CORE DISCUSSION PAPER

2008/57

Primal-dual Interior-Point Methods with Asymmetric Barriers

Yu. Nesterov *

September 2008

Abstract

In this paper we develop several polynomial-time interior-point methods (IPM) for solving nonlinear primal-dual conic optimization problem. We assume that the barriers for the primal and the dual cone are not conjugate. This broken symmetry does not allow to apply the standard primal-dual IPM. However, we show that in this situation it is also possible to develop very efficient optimization methods, which satisfy all desired qualities, including the infeasible-start features. Our technique is based on asymmetric primal-dual barrier augmented by squared residual of the primal-dual linear system.

Keywords: conic optimization, self-concordant barriers, polynomial-time methods, interior-point methods, path-following methods, potential-reduction methods, infeasible start.

*Center for Operations Research and Econometrics (CORE), Catholic University of Louvain (UCL), 34 voie du Roman Pays, 1348 Louvain-la-Neuve, Belgium; e-mail: nesterov@core.ucl.ac.be.

The research results presented in this paper have been supported by a grant "Action de recherche concertè ARC 04/09-315" from the "Direction de la recherche scientifique - Communautè française de Belgique". The scientific responsibility rests with its author(s).

1 Introduction

Motivation. The primal-dual long-step infeasible-start interior-point methods (IPM) are widely recognized as the most powerful polynomial-time schemes for solving nonlinear conic optimization problems (see [3, 6, 9, 11]). However, the practical implementations of these methods are known only for symmetric cones (e.g. [1, 7, 10]). The main reason for this is that only for symmetric cones (which are self-dual), the analytical form of the primal and the dual barriers coincide.

For other nonlinear cones, the situation is not so simple. Very often, we are able to construct a self-concordant barrier $F(x)$ for the primal cone K. However, in order to have a perfect symmetry in the primal-dual schemes, we need to choose the barrier $F_D(s)$ for the *dual cone*

$$K^* \;=\; \{s:\ \langle s, x\rangle \geq 0\ \forall x \in K\}$$

equal to the *Fenchel transform* of the primal barrier $F_P(x)$:

$$F_D(s) \;=\; F_P^*(s) \;=\; \max_x [-\langle s, x\rangle - F(x)]. \tag{1.1}$$

Thus, we need to assume that this (complex) nonlinear maximization problem is easily solvable, preferably in a closed form. This is not true even in the simplest situations. One of the most important examples is provided by the three-dimensional *power cone*:

$$\begin{aligned} K_\alpha &= \left\{x \in R^2_+ \times R:\ \left(x^{(1)}\right)^\alpha \left(x^{(2)}\right)^{1-\alpha} \geq \left|x^{(3)}\right|\right\}, \quad \alpha \in (0,1), \\ F_{P,\alpha}(x) &= -\ln\left[\left(x^{(1)}\right)^{2\alpha}\left(x^{(2)}\right)^{2(1-\alpha)} - \left(x^{(3)}\right)^2\right] - \ln x^{(1)} - \ln x^{(2)}. \end{aligned} \tag{1.2}$$

Then, for the dual cone we have

$$\begin{aligned} K^*_\alpha &= \left\{s \in R^2_+ \times R:\ \left(\tfrac{s^{(1)}}{\alpha}\right)^\alpha \left(\tfrac{s^{(2)}}{1-\alpha}\right)^{1-\alpha} \geq \left|s^{(3)}\right|\right\}, \\ F_{D,\alpha}(s) &= -\ln\left[\left(\tfrac{s^{(1)}}{\alpha}\right)^{2\alpha}\left(\tfrac{s^{(2)}}{1-\alpha}\right)^{2(1-\alpha)} - \left(s^{(3)}\right)^2\right] - \ln s^{(1)} - \ln s^{(2)}. \end{aligned} \tag{1.3}$$

Another interesting example is the conic hull of the epigraph of the exponent:

$$\begin{aligned} K_{P,e} &= \left\{x \in R \times R^2_+:\ x^{(1)} \geq x^{(2)} \ln \tfrac{x^{(2)}}{x^{(3)}}\right\}, \\ F_{P,e}(x) &= -\ln\left(x^{(1)} - x^{(2)} \ln \tfrac{x^{(2)}}{x^{(3)}}\right) - \ln x^{(2)} - \ln x^{(3)}, \\ K^*_{P,e} &= \left\{s \in R_+ \times R \times R_+:\ s^{(1)} + s^{(2)} \geq s^{(1)} \ln \tfrac{s^{(1)}}{s^{(3)}}\right\}, \\ F_{D,e}(x) &= -\ln\left(s^{(1)} + s^{(2)} - s^{(1)} \ln \tfrac{s^{(1)}}{s^{(3)}}\right) - \ln s^{(1)} - \ln s^{(3)}. \end{aligned} \tag{1.4}$$

In both examples, the corresponding Fenchel transform of the primal barrier is not computable in the closed form.

The main goal of this paper is to study our abilities in constructing the efficient primal-dual methods when the primal and the dual cones are endowed with *independent* self-concordant barriers[1)]. We propose several new potential-reduction and path-following strategies. Our technique is based on asymmetric primal-dual barrier augmented by squared residual of the primal-dual linear system. From the theoretical point of view, the most efficient among our schemes is an infeasible-start long-step path-following method. Its complexity estimate coincides with the best estimate known for this problem class. Note that even for symmetric cones, our schemes are new.

Contents. The paper is organized as follows. In Section 2, we introduce the primal-dual conic problem. The main difference from the standard setting is that the primal and dual barriers are independent. We introduce a measure for the quality of the primal-dual barrier with respect to some strictly feasible reference point. In Section 3 we study the efficiency of a feasible-start potential reduction method as applied to a homogenized primal-dual setting. We prove that it generates an ϵ-solution to our problem in $O(\nu_\Psi \ln \frac{1}{\epsilon})$ iterations, where ν_Ψ is the parameter of the primal-dual barrier. In Section 4 we propose two new infeasible-start potential-reduction methods. The first method is based on the quadratic potential, and the second one minimizes a homogeneous potential function. The potential functions of both methods include a squared residual of the homogenized linear system of the primal-dual conic problem. Thus, these methods generate minimization sequences inside full-dimensional convex cones, maintaining no linear constraints during the minimization process.

Finally, in Section 5 we study a new full-dimensional path-following scheme based on the quadratic potential function. We prove its complexity estimate of the order $O(\sqrt{\nu_\Psi} \ln \frac{1}{\epsilon})$ iterations. This scheme admits a cheap long-step *predictor-corrector* strategy, which can significantly improve its practical performance. In the last Section 6 we present and discuss the computational results.

Notation. In what follows, we denote by E (or other capital letter) a finite dimensional linear vector space. The dual space (the space of linear functions on E) is denoted by E^*. We denote the value of function $s \in E^*$ on $x \in E$ by $\langle s, x \rangle$. This notation is used for all linear spaces employed in the paper. Thus, the actual meaning of $\langle \cdot, \cdot \rangle$ can be understood from the spaces of the arguments. For example, for the coordinate space $E = R^n$ with $E^* = R^n$, this notation has the following sense:

$$\langle s, x \rangle \quad = \quad \sum_{i=1}^{n} s^{(i)} x^{(i)}, \quad x, s \in R^n.$$

For a linear operator $A : E \to H^*$ we denote by $A^* : H \to E^*$ the adjoint operator:

$$\langle Ax, y \rangle \quad = \quad \langle A^* y, x \rangle, \quad x \in E, \; y \in H.$$

For a positive-semidefinte self-adjoint linear operator $B : E \to E^*$ we denote

$$\|x\|_B \quad = \quad \langle Bx, x \rangle^{1/2}, \quad x \in E.$$

If B is positive definite, then we can define the corresponding dual norm:

$$\|s\|_B^* \quad = \quad \langle s, B^{-1} s \rangle^{1/2}, \quad s \in E^*.$$

[1)] For an alternative way of fighting with primal-dual asymmetry the reader can consult [2] or [5].

For the reader convenience, we recall some facts related to the self-concordant barriers. A convex cone $K \subset E$ is called *normal* if it is a closed convex pointed cone with nonempty interior. If K is normal, then the dual cone K^* is also normal. We say that function $F(x)$ is a *normal barrier* for cone K if it is *self-concordant* and logarithmically homogeneous:

$$F(\tau x) \;=\; F(x) - \nu_F \ln \tau, \quad x \in \operatorname{int} K,\ \tau > 0, \tag{1.5}$$

where ν_F is the *parameter* of the barrier (thus, we have ν_F-*normal barrier*). One of consequences of identity (1.5) is as follows:

$$\langle \nabla F(x), x \rangle \;=\; -\nu_F, \quad x \in \operatorname{int} K. \tag{1.6}$$

Recall that for self-concordant barriers the following inclusion holds:

$$\{u \in K :\ \langle \nabla^2 F(x)(u-x), u-x \rangle \le 1\} \;\subset\; K. \tag{1.7}$$

Another useful result is the *Theorem on Recession Direction*:

$$\langle \nabla^2 F(x)u, u \rangle^{1/2} \;\le\; \langle -\nabla F(x), u \rangle, \quad x \in \operatorname{int} K,\ u \in K. \tag{1.8}$$

If F is a ν_F-normal barrier for K, then its conjugate function $F^*(s)$ is a normal barrier for K^* with the same value of the parameter. The primal and dual normal barriers satisfy several useful relations. The most important one is as follows:

$$F(x) + F^*(s) \;\ge\; -\nu_F - \nu_F \ln \tfrac{\langle s,x \rangle}{\nu_F}, \tag{1.9}$$

where $x \in \operatorname{rint} K$ and $s \in \operatorname{rint} K^*$, and equality is attained if and only if $s = -\nabla F(x)$.

2 Asymmetric barrier for primal-dual cone

Consider the following conic optimization problem:

$$\text{Find } f_P^* \;=\; \min_x \{ \langle c, x \rangle,\ Ax = b,\ x \in K \}, \tag{2.1}$$

where $c \in E^*$, A is a nondegenerate operator from E to H^*, $b \in H^*$, and $K \subset E$ is a normal cone. The problem dual to (2.1) can be written as follows:

$$\text{Find } f_D^* \;=\; \max_{s,y} \{ \langle b, y \rangle,\ s + A^* y = c,\ s \in K^* \}. \tag{2.2}$$

We always assume that both primal and dual problems are *strictly feasible*:

$$\exists\, x_0 \in \operatorname{int} K,\ s_0 \in \operatorname{int} K^*,\ y_0 \in H :\quad Ax = b,\ s + A^* y_0 = c. \tag{2.3}$$

In this case, the problems (2.1), (2.2) have zero duality gap:

$$f_P^* \;=\; f_D^* \;\stackrel{\text{def}}{=}\; f^*.$$

Assume that we are able to compute values and derivatives of a ν_{F_P}-normal barrier $F_P(x)$ for the primal cone K. Then for problems (2.1), (2.2), we can introduce a *primal-dual central path* $(x(t), s(t), y(t)) \in K \times K^* \times H$:

$$Ax(t) = b, \quad s(t) + A^* y(t) = c, \quad s(t) = -\tfrac{1}{t} \nabla F_P(x(t)), \quad t > 0. \tag{2.4}$$

In view of assumption (2.3), this trajectory is well defined (see, for example, [3]).

Further, we assume that for the dual cone K^*, there is also a computable ν_{F_D}-normal barrier $F_D(s)$. Of course, the best choice would be to take

$$F_D(s) \;=\; F_P^*(s), \quad s \in \operatorname{int} K^*. \tag{2.5}$$

However, very often this is impossible since such a choice presumes a closed-from solvability of the nonlinear unconstrained optimization problem in (1.1). Therefore, in this paper we study the situation when F_D is an arbitrary normal barrier for the dual cone.

Thus, in our situation we have to treat the primal-dual problem

$$\begin{array}{rl} & \min\limits_{x,s,y} \langle c, x\rangle - \langle b, y\rangle \\ \text{s.t.} & Ax = b,\ s + A^*y = c, \\ & x \in K,\ s \in K^*, \end{array} \tag{2.6}$$

by the following normal barrier

$$\Psi(z) \;\equiv\; \Psi(x,s) \;=\; F_P(x) + F_D(s), \quad \nu_\Psi \;=\; \nu_{F_P} + \nu_{F_D},$$

$$z \;=\; (x,s) \in \operatorname{int} K \;\times\; \operatorname{int} K^* \;\stackrel{\text{def}}{=}\; \operatorname{int} \bar{K}.$$

Unfortunately, this barrier does not fit the standard problem settings addressed in the literature. Let us look, for example, at *generalized Karmarkar method* (see Section 4.3 in [6]). This technique can be applied to the following problem:

$$\text{Find } f^* \;=\; \min_w \{\langle d, w\rangle :\ \langle e, w\rangle = 1,\ w \in \hat{K}\}, \tag{2.7}$$

where $\hat{K}$ is a normal cone endowed with a θ-normal barrier. It is assumed that $f^* = 0$ and the feasible set of problem (2.7) is *bounded*.

Note that the primal-dual problem (2.6) can be rewritten in the format (2.7). Indeed, let us define

$$\hat{K} \;\equiv\; \operatorname{rint} \{w = (x,s,y,\tau) :\ Ax = \tau b,\ s + A^*y = \tau c,\ x \in K, s \in K^*, \tau \geq 0\},$$

$$\langle e, w\rangle \;\equiv\; \tau.$$

Then the corresponding optimization problem has zero optimal value. However, in this case the feasible set of problem (2.7) is *never bounded*. To see that, it is enough to consider the central path (2.4) as $t \to \infty$. It is interesting that the primal-dual problem (2.6) can be solved by a variant of Karmarkar method (see Section 5.1 in [3]), but only if we choose the barrier for the dual cone as in (2.5).

In the sequel, we are going to show that the barrier $\Psi(z)$ can be used for constructing different polynomial-time interior-point schemes. Our analysis is based on the following measure of the quality of this barrier *with respect to* the reference point $z_0 = (x_0, s_0)$ (for simplicity, we take the same point as in condition (2.3)). Let us define

$$\begin{array}{rcl} \kappa \;\equiv\; \kappa(z_0) & = & F_P(x_0) + F_D(s_0) \\ & & - \min\limits_{x,s} \{F_P(x) + F_D(s) :\ \langle s_0, x\rangle + \langle s, x_0\rangle = 2\langle s_0, x_0\rangle\}. \end{array} \tag{2.8}$$

Note that z_0 satisfies the equality constraint in the minimization problem in (2.8). Therefore $\kappa(z_0) \geq 0$.

Lemma 1 *For any $z \in \bar{K}$ we have*

$$\Psi(z) \;\geq\; \Psi(z_0) - \kappa(z_0) - \nu_\Psi \ln \tfrac{\langle s_0,x\rangle + \langle s,x_0\rangle}{2\langle s_0,x_0\rangle}. \tag{2.9}$$

Proof:
Indeed,

$$\begin{aligned} \Psi(z) &\overset{(1.5)}{=} F_P\left(\tfrac{2\langle s_0,x_0\rangle \cdot x}{\langle s_0,x\rangle + \langle s,x_0\rangle}\right) + F_D\left(\tfrac{2\langle s_0,x_0\rangle \cdot s}{\langle s_0,x\rangle + \langle s,x_0\rangle}\right) - \nu_\Psi \ln \tfrac{\langle s_0,x\rangle + \langle s,x_0\rangle}{2\langle s_0,x_0\rangle} \\ &\overset{(2.8)}{\geq} \Psi(z_0) - \kappa(z_0) - \nu_\Psi \ln \tfrac{\langle s_0,x\rangle + \langle s,x_0\rangle}{2\langle s_0,x_0\rangle}. \end{aligned}$$

□

3 Feasible-start potential-reduction IPM

Consider the following potential function:

$$\begin{aligned} \phi(w) &= (\nu_\Psi + 1)\ln[\langle c,x\rangle - \langle b,y\rangle] + F_P(x) + F_D(s) - \ln\tau, \quad w \in \operatorname{int}\mathcal{F}, \\ \mathcal{F} &= \{w = (x,s,y,\tau) \in K \times K^* \times H \times R_+ :\; Ax = \tau b,\; s + A^*y = \tau c\}. \end{aligned} \tag{3.1}$$

For any $w = (x,s,y,\tau) \in \operatorname{int}\mathcal{F}$ we have

$$0 \;<\; \langle s,x\rangle \;=\; \langle \tau c - A^*y, x\rangle \;=\; \tau\langle c,x\rangle - \langle Ax,y\rangle \;=\; \tau\cdot[\langle c,x\rangle - \langle b,y\rangle]. \tag{3.2}$$

Therefore, function ϕ is well defined.

Let us associate with $\mathcal{F}$ a full-dimensional convex cone $\hat{K}$. This means that we define a linear transformation $\mathcal{A}(u)$ such that

$$u \in \hat{K} \;\Leftrightarrow\; w = \mathcal{A}(u) \in \mathcal{F}.$$

Denote by $f(u)$ the restriction of the barrier $\Psi(z) - \ln\tau$ onto $\hat{K}$. Clearly, f is a $(\nu_\Psi + 1)$-self-concordant barrier for $\hat{K}$. Finally, denote by $\langle d,u\rangle$ the corresponding restriction of the linear function $\langle c,x\rangle - \langle b,y\rangle$. Then, in the new notation, we have the following potential function:

$$\varphi(u) \;=\; (\nu_\Psi + 1)\ln\langle d,u\rangle + f(u).$$

Moreover, we know that

$$\min_u\{\langle d,u\rangle :\; u \in \hat{K}\} \;=\; 0. \tag{3.3}$$

Denote by $u^* \in \hat{K}$ the point with $\langle d,u^*\rangle = 0$.

Let us fix some $u \in \hat{K}$. In our analysis, we will use the following convex upper approximation of function φ:

$$l_u(v) \;=\; (\nu_\Psi + 1)\cdot\left(\ln\langle d,u\rangle + \tfrac{\langle d,v-u\rangle}{\langle d,u\rangle}\right) \;+\; f(v).$$

Note that l_u is a self-concordant function, $l_u(u) = \varphi_\kappa(u)$, and $\nabla l(u) = \nabla\varphi_\kappa(u)$. Moreover,

$$\varphi(v) \;\leq\; l_u(v), \quad \forall\, v \in \hat{K}. \tag{3.4}$$

Lemma 2 *For any $u \in \hat{K}$, we have*

$$\|\nabla\varphi(u)\|^*_{\nabla^2 f(u)} \;\geq\; 1. \tag{3.5}$$

Proof:
Assume that $\|\nabla\varphi(u)\|^*_{\nabla^2 f(u)} < 1$. Then, the self-concordant function $l_u(v)$ attains its minimum at some point $v_* \in \operatorname{int}\hat{K}$ (see, for example, Theorem 4.1.11 in [4]):

$$\nabla l_u(v_*) \;=\; (\nu_\Psi + 1)\cdot\tfrac{d}{\langle d,u\rangle} \;+\; \nabla f(v_*) \;=\; 0.$$

Multiplying this equality by u^*, we obtain

$$\langle \nabla f(v_*), u^*\rangle \;=\; 0.$$

However, this is impossible since $u^* \in \hat{K}$, and $-\nabla f(v_*) \in \operatorname{int}\hat{K}^*$. □

Inequality (3.5) implies that one step of the Damped Newton method,

$$u_+ \;=\; u - \tfrac{[\nabla^2 f(u)]^{-1}\nabla\varphi(u)}{1+\|\nabla\varphi(u)\|^*_{\nabla^2 f(u)}}, \tag{3.6}$$

decreases the value of our potential function by an absolute constant:

$$\varphi(u_+) \;\leq\; l_u(u_+) \;\leq\; \varphi(u) - \omega, \tag{3.7}$$

where $\omega = 1 - \ln 2$ (see, for example, Theorem 4.1.12 in [4]). It remains to show that any point w with a small value of the potential φ can be transformed to a feasible solution of the primal-dual problem (2.6) with a small duality gap.

Consider $(x,s) \in \operatorname{int}\bar{K}$. Then

$$\begin{aligned}
\langle s_0, x\rangle + \langle s, x_0\rangle &\overset{(2.3)}{=} \langle c - A^*y_0, x\rangle + \langle s + A^*y - \tau c, x_0\rangle + \langle \tau c - A^*y, x_0\rangle \\
&\overset{(2.3)}{=} \langle c, x\rangle - \langle b, y\rangle + \langle \tau b - Ax, y_0\rangle + \langle s + A^*y - \tau c, x_0\rangle \\
&\quad + \tau[\langle c, x_0\rangle - \langle b, y_0\rangle].
\end{aligned} \tag{3.8}$$

In particular, for $(x,s) \in \operatorname{int}\hat{K}$, we have

$$\langle s_0, x\rangle + \langle s, x_0\rangle \;\overset{(3.2)}{=}\; \langle c, x\rangle - \langle b, y\rangle + \tau\langle s_0, x_0\rangle. \tag{3.9}$$

Therefore, for any $u \in \operatorname{int}\hat{K}$ with $(x,s,y,\tau) = \mathcal{A}(u)$ we have

$$\begin{aligned}
\varphi(u) &= (\nu_\Psi + 1)\ln[\langle c, x\rangle - \langle b, y\rangle] + F_P(x) + F_D(s) - \ln\tau \\
&\overset{(2.9)}{\geq} (\nu_\Psi + 1)\ln[\langle c, x\rangle - \langle b, y\rangle] + \Psi(z_0) - \kappa(z_0) - \nu_\Psi \ln\tfrac{\langle s_0,x\rangle + \langle s,x_0\rangle}{2\langle s_0,x_0\rangle} - \ln\tau \\
&\overset{(3.9)}{=} \Psi(z_0) - \kappa(z_0) + \ln\left[\langle c, \tfrac{x}{\tau}\rangle - \langle b, \tfrac{y}{\tau}\rangle\right] + \nu_\Psi \ln\tfrac{2\langle s_0,x_0\rangle\cdot[\langle c,x\rangle - \langle b,y\rangle]}{\langle c,x\rangle - \langle b,y\rangle + \tau\langle s_0,x_0\rangle}.
\end{aligned}$$

Hence,

$$\begin{aligned} \varphi(u) &\geq \Psi(z_0) - \kappa(z_0) + (\nu_\Psi + 1) \ln \tfrac{\langle c,x\rangle - \langle b,y\rangle}{\langle c,x\rangle - \langle b,y\rangle + \tau \langle s_0,x_0\rangle} \\ &\quad + (\nu_\Psi + 1) \ln \langle s_0, x_0\rangle + \nu_\Psi \ln 2. \end{aligned} \tag{3.10}$$

Now we can put all our observations together.

Theorem 1 *Let us choose $u_0 \in \hat{K}$ such that $\mathcal{A}(u_0) = w_0 \equiv (x_0, s_0, y_0, 1)$, and consider the damped Newton method*

$$u_{k+1} = u_k - \tfrac{[\nabla^2 f(u_k)]^{-1} \nabla \varphi(u_k)}{1 + \|\nabla \varphi(u_k)\|^*_{\nabla^2 f(u_k)}}, \quad k \geq 0, \tag{3.11}$$

with $(x_k, s_k, y_k, \tau_k) = \mathcal{A}(u_k)$. Then the projected point

$$\tilde{x}_k = \tfrac{x_k}{\tau_k}, \quad \tilde{s}_k = \tfrac{s_k}{\tau_k}, \quad \tilde{y}_k = \tfrac{y_k}{\tau_k}$$

is feasible for problem (2.6). Moreover, for any $k \geq 0$ we have

$$\tfrac{\langle s_0,x_0\rangle}{\langle c,\tilde{x}_k\rangle - \langle b,\tilde{y}_k\rangle} \geq 2 \exp\left\{ \tfrac{\omega \cdot k - \kappa(z_0) - \ln 2}{\nu_\Psi + 1} \right\} - 1. \tag{3.12}$$

Proof:
Note that $\varphi(u_0) = (\nu_\Psi + 1) \ln \langle s_0, x_0\rangle + \Psi(z_0)$. Hence,

$$\begin{aligned} &(\nu_\Psi + 1) \ln \langle s_0, x_0\rangle + \Psi(z_0) - k \cdot \omega \overset{(3.7)}{\geq} \varphi(u_k) \\ &\overset{(3.10)}{\geq} \Psi(z_0) - \kappa(z_0) - (\nu_\Psi + 1) \ln \left[1 + \tfrac{\langle s_0,x_0\rangle}{\langle c,\tilde{x}_k\rangle - \langle b,\tilde{y}_k\rangle} \right] \\ &+ (\nu_\Psi + 1) \ln \langle s_0, x_0\rangle + \nu_\Psi \ln 2. \end{aligned}$$

□

Note that the dependence of the estimate (3.12) on the value $\kappa(z_0)$ is quite moderate. This value is defined mainly by the local structure of the asymmetric primal-dual barrier function around the point z_0. Note that $\kappa(z_0) = 0$ if the starting point satisfies the following conditions:

$$s_0 = -\lambda \nabla F_P(x_0), \quad x_0 = -\lambda \nabla F_D(s_0)$$

with $\lambda = \frac{2}{\nu_\Psi} \langle s_0, x_0\rangle$.

4 Full-dimensional potential-reduction IPM

The potential-reduction interior-point method (3.11) has two disadvantages. Firstly, it requires a feasible point w_0 to start. Secondly, at each iteration we need to compute a displacement, which satisfies the linear equality constraints of the set $\mathcal{F}$ with a high precision. Any reasonable technique for such a computation needs projectors stored in

a matrix form. Hence, for large-scale sparse problems we can expect some additional memory limitations. In this section we describe another potential-reduction method, which is free from these shortcomings.

Note that the primal-dual problem (2.6) can be posed in the following form:

$$\begin{aligned} &\text{Find } w \in \mathcal{C} \equiv \{w = (x, s, y, \tau) \in K \times K^* \times H \times R_+\} \text{ such that} \\ &Ax - \tau b \;=\; 0, \quad s + A^* y - \tau c \;=\; 0, \quad \langle c, x\rangle - \langle b, y\rangle \;=\; 0. \end{aligned} \tag{4.1}$$

In order to measure the quality of approximate solutions to the linear system in (4.1), let us introduce two positive-definite self-adjoint operators:

$$B_H : \; H \to H^*, \quad B_E : \; E \to E^*,$$

Then we can define the convex quadratic function

$$\begin{aligned} \langle Qw, w\rangle \;&=\; \langle B_H^{-1}(Ax - \tau b), Ax - \tau b\rangle + \langle B_E^{-1}(s + A^* y - \tau c), s + A^* y - \tau c\rangle \\ &\quad + [\langle c, x\rangle - \langle b, y\rangle]^2 \;\stackrel{\text{def}}{=}\; \|w\|_Q^2, \end{aligned}$$

which vanishes at the solutions of the linear system in (4.1). Moreover, for any $w \in \mathcal{C}$ we have

$$\langle s_0, x\rangle + \langle s, x_0\rangle \;\stackrel{(3.8)}{\leq}\; \underbrace{[1 + \langle B_H y_0, y_0\rangle + \langle B_E x_0, x_0\rangle]^{1/2}}_{\stackrel{\text{def}}{=}\Omega} \cdot \|w\|_Q + \tau \cdot \delta_0, \tag{4.2}$$

where $\delta_0 \stackrel{\text{def}}{=} \langle c, x_0\rangle - \langle b, y_0\rangle \stackrel{(3.2)}{=} \langle s_0, x_0\rangle$. Denoting

$$F(w) \;=\; \Psi(z) - \ln\tau \;=\; F_P(x) + F_D(s) - \ln\tau, \quad \nu_F = \nu_\Psi + 1,$$

we can define the following *quadratic potential function*:

$$\Phi(w) \;=\; \tfrac{1}{2}\|w\|_Q^2 + F(w), \quad w \in \operatorname{int}\mathcal{C}. \tag{4.3}$$

Let us study its properties.

1. $\Phi(w)$ is a self-concordant function. Note that the quadratic function $\frac{1}{2}\|w\|_Q^2$ is strongly convex in its y-component. Since (x, s, τ)-part of the Hessian of the barrier $F(w)$ is positive definite, we conclude that the whole Hessian of the potential $\Phi(w)$ is nondegenerate at any $w \in \operatorname{int}\mathcal{C}$.

2. Potential $\Phi(w)$ has unbounded level sets. Therefore,

$$\|\nabla\Phi(z)\|^*_{\nabla^2\Phi(w)} \;\geq\; 1, \quad w \in \operatorname{int}\mathcal{C}. \tag{4.4}$$

3. We can define a *homogeneous* version of this potential:

$$\widetilde{\Phi}(w) \;=\; \min_{\lambda>0} \Phi(\lambda w) \;\leq\; \Phi(w), \quad w \in \operatorname{int}\mathcal{C}. \tag{4.5}$$

Note that

$$\Phi(\lambda w) \;=\; \tfrac{\lambda^2}{2} \cdot \|w\|_Q^2 + F(w) - \nu_F \ln\lambda.$$

Therefore, the optimal choice of λ is as follows:

$$\lambda \;=\; \lambda(w) \;\stackrel{\rm def}{=}\; \frac{\nu_F^{1/2}}{\|w\|_Q}.$$

Thus, we have a closed-form representation of the projected potential:

$$\widetilde{\Phi}(w) \;=\; \nu_F \ln \|w\|_Q + F(w) + \tfrac{\nu_F}{2}[1 - \ln \nu_F]. \tag{4.6}$$

This potential is a quasi-convex homogeneous function of degree zero.

4. For any point $w \in \operatorname{int} \mathcal{C}$, we can define

$$T(w) \;=\; \lambda(w) \cdot w \;=\; \frac{\nu_F^{1/2}}{\|w\|_Q} \cdot w.$$

Note that

$$\Phi(w) \;\geq\; \Phi(T(w)). \tag{4.7}$$

5. For any $w \in \operatorname{int} \mathcal{C}$ we have

$$\begin{aligned} \Phi(w) \;&=\; \tfrac{1}{2}\|w\|_Q^2 + \Psi(z) - \ln \tau \\ &\stackrel{(2.9)}{\geq} \tfrac{1}{2}\|w\|_Q^2 + \Psi(z_0) - \kappa(z_0) - \nu_\Psi \ln \tfrac{\langle s_0, x\rangle + \langle s, x_0\rangle}{2\delta_0} - \ln \tau \end{aligned}$$

Hence, varying w along the open ray $\{\lambda w, \lambda > 0\}$, we obtain

$$\widetilde{\Phi}(w) \;\geq\; \Psi(z_0) - \kappa(z_0) - \nu_\Psi \ln \tfrac{\langle s_0, x\rangle + \langle s, x_0\rangle}{2\delta_0} - \ln \tau + \min_{\lambda > 0} \left[\tfrac{\lambda^2}{2}\|w\|_Q^2 - \nu_F \ln \lambda \right]$$

Thus, the optimal choice is $\lambda = \lambda(w)$, and we get

$$\widetilde{\Phi}(w) \;\geq\; \Psi(z_0) - \kappa(z_0) - \nu_\Psi \ln \tfrac{\langle s_0, x\rangle + \langle s, x_0\rangle}{2\delta_0} - \ln \tau + \nu_F \ln \|w\|_Q + \tfrac{\nu_F}{2}[1 - \ln \nu_F].$$

In view of (4.2) we have

$$\begin{aligned} \langle s_0, x\rangle + \langle s, x_0\rangle \;&\leq\; \Omega \cdot \|w\|_Q + \tau \cdot \delta_0. \\ \tau \;&\leq\; \tfrac{1}{\delta_0}\left[\Omega \cdot \|w\|_Q + \tau \cdot \delta_0\right]. \end{aligned}$$

Hence,

$$\begin{aligned} \widetilde{\Phi}(w) \;&\geq\; \Psi(z_0) - \kappa(z_0) + \nu_\Psi \ln 2 + \nu_F \ln \delta_0 + \tfrac{\nu_F}{2}[1 - \ln \nu_F] - \nu_F \ln \left(\Omega + \tfrac{\tau \cdot \delta_0}{\|w\|_Q} \right) \\ &=\; \Psi(z_0) - \kappa(z_0) + \nu_\Psi \ln 2 + \tfrac{\nu_F}{2}[1 - \ln \nu_F] - \nu_F \ln \left(\tfrac{\Omega}{\delta_0} + \tfrac{\tau}{\|w\|_Q} \right). \end{aligned} \tag{4.8}$$

Thus, we can justify a rate of convergence of a simple potential-reduction scheme.

Theorem 2 *Let us choose an arbitrary $w_1 \in \operatorname{int} \mathcal{C}$ with $\lambda(w_1) = 1$. Let the sequence $\{w_k = (x_k, s_k, y_k, \tau_k)\}_{k=1}^{\infty}$ be generated by the rule*

$$\left.\begin{aligned} \bar{w}_k \;&=\; w_k - \tfrac{[\nabla^2 \Phi(w_k)]^{-1} \nabla \Phi(w_k)}{1 + \|\nabla \Phi(w_k)\|^*_{\nabla^2 \Phi(w_k)}} \\ w_{k+1} \;&=\; T(\bar{w}_k) \end{aligned}\right\}, \quad k \geq 1. \tag{4.9}$$

Then, for any $k \geq 1$ we have

$$\frac{\tau_k}{\|w_k\|_Q} \;\geq\; \exp\left\{\frac{(k-1)\omega}{\nu_F} + C_0\right\} - \frac{\Omega}{\langle s_0, x_0\rangle}, \tag{4.10}$$

where $C_0 = \frac{1}{\nu_F}\left[\Psi(z_0) - \kappa(z_0) + \nu_\Psi \ln 2 - \frac{\nu_F}{2}\ln \nu_F - F(w_1)\right]$.

Proof:
Indeed, in view of conditions of the theorem, we have

$$\Phi(w_1) \;=\; \tfrac{1}{2}\nu_F + F(w_1).$$

In view of the rules (4.9), we have:

$$\Phi(w_{k+1}) \;\leq\; \Phi(\bar{w}_k) \overset{(4.4)}{\leq} \Phi(w_k) - \omega.$$

Hence, for any $k \geq 1$ we have

$$\tfrac{1}{2}\nu_F + F(w_1) - (k-1)\cdot\omega \;\geq\; \Phi(w_k) \;\geq\; \widetilde{\Phi}(w_k)$$

$$\overset{(4.8)}{\geq} \Psi(z_0) - \kappa(z_0) + \nu_\Psi \ln 2 + \tfrac{\nu_F}{2}[1 - \ln \nu_F] - \nu_F \ln\left(\tfrac{\Omega}{\delta_0} + \tfrac{\tau_k}{\|w_k\|_Q}\right).$$

□

Corollary 1 *Define the projected point*

$$\tilde{x}_k \;=\; \tfrac{x_k}{\tau_k}, \quad \tilde{s}_k \;=\; \tfrac{s_k}{\tau_k}, \quad \tilde{y}_k \;=\; \tfrac{y_k}{\tau_k}.$$

Then, for k big enough, we have the following estimate:

$$\begin{aligned} &\max\left\{\|A\tilde{x}_k - b\|^*_{B_H},\; \|\tilde{s}_k + A^*\tilde{y}_k - c\|^*_{B_E},\; |\langle c, \tilde{x}_k\rangle - \langle b, \tilde{y}_k\rangle|\right\} \\ &\qquad \leq \left[\exp\left\{\tfrac{(k-1)\omega}{\nu_F} + C_0\right\} - \tfrac{\Omega}{\langle s_0, x_0\rangle}\right]^{-1}. \end{aligned} \tag{4.11}$$

A similar interior-point method can be developed on the basis of the homogeneous quadratic potential $\widetilde{\Phi}(w)$. Note that for any w and w_+ from $\operatorname{int}\mathcal{C}$ we have[2)]

$$\ln \|w_+\|_Q^2 \;=\; \ln\left(1 + \tfrac{\|w_+\|_Q^2 - \|w\|_Q^2}{\|w\|_Q^2}\right) + \ln\|w\|_Q^2 \;\leq\; q_w(w_+),$$

$$q_w(w_+) \;\overset{\text{def}}{=}\; \ln\|w\|_Q^2 + \tfrac{\|w_+\|_Q^2 - \|w\|_Q^2}{\|w\|_Q^2}.$$

[2)] For this function, it is possible to use a better convex upper bound:

$$\ln\langle Q(w+h), w+h\rangle \;\leq\; \ln\langle Qw, w\rangle + 2\tfrac{\langle Qw, h\rangle}{\langle Qw, w\rangle} + \tfrac{1}{\langle Qu, u\rangle}\langle\left[Q - \tfrac{Qww^TQ}{\langle Qw, w\rangle}\right]h, h\rangle.$$

However, we use a trivial bound for the sake of notation.

Thus, at any point $w \in \operatorname{int} \mathcal{C}$, we can form a convex upper approximation of the homogeneous potential. Indeed, define

$$\mathcal{Q}_w(w_+) \;=\; \tfrac{\nu_F}{2} q_u(w_+) + F(w).$$

Clearly, this is a self-concordant function. Note that

$$\begin{aligned} \widetilde{\Phi}(w) \;&=\; \mathcal{Q}_w(w), \quad \nabla \widetilde{\Phi}(w) \;=\; \nabla \mathcal{Q}_w(w), \\ \widetilde{\Phi}(w_+) \;&\leq\; \mathcal{Q}_w(w_+), \quad w_+ \in \operatorname{int} \mathcal{C}. \end{aligned} \tag{4.12}$$

Therefore, the local decrease of the upper approximation $\mathcal{Q}(w)$ forms a lower bound for the local decrease of the homogeneous potential. Let us show that it cannot be too small.

Lemma 3 *For any $w \in \operatorname{int} \mathcal{C}$ we have*

$$\|\nabla \mathcal{Q}_w(w)\|^*_{\nabla^2 \mathcal{Q}_w(w)} \;\geq\; 1. \tag{4.13}$$

Proof:
Assume that $\|\nabla \mathcal{Q}_w(w)\|^*_{\nabla^2 \mathcal{Q}_w(w)} < 1$. Then the self-concordant function $\mathcal{Q}_w(\cdot)$ attains its minimum at some point $\bar{w}$:

$$0 \;=\; \nabla \mathcal{Q}_w(\bar{w}) \;=\; \nu_F \tfrac{Q\bar{w}}{\|w\|_Q^2} + \nabla F(\bar{w}).$$

Note that $Qw^* = 0$ for any optimal solution $w^* \in W^*$. Hence, multiplying this equation by any nonzero $w^* \in W^*$, we get $\langle \nabla F(\bar{w}), w^* \rangle = 0$. This is impossible since the cone W^* is pointed. □

Now we can justify the following method:

1. Choose arbitrary $w_1 \in \operatorname{int} \mathcal{C}$.

2. For $k \geq 1$, iterate:

$$w_{k+1} = w_k - \frac{[\nabla^2 \mathcal{Q}_{w_k}(w_k)]^{-1} \nabla \widetilde{\Phi}(w_k)}{1 + \|\nabla \widetilde{\Phi}(w_k)\|^*_{\nabla^2 \mathcal{Q}_{w_k}(w_k)}}. \tag{4.14}$$

Theorem 3 *Let sequence $\{w_k \equiv (x_k, s_k, y_k, \tau_k)\}_{k=1}^{\infty}$ be generated by method (4.14). Define the projected point*

$$\tilde{x}_k \;=\; \tfrac{x_k}{\tau_k}, \quad \tilde{s}_k \;=\; \tfrac{s_k}{\tau_k}, \quad \tilde{y}_k \;=\; \tfrac{y_k}{\tau_k}.$$

Then, for k big enough, we have the following estimate:

$$\begin{aligned} &\max \left\{ \|A\tilde{x}_k - b\|^*_{B_H},\; \|\tilde{s}_k + A^*\tilde{y}_k - c\|^*_{B_E},\; |\langle c, \tilde{x}_k \rangle - \langle b, \tilde{y}_k \rangle| \right\} \\ &\qquad \leq\; \left[\exp\left\{ \tfrac{(k-1)\omega}{\nu_F} + C_1 \right\} - \tfrac{\Omega}{\langle s_0, x_0 \rangle} \right]^{-1}, \end{aligned} \tag{4.15}$$

where $C_1 = \frac{1}{\nu_F} [\Psi(z_0) - \kappa(z_0) + \nu_\Psi \ln 2 - F(w_1)] + \ln \frac{1}{\|w_1\|_Q}$.

Proof:
Indeed, in view of (4.12), one step of (4.14) coincides with one step of the Damped Newton Method as applies to $\mathcal{Q}_{w_k}(\cdot)$ at point w_k. Therefore,

$$\widetilde{\Phi}(w_{k+1}) \stackrel{(4.12)}{\leq} \mathcal{Q}_{w_k}(w_{k+1}) \stackrel{(4.13)}{\leq} \mathcal{Q}_{w_k}(w_k) - \omega \stackrel{(4.12)}{=} \widetilde{\Phi}(w_k) - \omega.$$

Thus, for any $k \geq 1$ we have

$$\nu_F \ln \|w_1\|_Q + F(w_1) + \tfrac{\nu_F}{2}[1 - \ln \nu_F] - (k-1) \cdot \omega \;=\; \widetilde{\Phi}(w_1) - (k-1) \cdot \omega$$

$$\geq\; \widetilde{\Phi}(w_k) \stackrel{(4.8)}{\geq} \Psi(z_0) - \kappa(z_0) + \nu_\Psi \ln 2 + \tfrac{\nu_F}{2}[1 - \ln \nu_F] - \nu_F \ln \left(\tfrac{\Omega}{\delta_0} + \tfrac{\tau_k}{\|w_k\|_Q} \right).$$

□

One of the advantages of the methods (4.9) and (4.14) consists in the lower requirements to the accuracy of computation of the Newton step. We will discuss this and other implementation issues in Section 6.

5 Path-following infeasible-start IPM

First of all, let us eliminate extra variables in the quadratic potential function. We will use notation $v = (x, s, \tau)$ for a truncated version of the vector w. Denote

$$\begin{aligned}
\langle \hat{Q} v, v \rangle &\stackrel{\text{def}}{=} \min_y \{ \langle Q w, w \rangle : \; w = (x, s, y, \tau) \}, \\
y(v) &\stackrel{\text{def}}{=} \arg\min_y \{ \langle Q w, w \rangle : \; w = (x, s, y, \tau) \}, \\
\widehat{\mathcal{C}} &= K \times K^* \times R_+, \\
\hat{F}(v) &= F_P(x) + F_D(s) - \ln \tau, \quad v \in \widehat{\mathcal{C}}, \quad \nu_{\hat{F}} = \nu_F, \\
\widehat{\Phi}(v) &= \tfrac{1}{2} \|v\|^2_{\hat{Q}} + \hat{F}(v), \quad v \in \operatorname{int} \widehat{\mathcal{C}}.
\end{aligned}$$

Note that now the cone $\widehat{\mathcal{C}}$ is normal. Therefore, the Hessian of the barrier $\hat{F}(\cdot)$ is positive definite at any feasible point.

The quadratic potential $\widehat{\Phi}(v)$ can be minimized also by a path-following scheme. Let us fix an arbitrary point $v_1 \in \operatorname{int} \widehat{\mathcal{C}}$ with

$$\lambda(v_1) \stackrel{\text{def}}{=} \frac{\nu_F^{1/2}}{\|v_1\|_{\hat{Q}}} \;=\; 1. \tag{5.1}$$

Now we can define the *central path* $v(\mu)$ by the following system of equations:

$$\nabla \widehat{\Phi}(v(\mu)) \;\equiv\; \hat{Q} v(\mu) + \nabla \hat{F}(v(\mu)) \;=\; \mu \cdot [\hat{Q} v_1 + \nabla \hat{F}(v_1)] \stackrel{\text{def}}{=} \mu \cdot g_1. \tag{5.2}$$

Clearly, $v(1) = v_1$. Our goal is to trace this trajectory as $\mu \to 0$.

Let us study the properties of the central path. From definition (5.2), we can compute its derivative:

$$v'(\mu) \;=\; [\nabla^2 \widehat{\Phi}(v(\mu))]^{-1} g_1. \tag{5.3}$$

Multiplying (5.2) by $v(\mu)$ we get

$$\|v(\mu)\|^2_{\hat{Q}} \;\overset{(1.6)}{=}\; \nu_F + \mu \langle g_1, v(\mu) \rangle.$$

Note that

$$\langle g_1, v'(\mu) \rangle \;\overset{(5.3)}{=}\; = \langle g_1, [\nabla^2 \widehat{\Phi}(v(\mu))]^{-1} g_1 \rangle > 0.$$

Therefore, for all $\mu \in (0, 1]$ we have

$$\langle g_1, v(\mu) \rangle \;\leq\; \langle g_1, v_1 \rangle \;\overset{(5.1),\,(1.6)}{=}\; 0.$$

Hence,

$$\|v(\mu)\|_{\hat{Q}} \;\leq\; \nu_F^{1/2}, \quad \mu \in (0, 1]. \tag{5.4}$$

On the other hand, let us multiply (5.2) by an optimal vector $v^* = (x^*, s^*, 1) \in V^*$. Then we obtain

$$\mu \cdot \langle -\nabla \hat{F}(v_1), v^* \rangle \;=\; \langle -\nabla \hat{F}(v(\mu)), v^* \rangle \;\overset{(1.8)}{\geq}\; \langle \nabla^2 \hat{F}(v(\mu)) v^*, v^* \rangle^{1/2}. \tag{5.5}$$

Note that

$$v(\mu) - \tfrac{v^*}{\langle \nabla^2 \hat{F}(v(\mu)) v^*, v^* \rangle^{1/2}} \;\overset{(1.7)}{\in}\; \widehat{\mathcal{C}}.$$

Denote $g_0 = (s_0, x_0, 0) \in \widehat{\mathcal{C}}^*$. Then, from the above inclusion we get

$$\langle g_0, v(\mu) \rangle \;\geq\; \tfrac{\langle g_0, v^* \rangle}{\langle \nabla^2 \hat{F}(v(\mu)) v^*, v^* \rangle^{1/2}} \;=\; \tfrac{\langle s_0, x^* \rangle + \langle s^*, x_0 \rangle}{\langle \nabla^2 \hat{F}(v(\mu)) v^*, v^* \rangle^{1/2}} \;\overset{(3.9)}{=}\; \tfrac{\langle s_0, x_0 \rangle}{\langle \nabla^2 \hat{F}(v(\mu)) v^*, v^* \rangle^{1/2}}.$$

Hence,

$$\tfrac{\langle s_0, x_0 \rangle}{\mu \cdot \langle -\nabla \hat{F}(v_1), v^* \rangle} \;\overset{(5.5)}{\leq}\; \tfrac{\langle s_0, x_0 \rangle}{\langle \nabla^2 \hat{F}(v(\mu)) v^*, v^* \rangle^{1/2}} \;\leq\; \langle g_0, v(\mu) \rangle$$

$$= \;\langle s_0, x(\mu) \rangle + \langle s(\mu), x_0 \rangle.$$

Note that the left-hand side of inequality (4.2) does not depend on y. Therefore, it can rewritten as

$$\langle s_0, x \rangle + \langle s, x_0 \rangle \;\leq\; \Omega \cdot \|v\|_{\hat{Q}} + \tau \cdot \langle s_0, x_0 \rangle.$$

Thus, the last two inequalities combined with (5.4) results in the following estimate.

Lemma 4 *For any $\mu \in (0, 1]$ we have*

$$\tfrac{\tau(\mu)}{\|v(\mu)\|_{\hat{Q}}} \;\geq\; \tfrac{1}{\nu_F^{1/2} \langle -\nabla \hat{F}(v_1), v^* \rangle \cdot \mu} - \tfrac{\Omega}{\langle s_0, x_0 \rangle}. \tag{5.6}$$

From this inequality, we can see that for μ small enough the scaled points

$$\tilde{x}(\mu) \;=\; \tfrac{x(\mu)}{\tau(\mu)} \in K, \quad \tilde{s}(\mu) \;=\; \tfrac{s(\mu)}{\tau(\mu)} \in K^*, \quad \tilde{y}(\mu) \;=\; \tfrac{y(v(\mu))}{\tau(\mu)}.$$

ensure a small residual of the linear system in problem (2.6) and a small duality gap. Thus, the only remaining question is how quickly we can decrease μ and keep our iterates in a small neighborhood of the central path. Fortunately, the answer on this question is already known. For the reader convenience, we reproduce here some results from Section 2 in [8] with a small adjustment of the notation.

Denote $\psi_\mu(v) = \widehat{\Phi}(v) - \mu\langle g_1, v\rangle$. Then $v(\mu) = \arg\min_v \psi_\mu(v)$. We will measure the proximity of point v to the central path by ensuring that the gradient $\nabla\psi_\mu(v)$ is small. For $v \in \operatorname{int}\widehat{\mathcal{C}}$, let us introduce two local metrics:

$$\begin{aligned} \sigma_v^*(g) &= \langle \nabla^2\hat{F}(v)[\nabla^2\widehat{\Phi}(v)]^{-1}g, [\nabla^2\widehat{\Phi}(v)]^{-1}g\rangle, \\ \theta_v^*(g) &= \langle g, [\nabla^2\hat{F}(v)]^{-1}g\rangle^{1/2}. \end{aligned}$$

Note that $\sigma_v^*(g) \le \theta_v^*(g)$. By Lemma 1 in [8],

$$\sigma_v^*(\nabla\widehat{\Phi}(v)) \;\le\; \nu_F^{1/2}, \quad v \in \operatorname{int}\widehat{\mathcal{C}}. \tag{5.7}$$

On the other hand, for the Newton iterate

$$v_+ \;=\; v - [\nabla^2\psi_\mu(v)]^{-1}\nabla\psi_\mu(v), \tag{5.8}$$

by Theorem 1 in [8] we have

$$\theta_{v_+}^*(\nabla\psi_\mu(v_+))) \;\le\; \left(\tfrac{\sigma_v^*(\nabla\psi_\mu(v))}{1-\sigma_v^*(\nabla\psi_\mu(v))}\right)^2. \tag{5.9}$$

Inequalities (5.7) and (5.9) form the basis for complexity analysis of the long-step path-following schemes.

Let us define the following system of neighborhoods of the central path:

$$\mathcal{N}_\beta(\mu) \;=\; \{v:\ \theta_v^*(\nabla\psi_\mu(v)) \le \beta\}, \quad \mu \in (0,1],\ \beta \in \left[0, \tfrac{3-\sqrt{5}}{2}\right).$$

We are going to employ two neighborhoods with sizes $\beta_0 \le \beta_1$. The points from the small neighborhoods $\mathcal{N}_{\beta_0}$ are used for making a long predictor step along the central path. Namely, for $v \in \mathcal{N}_{\beta_0}(\mu)$, we define the *predictor step*:

$$T_v(\alpha) \;=\; \underbrace{v - [\nabla^2\widehat{\Phi}(v)]^{-1}\nabla\psi_\mu(v)}_{\approx\, v(\mu)} - \alpha\underbrace{[\nabla^2\widehat{\Phi}(v)]^{-1}g_1}_{\approx\, v'(\mu) \text{ by (5.3)}}, \quad \alpha \ge 0. \tag{5.10}$$

We should try to choose the maximal α which keeps $T_v(\alpha) \in \mathcal{N}_{\beta_1}(\mu - \alpha)$. As far as the boundary of the large neighborhood is reached, we fix $\mu_+ = \mu - \alpha$ and apply the Newton method (5.8) for obtaining a point in $\mathcal{N}_{\beta_0}(\mu_+)$ (these are the *corrector steps*).

The following statement is an adaptation of Theorem 2 in [8]. For the reader's convenience, we present here its proof.

Theorem 4 *Assume that $v \in \mathcal{N}_{\beta_0}(\mu)$ for some $\mu > 0$. Let $\alpha \geq 0$ satisfy inequality*

$$\beta_0 + \tfrac{\alpha}{\mu}(\beta_0 + \sigma_v^*(\nabla\widehat{\Phi}(v))) \;\leq\; \tfrac{\sqrt{\beta_1}}{1+\sqrt{\beta_1}}.$$

Then $T_v(\alpha) \in \mathcal{N}_{\beta_1}(\mu - \alpha)$.

Proof:
Note that

$$\begin{aligned} \mu\sigma_v^*(g_1) &\leq \sigma^*(\nabla\psi_\mu(v)) + \sigma_v^*(\nabla\widehat{\Phi}(v)) \;\leq\; \theta^*(\nabla\psi_\mu(v)) + \sigma_v^*(\nabla\widehat{\Phi}(v)) \\ &\leq \beta_0 + \sigma_v^*(\nabla\widehat{\Phi}(v)). \end{aligned}$$

Therefore,

$$\begin{aligned} \sigma_v^*(\nabla\psi_{\mu-\alpha}(v)) &= \sigma_v^*(\nabla\psi_\mu(v) + \alpha g_1) \;\leq\; \sigma_v^*(\nabla\psi_\mu(v)) + \alpha\sigma_v^*(g_1) \\ &\leq \beta_0 + \tfrac{\alpha}{\mu}\left[\beta_0 + \sigma_v^*(\nabla\widehat{\Phi}(v))\right]. \end{aligned} \tag{5.11}$$

If the right-hand-side of this inequality does not exceed $\frac{\sqrt{\beta_1}}{1+\sqrt{\beta_1}}$, then, in view of inequality (5.9), the point

$$v_+ \;\stackrel{\rm def}{=}\; v - [\nabla^2\psi_\mu(v)]^{-1}\nabla\psi_{\mu-\alpha}(v)$$

belongs to the neighborhood $\mathcal{N}_{\beta_1}(\mu - \alpha)$. It remains to check that $v_+ = T_v(\alpha)$. □

In view of the presence of the corrector term in the definition (5.10), it is possible to have a convergent strategy with $\beta_1 = \beta_0$. In this case, we have no independent corrector steps at all. Let us write this *long-step infeasible-start path-following* IPM explicitly.

1. Choose arbitrary $v_1 \in \operatorname{int}\widehat{\mathcal{C}}$. Set $\mu_k = 1$ and $\beta = \frac{1}{9}$.

2. For $k \geq 1$ iterate:

a) Starting from $\alpha_k = \frac{5\,\mu_k}{4 + 36\sigma_v^*(\nabla\widehat{\Phi}(v))}$, find the maximal α_k satisfying $T_{v_k}(\alpha_k) \in \mathcal{N}_\beta(\mu_k - \alpha_k)$.

b) Set $v_{k+1} = T_{v_k}(\alpha_k)$, and $\mu_{k+1} = \mu_k - \alpha_k$.

(5.12)

In view of Theorem 4, method (5.12) generates a sequence of points in the neighborhood $\mathcal{N}_{\frac{1}{9}}(\cdot)$. Moreover,

$$\alpha_k \;\stackrel{(5.7)}{\geq}\; \tfrac{5\,\mu_k}{4 + 36\nu_F^{1/2}}.$$

Thus, for obtaining μ_k of the order ϵ, we need $O(\nu_F^{1/2}\ln\frac{1}{\epsilon})$ iterations. However, from the practical point of view, this strategy could be not very efficient. Indeed, by the

estimate (5.11) substituted in (5.9), we can see that

$$\theta^*_{T_{v_k}(\alpha_k)}(\nabla\psi_{\mu_k-\alpha_k}(T_{v_k}(\alpha_k))) \;\le\; \left(\frac{\beta_0\mu_k+\alpha_k\left[\beta_0+\sigma_v^*(\nabla\widehat{\Phi}(v))\right]}{(1-\beta_0)\mu_k-\alpha_k\left[\beta_0+\sigma_v^*(\nabla\widehat{\Phi}(v_k))\right]}\right)^2$$

$$= \;\frac{\beta_0^2}{(1-\beta_0)^2}+\frac{2\beta_0}{(1-\beta_0)^3}\cdot\left[\beta_0+\sigma_v^*(\nabla\widehat{\Phi}(v_k))\right]\cdot\frac{\alpha_k}{\mu_k}+O\left(\frac{\alpha_k^2}{\mu_k^2}\right).$$

Thus, for making long steps, we need to keep β_0 small with respect to β_1. In this case, the additional computational cost of the corrector stage can be compensated by the higher rate of convergence.

6 Numerical results

In the first part of this paper we have developed several infeasible-start interior-point methods for conic problems. All methods are justified by the worst-case complexity analysis. However, most of them are able to accelerate in a favorable situation. It is interesting to check how much this flexibility can help in solving the real-world problems. In this section we present and discuss the results of our preliminary testing on randomly generated conic problems.

Let us describe our random generator. The primal cone K is formed as a direct product of L small-dimensional cones K_i:

$$K \;=\; K_1 \times \ldots \times K_L,$$

where all K_i have the same analytic structure. Therefore, the dimension n of the primal problem (2.1), and the parameters of the primal and barriers ν_P and ν_D are defined by L. We consider the following basic cones.

- *Positive ray:* $n = L$, $\nu_P = \nu_D = L$.
- *3D-Lorentz cone:* $n = 3L$, $\nu_P = \nu_D = 2L$.
- *3D-Power cone:* $n = 3L$, $\nu_P = \nu_D = 4L$ (see (1.2) and (1.3)).
- *Conic hull of the epigraph of the exponent*: $n = 3L$, $\nu_P = \nu_D = 3L$ (see (1.4)).

The generation of random problem instances is performed in four steps.

1. Choose L (and therefore n). Choose the number of equality constraints m.
2. Generate randomly matrix $A \in R^{m\times n}$ with entries uniformly distributed in $[-1, 1]$.
3. Generate randomly $x_0 \in \operatorname{int} K$ and $s_0 \in \operatorname{int} K^*$. Define $c = s_0$ and $b = Ax_0$.
4. Choose the starting points $\bar{x}$ and $\bar{s}$ as the natural "unit" points of the cones K and K^* and $\bar{y}$ as the vector of all ones.

In our tests, we always choose $m = n/3$. We asked for accuracy 10^{-6} for the duality gap and for l_∞-residual of the primal-dual linear system.

Let us describe the computational results related to the path-following scheme. We implemented a simplified version of method (5.12), which follows the trajectory $w(\mu)$ defined by the equation

$$\nabla\Phi(w(\mu)) \;=\; \mu\nabla\Phi(\hat{w}), \quad \mu \in (0, 1],$$

where $\hat{w} = \lambda(\bar{w}) \cdot \bar{w}$ and $\bar{w} = (\bar{x}, \bar{s}, \bar{y}, 1)$.

In our experiments, the neighborhood parameters for the predictor-corrector process were chosen as

$$\beta_0 = 0.01, \quad \beta_1 = 0.75.$$

We were searching along the ray $v(\alpha) = T_v(\alpha)$ while

$$\theta^*_{v(\alpha)}(\nabla\psi_\mu(v(\alpha))) \leq \beta_1.$$

Since in our test problems the primal and dual cones are formed as direct products of L small-dimensional cones, the verification of this inequality is cheap. Cholesky decomposition of the Hessian of the augmented barrier is computed only once per iteration, at the point which lies on the boundary of β_1-neighborhood of the central path. This Hessian is used for the corrector process and for computation of the search direction at the predictor step.

The results of our experiments are given in Table 1. The first two columns of this table give the dimension of the problem. The remaining five columns present the number of iteration for different types of cones (positive orthant, Lorents cone, power cone of degree $\frac{1}{2}$ and $\frac{1}{3}$, and the conic hull of the epigraph of the exponent). We can see that for all cones the number of iterations of the long-step path-following scheme is very similar and quite small.

n	m	LP	Lorentz	$K_{\frac{1}{2}}$	$K_{\frac{1}{3}}$	K_{exp}
30	10	53	44	59	62	57
60	20	62	55	74	76	71
90	30	66	50	69	73	60
120	40	70	62	79	79	68
150	50	72	61	80	82	75
180	60	76	65	73	73	66
210	70	82	69	80	79	71
240	80	85	73	84	83	73
270	90	89	77	86	87	75
300	100	86	78	93	92	76
330	110	87	76	91	91	81
360	120	86	74	98	98	80
480	160	84	80	95	96	95
600	200	98	84	101	100	85
720	240	102	89	109	110	93
840	280	99	96	111	109	95
960	320	98	90	107	107	102

Table 1: Performance of long-step path-following method

At Figure 1 we plotted the average decrease of the potential function per each iteration of the path-following scheme. For short-step methods, this decrease cannot be better than

$O(\sqrt{\nu_F})$. However, our results clearly demonstrate its linear dependence on the parameter of the barrier. This confirms that in our experiments the long-step strategy really works.

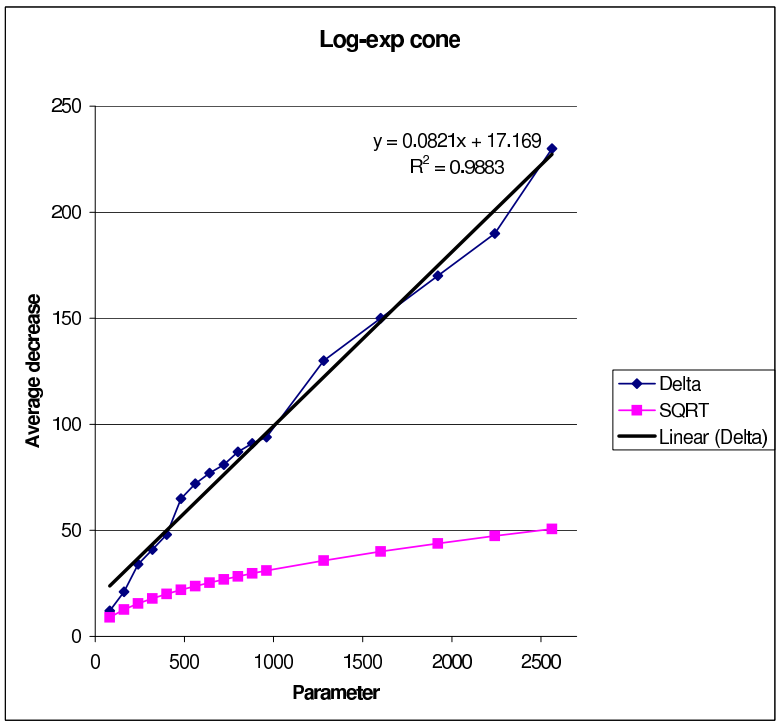

Figure 1: Average decrease of the potential for log-exp cone.

Finally, let us discuss the pure homogeneous potential-reduction strategy (4.14). Unfortunately, for nonlinear cones it was not very efficient. Usually, at the beginning we got a fast decrease of the potential, but later this decrease was becoming close to the theoretical bound. However, for pure Linear Programming problems by unknown reason this method was remarkably good. These results of our experiments are presented in Table 2. In our tests we implement an "exact" line search strategy by one-dimensional Newton method. This operation is cheap since for our test problems every step of this auxiliary scheme needs $O(n)$ operations. As a result, we get a very good average decrease of the potential at the main iterations of method (4.14). From the worst-case complexity analysis, we cannot guarantee more than a decrease by an absolute constant. However,

in our experiments we get it proportional to the dimension of the space of variables.

n	m	Iterations	Average $\Delta_{\tilde{\Phi}}$
30	10	10	$3.9 \cdot 10^1$
60	20	12	$7.2 \cdot 10^1$
90	30	16	$8.3 \cdot 10^1$
120	40	13	$1.4 \cdot 10^2$
150	50	13	$1.8 \cdot 10^2$
180	60	14	$2.1 \cdot 10^2$
210	70	16	$2.2 \cdot 10^2$
240	80	15	$2.7 \cdot 10^2$
270	90	15	$3.0 \cdot 10^2$
300	100	16	$3.1 \cdot 10^2$
330	110	16	$3.4 \cdot 10^2$
360	120	16	$3.8 \cdot 10^2$
480	160	16	$5.0 \cdot 10^2$
600	200	17	$6.2 \cdot 10^2$
720	240	17	$7.6 \cdot 10^2$
840	280	17	$8.8 \cdot 10^2$
960	320	18	$9.6 \cdot 10^2$

Table 2. Results of method (4.14) with line search for LP-problems.

References

[1] F. Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization . *SIAM Journal of Optimization*, **5**, 13 – 51 (1995).

[2] A. Nemirovski, L. Tuncel. Cone-free path-following and potential reduction polynomial time interior-point algorithms. *Mathematical Programming*, **102**(2), 261-294 (2005).

[3] Yu. Nesterov. Long-Step Strategies in Interior-Point Primal-Dual Methods. *Mathematical Programming*, **76**(1), 47 – 94 (1996).

[4] Yu. Nesterov. *Introductory Lectures on Convex Optimization*. Kluwer, Boston, 2004.

[5] Yu. Nesterov. Towards nonsymmetric conic optimization. CORE Discussion Paper 2006/28, Louvain-la-Neuve (2006).

[6] Yu. Nesterov, A. Nemirovskii. *Interior point polynomial methods in convex programming: Theory and Applications*, SIAM, Philadelphia, 1994.

[7] Yu. Nesterov, M. J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM Journal of Optimization*, **8**, 324 – 364 (1998).

[8] Yu.Nesterov, J.-Ph.Vial. Augmented self-concordant barriers and nonlinear optimization problems with finite complexity. *Mathematical Programming*, **99**(1), 149 – 174 (2004).

[9] J. Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization*. MPS/SIAM Series on Optimization 3. SIAM Publications, Philadelphia, 2001.

[10] M. J. Todd, K. C. Toh and R. H. Tutuncu. On the Nesterov-Todd direction in semidefinite programming. *SIAM Journal on Optimization*, **8**, 769 – 796 (1998).

[11] L. Tuncel. Generalization of primal-dual interior-point methods to convex optimization problems in conic form. *Foundations of computational mathematics*, **1**, 229 – 254 (2001).