

# On Conically Ordered Convex Programs

Shuzhong Zhang \*

May 2003; revised December 2004

## Abstract

In this paper we study a special class of convex optimization problems called *conically ordered convex programs* (COCP), where the feasible region is given as the level set of a vector-valued non-linear mapping, expressed as a nonnegative combination of convex functions. The nonnegativity of the vectors is defined using a pre-described conic ordering. The new model extends the ordinary convex programming models where the feasible sets are the level sets of convex functions, and it also extends the famous linear conic optimization models. We introduce a condition on the barrier function for the order-defining cone, termed as the *cone-consistent property*. The relationship between the cone-consistent barriers and the self-concordance barriers is investigated. We prove that if the order-defining cone admits a self-concordant and cone-consistent barrier function, and moreover, if the constraint functions are all convex quadratic then the overall composite barrier function is self-concordant. The problem is thus solvable in polynomial time, following Nesterov and Nemirovskii, by means of the path-following method. If the constraint functions are not quadratic, but *harmonically convex*, then we propose a variant of Iri-Imai type potential reduction method. In addition to the self-concordance and the cone-consistence conditions, we assume that the barrier function for the order-defining cone has the property that the image of the cone under its Hessian matrix is contained in its dual cone. All these conditions are satisfied by the familiar self-scaled cones. Under these conditions we show that the Iri-Imai type potential reduction algorithm converges in polynomial time. Duality issues related to this class of optimization problems are discussed as well.

**Keywords:** convex programming, conic ordering, self-concordant barriers, potential reduction.

---

\*Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong. Research supported by Hong Kong RGC Earmarked Grants CUHK4233/01E and CUHK4174/03E.

# 1 The problem formulation

Let  $\mathcal{K} \subseteq \mathfrak{R}^l$  be a solid closed convex cone. Let  $h_j \in \mathcal{K}$ ,  $j = 1, \dots, k$ , and  $h_0 \in \mathfrak{R}^l$ . Consider the following optimization problem, which we shall call *conically ordered convex program* (COCP) in this paper,

$$(P_{\mathcal{K}}) \quad \begin{aligned} &\text{minimize} && c^T x \\ &\text{subject to} && Ax = b \\ &&& h_0 - \sum_{j=1}^k f_j(x)h_j \in \mathcal{K}, \end{aligned}$$

where  $A \in \mathfrak{R}^{m \times n}$ ,  $c \in \mathfrak{R}^n$  and  $b \in \mathfrak{R}^m$ . If we use the conic ordering defined by  $\mathcal{K}$ , i.e., for  $y, z \in \mathfrak{R}^l$ ,  $y \preceq_{\mathcal{K}} z$  if and only if  $z - y \in \mathcal{K}$ , then  $(P_{\mathcal{K}})$  can be written as

$$\begin{aligned} &\text{minimize} && c^T x \\ &\text{subject to} && Ax = b \\ &&& \sum_{j=1}^k f_j(x)h_j \preceq_{\mathcal{K}} h_0. \end{aligned}$$

Note that all the results in this paper straightforwardly extend to the setting where more than one conic constraint are present, namely,

$$\begin{aligned} &\text{minimize} && c^T x \\ &\text{subject to} && Ax = b \\ &&& \sum_{j=1}^k f_j^i(x)h_j^i \preceq_{\mathcal{K}^i} h_0^i, \quad i = 1, \dots, p. \end{aligned}$$

To keep the analysis simple, however, we shall restrict ourselves to only one conic constraint in the subsequent discussion. It is evident that  $(P_{\mathcal{K}})$  is indeed a convex program as stated below.

**Lemma 1.1** *If  $f_i$ 's are all convex functions,  $i = 1, \dots, k$ , then  $(P_{\mathcal{K}})$  is a convex program.*

Since  $\mathcal{K}$  is solid, i.e.  $\mathcal{K} + (-\mathcal{K}) = \mathfrak{R}^l$ , and so any vector in  $\mathfrak{R}^l$  can be written as difference of two vectors in  $\mathcal{K}$ . Therefore, the ordinary linear conic constraint  $h_0 - \sum_{j=1}^k x_j h_j \in \mathcal{K}$ , where  $h_j \in \mathfrak{R}^l$  for all  $j$ , can be written as  $h_0 - \sum_{j=1}^k (x_j h_j^+ - x_j h_j^-) = h_0 - \sum_{j=1}^k x_j h_j^+ - \sum_{j=1}^k (-x_j) h_j^- \in \mathcal{K}$ , where  $h_j = h_j^+ - h_j^-$  and  $h_j^+, h_j^- \in \mathcal{K}$  for all  $j$ ; the latter expression is in the form of COCP. Obviously, if  $\mathcal{K} = \mathfrak{R}_+^l$  and  $h_j$ 's are unit vectors, then  $(P_{\mathcal{K}})$  is the usual convex program where the constraints are given by inequalities:  $f_j(x) \leq 0$ ,  $j = 1, \dots, k$ .

Another interesting special case of  $(P_{\mathcal{K}})$  is the following

$$(P_{\mathfrak{R}_+^l}) \quad \begin{aligned} &\text{minimize} && c^T x \\ &\text{subject to} && Ax = b \\ &&& \sum_{j=1}^k f_j(x) h_j \leq h_0 \end{aligned}$$

where  $f_j(x) = e^{\iota_j^T x + \iota_{j0}}$  with  $\iota_j \in \mathfrak{R}^n$  and  $\iota_{j0} \in \mathfrak{R}$ , and  $h_j \in \mathfrak{R}_+^l$ ,  $j = 0, 1, \dots, k$ . This problem is known as *geometric program*.

Note that in its original form, a general geometric program is given as

$$(GP) \quad \begin{aligned} &\text{minimize} && \sum_{j=1}^m c_j f_{0j}(t) \\ &\text{subject to} && \sum_{j=1}^m a_{ij} f_{ij}(t) \leq b_i, \quad i = 1, \dots, p \\ &&& f_{ij}(t) = b_i, \quad i = p + 1, \dots, p + q, \\ &&& t_1 > 0, \dots, t_n > 0 \end{aligned}$$

where  $c_j > 0$  and  $a_{ij} > 0$  and  $f_{ij}(t) = t_1^{\iota_1^{(ij)}} \cdots t_n^{\iota_n^{(ij)}}$ ,  $i = 0, 1, \dots, p + q$ ,  $j = 1, \dots, m$ . Such  $f_{ij}(t)$ 's are called *monomials*.

Using the variable transformation

$$x_1 = \log t_1, \quad x_2 = \log t_2, \quad \dots, \quad x_n = \log t_n$$

the monomial  $f_{ij}(t)$  becomes  $e^{(\iota^{(ij)})^T x}$  where  $(\iota^{(ij)})^T = [\iota_1^{(ij)}, \dots, \iota_n^{(ij)}]$ ,  $i = 0, 1, \dots, p + q$ ,  $j = 1, \dots, m$ . This way,  $(GP)$  is transformed into a COCP problem  $(P_{\mathfrak{R}_+^l})$ .

Next let us consider another special case of COCP

$$(PPSD) \quad \begin{aligned} &\text{minimize} && c^T x \\ &\text{subject to} && Ax = b \\ &&& H_0 - \sum_{j=1}^k f_j(x) H_j \succeq 0, \end{aligned}$$

where  $H_1 \succeq 0, \dots, H_k \succeq 0$  are  $l \times l$  positive semidefinite matrices, and  $f_1(x), \dots, f_k(x)$  are convex functions.

Consider the following barrier function

$$F(x) = -\log \det(H_0 - \sum_{j=1}^k f_j(x) H_j) \quad (1)$$

for the cone of  $l \times l$  positive semidefinite matrices  $\mathcal{S}_+^l$ .

A natural question to ask is: under what conditions is the barrier function  $F(x)$  defined in (1) *self-concordant*? For details on the theory of self-concordant barriers, one is referred to the excellent texts [5, 4, 7]. The notion of self-concordant barrier function was introduced by Nesterov and Nemirovskii in their seminal work [5]. For convenience, we shall include the definition below.

First, a barrier (convex) function  $F(x)$  for the cone  $\mathcal{K}$  is defined to have the property that  $F(x) < \infty$  for all  $x \in \text{int } \mathcal{K}$ , and  $F(x^k) \rightarrow \infty$  as  $x^k \rightarrow \bar{x}$  where  $\bar{x}$  is on the boundary of  $\mathcal{K}$ . Moreover, it is called *self-concordant* (Section 2.3.1 of [5]) if it further satisfies the property that

$$|\nabla^3 F(x)[u, u, u]| \leq C_1 (\nabla^2 F(x)[u, u])^{3/2} \quad (2)$$

and

$$|\nabla F(x)[u]| \leq C_2 (\nabla^2 F(x)[u, u])^{1/2} \quad (3)$$

for any  $x \in \text{int } \mathcal{K}$  and any direction  $u \in \mathfrak{R}^n$ . The positive constants  $C_1$  and  $C_2$  can be scaled so as to have only one degree of freedom. Usually  $C_1$  is set to be 2 and  $C_2$  is referred to as the *complexity parameter* of the cone  $\mathcal{K}$ . As noted by Renegar [7] (see also [12]), just like convexity, the self-concordant property is in fact a *line property*, meaning that if a function is self-concordant along every line (thus one-dimensional) restricted to its domain then the function is self-concordant in the whole domain, and vice versa.

For this purpose, let us consider a given  $x$  in the interior of the feasible domain of  $(P_{PSD})$ , and  $u \in \mathfrak{R}^n$  be any given direction. Consider the one-dimensional function

$$X(t) := H_0 - \sum_{j=1}^k f_j(x + tu) H_j$$

in the domain that  $X(t) \succ 0$ . Obviously,

$$\begin{aligned} X'(t) &= - \sum_{j=1}^k \nabla f_j(x + tu)[u] H_j = - \sum_{j=1}^k \nabla f_j(x + tu)^T u H_j, \\ X''(t) &= - \sum_{j=1}^k \nabla^2 f_j(x + tu)[u, u] H_j = - \sum_{j=1}^k u^T \nabla^2 f_j(x + tu) u H_j \leq 0, \\ X'''(t) &= - \sum_{j=1}^k \nabla^3 f_j(x + tu)[u, u, u] H_j, \end{aligned}$$

and

$$[X(t)^{-1}]' = -X(t)^{-1} X'(t) X(t)^{-1}.$$

Let

$$f(t) = -\log \det X(t).$$

Applying the chain-rule to the composite function  $f(t)$  we get

$$f'(t) = -\text{tr} \left( X(t)^{-1} X'(t) \right) \quad (4)$$

$$f''(t) = \text{tr} \left( X(t)^{-1} X'(t) X(t)^{-1} X'(t) \right) - \text{tr} \left( X(t)^{-1} X''(t) \right) \quad (5)$$

$$f'''(t) = 3\text{tr} \left( X(t)^{-1} X''(t) X(t)^{-1} X'(t) \right) - 2\text{tr} \left( X(t)^{-1} X'(t) \right)^3 - \text{tr} \left( X(t)^{-1} X'''(t) \right). \quad (6)$$

**Theorem 1.2** *If  $f_j(x)$ 's,  $j = 1, \dots, k$ , are convex and quadratic functions, then  $F(x)$  as defined in (1) is a self-concordant barrier.*

**Proof.** As discussed before, we need only to verify the self-concordant condition along a given line.

Let

$$M_1 := X(t)^{-1/2} X'(t) X(t)^{-1/2} \text{ and } M_2 := -X(t)^{-1/2} X''(t) X(t)^{-1/2} \succeq 0. \quad (7)$$

It follows from (4), (5) and (6) respectively that

$$f'(t) = \text{tr} M_1, \quad (8)$$

$$f''(t) = \text{tr} M_1^2 + \text{tr} M_2 = \|M_1\|_F^2 + \|M_2^{1/2}\|_F^2, \quad (9)$$

and

$$f'''(t) = -3\text{tr} M_1 M_2 - 2\text{tr} M_1^3 - \text{tr} \left( X(t)^{-1} X'''(t) \right). \quad (10)$$

By the Cauchy-Schwarz inequality, we have

$$|f'(t)| = |\text{tr} M_1| \leq \sqrt{l} (\text{tr} M_1^2)^{1/2} \leq \sqrt{l} (f''(t))^{1/2}$$

which is (3) with  $C_2 = \sqrt{l}$ .

Since  $f_j$ 's are all quadratic, we have  $X'''(t) \equiv 0$  and so (10) reduces to

$$f'''(t) = -3\text{tr} M_1 M_2 - 2\text{tr} M_1^3. \quad (11)$$

Clearly,  $|\text{tr} M_1^3|^{1/3} \leq (\text{tr} M_1^2)^{1/2} \leq (f''(t))^{1/2}$ . Hence,

$$|\text{tr} M_1^3| \leq (f''(t))^{3/2}. \quad (12)$$

Moreover,

$$\begin{aligned}
|\operatorname{tr} M_1 M_2| &\leq \|M_1\|_F \|M_2\|_F \\
&\leq \|M_1\|_F \|M_2^{1/2}\|_F^2 \\
&\leq \frac{1}{\sqrt{3}} \left( \|M_1\|_F^2 + \|M_2^{1/2}\|_F^2 \right)^{3/2} \\
&= \frac{1}{\sqrt{3}} (f''(t))^{3/2},
\end{aligned} \tag{13}$$

where the last step is due to (9). Therefore, from (11), (12) and (13) we obtain

$$|f'''(t)| \leq (2 + \sqrt{3})(f''(t))^{3/2},$$

which is (2) with  $C_1 = 2 + \sqrt{3}$ .

**Q.E.D.**

In case all  $f_j$ 's are affine linear, we have  $M_2 = 0$ , and  $|f'''(t)| \leq 2(f''(t))^{3/2}$ , which corresponds to the well-known fact that the logarithmic determinant barrier is self-concordant for *Linear Matrix Inequalities*.

We note that if all  $f_j$ 's are convex quadratic functions then the feasible set of  $(P_{PSD})$  can also be described by an enlarged system of linear matrix inequalities using Schur's complement lemma. In particular, suppose that  $f_j(x) = \sum_{i=1}^n (q_{ij}^T x - q_i)^2 + q_{0j}$ ,  $j = 1, \dots, k$ . Then,  $H_0 - \sum_{j=1}^k f_j(x) H_j \succeq 0$  if and only if

$$\begin{bmatrix}
H_0 - \sum_{j=1}^k q_{0j} H_j & (q_{11}^T x - q_1) H_1^{1/2} & \cdots & (q_{21}^T x - q_2) H_2^{1/2} & \cdots & (q_{nk}^T x - q_n) H_k^{1/2} \\
(q_{11}^T x - q_1) H_1^{1/2} & I & 0 & 0 & 0 & 0 \\
\vdots & 0 & I & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
(q_{nk}^T x - q_n) H_k^{1/2} & 0 & 0 & 0 & 0 & I
\end{bmatrix} \succeq 0.$$

However, the complexity parameter of the resulting LMI system becomes  $\sqrt{(kn+1)l}$ , instead of  $\sqrt{l}$ .

## 2 The cone-consistent barriers

The aim of this section is to derive the essential and general property of the barrier function that enabled Theorem 1.2 in the previous section.

Let us call a barrier function  $F(x)$  to be *consistent* with the cone  $\mathcal{K}$  (denoted as  $C_3$ -consistent in the sequel) if it satisfies the following condition:

$$\left(\nabla^2 F(x)[u, u]\right)^{1/2} \leq -C_3 \nabla F(x)[u] \quad (14)$$

for all  $x \in \text{int } \mathcal{K}$  and  $u \in \mathcal{K}$ , where  $C_3 > 0$  is a parameter.

It is obvious that if  $\mathcal{K} = \mathfrak{R}_+^n$  and  $F(x) = -\sum_{i=1}^n \log x_i$  then

$$\left(\nabla^2 F(x)[u, u]\right)^{1/2} = \sqrt{\sum_{i=1}^n \frac{u_i^2}{x_i^2}} \leq \sum_{i=1}^n \frac{u_i}{x_i} = -\nabla F(x)[u]$$

for  $x > 0$  and  $u \geq 0$ . So (14) is satisfied with  $C_3 = 1$ .

It is easy to verify that (14) is also satisfied by the logarithmic determinant barrier function for the positive semidefinite matrix cone. Similarly, it can be shown that the additional inequality (14) is satisfied by all self-scaled barrier functions for the self-scaled cones (see Nesterov and Todd [6]) with  $C_3 = 1$ ; in a different but equivalent framework, the self-scaled barrier function is the log determinant barrier for the symmetric cones (see Faybusovich [1] or Sturm [8, 9]).

One may therefore think that (14) is a property of the self-scaled cones only. However, this guess is incorrect. Let us consider the following cone in  $\mathfrak{R}^3$  as an epigraph of the  $l_4$ -norm in  $\mathfrak{R}^2$ :

$$\mathcal{K}_4 := \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mid x_1 \geq 0, x_1^4 - (x_2^4 + x_3^4) \geq 0 \right\}.$$

The above cone is not self-scaled. Also, the barrier function  $B_4(x) = -\log(x_1^4 - (x_2^4 + x_3^4))$  is not self-concordant, but it satisfies (14) with  $C_3 = 1$ . To see this, consider  $x = [x_1, x_2, x_3]^T \in \text{int } \mathcal{K}_4$ . We denote  $b_4(x) = x_1^4 - (x_2^4 + x_3^4) > 0$  and compute that

$$\nabla B_4(x) = -\frac{1}{b_4(x)} \begin{bmatrix} 4x_1^3 \\ -4x_2^3 \\ -4x_3^3 \end{bmatrix}$$

and

$$\nabla^2 B_4(x) = \nabla B_4(x) \nabla B_4(x)^T - \frac{12}{b_4(x)} \begin{bmatrix} x_1^2 & 0 & 0 \\ 0 & -x_2^2 & 0 \\ 0 & 0 & -x_3^2 \end{bmatrix}.$$

Take any  $y = [y_1, y_2, y_3]^T \in \mathcal{K}_4$ . We have

$$y^T \nabla^2 B_4(x) y = (-\langle \nabla B_4(x), y \rangle)^2 - \frac{12}{b_4(x)} (x_1^2 y_1^2 - x_2^2 y_2^2 - x_3^2 y_3^2).$$

Since  $x_1^2 > \sqrt{x_2^4 + x_3^4}$  and  $y_1^2 \geq \sqrt{y_2^4 + y_3^4}$ , we have

$$\begin{aligned} x_1^2 y_2^2 &\geq \sqrt{(x_2^4 + x_3^4)(y_2^4 + y_3^4)} \\ &\geq \sqrt{x_2^4 y_2^4 + x_2^4 y_3^4 + x_3^4 y_2^4 + x_3^4 y_3^4} \\ &\geq \sqrt{x_2^4 y_2^4 + 2x_2^2 y_2^2 x_3^2 y_3^2 + x_3^4 y_3^4} \\ &= x_2^2 y_2^2 + x_3^2 y_3^2. \end{aligned}$$

Therefore,

$$y^T \nabla^2 B_4(x) y \leq (-\langle \nabla B_4(x), y \rangle)^2.$$

Inequality (14) thus holds for the barrier function  $B_4(x)$  with  $C_3 = 1$ .

One may also suspect that the cone consistence condition (14) is implied by the self-concordance property. However, again this is not the case. In Zhang [12] it is shown that if  $b(x)$  is a convex quadratic function in  $\mathfrak{R}^n$ , then the barrier function  $B_h(\bar{x}) = -\log(q - pb(x/p))$  is a 1-logarithmically homogeneous self-concordant barrier for the convex cone

$$\mathcal{K}_h = \text{cl} \left\{ \bar{x} = \begin{bmatrix} p \\ q \\ x \end{bmatrix} \in \mathfrak{R}_{++} \times \mathfrak{R} \times \mathfrak{R}^n \mid q - pb(x/p) \geq 0 \right\}.$$

Consider  $y = [1, 1 + b(0), 0^T]^T \in \text{int } \mathcal{K}_h$ , and  $\bar{x} \in \text{int } \mathcal{K}_h$ . One computes that

$$\nabla B_h(\bar{x}) = -\frac{1}{q - pb(x/p)} \begin{bmatrix} -b(x/p) + \nabla b(x/p)^T(x/p) \\ 1 \\ -\nabla b(x/p) \end{bmatrix}$$

and

$$\begin{aligned} \nabla^2 B_h(\bar{x}) &= \frac{1}{(q - pb(x/p))^2} \begin{bmatrix} -b(x/p) + \nabla b(x/p)^T(x/p) \\ 1 \\ -\nabla b(x/p) \end{bmatrix} \cdot \begin{bmatrix} -b(x/p) + \nabla b(x/p)^T(x/p) \\ 1 \\ -\nabla b(x/p) \end{bmatrix}^T \\ &+ \frac{1}{p(q - pb(x/p))} \begin{bmatrix} (x/p)^T \nabla^2 b(x/p)(x/p) & 0 & -(x/p)^T \nabla^2 b(x/p) \\ 0 & 0 & 0 \\ -\nabla^2 b(x/p)(x/p) & 0 & \nabla^2 b(x/p) \end{bmatrix}. \end{aligned}$$

Therefore,

$$-\langle \nabla B_h(\bar{x}), y \rangle = \frac{1 + b(0) - b(x/p) + \nabla b(x/p)^T(x/p)}{q - pb(x/p)} > 0 \quad (15)$$



and

$$\nabla^2 B_h(\bar{x})[y, y] = (-\langle \nabla B_h(\bar{x}), y \rangle)^2 + \frac{(x/p)^T \nabla^2 b(x/p)(x/p)}{p(q - pb(x/p))}. \quad (16)$$

Thus, from (15) and (16) we have

$$\frac{\nabla^2 B_h(\bar{x})[y, y]}{(-\langle \nabla B_h(\bar{x}), y \rangle)^2} = 1 + \frac{q - pb(x/p)}{p} \times \frac{(x/p)^T \nabla^2 b(x/p)(x/p)}{(1 + b(0) - b(x/p) + \nabla b(x/p)^T(x/p))^2}.$$

The above quantity can not be uniformly bounded from above for all  $\bar{x} \in \text{int } \mathcal{K}_h$  if  $b(x)$  is strictly convex, because  $q$  can take arbitrarily large value while  $p$  and  $x$  are fixed. This implies that for the barrier function  $B_h(\bar{x})$ , the cone-consistence condition (14) is not satisfied, although the function is self-concordant and 1-logarithmically homogeneous.

These two examples show that the new condition (14) is a related but different property as compared with the self-concordance property, the self-scaled property, or the logarithmic homogeneous property of the barrier functions.

Consider  $(P_{\mathcal{K}})$ . Let  $F(\cdot)$  be a consistent self-concordant barrier function for the cone  $\mathcal{K}$  satisfying (2), (3) and (14) altogether.

Let

$$p(x) := h_0 - \sum_{j=1}^k f_j(x)h_j \in \text{int } \mathcal{K}$$

and

$$f(x) := F(p(x)). \quad (17)$$

Let  $\xi \in \mathfrak{R}^n$  be any given direction. For a chosen coordinate system we have

$$\nabla f(x)[\xi] = \langle \nabla F(p(x)), -\sum_{j=1}^k (\nabla f_j(x)^T \xi)h_j \rangle. \quad (18)$$

Consequently, by the chain-rule,

$$\begin{aligned} \nabla^2 f(x)[\xi, \xi] = & \\ \nabla^2 F(p(x)) & \left[ \sum_{j=1}^k (\nabla f_j(x)^T \xi)h_j, \sum_{j=1}^k (\nabla f_j(x)^T \xi)h_j \right] - \langle \nabla F(p(x)), \sum_{j=1}^k (\xi^T \nabla^2 f_j(x)\xi)h_j \rangle \end{aligned} \quad (19)$$

and

$$\begin{aligned} \nabla^3 f(x)[\xi, \xi, \xi] = & \nabla^3 F(p(x)) \left[ \sum_{j=1}^k (\nabla f_j(x)^T \xi)h_j, \sum_{j=1}^k (\nabla f_j(x)^T \xi)h_j, \sum_{j=1}^k (\nabla f_j(x)^T \xi)h_j \right] \\ & + 3\nabla^2 F(p(x)) \left[ \sum_{j=1}^k (\xi^T \nabla^2 f_j(x)\xi)h_j, \sum_{j=1}^k (\nabla f_j(x)^T \xi)h_j \right] \end{aligned}$$

$$-\left\langle \sum_{j=1}^k \nabla^3 f_j(x)[\xi, \xi, \xi] h_j, \nabla F(p(x)) \right\rangle. \quad (20)$$

Since  $f_j$ 's are convex functions,  $\xi^T \nabla^2 f_j(x) \xi \geq 0$  for all  $j = 1, \dots, k$ , and so

$$\sum_{j=1}^k (\xi^T \nabla^2 f_j(x) \xi) h_j \in \mathcal{K}.$$

Moreover, because  $-\nabla F(p(x)) \in \text{int } \mathcal{K}^*$  we obtain

$$\begin{aligned} \nabla^2 f(x)[\xi, \xi] &\geq \nabla^2 F(p(x)) \left[ \sum_{j=1}^k (\nabla f_j(x)^T \xi) h_j, \sum_{j=1}^k (\nabla f_j(x)^T \xi) h_j \right] \\ &\geq \frac{1}{C_2^2} \langle \nabla F(p(x)), -\sum_{j=1}^k (\nabla f_j(x)^T \xi) h_j \rangle^2 \\ &= \frac{1}{C_2^2} |\nabla f(x)[\xi]|^2 \end{aligned} \quad (21)$$

which is the second inequality required by the self-concordance (3).

Therefore we can use condition (14) to obtain from (19) that

$$\begin{aligned} \nabla^2 f(x)[\xi, \xi] &\geq \nabla^2 F(p(x)) \left[ \sum_{j=1}^k (\nabla f_j(x)^T \xi) h_j, \sum_{j=1}^k (\nabla f_j(x)^T \xi) h_j \right] \\ &\quad + \frac{1}{C_3} \sqrt{\nabla^2 F(p(x)) \left[ \sum_{j=1}^k (\xi^T \nabla^2 f_j(x) \xi) h_j, \sum_{j=1}^k (\xi^T \nabla^2 f_j(x) \xi) h_j \right]}. \end{aligned} \quad (22)$$

Denote

$$d_1 := \left( \nabla^2 F(p(x)) \right)^{1/2} \sum_{j=1}^k (\nabla f_j(x)^T \xi) h_j$$

and

$$d_2 := \left( \nabla^2 F(p(x)) \right)^{1/2} \sum_{j=1}^k (\xi^T \nabla^2 f_j(x) \xi) h_j.$$

We rewrite (22) as

$$\nabla^2 f(x)[\xi, \xi] \geq \|d_1\|^2 + \frac{1}{C_3} \|d_2\| \geq \left[ 3 \left( \frac{\|d_1\| \cdot \|d_2\|}{C_3} \right)^2 \right]^{\frac{1}{3}}. \quad (23)$$

**Theorem 2.1** *Suppose that  $F(\cdot)$  is a self-concordant barrier function for  $\mathcal{K}$ , and moreover it is consistent with  $\mathcal{K}$ . Suppose that  $f_j(x)$ 's are convex quadratic functions,  $j = 1, \dots, k$ . Then  $f(x)$  as defined in (17) is a self-concordant barrier function for  $(P_{\mathcal{K}})$ .*

**Proof.** If  $f_j$ 's are all quadratic functions then it follows from (20) that

$$\begin{aligned}
|\nabla^3 f(x)[\xi, \xi, \xi]| &\leq C_1(\nabla^2 f(x)[\xi, \xi])^{3/2} + 3d_1^T d_2 \\
&\leq C_1(\nabla^2 f(x)[\xi, \xi])^{3/2} + 3\|d_1\|\|d_2\| \\
&\leq C_1(\nabla^2 f(x)[\xi, \xi])^{3/2} + C_3\sqrt{3}(\nabla^2 f(x)[\xi, \xi])^{3/2} \\
&= (C_1 + C_3\sqrt{3})(\nabla^2 f(x)[\xi, \xi])^{3/2},
\end{aligned} \tag{24}$$

where the third inequality is due to (23). Inequalities (24) and (21) show the self-concordance of the function  $f(x)$ . **Q.E.D.**

### 3 A potential reduction approach

Theorem 2.1 suggests that if the conic barrier  $F$  is self-concordant and consistent with the cone, and  $f_j$ 's are all convex quadratic, then the composite barrier function  $f$  is self-concordant. Therefore we can apply the general theory of self-concordant barriers developed by Nesterov and Nemirovskii [5]. However, there are interesting cases that are not included in this model. Consider for instance the following *geometric semidefinite program*:

$$\begin{aligned}
(P_{PSD}) \quad &\text{minimize} && f_0(t) \\
&\text{subject to} && l_i \leq f_i(t) \leq u_i, \quad i = 1, \dots, k+m \\
&&& t_j > 0, \quad j = 1, \dots, n \\
&&& H_0 - \sum_{j=1}^k f_j(t)H_j \succeq 0,
\end{aligned}$$

where  $H_1 \succeq 0, \dots, H_k \succeq 0$  are  $l \times l$  positive semidefinite matrices, and  $f_i(t) = t_1^{i_1} \dots t_n^{i_n}$  is a monomial,  $i = 1, \dots, k+m$ . After the variable transformation  $x_j = \log t_j$ ,  $j = 1, \dots, n$ , the above problem can be turned into

$$\begin{aligned}
(P_{PSD}) \quad &\text{minimize} && t_0^T x \\
&\text{subject to} && \log l_i \leq t_i^T x \leq \log u_i, \quad i = 1, \dots, k+m \\
&&& H_0 - \sum_{j=1}^k e^{t_j^T x} H_j \succeq 0.
\end{aligned}$$

Clearly,  $(P_{PSD})$  is a conically ordered convex program. Unfortunately, in this case, the function  $e^{t_j^T x}$  is not quadratic, and so Theorem 2.1 does not apply: the barrier function  $f(x)$  is not self-concordant in general. However,  $e^{t_j^T x}$ 's are still well structured. In the above particular case,  $\nabla^2 e^{t_j^T x} = e^{t_j^T x} t_j t_j^T$ . Therefore, it holds that  $\nabla^2 e^{t_j^T x} \preceq \lambda \nabla^2 e^{t_j^T y}$  for all  $x$  and  $y$  in its feasible region, where  $\lambda = \max\{u_1/l_1, \dots, u_k/l_k\}$ .

This example motivates us to study algorithms that can solve the COCP problems without an overall easy computable self-concordant barriers. As a matter of fact, most known interior point methods for nonlinear convex programs with provable polynomial-time computational complexities are constructed within the framework of the central-path following using a self-concordant barrier function, except for a perhaps less known member of the IPM family called the Iri-Imai method [3], which is a variant of the potential reduction method. Zhang [11], and Sturm and Zhang [10] studied the Iri-Imai method (or the potential reduction method) applied to the so-called *harmonically convex programs*, and show that such algorithms also run in polynomial time. The main proof techniques in [11], and also proofs in this section, are inspired by a paper of Iri [2].

In this section, we shall study an Iri-Imai type potential reduction method for the conically ordered convex program  $(P_{\mathcal{K}})$ . First we mention the notion of harmonically convex functions; see [11].

**Definition 3.1** *A twice differentiable function  $f(\cdot)$  is called  $\lambda$ -harmonically convex on the convex domain  $\Omega$  if there is a constant  $\lambda$  such that  $0 \preceq \nabla^2 f(x) \preceq \lambda \nabla^2 f(y)$  holds for all  $x, y \in \Omega$ .*

It is easy to see that if  $\lambda > 0$  then  $\frac{1}{\lambda} \nabla^2 f(y) \preceq \nabla^2 f(x) \preceq \lambda \nabla^2 f(y)$  for all  $x, y \in \Omega$ . Obviously, the affine linear functions (0-harmonically convex), the convex quadratic functions (1-harmonically convex), and the strongly convex functions, are all examples of harmonically convex functions. It is shown in [11] that the (direct) sum of these functions characterizes the whole class of harmonically convex functions. If a convex program is to minimize a harmonically convex objective function, subject to harmonically convex inequality constraints, then [11] and [10] show that such problem can be solved by Iri-Imai type potential reduction algorithms in polynomial time. Note that even if  $f(x)$  is harmonically convex and negative on a domain  $\Omega$ , it does not follow that  $-\log(-f(x))$  is self-concordant. Consider for example  $f(x_1, x_2) = -x_1 + x_2^2 + \sin(x_2)$ , which is the sum of two harmonically convex functions,  $-x_1$  and  $x_2^2 + \sin(x_2)$ , hence harmonically convex. However,  $-\log(x_1 - x_2^2 - \sin(x_2))$  is not self-concordant.

Consider  $(P_{\mathcal{K}})$ . Let  $F(\cdot)$  be a barrier function for  $\mathcal{K}$ . Moreover, let us impose the following conditions.

**Condition 3.1**  $F(\cdot)$  is  $\nu$ -logarithmically homogeneous for  $\mathcal{K}$ , namely

$$F(ty) = F(y) - \nu \log t$$

for all  $y \in \text{int } \mathcal{K}$  and  $t > 0$ .

It is well known (see e.g. [5]) that if  $F$  is  $\nu$ -logarithmically homogeneous for  $\mathcal{K}$ , then the following

identities hold where  $y \in \text{int } \mathcal{K}$  and  $t > 0$ :

$$\nabla F(ty) = \frac{1}{t} \nabla F(y); \quad (25)$$

$$\nabla^2 F(ty) = \frac{1}{t^2} \nabla^2 F(y); \quad (26)$$

$$\nabla^2 F(y)y = -\nabla F(y); \quad (27)$$

$$(\nabla F(y))^T y = -\nu. \quad (28)$$

**Condition 3.2**  $F(\cdot)$  is a self-concordant barrier function for  $\mathcal{K}$ . In particular, the following two inequalities hold

$$|\nabla^3 F(y)[u, u, u]| \leq 2(\nabla^2 F(y)[u, u])^{3/2}$$

and

$$|\nabla F(y)[u]| \leq \sqrt{\nu}(\nabla^2 F(y)[u, u])^{1/2},$$

for all  $y \in \text{int } \mathcal{K}$  and  $u \in \text{span } \mathcal{K}$ .

Notice that the second inequality in Condition 3.2 follows from (27), (28), and also the Cauchy-Schwarz inequality; it is in fact implied by Condition 3.1.

Like (17), denote

$$f(x) := F(p(x)),$$

with  $p(x) = h_0 - \sum_{j=1}^k f_j(x)h_j$ . Clearly,  $f(\cdot)$  is a barrier function for the feasible set for  $(P_{\mathcal{K}})$ .

**Lemma 3.2** *Suppose that  $F(\cdot)$  is a  $\nu$ -logarithmically homogeneous barrier function for  $\mathcal{K}$  (Condition 3.1). Denote  $\mathcal{F}$  to be the feasible set for  $(P_{\mathcal{K}})$ . Then*

$$|f(x)[\xi]| \leq \sqrt{\nu} \left( \nabla^2 f(x)[\xi, \xi] \right)^{1/2}$$

for all  $x \in \text{rel int } \mathcal{F}$  and  $\xi \in \mathfrak{R}^n$ .

**Proof.** As  $x \in \text{rel int } \mathcal{F}$  we have  $p(x) = h_0 - \sum_{j=1}^k f_j(x)h_j \in \text{int } \mathcal{K}$ .

Relations (18) and (19) give that

$$\nabla f(x)[\xi] = \langle \nabla F(p(x)), -\sum_{j=1}^k (\nabla f_j(x)^T \xi) h_j \rangle$$

and

$$\nabla^2 f(x)[\xi, \xi] = \nabla^2 F(p(x)) \left[ \sum_{j=1}^k (\nabla f_j(x)^T \xi) h_j, \sum_{j=1}^k (\nabla f_j(x)^T \xi) h_j \right] - \langle \nabla F(p(x)), \sum_{j=1}^k (\xi^T \nabla^2 f_j(x) \xi) h_j \rangle.$$

Since  $-\nabla F(p(x)) \in \mathcal{K}^*$  and  $\sum_{j=1}^k (\xi^T \nabla^2 f_j(x) \xi) h_j \in \mathcal{K}$  due to the convexity of  $f_j$ 's, we obtain from the above equation that

$$\begin{aligned} \nabla^2 f(x)[\xi, \xi] &\geq \nabla^2 F(p(x)) \left[ \sum_{j=1}^k (\nabla f_j(x)^T \xi) h_j, \sum_{j=1}^k (\nabla f_j(x)^T \xi) h_j \right] \\ &\geq \frac{1}{\nu} \left| \langle \nabla F(p(x)), -\sum_{j=1}^k (\nabla f_j(x)^T \xi) h_j \rangle \right|^2 \\ &= \frac{1}{\nu} |f(x)[\xi]|^2, \end{aligned}$$

where the second step is due to Condition 3.1.

**Q.E.D.**

Suppose that  $(P_{\mathcal{K}})$  has a finite optimal value denoted by  $v^*$ . Let  $v \leq v^*$ . For a give parameter  $\gamma > 0$ , let us introduce

$$G(x; v) := (c^T x - v)^\gamma \exp(f(x)). \quad (29)$$

Note that  $G(x; v)$  is an extension of the so-called multiplicative barrier function of Iri and Imai; see [3]. The following theorem extends Iri and Imai's result.

**Theorem 3.3** *Suppose that  $F(\cdot)$  is a  $\nu$ -logarithmically homogeneous barrier function for  $\mathcal{K}$  (Condition 3.1). If  $\gamma \geq \nu + 1$  then  $G(x; v)$  is a convex function on  $\text{int } \mathcal{F}$ ; if  $\gamma > \nu + 1$  then  $G(x; v)$  is a strictly convex function on  $\text{int } \mathcal{F}$ .*

**Proof.** Differentiation yields

$$\frac{1}{G(x; v)} \nabla G(x; v) = \frac{\gamma}{c^T x - v} c + \nabla f(x) \quad (30)$$

$$\frac{1}{G(x; v)} \nabla^2 G(x; v) = \frac{1}{G(x; v)^2} \nabla G(x; v) \nabla G(x; v)^T - \frac{\gamma}{(c^T x - v)^2} c c^T + \nabla^2 f(x). \quad (31)$$

Take any  $0 \neq \eta \in \mathfrak{R}^n$ . Denote  $a := c^T \eta / (c^T x - v)$ ,  $b := \eta^T \nabla f(x)$  and  $e := \eta^T \nabla^2 f(x) \eta > 0$ . By Lemma 3.2 we have  $e \geq b^2 / \nu$ . Using (30) and (31) it follows that

$$\frac{1}{G(x; v)} \eta^T \nabla^2 G(x; v) \eta = (\gamma a + b)^2 - \gamma a^2 + e$$

$$\begin{aligned}
&= (\gamma^2 - \gamma)a^2 + 2\gamma ab + b^2 + e \\
&\geq (\gamma^2 - \gamma)a^2 + 2\gamma ab + (1 + 1/\nu)b^2.
\end{aligned}$$

The last expression is a quadratic form with respect to  $a$  and  $b$ , which is nonnegative as its discriminant is  $\gamma^2 - (\gamma^2 - \gamma)(1 + 1/\nu) = \gamma(\nu + 1 - \gamma)/\nu \leq 0$ . The quantity  $\eta^T \nabla^2 G(x; v) \eta$  is strict positive if  $\gamma > \nu + 1$ , because in this case the discriminant of the quadratic form is negative while  $e > 0$ . The theorem is thus proven. **Q.E.D.**

For simplicity of the analysis, from now on we shall set  $\gamma$  to be  $\nu + \sqrt{\nu}$ .

**Condition 3.3**  $F(\cdot)$  is  $\kappa$ -consistent ( $\kappa > 0$ ) with the cone  $\mathcal{K}$ , i.e.,

$$\left(\nabla^2 F(y)[u, u]\right)^{1/2} \leq -\kappa \nabla F(y)[u]$$

for all  $y \in \text{int } \mathcal{K}$  and  $u \in \mathcal{K}$ .

**Condition 3.4** All  $f_j(\cdot)$ 's,  $j = 1, \dots, k$ , are harmonically convex on  $\mathcal{F}$ . In particular, there is  $\lambda > 0$  such that

$$\nabla^2 f_j(x) \preceq \lambda \nabla^2 f_j(x)$$

for all  $x, y \in \mathcal{F}$  and  $j = 1, \dots, k$ .

**Condition 3.5**  $F(\cdot)$  satisfies  $\nabla^2 F(y)[u, v] \geq 0$  for all  $y \in \text{int } \mathcal{K}$  and  $u, v \in \mathcal{K}$ . In other words,  $\nabla^2 F(y)\mathcal{K} \subseteq \mathcal{K}^*$  for all  $y \in \text{int } \mathcal{K}$ .

A few comments are in order here, concerning the last condition. If  $\mathcal{K}$  is a self-scaled cone and  $F(\cdot)$  is the corresponding self-scaled barrier function, then as a part of the definition it holds that  $\nabla^2 F(y)\mathcal{K} = \mathcal{K}^*$  for all  $y \in \text{int } \mathcal{K}$ . In that respect, Condition 3.5 is naturally satisfied by all self-scaled cones with the corresponding self-scaled barrier functions. Thus, one may conjecture that Condition 3.5 is satisfied if and only if  $\mathcal{K}$  is a self-scaled cone. However, this is not the case. The key is that we only require  $\nabla^2 F(y)\mathcal{K} \subseteq \mathcal{K}^*$  here, rather than  $\nabla^2 F(y)\mathcal{K} = \mathcal{K}^*$  for all  $y \in \text{int } \mathcal{K}$ . In fact,  $\mathcal{K}$  can be any pointed closed convex cone. To see this, let  $\mathcal{C}$  be a second order cone that properly (strictly) contain  $\mathcal{K}$ . Therefore,  $\mathcal{C}^*$  is (strictly) properly contained in  $\mathcal{K}^*$ . Let  $F_c$  be the corresponding self-scaled barrier function for  $\mathcal{C}$ . Thus, for any  $y \in \text{int } \mathcal{K} \subset \text{int } \mathcal{C}$  we have  $\nabla^2 F_c(y)\mathcal{K} \subset \nabla^2 F_c(y)\mathcal{C} = \mathcal{C}^* \subset \mathcal{K}^*$ . For a given barrier function  $F(\cdot)$  for  $\mathcal{K}$ , consider  $F_\mu(\cdot) := F(\cdot) + \mu F_c(\cdot)$ , where  $\mu > 0$  is a given constant. Clearly,  $F_\mu(\cdot)$  is also a barrier function for  $\mathcal{K}$ . If  $\mu$  is chosen sufficiently large then  $\nabla^2 F_\mu(y)\mathcal{K} \subseteq \mathcal{K}^*$  for any  $y \in \text{int } \mathcal{K}$ .

**Lemma 3.4** Suppose that  $F(\cdot)$  satisfies Condition 3.5, and  $v_i, i = 1, \dots, m$ , are  $m$  arbitrary vectors in  $\mathcal{K}$ . Let  $y \in \text{int } \mathcal{K}$  and  $\nabla^2 F(y) \succ 0$  and  $\|x\|_{\nabla^2 F(y)} := \sqrt{x^T \nabla^2 F(y) x}$ . Then,

$$\max \left\{ \left\| \sum_{i=1}^m t_i v_i \right\|_{\nabla^2 F(y)} \mid 0 \leq t_i \leq \hat{t}_i, i = 1, \dots, m \right\} = \left\| \sum_{i=1}^m \hat{t}_i v_i \right\|_{\nabla^2 F(y)}.$$

**Proof.** Under Condition 3.5, we have  $v_i^T \nabla^2 F(y) v_j \geq 0$  for any  $i$  and  $j$ . The lemma is obvious since

$$\left\| \sum_{i=1}^m t_i v_i \right\|_{\nabla^2 F(y)}^2 = \sum_{i,j=1}^m t_i t_j v_i^T \nabla^2 F(y) v_j.$$

**Q.E.D.**

**Theorem 3.5** Suppose that  $F(\cdot)$  satisfies Conditions 3.1, 3.3, and 3.5, and  $f_j(\cdot)$ 's,  $j = 1, \dots, k$ , satisfy Condition 3.4. Let  $x \in \text{rel int } \mathcal{F}$  be any feasible but non-optimal point. Let  $\gamma = \nu + \sqrt{\nu}$ . Let the Newton direction at  $x$  be

$$d_N(x) := - \left( \nabla^2 G(x; v^*) \right)^{-1} \nabla G(x; v^*).$$

Then

$$-\nabla G(x; v^*)^T d_N(x) / G(x; v^*) = \nabla G(x; v^*)^T \left( \nabla^2 G(x; v^*) \right)^{-1} \nabla G(x; v^*) / G(x; v^*) \geq \delta,$$

where  $\delta := 1 / \sqrt{2 + 2\kappa^2 + 2\lambda\kappa} > 0$ .

**Proof.** Let

$$\delta(x) := \nabla G(x; v^*)^T \left( \nabla^2 G(x; v^*) \right)^{-1} \nabla G(x; v^*) / G(x; v^*).$$

Since  $\nabla^2 G(x; v^*) \succ 0$ , by the Cauchy-Schwarz inequality we know that for any  $\eta \neq 0$  it holds

$$\sqrt{\delta(x)} \geq \frac{\nabla G(x; v^*)^T \eta / G(x; v^*)}{\sqrt{\eta^T \nabla^2 G(x; v^*) \eta / G(x; v^*)}}. \quad (32)$$

Let us take  $\eta = x - x^* \neq 0$ . Then, noting  $c^T(x - x^*) / (c^T x - v^*) = 1$  and using (30) and (31) we have

$$\nabla G(x; v^*)^T (x - x^*) / G(x; v^*) = \gamma + \nabla f(x)^T (x - x^*). \quad (33)$$

By the mean-value theorem, there is  $\tilde{x}_j$  between  $x$  and  $x^*$ ,  $j = 1, \dots, k$ , such that

$$p(x) - p(x^*) = - \sum_{j=1}^k \nabla f_j(x)^T (x - x^*) h_j + \frac{1}{2} \sum_{j=1}^k (x - x^*)^T \nabla^2 f_j(\tilde{x}_j) (x - x^*) h_j.$$



By letting

$$\tilde{u} = \frac{1}{2} \sum_{j=1}^k (x - x^*)^T \nabla^2 f_j(\tilde{x}_j) (x - x^*) h_j \in \mathcal{K} \quad (34)$$

we have

$$\begin{aligned} \nabla f(x)^T (x - x^*) &= \nabla F(p(x)) \left[ - \sum_{j=1}^k \nabla f_j(x)^T (x - x^*) h_j \right] \\ &= \nabla F(p(x)) [p(x) - p(x^*) - \tilde{u}] \\ &= -\nu - \nabla F(p(x)) [p(x^*) + \tilde{u}], \end{aligned} \quad (35)$$

where in the last step we used the property (28).

Using the estimation (35) in (33) and the fact that  $p(x), p(x^*), \tilde{u} \in \mathcal{K}$  we have

$$\begin{aligned} \nabla G(x; v^*)^T (x - x^*) / G(x; v^*) &= \sqrt{\nu} - \nabla F(p(x)) [p(x^*) + \tilde{u}] \\ &\geq \sqrt{\nu} + (1/\kappa) \sqrt{\nabla^2 F(p(x)) [p(x^*) + \tilde{u}, p(x^*) + \tilde{u}]}. \end{aligned} \quad (36)$$

Similarly, since  $-\nabla F(p(x)) [p(x^*)] \geq 0$ , we have

$$\nabla G(x; v^*)^T (x - x^*) / G(x; v^*) \geq \sqrt{\nu} - \nabla F(p(x)) [\tilde{u}] \geq \sqrt{\nu} + (1/\kappa) \sqrt{\nabla^2 F(p(x)) [\tilde{u}, \tilde{u}]}. \quad (37)$$

On the other hand,

$$\begin{aligned} &(x - x^*)^T \nabla^2 G(x; v^*) (x - x^*) / G(x; v^*) - \left( \nabla G(x; v^*)^T (x - x^*) / G(x; v^*) \right)^2 \\ &= -\gamma + (x - x^*)^T \nabla^2 f(x) (x - x^*) \\ &= -\gamma + \nabla^2 F(p(x)) \left[ \sum_{j=1}^k \nabla f_j(x)^T (x - x^*) h_j, \sum_{j=1}^k \nabla f_j(x)^T (x - x^*) h_j \right] \\ &\quad - \nabla F(p(x)) \left[ \sum_{j=1}^k (x - x^*)^T \nabla^2 f_j(x)^T (x - x^*) h_j \right] \\ &\leq -\gamma + \nabla^2 F(p(x)) [p(x) - p(x^*) - \tilde{u}, p(x) - p(x^*) - \tilde{u}] \\ &\quad + \sqrt{\nu} \sqrt{\nabla^2 F(p(x)) \left[ \sum_{j=1}^k (x - x^*)^T \nabla^2 f_j(x)^T (x - x^*) h_j, \sum_{j=1}^k (x - x^*)^T \nabla^2 f_j(x)^T (x - x^*) h_j \right]} \\ &\leq -\gamma + 2 \left( \nabla^2 F(p(x)) [p(x), p(x)] + \nabla^2 F(p(x)) [p(x^*) + \tilde{u}, p(x^*) + \tilde{u}] \right) + 2\lambda \sqrt{\nu} \sqrt{\nabla^2 F(p(x)) [\tilde{u}, \tilde{u}]} \\ &\leq -\gamma + 2\nu + 2\kappa^2 \left( \nabla G(x; v^*)^T (x - x^*) / G(x; v^*) \right)^2 + 2\lambda \kappa \sqrt{\nu} \left( \nabla G(x; v^*)^T (x - x^*) / G(x; v^*) \right) \\ &\leq (1 + 2\kappa^2 + 2\lambda \kappa) \left( \nabla G(x; v^*)^T (x - x^*) / G(x; v^*) \right)^2, \end{aligned} \quad (38)$$

where the first step is due to (31) (pre and post multiplying both sides by  $(x - x^*)^T$  and  $x - x^*$  respectively), and at the fourth step we used  $(x - x^*)^T \nabla^2 f_j(x)^T (x - x^*) \leq \lambda (x - x^*)^T \nabla^2 f_j(\tilde{x}_j)^T (x - x^*)$ ,  $j = 1, \dots, k$ , due to Condition 3.4, and then applied Lemma 3.4, and finally at the fifth and the sixth steps we used (36) and (37).

The theorem follows by combining (32) and (38) to obtain  $\delta(x) \geq \delta$ .

**Q.E.D.**

Now let us study the effect of applying the Newton method (with line-minimization) to reduce the value of  $G(x; v)$  where  $v \leq v^*$ . Denote

$$d := - \left( \nabla^2 G(x; v) \right)^{-1} \nabla G(x; v).$$

Multiplying  $d^T$  on both sides of (30), and pre-multiplying  $d^T$  and post-multiplying  $d$  on both sides of (31), and denoting  $a := c^T d / (c^T x - v)$  and, if there is no confusion, still denoting  $\delta(x) := \nabla G(x; v)^T (\nabla^2 G(x; v))^{-1} \nabla G(x; v) / G(x; v)$ , we obtain

$$-\delta(x) = \gamma a + d^T \nabla f(x) \tag{39}$$

$$\delta(x) - \delta(x)^2 = -\gamma a^2 + d^T \nabla^2 f(x) d. \tag{40}$$

Eliminating  $a$  from (39) and (40) we get

$$\gamma \delta(x)(1 - \delta(x)) + (\delta(x) + d^T \nabla f(x))^2 = \gamma d^T \nabla^2 f(x) d. \tag{41}$$

Relations (18) and (19) yield

$$\nabla f(x)^T d = \nabla F(p(x)) \left[ - \sum_{j=1}^k d^T \nabla f_j(x) h_j \right] \tag{42}$$

$$\begin{aligned} d^T \nabla^2 f(x) d &= \nabla^2 F(p(x)) \left[ \sum_{j=1}^k d^T \nabla f_j(x) h_j, \sum_{j=1}^k d^T \nabla f_j(x) h_j \right] \\ &\quad - \nabla F(p(x)) \left[ \sum_{j=1}^k d^T \nabla^2 f_j(x) d h_j \right]. \end{aligned} \tag{43}$$

Since  $F(\cdot)$  is  $\nu$ -logarithmically homogeneous and  $f_j$ 's are convex, we obtain from (41) and (43) that

$$\begin{aligned} \gamma \delta(x)(1 - \delta(x)) + (\delta(x) + d^T \nabla f(x))^2 &\geq \gamma \nabla^2 F(p(x)) \left[ \sum_{j=1}^k d^T \nabla f_j(x) h_j, \sum_{j=1}^k d^T \nabla f_j(x) h_j \right] \\ &\geq \frac{\gamma}{\nu} \left( \nabla F(p(x)) \left[ - \sum_{j=1}^k d^T \nabla f_j(x) h_j \right] \right)^2 \\ &= \frac{\gamma}{\nu} \left( d^T \nabla f(x) \right)^2. \end{aligned} \tag{44}$$

Noting that  $\gamma = \nu + \sqrt{\nu}$ , because  $d^T \nabla f(x)$  is a real number, checking the discriminant of the above quadratic inequality in terms of  $d^T \nabla f(x)$  we conclude that

$$\delta(x) \leq \frac{\nu + \sqrt{\nu}}{\nu - 1}. \quad (45)$$

The quadratic inequality (44) can be rearranged as

$$\alpha_1(x) \leq d^T \nabla f(x) \leq \alpha_2(x), \quad (46)$$

with

$$\begin{aligned} \alpha_1(x) &:= \sqrt{\nu} \delta(x) - \sqrt{\sqrt{\nu}(\gamma - (\nu - 1)\delta(x)) \delta(x)} \\ \alpha_2(x) &:= \sqrt{\nu} \delta(x) + \sqrt{\sqrt{\nu}(\gamma - (\nu - 1)\delta(x)) \delta(x)}. \end{aligned}$$

Simplifying, we get

$$\alpha_2(x) \leq \frac{\nu + \nu\sqrt{\nu}}{\nu - 1} + \frac{\nu + \sqrt{\nu}}{2} \sqrt{\frac{\sqrt{\nu}}{\nu - 1}} =: \alpha_2 \text{ and } |\alpha_1(x)| \leq \alpha_2(x) \leq \alpha_2.$$

Now we use the  $\kappa$ -consistency, Condition 3.3, to obtain from (41) and (43) that

$$\begin{aligned} & \frac{1}{\kappa} \sqrt{\nabla^2 F(p(x)) \left[ \sum_{j=1}^k d^T \nabla^2 f_j(x) dh_j, \sum_{j=1}^k d^T \nabla^2 f_j(x) dh_j \right]} \\ & \leq -\nabla F(p(x)) \left[ \sum_{j=1}^k d^T \nabla^2 f_j(x) dh_j \right] \\ & \leq d^T \nabla^2 f(x) d \\ & = \frac{\gamma \delta(x) (1 - \delta(x)) + (\delta(x) + d^T \nabla f(x))^2}{\gamma} \\ & \leq \frac{\gamma \delta(x) + 2\delta(x)^2 + 2(d^T \nabla f(x))^2}{\gamma} \\ & \leq \frac{\frac{\gamma^2}{\nu-1} + 2\frac{\gamma^2}{(\nu-1)^2} + 2\alpha_2^2}{\gamma} =: \alpha_3, \end{aligned} \quad (47)$$

where the second step is due to (43), the third step is due to (41), and the last step is due to (45) and (46).

This yields

$$\nabla^2 F(p(x)) \left[ \sum_{j=1}^k d^T \nabla^2 f_j(x) dh_j, \sum_{j=1}^k d^T \nabla^2 f_j(x) dh_j \right] \leq \kappa^2 \alpha_3^2. \quad (48)$$

Along the same line, from (43) and (47) we get

$$\nabla^2 F(p(x)) \left[ \sum_{j=1}^k d^T \nabla f_j(x) h_j, \sum_{j=1}^k d^T \nabla f_j(x) h_j \right] \leq d^T \nabla^2 f(x) d \leq \alpha_3. \quad (49)$$

Consider an iterative point  $x + td$  along the Newton direction  $d$  with step size  $t > 0$ . Using the mean-value theorem, there are  $\tilde{x}_j \in [x, x + td]$ ,  $j = 1, \dots, k$ , such that

$$\begin{aligned} p(x + td) - p(x) &= - \sum_{j=1}^k (f_j(x + td) - f_j(x)) h_j \\ &= -t \sum_{j=1}^k d^T \nabla f_j(x) h_j - \frac{t^2}{2} \sum_{j=1}^k d^T \nabla^2 f_j(\tilde{x}_j) d h_j. \end{aligned} \quad (50)$$

Let  $\tilde{u} := \sum_{j=1}^k d^T \nabla^2 f_j(\tilde{x}_j) d h_j$ .

Remember that the norm induced by the Hessian of  $F$  at  $y$  is denoted by  $\|u\|_{\nabla^2 F(y)} = \sqrt{u^T \nabla^2 F(y) u}$ .

The lemma below is well known and it is a crucial property of the self-concordant barrier functions; see [4] and [7].

**Lemma 3.6** *Suppose that  $F(\cdot)$  is self-concordant, i.e., it satisfies Condition 3.2. Suppose that  $p(x) \in \text{int } \mathcal{K}$ . If  $\|y - p(x)\|_{\nabla^2 F(p(x))} < 1$  then  $y \in \text{int } \mathcal{K}$ . Moreover,*

$$\|u\|_{\nabla^2 F(y)} \leq \frac{\|u\|_{\nabla^2 F(p(x))}}{1 - \|y - p(x)\|_{\nabla^2 F(p(x))}},$$

for all  $u$ .

In this notation, (48) can be written as

$$\left\| \sum_{j=1}^k d^T \nabla^2 f_j(x) d h_j \right\|_{\nabla^2 F(p(x))} \leq \kappa \alpha_3, \quad (51)$$

and (49) as

$$\left\| \sum_{j=1}^k d^T \nabla f_j(x) h_j \right\|_{\nabla^2 F(p(x))} \leq \sqrt{\alpha_3}. \quad (52)$$

**Lemma 3.7** *Suppose that  $F(\cdot)$  satisfies Conditions 3.1, 3.2, 3.3, and 3.5, and  $f_j(\cdot)$ 's,  $j = 1, \dots, k$ , satisfy Condition 3.4. Let  $x \in \text{rel int } \mathcal{F}$  be any feasible but non-optimal point. Let  $\gamma = \nu + \sqrt{\nu}$ . Let the Newton direction at  $x$  be  $d = -(\nabla^2 G(x; v))^{-1} \nabla G(x; v)$  with  $v \leq v^*$ . Let*

$$\alpha_4 := \frac{2}{\sqrt{\alpha_3}(1 + \sqrt{1 + 2\lambda\kappa})}.$$

Then, for any step-size  $0 < t < \alpha_4$ , the iterative point  $p(x + td) \in \mathcal{F}$ .

**Proof.** It follows from (50) that

$$p(x + td) - p(x) = -t \sum_{j=1}^k d^T \nabla f_j(x) h_j - \frac{t^2}{2} \tilde{u}.$$

By Lemma 3.4 and Condition 3.4, also noting (51) we have

$$\|\tilde{u}\|_{\nabla^2 F(p(x))} \leq \lambda \left\| \sum_{j=1}^k d^T \nabla^2 f_j(x) d h_j \right\|_{\nabla^2 F(p(x))} \leq \lambda \kappa \alpha_3.$$

Using the above estimation and (52) we know that if  $0 < t < \alpha_4$  then

$$\begin{aligned} \|p(x + td) - p(x)\|_{\nabla^2 F(p(x))} &\leq t \left\| \sum_{j=1}^k d^T \nabla f_j(x) h_j \right\|_{\nabla^2 F(p(x))} + \frac{t^2}{2} \|\tilde{u}\|_{\nabla^2 F(p(x))} \\ &\leq t\sqrt{\alpha_3} + \frac{t^2}{2} \lambda \kappa \alpha_3 < 1, \end{aligned}$$

because  $\alpha_4$  is the largest root of the quadratic equation

$$t\sqrt{\alpha_3} + \frac{t^2}{2} \lambda \kappa \alpha_3 - 1 = 0.$$

By Lemma 3.6, the result follows. **Q.E.D.**

Remark that, in a similar way one can show that if  $0 < t < \alpha_5$  with

$$\alpha_5 := \frac{1}{\sqrt{\alpha_3}(1 + \sqrt{1 + \lambda \kappa})},$$

then

$$t \left\| \sum_{j=1}^k d^T \nabla f_j(x) h_j \right\|_{\nabla^2 F(p(x))} + \frac{t^2}{2} \|\tilde{u}\|_{\nabla^2 F(p(x))} \leq t\sqrt{\alpha_3} + \frac{t^2}{2} \lambda \kappa \alpha_3 < 1/2. \quad (53)$$

**Theorem 3.8** *Suppose that  $F(\cdot)$  satisfies Conditions 3.1, 3.2, 3.3, and 3.5, and  $f_j(\cdot)$ 's,  $j = 1, \dots, k$ , satisfy Condition 3.4. Let  $x \in \text{rel int } \mathcal{F}$  be any feasible but non-optimal point. Let  $\gamma = \nu + \sqrt{\nu}$ . Let the Newton direction at  $x$  be  $d = -(\nabla^2 G(x; v))^{-1} \nabla G(x; v)$ . Suppose that  $-\nabla G(x; v)^T d / G(x; v) \geq \delta$ . Then there exists  $\beta > 0$  such that*

$$\min_{0 \leq t \leq \alpha_5} (\log G(x + td; v) - \log G(x; v)) \leq -\beta.$$

**Proof.** Denote  $x(t) := x + td$ , and let  $a := \frac{c^T d}{c^T x - v}$ . Lemma 3.7 and the remark thereafter assure that if  $0 < t < \alpha_5 (< \alpha_4)$  then  $x(t) \in \mathcal{F}$ , and  $\|p(x(t)) - p(x)\|_{\nabla^2 F(p(x))} < 1/2$ .

Now, consider a fixed  $0 < t < \alpha_5$ . By the mean-value theorem, there is  $0 \leq \tilde{t} \leq t < \alpha_5$  such that the following estimation holds

$$\begin{aligned}
& \log G(x + td; v) - \log G(x; v) = \gamma \log(1 + at) + f(x(t)) - f(x) \\
& = (\gamma a + \nabla f(x)^T d)t + \frac{t^2}{2} \left( -\gamma \frac{a^2}{(1 + at)^2} + d^T \nabla^2 f(x(\tilde{t}))d \right) \\
& \leq -\delta(x)t + \frac{t^2}{2} \times \nabla^2 F(p(x(\tilde{t}))) \left[ \sum_{j=1}^k d^T \nabla f_j(x(\tilde{t}))h_j, \sum_{j=1}^k d^T \nabla f_j(x(\tilde{t}))h_j \right] \\
& \quad + \frac{t^2}{2} \times \sqrt{\nu} \sqrt{\nabla^2 F(p(x(\tilde{t}))) \left[ \sum_{j=1}^k d^T \nabla^2 f_j(x(\tilde{t}))dh_j, \sum_{j=1}^k d^T \nabla^2 f_j(x(\tilde{t}))dh_j \right]} \\
& \leq -\delta t + \frac{t^2}{2} \times \left[ \frac{\left\| \sum_{j=1}^k d^T \nabla f_j(x(\tilde{t}))h_j \right\|_{\nabla^2 F(p(x))}^2}{\left(1 - \|p(x(\tilde{t})) - p(x)\|_{F(p(x))}\right)^2} + \sqrt{\nu} \times \frac{\left\| \sum_{j=1}^k d^T \nabla^2 f_j(x(\tilde{t}))dh_j \right\|_{\nabla^2 F(p(x))}}{1 - \|p(x(\tilde{t})) - p(x)\|_{\nabla^2 F(p(x))}} \right].
\end{aligned} \tag{54}$$

Notice that

$$\begin{aligned}
& \left\| \sum_{j=1}^k d^T \nabla f_j(x(\tilde{t}))h_j \right\|_{\nabla^2 F(p(x))}^2 \\
& \leq \left( \left\| \sum_{j=1}^k d^T \nabla f_j(x)h_j \right\|_{\nabla^2 F(p(x))} + \frac{\lambda}{2} \left\| \sum_{j=1}^k d^T \nabla^2 f_j(x)dh_j \right\|_{\nabla^2 F(p(x))} \right)^2 \\
& \leq 2 \left( \left\| \sum_{j=1}^k d^T \nabla f_j(x)h_j \right\|_{\nabla^2 F(p(x))}^2 + \frac{\lambda^2}{4} \left\| \sum_{j=1}^k d^T \nabla^2 f_j(x)dh_j \right\|_{\nabla^2 F(p(x))}^2 \right) \\
& \leq 2\alpha_3 + \lambda^2 \kappa^2 \alpha_3^2 / 2,
\end{aligned}$$

where the last step follows from (51) and (52), and

$$\left\| \sum_{j=1}^k d^T \nabla^2 f_j(x(\tilde{t}))dh_j \right\|_{\nabla^2 F(p(x))} \leq \lambda \left\| \sum_{j=1}^k d^T \nabla^2 f_j(x)dh_j \right\|_{\nabla^2 F(p(x))} \leq \lambda \kappa \alpha_3.$$

The coefficient of the second order term in (54) can be further estimated as

$$\frac{\left\| \sum_{j=1}^k d^T \nabla f_j(x(\tilde{t}))h_j \right\|_{\nabla^2 F(p(x))}^2}{\left(1 - \|p(x(\tilde{t})) - p(x)\|_{\nabla^2 F(p(x))}\right)^2} + \sqrt{\nu} \times \frac{\left\| \sum_{j=1}^k d^T \nabla^2 f_j(x(\tilde{t}))dh_j \right\|_{\nabla^2 F(p(x))}}{1 - \|p(x(\tilde{t})) - p(x)\|_{\nabla^2 F(p(x))}}$$

$$\leq 8\alpha_3 + 2\lambda^2\kappa^2\alpha_3^2 + 2\lambda\kappa\alpha_3\sqrt{\nu} =: \alpha_6.$$

Using these estimations we rewrite (54) as follows

$$\log G(x + td; v) - \log G(x; v) \leq -\delta t + \frac{\alpha_6}{2}t^2.$$

Therefore, if we take the step-length as  $\bar{t} := \min\{\frac{\delta}{\alpha_6}, \alpha_5\}$ , then  $\log G(x + \bar{t}d; v) - \log G(x; v) \leq -\beta$ , where  $\beta := \delta\bar{t} - \frac{\alpha_6}{2}\bar{t}^2$ . **Q.E.D.**

If  $\mathcal{K}$  is a symmetric cone, say  $\mathcal{K} = \mathcal{S}_+^l$  and  $F(p) = -\log \det p$ , then we can improve the bound in (54). In that case, we have

$$\begin{aligned} -\nabla F(p)[h] &= \operatorname{tr} p^{-1}h \\ \nabla^2 F(p)[h, h] &= \operatorname{tr}(p^{-1}hp^{-1}h). \end{aligned}$$

Therefore, if  $\|\tilde{p} - p\|_{\nabla^2 F(p)} = \sqrt{\nabla^2 F(p)[\tilde{p} - p, \tilde{p} - p]} = \|p^{-\frac{1}{2}}(\tilde{p} - p)p^{-\frac{1}{2}}\|_F = \|p^{-\frac{1}{2}}\tilde{p}p^{-\frac{1}{2}} - I\|_F < 1/2$  then  $\frac{1}{2}p \preceq \tilde{p} \preceq \frac{3}{2}p$ , and since  $h \in \mathcal{K}$  we have

$$0 < -\nabla F(\tilde{p})[h] = \operatorname{tr} \tilde{p}^{-1}h \leq 2\operatorname{tr} p^{-1}h = -2\nabla F(p)[h].$$

Instead of the estimations in (54), we now have

$$\begin{aligned} \log G(x + td; v) - \log G(x; v) &= \gamma \log(1 + at) + f(x(t)) - f(x) \\ &= (\gamma a + \nabla f(x)^T d)t + \frac{t^2}{2} \left( -\gamma \frac{a^2}{(1 + at)^2} + d^T \nabla^2 f(x(\tilde{t}))d \right) \\ &\leq -\delta(x)t + \frac{t^2}{2} \times \nabla^2 F(p(x(\tilde{t}))) \left[ \sum_{j=1}^k d^T \nabla f_j(x(\tilde{t}))h_j, \sum_{j=1}^k d^T \nabla f_j(x(\tilde{t}))h_j \right] \\ &\quad - \frac{t^2}{2} \times \nabla F(p(x(\tilde{t}))) \left[ \sum_{j=1}^k d^T \nabla^2 f_j(x(\tilde{t}))dh_j \right] \\ &\leq -\delta t + \frac{t^2}{2} \times \left[ \frac{\left\| \sum_{j=1}^k d^T \nabla f_j(x(\tilde{t}))h_j \right\|_{\nabla^2 F(p(x))}^2}{\left(1 - \|p(x(\tilde{t})) - p(x)\|_{F(p(x))}\right)^2} - 2\lambda \nabla F(p(x)) \left[ \sum_{j=1}^k d^T \nabla^2 f_j(x)dh_j \right] \right] \\ &\leq -\delta t + \frac{t^2}{2} \times \left( 8\alpha_3 + 2\lambda^2\kappa^2\alpha_3^2 + 2\lambda\alpha_3 \right). \end{aligned}$$

By letting  $\alpha_7 := 8\alpha_3 + 2\lambda^2\kappa^2\alpha_3^2 + 2\lambda\alpha_3$ , and  $\bar{t}' := \min\{\frac{\delta}{\alpha_7}, \alpha_5\}$  we have the estimation that  $\log G(x + \bar{t}'d; v) - \log G(x; v) \leq -\beta'$  with  $\beta' := \delta\bar{t}' - \frac{\alpha_7}{2}(\bar{t}')^2$ .

One easily verifies that all the constants,  $\alpha_2$  to  $\alpha_7$ ,  $\delta$ ,  $\beta$  and  $\beta'$ , are rationally related to the problem parameters  $\nu$ ,  $\kappa$  and  $\lambda$ . This gives rise to a polynomial-time complexity proof for the following extended version of the Iri-Imai method (or, potential reduction method).

### An Iri-Imai Type Potential Reduction Method for COCP

*Input data:* An initial  $x^0 \in \text{rel int } \mathcal{F}$  and an initial lower bound of the optimal value  $v^0 \leq v^*$ .

*Output:* A feasible solution  $x \in \text{rel int } \mathcal{F}$  such that  $c^T x - v^* \leq \epsilon$  where  $\epsilon > 0$  is the desired precision.

**Step 0.** Let  $x := x^0$  and  $v := v^0$ . Go to Step 1.

**Step 1.** Check whether  $c^T x - v \leq \epsilon$  or not. If *yes*, stop with  $x$  as the output; if *no*, continue with Step 2.

**Step 2.** Compute the Newton direction at  $x$

$$d := - \left( \nabla^2 G(x; v) \right)^{-1} \nabla G(x; v).$$

If  $-\nabla G(x; v)^T d / G(x; v) < \delta$  then increase the value  $v$  until  $-\nabla G(x; v)^T d / G(x; v) = \delta$ ; otherwise continue with Step 3.

**Step 3.** Apply the line-minimization for  $\log G(x; v)$  along the direction  $d$ , starting from  $x$ . Let

$$\bar{t} = \operatorname{argmin}_{t \geq 0} \log G(x + td; v).$$

Let  $x := x + \bar{t}d$ , and go to Step 1.

The next theorem is a consequence of Theorem 3.8 and the remarks thereafter.

**Theorem 3.9** *Suppose that  $F(\cdot)$  satisfies Conditions 3.1, 3.2, 3.3, and 3.5, and  $f_j(\cdot)$ 's,  $j = 1, \dots, k$ , satisfy Condition 3.4. Let  $\gamma = \nu + \sqrt{\nu}$ . Then the above described Iri-Imai type potential reduction algorithm terminates in a number of steps, which is bounded by a polynomial in terms of  $\nu$ ,  $\kappa$ ,  $\lambda$  and  $\log \frac{1}{\epsilon}$ , provided that (i) an optimal solution for  $(P_\kappa)$  exists; (ii)  $f(x) \geq -L$  for all  $x \in \mathcal{F}$ , where  $|L|$  is bounded by a polynomial in  $\nu$ ,  $\kappa$  and  $\lambda$ ; (iii) the initial potential function value  $\log G(x^0; v^0)$  is bounded by a polynomial in  $\nu$ ,  $\kappa$ ,  $\lambda$ , and  $\log \frac{1}{\epsilon}$ .*



**Proof.** Let the iterates produced by the algorithm be  $\{x^0, x^1, \dots, x^j, \dots\}$ , and the sequence of the lower bounds be  $\{v^0, v^1, \dots, v^j, \dots\}$ . By Theorem 3.8 we have

$$\log G(x^j; v^j) \leq \log G(x^{j-1}; v^{j-1}) - \beta \leq \log G(x^0; v^0) - j\beta$$

where  $1/\beta$  is bounded from above by a polynomial in  $\nu$ ,  $\kappa$  and  $\lambda$ . Thus,

$$\log(c^T x^j - v^j) \leq \frac{\log G(x^0; v^0) - j\beta - f(x^j)}{\gamma} \leq \frac{\log G(x^0; v^0) - j\beta + L}{\gamma}. \quad (55)$$

Therefore, if

$$j \geq \frac{\log G(x^0; v^0) + L + \gamma \log(1/\epsilon)}{\beta}$$

then it follows from (55) that

$$c^T x^j - v^j \leq \epsilon.$$

**Q.E.D.**

We observe that Theorem 3.9 provides explicit computational complexity bounds for some new classes of convex programming problems. To give one such example, let us call a convex function  $f(x)$  to be *quadratically dominated* if it can be split into two parts:  $f(x) = q(x) + h(x)$  where  $q(x)$  is convex quadratic and  $h(x)$  is any other function whose Hessian is dominated by  $\nabla^2 q(x) = Q \succeq 0$ , i.e., there is  $0 \leq \theta < 1$  such that  $-\theta Q \preceq \nabla^2 h(x) \preceq \theta Q$ . Our example at the beginning of Section 3,  $f(x) = x^2 + \sin x$ , is such a convex quadratically dominated function. Now consider the following COCP:  $\mathcal{K}$  is a symmetric cone ( $\kappa = 1$ );  $f_j$ 's are convex quadratically dominated functions (all with a uniform and constant bound  $\theta$ ). Since convex quadratically dominated functions are harmonically convex with  $\lambda = (1 + \theta)/(1 - \theta)$ , we have  $\lambda = O(1)$ , as  $\theta$  is assumed to be fixed here. It follows that

$$\begin{aligned} \alpha_2 &= \frac{\nu + \nu\sqrt{\nu}}{\nu-1} + \frac{\nu + \sqrt{\nu}}{2} \sqrt{\frac{\sqrt{\nu}}{\nu-1}} && \longrightarrow O(\sqrt{\nu}) \\ \alpha_3 &= \frac{\frac{\gamma^2}{\nu-1} + 2\frac{\gamma^2}{(\nu-1)^2} + 2\alpha_2^2}{\gamma} && \longrightarrow O(1) \\ \alpha_4 &= \frac{2}{\sqrt{\alpha_3(1 + \sqrt{1 + 2\lambda\kappa})}} && \longrightarrow O(1) \\ \alpha_5 &= \frac{1}{\sqrt{\alpha_3(1 + \sqrt{1 + \lambda\kappa})}} && \longrightarrow O(1) \\ \alpha_7 &= 8\alpha_3 + 2\lambda^2\kappa^2\alpha_3^2 + 2\lambda\alpha_3 && \longrightarrow O(1) \\ \delta &= 1/\sqrt{2 + 2\kappa^2 + 2\lambda\kappa} && \longrightarrow O(1) \\ \bar{\ell} &= \min\left\{\frac{\delta}{\alpha_7}, \alpha_5\right\} && \longrightarrow O(1) \\ \beta' &= \delta\bar{\ell} - \frac{\alpha_7}{2}(\bar{\ell})^2 && \longrightarrow O(1). \end{aligned}$$

Therefore, the Iri-Imai type potential reduction algorithm would solve this problem in  $O(\nu \log \frac{1}{\epsilon})$  number of iterations to reach an  $\epsilon$ -optimal solution if  $\epsilon > 0$  is small enough, as asserted by Theorem 3.9. Similarly, the *geometric semidefinite program* ( $P_{PSD}$ ) that we introduced at the beginning of this section can be solved in  $O((m + k + l) \log \frac{1}{\epsilon})$  number of iterations if  $\max_{1 \leq i \leq k} u_i/l_i$  is a constant.

Even if  $\mathcal{K}$  is not a symmetric cone (assuming  $\kappa$  is fixed however) and  $f_j$ 's are either quadratically dominated convex functions or they are generated from monomials, since  $\alpha_6 = 8\alpha_3 + 2\lambda^2\kappa^2\alpha_3^2 + 2\lambda\kappa\alpha_3\sqrt{\nu} = O(\sqrt{\nu})$  we can still apply Theorem 3.8 and Theorem 3.9 to conclude that these problems can be solved in no more than  $O(\nu^{1.5} \log \frac{1}{\epsilon})$  number of iterations to reach an  $\epsilon$ -optimal solution.

## 4 Duality and lower bounds

What remains to be discussed is how to obtain a lower bound  $v \leq v^*$  in order to apply the Iri-Imai type potential reduction method as introduced in the previous section. This naturally leads to the duality issues. Consider the following COCP problem:

$$(P_{\mathcal{K}}) \quad \begin{aligned} &\text{minimize} && c^T x \\ &\text{subject to} && Ax = b \\ &&& h_0 - \sum_{j=1}^k f_j(x)h_j \in \mathcal{K}, \end{aligned}$$

where  $A \in \mathfrak{R}^{m \times n}$ ,  $c \in \mathfrak{R}^n$  and  $b \in \mathfrak{R}^m$ ,  $h_1, h_2, \dots, h_k \in \mathcal{K}$ , and  $f_j(\cdot)$ 's are convex functions.

For this purpose, let us note that the conjugate of a convex function  $f(x)$ , which is also a convex function, is defined as

$$f^*(s) = \sup\{(-s)^T x - f(x) \mid x \in \text{dom } f\},$$

where ‘dom  $f$ ’ stands for the domain of the function  $f$ .

Define a closed convex cone as follows

$$\mathcal{C} := \text{cl} \left\{ \left[ \begin{array}{c} p \\ q \\ x \end{array} \right] \mid p > 0, q - pf(x/p) \geq 0 \right\}.$$

Then, the dual of  $\mathcal{C}$  can be computed as

$$\mathcal{C}^* = \text{cl} \left\{ \left[ \begin{array}{c} u \\ v \\ s \end{array} \right] \mid v > 0, u - vf^*(s/v) \geq 0 \right\}.$$

One is referred to [12] for discussions on this and other related issues.

Let

$$\mathcal{K}_j := \text{cl} \left\{ \left[ \begin{array}{c} p_j \\ q_j \\ x_j \end{array} \right] \mid p_j > 0, q_j - p_j f_j(x_j/p_j) \geq 0 \right\}, j = 1, \dots, k.$$

Then,

$$\mathcal{K}_j^* = \text{cl} \left\{ \left[ \begin{array}{c} u_j \\ v_j \\ s_j \end{array} \right] \middle| v_j > 0, u_j - v_j f_j^*(s_j/v_j) \geq 0 \right\}, j = 1, \dots, k.$$

In light of these new cones, we may rewrite our original COCP as

$$\begin{aligned} (P_{\mathcal{K}}) \quad & \text{minimize} && c^T x \\ & \text{subject to} && Ax = b \\ & && \left[ \begin{array}{c} 1 \\ q_j \\ x \end{array} \right] \in \mathcal{K}_j, j = 1, \dots, k \\ & && h_0 - \sum_{j=1}^k q_j h_j \in \mathcal{K}. \end{aligned}$$

The above description is in the form of *linear* conic optimization. Hence its dual problem can be found by standard procedures. Let us derive the dual using the Lagrangian multipliers. In particular,

let  $y$  be the multipliers for the first set of equality constraints, and  $\left[ \begin{array}{c} u_j \\ v_j \\ s_j \end{array} \right] \in \mathcal{K}_j^*$  be the multipliers for the second set of (conic) constraints. Let  $z \in \mathcal{K}^*$  be the multiplier for the last conic constraint.

Then the Lagrangian function is

$$L(x, q; y, u, v, s, z) = c^T x - y^T (Ax - b) - \sum_{j=1}^k (u_j + v_j q_j + s_j^T x) - z^T (h_0 - \sum_{j=1}^k q_j h_j),$$

from which the dual problem can be derived as

$$\begin{aligned} (D_{\mathcal{K}}) \quad & \text{maximize} && b^T y - h_0^T z - \sum_{j=1}^k u_j \\ & \text{subject to} && A^T y + \sum_{j=1}^k s_j = c \\ & && \left[ \begin{array}{c} u_j \\ h_j^T z \\ s_j \end{array} \right] \in \mathcal{K}_j^*, j = 1, \dots, k \\ & && z \in \mathcal{K}^*. \end{aligned}$$

This problem can be further explicitly expressed as

$$\begin{aligned}
(D_{\mathcal{K}}) \quad & \text{maximize} \quad b^T y - h_0^T z - \sum_{j=1}^k (h_j^T z) f_j^*(s_j/h_j^T z) \\
& \text{subject to} \quad A^T y + \sum_{j=1}^k s_j = c \\
& \quad \quad \quad z \in \mathcal{K}^*.
\end{aligned}$$

Therefore, a lower bound  $v \leq v^*$  for  $(P_{\mathcal{K}})$  could be obtained from a finite feasible solution for the above problem  $(D_{\mathcal{K}})$ .

If  $f_j(x) = \frac{1}{2}x^T Q_j x$  with  $Q_j \succ 0$  for all  $j$ , then  $f_j^*(s) = \frac{1}{2}s^T Q_j^{-1} s$ , and the dual problem becomes

$$\begin{aligned}
& \text{maximize} \quad b^T y - h_0^T z - \frac{1}{2} \sum_{j=1}^k s_j^T Q_j^{-1} s_j / h_j^T z \\
& \text{subject to} \quad A^T y + \sum_{j=1}^k s_j = c \\
& \quad \quad \quad z \in \mathcal{K}^*.
\end{aligned}$$

In the case of ‘conically ordered geometric programming’, where  $f_j(x) = e^{\iota_j^T x}$  for all  $j$ . It is easy to compute that

$$f_j^*(s_j) = \begin{cases} +\infty, & \text{if } s_j \neq -\frac{\|s_j\|}{\|\iota_j\|} \iota_j \\ \frac{\|s_j\|}{\|\iota_j\|} \log \frac{\|s_j\|}{\|\iota_j\|} - \frac{\|s_j\|}{\|\iota_j\|}, & \text{if } s_j = -\frac{\|s_j\|}{\|\iota_j\|} \iota_j. \end{cases}$$

Therefore, its dual is

$$\begin{aligned}
& \text{maximize} \quad b^T y - h_0^T z - \sum_{j=1}^k w_j \left( \log w_j - \log h_j^T z \right) + \sum_{j=1}^k w_j \\
& \text{subject to} \quad A^T y - [\iota_1, \dots, \iota_k] w = c \\
& \quad \quad \quad z \in \mathcal{K}^*, w \geq 0.
\end{aligned}$$

## References

- [1] L. Faybusovich, *Jordan algebra, symmetric cones and interior point methods*, Research Report, University of Notre Dame, 1995.
- [2] M. Iri, *A proof of the polynomiality of the Iri-Imai method*, Journal of Complexity 9, 269 – 290, 1993.
- [3] M. Iri and H. Imai, *A multiplicative barrier function method for linear programming*, Algorithmica 1, 455 – 482, 1986.

- [4] Yu. Nesterov, *Introductory lectures on convex programming*, Kluwer Academic Publishers, Dordrecht, 2004.
- [5] Yu. Nesterov and A. Nemirovskii, *Interior point polynomial methods in convex programming*, Studies in Applied Mathematics 13, SIAM, Philadelphia, PA, 1994.
- [6] Yu. Nesterov and M.J. Todd, *Self-scaled barriers and interior-point methods for convex programming*, Mathematics of Operations Research 22, 1 – 42, 1997.
- [7] J. Renegar, *A mathematical view of interior-point methods in convex optimization*, MPS-SIAM Series on Optimization, 2001.
- [8] J.F. Sturm, *Unifying semi-definite and second-order cone programming: symmetric cones*, Research Report, 1998.
- [9] J.F. Sturm, *Similarity and other special relations for symmetric cones*, Linear Algebra and its Applications 312, 135 – 154, 2000.
- [10] J.F. Sturm and S. Zhang, *A potential reduction method for harmonically convex programming*, Journal of Optimization Theory and Applications 84, 181 – 205, 1995.
- [11] S. Zhang, *Convergence property of the Iri-Imai algorithm for some smooth convex programming problems*, Journal of Optimization Theory and Applications 82, 121 – 138, 1994.
- [12] S. Zhang, *A new self-dual embedding method for convex programming*, Journal of Global Optimization 29, 479 – 496, 2004.