Dual versus Primal-Dual Interior-Point Methods for Linear and Conic Programming

M. J. Todd*

August 30, 2004

Abstract

We observe a curious property of dual versus primal-dual path-following interior-point methods when applied to unbounded linear or conic programming problems in dual form. While primal-dual methods can be viewed as implicitly following a central path to detect primal infeasibility and dual unboundedness, dual methods are implicitly moving *away* from the analytic center of the set of infeasibility/unboundedness detectors.

Dedicated to Clovis Gonzaga on his 60th birthday.

*School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York 14853, USA (`miketodd@cs.cornell.edu`). This author was supported in part by NSF through grant DMS-0209457 and ONR through grant N00014-02-1-0057.

1 Introduction

Path-following interior-point methods have been shown to be very successful algorithms for linear and conic programming problems: see, for instance, Gonzaga [3], Wright [9], and Ye [10]. These algorithms aim to follow the so-called central path towards optimal solutions. This path only exists when both primal and dual problems have strictly feasible solutions, but the methods are surprisingly successful in detecting infeasibility and unboundedness in case they are present. Recently, the author [8] provided some rationale for this success when using primal-dual algorithms: roughly, the methods could be viewed as implicitly following another well-defined central path towards optimal solutions of auxiliary problems that demonstrated the infeasibility and unboundedness of the original problems. Here we investigate dual path-following algorithms from this viewpoint, and find that their behavior is radically different.

Suppose we wish to solve the linear programming problem in dual form

$$\begin{array}{lll}(D) & \text{maximize} & b^T y, \\ & & A^T y + s = c, \quad s \geq 0.\end{array}$$

Here A, an $m \times n$ matrix, $b \in \mathbf{R}^m$, and $c \in \mathbf{R}^n$ form the data; $(y, s) \in \mathbf{R}^m \times \mathbf{R}^n$ constitutes the variables of the problem. We will assume without real loss of generality that A has full row rank. We call (D) the dual problem although we consider it of primary importance, because it is the dual of the standard form problem (called the primal)

$$\begin{array}{lll}(P) & \text{minimize} & c^T x, \\ & & Ax = b, \quad x \geq 0,\end{array}$$

which has been the focus of most treatments of interior-point methods for linear programming.

We prefer to concentrate on (D) because dual (or less tautologically, non-primal-dual) interior-point methods are typically applied to problems in this form. Examples include the dual path-following algorithm of Renegar [7], the dual affine-scaling method of Adler, Karmarkar, Resende, and Veiga [1], and the dual potential-reduction algorithm of Benson, Ye, and Zhang [2] for semidefinite programming. In addition, there are more theoretical reasons to consider (D): there may be a self-concordant barrier for $\mathcal{F}_Y(D) := \{y \in \mathbf{R}^m : c - A^T y \geq 0\}$ whose complexity value (or parameter) is much less than n, that of the standard logarithmic barrier function for the nonnegative orthant $\mathbf{R}^n_+$. One example is the universal barrier for the L_1-ball in $\mathbf{R}^m$ considered by Güler [5], although it is not effectively computable. However, for most of the paper, we consider the standard barrier $-\ln(s)$ for $\mathbf{R}^n_+$ and the corresponding barrier $-\ln(c - A^T y)$ for $\mathcal{F}_Y(D)$, where $\ln(v)$ for a vector v denotes the sum of the natural logarithms of the components of v. This is for simplicity of development and for comparison to the primal-dual method.

Path-following interior-point methods can be viewed as approximating solutions to the dual barrier problem

$$\begin{array}{llcccc}(BD_\mu) & \text{maximize} & b^T y & + & \mu \ln(s) & \\ & & A^T y & + & s & = & c, \\ & & & & (s > 0), & \end{array}$$

as μ decreases towards zero. Here the final constraint is in parentheses because $\ln(s)$ approaches $-\infty$ if s approaches the boundary of the positive orthant. We can also extend the definition of $\ln(v)$ to make it $-\infty$ if v is not positive. Closely related is the primal barrier problem

$$\begin{array}{lllcl} (BP_\mu) & \text{minimize} & c^T x & - & \mu \ln(x) \\ & & Ax & & = \; b, \\ & & & & (x > 0). \end{array}$$

The optimality conditions for (BD_μ) are the existence of Lagrange multipliers $x \in \mathbf{R}^n$ such that

$$\begin{array}{rcrcll} & A^T y & + \; s & = & c, & s > 0, \\ Ax & & & = & b, & \\ x & & - \; \mu s^{-1} & = & 0 & \end{array} \tag{1.1}$$

(where s^{-1} denotes the vector of reciprocals of the components of s), so that necessarily $x = \mu s^{-1} > 0$, while those for (BP_μ) can be written as the existence of $(y, s) \in \mathbf{R}^m \times \mathbf{R}^n$ so that

$$\begin{array}{rcrcll} & A^T y & + \; s & = & c, & \\ Ax & & & = & b, & x > 0, \\ -\mu x^{-1} & & + \; s & = & 0, & \end{array} \tag{1.2}$$

so that necessarily $s = \mu x^{-1} > 0$. Note that these conditions are equivalent, and can also be written in the more symmetrical form

$$\begin{array}{rcrcll} & A^T y & + \; s & = & c, & s > 0, \\ Ax & & & = & b, & x > 0, \\ & XSe & & = & \mu e, & \end{array} \tag{1.3}$$

where X and S are the diagonal matrices containing the components of x and s respectively, and $e \in \mathbf{R}^n$ denotes a vector of ones.

In order for solutions of (1.1)–(1.3) to exist for a particular positive μ we must have strictly feasible solutions (with $x > 0$ and $s > 0$) to (P) and (D), and these conditions turn out to be also sufficient for unique solutions $(x(\mu), y(\mu), s(\mu))$ to (1.1)–(1.3) to exist for all positive μ: see, e.g., Wright [9]. In this case, the set of such solutions is called the (primal-dual) *central path*.

While the nonlinear systems above are equivalent, the corresponding Newton systems yield different search directions $(\Delta x, \Delta y, \Delta s)$. Primal-dual path-following methods move in the directions that solve the Newton system for (1.3) from a current iterate (x, y, s) with $x > 0$ and $s > 0$, replacing μ with $\sigma s^T x/n$ for some $\sigma \in [0, 1]$. If the initial (and hence all subsequent) iterates x and (y, s) are feasible in (P) and (D) respectively, this is called a feasible-interior-point method, otherwise an infeasible-interior-point method (IIPM). Attractive theoretical convergence results are available for both feasible- and infeasible-interior-point methods when strictly feasible solutions exist for both (P) and (D): see again Wright [9], e.g.

We are concerned with the case where (D) has a strictly feasible solution, but (P) is infeasible, so that (D) is unbounded. Then solutions to (1.1)–(1.3) do not exist, and so following the (nonexistent) central path seems an exercise in futility. Nevertheless, primal-dual IIPMs seem very successful, in that in such a case the iterates (y, s), when scaled by

$b^T y$, provide approximate certificates $(\bar{y}, \bar{s})$ of primal infeasibility and dual unboundedness: $A^T\bar{y} + \bar{s} \approx 0$, $b^T\bar{y} = 1$, and $\bar{s} \geq 0$.

Indeed, the author [8] provided some justification for this success by showing that, under suitable conditions, the primal-dual IIPM for (P) and (D) is implicitly applying a similar method to the pair of dual linear programming problems below, which do have strictly feasible solutions (and hence an associated central path) under the slightly stronger assumption that (P) is strictly infeasible (see below), and whose solutions provide such a primal infeasibility/dual unboundedness certificate.

The dual problem has constraints corresponding to the conditions for primal infeasibility, and an objective function depending on the x-component of the initial iterate:

$$
\begin{array}{llrcrcl}
(\bar{D}) & \max & (Ax_0)^T\bar{y} & & & & \\
 & & A^T\bar{y} & + & \bar{s} & = & 0, \\
 & & b^T\bar{y} & & & = & 1, \\
 & & & & \bar{s} & \geq & 0,
\end{array}
$$

with dual

$$
\begin{array}{llrcrcl}
(\bar{P}) & \min & & & \bar{\zeta} & & \\
 & & A\bar{x} & + & b\bar{\zeta} & = & Ax_0, \\
 & & \bar{x} \geq 0. & & & &
\end{array}
$$

Corresponding to every iterate (x, y, s) for (P) and (D), there is a corresponding "shadow iterate" $(\bar{x}, \bar{\zeta}, \bar{y}, \bar{s})$ for $(\bar{P})$ and $(\bar{D})$: we obtain $(\bar{y}, \bar{s})$ by scaling by $b^T y$ as above, while a different scaling of x produces $\bar{x}$ and $\bar{\zeta}$. Interestingly, while we assume that (y, s) is feasible for (D) while necessarily x is infeasible for (P), the reverse is true for the shadow iterates: $(\bar{y}, \bar{s})$ is only approximately feasible for $(\bar{D})$, while $(\bar{x}, \bar{\zeta})$ is exactly feasible for $(\bar{P})$ (for example, $\bar{x} = x_0$, $\bar{\zeta} = 0$ is the initial shadow iterate). We then compare the sequence of original and shadow iterates. The precise form of the result can be found in [8].

Our aim here is to investigate this situation when a dual path-following method is used. We find that, in strong contrast to the primal-dual case, the associated scaled iterates are moving *away* from a central primal infeasibility/dual unboundedness certificate: more precisely, while they are trying to move towards satisfying the linear equations defining an infeasibility/unboundedness certificate, they are moving away from the certificate that minimizes a natural barrier function. Hence dual methods may be more likely to approach the boundary of the nonnegative orthant more closely and hence run into numerical difficulties.

In Section 2 we define a natural centering problem for finding a primal infeasibility/dual unboundedness certificate, and obtain the necessary and sufficient conditions for this problem to have a solution. The following section is devoted to comparing the dual path-following method for (D) and the corresponding scaled iterates with the iterates of Newton's method for this centering problem. In Section 4, we consider the dual affine-scaling method instead of the dual path-following method and obtain a similar result, and then show that our results also hold for general conic programming problems.

Let us make a small parenthetical remark. In [4], Gonzaga and Todd show another perhaps surprising difference between dual and primal-dual methods: dual potential-reduction algorithms cannot assure an R-linear rate of convergence greater than one, although R-quadratic convergence is possible for primal-dual potential-reduction methods.

2 The Centering Problem

We assume henceforth that (P) is *strictly infeasible*, so that there is some $(\bar{y}, \bar{s})$ with

$$A^T\bar{y} + \bar{s} = 0, \qquad b^T\bar{y} = 1, \qquad \bar{s} > 0.$$

This implies that (D) is strictly feasible, since a sufficiently large multiple of $(\bar{y}, \bar{s})$ can be added to $(0, c)$ to make its s-component positive, and unbounded, since increasing this multiple sends the objective function to infinity. In [8] we noted that (P) is strictly infeasible iff it is infeasible and for every $\tilde{b}$, $\{x : Ax = \tilde{b}, x \geq 0\}$ is either empty or bounded. Here, since our primary interest is in (D), we note that this condition holds iff (D) is unbounded, as is $\max\{s_j : A^Ty + s = c, s \geq 0\}$ for every j.

Consider the *centering* problem

$$\begin{array}{llccccc} (CD) & \text{maximize} & & & \ln(\bar{s}) & & \\ & & A^T\bar{y} & + & \bar{s} & = & 0, \\ & & b^T\bar{y} & & & = & 1 \\ & & & & (\bar{s} & > & 0). \end{array}$$

The optimal solution to this problem (if it exists) is the analytic center of the set of primal infeasibility/dual unboundedness certificates. Necessary and sufficient conditions for $(\bar{y}, \bar{s})$ to solve (CD) are that there exists $(\bar{x}, \bar{\zeta}) \in \mathbf{R}^n \times \mathbf{R}$ such that

$$\begin{array}{ccccccccl} & & & A^T\bar{y} & + & \bar{s} & = & c, & \bar{s} > 0, \\ & & & b^T\bar{y} & & & = & 1, & \\ A\bar{x} & + & b\bar{\zeta} & & & & = & 0, & \\ \bar{x} & & & & - & \bar{s}^{-1} & = & 0. & \end{array} \tag{2.1}$$

For what follows, we do not require that (CD) have an optimal solution, but it helps to interpret our results, so we provide the following characterization result.

Proposition 2.1 *(CD) has an optimal solution, or equivalently (2.1) has a solution, iff, for every $\tilde{b}$ sufficiently close to b, $\min\{\tilde{b}^Ty : A^Ty + s = c, s \geq 0\}$ has an optimal solution.*

Proof: We can view (CD) as a barrier problem (with $\mu = 1$) for a related problem with the same constraints (except $\bar{s} \geq 0$) and zero objective function, so a solution to (CD) or equivalently to (2.1) exists iff there is a strictly feasible solution to

$$A\bar{x} + b\bar{\zeta} = 0, \qquad \bar{x} \geq 0.$$

Such a solution cannot have $\bar{\zeta} \leq 0$, since if it were negative we immediately get a feasible solution to (P), while it if were zero we could find one by adding a sufficiently large multiple of this solution to any x satisfying $Ax = b$. Hence (CD) has an optimal solution iff there is a strictly feasible solution to

$$Ax = -b, \qquad x \geq 0. \tag{2.2}$$

Since A has full row rank, there are solutions to $Ax = \pm e_i$ for each i, where e_i is the ith coordinate vector in $\mathbf{R}^m$, and hence a strictly feasible solution to (2.2) implies that there are feasible solutions to $Ax = -\tilde{b}, x \geq 0$ for all $\tilde{b}$ sufficiently close to b. Conversely, if the latter holds, there is a solution to $Ax = -b - \epsilon Ae, x \geq 0$, for some positive ϵ, and hence a strictly feasible solution to (2.2).

Now (D) has a feasible solution, so we conclude that this condition is equivalent to the existence of optimal solutions to $\min\{\tilde{b}^T y : A^T y + s = c, s \geq 0\}$ for all $\tilde{b}$ sufficiently close to b. □

3 The "Anti-Newton" Step

Let us now compare Newton steps for (BD_μ) and (CD). We suppose we are given a strictly feasible solution (y, s) for (D), and assume that

$$\beta := b^T y > 0, \tag{3.1}$$

so that the "shadow iterate"

$$(\bar{y}, \bar{s}) := \frac{1}{\beta}(y, s) \tag{3.2}$$

satisfies $b^T \bar{y} = 1, \bar{s} > 0$ and, for $\beta >> \|c\|$, is approximately strictly feasible for (CD), since $A^T \bar{y} + \bar{s} = c/\beta \approx 0$.

The dual path-following or dual barrier method takes a damped Newton step for (BD_μ) from (y, s). Thus it moves in the directions $(\Delta y, \Delta s)$ which, together with some $x_+ \in \mathbf{R}^n$, solve the Newton system

$$\begin{array}{rcrcl} & A^T \Delta y & + & \Delta s & = & 0, \\ Ax_+ & & & & = & b, \\ x_+ & & + & \mu S^{-2} \Delta s & = & \mu s^{-1}. \end{array} \tag{3.3}$$

Alternatively, if we write $x + \Delta x$ for x_+, this is the Newton system for the optimality conditions (1.1) for (BD_μ) from (x, y, s), for any $x \in \mathbf{R}^n$. We prefer the system (3.3) written as above to stress that $(\Delta y, \Delta s)$ is independent of the iterate x, since x appears linearly in (1.1).

The result of taking a step of length $\alpha > 0$ in these directions is

$$(y_+, s_+) := (y, s) + \alpha(\Delta y, \Delta s) \tag{3.4}$$

with objective function

$$\beta_+ := b^T y_+ = b^T y + \alpha b^T \Delta y =: \beta + \alpha \Delta \beta,$$

and the associated shadow iterate, if we assume that $\Delta \beta > 0$ (we will discuss this below), is

$$\begin{array}{rcl} (\bar{y}_+, \bar{s}_+) & = & \frac{1}{\beta + \alpha \Delta \beta}(y + \alpha \Delta y, s + \alpha \Delta s) \\ & = & \frac{1}{\beta}(y, s) + \frac{\alpha \Delta \beta}{\beta + \alpha \Delta \beta} \left(\frac{\Delta y}{\Delta \beta} - \frac{y}{\beta}, \frac{\Delta s}{\Delta \beta} - \frac{s}{\beta} \right) \\ & = & (\bar{y}, \bar{s}) + \bar{\alpha}(\Delta \bar{y}, \Delta \bar{s}), \end{array} \tag{3.5}$$

where

$$\bar{\alpha} := \frac{\alpha \Delta \beta}{\beta + \alpha \Delta \beta}, \quad \Delta \bar{y} := \frac{\Delta y}{\Delta \beta} - \bar{y}, \quad \Delta \bar{s} = \frac{\Delta s}{\Delta \beta} - \bar{s}. \tag{3.6}$$

Our aim is to see to what extent the implicit shadow directions $\Delta\bar{y}$ and $\Delta\bar{s}$ defined above satisfy the Newton system for (CD): for some $(\bar{x}_+, \bar{\zeta}_+) \in \mathbf{R}^n \times \mathbf{R}$,

$$\begin{array}{rcrcrcl} & & A^T\Delta\bar{y} & + & \Delta\bar{s} & = & -A^T\bar{y} - \bar{s}, \\ & & b^T\Delta\bar{y} & & & = & 1 - b^T\bar{y} = 0, \\ A\bar{x}_+ & + & b\bar{\zeta}_+ & & & = & 0, \\ \bar{x}_+ & & & + & \bar{S}^{-2}\Delta\bar{s} & = & \bar{s}^{-1}. \end{array} \tag{3.7}$$

Note that again we have used the "+" subscript for $\bar{x}$ and $\bar{\zeta}$, since they appear linearly in the optimality conditions (2.1) for (CD).

Theorem 3.1 *Under the assumptions that β and $\Delta\beta$ are positive, the directions $\Delta\bar{y}$ and $\Delta\bar{s}$ of (3.6), together with $\bar{x}_+ := (\beta^2/\mu\Delta\beta)x_+$ and $\bar{\zeta}_+ := -\beta^2/\mu\Delta\beta$, satisfy the first three equations of (3.7) along with*

$$\bar{x}_+ + \bar{S}^{-2}\Delta\bar{s} = -\left(1 - \frac{\beta}{\Delta\beta}\right)\bar{s}^{-1}. \tag{3.8}$$

Note especially the negative sign on the right-hand side of (3.8): this is the reason we think of these directions as approximations to "anti-Newton" directions, as we discuss below.

Proof: We find

$$A^T\Delta\bar{y} + \Delta\bar{s} = \frac{1}{\Delta\beta}(A^T\Delta y + \Delta s) - \frac{1}{\beta}(A^Ty + s) = 0 - A^T\bar{y} - \bar{s}$$

as desired, using the first equation of (3.3). Similarly, $b^T\Delta\bar{y} = 0$ follows from the definitions of β, $\Delta\beta$, and $\Delta\bar{y}$; and $A^T\bar{x}_+ + b\bar{\zeta}_+ = 0$ from the definitions of $\bar{x}_+$ and $\bar{\zeta}_+$ and the second equation of (3.3). Finally,

$$\begin{aligned} \bar{x}_+ + \bar{S}^{-2}\Delta\bar{s} &= \frac{\beta^2}{\mu\Delta\beta}x_+ + \beta^2 S^{-2}\left(\frac{\Delta s}{\Delta\beta} - \frac{s}{\beta}\right) \\ &= \frac{\beta^2}{\mu\Delta\beta}(x_+ + \mu S^{-2}\Delta s) - \beta s^{-1} \\ &= \frac{\beta^2}{\mu\Delta\beta}\mu s^{-1} - \bar{s}^{-1} \\ &= -\left(1 - \frac{\beta}{\Delta\beta}\right)\bar{s}^{-1}, \end{aligned}$$

where the first equation follows from the definitions and the third from the third equation of (3.3). □

We now discuss the assumption that $\Delta\beta > 0$ and the interpretation of the theorem. We note that the solution to (3.3) yields

$$\Delta y = (AS^{-2}A^T)^{-1}(\mu^{-1}b - As^{-1})$$

so that

$$\Delta\beta = \mu^{-1}b^T(AS^{-2}A^T)^{-1}b - b^T(AS^{-2}A^T)^{-1}As^{-1}.$$

Since $(AS^{-2}A^T)^{-1}$ is positive definite, we see that $\Delta\beta \to +\infty$ as $\mu \downarrow 0$, so that $\Delta\beta > 0$ for sufficiently small positive μ. Further, for such μ, $1 - \beta/\Delta\beta \approx 1$, so that $(\Delta y, \Delta s)$ approximately solves the Newton system (3.7), *except* with the sign of the last right-hand side reversed. Since this equation corresponds to taking a Newton step towards the minimizer (or maximizer!) of the centering objective function in (CD), we can interpret the theorem as stating that the shadow search directions, while moving towards feasibility in (CD), move *directly away from* the minimizer of a quadratic approximation to the barrier function: we therefore think of the solutions to (3.7) with the sign of the last right-hand side reversed as "anti-Newton" directions.

It might seem somewhat surprising that the behavior of the primal-dual IIPM as described in the introduction is somehow "good," while that of the dual path-following method is "bad." After all, if the current primal-dual iterate (x, y, s) satisfies $x = \mu s^{-1}$ (of course, this does not mean we are on the central path, since x is infeasible), then the primal-dual IIPM's (y, s)-directions coincide with those of the dual path-following method. This is not a contradiction. Even with the feasible interior-point method, the search direction when on the central path is the opposite of the centering direction. But the good behavior of path-following methods (when the central path exists) is that they have a tendency from the centering part of the step to approach the path, even if it is leading away from a central point. On the other hand, our analysis above shows that the shadow iterates corresponding to the dual barrier method are moving in a sense radially away from a central point, and hence are not converging to any interior point. Thus in the primal-dual method, the primal iterates exert a stabilizing influence on the corresponding shadow iterates, while, as we have seen, in the dual path-following method the (y, s) iterates are independent of the x iterates.

4 Extensions

In this section, we consider the dual affine-scaling directions for (D) and also extensions of our results to more general conic programming problem, such as second-order cone and semidefinite programming problems.

The dual affine-scaling directions for (D) at the strictly feasible point (y, s) are the solutions, together with $x_+^a \in \mathbf{R}^n$, to

$$\begin{array}{rcrcrcl} & & A^T\Delta y^a & + & \Delta s^a & = & 0, \\ Ax_+^a & & & & & = & b, \\ x_+^a & & & + & S^{-2}\Delta s^a & = & 0. \end{array} \tag{4.1}$$

We can view these directions either as steepest ascent for (D), with the metric for s defined by the Hessian of the barrier function at the current iterate, or as the limits of $(\mu\Delta y, \mu\Delta s)$, with $(\Delta y, \Delta s)$ defined by (3.3), as $\mu \downarrow 0$.

By seeing the effect on the shadow iterate of taking a step in these directions, we can define the dual affine-scaling shadow directions $(\Delta\bar{y}^a, \Delta\bar{s}^a)$ from these exactly as in the previous section, with $(\Delta y^a, \Delta s^a)$ replacing $(\Delta y, \Delta s)$ and $\Delta\beta^a := b^T\Delta y^a = b^T(AS^{-2}A^T)^{-1}b > 0$ replacing $\Delta\beta$ in (3.5) and (3.6).

Following the proof of the previous section, we can easily establish

Theorem 4.1 *Under the assumption that $\beta > 0$, the dual affine-scaling shadow directions $(\Delta\bar{y}^a, \Delta\bar{s}^a)$, together with $\bar{x}^a_+ := (\beta^2/\Delta\beta^a)x^a_+$ and $\bar{\zeta}^a_+ := -\beta^2/\Delta\beta^a$, satisfy the first three equations of (3.7) and*

$$\bar{x}^a_+ + \bar{S}^{-2}\Delta\bar{s} = -\bar{s}^{-1},$$

and hence are exactly the anti-Newton directions for (CD). □

For the primal-dual method, the results in [8] show immediately that the shadow directions corresponding to the primal-dual affine-scaling directions for (D) and (P) are exactly the primal-dual affine-scaling directions for $(\bar{D})$ and $(\bar{P})$; we merely set σ equal to zero.

Finally, we show that these results extend to the dual path-following method or dual affine-scaling method applied to any conic programming problem, as we showed for the primal-dual IIPM in [8]. Suppose our dual problem is replaced by

$$(D) \quad \text{maximize} \quad b^T y, \\ A^* y + s = c, \quad s \in K^*,$$

where K^* is a closed convex solid pointed cone in a Euclidean space E with inner product $\langle \cdot, \cdot \rangle$, and A^* is the adjoint of a linear map A from E to $\mathbf{R}^m$. This is the dual of the primal problem

$$(P) \quad \text{minimize} \quad \langle c, x \rangle, \\ Ax = b, \quad x \in K,$$

where K is the cone dual to K^*, $\{x \in E : \langle s, x \rangle \geq 0 \text{ for all } s \in K^*\}$. Two cases of interest are where K^* is a Cartesian product of second-order cones of the form $\{(\tau, t) \in \mathbf{R} \times \mathbf{R}^p : \tau \geq \|t\|\}$, leading to second-order cone programming, and where K^* is the cone of positive semidefinite matrices of some order q (or possibly a product of such cones), leading to semidefinite programming. In both these cases, the dual cone coincides with the original cone.

We suppose we have a logarithmically homogeneous self-concordant barrier function F_* for K^*; this is a strictly convex function, finite on $\operatorname{int} K^*$ and converging to $+\infty$ as its argument approaches a point on the boundary, that satisfies certain bounds on its derivatives introduced by Nesterov and Nemirovski [6]. For our purposes, all that is important is that F_* satisfies

$$F'_*(\tau s) = \tau^{-1} F'_*(s), \quad F''_*(\tau s) = \tau^{-2} F''_*(s), \quad F''_*(s)s = -F'_*(s), \tag{4.2}$$

for any $s \in \operatorname{int} K^*$ and any positive τ.

We note the corresponding changes in our problems and equation systems above. First, in both (BD_μ) and (CD), the implicit constraint $s > 0$ becomes the implicit constraint $s \in \operatorname{int} K^*$, and the objective functions become $b^T y - \mu F_*(s)$ and $-F_*(s)$ respectively; also, A^T is replaced by A^*. The optimality conditions for these problems are similarly slightly modified: again $s > 0$ becomes $s \in \operatorname{int} K^*$, and A^T is replaced by A^*. Also, the last equations are replaced by

$$x + \mu F'_*(s) = 0$$

for (1.1) and

$$\bar{x} + F'_*(\bar{s}) = 0$$

for (2.1). Finally, the direction-defining systems change as follows: A^T is replaced by A^*, and the last equation of (3.3) becomes

$$x_+ + \mu F_*''(s)\Delta s = -\mu F_*'(s), \tag{4.3}$$

while the last equation of (3.7) becomes

$$\bar{x}_+ + F_*''(\bar{s})\Delta\bar{s} = -F_*'(\bar{s}), \tag{4.4}$$

and the last equation of (4.1) becomes

$$x_+^a + F_*''(s)\Delta s = 0. \tag{4.5}$$

We again assume that (P) is strictly infeasible, so that there exists $(\bar{y}, \bar{s})$ with $A^*\bar{y} + \bar{s} = 0$, $b^T\bar{y} = 1$, and $s \in \operatorname{int} K^*$. Then the analog of Proposition 2.1 remains true, with essentially the same proof. More importantly, the analogs of Theorems 3.1 and 4.1 remain true. For the first, establishing the first three equations is straightforward. Also,

$$\begin{aligned}
\bar{x}_+ + F_*''(\bar{s})\Delta\bar{s} &= \frac{\beta^2}{\mu\Delta\beta}x_+ + \beta^2 F_*''(s)\left(\frac{\Delta s}{\Delta\beta} - \frac{s}{\beta}\right) \\
&= \frac{\beta^2}{\mu\Delta\beta}(x_+ + \mu F_*''(s)\Delta s) - \beta F_*''(s)s \\
&= \frac{\beta^2}{\mu\Delta\beta}(-\mu F_*'(s)) + \beta F_*'(s) \\
&= -\left(1 - \frac{\beta}{\Delta\beta}\right)(-F_*'(\bar{s})),
\end{aligned}$$

where we have repeatedly used (4.2) as well as (4.3). Similarly, for the second, the only complication is the last equation, and we find

$$\begin{aligned}
\bar{x}_+^a + F_*''(\bar{s})\Delta\bar{s}^a &= \frac{\beta^2}{\Delta\beta^a}x_+^a + \beta^2 F_*''(s)\left(\frac{\Delta s^a}{\Delta\beta^a} - \frac{s}{\beta}\right) \\
&= \frac{\beta^2}{\Delta\beta^a}(x_+^a + F_*''(s)\Delta s^a) - \beta F_*''(s)s \\
&= \beta F_*'(s) \\
&= -(-F_*'(\bar{s})),
\end{aligned}$$

using (4.2) and (4.5).

Hence once again, comparing with (4.4), we find that the corresponding shadow iterates are moving in (approximate) anti-Newton directions.

References

[1] I. Adler, N. K. Karmarkar, M. G. C. Resende, and G. Veiga. An implementation of Karmarkar's algorithm for linear programming. *Mathematical Programming*, 44:297–335, 1989.

[2] S. J. Benson, Y. Ye, and X. Zhang. Solving large-scale sparse semidefinite programs for combinatorial optimization. *SIAM Journal on Optimization*, 10:443-461, 2000.

[3] C. C. Gonzaga. Path following methods for linear programming. *SIAM Review*, 34(2):167–227, 1992.

[4] C. C. Gonzaga and M. J. Todd. An $O(\sqrt{n}L)$–iteration large–step primal–dual affine algorithm for linear programming. *SIAM Journal on Optimization*, 2:349–359, 1992.

[5] O. Güler. Barrier functions in interior-point methods. *Mathematics of Operations Research*, 21:860–885, 1996.

[6] Yu. E. Nesterov and A. S. Nemirovski. *Interior Point Polynomial Methods in Convex Programming : Theory and Algorithms*. SIAM Publications. SIAM, Philadelphia, USA, 1994.

[7] J. Renegar. A polynomial–time algorithm, based on Newton's method, for linear programming. *Mathematical Programming*, 40:59–93, 1988.

[8] M. J. Todd. Detecting infeasibility in infeasible-interior-point methods for optimization. In Felipe Cucker, Ron DeVore, Peter Olver, and Endre Suli, editors, *Foundations of Computational Mathematics, Minneapolis 2002*, pages 157–192. Cambridge University Press, 2004.

[9] S. J. Wright. *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, 1997.

[10] Y. Ye. *Interior Point Algorithms: Theory and Analysis*. John Wiley, New York, 1997.