \tag{19}$$

Note here that $\|v\|_x = \|H(x)^{\frac{1}{2}}v\|$ with $H(x)$ and $\|\cdot\|$ being the ones induced by e.

This lemma gives a lower bound on the minimum eigenvalue for the Hessian of f computed in the norm induced by e at any point x such that $|x-e|<1$:

$$\lambda_{\min}(H(x)) = \inf_{v\neq 0} \frac{\|H(x)^{\frac{1}{2}}v\|^2}{\|v\|^2} = \inf_{v \neq 0} \frac{\|v\|_x^2}{\|v\|^2} > \frac{1}{4}. \tag{20}$$

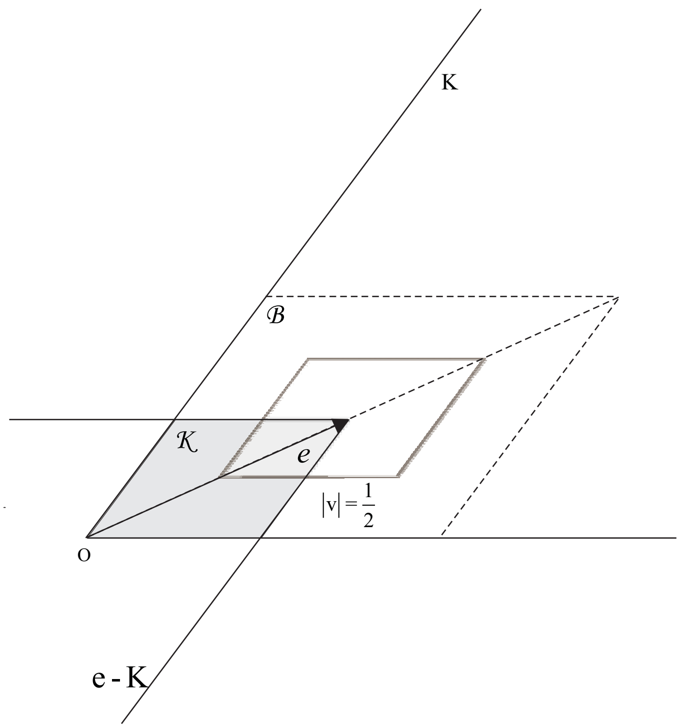

Figure 1: The sets $\mathcal{B}$, $\mathcal{K}$ and the level set $|v| = \frac{1}{2}$.

Now let's take a look at the domain described by $|x-e|<1$. We claim that:

Lemma 4.2 $\mathcal{K}$:=*int*$(K) \cap (e-K) \subseteq \{x \in$ *int*$(K) : |x-e|<1\}$.

Proof. Let $y \in$ int$(K) \cap (e-K)$. Then $y = e - z$, with $z \in K$. The point $y - e \in \mathcal{B}$ because $e + y - e = y \in \text{int}(K)$ and $e - (y-e) = z + e \in \text{int}(K)$.

Let y' be the point of intersection between ∂K and the line that goes through e and has the direction $y - e$. Then $y' - e = \tau(y-e)$ for some $\tau > 1$. The middle point between y and y' is clearly a point interior to K. Moreover,

$$e + \frac{y+y'}{2} - e = \frac{y+y'}{2} \in \text{int}(K)$$

and

$$e - \frac{y+y'}{2} + e = e + \frac{e-y}{2} + \frac{e-y'}{2} \in \text{int}(K).$$

So

$$\frac{1+\tau}{2}(e-y) = e - \frac{y+y'}{2} \in \mathcal{B}.$$

Then:

$$|y-e| = \inf\{t \geq 0 : \frac{1}{t}(e-y) \in \mathcal{B}\} \leq \frac{2}{\tau+1} < 1.$$

So $|y-e| < 1$. □

As a direct consequence of the previous analysis:

COROLLARY 4.1 *Let $f \in SCB$ be intrinsically self-conjugate. Then, for any $x \in int(K) \cap (e-K)$:*

$$\lambda_{\min}(H(x)) > \frac{1}{4}. \tag{21}$$

5. Analytic centers Let $(X, \langle\cdot,\cdot\rangle_X)$ and $(Y, \langle\cdot,\cdot\rangle_Y)$ be two Hilbert spaces of finite dimensions: $\dim X = n$, $\dim Y = m$. In X consider a full-dimensional self-scaled cone K, pointed at zero (i.e. $K \cap -K = \{0\}$) with the corresponding intrinsically self-conjugate barrier functional $f : X \to I\!R$. Let $A : X \to Y$ be a surjective linear operator.

The analytic center (the *AC*) of the domain $\mathcal{F}_P := \{x \in K : Ax = 0\}$ with respect to $f(x) + \langle c, x\rangle_X$ is the exact solution to the problem:

$$\begin{array}{rll} \min & f(x) + \langle c, x\rangle & \\ \text{subject to} & Ax = 0 & \qquad (P_1), \\ & x \in K. & \end{array}$$

Alternatively, the analytic center can be defined using the dual formulation of the previous problem. The analytic center of $\mathcal{F}_D := \{s \in K : A^*y + s = c\}$ with respect to $f_e^*(s)$ is the solution to:

$$\begin{array}{rll} \min & f_e^*(s) & \\ \text{subject to} & A^*y + s = c, & \qquad (D_1) \\ & s \in K. & \end{array}$$

One last thing to note here. The functional f is intrinsic self-conjugate. Then, by definition, $f_e^*(s) - f(s)$ is constant. So minimizing $f_e^*(s)$ is equivalent to minimizing $f(s)$. In what will follow we will keep using the notation $f^*(s)$ although we are actually using $f_e^*(s)$.

Both problems have the same set of KKT conditions. Hence, for any analytic center the next equalities hold:

$$\begin{array}{rcl} g(x) + s & = & 0, \\ g(s) + x & = & 0, \\ Ax & = & 0, \\ A^*y + s & = & c, \\ x, s & \in & K. \end{array} \tag{22}$$

For simplicity we will say that x or y or s is an analytic center if they are the components of an analytic center.

We can introduce the notion of θ - analytic center by relaxing some of the previous equalities. First we will define this notion, then the following lemma will give an insight for this definition.

DEFINITION 5.1 *(x, y, s) is a θ - analytic center for $\mathcal{F}_P$, $\mathcal{F}_D$ iff $x \in \mathcal{F}_P$, $s \in \mathcal{F}_D$ and*

$$\|I - H(x)^{-\frac{1}{2}} H(s)^{-\frac{1}{2}}\| \leq \frac{\theta}{\sqrt{\vartheta_f}}. \tag{23}$$

LEMMA 5.1 *Let (x, y, s) be a θ - analytic center. Then:*

$$\|x + g(s)\|_{-g(s)} \leq \theta,$$
$$\|s + g(x)\|_{-g(x)} \leq \theta. \tag{24}$$

PROOF. We will prove only the first inequality. The second inequality then follows from Theorem 3.7.1 in Renegar [14]. Note that the inner product $\langle \cdot, \cdot \rangle_X$ is the one induced by e. Using *Theorem 3.4*:

$$\|x + g(s)\|^2_{-g(s)} = \quad \langle x + g(s), H(s)^{-1}(x + g(s)) \rangle_X.$$

Next we will use the fact that, from *Theorems 3.1* and *3.5* and (4), for any $x \in K$:

$$g(x) = H(x)^{\frac{1}{2}} e \text{ and } x = -g_x(x) = -H(x)^{-1} g(x).$$

Based on these:

$$\begin{aligned}
&\|x + g(s)\|^2_{-g(s)} = \langle -H(x)^{-1} g(x) + g(s), H(s)^{-1}(-H(x)^{-1} g(x) + g(s)) \rangle_X \\
&= \langle -H(x)^{-1} H(x)^{\frac{1}{2}} e + H(s)^{\frac{1}{2}} e, H(s)^{-1}(-H(x)^{-1} H(x)^{\frac{1}{2}} e + H(s)^{\frac{1}{2}} e) \rangle_X \\
&= \langle -H(s)^{-\frac{1}{2}} H(x)^{-\frac{1}{2}} e + e, -H(s)^{-\frac{1}{2}} H(x)^{-\frac{1}{2}} e + e \rangle_X.
\end{aligned}$$

So:

$$\|x + g(s)\|_{-g(s)} \leq \|I - H(s)^{-\frac{1}{2}} H(x)^{-\frac{1}{2}}\| \|e\| \leq \theta.$$

□

The motivation for using this definition for a θ - analytic center should be clear if we compare it with the usual definition used in linear programming for a θ - analytic center:

$$\|e - xs\| \leq \theta.$$

with e being the vector of all ones and xs the Hadamard product of the vectors x and s.

Using the fact that in linear programming case the hessian is given by $H(x) = diag(x^{-2})$ (where $diag(x)$ is the diagonal matrix with the vector x being the diagonal and x^{-1} is the vector with components x_i^{-1}) our definition reduces to:

$$\|I - H(x)^{-\frac{1}{2}} H(s)^{-\frac{1}{2}}\| = \|diag(e - xs)\| = \max_i (1 - x_i s_i) \leq \frac{\theta}{\sqrt{\vartheta_f}}.$$

Note here that, in the linear case, our definition for a θ analytic center becomes the standard definition form the LP case, the only difference being that we use the infinity norm instead of the Euclidean norm. Using *Lemma 5.1* it is clear that our definition is close to the one used in the linear programming case:

$$\|x + g(s)\|^2_{-g(s)} = (x - s^{-1})^T diag(s^2)(x - s^{-1}) = \|e - xs\|^2.$$

LEMMA 5.2 *If (x, y, s) is the analytic center for the intrinsically self-conjugate barrier functional f then $H(s)H(x) = I$.*

PROOF. Note that $g^* \equiv g$ and $H^* \equiv H$ because f is an intrinsically self-conjugate functional. Because (x, y, s) is an analytic center then $s = -g(x)$ so, using *Theorem 3.4* we get that $H(s) = H^*(s) = H(x)^{-1}$. Hence the conclusion. □

In a linear programming formulation $H(s)H(x) = I$ translates into $x_i s_i = 1$ for all i. This is the exact expression that defines the exact analytic center in the linear programming case.

LEMMA 5.3 *Let f be an intrinsically self-conjugate barrier functional defined on a self-scaled cone K. Let $x, s \in K$ such that $x = -g(s)$. Then:*

$$f(x) + f^*(s) = -\vartheta_f. \tag{25}$$

Proof. Because f is self-conjugate we have $g^*(s) = g(s)$. Renegar proved in Theorem 3.3.4 of [14] that regardless of the inner product, the conjugate functional satisfies:

$$f^*(s) = \langle g^*(s), s\rangle - f(-g^*(s)).$$

So $f^*(s) = \langle g(s), s\rangle - f(x) = -\vartheta_f - f(x)$. □

Suppose that x is a feasible point in $\mathcal{F}_P$. If f is a self-concordant functional, then, by definition, $\|\Delta x\|_x \leq 1$ implies that $x + \Delta x$ is feasible.

This inequality describes an ellipsoid around the point x (also known as Dikin's ellipsoid). This ellipsoid defines a region around the point x where $x+\Delta x$ is feasible too. The following lemmas will give sufficient conditions on Δx and Δs to get $x + \Delta x$, $s + \Delta s$ feasible, given that x and s are feasible in $\mathcal{F}_P$ and $\mathcal{F}_D$ respectively.

Lemma 5.4 *Let* $\mathcal{E}_P = \{\Delta x \in X : A\Delta x = 0, \|\Delta x\|_x \leq 1\}$. *Let* (x, y, s) *be a* θ *- analytic center. Then:*

$$(1 + \frac{\theta}{\sqrt{\vartheta_f}})^{-1}\mathcal{E}_P \subseteq \{\Delta x \in X : A\Delta x = 0, \|\Delta x\|_{H(s)^{-1}} \leq 1\} \subseteq (1 - \frac{\theta}{\sqrt{\vartheta_f}})^{-1}\mathcal{E}_P.$$

Proof.

$$\begin{aligned}
\|\Delta x\|_{H(s)^{-1}} &= \|H(s)^{-\frac{1}{2}}\Delta x\|_X = \|H(s)^{-\frac{1}{2}}H(x)^{-\frac{1}{2}}H(x)^{\frac{1}{2}}\Delta x\|_X \\
&\leq \|H(s)^{-\frac{1}{2}}H(x)^{-\frac{1}{2}}\|\|H(x)^{\frac{1}{2}}\Delta x\|_X \leq (1 + \frac{\theta}{\sqrt{\vartheta_f}})\|H(x)^{\frac{1}{2}}\Delta x\|_X \\
&= (1 + \frac{\theta}{\sqrt{\vartheta_f}})\|\Delta x\|_x.
\end{aligned}$$

Also:

$$\begin{aligned}
\|\Delta x\|_x &= \|H(x)^{\frac{1}{2}}\Delta x\|_X = \|H(x)^{\frac{1}{2}}H(s)^{\frac{1}{2}}H(s)^{-\frac{1}{2}}\Delta x\|_X \\
&\leq \|H(x)^{\frac{1}{2}}H(s)^{\frac{1}{2}}\|\|\Delta x\|_{H(s)^{-1}} \leq \frac{1}{(1 - \frac{\theta}{\sqrt{\vartheta_f}})}\|\Delta x\|_{H(s)^{-1}}.
\end{aligned}$$

□

A similar result holds for the Dikin's ellipsoid around s.

Let $\mathcal{E}_D = \{\Delta s \in X : \Delta s = -A^*\Delta y, \|\Delta s\|_s \leq 1\}$.

Lemma 5.5 *Let* (x, y, s) *be a* θ *- analytic center. Then:*

$$(1 + \frac{\theta}{\sqrt{\vartheta_f}})^{-1}\mathcal{E}_D \subseteq \{\Delta s : \Delta s = -A^*\Delta y, \|\Delta s\|_{H(x)^{-1}} \leq 1\} \subseteq (1 + \frac{\theta}{\sqrt{\vartheta_f}})^{-1}\mathcal{E}_D.$$

Because analytic centers are minimizers of convex functionals defined on closed, bounded, convex sets, the method of choice for computing them is the Newton method. The Newton step is defined to be the vector $H(x)^{-1}g(x)$. This vector is the same as $g_x(x)$. Immediately we can see that, for logarithmically homogeneous barrier functionals the Newton step has constant length if measured in the norm induced by x: $\|g_x(x)\|_x = \sqrt{\vartheta_f}$. One advantage of using a self concordant barrier functional is that the Newton step doesn't change when the local inner product changes. This gives us more flexibility in the way we choose the local inner product.

When computing approximate analytic centers we need a way of estimating distances to the exact analytic center. When working with general functionals it is impossible to achieve this without knowing the exact analytic center. This problem is eliminated when using self-concordant functionals. This is because we can use local inner products instead of the original one. We can compute the distance between two points x and y without knowing y. All we need to do is to use $\|x - y\|_y$ to measure the distance.

All these properties will play an important role when we will analyze the complexity of recovering the analytic center after adding cuts.

6. Assumptions and Notations We assume that all the operators $A_i : X_i \to Y$, $i \geq 1$ defining the cuts are injective, hence the adjoint operators A_i^* are surjective. Also, wlog we assume that $\|A_i\| = 1$. The fact that A_i is injective gives also a bound on how many cuts we can add at a certain moment: $p_i \leq m$.

For each space $(X_i, \langle\cdot,\cdot\rangle_i)$ we will use the local norm induced by an arbitrary element $e_i \in \mathrm{int}(K_i)$. So whenever we use $\langle\cdot,\cdot\rangle_i$ we will actually mean $\langle\cdot,\cdot\rangle_{e_i}$. If there is no danger for confusion, we will also use $\langle\cdot,\cdot\rangle$ instead of $\langle\cdot,\cdot\rangle_{e_i}$.

The following assumptions are not critical for our analysis. We use them just to keep the analysis simpler and easier to understand. The analysis would be the same without these assumptions but the notation would be more complicated.

We assume that $\|H_i(e_i)^{-1}\| = 1$ for $i \geq 0$, where H_i are the hessians corresponding to the intrinsically self-conjugate functionals that are generated by the algorithm. The hessians are computed in the original inner product (not the scaled one). To ensure this, it is enough to pick an arbitrary $e_i' \in \mathrm{int}(K_i)$. Then take $e_i := \|H_i^{-1}(e_i')\|^{-\frac{1}{2}} e_i'$. Because f_i is logarithmically homogeneous (hence $H_i^{-1}(tx) = t^2 H_i^{-1}(x)$) for e_i we have $\|H_i^{-1}(e_i)\| = 1$. We can scale e_0 in a similar way to get $\|H_0(e_0)\| = 1$.

Let $\sigma_i := \sqrt{\frac{p_i}{\vartheta_i}} e_i$. The length of this vector, measured in the local inner product induced by e_i is $\|\sigma_i\| = \sqrt{p_i}$. Without loss of generality, we can assume that $f_i(\sigma_i) = 0$. We can do this easily. If f_i evaluated at this point is different from zero, then we can replace $f_i(x)$ by $f_i(x) - f_i(\sigma_i)$. Note that we can do this because the sum between a constant and an intrinsically self-conjugate barrier functional is an intrinsically self-conjugate barrier functional.

7. The Algorithm The algorithm starts with the initial set

$$\Omega_0 := \{y \in Y : -\tilde{c}_0 \preceq_{\tilde{K}_0} y \preceq_{\tilde{K}_0} \tilde{c}_0 \text{ with } \tilde{c}_0 \in int(\tilde{K}_0)\}$$

as the first outer-approximation of Γ. The cone $\tilde{K}_0$ is a self-scaled cone in $(\tilde{X}_0, \langle\cdot,\cdot\rangle_0)$ - an m - dimensional Hilbert space.

Let $X_0 := \tilde{X}_0 \oplus \tilde{X}_0$, $K_0 := \tilde{K}_0 \oplus \tilde{K}_0$ and let f_0 be the intrinsically self-conjugate barrier functional corresponding to K_0, $f_0 : \mathrm{int}(K_0) \to I\!R$. The set Ω_0 can be described by

$$\Omega_0 := \{y \in Y : A_0^* y + s = c_0 \text{ with } s \in K_0\}.$$

Here, A_0 is a linear operator defined on X_0, $A_0 : X_0 \to Y$ such that, its adjoint $A_0^* : Y \to X_0$ describes Ω_0 (i.e. $A_0 := I_m \oplus (-I_m)$).

Let $\tilde{e}_0 \in \mathrm{int}(\tilde{K}_0)$ be an arbitrary point chosen such that the inverse of the hessian $\tilde{H}_0$ of $\tilde{f}_0$ has unit norm at $\tilde{e}_0$: $\|\tilde{H}_0^{-1}(\tilde{e}_0)\| = 1$. Let's take $e_0 := \tilde{e}_0 \oplus \tilde{e}_0$. Then $e_0 \in K_0$ and $\|H_0^{-1}(e_0)\| = 1$ too.

Now, we change the inner product to be the one induced by e_0. Because of this change, the adjoint of the operator A_0 changes from A_0^* to $H_0(e_0)^{-1} A_0^*$. In order not to complicate the notation, we will define A_0^* to be the adjoint of A_0 in the new inner product. Also, we will use c_0 instead of the scaled vector $H_0(e_0)^{-1} c_0$.

Using this new notation, the set Ω_0 has the same description as before:

$$\Omega_0 = \{y \in Y : A_0^* y + s = c_0, \text{ with } s \in K_0\}.$$

Let (x_0, y_0, s_0) be the θ - analytic center corresponding to f_0. In order to obtain this point, we can take a sequence of primal-dual Newton steps, starting at the strictly feasible point $(e_0, 0, c_0) \in K_0 \times \Omega_0 \times K_0$. Note that e_0 and c_0 are strictly interior to K_0. Also, the origin is a point strictly feasible in Ω_0.

Once at y_0, the oracle is called. If $y_0 \in \Gamma$ the oracle returns y_0 and the algorithm stops with the solution to our problem. If $y_0 \notin \Gamma$, the oracle returns p_1 - central cuts. That is, the oracle returns a p_1-dimensional Hilbert space $(X_1, \langle\cdot,\cdot\rangle_1)$ together with a self-scaled cone K_1, the corresponding intrinsically self-conjugate barrier functional $f_1 : K_1 \to I\!R$ and a linear injective operator $A_1 : X_1 \to Y$ such that

$$\Gamma \subseteq \{y \in Y : A_1^* y + s = A_1^* y_0 \text{ with } s \in K_1\}.$$

The equality $A_1^* y + s = A_1^* y_0$ defines a central cut. It is called central because the point $(y, s) := (y_0, 0)$ lies on the cut with s being the vertex of the cone K_1.

We change the inner product on the space X_1 with a local one induced by a vector $e_1 \in \text{int}(K_1)$ chosen arbitrarily such that the norm of the hessian of f_1 computed in the original norm at e_1 is unitary. Also we change the functional f_1 by adding a constant such that the modified functional:

$$f_1(\sqrt{\frac{p_1}{\vartheta_1}} e_1) = 0.$$

(as already discussed in *Section 6*).

Now we build the new instance of the algorithm. First, let $\bar{X}_1 := X_0 \oplus X_1$ be an $(2m+p_1)$ - dimensional Hilbert space with the inner product induced by the inner products of X_0 and X_1. Let $\bar{K}_1 := K_0 \oplus K_1$ be the new self-scaled cone with the corresponding intrinsically self-conjugate barrier functional $\bar{f}_1 := f_0 \oplus f_1$. After adding the new cuts Ω_0 becomes

$$\Omega_1 := \Omega_0 \cap \{y \in Y : A_1^* y + s = A_1^* y_0 \text{ with } s \in K_1\}.$$

For the new instance of the algorithm, the old θ - analytic center (x_0, y_0, s_0) becomes $(x_0 \oplus 0_{p_1}, y_0, s_0 \oplus 0_{p_1})$ (with 0_{p_1} being the zero vector in X_1).

The point y_0 lies on the boundary of the new set Ω_1. First we will take a step to recover strict feasibility for this point. After that we generate a sequence of Newton steps that will take the point to (x_1, y_1, s_1), the θ - analytic center of the new domain Ω_1. We will discuss more about these steps in sections 8, 9 and 10.

At this point we call the oracle again. If $y_1 \in \Gamma$, we stop with the solution to our problem. If $y_1 \notin \Gamma$, the oracle returns p_2 central cuts that are added to the old instance of the algorithm, generating a new set Ω_2. Then the algorithm proceeds as before.

We will prove that the algorithm must stop with a solution after a sufficiently large number of cuts has been added. The analysis of the number of cuts generated is presented in sections 9, 11, and 12.

After i iterations, the i-th instance of the algorithm is described by a Hilbert space $\bar{X}_i = \bigoplus_{j=0}^{i} X_j$ together with a self-scaled cone $\bar{K}_i = \bigoplus_{j=0}^{i} K_j$, the domain of an intrinsically self-conjugate barrier functional $\bar{f}_i = \bigoplus_{j=0}^{i} f_j$. The current set Ω_i is described by the linear operator $\bar{A}_i : \bar{X}_i \to Y$, with $\bar{A}_i = \bigoplus_{j=0}^{i} A_j$. All linear operators $A_j : X_j \to Y$, $j \geq 1$, are injective and the inner products considered in the p_j-dimensional Hilbert spaces X_j are the ones induced by fixed elements $e_j \in \text{int}(K_j)$. These vectors e_j are strictly interior to the respective cones K_j and $\|H_j(e_j)^{-1}\| = 1, \forall j \geq 1$ (here H_j is the hessian of f_j computed in the original norm of X_j, not in the local norm induced by e_j).

The idea behind our analysis is quite simple. As the algorithm proceeds, the sequence of sets Ω_i is generated. We will use the exact analytic centers s_i^c of these sets. It is not necessary to generate exact analytic centers, we only consider them for the analysis to go through. The main steps are:

- Get an upper bound UB_i for $f_i^*(s_i^c)$, for any i
- Compare two consecutive f_i^* at the corresponding AC s_i^c:

$$f_{i+1}^*(s_{i+1}^c) \geq f_i^*(s_i^c) + LB_i$$

- After k steps :

$$UB_k \geq f_0^*(s_0^c) + \sum_{i=0}^{k-1} LB_i$$

- We prove that $UB_k \to \infty$ slower than $\sum_{i=0}^{k-1} LB_i$ does
- The algorithm stops and concludes the problem is infeasible as soon as

$$UB_k < f_0^*(s_0^c) + \sum_{i=0}^{k-1} LB_i$$

8. The recovery of feasibility In this section we will study the impact of the central cuts added through an θ - analytic center and how feasibility can be restored.

Consider an instance of the algorithm described by an intrinsically self-conjugate functional f_1 defined on a Hilbert space $(X_1, \langle\cdot,\cdot\rangle_1)$ with the corresponding full-dimensional self-scaled cone K_1 pointed at zero. We consider here the case $i = 1$ for notational convenience. This analysis applies to any stage i of the algorithm. The outer-approximation of the domain of interest Γ in this instance is

$$\Omega_1 = \{y \in Y : A_1^* y + s = c_1, s \in K_1\},$$

with $A_1 : X_1 \to Y$ a linear operator. Let (x_1, y_1, s_1) be the θ - analytic center for $\mathcal{F}_P$, $\mathcal{F}_D$. So its components must verify:

$$A_1 x_1 = 0, \tag{26}$$

$$A_1^* y_1 + s_1 = c_1, \tag{27}$$

$$x_1, s_1 \in K_1 \text{ and } y_1 \in Y. \tag{28}$$

We add p central cuts at this point: $A_2^* y + s = c_2$ with $A_2^* y_1 = c_2$. The operator A_2 is defined on a p - dimensional Hilbert space $(X_2, \langle\cdot,\cdot\rangle_2)$. We assume that A_2 is injective and linear.

The outer-approximation domain Ω_1 becomes

$$\Omega_2 := \Omega_1 \cap \{y \in Y : A_2^* y + s = c_2, s \in K_2\}.$$

K_2 is a self-scaled cone in X_2 and let $f_2 : X_2 \to I\!R$ be the corresponding intrinsically self-conjugate functional.

After adding the cuts, the primal and dual feasible sets $\mathcal{F}_P$ and $\mathcal{F}_D$ are changed:

$$\mathcal{F}_P := \{x \oplus \beta : A_1 x + A_2 \beta = 0 \text{ with } x \in K_1, \beta \in K_2\}$$

and

$$\mathcal{F}_D := \{s \oplus \gamma : A_1^* y + s = c_1, A_2^* y + \gamma = c_2 \text{ with } s \in K_1, \gamma \in K_2, y \in Y\}.$$

Let $f := f_1 \oplus f_2$, $X := X_1 \oplus X_2$. After adding the cuts the old point (x_1, y_1, s_1) becomes (x_2, y_2, s_2):

$$x_2 = x_1 \oplus \beta, y_2 = y_1, s_2 = s_1 \oplus \gamma,$$

with y_2 on the boundary of the new domain Ω_2. At this new point, $\beta = 0$ and $\gamma = 0$ hence both f and f^* are infinitely large. One step to recover feasibility is needed. Let this step be: $\Delta x \oplus \beta$, Δy and $\Delta s \oplus \gamma$. The new point must be feasible in $\mathcal{F}_P$ and $\mathcal{F}_D$ so:

$$A_1(x_1 + \Delta x) + A_2 \beta = 0, \tag{29}$$

$$A_1^*(y_1 + \Delta y) + s_1 + \Delta s = c_1, \tag{30}$$

$$A_2^*(y_1 + \Delta y) + \gamma = c_2, \tag{31}$$

with $x_1, x_1 + \Delta x, s_1, s_1 + \Delta s \in K_1$ and $\beta, \gamma \in K_2$. So, in order to get back feasibility we need to have:

$$A_1 \Delta x + A_2 \beta = 0, \tag{32}$$

$$A_1^* \Delta y + \Delta s = 0, \tag{33}$$

$$A_2^* \Delta y + \gamma = 0. \tag{34}$$

In moving away from the boundary of Ω_2 we should try to minimize as much as possible the contribution of β and γ to the potential functions. One way of doing this is to set-up the next problems:

$$\begin{array}{rl} \min & f_2(\beta) \\ \text{subject to} & A_1 \Delta x + A_2 \beta = 0 \\ & \beta \in K_2 \end{array}$$

and

$$\begin{array}{rl} \min & f_2^*(\gamma) \\ \text{subject to} & A_2^* \Delta y + \gamma = 0 \\ & \gamma \in K_2 \end{array}$$

These two formulations do not describe completely our problem. What is needed is a constraint that insures that $x_1 + \Delta x$ and $s_1 + \Delta s$ stay feasible in K_1 too. Using the analysis of the Dikin's Ellipsoids we have already made, it is enough to add $\|\Delta x\|_{H_1(s_1)^{-1}} \leq 1 - \frac{\theta}{\sqrt{\vartheta_f}}$ and $\|\Delta s\|_{s_1} \leq 1$ to keep $x_1 + \Delta x$ and $s_1 + \Delta s$ feasible in K_1. Next we will analyze the problems using $\|\Delta x\|_{H_1(s_1)^{-1}} \leq 1$. We do this to keep the analysis clear. Later we will scale the steps by $\alpha < 1 - \frac{\theta}{\sqrt{\vartheta_f}}$ so the feasibility will be preserved.

So a good choice is to take β and γ to be the solutions to the following problems:

$$\begin{array}{rl} \min & f_2(\beta) \\ \text{subject to} & A_1\Delta x + A_2\beta = 0, \\ & \|\Delta x\|_{H_1(s_1)^{-1}} \leq 1, \\ & \beta \in K_2 \end{array} \qquad (P_2)$$

and

$$\begin{array}{rl} \min & f_2^*(\gamma) \\ \text{subject to} & A_2^*\Delta y + \gamma = 0, \\ & \|\Delta s\|_{s_1} \leq 1, \\ & \gamma \in K_2. \end{array} \qquad (D_2)$$

These two problems are well posed. The feasible regions are not empty because A_1 is surjective, A_2 is injective and the equality constraints are homogeneous. The objectives are convex functionals so, if the minimum exists, it is unique. The cone K_2 is the domain for both f_2 and f_2^*. The second constraint in each problem ensures that the feasible sets don't contain rays. So both problems have a unique optimal value.

This approach is similar to the one proposed for the linear programming case by Goffin and Vial in [2]. It is a generalization of the approach used by Mitchell and Todd in [7] for the case $p_2 = 1$ (only one cut is added at each iteration).

Now, let's analyze (P_2). The KKT conditions are:

$$g_2(\beta) + A_2^*\lambda = 0, \tag{35}$$
$$A_1^*\lambda + \nu H_1(s_1)^{-1}\Delta x = 0, \tag{36}$$
$$\nu(1 - \langle \Delta x, H_1(s_1)^{-1}\Delta x\rangle_1) = 0, \tag{37}$$
$$A_1\Delta x + A_2\beta = 0. \tag{38}$$

If we take

$$\Delta x = -H_1(s_1)A_1^*(A_1H_1(s_1)A_1^*)^{-1}A_2\beta, \tag{39}$$
$$\nu = \vartheta_{f_2}, \tag{40}$$
$$\lambda = \vartheta_{f_2}(A_1H_1(s_1)A_1^*)^{-1}A_2\beta \tag{41}$$

both equations (36) and (38) are verified.

For β we use the approach used by Goffin and Vial in [2] and we will take it to be the solution to the next problem:

$$\min_{\beta\in K_2} \frac{\vartheta_{f_2}}{2}\langle \beta, V\beta\rangle_2 + f_2(\beta) \tag{42}$$
$$\text{with } V = A_2^*(A_1H_1(s_1)A_1^*)^{-1}A_2. \tag{43}$$

The optimality condition for this minimization problem is given by:

$$\vartheta_{f_2}V\beta + g_2(\beta) = 0. \tag{44}$$

It is easy to verify that the equation (35) holds true for β the solution for problem (42). For equation (37) it is enough to note that

$$\|\Delta x\|^2_{H_1(s_1)^{-1}} = \langle \Delta x, H_1(s_1)^{-1}\Delta x\rangle_1 = \langle \beta, V\beta\rangle_2 = -\frac{1}{\vartheta_{f_2}}\langle \beta, g_2(\beta)\rangle_2 = 1. \tag{45}$$

Now, let's look at the second problem (D_2). The optimality conditions are:

$$\begin{aligned} g_2^*(\gamma) + \mu &= 0, & (46)\\ A_2\mu + \nu A_1 H_1(s_1) A_1^* \Delta y &= 0, \ \nu \geq 0, & (47)\\ A_2^* \Delta y + \gamma &= 0, & (48)\\ \nu(1 - \langle \Delta y, A_1 H_1(s_1) A_1^* \Delta y \rangle_Y) &= 0. & (49)\end{aligned}$$

The solution to this problem is given by:

$$\begin{aligned} \Delta y &= -(A_1 H_1(s_1) A_1^*)^{-1} A_2 \beta, & (50)\\ \gamma &= V\beta = A_2^*(A_1 H_1(s_1) A_1^*)^{-1} A_2 \beta, & (51)\\ \mu &= \vartheta_{f_2} \beta, & (52)\\ \nu &= \vartheta_{f_2}. & (53)\end{aligned}$$

Here β is the solution of problem (42). The equations (48) and (47) are obviously satisfied. For equation (49):

$$\langle \Delta y, A_1 H_1(s_1) A_1^* \Delta y \rangle_Y = \langle \beta, A_2^*(A_1 H_1(s_1) A_1^*)^{-1} A_2 \beta \rangle_2 = \langle \beta, V\beta \rangle_2 = 1.$$

Finally, for equation (46) it is enough to notice that $-g_2(-g_2(\beta))$ is equal to both β (because $-g_2$ is an involution, as can be seen from *Theorem 3.4)* and $-g_2(\vartheta_{f_2}\gamma)$ (as given by equation (44)). Using the fact that f_2 is logarithmically homogeneous, the conclusion follows immediately:

$$g_2(\gamma) = \vartheta_{f_2} g_2(\vartheta_{f_2}\gamma) = \vartheta_{f_2} g_2(-g_2(\beta)) = -\vartheta_{f_2}\beta = -\mu. \tag{54}$$

Instead of full steps Δx, Δs, some scaled steps $\alpha\Delta x$, $\alpha\Delta s$ are taken. The next lemma gives a characterization of such scaled steps.

Lemma 8.1 *Let Δx and Δs be the steps considered in the problems (P_2) and (D_2). For any $\alpha < (1 - \frac{\theta}{\sqrt{\vartheta_f}})\zeta$ with $0 < \zeta < 1$:*

$$\|\alpha \Delta x\|_{x_1} < \zeta \text{ and } \|\alpha \Delta s\|_{s_1} < \zeta.$$

Proof. Here we will use *Lemma 5.4.*

$$\|\alpha \Delta x\|_{x_1} = \alpha \|\Delta x\|_{x_1} \leq \alpha \frac{1}{1 - \frac{\theta}{\sqrt{\vartheta_f}}} \|\Delta x\|_{H(s_1)^{-1}} < \zeta.$$

The second inequality is immediate:

$$\|\alpha \Delta s\|_{s_1} = \alpha \|\Delta s\|_{s_1} \leq \alpha < \zeta.$$

□

We have that $g_2(\beta) = -\vartheta_{f_2}\gamma$, with $\beta, \gamma \in K_2$. So we can use *Lemma 5.3:*

$$f_2(\beta) + f_2^*(\vartheta_{f_2}\gamma) = -\vartheta_{f_2}. \tag{55}$$

The fact that f_2 and f_2^* are logarithmically homogeneous implies:

$$\begin{aligned} f_2(\alpha\beta) + f_2^*(\alpha\gamma) &= f_2(\beta) + f_2^*(\gamma) - 2\vartheta_{f_2} \ln \alpha & (56)\\ &= -\vartheta_{f_2} - 2\vartheta_{f_2} \ln \alpha + f_2^*(\gamma) - f_2^*(\vartheta_{f_2}\gamma). & (57)\end{aligned}$$

So we proved that:

$$f_2(\alpha\beta) + f_2^*(\alpha\gamma) = -\vartheta_{f_2} - 2\vartheta_{f_2} \ln \alpha + \vartheta_{f_2} \ln \vartheta_{f_2}. \tag{58}$$

This equality provides a measure of the influence the added cut has over the self-concordant barrier functional.

9. Potentials In analyzing the complexity of the algorithm (for both local and global convergence) we will make use of primal-dual potentials. The way potentials change from one analytic center to the next one will give us a measure for the total number of cuts that can be introduced before the algorithm stops with a solution. We will also use potential functionals in finding the number of steps required to get to the θ - analytic center after new cuts are added in the problem.

DEFINITION 9.1 *For an instance of the algorithm described by the functional f, the vector c and the linear operator A, we define the primal-dual potential to be:*

$$\Phi_{PD}(x,s) := \langle c,x\rangle + f(x) + f^*(s) \text{ for any } x,s \in K.$$

It is customary to call $\langle c,x\rangle + f(x)$ the primal potential and $f^*(s)$ the dual potential. Note that if $Ax = 0$ and $s = c - A^*y$ for some y then $\langle c,x\rangle = \langle s,x\rangle$.

Let (x_1, y_1, s_1) be the current θ - analytic center with the corresponding primal-dual potential:

$$\phi_1 := \langle c_1, x_1\rangle_1 + f_1(x_1) + f_1^*(s_1).$$

After adding the cuts described by f_2, A_2 and c_2 we take a scaled step to get back into the feasible region. At this new point, the primal-dual potential is:

$$\begin{aligned}\phi_{new} := \langle c_1, x_1 &+ \alpha\Delta x\rangle_1 + \langle c_2, \alpha\beta\rangle_2 + f_1(x_1+\alpha\Delta x) + f_2(\alpha\beta)\\ &+ f_1^*(s_1+\alpha\Delta s) + f_2^*(\alpha\gamma).\end{aligned}$$

Using equation (58) the new potential can be written as

$$\phi_{new} = \phi_1 + \vartheta_{f_2}\ln\frac{\vartheta_{f_2}}{\alpha^2} - \vartheta_{f_2} + \alpha(\langle c_1,\Delta x\rangle_1 + \langle c_2,\beta\rangle_2) + F, \tag{59}$$

with

$$F = f_1(x_1+\alpha\Delta x) - f_1(x_1) + f_1^*(s_1+\alpha\Delta s) - f_1^*(s_1). \tag{60}$$

Because the cuts are central: $A_2^* y_1 = c_2$, hence

$$\langle\beta, c_2\rangle_2 = \langle A_2\beta, y_1\rangle_Y = -\langle A_1\Delta x, y_1\rangle_Y = -\langle\Delta x, A_1^* y_1\rangle_1.$$

Therefore,

$$\langle c_1,\Delta x\rangle_1 + \langle c_2,\beta\rangle_2 = \langle c_1 - A_1^* y_1, \Delta x\rangle_1 = \langle s_1,\Delta x\rangle_1.$$

So, finally:

$$\phi_{new} = \phi_1 + \alpha\langle s_1,\Delta x\rangle_1 + \vartheta_{f_2}\ln\frac{\vartheta_{f_2}}{\alpha^2} - \vartheta_{f_2} + F.$$

Now let's evaluate $F + \alpha\langle s_1,\Delta x\rangle_1$. Let's start with $\alpha\langle s_1,\Delta x\rangle_1 + f_1(x_1+\alpha\Delta x) - f_1(x_1)$. Note that the recovery step is scaled by $\alpha < (1 - \frac{\theta}{\sqrt{\vartheta_f}})\zeta$ so we can use the inequality (9):

$$\begin{aligned}\alpha\langle s_1,\Delta x\rangle_1 &+ f_1(x_1+\alpha\Delta x) - f_1(x_1)\\ &\leq \alpha\langle s_1,\Delta x\rangle_1 + \alpha\langle g_1(x_1),\Delta x\rangle_1 + \frac{1}{2}\|\alpha\Delta x\|_x^2 + \frac{\|\alpha\Delta x\|_x^3}{3(1-\|\alpha\Delta x\|_x)}.\end{aligned}$$

Now:

$$\begin{aligned}\langle s_1 + g_1(x_1), \Delta x\rangle_1 &= \langle s_1 + g_1(x_1), H_1(x_1)^{-\frac{1}{2}} H_1(x_1)^{\frac{1}{2}} \Delta x\rangle_1 \\ &\le \|H_1(x_1)^{-\frac{1}{2}}(s_1 + g_1(x_1))\|_1 \|\Delta x\|_{x_1} \\ &= \|s_1 + g_1(x_1)\|_{H_1(x_1)^{-1}} \|\Delta x\|_{x_1} \\ &= \|s_1 + g_1(x_1)\|_{-g_1(x_1)} \|\Delta x\|_{x_1} \\ &\le \theta \|\Delta x\|_{x_1}.\end{aligned}$$

Here, we used *Lemma 5.1* and the fact that $H(x)^{-1} = H(-g(x))$ (see *Theorem 3.4).*

So:

$$\begin{aligned}\alpha\langle s_1, \Delta x\rangle_1 \quad &+ \quad f_1(x_1 + \alpha\Delta x) - f_1(x_1) \le \\ &\le \quad \theta\|\alpha\Delta x\|_{x_1} + \frac{1}{2}\|\alpha\Delta x\|_{x_1}^2 + \frac{\|\alpha\Delta x\|_{x_1}^3}{3(1 - \|\alpha\Delta x\|_{x_1})}.\end{aligned}$$

Next we use the fact that the function $\theta x + \frac{1}{2}x^2 + \frac{x^3}{3(1-x)}$ is increasing on the open interval $(0, 1)$ and the recovery step is scaled to satisfy $\|\alpha\Delta x\|_{x_1} \le \zeta < 1$. This implies that

$$\alpha\langle s_1, \Delta x\rangle_1 + f_1(x_1 + \alpha\Delta x) - f_1(x_1) \le \theta\zeta + \frac{1}{2}\zeta^2 + \frac{\zeta^3}{3(1-\zeta)}. \tag{61}$$

A similar analysis for the second part of F: $f_1^*(s_1 + \alpha\Delta s) - f_1^*(s_1)$ gives:

$$f_1^*(s_1 + \alpha\Delta s) - f_1^*(s_1) \quad \le \theta\zeta + \tfrac{1}{2}\zeta^2 + \tfrac{\zeta^3}{3(1-\zeta)}. \tag{62}$$

Using inequalities (61) and (62) we get:

$$\phi_{new} \le \phi_1 + \vartheta_{f_2} \ln \frac{\vartheta_{f_2}}{\alpha^2} - \vartheta_{f_2} + 2\theta\zeta + \zeta^2 + \frac{2\zeta^3}{3(1-\zeta)}. \tag{63}$$

This upper bound on the primal-dual potential function will enable us to show a bound of $O(\vartheta_{f_2} \ln(\vartheta_{f_2}))$ on the number of Newton steps required to obtain a new θ - analytic center in *Theorem 10.4* in the next section. First, we conclude this section with an upper bound on the potential function value of a θ - analytic center, which thus provides an upper bound on ϕ_1.

THEOREM 9.1 *Let (x, y, s) be a θ - analytic center corresponding to an instance of the algorithm described by the functional f, the linear operator A and the vector c. Then,*

$$\Phi_{PD}(x, s) \le \frac{\theta^3}{3(1-\theta)} + \frac{\theta^2}{2}. \tag{64}$$

PROOF. Because (x, y, s) is a θ - analytic center we can use *Lemma 5.1* to get

$$\|x + g(s)\|_{-g(s)} \le \theta.$$

This inequality implies, using (3), that $x \in B_{-g(s)}(-g(s), \theta)$. Because $\theta < 1$ we can use *Theorem 3.2* to get:

$$\left| f(x) - f(-g(s)) + \langle -g(-g(s)), x + g(s)\rangle - \frac{1}{2}G(x,s)^2 \right| \le \frac{G(x,s)^3}{3\,(1 - G(x,s))}.$$

where $G(x, s) = \|x + g(s)\|_{-g(s)}$.

Because f is an intrinsically self-conjugate barrier functional we have:

$$f^*(s) = \langle g(s), s\rangle - f(-g(s)) \text{ and } -g(-g(s)) = s.$$

Using these equalities together with the fact that $\langle x, s\rangle = \langle c, x\rangle$ we can write:

$$f(x) + f^*(s) + \langle c, x\rangle \le \frac{1}{2}G(x,s)^2 + \frac{G(x,s)^3}{3(1 - G(x,s))}.$$

The functional $G(x, s)$ is bounded above by θ. Using this together with the fact that the function $\frac{1}{2}x^2 + \frac{x^3}{3(1-x)}$ is increasing for $0 < x < 1$, we get the desired conclusion. □

10. Complexity of recovering the θ - analytic center After the current point is moved back in the feasible region obtained from the old one by adding central cuts, a sequence of steps is required to get in the vicinity of the analytic center of the new domain. One way of obtaining such a point is to take some Newton steps. In this section we will prove that one way to achieve this is to use two different sequences of steps. We will use potential functionals in this analysis.

At the beginning, when the point is still far away from the analytic center, the directions used are the Nesterov-Todd directions. These directions were first used in interior-point algorithms in linear programming. Nesterov and Todd generalized them later for the general case of conic programming (see [4], [8] for more details). These directions will ensure that the primal-dual potential decreases by a fixed amount at each iteration. Once close enough to the analytic center, a different sequence of steps will bring the point to a θ - analytic center.

As before, let the primal-dual potential functional be:

$$\Phi(x,s) := \langle x,s\rangle + f(x) + f^*(s).$$

Before defining the Nesterov-Todd direction we will introduce some notation. Let L denote the null space of A (the surjective operator defining the feasible region) and $L^\perp$ the corresponding orthogonal space. Let $P_{L,v}(u)$ be the orthogonal projection of u onto L in the local inner product induced by v.

Let (x,y,s) be the current point with w the corresponding scaling point for the ordered pair (x,s) (i.e. $H(w)x = s$). Such a point is uniquely defined by x and s. Similarly we take w^* to be the scaling point for the ordered pair (s,x) (i.e. $H(w^*)s = x$). Renegar in [14] gives a detailed discussion about scaling points and their properties.

With these notations, the primal and dual Nesterov-Todd directions are given by:

$$d_x := -P_{L,w}(x + g_w(x)), \tag{65}$$
$$d_s := -P_{L^\perp,w^*}(s + g_{w^*}(s)). \tag{66}$$

Note here that if we use the inner products induced by x and s instead of the ones induced by w and w^*, the Nesterov-Todd directions become the usual Newton directions.

One important property of these directions is that they provide an orthogonal decomposition w.r.t. $\langle\cdot,\cdot\rangle_w$ for $-(x+g_w(x))$ (see [14]) :

$$d_x + H(w)^{-1}d_s = -(x + g_w(x)). \tag{67}$$

Using the local inner product induced by w we define for all $\tilde{x},\tilde{s} \in int(K)$:

$$\Phi_w(\tilde{x},\tilde{s}) = \langle\tilde{x},\tilde{s}\rangle_w + f(\tilde{x}) + f^*(\tilde{s}). \tag{68}$$

Now for any $x, w \in K$, $f(H(w)x) = f(x) + 2(f(w) - f(e))$ (see [14], formula (3.34)) so

$$f(\tilde{s}) = f(H(w)(H(w)^{-1}\tilde{s})) = f(H(w)^{-1}\tilde{s}) + 2f(w) - 2f(e).$$

Combining the previous expressions we conclude that:

$$\Phi_w(\tilde{x}, H(w)^{-1}\tilde{s}) = \Phi(\tilde{x},\tilde{s}) + 2f(e) - 2f(w). \tag{69}$$

Our goal is to prove that by taking a scaled Nesterov-Todd step, the primal-dual potential functional decreases by a constant value. We will use

$$\phi(t) := \Phi(x + td_x, s + td_s) \tag{70}$$

to find the scaling parameter t that minimizes the primal-dual potential.

Let's define:

$$\breve{\phi}_w(t) := \Phi_w(x + t\breve{d}_x, x + tH(w)^{-1}\breve{d}_s), \tag{71}$$
$$\breve{\phi}(t) := \Phi(x + t\breve{d}_x, x + t\breve{d}_s) \tag{72}$$

with $\breve{d}_x$, $\breve{d}_s$, the scaled vectors:

$$(\breve{d}_x, \breve{d}_s) := \frac{1}{\|H_w(x)\|_w^{\frac{1}{2}} \|x + g_w(x)\|_w}(d_x, d_s) \tag{73}$$

Now:

$$\begin{aligned}
\breve{\phi}_w(t) &:= \Phi_w(x + t\breve{d}_x, x + tH(w)^{-1}\breve{d}_s) \\
&= \Phi_w(x + t\breve{d}_x, H(w)^{-1}(H(w)x + t\breve{d}_s)) \\
&= \Phi_w(x + t\breve{d}_x, H(w)^{-1}(s + t\breve{d}_s)) \\
&= \Phi(x + t\breve{d}_x, s + t\breve{d}_s) + 2f(e) - 2f(w) \\
&= \breve{\phi}(t) + 2f(e) - 2f(w).
\end{aligned}$$

Using the approach from [14], let's denote:

$$\begin{aligned}
\psi_1(t) &:= \langle x + t\breve{d}_x, x + tH(w)^{-1}\breve{d}_s\rangle_w, & (74)\\
\psi_2(t) &:= f(x + t\breve{d}_x), & (75)\\
\psi_3(t) &:= f(x + tH(w)^{-1}\breve{d}_s) - \vartheta_f - 2f(e). & (76)\\
& & (77)
\end{aligned}$$

Using the fact that $f^*(s) = f(s) - (\vartheta_f + 2f(e))$ (as given in Lemma 3.3), we have:

$$\breve{\phi}_w(t) = \psi_1(t) + \psi_2(t) + \psi_3(t). \tag{78}$$

Because $\langle \breve{d}_x, H(w)^{-1}\breve{d}_s\rangle_w = 0$, the first functional $\psi_1(t)$ can be written as

$$\begin{aligned}
\psi_1(t) &= \psi_1(0) + t\langle x, \breve{d}_x + H(w)^{-1}\breve{d}_s\rangle_w \\
&= \psi_1(0) - t\frac{\langle x, x + g_w(x)\rangle_w}{\|H_w(x)\|_w^{\frac{1}{2}} \|x + g_w(x)\|_w}.
\end{aligned}$$

Renegar proved in [14] (Theorem 3.8.2) that:

$$\psi_2(t) \le \psi_2(0) + t\langle g_w(x), \breve{d}_x\rangle_w + \frac{t^2}{1-t}, \tag{79}$$

$$\psi_3(t) \le \psi_3(0) + t\langle g_w(x), H(w)^{-1}\breve{d}_s\rangle_w + \frac{t^2}{1-t}. \tag{80}$$

Using all these relations we can relate $\breve{\phi}_w(t)$ and $\breve{\phi}_w(0)$:

$$\breve{\phi}_w(t) \le \breve{\phi}_w(0) - t\frac{\|x + g_w(x)\|_w}{\|H_w(x)\|_w^{\frac{1}{2}}} + \frac{2t^2}{1-t}. \tag{81}$$

Then, immediately:

$$\breve{\phi}(t) \le \breve{\phi}(0) - t\frac{\|x + g_w(x)\|_w}{\|H_w(x)\|_w^{\frac{1}{2}}} + \frac{2t^2}{1-t}.$$

or

$$\Phi(x + t\breve{d}_x, s + t\breve{d}_s) \le \Phi(x, s) - t\frac{\|x + g_w(x)\|_w}{\|H_w(x)\|_w^{\frac{1}{2}}} + \frac{2t^2}{1-t}.$$

Next we will introduce a variant of Theorem 3.5.11 from [14]:

THEOREM 10.1 *Let K be a self-scaled cone. If $x, w \in int(K)$ then:*

$$\|x + g_w(x)\|_w \ge \max\{\|H_w(x)^{\frac{1}{2}}\|_w, \|H_w(x)^{-\frac{1}{2}}\|_w\} \min\{\frac{1}{5}, \frac{4}{5}\|x - w\|_w\}. \tag{82}$$

We are ready now to prove the following theorem:

THEOREM 10.2 *If* $\|x - w\|_w \geq \frac{1}{4}$ *then:*

$$\Phi(x + t\breve{d}_x, s + t\breve{d}_s) \leq \Phi(x, s) - \frac{1}{250}. \tag{83}$$

PROOF. The proof is based on the previous analysis and the fact that

$$\min_{0<t<1}\left(\frac{2t^2}{1-t} - \frac{t}{5}\right) < -\frac{1}{250}.$$

□

We know that, if (x, y, s) is the exact analytic center, then $x = w$. Also, the exact analytic center is the minimizer for the primal-dual potential functional $\Phi(x, s)$. Theorem 10.2 says that, as long as the point is sufficiently far away from the exact analytic center, the primal-dual potential is guaranteed to decrease by a constant quantity.

Because of the assumption made about the problem, the analytic center exists, so the primal-dual potential functional has a strictly feasible minimizer. This implies that, after a number of scaled Nesterov-Todd steps for the current point (x, y, s), $\|x - w\|_w < \frac{1}{4}$.

As soon as this happens, we will switch from using Nesterov-Todd steps to a new kind of step, suggested in [14]:

$$\delta_x := 2P_{L,w}(w - x), \tag{84}$$

$$\delta_s := 2P_{L^\perp,w^*}(w^* - s), \tag{85}$$

where $P_{L,w}$ is the orthogonal projection onto L (in the local product $\langle \cdot, \cdot \rangle_w$).

The key element here is the following theorem (see *Theorems 3.7.2 and 3.7.1* from [14] for a detailed proof):

THEOREM 10.3 *If at the current point* (x, y, s)*:*

$$\|x - w\|_w < \alpha < \frac{1}{4}$$

then at the new point $(x_+, s_+) := (x + \delta_x, s + \delta_s)$*:*

$$\|s_+ + g(x_+)\|_{-g(x_+)} < (1+\alpha)\frac{\alpha^2}{1-\alpha} < \frac{1}{5}. \tag{86}$$

If w_+ *is the scaling point for the ordered pair* (x_+, s_+)*, then:*

$$\|x_+ - w_+\|_{w_+} < \frac{5\alpha^2(1+\alpha)}{4(1-\alpha)} < 3\alpha^2 < \frac{1}{5}. \tag{87}$$

It is easy to see, using Theorem 10.3 that, as soon

$$\|x - w\|_w < \frac{1}{4}, \tag{88}$$

the sequence of points generated by using the new steps will converge quadratically to the exact analytic center. In practical terms, if the parameter θ defining the θ - analytic center is of order 10^{-10}, then we need only 6 such steps to get to a θ - analytic center.

Let (x^c, y^c, s^c) denote the exact analytic center. *Lemma 5.3* gives a connection between the values of $f(x^c)$ and $f^*(s^c)$:

$$f(x^c) + f^*(s^c) = -\vartheta_f.$$

Now let's analyze $\langle x^c, c\rangle$:

$$\langle x^c, c\rangle = \langle x^c, A^* y^c + s^c\rangle = \langle x^c, s^c\rangle = \langle x^c, -g(x^c)\rangle = \vartheta_f.$$

It follows that the primal-dual potential functional $\Phi(x, s)$ has value zero at the exact analytic center. The following upper bound on the number of Newton steps required to obtain a new θ - analytic center after the addition of the cuts is then a consequence of (63) and *Theorems 10.2* and *10.3*. By taking arbitrarily $\theta = 0.9$ and $\zeta = 0.9$, we can choose the step length $\alpha = 0.09$.

THEOREM 10.4 *After the addition of new cuts with a barrier functional with complexity value ϑ_{f_2}, a new θ - analytic center can be obtained in $O(\vartheta_{f_2}\ln(\vartheta_{f_2}))$ Newton steps.*

11. An upper bound on the dual barrier functional In this section we will derive an upper bound on the value of the dual potential f_i^* evaluated at the analytic center of the set Ω_i. This bound together with the fact that the values of the potential functionals keep increasing as the algorithm proceeds will help us prove that the algorithm will eventually stop with a solution.

Let (x^k, y^k, s^k) be the exact analytic center of Ω_k (the outer-approximation set of Γ after k iterations). This analytic center corresponds to the self-concordant barrier functional $f := \bigoplus_{i=0}^{k} f_i$ and the cone $K := \bigoplus_{i=0}^{k} K_i$ that is in the space $X = \bigoplus_{i=0}^{k} X_i$. Ω_k is described by the operator $A := \bigoplus_{i=0}^{k} A_i$, and the vector $c := \bigoplus_{i=0}^{k} c_i$. Our initial assumption that Γ contains a closed ball of radius ε implies that:

$$\mathcal{M} := \{y \in Y : y \in \Omega_k, B_Y(y,\varepsilon) \subset \Omega_k\} \neq \emptyset.$$

Because (x^k, y^k, s^k) is the analytic center of Ω_k, $s^k = c - A^*y^k$ is the minimizer of f^* over the set of all feasible points. Then,

$$f^*(s^k) \le f^*(s), \forall s \in \mathcal{M}_s := \{s : s = c - A^*y \text{ with } y \in \mathcal{M}\}.$$

LEMMA 11.1 *Let s be an arbitrary point in the set $\mathcal{M}_s$, with $s_i \in K_i$ the corresponding components. Then the distance (measured using the local inner product) from s_i to the boundary of the cone K_i, for $i \ge 1$, satisfies:*

$$d(s_i, \partial K_i) \ge \varepsilon\sqrt{\lambda_{\min}(A_i^*A_i)}. \tag{89}$$

Here $\lambda_{\min}(A_i^*A_i)$ is the minimum eigenvalue of $A_i^*A_i$. For the initial case $i = 0$:

$$d(s_0, \partial K_0) \ge \varepsilon. \tag{90}$$

PROOF. Let $s \in \mathcal{M}_s$ with the corresponding $y \in \mathcal{M}$ ($s = c - A^*y$). So:

$$y + \varepsilon u \in \Omega_k, \forall u \in Y, \|u\|_Y = 1. \tag{91}$$

The point s is strictly interior to the cone K. This implies that each of its components s_i is strictly interior to its corresponding cone K_i.

Then

$$\exists s_\varepsilon \in K_i \text{ such that } A_i^*(y + \varepsilon u) + s_\varepsilon = c_i. \tag{92}$$

At the same time:

$$s_i := c_i - A_i^*y \in K_i.$$

Using the last two relations we conclude that:

$$s_\varepsilon = s_i - \varepsilon A_i^*u \text{ is feasible }, \forall u \in Y, ||u||_Y = 1.$$

Our goal is to get an estimate for the distance between s_i and the boundary of K_i. Two cases arise, one for $i = 0$ and one for $i \ge 1$. The difference between these two cases is that A_i is injective only for $i \ge 1$. However, for $i = 0$ the operator A_0 is the $\oplus$ - sum of two bijective operators I and $-I$. So, this case can be treated the same way as the general case if we are using the components of A_0.

Now let's consider the case $i \ge 1$. Let v be a vector parallel to the direction which projects s_i onto ∂K_i. The operator A_i^* is surjective so there exists a vector $u \in Y$, with $\|u\|_Y = 1$ such that A_i^*u is parallel to v (for the case when $\dim(X_i) = 1$, this means $A_i^*u \neq 0$). We observe here that we can take u

to be a vector in the range of A_i (because any component of u from $\mathrm{Ker}(A_i^*)$ will have no contribution to $A_i^* u$). The size of $A_i^* u$ gives a lower bound for the distance from s_i to ∂K_i.

A lower bound for the size of $\|A_i^* u\|$ is given by the solution to the next problem:

$$\begin{array}{rl} \min & \|A_i^* u\| \\ \text{such that} & u \in \mathrm{Range}(A_i), \\ & \|u\|_Y = 1. \end{array}$$

We can reformulate this problem as:

$$\begin{array}{rl} \min & \|A_i^* A_i w\| \\ \text{such that} & \|A_i w\|_Y = 1, \\ & w \in X_i. \end{array}$$

The operator $A_i^* A_i : X_i \to X_i$ is positive definite (because A_i is injective hence $Ker(A_i) = \{0\}$).

Let $\{w_1, w_2, \ldots, w_{p_i}\}$ be an orthogonal basis formed by eigenvectors of $A_i^* A_i$ with the corresponding eigenvalues λ_i. Any vector $w \in X_i$ can be written as:

$$w = \sum_{j=1}^{p_i} \alpha_j w_j.$$

Using this decomposition:

$$\|A_i w\|_Y^2 = \sum_{j=1}^{p_i} \alpha_j^2 \lambda_j = 1 \tag{93}$$

and

$$\|A_i^* A_i w\|^2 = \sum_{j=1}^{p_i} \alpha_j^2 \lambda_j^2. \tag{94}$$

The equalities (93) and (94) imply:

$$\|A_i^* A_i w\|^2 = \sum_{j=1}^{p_i} \alpha_j^2 \lambda_j^2 \geq \lambda_{\min} \sum_{j=1}^{p_i} \alpha_j^2 \lambda_j = \lambda_{\min}(A_i^* A_i).$$

Now we can conclude that:

$$\min\{\|A_i^* u\| : \|u\|_Y = 1 \text{ and } u \in \mathrm{Range}(A_i)\} \geq \sqrt{\lambda_{\min}(A_i^* A_i)}.$$

So, the distance from s_i to the boundary of the cone K_i is greater than or equal to $\varepsilon \sqrt{\lambda_{\min}(A_i^* A_i)}$.
□

Next we will analyze the implications of the assumption we made that $f_i(\sigma_i) = 0$ where σ_i is a vector of norm $\sqrt{p_i}$ described by $\sigma_i = \sqrt{\frac{p_i}{\vartheta_i}} e_i$, e_i being the vector in X_i that induces the scaled inner product.

Lemma 11.2 *Let $\sigma_i \in \partial B(0, \sqrt{\dim(X_i)}) \cap K_i$ be the point where $f_i(\sigma_i) = 0$. Then $f_i^*(\sigma_i) = \vartheta_i(\ln \frac{\vartheta_i}{p_i} - 1)$, for all $i \geq 0$ (we take here $p_0 = 2m$).*

Proof. If we use *Lemma 3.3* together with $f_i(\sigma_i) = 0$:

$$f_i^*(\sigma_i) = f_i(\sigma_i) - \vartheta_i - 2 f_i(e_i) = -\vartheta_i - 2 f_i(e_i).$$

The functional f_i is logarithmically homogeneous and $e_i = \sqrt{\frac{\vartheta_i}{p_i}} \sigma_i$. So

$$f_i(e_i) = f_i(\sigma_i) - \frac{\vartheta_i}{2} \ln \frac{\vartheta_i}{p_i} = -\frac{\vartheta_i}{2} \ln \frac{\vartheta_i}{p_i}.$$

The conclusion follows immediately. □

LEMMA 11.3 *At any instance k of the algorithm described by the Hilbert space X with the corresponding cone K and barrier functional f there exists a point $x \in \partial B(0, \sqrt{\dim(X)}) \cap K$ such that $f(x) = 0$.*

PROOF. Let $x \in X$ be the vector with components $x_i = \sigma_i$, for $i \geq 0$. Clearly, $x \in K$. Also $f(x) = f_0(\sigma_0) + f_1(\sigma_1) + \ldots + f_k(\sigma_k)$. Then, immediately we can see that $f(x) = 0$.

Because $\|x\|^2 = \sum_{i=0}^{k} \|\sigma_i\|_i^2 = \sum_{i=0}^{k} \dim(X_i)$, it follows that $x \in \partial B(0, \sqrt{\dim(X)})$. □

Now we can prove the main result of this section:

THEOREM 11.1 *At any instance k of the algorithm described by the space X, the cone K and the functional f (where $X := X_0 \oplus X_1 \oplus \ldots \oplus X_k$, $K := K_0 \oplus K_1 \oplus \ldots \oplus K_k$ and $f^*(s) = \sum_{i=0}^{k} f_i^*(s_i), s_i \in K_i$), for all $\bar{s} \in \mathcal{M}_s$,*

$$f^*(\bar{s}) \leq \sum_{i=0}^{k} \vartheta_{f_i} \ln \frac{\vartheta_{f_i}}{\varepsilon_i}$$

where $\varepsilon_i = \varepsilon\sqrt{\lambda_{\min}(A_i^* A_i)}$ for $i \geq 1$ and $\varepsilon_0 = \varepsilon$. In particular, if s_{AC} is the analytic center,

$$f^*(s_{AC}) \leq \sum_{i=0}^{k} \vartheta_{f_i} \ln \frac{\vartheta_{f_i}}{\varepsilon_i}.$$

PROOF. Let $\bar{s}$ be a point in K such that the distance from $\bar{y}$ to the boundary of Ω_k is greater than or equal to ε (i.e. $B_Y(\bar{y}, \varepsilon) \subset \Omega_k$). We have:

$$f^*(\bar{s}) = \sum_{i=0}^{k} f_i^*(\bar{s}_i),$$

where $\bar{s}_i$ are the components of $\bar{s}$ from K_i, $\bar{s}_i \in K_i$.

Using *Lemma 11.1* we get $B_i(\bar{s}_i, \varepsilon_i) \subset K_i$.

For each f_i we know that there exists a point $\sigma_i \in K_i \cap \partial B_i(0, \sqrt{p_i})$ such that $f_i(\sigma_i) = 0$. It is easy to see that the point $\frac{\varepsilon_i}{\sqrt{p_i}}\sigma_i \in K_i \cap \partial B_i(0, \varepsilon_i)$, so $\bar{s}_i - \frac{\varepsilon_i}{\sqrt{p_i}}\sigma_i \in K_i$. Using *Lemma 3.1*, *Lemma 11.2* and the fact that the functional f_i^* is logarithmically homogeneous we have:

$$f_i^*(\bar{s}_i) \leq f_i^*(\frac{\varepsilon_i}{\sqrt{p_i}}\sigma_i) = f_i^*(\sigma_i) - \vartheta_{f_i} \ln \frac{\varepsilon_i}{\sqrt{p_i}} = \vartheta_{f_i} \ln \frac{\vartheta_{f_i}}{\varepsilon_i\sqrt{p_i}} - \vartheta_{f_i} \leq \vartheta_{f_i} \ln \frac{\vartheta_{f_i}}{\varepsilon_i}.$$

So

$$f^*(\bar{s}) \leq \sum_{i=0}^{k} \vartheta_{f_i} \ln \frac{\vartheta_{f_i}}{\varepsilon_i}.$$

The last statement of the theorem is immediate because s_{AC} is the analytic center, hence it minimizes f^* over $\mathcal{M}_s$. □

COROLLARY 11.1 *Let $\Lambda := \min_{i=1,\ldots,k} \sqrt{\lambda_{\min}(A_i^* A_i)}$. Then:*

$$f^*(s_{AC}) \leq \sum_{i=1}^{k} \vartheta_{f_i} \ln \frac{\vartheta_{f_i}}{\varepsilon\Lambda} + \vartheta_{f_0} \ln \frac{\vartheta_{f_0}}{\varepsilon}. \tag{95}$$

12. Complexity Analysis In this section we will derive an upper bound for the number of cuts that may be added to the problem before we are guaranteed to have a solution.

First we start by getting a lower bound for the minimum eigenvalue of the hessian of any potential functional evaluated at any feasible point.

Let $\bar{s} \in \text{int}(K)$ be any strictly feasible point for the k-th iteration of the algorithm. At this stage, the dual potential is given by

$$f^*(s) := f_0^*(s_0) + f_1^*(s_1) + f_2^*(s_2) + \ldots + f_k^*(s_k)$$

where $s = \bigoplus_{i=0}^{k} s_i$, with $s_i \in K_i$, $i = 0, \ldots, k$.

$H(\bar{s})$, the hessian of the barrier functional f^*, has a block diagonal matrix representation, each block corresponding to a hessian $H_i(\bar{s}_i)$. Because of this structure, the minimum eigenvalue of $H(\bar{s})$ is equal to the minimum of all eigenvalues of $H_i(\bar{s}_i)$, $i = 0, \ldots, k$.

Now let's look at the hessian $H_i(\bar{s}_i)$, $\bar{s}_i \in \text{int}(K_i)$. The norm used is the one induced by a vector $e_i \in K_i$. In this norm $\|e_i\| = \sqrt{\vartheta_{f_i}}$ (see (16)). Moreover, the distance (measured in the norm induced by e_i) from e_i to the boundary of the cone K_i is greater than or equal to 1. Let

$$d = \max\{\|z\| : z \in \partial K_i \cap (\bar{s}_i - K_i)\},$$

$\bar{s}_i$ is strictly interior to K_i so $d \neq 0$. We define $\bar{s}_d := \frac{1}{d}\bar{s}_i$.

The next lemma will give a description for the position of $\bar{s}_d$ in the cone K_i.

Lemma 12.1 $\bar{s}_d \in \mathcal{K} := \{\mathit{int}(K_i) \cap (e_i - K_i)\}$.

Proof. Suppose $\bar{s}_d \notin \mathcal{K}$. Then, because the origin is on the boundary of the convex set $\mathcal{K}$, the line containing both $\bar{s}_d$ and the origin intersects the boundary of $e_i - K_i$ in a unique point s_e, with $\|s_e\| < \|\bar{s}_d\|$. Let $\mathcal{P}$ be the plane determined by e_i and $\bar{s}_d$ together with the origin. Then $\mathcal{P} \cap K_i = \{OA, OB\}$, with OA and OB being two rays of the cone K_i (see *Fig.* 2). Take OA and OB such that e_i is in the angle determined by OA and $O\bar{s}_d$. Next:

$$\begin{aligned} C_1 &= OA \cap \partial(\bar{s}_d - K_i), \\ C_2 &= OB \cap \partial(\bar{s}_d - K_i), \\ D_1 &= OA \cap \partial(e_i - K_i), \\ D_2 &= OB \cap \partial(e_i - K_i). \end{aligned}$$

With these notations we have:

$$1 \geq \|OC_2\| > \|OD_2\| = \|D_1 e_i\| \geq d(e_i, AO) \geq 1. \tag{96}$$

So we arrived at a contradiction. This means that $\bar{s}_d \in \mathcal{K}$.

□

Note here that

$$\|\bar{s}_d\|^2 = \|OC_1\|^2 + \|OC_2\|^2 + 2\langle OC_1, OC_2\rangle \geq \|OC_1\|^2. \tag{97}$$

This inequality holds for any point $C_1 \in K_i \cap \partial(\bar{s}_d - K_i)$. Hence $\|\bar{s}_d\| \geq 1$. This implies that $\|\bar{s}_i\| \geq d$.

Now, as already proved in *Lemma 4.2*, any point $z \in \mathcal{K}$ has the property that $|z - e_i| < 1$. For such a point *Corollary 4.1* shows that the minimum eigenvalue of $H_i(z)$ is greater than $\frac{1}{4}$. So:

$$\lambda_{\min}(H_i(\bar{s}_d)) > \frac{1}{4}. \tag{98}$$

Next:

$$\lambda_{\min}(H_i(\bar{s}_d)) = \lambda_{\min}(H_i(\frac{1}{d}\bar{s}_i)) = d^2 \lambda_{\min}(H_i(\bar{s}_i))$$

so

$$\lambda_{\min}(H_i(\bar{s}_i)) > \frac{1}{4d^2} \geq \frac{1}{4\|\bar{s}_i\|^2}. \tag{99}$$

In order to get a lower bound for the minimum eigenvalue of $H_i(\bar{s}_i)$ we need to find an upper bound for $\|\bar{s}_i\|$.

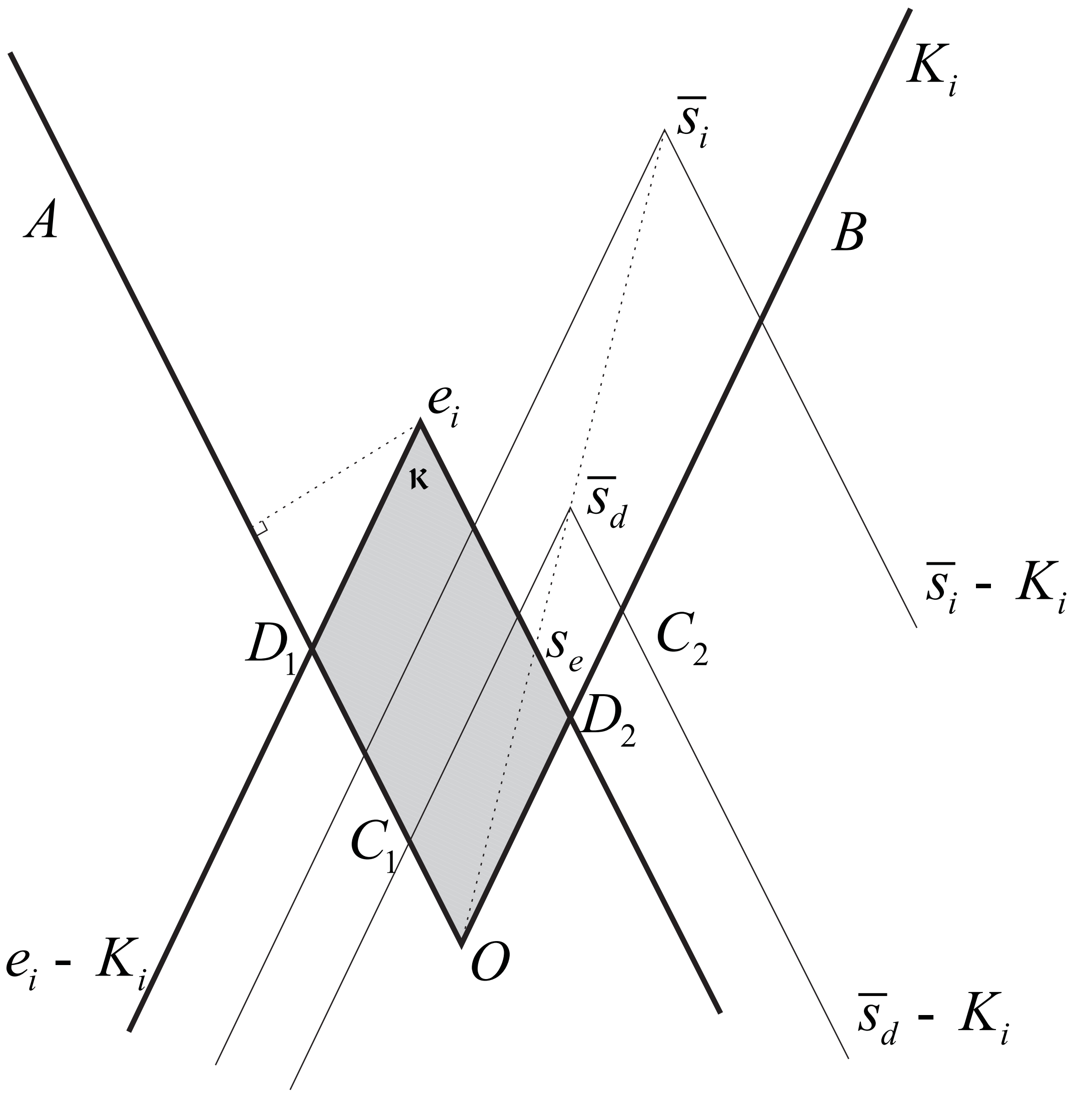

Figure 2: Position of $\bar{s}_d$ relative to $\mathcal{K}$.

Because $\bar{s}$ is feasible, we have that $A^*\bar{y} + \bar{s} = c$, for some $\bar{y} \in \Omega_k$. Here $A = A_0 \oplus A_1 \oplus A_2 \oplus \ldots \oplus A_k$ and $c = c_0 \oplus c_1 \oplus c_2 \oplus \ldots \oplus c_k$. So, componentwise, for each $i = 0, \ldots, k$, $\bar{s}_i = c_i - A_i^*\bar{y}$.

Two different cases arise: one corresponding to $i = 0$ (this is right at the beginning, before adding any cuts to the initial set Ω_0) and one corresponding to $i > 0$.

Let's look at the second case. In this case there exists at least one previous θ - analytic center. Let's denote it $(\hat{x}, \hat{y}, \hat{s})$. The cuts added through this point have the property: $A_i^*\hat{y} = c_i$. We mention here one more time that the inner product used is the local one induced by $e_i \in K_i$ for which $\|H_i(e_i)^{-1}\| = 1$. The norm of A_i in the original inner product is one so

$$\|\bar{s}_i\|_{e_i}^2 = \|A_i^*\bar{y} - A_i^*\hat{y}\|_{e_i}^2 \leq \|H_i(e_i)^{-1}\|\|A_i^*\|^2\|\bar{y} - \hat{y}\|^2 = \|\bar{y} - \hat{y}\|^2.$$

In the above sequence of inequalities, the index e_i is for the norms induced by the local inner product. If the index e_i is missing, then the inner product used is the original one.

This implies:

$$\|\bar{s}_i\|_{e_i} \leq \|\bar{y} - \hat{y}\|. \tag{100}$$

Now, both $\hat{y}$ and $\bar{y}$ are in Ω_i which is a subset of the initial set Ω_0. The next lemma will give a bound for the size of any point $y \in \Omega_0$.

LEMMA 12.2 *Let $\Omega_0 := \{y \in Y : y + s_1 = \tilde{c}_0, -y + s_2 = \tilde{c}_0$ with $\tilde{c}_0, s_1, s_2 \in \tilde{K}_0\}$. Then $\|y\| \leq \|\tilde{c}_0\|$ for any $y \in \Omega_0$.*

PROOF. The proof is rather immediate. If we take the square of the equalities defining Ω_0,

$$\|\tilde{c}_0\|^2 = \|y\|^2 + \|s_1\|^2 + 2\langle y, s_1\rangle,$$
$$\|\tilde{c}_0\|^2 = \|y\|^2 + \|s_2\|^2 - 2\langle y, s_2\rangle.$$

This implies

$$2\|\tilde{c}_0\|^2 = 2\|y\|^2 + \|s_1\|^2 + \|s_2\|^2 + 2\langle y, s_1 - s_2\rangle. \tag{101}$$

Because $\tilde{K}_0$ is a self-scaled cone, and $s_1, s_2 \in \tilde{K}_0$, their inner product is positive:

$$\langle s_1, s_2\rangle \geq 0.$$

Using this observation together with the fact that $s_1 + s_2 = 2\tilde{c}_0$, we get a bound on the sum of norms of s_1 and s_2:

$$\|s_1\|^2 + \|s_2\|^2 \leq 4\|\tilde{c}_0\|^2.$$

Using this inequality and the fact that $s_1 - s_2 = -2y$ the conclusion follows immediately:

$$\|y\| \leq \|\tilde{c}_0\|.$$

□

Using the previous lemma and (100) we finally get:

$$\|\bar{s}_i\| \leq 2\|\tilde{c}_0\| \text{ for any } i > 0.$$

For the case $i = 0$ we have that $\bar{s}_0 \in K_0 := \tilde{K}_0 \oplus \tilde{K}_0$. So we can decompose $\bar{s}_0$ in two parts: s_1 and s_2, both elements in $\tilde{K}_0$. For $\bar{s}_0$ there exists $y \in \Omega_0$ such that:

$$y + s_1 = \tilde{c}_0,$$
$$-y + s_2 = \tilde{c}_0.$$

We know that $s_1, s_2, \tilde{c}_0 \in \tilde{K}_0$, with $\tilde{K}_0$ a self-conjugate cone. Using this and the fact that $s_1 + s_2 = 2\tilde{c}_0$ it follows that

$$\|\bar{s}_0\|^2 = \|s_1\|^2 + \|s_2\|^2 \leq 4\|\tilde{c}_0\|^2. \tag{102}$$

Hence,

$$\|\bar{s}_0\| \leq 2\|\tilde{c}_0\|.$$

It follows from inequality (99) that the smallest eigenvalue of the hessian can be bounded away from zero.

LEMMA 12.3 *For any strictly feasible point* $\bar{s} \in K$*:*

$$\lambda_{\min}(H(\bar{s})) \geq \frac{1}{16\|\tilde{c}_0\|^2}. \tag{103}$$

Now we compare the value of the dual - potential functionals f^* at two consecutive analytic centers.

Let's consider Ω and $\tilde{\Omega}$ to be two consecutive outer-approximations of Γ. These two sets correspond to two instances of the algorithm described by (f, X, K, A, c) and $(\tilde{f}, \tilde{X}, \tilde{K}, \tilde{A}, \tilde{c})$. The second instance is obtained from the first one by adding central cuts through the θ - analytic center of Ω. Let these cuts be described by: $(\hat{f}, \hat{X}, \hat{K}, \hat{A}, \hat{c})$. So $\tilde{f} = f \oplus \hat{f}$, $\tilde{X} = X \oplus \hat{X}$ and so on. Let $(x^c_\theta, y^c_\theta, s^c_\theta)$ be the θ - analytic center for f. After adding the cuts right through $(x^c_\theta, y^c_\theta, s^c_\theta)$ a scaled step is taken to recover feasibility. Let $(x(\alpha), y(\alpha), s(\alpha))$ be the point right after this step is taken so

$$\begin{aligned} x(\alpha) &= (x^c_\theta + \alpha\Delta x) \oplus (\alpha\beta), \\ y(\alpha) &= y^c_\theta + \alpha\Delta y, \\ s(\alpha) &= (s^c_\theta + \alpha\Delta s) \oplus (\alpha\gamma). \end{aligned}$$

Using all these notations we are ready to prove the following theorem, which gives a bound for the change in the barrier functional evaluated at two consecutive exact analytic centers.

THEOREM 12.1 *Let* (x^c, y^c, s^c) *and* $(\tilde{x}^c, \tilde{y}^c, \tilde{s}^c)$ *be two consecutive analytic centers for the domains* Ω *and* $\tilde{\Omega}$*. Then,*

$$\tilde{f}^*(\tilde{s}^c) \geq f^*(s^c) - \hat{f}(\alpha\beta) - \theta\zeta - \frac{1}{2}\zeta^2 - \frac{\zeta^3}{3(1-\zeta)} - \frac{\theta^3}{3(1-\theta)} - \frac{1}{2}\theta^2. \tag{104}$$

PROOF. Because $(\tilde{x}^c, \tilde{y}^c, \tilde{s}^c)$ is an exact analytic center for $\tilde{f}$, $\tilde{x}^c$ minimizes the value of $\tilde{f}(x) + \langle x, \tilde{c}\rangle$

$$\tilde{f}(\tilde{x}^c) \leq \tilde{f}(x(\alpha)) + \langle x(\alpha), \tilde{c}\rangle - \langle \tilde{x}^c, \tilde{c}\rangle. \tag{105}$$

Since $\Phi(\tilde{x}^c, \tilde{s}^c) = 0$, we can rewrite inequality (105) as:

$$\tilde{f}^*(\tilde{s}^c) \geq -f(x^c_\theta + \alpha\Delta x) - \hat{f}(\alpha\beta) - \langle x^c_\theta + \alpha\Delta x, c\rangle - \langle \alpha\beta, \hat{c}\rangle.$$

We can use now the bound on $f(x^c_\theta + \alpha\Delta x)$ given by the inequality (61). Before doing this let's notice that:

$$\begin{aligned} \langle \hat{c}, \beta\rangle &= \left\langle \hat{A}^* y^c_\theta, \beta\right\rangle = \left\langle y^c_\theta, \hat{A}\beta\right\rangle = -\langle y^c_\theta, A\Delta x\rangle = \\ &= -\langle A^* y^c_\theta, \Delta x\rangle = \langle s^c_\theta, \Delta x\rangle - \langle c, \Delta x\rangle. \end{aligned}$$

So

$$\tilde{f}^*(\tilde{s}^c) \geq -f(x^c_\theta) - \hat{f}(\alpha\beta) - \langle x^c_\theta, c\rangle - \theta\zeta - \frac{1}{2}\zeta^2 - \frac{\zeta^3}{3(1-\zeta)}.$$

In order to get the desired result we have to use *Theorem 9.1* and use the fact that $f^*(s^c_\theta) \geq f^*(s^c)$ (this is because (x^c, y^c, s^c) is an exact analytic center). □

The step required to move the point back in the feasible region after the cuts are added depends upon the vector β. This vector is the solution to the minimization problem (42). So, using the fact that $\langle \beta, V\beta\rangle = 1$ (from equation (45)),

$$\hat{f}(\beta) \leq \hat{f}(\beta') + \frac{\vartheta_{\hat{f}}}{2}\langle \beta', V\beta'\rangle - \frac{\vartheta_{\hat{f}}}{2}, \text{ for any } \beta' \in \hat{K} \tag{106}$$

with V given by:

$$V = \hat{A}^*(AH(s_1)A^*)^{-1}\hat{A}. \tag{107}$$

Taking in account all these observations, the fact that $\hat{f}$ is logarithmically homogeneous and $\alpha < 1$, the previous theorem can be restated as:

$$\tilde{f}^*(\tilde{s}^c) \geq f^*(s^c) - \hat{f}(\beta') - \frac{\vartheta_{\hat{f}}}{2}\langle \beta', V\beta' \rangle + \frac{\vartheta_{\hat{f}}}{2} + \vartheta_{\hat{f}} \ln \alpha - \mathcal{F}(\theta, \zeta) \tag{108}$$

with

$$\mathcal{F}(\theta, \zeta) = \theta\zeta + \frac{1}{2}\zeta^2 + \frac{\zeta^3}{3(1-\zeta)} + \frac{\theta^3}{3(1-\theta)} + \frac{1}{2}\theta^2 \tag{109}$$

for any $\beta' \in \hat{K}$.

Now we are ready to get an estimate for the number of cuts required to be added in order to find an interior point in Γ. Before this we will reintroduce some notation. Let $(X_i, \langle \cdot, \cdot \rangle_i)$, K_i, A_i and f_i, $i = 0, \dots, k$, be the elements that describe the initial instance of the algorithm and the cuts that are added during the first k - iterations of the algorithm. Let $\bar{X}_i = \overset{i}{\underset{j=0}{\oplus}} X_j$, $\bar{K}_i = \overset{i}{\underset{j=0}{\oplus}} K_j$, $\bar{A}_i = \overset{i}{\underset{j=0}{\oplus}} A_j$, $\bar{f}_i = \overset{i}{\underset{j=0}{\oplus}} f_j$ be the elements that describe the instance of the algorithm after adding the i-th cut. Let $\bar{s}_i$ and s_i^θ be the exact analytic center and a θ - analytic center of the domain Ω_i respectively.

After i iterations of the algorithm, using formula (108) we get:

$$\bar{f}_i^*(\bar{s}_i) \geq \bar{f}_{i-1}^*(\bar{s}_{i-1}) - f_i(\beta_i') - \frac{\vartheta_{f_i}}{2}(\langle \beta_i', V_i\beta_i' \rangle_i - 1) + \vartheta_{f_i} \ln \alpha - \mathcal{F}(\theta, \zeta), \tag{110}$$

where $V_i = A_i^*(\bar{A}_{i-1}\bar{H}_{i-1}(\bar{s}_{i-1})\bar{A}_{i-1}^*)^{-1}A_i$, β_i' is any point in the interior of K_i and $\mathcal{F}(\theta, \zeta)$ is given in (109).

One of the assumptions we made about the functionals f_i was that, for each of them, there exists a point $\sigma_i \in K_i$ ($\sigma_i := \sqrt{\frac{p_i}{\vartheta_i}} e_i$)with norm equal to $\sqrt{p_i}$ such that $f_i(\sigma_i) = 0$.

Notice here that unlike β_i (the exact solution for problem (42)) for which both β_i and $V_i\beta_i$ have to be in K_i, the only requirement for β_i' is to be an element from K_i. This gives us more choices for picking a suitable vector.

Now we can choose β_i' to be:

$$\beta_i' = \frac{e_i}{\sqrt{\langle e_i, V_i e_i \rangle_i}}. \tag{111}$$

Clearly $\langle \beta_i', V_i\beta_i' \rangle_i - 1 = 0$. Moreover, using the fact that f_i is logarithmically homogeneous:

$$f_i(\beta_i') = f_i(\sigma_i) + \frac{\vartheta_{f_i}}{2} \ln\langle \sigma_i, V_i\sigma_i \rangle_i = \frac{\vartheta_{f_i}}{2} \ln\langle \sigma_i, V_i\sigma_i \rangle_i. \tag{112}$$

The inequality (110) can be further simplified to:

$$\bar{f}_i^*(\bar{s}_i) \geq \bar{f}_{i-1}^*(\bar{s}_{i-1}) - \frac{\vartheta_{f_i}}{2} \ln\langle \sigma_i, V_i\sigma_i \rangle_i + \vartheta_{f_i} \ln \alpha - \mathcal{F}(\theta, \zeta). \tag{113}$$

Now, in each space X_i we can choose an orthonormal basis $\{e_j^i\}_{j=1,\dots,p_i}$ such that

$$\sigma_i = \sum_{j=1}^{p_i} e_j^i, \tag{114}$$

To do this it is enough to pick an orthonormal basis and then rotate it until $\sum_{j=1}^{p_i} e_j^i$ overlaps with σ_i. It is clear that $\|\sigma_i\|_i = \sqrt{p_i}$, for any $i \geq 0$. Let's look now at $\langle \sigma_i, V_i\sigma_i\rangle_i$:

$$\langle \sigma_i, V_i\sigma_i\rangle_i = \sum_{j=1}^{p_i}\sum_{l=1}^{p_i}\langle e_j^i, V_i e_l^i\rangle_i = \sum_{j=1}^{p_i}\sum_{l=1}^{p_i}\langle e_j^i, e_l^i\rangle_{V_i} \leq \sum_{j=1}^{p_i}\sum_{l=1}^{p_i}\|e_j^i\|_{V_i}\|e_l^i\|_{V_i}. \tag{115}$$

Using the mean inequality:

$$\langle \sigma_i, V_i\sigma_i\rangle_i \leq \sum_{j=1}^{p_i}\sum_{l=1}^{p_i}\frac{\|e_j^i\|_{V_i}^2 + \|e_l^i\|_{V_i}^2}{2} = p_i\sum_{j=1}^{p_i}\|e_j^i\|_{V_i}^2. \tag{116}$$

So:

$$\bar{f}_i^*(\bar{s}_i) \geq \bar{f}_{i-1}^*(\bar{s}_{i-1}) - \frac{\vartheta_{f_i}}{2}\ln(p_i\sum_{j=1}^{p_i}\|e_j^i\|_{V_i}^2) + \vartheta_{f_i}\ln\alpha - \mathcal{F}(\theta,\zeta). \tag{117}$$

This inequality gives a relationship between the dual potential functionals evaluated at two consecutive exact analytic centers. A direct relationship between the potential at the initial analytic center $\bar{s}_0$ and the potential at the k-th analytic center $\bar{s}_k$ can be easily obtained by taking the sum of the previous inequalities from $i = 1$ to $i = k$:

$$\bar{f}_k^*(\bar{s}_k) \geq \bar{f}_0^*(\bar{s}_0) - \sum_{i=1}^{k}(\frac{\vartheta_{f_i}}{2}\ln(p_i\sum_{j=1}^{p_i}\|e_j^i\|_{V_i}^2)) + \ln\alpha\sum_{i=1}^{k}\vartheta_{f_i} - k\mathcal{F}(\theta,\zeta). \tag{118}$$

Let $P = \max_{i=1,\dots,k} p_i$ (i.e. at each stage we do not add more than P cuts). Then:

$$\bar{f}_k^*(\bar{s}_k) \geq \bar{f}_0^*(\bar{s}_0) - \frac{1}{2}(\ln P - \ln\alpha^2)\sum_{i=1}^{k}\vartheta_{f_i} - \sum_{i=1}^{k}(\frac{\vartheta_{f_i}}{2}\ln\sum_{j=1}^{p_i}\|e_j^i\|_{V_i}^2) - k\mathcal{F}(\theta,\zeta). \tag{119}$$

We can simplify this inequality by using the concavity of the logarithm function together with the fact that $\vartheta_{f_i} \geq 1$:

$$\bar{f}_k^*(\bar{s}_k) \geq \bar{f}_0^*(\bar{s}_0) - \frac{1}{2}(\ln P - \ln\alpha^2)\sum_{i=1}^{k}\vartheta_{f_i} - \frac{\sum_{l=1}^{k}\vartheta_{f_l}}{2}\ln\frac{\sum_{i=1}^{k}(\vartheta_{f_i}\sum_{j=1}^{p_i}\|e_j^i\|_{V_i}^2)}{\sum_{t=1}^{k}\vartheta_{f_t}} - k\mathcal{F}(\theta,\zeta).$$

For any i: $\vartheta_{f_i} \geq 1$. So $\sum_{i=1}^{k} p_i \leq P\sum_{i=1}^{k}\vartheta_{f_i}$. Let $\Theta := \max_{i=1,\dots,k}\vartheta_{f_i}$. Then:

$$\bar{f}_k^*(\bar{s}_k) \geq \bar{f}_0^*(\bar{s}_0) - \frac{1}{2}(\ln P - \ln\alpha^2)\sum_{i=1}^{k}\vartheta_{f_i} - \frac{\sum_{l=1}^{k}\vartheta_{f_l}}{2}\ln P\Theta\frac{\sum_{i=1}^{k}(\sum_{j=1}^{p_i}\|e_j^i\|_{V_i}^2)}{\sum_{t=1}^{k}p_t} - k\mathcal{F}(\theta,\zeta).$$

So:

$$\bar{f}_k^*(\bar{s}_k) \geq \bar{f}_0^*(\bar{s}_0) - \frac{\sum_{l=1}^{k}\vartheta_{f_l}}{2}(2\ln P + \ln\frac{\Theta}{\alpha^2} + \ln\frac{\sum_{i=1}^{k}(\sum_{j=1}^{p_i}\|e_j^i\|_{V_i}^2)}{\sum_{t=1}^{k}p_t}) - k\mathcal{F}(\theta,\zeta). \tag{120}$$

By taking arbitrarily $\theta \leq 0.9$ and $\zeta \leq 0.9$, the value of $\mathcal{F}(\theta,\zeta)$ can be made smaller than 6.5. Then, for this choice of θ and ζ, $k\mathcal{F}(\theta,\zeta) \leq 7k \leq 7\sum_{l=1}^{k}\vartheta_{f_l}$.

So:

$$\bar{f}_k^*(\bar{s}_k) \geq \bar{f}_0^*(\bar{s}_0) - \frac{\sum_{l=1}^{k}\vartheta_{f_l}}{2}(2\ln P + 14 + \ln\frac{\Theta}{\alpha^2} + \ln\frac{\sum_{i=1}^{k}(\sum_{j=1}^{p_i}\|e_j^i\|_{V_i}^2)}{\sum_{t=1}^{k}p_t}). \tag{121}$$

Now we have to get an estimate for: $\sum_{i=1}^{k}(\sum_{j=1}^{p_i}\|e_j^i\|_{V_i}^2)$. We will take the same approach used by Ye in [17]. Because of the specifics of our problem, we will present here the entire scheme.

Let $C_0 := 16\|\tilde{c}_0\|^2$. Each term $\|e_j^i\|_{V_i}^2$ can be bounded from above if we use *Lemma 12.3*:

$$\|e_j^i\|_{V_i}^2 \leq C_0\langle e_j^i, A_i^*(\bar{A}_{i-1}\bar{A}_{i-1}^*)^{-1}A_ie_j^i\rangle_i.$$

Let $\mathcal{A}_i$ be the matrix representation of the operator A_i with respect to the basis $\{e_j^i\}$, $j=1,\ldots,p_i$ for $i=1,\ldots,k$. Let $\mathcal{A}_0$ be the matrix representation for A_0 with respect to an orthonormal basis $\{e_j^0\}, j=1,\ldots,2m$ of X_0. The corresponding matrix representation for $\bar{A}_i$ is given by the $m\times(2m+\sum_{i=1}^{k}p_i)$ block matrix $\bar{\mathcal{A}}_i = [\mathcal{A}_0,\mathcal{A}_1,\ldots,\mathcal{A}_k]$. Let a_j^i be the j-th column of $\mathcal{A}_i$. Using this notation we have:

$$\sum_{i=1}^{k}(\sum_{j=1}^{p_i}\|e_j^i\|_{V_i}^2) \leq C_0\sum_{i=1}^{k}\sum_{j=1}^{p_i}({a_j^i}^T(\sum_{l=0}^{i-1}\mathcal{A}_l\mathcal{A}_l^T)^{-1}a_j^i).$$

Let $B_0 := \mathcal{A}_0\mathcal{A}_0^T$ and $B_{i+1} = B_i + \mathcal{A}_{i+1}\mathcal{A}_{i+1}^T$, for $i \geq 0$.

With this notation:

$$\sum_{i=1}^{k}(\sum_{j=1}^{p_i}\|e_j^i\|_{V_i}^2) \leq C_0\sum_{i=1}^{k}\sum_{j=1}^{p_i}({a_j^i}^T(B_{i-1})^{-1}a_j^i). \tag{122}$$

Lemma 12.4 *Let* $C_B = \frac{1}{1+(P+2)\|B_0^{-1}\|}$. *Then*

$$\sum_{i=1}^{k}(\sum_{j=1}^{p_i}\|e_j^i\|_{V_i}^2) \leq \frac{4C_0}{3C_B}(2m\ln\frac{tr(B_0)+\sum_{i=1}^{k}p_i}{2m} - \ln(\det B_0)). \tag{123}$$

Proof. Notice that: $\mathcal{A}_i\mathcal{A}_i^T = \sum_{j=1}^{p_i}a_j^i{a_j^i}^T$. If we denote:

$$\omega^2 = {a_1^{i+1}}^T(B_i + \sum_{j=2}^{p_{i+1}}a_j^{i+1}{a_j^{i+1}}^T)^{-1}a_1^{i+1}$$

then, as shown by Ye in [17],

$$\det B_{i+1} = \det(B_i + \sum_{j=1}^{p_{i+1}}a_j^{i+1}{a_j^{i+1}}^T) = (1+\omega^2)\det(B_i + \sum_{j=2}^{p_{i+1}}a_j^{i+1}{a_j^{i+1}}^T).$$

We know from the initial assumptions that $\|A_i\| = 1$ for all $i > 0$ so:

$$\|a_j^{i+1}\| \leq \|\mathcal{A}_{i+1}\| = \|A_{i+1}\| = 1.$$

We can rewrite ω^2 as:

$$\omega^2 = {a_1^{i+1}}^T B_i^{-\frac{1}{2}} (I + \sum_{j=2}^{p_{i+1}} B_i^{-\frac{1}{2}} a_j^{i+1} {a_j^{i+1}}^T B_i^{-\frac{1}{2}})^{-1} B_i^{-\frac{1}{2}} a_1^{i+1}.$$

Next, for any y with $\|y\| = 1$:

$$\begin{aligned} y^T(I + \sum_{j=2}^{p_{i+1}} B_i^{-\frac{1}{2}} a_j^{i+1} {a_j^{i+1}}^T B_i^{-\frac{1}{2}})y &= 1 + \sum_{j=2}^{p_{i+1}} (y^T B_i^{-\frac{1}{2}} a_j^{i+1})^2 \\ &\leq 1 + \sum_{j=2}^{p_{i+1}} \|B_i^{-\frac{1}{2}} a_j^{i+1}\|^2 \\ &= 1 + \sum_{j=2}^{p_{i+1}} {a_j^{i+1}}^T B_i^{-1} a_j^{i+1} \\ &\leq 1 + \sum_{j=2}^{p_{i+1}} {a_j^{i+1}}^T B_0^{-1} a_j^{i+1} \\ &\leq 1 + \|B_0^{-1}\| \sum_{j=2}^{p_{i+1}} \|a_j^{i+1}\|^2 \\ &\leq 1 + (p_{i+1} - 1)\|B_0^{-1}\| \\ &\leq 1 + (P+2)\|B_0^{-1}\|. \end{aligned}$$

where the last line is a conservative overestimate chosen to simplify the following exposition. So the maximum eigenvalue of $I + \sum_{j=2}^{p_{i+1}} B_i^{-\frac{1}{2}} a_j^{i+1} {a_j^{i+1}}^T B_i^{-\frac{1}{2}}$ is less than or equal to $1+(P+2)\|B_0^{-1}\|$. This allows us to write:

$$\omega^2 \geq \frac{1}{1 + (P+2)\|B_0^{-1}\|} {a_1^{i+1}}^T B_i^{-1} a_1^{i+1}. \tag{124}$$

Hence:

$$\det B_{i+1} \geq (1 + \frac{{a_1^{i+1}}^T B_i^{-1} a_1^{i+1}}{1 + (P+2)\|B_0^{-1}\|}) \det(B_i + \sum_{j=2}^{p_{i+1}} a_j^{i+1} {a_j^{i+1}}^T). \tag{125}$$

Repeating this process inductively, we finally get

$$\ln\det B_{i+1} \geq \sum_{j=1}^{p_{i+1}} \ln(1 + \frac{{a_j^{i+1}}^T B_i^{-1} a_j^{i+1}}{1 + (P+2)\|B_0^{-1}\|}) + \ln\det(B_i).$$

or, using C_B:

$$\ln\det B_{i+1} \geq \sum_{j=1}^{p_{i+1}} \ln(1 + C_B {a_j^{i+1}}^T B_i^{-1} a_j^{i+1}) + \ln\det(B_i). \tag{126}$$

We know that $B_i - B_0$ is a positive semidefinite matrix for any $i \geq 1$. So:

$${a_j^{i+1}}^T B_i^{-1} a_j^{i+1} \leq {a_j^{i+1}}^T B_0^{-1} a_j^{i+1} \leq \|B_0^{-1}\| \|a_j^{i+1}\|^2 \leq \|B_0^{-1}\|.$$

Based on this, it is clear that, for any $P \geq 1$:

$$C_B {a_j^{i+1}}^T B_i^{-1} a_j^{i+1} \leq \frac{\|B_0^{-1}\|}{1 + (P+2)\|B_0^{-1}\|} < \frac{1}{3} < 1.$$

Now, the inequality $\ln(1+x) \geq x - \frac{x^2}{2(1-x)}$ holds true for any $x \in [0,1)$. Using it and the fact that the function $1 - \frac{x}{2(1-x)}$ is decreasing, we get, for any $i = 0, \ldots, k-1$

$$\ln(1 + C_B {a_j^{i+1}}^T B_i^{-1} a_j^{i+1}) \geq C_B {a_j^{i+1}}^T B_i^{-1} a_j^{i+1} (1 - \frac{1/3}{2(1-1/3)})$$

or

$$\ln(1 + C_B {a_j^{i+1}}^T B_i^{-1} a_j^{i+1}) \geq \frac{3}{4} C_B {a_j^{i+1}}^T B_i^{-1} a_j^{i+1}.$$

So, for any $i = 0, \dots, k-1$

$$\ln \det B_{i+1} \geq \ln \det(B_i) + \frac{3}{4} C_B \sum_{j=1}^{p_{i+1}} {a_j^{i+1}}^T B_i^{-1} a_j^{i+1}.$$

After we add the inequalities corresponding to $i = 0$ to $i = k-1$, we get:

$$\ln \det B_k \geq \ln \det(B_0) + \frac{3}{4} C_B \sum_{i=1}^{k} \sum_{j=1}^{p_i} {a_j^i}^T B_{i-1}^{-1} a_j^i.$$

Now:

$$\text{tr}(B_k) = \sum_{i=0}^{k} \text{tr}(\mathcal{A}_i \mathcal{A}_i^T) = \text{tr}(B_0) + \sum_{i=1}^{k} \sum_{j=1}^{p_i} \|a_j^i\|^2 \leq \text{tr}(B_0) + \sum_{i=1}^{k} p_i.$$

Using the mean inequality (for the sum and the product of eigenvalues of B_k):

$$\ln(\det B_k) \leq m \ln \frac{\text{tr}(B_0) + \sum_{i=1}^{k} p_i}{m}.$$

The conclusion follows immediately from (122).

Note that $\ln(\det B_k)$ is well defined since B_k is positive definite being the sum of the positive definite matrix $\mathcal{A}_0 \mathcal{A}_0^T$ and positive semidefinite matrices $\mathcal{A}_i \mathcal{A}_i^T$. □

Using *Lemma 12.4* and inequality (121) we get:

$$\bar{f}_k^*(\bar{s}_k) \geq \bar{f}_0^*(\bar{s}_0)$$

$$-0.5 \sum_{l=1}^{k} \vartheta_{f_l} (\ln \frac{4 C_0 \Theta P^2}{3 C_B \alpha^2} + 14 + \ln(m \ln \frac{\text{tr}(B_0) + \sum_{i=1}^{k} p_i}{m} - \ln(\det B_0)) - \ln(\sum_{i=1}^{k} p_i)).$$

Corollary 11.1 gives an upper bound for $\bar{f}_k^*(\bar{s}_k)$:

$$\bar{f}_k^*(\bar{s}_k) \leq \sum_{i=1}^{k} \vartheta_{f_i} \ln \frac{\vartheta_{f_i}}{\varepsilon \Lambda} + \vartheta_{f_0} \ln \frac{\vartheta_{f_0}}{\varepsilon}.$$

Theorem 12.2 *The algorithm stops with a solution as soon as:*

$$(\sum_{l=1}^{k} \vartheta_{f_l})(\ln(H(m \ln \tfrac{1}{m}(tr(B_0) + \sum_{i=1}^{k} p_i) - \ln(\det B_0)) - \ln(\sum_{i=1}^{k} p_i)) \leq 2 \bar{f}_0^*(\bar{s}_0) - 2\vartheta_{f_0} \ln \tfrac{\vartheta_{f_0}}{\varepsilon}.$$

with $H = \frac{4 C_0 \Theta^3 P^2 e^{14}}{3 \varepsilon^2 \Lambda^2 C_B \alpha^2}$. *The number of cuts added is at most* $O^*(\frac{m^2 P^3 \Theta^3}{\varepsilon^2 \Lambda^2})$ *(here* O^* *means that terms of low order are ignored). Here we assumed that* $\|\tilde{c}_0\|$ *has the size of order* $\sqrt{m}$. *Also we used the fact that* $\|B_0^{-1}\|$ *has size* $O(1)$.

PROOF. This result follows directly from the previous analysis. Note here that $C_B = \frac{1}{1+(P+2)\|B_0^{-1}\|}$ has a contribution in the complexity result. □

This result is similar to the ones for linear, semidefinite, or second order cone programming. Θ and Λ are the only extra terms. The reason for this is straightforward. In the linear or semidefinite case the potential functions are separable. In general this is not necessarily the case. This explains the presence of Θ which characterizes the barrier functional as a whole. The only assumption we made on the cuts that are added was that the operators describing them have unit norms. This assumption is not critical. We use it only to keep the analysis simple. In the linear programming approach a similar assumption often made is that the matrices describing the cuts are assumed to have columns of norm one. This gives more structure to the cuts. In our general case we cannot work at the "column" level, so we had to use an overall characterization of the cuts. The parameter Λ characterizes the quality of the cuts that are generated by the oracle.

13. Conclusions and future work In this paper we proposed and analyzed an algorithm for solving feasibility problems that arise in conic programming. The approach is based on an analytic center cutting plane method. We generalized here the particular cases of linear programming, second order cone programming and semidefinite programming. Our algorithm can be easily adjusted to these particular cases.

The assumptions we made about the problem are usual ones. Although we are dealing with a general case we didn't need to impose any extra conditions on the problems. The feasibility problems have convex, closed, bounded, fully dimensional sets of interest. These sets are described by an oracle that either recognizes that a point is strictly interior to the set or returns a set of violated constraints. Multiple cuts are added centrally when the current point is infeasible. These cuts can be linear, quadratic, semidefinite or any combination of these types.

The complexity results are similar to the ones obtained for less general cases. We proved that our algorithm generates no more than $O^*(\frac{m^2P^3\Theta^3}{\varepsilon^2\Lambda^2})$ analytic centers before a solution is obtained. This result compares favorably with $O^*(\frac{m^2P^2}{\varepsilon^2})$ (obtained for the linear case) and $O(\frac{m^3P^2}{\varepsilon^2})$ (for the semidefinite case). The extra terms we have are Θ and Λ, which characterize the self-concordant functionals and the cuts that are introduced, respectively.

Results in the literature for semidefinite and SOCP problems depend on a condition number for the added cuts. Our parameter Λ can be regarded as a similar condition number, and it would be of interest to relate the various measures more closely.

Our proof for the number of Newton steps required to obtain a new θ - analytic center depended on an exact solution to problems (P_2) and (D_2), found by solving (42). It is not necessary to solve this problem to optimality; all that is required is that the new value of the primal-dual potential function be bounded by a function of size $O(\vartheta_{f_2}\ln(\vartheta_{f_2}))$. The bound on the number of cuts added in *Theorem 12.2* is independent of the solution of (42).

Open questions remain to be addressed in future work. It would be interesting to analyze how the algorithm changes if deep cuts are used (instead of central ones) or if some of them are dropped. In our analysis the operators describing the cuts had to be injective. This requirement limits the size of second order cones that can be added by the oracle.

Acknowledgments. The research of the first author was supported in part by the National Science Foundation's VIGRE Program (Grant DMS-9983646). The research of the second author was supported in part by NSF grant numbers CCR-9901822 and DMS-0317323. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] D. S. Atkinson and P. M. Vaidya, *A cutting plane algorithm for convex programming that uses analytic centers*, Mathematical Programming **69** (1995), 1–43.

[2] J.-L. Goffin and J.-P. Vial, *Multiple cuts in the analytic center cutting plane method*, SIAM Journal

on Optimization **11** (2001), no. 1, 266–288.

[3] ______, *Convex nondifferentiable optimization: a survey focussed on the analytic center cutting plane method*, Optimization Methods and Software **17** (2002), no. 5, 805–867.

[4] M. Kojima, S. Mizuno, and A. Yoshise, *A primal-dual interior point method for linear programming*, Progress in Mathematical Programming (New York) (N. Megiddo, ed.), Springer-Verlag, 1989, pp. 29–47.

[5] J. E. Mitchell, *Polynomial interior point cutting plane methods*, Optimization Methods and Software **18** (2003), no. 5, 507–534.

[6] J. E. Mitchell and S. Ramaswamy, *A long-step, cutting plane algorithm for linear and convex programming*, Annals of Operations Research **99** (2000), 95–122.

[7] J. E. Mitchell and M. J. Todd, *Solving combinatorial optimization problems using Karmarkar's algorithm*, Mathematical Programming **56** (1992), 245–284.

[8] R. D. C. Monteiro and I. Adler, *Interior path following primal-dual algorithms. Part I: Linear programming*, Mathematical Programming **44(1)** (1989), 27–41.

[9] Y. E. Nesterov and A. S. Nemirovsky, *Interior point polynomial methods in convex programming: Theory and algorithms*, SIAM Publications, SIAM, Philadelphia, USA, 1993.

[10] M. R. Oskoorouchi and J.-L. Goffin, *The analytic center cutting plane method with semidefinite cuts*, SIAM Journal on Optimization **13** (2003), no. 4, 1029–1053.

[11] ______, *An interior point cutting plane method for convex feasibility problem with second-order cone inequalities*, Tech. report, College of Business Administration, California State University San Marcos, California, April 2003.

[12] M. R. Oskoorouchi and J. E. Mitchell, *A second-order cone cutting surface method: complexity and application*, Tech. report, College of Business Administration, California State University San Marcos, California, May 2005.

[13] O. Péton and J.-P. Vial, *Multiple cuts with a homogeneous analytic center cutting plane method*, Computational Optimization and Applications **24** (2003), no. 1, 37–61.

[14] J. Renegar, *A mathematical view of interior-point methods in convex optimization*, MPS-SIAM Series on Optimization, no. 1, SIAM, Philadelphia, 2001.

[15] G. Sonnevend, *New algorithms in convex programming based on a notion of center and on rational extrapolations*, International Series of Numerical Mathematics **84** (1988), 311–327.

[16] K. C. Toh, G. Zhao, and J. Sun, *A multiple-cut analytic center cutting plane method for semidefinite feasibility problems*, SIAM Journal on Optimization **12** (2002), no. 4, 1126–1146.

[17] Y. Ye, *Complexity analysis of the analytic center cutting plane method that uses multiple cuts*, Mathematical Programming **78** (1997), 85–104.