A Potential Reduction Method for Harmonically Convex Programming[1]

J. F. Sturm[2] and S. Zhang[3]

Communicated by Z.Q. Luo

1. The authors like to thank Dr. Hans Nieuwenhuis for carefully reading this paper and the anonymous referees for the worthy suggestions.

2. Student, Department of Econometrics, University of Groningen, Groningen, The Netherlands.

3. Assistant Professor, Department of Econometrics, University of Groningen, Groningen, The Netherlands.

Abstract

In this paper we introduce a potential reduction method for harmonically convex programming. We show that, if the objective function and the m constraint functions are all k-harmonically convex in the feasible set, then the number of iterations needed to find an ϵ-optimal solution is bounded by a polynomial in m, k and $\log(1/\epsilon)$. The method requires either the optimal objective value of the problem or an upper bound of the harmonic constant k as a working parameter. Moreover, we discuss the relation between the harmonic convexity condition used in this paper and some other convexity and smoothness conditions used in the literature.

Key Words: Convex programming, harmonic convexity, potential reduction methods, polynomiality.

1 Introduction

In this paper we analyze an interior point method for solving the following convex programming problem:

$$\begin{array}{lll} \text{(CP)} & \min & f(x) \\ & \text{s.t.} & g_i(x) \le 0 \ , \ i = 1, 2, \cdots, m \\ & & x \in \mathcal{R}^n \end{array}$$

where f and $g_1, g_2, \cdots, g_m$ are convex functions in $\mathcal{R}^n$.

The first interior point method for solving (CP), to the best knowledge of the authors, was discussed by Frisch (Ref. 1) in 1955. Due to numerical difficulties, his method, known as the classical logarithmic barrier method, has become outdated. In 1967, Huard introduced the so-called center method (Ref. 2). The center method solves (CP) by approximately computing analytical centers of the shrinking convex regions (containing the optimal point) given as follows

$$\{x \in \mathcal{R}^n \mid g_1(x) \le 0, g_2(x) \le 0, \cdots, g_m(x) \le 0, f(x) \le z_u\}$$

for a decreasing sequence of upper bounds z_u.

In 1986, the center method was used by Renegar (Ref. 3) to solve linear programs. In the linear programming case, Renegar showed that if the constraint $f(x) \le z_u$ is replicated m times, only $\mathcal{O}(\sqrt{m+n}L)$ center-iterations are needed, where L denotes the input length of the linear program. Jarre (Refs. 4-5), Mehrotra and Sun (Ref. 6) and Nesterov and Nemirovsky (Ref. 7) generalized the results of Renegar to classes of smooth convex programming that include quadratically constrained convex quadratic programming.

In practice, however, Renegar's so-called short step center method is hopelessly slow. To overcome this shortcoming, Jarre, Sonnevend and Stoer (Ref. 8) used a predictor corrector method in their implementation.

Another way to improve practical performance is to use a so-called large step center method and to allow line-searches along the Newton direction, as in Den Hertog (Ref. 9) and Den Hertog, Roos and Terlaky (Ref. 10). This large step center method has an $\mathcal{O}(mL)$ iteration bound for linear programming.

The same iteration bound holds for Karmarkar's well known potential reduction algorithm (Ref. 11), which is the first polynomial interior point method for linear programming. It should be noted however, that interior point methods already existed, cf. Frisch (Ref. 1), and polynomial time solvability of linear programming was already proved by Khachian (Ref. 12). In practice, the performance of the potential reduction method of Karmarkar and the large step center method are comparable.

Because of its projective nature, Karmarkar's method cannot easily be generalized to solve (CP). Nesterov and Nemirovsky (Ref. 7), however, succeeded in generalizing Gonzaga's conical projective potential reduction method (Ref. 13) to solve some convex programming problems.

Iri and Imai (Ref. 14) showed in 1986 that the exponential of the potential function for linear programming, the so-called multiplicative barrier function, is convex and can therefore be minimized using Newton's method. In 1991, Iri (Ref. 15) obtained an $\mathcal{O}(mL)$ iteration bound for this method. The results of Iri were generalized to harmonically convex programming by Zhang (Ref. 16).

In this paper we will present a potential reduction method for solving harmonically convex programming problems. The harmonic convexity condition was used by Mehrotra and Sun (Ref. 6) and was formally defined and studied by Zhang (Ref. 16). The analysis of the potential reduction method proposed in this paper is simpler than that of Iri and Imai's method, cf. Zhang (Ref. 16). In the linear programming case, the method turns out to be similar to those discussed in Gonzaga (Ref. 17) and Ye (Ref. 18). It solves linear programs in $\mathcal{O}(mL)$ iterations, and if a primal-dual potential function is incorporated, then the complexity is reduced to $\mathcal{O}(\sqrt{m}L)$, cf. (Refs. 18-19). We will prove in this paper that for the potential reduction method a similar iteration bound holds true for harmonically convex programming. The complexity bound we obtain for this new method is better than the one obtained in (Ref. 16) in terms of the harmonic constant. We then investigate and compare the conditions imposed on the functions by various authors to guarantee the convergence of their interior point methods. We show that on a compact set, the class of harmonically convex functions is larger than the classes of smooth convex functions analyzed in Jarre (Refs.4-5), Nesterov and Nemirovsky (Ref. 7), Den Hertog (Ref. 9) and Den

Hertog, Roos and Terlaky (Ref. 10).

This paper is organized as follows. In the next section we will introduce some notions, such as the harmonic convexity and ρ-feasibility, and some assumptions and elementary lemmas necessary for our analysis. In Section 3 we will discuss the potential reduction method for harmonically convex programs. To make the analysis simpler we will first discuss in Section 3 the case where the objective is a linear function. Two alternative algorithms are presented in the same section treating separately two cases: 1) the optimal value of the problem is known; 2) an upper bound of the harmonic constant k is known. The analysis is carried on further in Section 4 for the nonlinear (but harmonically convex) objective case. The same convergence results as presented in Section 3 are obtained in Section 4. In Section 5 we compare the harmonic convexity condition with the Relative Lipschitz condition and the self-concordancy condition, which are commonly used in the literature. Finally we conclude the paper in Section 6.

2 Preliminaries

The feasible region for the convex program (CP) is denoted by

$$\mathcal{F} := \{x \in \mathcal{R}^n \mid g_i(x) \leq 0 \text{ for } i = 1, 2, \cdots, m\}.$$

With respect to the interior of $\mathcal{F}$, defined by $\mathcal{F}^0 := \{x \mid g_i(x) < 0 \text{ for } i = 1, 2, \cdots, m\}$, we make the following assumption.

Assumption 2.1. Slater Condition. $\mathcal{F}^0 \neq \emptyset$.

In addition, we assume that the functions f and $g_1, g_2, \cdots, g_m$ are twice continuously differentiable and satisfy the so-called harmonic convexity condition, defined below. First we note that two matrices A and B are said to satisfy $A \preceq B$ if and only if $B - A$ is a positive semi-definite matrix.

Definition 2.1 A twice differentiable convex function $g : \mathcal{X} \rightarrow \mathcal{R}$ is said to be k-harmonically convex on its convex domain $\mathcal{X}$, where $k \geq 0$ is a constant, iff for any $x, y \in \mathcal{X}$ it holds that

$$\nabla^2 g(x) \preceq k \nabla^2 g(y).$$

The constant k is called a harmonic constant.

Clearly, linear functions are 0-harmonically convex and convex quadratic functions are 1-harmonically convex. Note that if a function is k-harmonically convex, then it is also k'-harmonically convex for any $k' \geq k$. Hence, we do not lose generality by assuming $k \geq 1$ in the sequel. A harmonically convex programming problem (CP) is defined as follows.

Definition 2.2 If the functions f and $g_1, g_2, \cdots, g_m$ are twice differentiable k-harmonically convex functions on $\mathcal{F}$ then (CP) is called a harmonically convex programming problem.

Now we introduce an important concept in our analysis which we call ρ-feasibility.

Definition 2.3 For an interior solution $x \in \mathcal{F}^0$, a point $x+v$ is called ρ-feasible with respect to x iff

$$g_i(x+v) \leq (1-\rho)g_i(x) \text{ for } i = 1, 2, \cdots, m$$

where $\rho \in [0, 1]$.

For a convex program (CP) satisfying the Slater condition (Assumption 2.1), the analytic center of $\mathcal{F}$ is defined to be the subset of $\mathcal{F}$ on which the center function $\chi(x)$,

$$\chi(x) := -\sum_{i=1}^{m} \log(-g_i(x))$$

is minimized (see Huard (Ref. 2)). Note that χ is convex, because its Hessian matrix

$$\nabla^2\chi(x) = \sum_{i=1}^{m} \nabla^2 g_i(x)/[-g_i(x)] + \sum_{i=1}^{m} \nabla g_i(x)\nabla g_i(x)^{\mathrm{T}}/g_i(x)^2$$

is positive semi-definite.

The following lemma gives a sufficient condition for a step to be feasible.

Lemma 2.1. Sufficient Condition for Feasibility. Let $v \in \mathcal{R}^n$, $x \in \mathcal{F}^0$ and $\rho \in [0, 1)$. Let $\beta := 2\rho/(1+\sqrt{1+2k\rho})$. If $\sqrt{v^{\mathrm{T}}\nabla^2\chi(x)v} \leq \beta$ then $x+v$ is ρ-feasible with respect to x.

Proof. See also Corollary 3.3 of Mehrotra and Sun (Ref. 6) for the case $\beta = 0.5/\sqrt{k}$.

Let $r > 0$ be such that $x + rv \in \mathcal{F}$ and let $i \in \{1, 2, \cdots, m\}$. By Taylor's formula there exists a t, $0 < t < r$ such that

$$\begin{aligned} g_i(x + rv) &= g_i(x) + r\nabla g_i(x)^{\mathrm{T}} v + (r^2/2) v^{\mathrm{T}} \nabla^2 g_i(x + tv) v \\ &= (-1 + r\nabla g_i(x)^{\mathrm{T}} v/[-g_i(x)] \\ &\quad + (r^2/2) v^{\mathrm{T}} \nabla^2 g_i(x + tv) v/[-g_i(x)])(-g_i(x)). \end{aligned}$$

Since

$$\begin{aligned} (\nabla g_i(x)^{\mathrm{T}} v/[-g_i(x)])^2 &= v^{\mathrm{T}} \nabla g_i(x) \nabla g_i(x)^{\mathrm{T}} v/g_i(x)^2 \\ &\leq v^{\mathrm{T}} \nabla^2 \chi(x) v \end{aligned}$$

and, using the harmonic convexity of g_i,

$$v^{\mathrm{T}} \nabla^2 g_i(x + tv) v/[-g_i(x)] \leq k v^{\mathrm{T}} \nabla^2 \chi(x) v,$$

we get

$$\begin{aligned} g_i(x + rv)/[-g_i(x)] &\leq -1 + r\sqrt{v^{\mathrm{T}} \nabla^2 \chi(x) v} + (r^2 k/2) v^{\mathrm{T}} \nabla^2 \chi(x) v \\ &\leq -1 + r\beta + r^2 \beta^2 k/2 \\ &= -1 + r\rho + r(r-1)k\rho^2/(1 + k\rho + \sqrt{1 + 2k\rho}). \quad (1) \end{aligned}$$

Now let r' be such that $g_i(x + r'v) = 0$, i.e. $x + r'v$ lies on the boundary of $\mathcal{F}$. It follows from (1) that

$$0 = g_i(x + r'v)/[-g_i(x)] = -1 + r'\rho + r'(r' - 1)k\rho^2/(1 + k\rho + \sqrt{1 + 2k\rho}),$$

which implies $r' > 1$. Taking $r = 1 < r'$ in (1), we obtain

$$g_i(x + v) \leq (1 - \rho) g_i(x).$$

Since the above relation holds for any $i \in \{1, 2, \cdots, m\}$, the lemma is proved. $\Box$

Corollary 2.1 If $\mathcal{F}$ is bounded, then the Hessian of the center function, $\nabla^2\chi(x)$ is positive definite for any $x \in \mathcal{F}^0$.

Proof. Suppose to the contrary that there exist $x \in \mathcal{F}^0$ and $v \neq 0$ such that $v^{\mathrm{T}}\nabla^2\chi(x)v = 0$ then, using Lemma 2.1, $x + tv \in \mathcal{F}$ for any $t \in \mathcal{R}^n$ which contradicts the boundedness of $\mathcal{F}$. $\Box$

To simplify our analysis we make the following assumption.

Assumption 2.2 $\mathcal{F}$ is bounded.

Because the constraint functions $g_i(x)$ are convex and continuous, it follows that $\mathcal{F}$ is a closed set. Therefore Assumption 2.2 implies that $\mathcal{F}$ is a compact set.

Now suppose a lower bound z^0 for the optimal value of (CP) is known, i.e. for all $x \in \mathcal{F}$ it holds that

$$f(x) \geq z^0.$$

We will assume in this paper that an initial interior point of $\mathcal{F}$ is available. To introduce a proper stopping criterion we use the following definition.

Definition 2.4 A solution $x \in \mathcal{F}$ is said to be an ϵ-optimal solution to (CP) with $\epsilon > 0$ iff for any $y \in \mathcal{F}$

$$f(x) \leq f(y) + \epsilon.$$

We show in this paper that if either the harmonic constant k, or the optimal objective value is known, the scaling potential reduction method discussed in the next section finds a so-called ϵ-optimal solution with the number of main iterations bounded by a polynomial of m, $\log(1/\epsilon)$ and k.

3 Scaling Potential Reduction Method

A potential function for (CP) is defined by

$$\phi(x, z) := q \log(f(x) - z) + \chi(x)$$

where $q > 0$, and z is a lower bound for the optimal value of (CP), i.e. for any $x \in \mathcal{F}^0$ it holds that $f(x) \geq z$.

By a potential reduction method one tries to construct iteratively the interior solution x' based on the current interior point $x \in \mathcal{F}^0$, in such a way that a certain reduction can be guaranteed in the value of the potential function ϕ.

The potential function was introduced by Karmarkar (Ref. 11) in 1984 for solving linear programming problems. The idea of using such a potential function, however, can be generalized to convex programming. For the convex program (CP), because the center function χ is bounded from below by its value in the center of $\mathcal{F}$, reducing the potential function ϕ by at least a constant amount at each iteration will decrease the gap $f(x) - z$ in the long run also. This can be seen from the following lemma.

Lemma 3.1 Suppose that Assumptions 2.1 and 2.2 are satisfied, and an initial interior point x^0, and an initial lower bound z^0 are available. Let

$$N := \max_{1\le i\le m} \max_{x\in\mathcal{F}}(-g_i(x))$$

and let $\delta > 0$ be a fixed quantity. If an algorithm reduces the potential function ϕ(x,z) by at least δ at each iteration, then the algorithm can be used to find an ϵ-optimal solution to (CP) in $\mathcal{O}((\phi(x^0, z^0) + q\log(1/\epsilon) + m\log(N))/\delta)$ iterations.

Proof. See Theorem 3.4 in Zhang (Ref. 16); the details are omitted here. □

Notice again that in this paper we assume that the initial interior solution (x^0,z^0) is available.

Like most of the interior point methods, the iterative procedure of the method discussed in this paper is based on reducing the potential value by searching along a certain direction from the current point. Therefore a good search direction is crucial for potential reduction methods. Now we shall describe the search direction used in our method.

It is easy to compute that

$$\nabla\phi(x, z) = q\nabla f(x)/[f(x) - z] + \nabla\chi(x)$$

$$\nabla\chi(x) = \sum_{i=1}^{m} \nabla g_i(x)/[-g_i(x)]$$

and

$$\nabla^2\phi(x,z) = q\nabla^2 f(x)/[f(x)-z] - q\nabla f(x)\nabla f(x)^{\mathrm{T}}/[f(x)-z]^2 + \nabla^2\chi(x).$$

By letting

$$S(x) := q\nabla^2 f(x)/[f(x)-z] + \nabla^2\chi(x)$$

it follows that

$$\nabla^2\phi(x,z) \preceq S(x).$$

In our approach we take the search direction given by $d := -S(x)^{-1}\nabla\phi(x,z)$.

Clearly, since $S(x)$ is always positive definite (cf. Corollary 2.1), the direction d is well defined and is indeed a descent direction for ϕ.

Now we introduce a relevant quantity h defined by

$$h := \sqrt{\nabla\phi(x,z)^{\mathrm{T}} S(x)^{-1}\nabla\phi(x,z)}.$$

As we will see later, this quantity plays an important role in the analysis.

To simplify the analysis, we first investigate the case where f is a linear function, i.e.

$$f(x) = c^{\mathrm{T}}x.$$

Note that actually this assumption can be made without loss of generality, since, in order to solve (CP), it suffices to solve

$$\min\{x_{n+1} \mid (x_1,\cdots,x_n) \in \mathcal{F}, f(x_1,\cdots,x_n) \le x_{n+1}\}$$

which clearly has a linear objective.

In this case, the search direction will be given by $d := -\nabla^2\chi(x)^{-1}\nabla\phi(x,z)$. Remark here that for linear programming, the matrix $\nabla^2\chi(x)^{-1}$ is also used in the dual affine scaling method to obtain a search direction. We first prove the following lemma.

Lemma 3.2 Let $\rho \in (0,1)$, $\beta := 2\rho/(1+\sqrt{1+2k\rho})$ and $t \leq \beta/h$. Then

$$d^{\mathrm{T}}\nabla^2\phi(x+td,z)d \leq [(2+k(1-\rho)+2k^2\beta^2)/(1-\rho)^2]h^2$$

Proof. Using Lemma 2.1 we obtain

$$\begin{aligned} d^{\mathrm{T}}\nabla^2\phi(x+td,z)d &\leq \sum_{i=1}^{m} d^{\mathrm{T}}\nabla^2 g_i(x+td)d/[-g_i(x+td)] \\ &\quad +\sum_{i=1}^{m}\nabla g_i(x+td)^{\mathrm{T}}d)^2/g_i(x+td)^2 \\ &\leq [k/(1-\rho)]\sum_{i=1}^{m} d^{\mathrm{T}}\nabla^2 g_i(x)d/[-g_i(x)] \\ &\quad +(1-\rho)^{-2}\sum_{i=1}^{m}(\nabla g_i(x+td)^{\mathrm{T}}d)^2/g_i(x)^2. \end{aligned} \tag{2}$$

By the mean value theorem we know that there exist some $\bar{t}_i$, $0 \leq \bar{t}_i \leq t$, such that

$$\begin{aligned} \sum_{i=1}^{m}(\nabla g_i(x+td)^{\mathrm{T}}d)^2/g_i(x)^2 &= \sum_{i=1}^{m}[\nabla g_i(x)^{\mathrm{T}}d + td^{\mathrm{T}}\nabla^2 g_i(x+\bar{t}_i d)d]^2/g_i(x)^2 \\ &\leq 2\sum_{i=1}^{m}[\nabla g_i(x)^{\mathrm{T}}d]^2/g_i(x)^2 \\ &\quad +2k^2t^2h^2\sum_{i=1}^{m} d^{\mathrm{T}}\nabla^2 g_i(x)d/[-g_i(x)] \\ &\leq 2h^2 + 2k^2\beta^2h^2. \end{aligned}$$

Combining (2) with the above inequality the lemma follows. □

The reduction in potential function value can be obtained by doing a line-search in $\mathcal{F}^0$ along the direction d. It can be estimated using Taylor's formula as it is shown in the following lemma. To simplify the notation, if not specified we let $\phi(x)$ denote $\phi(x, z)$ in the sequel.

Lemma 3.3. Reduction in ϕ by Updating x. Let $\rho = (1/8k)\min(1, h)$, $\beta := 2\rho/(1 + \sqrt{1 + 2k\rho})$ and $t := \beta/h$. Then

$$\phi(x) - \phi(x + td) > (0.06/k)\min(h, h^2)$$

Proof. Let $\gamma := 2 + k(1 - \rho) + 2k^2\beta^2$.

Using Taylor's formula and Lemma 3.2, there exists a $\bar{t}$, $0 \le \bar{t} \le t$ such that

$$\begin{aligned}
\phi(x) - \phi(x + td) &= -t\nabla\phi(x)^{\mathrm{T}}d - (t^2/2)d^{\mathrm{T}}\nabla^2\phi(x + \bar{t}d)d \\
&\ge t\nabla\phi(x)^{\mathrm{T}}\nabla^2\chi(x)^{-1}\nabla\phi(x) - \gamma t^2h^2/2(1 - \rho)^2 \\
&= th^2 - \gamma t^2h^2/2(1 - \rho)^2 = [2h - \gamma\beta/(1 - \rho)^2]\beta/2. \quad (3)
\end{aligned}$$

Since

$$\beta = 2\rho/(1 + \sqrt{1 + 2k\rho}) = [0.25/k(1 + \sqrt{1 + 0.25\min(1, h)})]\min(1, h)$$

we have

$$(1/12k)\min(1, h) \le \beta \le (1/8k)\min(1, h). \quad (4)$$

Using $k \geq 1$ it follows that

$$\begin{aligned}\gamma\beta &= (2 + k(1-\rho) + 2k^2\beta^2)\beta \\ &\leq [(2 + k + 1/32)/8k] \min(1, h) \\ &\leq [(3 + 1/32)/8] \min(1, h) < 0.5(1-\rho)^2 \min(1, h).\end{aligned}$$

Combining the above inequality with (3) and (4) and noting that $3/48 > 0.06$, the lemma follows. □

The previous lemma showed that at each iteration, the potential function is reduced by at least a constant amount of $0.03/k$, as long as $h^2 \geq 1/2$. We proceed by showing that if $h^2 < 1/2$, the lower bound z can be updated, which will cause a reduction in potential function value by even a considerably larger amount than $0.03/k$.

Let x^* be an optimal solution of (CP) and let $w_1, w_2, \cdots, w_m$ be

$$w_i := 1 - \nabla g_i(x)^{\mathrm{T}}(x^* - x)/[-g_i(x)] \text{ for } i = 1, \cdots, m.$$

Further using the mean value theorem, there exist $t_i \in [0, 1]$, $i = 1, \cdots, m$, such that

$$\begin{aligned}0 &\geq g_i(x^*)/[-g_i(x)] \\ &= -w_i + (1/2)(x^* - x)^{\mathrm{T}}\nabla^2 g_i(x + t_i(x^* - x))(x^* - x)/[-g_i(x)] \\ &\geq -w_i + (1/2k)(x^* - x)^{\mathrm{T}}\nabla^2 g_i(x)(x^* - x)/[-g_i(x)]\end{aligned}$$

and therefore

$$0 \leq (x^* - x)^{\mathrm{T}}\nabla^2 g_i(x)(x^* - x)/[-g_i(x)] \leq 2kw_i. \tag{5}$$

Observe, by the Cauchy-Schwartz inequality, that

$$h = \sqrt{\nabla\phi(x)^{\mathrm{T}}\nabla^2\chi(x)^{-1}\nabla\phi(x)} \geq -\nabla\phi(x)^{\mathrm{T}}(x^* - x)/\sqrt{(x^* - x)^{\mathrm{T}}\nabla^2\chi(x)(x^* - x)}. \tag{6}$$

Moreover, notice that

$$-\nabla\phi(x)^{\mathrm{T}}(x^* - x) = -qc^{\mathrm{T}}(x^* - x)/(c^{\mathrm{T}}x - z) - m + \sum_{i=1}^{m} w_i \tag{7}$$

and using (5),

$$\begin{aligned}(x^* - x)^{\mathrm{T}}\nabla^2\chi(x)(x^* - x) &\leq 2k\sum_{i=1}^{m} w_i + \sum_{i=1}^{m}(1 - w_i)^2 \\ &\leq m + 2k\sum_{i=1}^{m} w_i + \sum_{i=1}^{m} w_i^2. \end{aligned} \tag{8}$$

Observing that

$$\begin{aligned}(\sqrt{m} + k - 1 + \sum_{i=1}^{m} w_i)^2 &= (\sqrt{m} + k - 1)^2 + 2(\sqrt{m} + k - 1)\sum_{i=1}^{m} w_i + (\sum_{i=1}^{m} w_i)^2 \\ &\geq m + 2k\sum_{i=1}^{m} w_i + \sum_{i=1}^{m} w_i^2,\end{aligned}$$

where we used $m \geq 1$ and $k \geq 1$, it follows from (7), (8) and (6) that

$$h \geq \frac{q(c^{\mathrm{T}}x - c^{\mathrm{T}}x^*)/(c^{\mathrm{T}}x - z) - m + \sum_{i=1}^{m} w_i}{\sqrt{m} + k - 1 + \sum_{i=1}^{m} w_i}. \tag{9}$$

Similarly, we notice that

$$\begin{aligned}k(\sqrt{m} + \sum_{i=1}^{m} w_i)^2 &= km + 2k\sqrt{m}\sum_{i=1}^{m} w_i + k(\sum_{i=1}^{m} w_i)^2 \\ &\geq m + 2k\sum_{i=1}^{m} w_i + \sum_{i=1}^{m} w_i^2.\end{aligned}$$

Therefore, it also follows from (7), (8) and (6) that

$$h \geq \frac{q(c^{\mathrm{T}}x - c^{\mathrm{T}}x^*)/(c^{\mathrm{T}}x - z) - m + \sum_{i=1}^{m} w_i}{\sqrt{k}(\sqrt{m} + \sum_{i=1}^{m} w_i)}. \tag{10}$$

If we use the optimal value $z^* := c^{\mathrm{T}}x^*$ as a lower bound, then we obtain some lower bounds on h as given in Lemma 3.4.

Lemma 3.4 If z $= c^{\mathrm{T}}x^*$ then

$$h \geq \min[1, (q - m)/(\sqrt{m} + k - 1)]$$

and

$$h \geq (1/\sqrt{k}) \min(1, (q - m)/\sqrt{m}).$$

Proof. The two inequalities follow directly from (9) and (10) by replacing z with $c^{\mathrm{T}}x^*$. (See also Zhang (Ref. 16), or Iri (Ref. 15).) □

Similar to Iri and Imai's algorithm for solving convex programs as discussed in Zhang (Ref. 16), we first assume that the optimal value of (CP) z^* ($z^* = c^{\mathrm{T}}x^*$) is known. This results in the following algorithm.

Algorithm 3.1. Scaling Potential Reduction Algorithm. The algorithm requires f, g_1, $\cdots$, g_m, x^0, z^* and ϵ as input.

Step 0 Set $q = m + \sqrt{m}$, $i = 0$.

Step 1 If $f(x^i) - z^* \leq \epsilon$, stop.

Step 2 Set $d = -S(x^i)^{-1}\nabla\phi(x^i, z^*)$.

Step 3 Compute x^{i+1} by minimizing $\phi(x, z^*)$ along d.

Step 4 Set $i = i + 1$ and return to Step 1.

Concerning the complexity of Algorithm 1, we have the following result.

Theorem 3.1 For a linear objective function $f(x) = c^{\mathrm{T}}x$, Algorithm 1 finds an ϵ-optimal solution of (CP) in $\mathcal{O}((\phi(x^0, z^*) + m\log(N/\epsilon))k^2)$ iterations.

Proof. The theorem follows immediately by using Lemma 3.4, Lemma 3.3 and Lemma 3.1. □

Remark 3.1 For linear programming, a 2^{-2L}-optimal solution is sufficient to obtain a vertex optimal solution in $\mathcal{O}(m^3)$ elementary operations by a purification procedure, where L is the input size of the problem. Hence, the iteration bound is of order $\mathcal{O}(mL)$, which is the same as in Karmarkar (Ref. 11) and Iri (Ref. 15).

Remark 3.2 For linear programming, $S(x)$ is known as the dual affine scaling matrix, and the corresponding direction d is one of the directions discussed in Gonzaga (Ref. 17).

Remark 3.3 If we would fix $q = m + \sqrt{m} + k$ in Step 0, then we could have an $\mathcal{O}((\phi(x^0, z^0) + (m + k)\log(1/\epsilon) + m\log(N))k)$ iteration bound for Algorithm 3.1, which is better than the iteration bound in Theorem 3.1 in terms of k. The advantage of Algorithm 3.1 however is that the harmonic constant k is not assumed to be known.

Remark 3.4 Because of the harmonic convexity, the Hessian matrices computed at different points in the domain are related in terms of the harmonic constant. Hence, in order to prove Theorem 3.1, the Hessian matrices $\nabla^2 g_1, \nabla^2 g_2, \cdots, \nabla^2 g_m$ entering the scaling matrix $S(x)$ do not need to be computed at the current interior point x. They can once be computed at an arbitrary point in $\mathcal{F}^0$ and be used in later iterations.

By assuming $z = z^*$, we were able to bound h from below by some positive quantity in Lemma 3.4. For fixed $z < z^*$, the potential function attains its minimum in $\mathcal{F}^0$, so in this case, h can be as small as zero. We now fix $q = m + \sqrt{m} + k$ so that it follows from (9) that for $h < 1$,

$$\begin{aligned} h &\geq 1 - \frac{m + \sqrt{m} + k - 1 - q(c^{\mathrm{T}}x - z^*)/(c^{\mathrm{T}}x - z)}{\sqrt{m} + k - 1 + \sum_{i=1}^{m} w_i} \\ &\geq 1 - \frac{m + \sqrt{m} + k - 1 - q(c^{\mathrm{T}}x - z^*)/(c^{\mathrm{T}}x - z)}{\sqrt{m} + k - 1} \\ &= q(c^{\mathrm{T}}x - z^*)/(\sqrt{m} + k - 1)(c^{\mathrm{T}}x - z) - m/(\sqrt{m} + k - 1). \end{aligned}$$

This implies that

$$z^* \geq c^{\mathrm{T}}x - (1/q)[m + h(\sqrt{m} + k - 1)](c^{\mathrm{T}}x - z). \tag{11}$$

We can use this relation to update the lower bound z.

Lemma 3.5. Reduction after Updating *z*. If $h < 1$ and $q = m + \sqrt{m} + k$ and

$$z' := c^{\mathrm{T}}x - (1/q)[m + h(\sqrt{m} + k - 1)](c^{\mathrm{T}}x - z)$$

then

$$\phi(x, z) - \phi(x, z') \geq (1 - h)(\sqrt{m} + k - 1)$$

In the proof of Lemma 3.5 we will use the following well-known lemma due to Karmarkar.

Lemma 3.6 If $| x |< 1$ then

$$-x \geq \log(1 - x) \geq -x - x^2/2(1- | x |)$$

Proof. See, e.g., Karmarkar (Ref. 11) or Freund (Ref. 20). □

Proof of Lemma 3.5

From (11) it follows that

$$\begin{aligned}
\phi(x, z') - \phi(x, z) &= q \log((c^{\mathrm{T}}x - z')/(c^{\mathrm{T}}x - z)) \\
&= q \log([m + h(\sqrt{m} + k - 1)]/q) \\
&\leq q \log(1 - (1 - h)(\sqrt{m} + k - 1)/q) \\
&\leq -(1 - h)(\sqrt{m} + k - 1)
\end{aligned}$$

where the last inequality follows from Lemma 3.6. The lemma is proved. □

Now we are ready to present the next potential reduction algorithm for solving (CP) without assuming the optimal value to be known (but assuming a known upper bound on the harmonic constant).

Algorithm 3.2. Scaling Potential Reduction Algorithm. The algorithm requires f, g_1, $\cdots$, g_m, x^0, z^0, k and ϵ as input.

Step 0 Set $q = m + \sqrt{m} + k$, $i = 0$.

Step 1 If $f(x^i) - z^i \leq \epsilon$, stop.

Step 2 Set $d = -S(x^i)^{-1}\nabla\phi(x^i, z^i)$ and $h = \sqrt{-\nabla\phi(x^i, z^i)^{\mathrm{T}} d}$.

Step 3 If $h^2 < 1/2$ go to Step 6. Otherwise go to Step 4.

Step 4 Compute x^{i+1} by minimizing $\phi(x, z^i)$ along d. Set $z^{i+1} = z^i$.

Step 5 Set $i = i + 1$ and return to Step 1.

Step 6 Set $z^{i+1} = f(x^i) - (1/q)[m + h(\sqrt{m} + k - 1)](f(x^i) - z^i)$. Set $x^{i+1} = x^i$.

Step 7 Set $i = i + 1$ and return to Step 1.

Concerning the complexity of Algorithm 2, we have the following theorem.

Theorem 3.2 For a linear objective function $f(x) = c^{\mathrm{T}}x$ Algorithm 2 finds an ϵ-optimal solution of (CP) in $\mathcal{O}((\phi(x^0, z^0) + (m + k)\log(1/\epsilon) + m\log N)k)$ iterations.

Proof. The theorem follows from Lemma 3.5, Lemma 3.3 and Lemma 3.1.

□

Remark 3.5 It is easy to modify Algorithm 3.2 by introducing primal variables in the potential functions to solve linear programs in $\mathcal{O}(\sqrt{m}L)$ iterations. See, e.g., Ye (Ref. 18).

4 Nonlinear Objective Case

In this section we will discuss the convergence of the potential reduction method applied directly to a convex program with a nonlinear objective function (without adding constraints or variables).

Recall that the scaling matrix is defined by

$$S(x) := q\nabla^2 f(x)/[f(x) - z] + \nabla^2 \chi(x),$$

the search-direction is $d := -S(x)^{-1}\nabla\phi(x, z)$, and the quantity h is defined by

$$h := \sqrt{\nabla\phi(x, z)^T S(x)^{-1} \nabla\phi(x, z)}.$$

Similar to Lemma 3.2 we now show the following lemma.

Lemma 4.1 For a fixed ρ such that $0 < \rho \leq 0.145$, let $\beta := 2\rho/(1 + \sqrt{1 + 2k\rho})$, $t \leq \beta/h$. Then

$$d^T \nabla^2 \phi(x + td)d \leq 2[(1 + k(1 - \rho) + k^2\beta^2)/(1 - \rho)^2]h^2.$$

Before proving Lemma 4.1 we first give a bound on the possible improvement in objective function value along the direction d with a certain step size.

Lemma 4.2 For a fixed ρ such that $0 < \rho \leq 0.145$, let $\beta \leq 2\rho/(1+\sqrt{1+2k\rho})$ and $t \leq \beta/h$. Then

$$f(x+td) - z \geq [(1-\rho)/2](f(x)-z).$$

Proof. Let $\beta_1 := 2/(1+\sqrt{1+2k})$. By Lemma 2.1 we know that $x+(\beta_1/h)d$ is a feasible solution.

By the subgradient inequality we have

$$f(x+td) \geq f(x) + t\nabla f(x)^{\mathrm{T}} d. \tag{12}$$

Using Taylor's formula we have

$$(\beta_1/\beta)t\nabla f(x)^{\mathrm{T}} d = f(x+t(\beta_1/\beta)d) - f(x) - (t^2\beta_1^2/2\beta^2)d^{\mathrm{T}}\nabla^2 f(x+\bar{t}d)d \tag{13}$$

where $0 \leq \bar{t} \leq t$.

Noticing that $f(x+t(\beta_1/\beta)d) \geq z$ and

$$\begin{aligned} d^{\mathrm{T}}\nabla^2 f(x+\bar{t}d)d &\leq kd^{\mathrm{T}}\nabla^2 f(x)d \\ &\leq kd^{\mathrm{T}}S(x)d[f(x)-z]/q \\ &= kh^2[f(x)-z]/q \end{aligned}$$

we obtain from the above inequalities and (12) and (13) that

$$\begin{aligned} f(x+td) &\geq f(x) + (\beta/\beta_1)[z - f(x) - k\beta_1^2(f(x)-z)/2q] \\ &\geq f(x) - \sqrt{\rho}(1+1/q)(f(x)-z), \end{aligned}$$

where the last inequality follows from $\beta/\beta_1 = \rho(1+\sqrt{1+2k})/(1+\sqrt{1+2k\rho}) \leq \sqrt{\rho}$ and $k\beta_1^2/2 \leq 1$.

Noting that that the smallest root of $y^2 - 3y + 1$ is larger than $\sqrt{0.145}$, we get for $\rho \leq 0.145$:

$$2 - 3\sqrt{\rho} \geq 1 - \rho$$

and therefore using $q \geq 2$,

$$f(x+td) \geq f(x) - (3/2)\sqrt{\rho}(f(x) - z) \geq z + (1/2)(1-\rho)(f(x) - z).$$

□

Proof of Lemma 4.1

Using Lemma 4.2, analogously to the proof of Lemma 3.2, we have

$$\begin{aligned}
d^{\mathrm{T}}\nabla^2\phi(x+td)d &\leq qd^{\mathrm{T}}\nabla^2 f(x+td)d/[f(x+td)-z] \\
&\quad + \sum_{i=1}^{m} d^{\mathrm{T}}\nabla^2 g_i(x+td)d/[-g_i(x+td)] \\
&\quad + \sum_{i=1}^{m} (\nabla g_i(x+td)^{\mathrm{T}} d)^2/g_i(x+td)^2 \\
&\leq [2k/(1-\rho)]d^{\mathrm{T}}(q\nabla^2 f(x)/[f(x)-z] + \sum_{i=1}^{m} \nabla^2 g_i(x)/[-g_i(x)])d \\
&\quad + 2[(1+k^2\beta^2)/(1-\rho)^2]h^2 \\
&\leq 2[1 + k(1-\rho) + k^2\beta^2]h^2/(1-\rho)^2,
\end{aligned}$$

so the lemma is proved. □

The reduction in potential function value by doing a line-search in $\mathcal{F}^0$ along d can now easily be estimated.

Lemma 4.3. Reduction in ϕ by Updating x. Let $\rho = (1/8k)\min(1,h)$, $\beta := 2\rho/(1+\sqrt{1+2k\rho})$ and $t := \beta/h$. Then

$$\phi(x) - \phi(x+td) > (0.05/k)\min(1,h^2).$$

Proof.

Let $\gamma := 2 + 2k(1-\rho) + 2k^2\beta^2$.

Analogously to the proof of Lemma 3.3, using $k \geq 1$ and (4) we obtain

$$\begin{aligned} \gamma\beta &= (2 + 2k(1-\rho) + 2k^2\beta^2)\beta \\ &\leq [(2 + 2k + 1/32)/8k]\min(1,h) \\ &\leq [(4 + 1/32)/8]\min(1,h) < 0.66(1-\rho)^2\min(1,h) \end{aligned}$$

combining this with (3) and (4), and noting that $(2-0.66)/24 > 0.05$, the lemma follows. □

Recall that x^* denotes an optimal solution of (CP) and that

$$w_i := 1 - \nabla g_i(x)^{\mathrm{T}}(x^* - x)/[-g_i(x)] \text{ for } i = 1, \cdots, m.$$

In addition we define now the nonnegative quantity w_0 as follows

$$w_0 := [f(x^*) - z]/[f(x) - z] - (1 + \nabla f(x)^{\mathrm{T}}(x^* - x)/[f(x) - z]).$$

Using the mean value theorem and the k-harmonic convexity of f we have the following inequalities (where $t \in [0,1]$)

$$\begin{aligned} 0 &= w_0 - (1/2)(x^* - x)^{\mathrm{T}}\nabla^2 f(x + t(x^* - x))(x^* - x)/[f(x) - z] \\ &\leq w_0 - (1/2k)(x^* - x)^{\mathrm{T}}\nabla^2 f(x)(x^* - x)/[f(x) - z] \end{aligned}$$

and therefore

$$(x^* - x)^{\mathrm{T}} \nabla^2 f(x)(x^* - x)/[f(x) - z] \leq 2kw_0. \tag{14}$$

Because

$$\begin{aligned} -\nabla\phi(x)(x^* - x) &= qw_0 - q[f(x^*) - z]/[f(x) - z] + q - m + \sum_{i=1}^{m} w_i \\ &= q[f(x) - f(x^*)/[f(x) - z] + qw_0 - m + \sum_{i=1}^{m} w_i \end{aligned} \tag{15}$$

and using (5) and (14),

$$\begin{aligned} (x^* - x)^{\mathrm{T}} S(x)(x^* - x) &\leq 2qkw_0 + 2k\sum_{i=1}^{m} w_i + \sum_{i=1}^{m}(1 - w_i)^2 \\ &\leq m + 2k(qw_0 + \sum_{i=1}^{m} w_i) + \sum_{i=1}^{m} w_i^2 \\ &\leq mk + 2k\sqrt{m}(qw_0 + \sum_{i=1}^{m} w_i) + k(qw_0 + \sum_{i=1}^{m} w_i)^2 \\ &= k(\sqrt{m} + qw_0 + \sum_{i=1}^{m} w_i)^2, \end{aligned} \tag{16}$$

it follows using (6) that

$$h \geq \frac{q - m + qw_0 + \sum_{i=1}^{m} w_i - q[f(x^*) - z]/[f(x) - z]}{\sqrt{k}(\sqrt{m} + qw_0 + \sum_{i=1}^{m} w_i)}.$$

Based on this lower bound on h, similar to Theorem 3.1, we have the following result.

Theorem 4.1 In $\mathcal{O}((\phi(x^0, z^*) + m\log(N/\epsilon))k^2)$ iterations, Algorithm 3.1 finds an ϵ-optimal solution of (CP) .

Proof. See the proof for Theorem 3.1. □

From (16) it follows that

$$\begin{aligned}(x^* - x)^{\mathrm{T}} S(x)(x^* - x) &\leq m + 2k(qw_0 + \sum_{i=1}^{m} w_i) + \sum_{i=1}^{m} w_i^2 \\ &\leq (\sqrt{m} + k - 1) + 2(\sqrt{m} + k - 1)(qw_0 + \sum_{i=1}^{m} w_i) \\ &\quad +(qw_0 + \sum_{i=1}^{m} w_i)^2 \\ &= (\sqrt{m} + k - 1 + qw_0 + \sum_{i=1}^{m} w_i)^2.\end{aligned}$$

Using the same arguments as for the linear objective case, we thus obtain from (6), (15) and (16) that

$$h \geq \frac{q[f(x) - f(x^*)/[f(x) - z] + qw_0 - m + \sum_{i=1}^{m} w_i}{\sqrt{m} + k - 1 + qw_0 + \sum_{i=1}^{m} w_i}. \tag{17}$$

If we fix $q = m + \sqrt{m} + k$ and if $h < 1$ then it follows from (17) that

$$f(x^*) \geq f(x) - (1/q)[m + h(\sqrt{m} + k - 1)](f(x) - z).$$

This yields the following theorem.

Theorem 4.2 Algorithm 2 finds an ϵ-optimal solution of (CP) in $\mathcal{O}((\phi(x^0, z^0) + (m + k)\log(1/\epsilon) + m \log N)k)$ iterations.

Proof. Cf. the proof for Theorem 3.2. □

5 Harmonic Convexity

In this section we discuss the convexity condition underlying our analysis in comparison with those used by some other authors.

Recall that a convex function is called strongly convex (or uniformly convex) on a compact domain iff the Hessian matrix exists and its eigenvalues are bounded by some positive constants from both below and above (Ref. 21). It is easy to see that strongly convex functions are also harmonically convex. Moreover, we have the following characterization of harmonic convexity.

Theorem 5.1. Characterization of Harmonic Convexity. Let $\mathcal{X} \subseteq \mathcal{R}^n$. A function $g : \mathcal{X} \to \mathcal{R}$ is harmonically convex iff there exists a regular matrix T and a partition $x = \begin{bmatrix} x' \\ x'' \end{bmatrix}$ such that $\overline{g}(x) := g(Tx)$ is strongly convex in x' and linear in x''.

Proof. See Zhang (Ref. 16). □

For a harmonically convex function, the harmonic constant k can be computed as the ratio of the largest and the smallest nonzero eigenvalue of its Hessian on its domain.

Instead of Assumption 2.2, Nesterov and Nemirovsky (Ref. 7), Jarre (Ref. 5) and Den Hertog (Ref. 9) required that the barrier functions $-\log(-g_i(x))$ are all self-concordant on $\mathcal{F}^0, i = 1, \cdots, m$. The concept of self-concordancy described below was introduced by Nesterov and Nemirovsky (Ref. 7) in 1989.

Definition 5.1 Let $\kappa \geq 0$ and $\mathcal{X} \subseteq \mathcal{R}^n$. A three-times continuously differentiable convex function $\varphi : \mathcal{X} \to \mathcal{R}$ is said to be κ-self-concordant on its domain

$\mathcal{X}$ iff for any $x \in \mathcal{X}$ and $v \in \mathcal{R}^n$ it holds that

$$\left|(d^3/dt^3)\varphi(x+tv)\mid_{t=0}\right| \le 2\kappa[(d^2/dt^2)\varphi(x+tv)\mid_{t=0}]^{3/2}.$$

On a compact set, self-concordancy is more stringent than harmonic-convexity, as it is shown in the following theorem.

Theorem 5.2 If a function φ is κ-self-concordant on a compact set $\mathcal{X}$, then it is also k-harmonically convex on $\mathcal{X}$ with the harmonic constant

$$k := \exp\left(2\kappa \max_{x,y\in\mathcal{X},0\le t\le 1} \sqrt{(y-x)^{\mathrm{T}}\nabla^2\varphi(tx+(1-t)y)(y-x)}\right).$$

Before proving Theorem 5.2 we introduce the following lemma.

Lemma 5.1 If φ is κ-self-concordant then

$$\left|(d^3/dt_1dt_2dt_3)\varphi(x+t_1u+t_2v+t_3w)\mid_{t_1=t_2=t_3=0}\right| \le$$
$$2\kappa\sqrt{u^{\mathrm{T}}\nabla^2\varphi(x)u}\sqrt{v^{\mathrm{T}}\nabla^2\varphi(x)v}\sqrt{w^{\mathrm{T}}\nabla^2\varphi(x)w}.$$

Proof. See Lemma A.2 of Jarre (Ref. 5). □

Proof of Theorem 5.2

Let $\Omega := \max_{x,y\in\mathcal{X},0\le t\le 1}\sqrt{(y-x)^{\mathrm{T}}\nabla^2\varphi(tx+(1-t)y)(y-x)}$. By continuity of the Hessian matrix $\nabla^2\varphi$ we know that Ω is a finite number. Let $x, y \in \mathcal{X}$, $v \in \mathcal{R}^n$ and $t \in [0,1]$. Define $d := y-x$ and $r(t) := v^{\mathrm{T}}\nabla^2\varphi(x+td)v$. Note that

$$r(t) = (d^2/ds^2)\varphi(x+sv+td)\mid_{s=0}$$

and therefore, using Lemma 5.1, it follows that

$$(d/d\tau)r(t+\tau)|_{\tau=0} \leq 2\kappa r(t)\sqrt{d^{\mathrm{T}}\nabla^2\varphi(x+td)d} \leq 2\kappa\Omega r(t).$$

This implies that

$$\log(v^{\mathrm{T}}\nabla^2\varphi(y)v/v^{\mathrm{T}}\nabla^2\varphi(x)v) = \log(r(1)) - \log(r(0)) \leq 2\kappa\Omega = \log(k),$$

and so the theorem is proved. □

A constraint function g_1 having a self-concordant barrier is not necessarily convex. But, whenever there exists some $\gamma > 0$ for such a constraint function such that its barrier $-\log(\gamma - g_1(x))$ is self-concordant on $\mathcal{F}$, then it follows from Theorem 5.2 that

$$\overline{g}_1(x) := -\log((\gamma - g_1(x))/\gamma)$$

is harmonically convex on $\mathcal{F}$. Clearly, in this case the constraint function $g_1(x)$ can be replaced by $\overline{g}_1(x)$ without affecting the feasible region of the convex program (CP).

Another convexity and smoothness assumption on (CP), different from harmonic convexity, was used by Jarre (Ref. 4) and Den Hertog, Roos and Terlaky (Ref. 10). That assumption says that the constraint functions $g_i(x)$, $i = 1, \cdots, m$ and the objective function $f(x)$ should satisfy the so-called relative Lipschitz condition on $\mathcal{F}^0$. The relative Lipschitz condition, given below, was introduced by Jarre (Ref. 4) in 1988.

Definition 5.2 A twice continuously differentiable function g_1 is said to be M-relative Lipschitz convex on its domain $\mathcal{X} \subseteq \{x \in \mathcal{R}^n \mid g(x) < 0\}$ iff there exists a nonnegative constant M such that for any $x, y \in \mathcal{X}$

$$-MH(x,y)\nabla^2 g_1(y) \preceq \nabla^2 g_1(x) - \nabla^2 g_1(y) \preceq MH(x,y)\nabla^2 g_1(y)$$

where

$$H(x,y) := \sqrt{-(y-x)^{\mathrm{T}}\nabla^2 \log(-g_1(y))(y-x)}.$$

On a compact set, the relative Lipschitz condition is clearly more stringent than the harmonic convexity condition, as it is shown in the following theorem.

Theorem 5.3 If a function g_1 is M-relative Lipschitz convex on a compact set $\mathcal{X}$, then it is also k-harmonically convex on $\mathcal{X}$ with the harmonic constant $k := 1 + M \max_{x,y \in \mathcal{X}} H(x,y)$.

Proof. The theorem follows immediately after noticing the relation

$$\nabla^2 g_1(x) \preceq (1 + MH(x,y))\nabla^2 g_1(y)$$

for all x and $y \in \mathcal{X}$. □

Moreover, Jarre (Ref. 5) showed that if a three times continuously differentiable function g_1 is M-relative Lipschitz on $\mathcal{F}^0$, then its barrier function $-\log(-g_1(x))$ is $(1+M)$-self-concordant.

As the following example shows, a harmonically convex function may be neither relative Lipschitz convex nor self-concordant.

Example 5.1 Consider the function $g : \mathcal{R} \to \mathcal{R}$, $g(x) := x^2(1 + \sqrt{| x |}) - 2$ for which

$$(d^2/dx^2)g(x) = 2 + (15/4)\sqrt{| x |}$$

Clearly, $g \in C^2$. Moreover, it is harmonically convex on $\{x \mid g(x) < 0\} = (-1, 1)$ and d^3g/dx^3 is continuous on $(0, 1)$. However, $\lim_{x \downarrow 0}(d^3/dx^3)g(x) = \infty$ and therefore neither g nor $-\log(-g(x))$ is self concordant on $(0, 1)$. This in turn implies that the relative Lipschitz condition cannot be fulfilled on $(0, 1)$.

Unfortunately, these three convexity and smoothness conditions do not even include some simple convex polynomials, as it is shown in the following example.

Example 5.2 Consider the function $g : \mathcal{R} \to \mathcal{R}$, $g(x) := x^4 - 1$ for which

$$(d^2/dx^2)g(x) = 12x^2$$

Clearly, g is convex, but not harmonically convex on $\{x \mid g(x) < 0\} = (-1, 1)$. Note that if g would be the only constraint, then the Hessian matrix of the center function, $\nabla^2\chi(x)$, could be singular in the interior of the feasible region. Therefore, the algorithm proposed in this paper could break down if the harmonic convexity condition is not fulfilled.

In (Ref. 19) it is shown how to modify the scaling potential reduction method in order to solve convex programs with functions that are convex, but not harmonically convex.

It is still an open question whether it is possible that on a bounded but *not* closed set a function is self-concordant or relative Lipschitz convex, but not harmonically convex.

6 Conclusions

Compared with other papers on the interior point approach for convex programming, the convergence and complexity analysis in this paper is rather simple. Yet, we showed that on a compact domain the condition of harmonic convexity is less restrictive than the smoothness conditions introduced by Jarre (Ref. 4) and Nesterov and Nemirovsky (Ref. 7).

Jarre (Ref. 4) and Den Hertog, Roos and Terlaky (Ref. 10) require in the algorithms an available relative Lipschitz constant as a working parameter. The same is true with the self-concordancy constant in Nesterov and Nemirovsky (Ref. 7), Jarre (Ref. 5) and Den Hertog (Ref. 9), and the harmonic constant in Mehrotra and Sun (Ref. 6) and Sun and Qi (Ref. 22). In Zhang (Ref. 16), by supposing the harmonic convexity condition, instead of requiring a known harmonic constant in the algorithm, the optimal objective value is assumed to be known. In this paper, we leave the two alternatives open: either the optimal objective value or an harmonic constant is known.

The Iri-Imai method analyzed in Zhang (Ref. 16) is closely related to the scaling gradient method proposed in this paper. Instead of using S as scaling matrix, the Hessian matrix of $\Phi(x) := \exp(\phi(x, z^*))$, where z^* is the optimal objective value, is used in Iri and Imai's method, so that the search direction is actually the Newton direction for minimizing $\Phi(x)$. The matrix $S(x)$ and the Hessian of $\Phi(x)$ satisfy the following relation for $q > m + 1$:

$$\nabla^2\Phi(x)/\Phi(x) = \nabla^2\phi(x, z^*) + \nabla\phi(x, z^*)\nabla\phi(x, z^*)^{\mathrm{T}} \succeq [(q - m - 1)/(q - 1)]S(x).$$

If we let $\overline{S}(x) := [(q - 1)/(q - m - 1)\Phi(x)]\nabla^2\Phi(x)$ as the scaling matrix, Lemma 4.1 and Lemma 4.3 remain true. Moreover, we obtain in this way the Iri-Imai direction. Now using Theorem 3.1 in Zhang (Ref. 16) which states that

$$\Phi(x)\nabla\phi(x, z^*)^{\mathrm{T}}\nabla^2\Phi(x)^{-1}\nabla\phi(x, z^*) \geq 1/(2k + 2)$$

it follows for $q = 2m + 1$, in which case $\overline{S}(x) \succeq (1/2)S(x)$ and $\overline{S}(x)^{-1} = (1/2)\Phi(x)\nabla^2\Phi(x)^{-1}$, that Iri and Imai's method solves (CP) in $\mathcal{O}((\phi(x^0, z^0) + m\log(1/\epsilon) + m\log(N))k^2)$ iterations, using a similar argument as in Theorem 3.1. This bound is of order $\mathcal{O}(k^2)$ smaller than the one given in Zhang (Ref. 16).

By assuming the harmonic convexity, Mehrotra and Sun (Ref. 6) obtained an $\mathcal{O}(\sqrt{m}k^6\log(1/\epsilon))$ iteration bound for their short step path following method. Very recently, Sun and Qi (Ref. 22) proposed an algorithm with the same iteration bound as in (Ref. 6) but without requiring second order differentiability. Instead, they use a condition, which is similar to the harmonic convexity condition, on generalized Hessians. This alleviation can be made for our analysis too. Remark here that one can use a second order Taylor expansion theorem employing the generalized Hessian (cf. Proposition 2.1 in Sun and Qi (Ref. 22)). In order to keep (6) valid in this case, however, the generalized Hessians used in the scaling matrix should be positive semi-definite.

References

1. FRISCH, K.R., *The Logarithmic Potential Method of Convex Programming*, Memorandum, Institute of Economics, Oslo, Norway, 1955.

2. HUARD, P., *Resolution of Mathematical Programming with Nonlinear Constraints by the Method of Centres*, Nonlinear Programming, Edited by J. Abadie, North Holland Publishing Company, Amsterdam, The Netherlands, pp. 207-219, 1967.

3. RENEGAR, J., *A Polynomial-Time Algorithm, based on Newton's Method, for Linear Programming*, Mathematical Programming, Vol. 40, pp.59-93, 1988.

4. JARRE, F., *On the Method of Centres for solving Smooth Convex Programs*, Lecture Notes in Mathematics, Edited by S. Dolecki, Springer Verlag, Berlin, Germany, Vol. 1405, pp. 69-85, 1989.

5. JARRE, F., *Interior-Point Methods for Convex Programming*, Applied Mathematics and Optimization, Vol. 26, pp.287-311, 1992.

6. MEHROTRA, S. and SUN, J., *An Interior Point Algorithm for solving Smooth Convex Programs based on Newton's method*, Contemporary Mathematics, Vol. 114, pp. 265-284, 1990.

7. NESTEROV, J.E. and NEMIROVSKY, A.S., *Self-Concordant Functions and Polynomial Time Methods in Convex Programming*, Central Economical and Mathematical Institute, USSR Academy of Sciences, Moscow, USSR, 1989.

8. JARRE, F., SONNEVEND, G. and STOER, J., *An Implementation of the Method of Analytic Centers*, Analysis and Optimization of Systems, Edited by A. Bensoussan and J.L. Lions, Lecture Notes in Control and Information Sciences, Springer Verlag, Berlin, Germany, Vol. 111, pp.297-307, 1988.

9. DEN HERTOG, D., *Interior Point Approach to Linear, Quadratic and Convex Programming. Algorithms and Complexity*, Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands, 1992.

10. DEN HERTOG, D., ROOS, C., TERLAKY, T., *A Large-Step Analytic Center Method for a Class of Smooth Convex Programming Problems*, SIAM Journal on Optimization, Vol. 2, pp.55-70, 1992.

11. KARMARKAR, N.K., *A New Polynomial-Time Algorithm for Linear Programming*, Combinatorica, Vol. 4, pp.373-395, 1984.

12. KHACHIAN, L.G., *A Polynomial Algorithm in Linear Programming*, Soviet Mathematics Doklady, Vol. 20, pp. 191-194, 1979.

13. GONZAGA, C.C., *Conical Projection Algorithms for Linear Programming*, Mathematical Programming, Vol. 43, pp.151-173, 1989.

14. IRI, M., IMAI, H., *A Multiplicative Barrier Function Method for Linear Programming*, Algorithmica, Vol. 1, pp.455-482, 1986.

15. IRI, M., *A Proof of the Polynomiality of the Iri-Imai Method*, Journal of Complexity, Vol. 9, pp.269-290, 1993.

16. ZHANG, S., *Convergence Properties of the Iri-Imai Algorithm for a Class of Smooth Convex Programming Problems*, Journal of Optimization Theory and Applications (to appear).

17. GONZAGA, C.C., *Interior Point Algorithms for Linear Programming Problems with Inequality Constraints*, Mathematical Programming, Vol. 52, pp.209-226, 1991.

18. YE, Y, *An $\mathcal{O}(n^3L)$ Potential Reduction Algorithm for Linear Programming*, Mathematical Programming, Vol. 50, pp.239-258, 1991.

19. STURM, J.F., *The Scaling Potential Reduction Method*, M.S. Thesis, University of Groningen, 1993.

20. FREUND, R., *Polynomial-time Algorithms for Linear Programming based only on Primal Scaling and Projected Gradients of a Potential Function*, Mathematical Programming, Vol. 51, pp.203-222, 1991.

21. AVRIEL, M., DIEWERT, W.E., SCHAIBLE, S., and ZANG, I., *Generalized Concavity*, Plenum Press, New York, New York, 1988.

22. SUN, J., QI, L., *An Interior Point Algorithm of* $\mathcal{O}(\sqrt{m} \mid \log \epsilon \mid)$ *Iterations for* C^1*-Convex Programming*, Mathematical Programming, Vol. 57, pp.239-257, 1992.