

A POLYNOMIAL-TIME INTERIOR-POINT METHOD FOR CONIC OPTIMIZATION, WITH INEXACT BARRIER EVALUATIONS*

SIMON P. SCHURR[†], DIANNE P. O’LEARY[‡], AND ANDRÉ L. TITS[§]

Abstract. We consider a primal-dual short-step interior-point method for conic convex optimization problems for which exact evaluation of the gradient and Hessian of the primal and dual barrier functions is either impossible or prohibitively expensive. As our main contribution, we show that if approximate gradients and Hessians of the primal barrier function can be computed, and the relative errors in such quantities are not too large, then the method has polynomial worst-case iteration complexity. (In particular, polynomial iteration complexity ensues when the gradient and Hessian are evaluated exactly.) In addition, the algorithm requires no evaluation—or even approximate evaluation—of quantities related to the barrier function for the dual cone, even for problems in which the underlying cone is not self-dual.

Key words. convex cone, conic optimization, self-concordant barrier function, interior-point method, inexact algorithm, polynomial complexity

AMS subject classifications. 90C51, 90C25

DOI. 10.1137/080722825

1. Introduction. Interior-point methods (IPMs) are the algorithms of choice for solving many convex optimization problems, including semidefinite programs (SDPs) and second-order cone programs (SOCPs) (e.g., [20, 1, 3]). The bulk of the work at each iteration of an IPM consists of the evaluation of the first and second derivatives of a suitably selected barrier function and the solution of the linear system of equations formed by Newton’s method. The most efficient algorithms require barrier functions for both the primal and dual problems.

The success of IPMs on SDPs and SOCPs has not been matched for more general classes of conic convex optimization problems. The primary hindrance is the difficulty of explicitly evaluating the derivatives of the associated barrier functions, particularly those for the dual problem. The construction of the universal barrier function [20] provides a definition for these derivatives but does not necessarily prescribe a polynomial-time evaluation procedure.

Even for well-studied problem classes, evaluation of the derivatives can be quite expensive. For example, in SDP, where the variables are symmetric positive semidefinite matrices, derivatives of the barrier function are usually evaluated by computing inverses or Cholesky decompositions of matrices. Since these linear algebra tasks can be expensive when the matrices are large or dense, it might be preferable to instead compute approximate inverses or approximate Cholesky decompositions, and hence

*Received by the editors April 29, 2008; accepted for publication (in revised form) January 7, 2009; published electronically May 14, 2009. This work was supported by NSF grant DMI0422931 and DoE grant DEFG0204ER25655. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation or those of the US Department of Energy.

<http://www.siam.org/journals/siopt/20-1/72282.html>

[†]Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, N2L 3G1 Canada (spschurr@uwaterloo.ca).

[‡]Computer Science Department and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742 (oleary@cs.umd.edu).

[§]Department of Electrical and Computer Engineering and Institute for Systems Research, University of Maryland, College Park, MD 20742 (andre@umd.edu).

an approximate gradient and Hessian of the barrier function, if doing so does not greatly increase the number of IPM iterations.

Therefore, in this work, we consider applying a simple primal-dual short-step IPM to conic convex optimization problems for which exact evaluation of the gradient and Hessian is either impossible or prohibitively expensive. Note that this algorithm does not require evaluation of the gradient and Hessian of the dual barrier function. (This property is shared by an algorithm recently proposed by Nesterov [19].) As our main contribution, we show that if estimates of the gradient and Hessian of the primal barrier function are known, and their relative errors are not too large, then the algorithm has the desirable property of polynomial worst-case iteration complexity.

There is an extensive body of literature on inexact IPMs for linear optimization problems, and a smaller body of work on inexact methods for nonlinear convex optimization problems. We note the paper [29], in which it was shown that a simple primal-dual potential reduction method for linear optimization problems can in principle be applied to conic optimization problems. It is shown how the assumptions made on the availability of a self-concordant barrier function and its derivatives are crucial to the convergence and complexity properties of the method. The inverse Hessian of an appropriate self-concordant barrier function is approximated by a matrix that may be updated according to a quasi-Newton scheme. We take a different approach to such approximations.

In the IPM literature, the term “inexact algorithm” typically refers to algorithms in which approximate right-hand sides of the linear system are computed. In some schemes approximate coefficient matrices are also used. The inexact algorithm we consider here can be interpreted in this way, although our primary concern is to estimate the gradient and Hessian of a suitable self-concordant barrier function, rather than the coefficient matrix and right-hand side of the linear system per se. In, e.g., [15, 8, 12] infeasible-point methods for linear optimization are presented in which the linear system is solved approximately. In [13] an inexact primal-dual path-following method is proposed for convex quadratic optimization problems. In [2] an inexact algorithm is presented for monotone horizontal linear complementarity problems. (This class of complementarity problems includes as special cases linear optimization problems and convex quadratic optimization problems.) Inexact methods for other classes of nonlinear convex problems have also been studied. In [32] an inexact primal-dual method is presented for SDP. We also mention the paper [28] in which an inexact primal-dual path-following algorithm is proposed for a class of convex quadratic SDPs. Only the equations resulting from the linearization of the complementarity condition are perturbed, and a polynomial worst-case complexity result is proven. Another source of inexactness is the use of an iterative method such as the preconditioned conjugate gradient method to solve the linear system, and this has been proposed (e.g., in [16, 11, 14, 6, 9, 30]) and analyzed (e.g., in [4, 5, 10]) for certain classes of convex optimization problems.

This paper is organized as follows. We first present some relevant properties of barrier functions in section 2. In section 3 we consider a primal-dual short-step algorithm that uses inexact values of the primal barrier gradient and Hessian. We show in section 4 that if the gradient and Hessian estimates are accurate enough, then the algorithm converges in a polynomial number of iterations. In section 5 we present some applications of our results. Finally, section 6 is devoted to a few concluding remarks.

Much of the work reported here is derived from the dissertation [27]. Additional details, including results obtained by assuming the errors in the gradient and Hessian approximations are dependent, which can occur in practice, may be found there.

2. Notation and preliminaries. A *cone* K is a nonempty set such that $\alpha x \in K$ for all scalars $\alpha \geq 0$ and $x \in K$. Hence K always includes the origin. A cone that is closed, convex, solid (i.e., has nonempty interior), and pointed (i.e., contains no lines) is said to be *proper*.¹ Given an inner product $\langle \cdot, \cdot \rangle$ on \mathcal{R}^n , the *dual* of a set S is given by

$$S^* = \{y : \langle x, y \rangle \geq 0 \ \forall x \in S\}.$$

It can be easily verified that the dual of any set is a closed convex cone, so we may refer to S^* as the dual cone of S . Denote by $\|\cdot\|$ the vector norm induced by $\langle \cdot, \cdot \rangle$, and given a linear mapping A , let $\|A\|$ be the operator norm induced by this vector norm: $\|A\| = \sup_{x \neq 0} \|Ax\|/\|x\|$. Note that for any inner product, $\|A\|$ is the largest singular value of A ; i.e., $\|A\|$ is the spectral norm. The adjoint of A is denoted by A^* , and its inverse by A^{-*} . It will be clear from the context whether $*$ denotes a dual cone or an adjoint map. If A is self-adjoint and positive definite, $A^{1/2}$ denotes its positive definite square root.

Given a set $S \subset \mathcal{R}^n$ and a smooth function $F : S \rightarrow \mathcal{R}$, let $F^{(k)}(x)$ denote the k th derivative of F at $x \in \text{int}(S)$. Following standard practice in the optimization literature, given directions h_1, \dots, h_j , let $F^{(k)}(x)[h_1, \dots, h_j]$ denote the result of applying $F^{(k)}(x)$ successively to h_1, \dots, h_j . For example,

$$(2.1) \quad F''(x)[h_1, h_2] = (F''(x)h_1)h_2 = \langle F''(x)h_1, h_2 \rangle \in \mathcal{R},$$

and

$$(2.2) \quad F'''(x)[h_1, h_2] = (F'''(x)h_1)h_2 \in \mathcal{L}(\mathcal{R}^n),$$

where $\mathcal{L}(\mathcal{R}^n)$ is the space of linear functionals over \mathcal{R}^n , which is isomorphic to \mathcal{R}^n itself.

In the remainder of this section, we establish properties of barrier functions needed in our analysis.

2.1. Self-concordant barrier functions. Let $S \subset \mathcal{R}^n$ be a closed convex set with nonempty interior. The function $F : \text{int}(S) \rightarrow \mathcal{R}$ is said to be a *ψ -self-concordant barrier function* (or simply a *self-concordant barrier*) for S if it is closed (i.e., its epigraph is a closed set), is three times continuously differentiable, and satisfies

$$(2.3) \quad |F'''(x)[h, h, h]| \leq 2(F''(x)[h, h])^{3/2} \quad \forall x \in \text{int}(S) \text{ and } h \in \mathcal{R}^n,$$

and

$$\psi = \sup_{x \in \text{int}(S)} \omega(F, x)^2 < \infty,$$

where

$$\omega(F, x) = \inf_t \{t : |F'(x)[h]| \leq t(F''(x)[h, h])^{1/2} \ \forall h \in \mathcal{R}^n\}.$$
²

If furthermore $F''(x)$ is positive definite for all $x \in \text{int}(S)$, then F is said to be *non-degenerate*. Note that property (2.3) implies the convexity of F , and that convexity

¹Some authors call such a cone *full* or *regular*.

²The quantity $\omega(F, x)$ is often referred to as the *Newton decrement*.

and closedness of F imply that F satisfies the *barrier property*, i.e., for every sequence $\{x^i\} \subset \text{int}(S)$ converging to a boundary point of S , $F(x^i) \rightarrow \infty$. (See [18, proof of Theorem 4.1.4] for details.)

It is known that a self-concordant barrier for a proper cone is necessarily nondegenerate [18, Theorem 4.1.3], but for the sake of clarity we will sometimes refer to such functions as nondegenerate self-concordant barriers.

The quantity ψ is known as the *complexity parameter* of F . If $S \subset \mathcal{R}^n$ is a closed convex set with nonempty interior, and F is a ψ -self-concordant barrier for S , then $\psi \geq 1$ [20, Corollary 2.3.3]. Let F be a nondegenerate self-concordant barrier for $S \subset \mathcal{R}^n$, and let $x \in \text{int}(S)$. Then $F''(x)$ and its inverse induce the following “local” norms on \mathcal{R}^n :

$$(2.4) \quad \|h\|_{x,F} = \langle h, F''(x)h \rangle^{1/2}, \quad \|h\|_{x,F}^* = \langle h, F''(x)^{-1}h \rangle^{1/2}, \quad h \in \mathcal{R}^n.$$

These norms are dual to each other. The following result—which uses and builds upon [20, Theorem 2.1.1]—characterizes the local behavior of the Hessian of a self-concordant barrier.

LEMMA 2.1. *Let F be a nondegenerate self-concordant barrier for the closed convex set S having nonempty interior.*

(a) *If $x, y \in \text{int}(S)$ are such that $r := \|x - y\|_{x,F} < 1$, then for all h ,*

$$(2.5) \quad (1 - r)\|h\|_{x,F} \leq \|h\|_{y,F} \leq \frac{1}{1 - r}\|h\|_{x,F},$$

$$(2.6) \quad (1 - r)\|h\|_{x,F}^* \leq \|h\|_{y,F}^* \leq \frac{1}{1 - r}\|h\|_{x,F}^*.$$

(b) *For every $x \in \text{int}(S)$, $\|x - y\|_{x,F} < 1$ implies that $y \in \text{int}(S)$.*

Proof. The inequalities in (2.5) and the result in (b) are from [20, Theorem 2.1.1]. The inequalities in (2.6) follow from those in (2.5); see, e.g., [27, Lemma 3.2.9]. \square

Although Lemma 2.1 holds for self-concordant barrier functions on general convex sets, we will be most interested in barriers on proper cones, so subsequent results will be phrased with this in mind. Given a proper cone K and a function $F : \text{int}(K) \rightarrow \mathcal{R}$, the *conjugate function*, $F_* : \text{int}(K^*) \rightarrow \mathcal{R}$, of F is given by³

$$(2.7) \quad F_*(s) := \sup_{x \in \text{int}(K)} \{-\langle x, s \rangle - F(x)\}, \quad s \in \text{int}(K^*).$$

Since $F_*(s)$ is the pointwise supremum of linear functions of s , F_* is convex on $\text{int}(K^*)$. It was shown in, e.g., [25, Theorem 3.3.1], that if F is a nondegenerate self-concordant barrier for K , then F_* is a nondegenerate self-concordant barrier for K^* . In light of this, the next result follows directly from Lemma 2.1.

LEMMA 2.2. *Let F be a nondegenerate self-concordant barrier for the proper cone K .*

(a) *If $t, s \in \text{int}(K^*)$ are such that $r := \|t - s\|_{t,F_*} < 1$, then for all h ,*

$$(1 - r)\|h\|_{t,F_*} \leq \|h\|_{s,F_*} \leq \frac{1}{1 - r}\|h\|_{t,F_*}.$$

(b) *For every $s \in \text{int}(K^*)$, $\|s - t\|_{s,F_*} < 1$ implies that $t \in \text{int}(K^*)$.*

³Strictly speaking, in the classical definition of a conjugate function, the domain of F_* is not restricted to $\text{int}(K^*)$. We include such a restriction here so that F_* is finite, which is the only case of interest to us. The definition of a conjugate function used here is found in, e.g., [22, 23, 25]. It is slightly different from that found elsewhere in the literature, including [20]. The difference is the minus sign in front of the $\langle x, s \rangle$ term, which turns the domain of F_* from $\text{int}(-K^*)$ into $\text{int}(K^*)$.

The following technical results will be used in the analysis of the interior-point method to be presented later.

LEMMA 2.3. *Let F be a nondegenerate self-concordant barrier for the proper cone K , and let $x \in \text{int}(K)$. Then for all h_1 and h_2 ,*

$$\|F'''(x)[h_1, h_2]\|_{x,F}^* \leq 2\|h_1\|_{x,F}\|h_2\|_{x,F}.$$

Proof. Since F is a nondegenerate self-concordant barrier function, it follows from [20, Proposition 9.1.1] that $|F'''(x)[h_1, h_2, h_3]| \leq 2\|h_1\|_{x,F}\|h_2\|_{x,F}\|h_3\|_{x,F}$, for all $x \in \text{int}(K)$ and all h_1, h_2, h_3 . (In fact, only property (2.3) in the definition of a self-concordant barrier function is used to prove this result.) Now by definition of a dual norm,

$$\begin{aligned} \|F'''(x)[h_1, h_2]\|_{x,F}^* &= \max_{y:\|y\|_{x,F}=1} \langle y, F'''(x)[h_1, h_2] \rangle \\ &= \max_{y:\|y\|_{x,F}=1} F'''(x)[h_1, h_2, y] \\ &\leq \max_{y:\|y\|_{x,F}=1} 2\|h_1\|_{x,F}\|h_2\|_{x,F}\|y\|_{x,F} \\ &= 2\|h_1\|_{x,F}\|h_2\|_{x,F}. \quad \square \end{aligned}$$

COROLLARY 2.4. *Let F be a nondegenerate self-concordant barrier for the proper cone K , let $x \in \text{int}(K)$ and h be such that $\beta := \|h\|_{x,F} < 1$, and let $\alpha \in [0, 1]$. Then*

$$\|F'''(x + \alpha h)[h, h]\|_{x+\alpha h, F}^* \leq \frac{2\beta^2}{(1-\beta)(1-\alpha\beta)}.$$

Proof. Since $\alpha\beta < 1$, it follows from Lemma 2.1(b) that $x + \alpha h \in \text{int}(K)$, and then from (2.5) that

$$(2.8) \quad \|h\|_{x+\alpha h, F} \leq \frac{\beta}{1-\alpha\beta}.$$

Moreover, $\|(1-\alpha)h\|_{x+\alpha h, F} \leq (1-\alpha)\beta/(1-\alpha\beta) < 1$, so (2.6) followed by Lemma 2.3 and then (2.8) yields

$$\begin{aligned} \|F'''(x + \alpha h)[h, h]\|_{x+\alpha h, F}^* &\leq \frac{1}{1 - \|(1-\alpha)h\|_{x+\alpha h, F}} \|F'''(x + \alpha h)[h, h]\|_{x+\alpha h, F}^* \\ &\leq \frac{1}{1 - (1-\alpha)\|h\|_{x+\alpha h, F}} 2\|h\|_{x+\alpha h, F}^2 \\ &\leq \frac{1}{1 - (1-\alpha)\frac{\beta}{1-\alpha\beta}} 2\left(\frac{\beta}{1-\alpha\beta}\right)^2 \\ (2.9) \quad &= \frac{2\beta^2}{(1-\beta)(1-\alpha\beta)}. \quad \square \end{aligned}$$

2.2. Logarithmically homogeneous barrier functions. An important class of barrier functions for convex cones is that of logarithmically homogeneous barriers. (The terminology reflects the fact that the exponential of such a function is homogeneous.) These barrier functions were first defined in [20, Definition 2.3.2] and occur widely in practice. Given a proper cone K , a convex twice continuously differentiable function $F : \text{int}(K) \rightarrow \mathcal{R}$ is said to be a ν -logarithmically homogeneous barrier for

K if $\nu \geq 1$, F satisfies the barrier property for K , and the *logarithmic-homogeneity relation*

$$(2.10) \quad F(tx) = F(x) - \nu \log(t) \quad \forall x \in \text{int}(K), t > 0$$

holds. The following well-known properties of logarithmically homogeneous barriers are the same or similar to those in [20, Proposition 2.3.4], so we omit proofs.

LEMMA 2.5. *Let K be a proper cone, and $F : \text{int}(K) \rightarrow \mathcal{R}$ be a ν -logarithmically homogeneous barrier for K . For all $x \in \text{int}(K)$ and $t > 0$:*

- (a) $F'(tx) = \frac{1}{t}F'(x)$ and $F''(tx) = \frac{1}{t^2}F''(x)$;
- (b) $\langle F'(x), x \rangle = -\nu$;
- (c) $F''(x)x = -F'(x)$;
- (d) $\langle F''(x)x, x \rangle = \nu$;
- (e) *If in addition $F''(x)$ is positive definite for all $x \in \text{int}(K)$, then $\|F'(x)\|_{x,F}^* = \nu^{1/2}$.*

It was noted in [20, Corollary 2.3.2] that if F is a ν -logarithmically homogeneous barrier for the proper cone K , and F is three times differentiable and satisfies (2.3), then F is a ν -self-concordant barrier for K . Such an F is called a ν -logarithmically homogeneous (nondegenerate) self-concordant barrier for K .

The following result relates the norm induced by the Hessian of F to that induced by the Hessian of F_* . Here and throughout, we write $F_*''(\cdot)$ as shorthand for $(F_*(\cdot))''$.

LEMMA 2.6. *For some ν , let F be a ν -logarithmically homogeneous self-concordant barrier for the proper cone K , and let $x \in \text{int}(K)$. Then for all $\mu > 0$ and all h ,*

$$\|h\|_{-\mu F'(x), F_*} = \frac{1}{\mu} \|h\|_{x,F}^*$$

Proof. It follows from, e.g., [25, Theorem 3.3.1] that F_* is a ν -logarithmically homogeneous self-concordant barrier for K^* . Using Lemma 2.5(a), we obtain that for all $x \in \text{int}(K)$ and $\mu > 0$,

$$F_*''(-\mu F'(x)) = \frac{1}{\mu^2} F_*''(-F'(x)) = \frac{1}{\mu^2} F''(x)^{-1},$$

where the last equality follows from the relation $F_*''(-F'(x)) = F''(x)^{-1}$; see, e.g., [25, Theorem 3.3.4]. (This relation holds for all $x \in \text{int}(K)$, and uses the fact that $F'(x) \in -\text{int}(K^*)$.) So for all h , $\|h\|_{-\mu F'(x), F_*} = \|h\|_{x, \mu^2 F}^* = \frac{1}{\mu} \|h\|_{x,F}^*$. \square

To our knowledge, the following key relation between the norms induced by the Hessians of F and F_* is new. It shows that in a certain region (that we will later identify as a neighborhood of the central path), the dual norm induced by the Hessian of F is comparable to the norm induced by the Hessian of F_* .

LEMMA 2.7. *For some ν , let F be a ν -logarithmically-homogeneous self-concordant barrier for the proper cone K . Let $\theta \in (0, 1)$, and suppose that $x \in \text{int}(K)$ and $s \in \text{int}(K^*)$ satisfy*

$$(2.11) \quad \|s + \mu F'(x)\|_{x,F}^* \leq \theta \mu,$$

where $\mu = \langle x, s \rangle / \nu$. Then for any h ,

$$\frac{1 - \theta}{\mu} \|h\|_{x,F}^* \leq \|h\|_{s,F_*} \leq \frac{1}{(1 - \theta)\mu} \|h\|_{x,F}^*$$

Proof. As noted previously, $F'(x) \in -\text{int}(K^*)$ for all $x \in \text{int}(K)$. Moreover, $\mu > 0$. So $t := -\mu F'(x) \in \text{int}(K^*)$. Since $x \in \text{int}(K)$ and $\mu > 0$, we may invoke Lemma 2.6. Following this, we use (2.11) to obtain

$$r := \|t - s\|_{t, F_*} = \|\mu F'(x) + s\|_{-\mu F'(x), F_*} = \frac{1}{\mu} \|\mu F'(x) + s\|_{x, F}^* \leq \theta < 1.$$

Now applying Lemma 2.2(a) and using $r \leq \theta$, we obtain for any h ,

$$(2.12) \quad (1 - \theta) \|h\|_{-\mu F'(x), F_*} \leq \|h\|_{s, F_*} \leq \frac{1}{1 - \theta} \|h\|_{-\mu F'(x), F_*}.$$

Applying Lemma 2.6 again gives the required result. \square

3. An inexact primal-dual interior-point method for conic optimization. In this section we define our problem, give an overview of the short-step primal-dual IPM, and then present an algorithm that does not require exact evaluation of the primal or dual barrier functions or their derivatives.

3.1. Conic optimization problem. Any primal-dual pair of convex optimization problems can be written in the following conic form:

$$(3.1) \quad v_P = \inf_x \{ \langle c, x \rangle : Ax = b, x \in K \},$$

$$(3.2) \quad v_D = \sup_{w, s} \{ \langle b, w \rangle : A^*w + s = c, s \in K^* \},$$

where $A : \mathcal{R}^n \rightarrow \mathcal{R}^m$, $b \in \mathcal{R}^m$, $c \in \mathcal{R}^n$, and $K \subset \mathcal{R}^n$ is a closed convex cone. A point x is said to be *strongly feasible* for (3.1) if $Ax = b$ and $x \in \text{int}(K)$, and a pair (w, s) is said to be *strongly feasible* for (3.2) if $A^*w + s = c$ and $s \in \text{int}(K^*)$. In the case that K is the nonnegative orthant, (3.1) and (3.2) are called linear optimization problems. If K is the second-order cone $\{x : x_n \geq \|x_{1:n-1}\|_2\}$, then (3.1) and (3.2) are called second-order cone optimization problems.⁴ If K is the proper cone of positive semidefinite matrices, considered as a subset of the space of symmetric matrices, then (3.1) and (3.2) are called semidefinite programs.⁵

Assumption 3.1.

- (a) $m > 0$ and A is onto;
- (b) K is a proper cone;
- (c) The optimization problems (3.1) and (3.2) each satisfy the generalized Slater constraint qualification; i.e., there exists a strongly feasible x for (3.1) and a strongly feasible pair (w, s) for (3.2).

It follows from Assumption 3.1(b) that K^* is also a proper cone. Assumption 3.1(c) can be made without loss of generality, since if it fails to hold, one can embed (3.1)–(3.2) in a higher-dimensional conic problem for which the assumption does hold; see, e.g., [21]. Under Assumption 3.1, the optimal primal and dual solution sets are nonempty and bounded (see [17, Theorem 1]), and the duality gap is zero, i.e., $v_P = v_D$; see, e.g., [26, Corollary 30.5.2].

Consider the *barrier problem* associated with (3.1), viz.,

$$(3.3) \quad v_P(\mu) = \inf_x \{ \langle c, x \rangle + \mu F(x) : Ax = b \},$$

⁴The second-order cone is also known as the Lorentz cone or ice-cream cone.

⁵The space of symmetric matrices of order n is, of course, isomorphic to $\mathcal{R}^{n(n+1)/2}$.

where $\mu > 0$ is called the *barrier parameter*, and the function $F : \text{int}(K) \rightarrow \mathcal{R}$ satisfies the following assumption.

Assumption 3.2. The function F is a ν -logarithmically homogeneous self-concordant barrier for K .

Such an F is known to exist for every proper cone K [20, Theorem 2.5.1]. As mentioned previously, the conjugate function F_* of such an F is a ν -logarithmically homogeneous self-concordant barrier for K^* , and hence is a suitable barrier for (3.2). The resulting dual barrier problem is

$$(3.4) \quad v_D(\mu) = \sup_{w,s} \{ \langle b, w \rangle - \mu F_*(s) : A^*w + s = c \}.$$

It was shown in [17, Lemma 1] that under the generalized Slater constraint qualification, (3.3) and (3.4) have a unique minimizer for each $\mu > 0$.

3.2. Primal-dual interior-point methods for conic problems. Of special importance in primal-dual interior-point algorithms is the set of triples (x, w, s) such that for some $\mu > 0$, x is a minimizer of the primal barrier problem and (w, s) is a minimizer of the dual barrier problem. This set is called the *primal-dual central path*. To exploit the fact that the primal-dual central path is a curve culminating at a point in the primal-dual optimal solution set, we use an iterative algorithm whose iterates stay close to the primal-dual central path while also converging to the primal-dual optimal solution set.

It is well known that under Assumption 3.1 the primal-dual central path consists of the triples satisfying the following system:

$$(3.5) \quad \begin{aligned} Ax &= b \\ A^*w + s &= c \\ \mu F'(x) + s &= 0 \\ x \in \text{int}(K), \quad s &\in \text{int}(K^*). \end{aligned}$$

For any (x, w, s) on the primal-dual central path the pointwise duality gap is

$$(3.6) \quad \langle c, x \rangle - \langle b, w \rangle = \langle x, s \rangle = \langle x, -\mu F'(x) \rangle = -\mu \langle x, F'(x) \rangle = \nu \mu,$$

where the last equality follows from Lemma 2.5(b). Therefore to follow the primal-dual central path to the optimal primal-dual solution set, we decrease the positive *duality measure*

$$(3.7) \quad \mu = \frac{\langle x, s \rangle}{\nu},$$

of the triple (x, w, s) to zero, at which point x solves (3.1) and (w, s) solves (3.2).

Applying Newton’s method to the system of equations in (3.5), we obtain the linear system

$$(3.8) \quad \begin{bmatrix} A & 0 & 0 \\ 0 & A^* & I \\ \mu F''(x) & 0 & I \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta w \\ \Delta s \end{bmatrix} = \begin{bmatrix} b - Ax \\ c - A^*w - s \\ -\tau \mu F'(x) - s \end{bmatrix},$$

with $\tau = 1$.⁶

⁶Strictly speaking, A , A^* , and $F''(x)$, and hence the block matrix on the left-hand side of (3.8), are linear mappings rather than matrices, but the meaning of (3.8) is clear.

It is well known that when F is a logarithmically homogeneous barrier for K , the primal-dual central path is unchanged when the condition $\mu F'(x) + s = 0$ in (3.5) is replaced by $\mu F'_*(s) + x = 0$. Such a replacement gives rise to a “dual linearization” analogous to the “primal linearization” in (3.8). In general these two approaches yield different Newton directions. The advantage of using (3.8) is that evaluation of F_* and its derivatives is not required. On the other hand, the direction obtained from (3.8) is biased in favor of solving the primal barrier problem.

Since we intend for the algorithm stated below to produce iterates staying close to the primal-dual central path while making progress towards the optimal primal-dual solution set, it is necessary to quantify what it means for a point to lie close to the central path. Given $\theta \in (0, 1)$, define the $\mathcal{N}(\theta)$ neighborhood of the primal-dual central path by

$$(3.9) \quad \mathcal{N}(\theta) := \left\{ (x, w, s) : (x, w, s) \text{ is strongly feasible for (3.1)–(3.2), and} \right. \\ \left. \|s + \mu F'(x)\|_{x,F}^* \leq \theta\mu, \text{ where } \mu = \frac{\langle x, s \rangle}{\nu} \right\}.$$

The neighborhood $\mathcal{N}(\theta)$ defined in (3.9) was used in [23, Section 6] for optimization over the class of *self-scaled* cones. In the case that K is the nonnegative orthant and $F(x) = -\sum_i \log(x_i)$ is the standard logarithmic barrier function, we have $\|s + \mu F'(x)\|_{x,F}^* = \|Xs - \mu e\|_2$, so $\mathcal{N}(\theta)$ is the familiar $\mathcal{N}_2(\theta)$ neighborhood used in linear optimization; see, e.g., [31, p. 9]. Note that points in the set $\mathcal{N}(\theta)$ satisfy all conditions in (3.5) with the possible exception of the system of equations $s + \mu F'(x) = 0$, whose residual is small (as measured in the $\|\cdot\|_{x,F}^*$ norm).

We will use a primal-dual short-step algorithm to solve (3.1) and (3.2). Short-step algorithms date back to an important paper of Renegar [24], in which a polynomial-time primal algorithm for linear optimization was given. The name “short-step” arises from the fact that this class of algorithms generates at each iteration Newton steps that are “short” enough to be feasible without the need for a line search. This is an advantage in conic optimization, since line searches may be expensive and difficult for many classes of cones K . The major downside is that such Newton steps are usually too conservative; in practice larger steps are often possible, leading to a faster reduction in the duality measure, and hence a smaller number of iterations.

A standard short-step primal-dual feasible-point algorithm uses (3.8) to compute a sequence of steps for a sequence of μ values converging to zero. The algorithm uses two parameters. One is $\theta \in (0, 1)$, which stipulates the width of the neighborhood $\mathcal{N}(\theta)$ inside which the iterates are constrained to lie, and the other is a *centering parameter* $\tau \in (0, 1)$. (See, e.g., [31, p. 8], where the parameter is denoted by σ .)

We terminate the iteration upon finding an ε -optimal solution of (3.1)–(3.2), which is a feasible primal-dual point whose duality measure is no greater than ε . We show in section 4 that by choosing τ and θ appropriately, it can be ensured that all iterates of the algorithm indeed stay in the neighborhood $\mathcal{N}(\theta)$, even if $F'(x)$ and $F''(x)$ in (3.8) and (3.9) are not computed exactly, provided close enough approximations are used. Moreover, the duality measure decreases geometrically at each iteration, insuring convergence to an optimal solution in a reasonable number of iterations.

3.3. An inexact interior-point method. The interior-point method we propose uses *inexact barrier function evaluations*; i.e., for a given $x \in \text{int}(K)$, the quantities $F'(x)$ and $F''(x)$ in (3.8) are computed inexactly. Denote these estimates by

$F_1(x)$ and $F_2(x)$, respectively. Our analysis is independent of the estimation algorithm and depends only on the magnitude of the approximation errors. The core of the algorithm is the following update procedure (cf. (3.8)).

Iteration Short_step
 Input: $\theta, \tau \in (0, 1)$ and a triple $(x, w, s) \in \mathcal{N}(\theta)$.
 (1) Let $\mu = \langle x, s \rangle / \nu$ and solve the linear system

$$(3.10) \quad \begin{bmatrix} A & 0 & 0 \\ 0 & A^* & I \\ \mu F_2(x) & 0 & I \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta w \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\tau \mu F_1(x) - s \end{bmatrix},$$

where $F_2(x)$ is self-adjoint and positive definite.
 (2) Set $x_+ = x + \Delta x$, $w_+ = w + \Delta w$, and $s_+ = s + \Delta s$.

Remark 3.3. Iteration Short_step requires the availability of an initial triple (x, w, s) in $\mathcal{N}(\theta)$. This requirement is inherited from other short-step methods. One method of finding such a point is to repeatedly perform Iteration Short_step with $\tau = 1$, but with μ kept constant at, say, its initial value. In other words, we seek to move directly towards a point on the central path. However, to our knowledge, even given a strictly feasible initial triple and using exact computation of the gradient and Hessian of the barrier function, no algorithm has been proposed that reaches an appropriate neighborhood of the central path within a polynomial iteration bound independent of the nearness of the initial point to the boundary of the cone. Accordingly, this question requires further study beyond the scope of this paper.

A noteworthy feature of Iteration Short_step is that its implementation does not require evaluating, or even estimating, numerical values of the dual barrier function F_* or its derivatives. For suitable values of θ and τ , and with $F_1(x)$ and $F_2(x)$ taken to be $F'(x)$ and $F''(x)$, respectively, it is known (at least in the SDP case) to produce an algorithm with polynomial iteration complexity; see, e.g., [1, section 6.5]. As we show below, such a property is preserved when one uses inexact evaluations of the derivatives of F , which can yield a substantial computational advantage.

We now study the effect of errors in our estimates of $F'(x)$ and $F''(x)$ on the convergence of the interior-point method.

4. Convergence analysis of the inexact algorithm. In this section we prove that the inexact algorithm produces a strongly feasible set of iterates and converges in a polynomial number of iterations.

4.1. Feasibility and convergence of the iterates. Denote the errors in the known estimates of the gradient and Hessian of $F(x)$ by

$$E_1(x) = F_1(x) - F'(x), \quad E_2(x) = F_2(x) - F''(x).$$

By definition $E_2(x)$ is a self-adjoint linear mapping. Throughout this section, we will assume that the errors $E_1(x)$ and $E_2(x)$ are “small enough”. Specifically, for some ϵ_1 and ϵ_2 ,

$$(4.1) \quad \frac{\|E_1(x)\|_{x,F}^*}{\|F'(x)\|_{x,F}^*} \equiv \frac{\|E_1(x)\|_{x,F}^*}{\nu^{1/2}} \leq \epsilon_1 < 1 \quad \forall x \in \text{int}(K),$$

$$(4.2) \quad \|F''(x)^{-1/2} E_2(x) F''(x)^{-1/2}\| \leq \epsilon_2 < 1 \quad \forall x \in \text{int}(K),$$

well defined since $F''(x)$ is positive definite for $x \in \text{int}(K)$. The quantity $\frac{\|E_1(x)\|_{x,F}^*}{\|F''(x)\|_{x,F}^*}$ is the relative error in $F_1(x)$, measured in the $\|\cdot\|_{x,F}^*$ norm, and $\|F''(x)^{-1/2}E_2(x)F''(x)^{-1/2}\|$ is the absolute error in $F_2(x)$ relative to the Hessian of F . It is also an upper bound on the relative error in $F_2(x)$, measured in the norm $\|\cdot\|$:

$$\begin{aligned} \frac{\|E_2(x)\|}{\|F''(x)\|} &= \frac{\|F''(x)^{1/2}F''(x)^{-1/2}E_2(x)F''(x)^{-1/2}F''(x)^{1/2}\|}{\|F''(x)\|} \\ &\leq \frac{\|F''(x)^{1/2}\| \|F''(x)^{-1/2}E_2(x)F''(x)^{-1/2}\| \|F''(x)^{1/2}\|}{\|F''(x)\|} \\ &= \frac{\|F''(x)^{1/2}\|^2}{\|F''(x)\|} \|F''(x)^{-1/2}E_2(x)F''(x)^{-1/2}\| \\ &= \|F''(x)^{-1/2}E_2(x)F''(x)^{-1/2}\|. \end{aligned}$$

The last equality follows from the relation $\|B^2\| = \lambda_{\max}(B^2) = \lambda_{\max}(B)^2 = \|B\|^2$ for a positive semidefinite self-adjoint operator B , where $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue.

We now show that under the assumption in (4.2), the eigenvalues of $F_2(x)$ are close to those of $F''(x)$. We will use \preceq to denote a partial ordering with respect to the cone of positive semidefinite matrices. Namely, for symmetric matrices M_1 and M_2 , $M_1 \preceq M_2$ if and only if $M_2 - M_1$ is positive semidefinite.

LEMMA 4.1. *Suppose that F'' and its estimate F_2 satisfy (4.2). Then*

$$(1 - \epsilon_2)F''(x) \preceq F_2(x) \preceq (1 + \epsilon_2)F''(x) \quad \forall x \in \text{int}(K).$$

Moreover $F_2(x)$ is positive definite for all $x \in \text{int}(K)$.

Proof. The assumption in (4.2) and the definition of a norm induced by an inner product implies that for all $x \in \text{int}(K)$, each eigenvalue of the self-adjoint operator $F''(x)^{-1/2}E_2(x)F''(x)^{-1/2}$ is bounded by ϵ_2 , so

$$-\epsilon_2 I \preceq F''(x)^{-1/2}E_2(x)F''(x)^{-1/2} \preceq \epsilon_2 I.$$

Multiplying this matrix inequality on the left and right by the positive definite matrix $F''(x)^{1/2}$ preserves the partial ordering \preceq :

$$-\epsilon_2 F''(x) \preceq E_2(x) \preceq \epsilon_2 F''(x).$$

Adding $F''(x)$ to each quantity in this matrix inequality gives the required result. Since $F''(x)$ is positive definite, so F_2 is also. \square

Since $F_2(x)$ is positive definite for all $x \in \text{int}(K)$, it follows that $F_2(x)^{-1}$ is well defined. Due to this and A being onto (Assumption 3.1), (3.10) has a unique solution.

The outline for the remainder of this section is as follows. After giving some technical results, we prove that for a primal-dual iterate (x, w, s) lying inside a neighborhood $\mathcal{N}(\theta)$ of the central path, the Newton steps Δx and Δs in (3.10) are bounded by a constant and a constant times the duality measure μ , respectively, in the appropriate norms (Lemma 4.3 and Corollary 4.4). We then show that if these constants are appropriately bounded, then a full Newton step produces a strongly feasible primal-dual point (Lemma 4.5). Next we derive lower and upper bounds on the rate of decrease of the duality measure at each iteration (Lemma 4.6). The lower bound is used to show that all iterates stay in the $\mathcal{N}(\theta)$ neighborhood of the central path (Lemmas 4.7, 4.8, and 4.9). Thus Iteration Short_step can be repeated ad infinitum,

and the resulting algorithm belongs to the class of “path-following” methods. It also belongs to the class of feasible-point algorithms: All iterates are (strongly) feasible. Finally, using the upper bound on the rate of decrease of the duality measure, we show that the algorithm has polynomial iteration complexity.

Since $F''(x)$ and $F_2(x)$ are positive definite on $\text{int}(K)$, the positive definite square roots of these two matrices are well defined. Define

$$(4.3) \quad D(x) = F''(x)^{1/2} F_2(x)^{-1/2} = F''(x)^{1/2} (F''(x) + E_2(x))^{-1/2}$$

for all $x \in \text{int}(K)$. In this section we will use the following implication of the Cauchy-Schwarz inequality: given $x \in \text{int}(K)$ and $y, z \in \mathcal{R}^n$,

$$|\langle y, z \rangle| \leq \|y\|_{x,F}^* \|z\|_{x,F}.$$

The following technical results will be used in our analysis.

LEMMA 4.2. *Suppose that F_1 and F_2 satisfy (4.1) and (4.2), and let $x \in \text{int}(K)$ and $z \in \mathcal{R}^n$. Let $D(x)$ be defined as in (4.3). Then*

$$(4.4a) \quad \|D(x)\|^2 \leq \frac{1}{1 - \epsilon_2},$$

$$(4.4b) \quad \|D(x)^{-1}\|^2 \leq 1 + \epsilon_2,$$

$$(4.4c) \quad |\langle F'(x), z \rangle| \leq \nu^{1/2} \|z\|_{x,F},$$

$$(4.4d) \quad |\langle E_1(x), z \rangle| \leq \epsilon_1 \nu^{1/2} \|z\|_{x,F},$$

$$(4.4e) \quad |\langle E_1(x), x \rangle| \leq \epsilon_1 \nu,$$

$$(4.4f) \quad |\langle x, E_2(x)z \rangle| \leq \epsilon_2 \nu^{1/2} \|z\|_{x,F}.$$

Proof. First, since $\|D(x)\|$ is the square root of the largest eigenvalue of $D(x)D(x)^*$, we have

$$\begin{aligned} \|D(x)\|^2 &= \|D(x)D(x)^*\| = \|F''(x)^{1/2} (F''(x) + E_2(x))^{-1} F''(x)^{1/2}\| \\ &= \|(I + F''(x)^{-1/2} E_2(x) F''(x)^{-1/2})^{-1}\| \\ &\leq \frac{1}{1 - \|F''(x)^{-1/2} E_2(x) F''(x)^{-1/2}\|} \\ &\leq \frac{1}{1 - \epsilon_2}, \end{aligned}$$

where the inequalities follow from (4.2). This proves (4.4a). Next, the invertible mapping D satisfies

$$\begin{aligned} \|D(x)^{-1}\|^2 &= \|D(x)^{-*} D(x)^{-1}\| = \|F''(x)^{-1/2} (F''(x) + E_2(x)) F''(x)^{-1/2}\| \\ &= \|I + F''(x)^{-1/2} E_2(x) F''(x)^{-1/2}\| \\ &\leq 1 + \|F''(x)^{-1/2} E_2(x) F''(x)^{-1/2}\| \\ &\leq 1 + \epsilon_2, \end{aligned}$$

where the inequalities again follow from (4.2), proving (4.4b). Further,

$$\begin{aligned} |\langle F'(x), z \rangle| &\leq \|F'(x)\|_{x,F}^* \|z\|_{x,F} \\ &= \nu^{1/2} \|z\|_{x,F}, \end{aligned}$$

where we have used Lemma 2.5(e), proving (4.4c). Now we have

$$\begin{aligned} |\langle E_1(x), z \rangle| &\leq \|E_1(x)\|_{x,F}^* \|z\|_{x,F} \\ &\leq \epsilon_1 \nu^{1/2} \|z\|_{x,F}, \end{aligned}$$

where the last inequality follows from (4.1), proving (4.4d). The inequality in (4.4e) follows from (4.4d) with $z = x$, where Lemma 2.5(d) has also been used. Finally, using Lemma 2.5(d) and (4.2), we have

$$\begin{aligned} |\langle x, E_2(x)z \rangle| &\leq \|E_2(x)^* x\|_{x,F}^* \|z\|_{x,F} \\ &= \|E_2(x)^* F''(x)^{-1/2} \cdot F''(x)^{1/2} x\|_{x,F}^* \|z\|_{x,F} \\ &\leq \|F''(x)^{-1/2} E_2(x) F''(x)^{-1/2}\| \|F''(x)^{1/2} x\| \|z\|_{x,F} \\ &\leq \epsilon_2 \nu^{1/2} \|z\|_{x,F}, \end{aligned}$$

proving (4.4f). \square

Define the following three constants, which depend on the IPM parameters θ and τ , the maximum errors ϵ_1 and ϵ_2 , and the complexity parameter $\nu \geq 1$.

$$\begin{aligned} (4.5) \quad \beta_0 &:= \frac{\theta + (1 - \tau + \tau\epsilon_1)\nu^{1/2}}{(1 - \epsilon_2)^{1/2}}, \\ \beta_1 &:= \frac{\beta_0}{(1 - \epsilon_2)^{1/2}}, \\ \beta_2 &:= \max \left\{ \beta_1, \beta_0 \frac{(1 + \epsilon_2)^{1/2}}{1 - \theta} \right\}. \end{aligned}$$

LEMMA 4.3. *Suppose that F_1 and F_2 satisfy (4.1) and (4.2). Let $\theta \in (0, 1)$, and suppose that $(x, w, s) \in \mathcal{N}(\theta)$ for some w , and that $x, s, \Delta x$, and Δs satisfy (3.10) for some Δw . Then*

$$\mu^2 \|F_2(x)^{1/2} \Delta x\|^2 + \|F_2(x)^{-1/2} \Delta s\|^2 \leq \mu^2 \beta_0^2.$$

Proof. Premultiplying the third block equation in (3.10) by $(\mu F_2(x))^{-1/2}$ yields:

$$(4.6) \quad (\mu F_2(x))^{1/2} \Delta x + (\mu F_2(x))^{-1/2} \Delta s = -(\mu F_2(x))^{-1/2} (\tau \mu F_1(x) + s).$$

It is seen from (3.10) that Δx lies in the nullspace of A , while Δs lies in the range of A^* . Therefore Δx is orthogonal to Δs . Taking the square of the norm of each side of (4.6) and using this orthogonality, we obtain

$$\|(\mu F_2(x))^{1/2} \Delta x\|^2 + \|(\mu F_2(x))^{-1/2} \Delta s\|^2 = \|(\mu F_2(x))^{-1/2} (\tau \mu F_1(x) + s)\|^2.$$

Multiplying this equation by μ yields

$$(4.7) \quad \mu^2 \|F_2(x)^{1/2} \Delta x\|^2 + \|F_2(x)^{-1/2} \Delta s\|^2 = \|F_2(x)^{-1/2} (\tau \mu F_1(x) + s)\|^2.$$

Let us now bound the right-hand side of (4.7). In the following, (4.8a) follows from (4.3), and (4.8b) follows from (4.4a). The inequality in (4.8c) follows from $(x, w, s) \in$

$\mathcal{N}(\theta)$, (4.1), and Lemma 2.5(e).

$$\begin{aligned}
 (4.8a) \quad & \|F_2(x)^{-1/2}(\tau\mu F_1(x) + s)\| \\
 &= \|D(x)^* F''(x)^{-1/2}(\tau\mu F_1(x) + s)\| \\
 (4.8b) \quad & \leq \frac{1}{(1 - \epsilon_2)^{1/2}} \|\tau\mu F_1(x) + s\|_{x,F}^* \\
 &= \frac{1}{(1 - \epsilon_2)^{1/2}} \|(s + \mu F'(x)) + \tau\mu E_1(x) - \mu(1 - \tau)F'(x)\|_{x,F}^* \\
 & \leq \frac{1}{(1 - \epsilon_2)^{1/2}} (\|s + \mu F'(x)\|_{x,F}^* + \tau\mu \|E_1(x)\|_{x,F}^* + \mu(1 - \tau) \|F'(x)\|_{x,F}^*) \\
 (4.8c) \quad & \leq \frac{1}{(1 - \epsilon_2)^{1/2}} (\theta\mu + \tau\mu\epsilon_1\nu^{1/2} + \mu(1 - \tau)\nu^{1/2}) \\
 &= \mu\beta_0.
 \end{aligned}$$

Combining this with (4.7) yields the required result. \square

COROLLARY 4.4. *Suppose that F_1 and F_2 satisfy (4.1) and (4.2). Let $\theta \in (0, 1)$, and suppose that $(x, w, s) \in \mathcal{N}(\theta)$ for some w , and that $x, s, \Delta x$, and Δs satisfy (3.10) for some Δw . Then*

$$\begin{aligned}
 \|F_2(x)^{1/2}\Delta x\| &\leq \beta_0, \\
 \|F_2(x)^{-1/2}\Delta s\| &\leq \mu\beta_0, \\
 \|\Delta x\|_{x,F} &\leq \beta_1, \\
 \|\Delta s\|_{x,F}^* &\leq (1 + \epsilon_2)^{1/2}\mu\beta_0.
 \end{aligned}$$

Proof. The first two bounds follow immediately from Lemma 4.3. Further, using (4.3) and (4.4a), we have

$$\|\Delta x\|_{x,F} = \|D(x)F_2(x)^{1/2}\Delta x\| \leq \|D(x)\| \|F_2(x)^{1/2}\Delta x\| \leq (1 - \epsilon_2)^{-1/2}\beta_0 = \beta_1.$$

This proves the third inequality. To prove the last, similarly use (4.3) and (4.4b) to obtain

$$\begin{aligned}
 \|\Delta s\|_{x,F}^* &= \|D(x)^{-*}F_2(x)^{-1/2}\Delta s\| \leq (1 + \epsilon_2)^{1/2}\|F_2(x)^{-1/2}\Delta s\| \\
 &\leq (1 + \epsilon_2)^{1/2}\mu\beta_0. \quad \square
 \end{aligned}$$

Iteration `Short_step` takes a full primal-dual Newton step from a triple $(x, w, s) \in \mathcal{N}(\theta)$. In the next series of lemmas we show under a condition on the parameters θ, τ, ϵ_1 , and ϵ_2 , not only that such a step is strongly feasible, justifying step (2) in Iteration `Short_step`, but further, that the new iterate remains in the $\mathcal{N}(\theta)$ neighborhood of the central path, so Iteration `Short_step` can be repeated ad infinitum. We also characterize how the duality measure changes after taking a full Newton step. The following condition, in which dependence of β_2 on $\theta, \tau, \epsilon_1, \epsilon_2$, and ν is emphasized, will play a key role:

$$(4.9) \quad \beta_2(\theta, \tau, \epsilon_1, \epsilon_2, \nu) < 1.$$

LEMMA 4.5. *Suppose that F_1 and F_2 satisfy (4.1) and (4.2). Let $\theta, \tau, \epsilon_1, \epsilon_2$, and ν be such that (4.9) holds, and let $(x, w, s) \in \mathcal{N}(\theta)$. Then the point (x_+, w_+, s_+) generated by Iteration `Short_step` is strongly feasible.*

Proof. The equality constraints in (3.1) and (3.2) are satisfied by (x_+, w_+, s_+) , since they are satisfied by (x, w, s) (by virtue of this triple lying in $\mathcal{N}(\theta)$), and any step from (x, w, s) in the direction $(\Delta x, \Delta w, \Delta s)$ will satisfy these constraints due to the first two block equations in (3.10). We now show that $x_+ \in \text{int}(K)$ and $s_+ \in \text{int}(K^*)$. Since F is a nondegenerate self-concordant barrier (Assumption 3.2), Lemma 2.1(b) is applicable: If $\|\Delta x\|_{x,F} < 1$, then $x + \Delta x \in \text{int}(K)$. By Corollary 4.4, a sufficient condition for $\|\Delta x\|_{x,F} < 1$ is $\beta_1 < 1$, and this holds since $\beta_2 < 1$. Similarly, in light of Lemma 2.2(b), if $\|\Delta s\|_{s,F_*} < 1$, then $s + \Delta s \in \text{int}(K^*)$. It remains to show that $\|\Delta s\|_{s,F_*} < 1$ under the assumption (4.9). Since $(x, w, s) \in \mathcal{N}(\theta)$, we may apply Lemma 2.7 with $h = \Delta s$ to obtain

$$\|\Delta s\|_{s,F_*} \leq \frac{\|\Delta s\|_{x,F}^*}{(1 - \theta)\mu} \leq \frac{(1 + \epsilon_2)^{1/2}\beta_0}{1 - \theta},$$

where the last inequality follows from Corollary 4.4. By the definition of β_2 in (4.5), we have $\|\Delta s\|_{s,F_*} \leq \beta_2 < 1$. \square

It follows from Lemma 4.5 that a full step in the Newton direction for (3.10) is strongly feasible. We now study the change in the duality measure after taking a full primal-dual Newton step.

LEMMA 4.6. *Suppose that F_1 and F_2 satisfy (4.1) and (4.2). Let $\theta, \tau, \epsilon_1, \epsilon_2$, and ν be such that (4.9) holds, and let $(x, w, s) \in \mathcal{N}(\theta)$. Then the duality measure μ_+ of (x_+, w_+, s_+) (see (3.7)) in Iteration Short_step satisfies*

$$(4.10) \quad \underline{\delta}\mu \leq \mu_+ \leq \bar{\delta}\mu,$$

where

$$(4.11) \quad \underline{\delta} = \tau - \tau\epsilon_1 - \frac{1 - \tau + \tau\epsilon_1 + \epsilon_2}{\nu^{1/2}}\beta_1 - \frac{1 - \epsilon_2}{\nu}\beta_1^2,$$

$$(4.12) \quad \bar{\delta} = \tau + \tau\epsilon_1 + \frac{1 - \tau + \tau\epsilon_1 + \epsilon_2}{\nu^{1/2}}\phi - \frac{1 - \epsilon_2}{\nu}\phi^2,$$

with

$$(4.13) \quad \phi = \min \left\{ \beta_1, \frac{(1 - \tau + \tau\epsilon_1 + \epsilon_2)\nu^{1/2}}{2(1 - \epsilon_2)} \right\}.$$

Proof. Recalling that for the direction obtained from (3.8), Δx is orthogonal to Δs , we have

$$(4.14) \quad \begin{aligned} \nu\mu_+ &= \langle x_+, s_+ \rangle \\ &= \langle x + \Delta x, s + \Delta s \rangle \\ &= \langle x, s + \Delta s \rangle + \langle \Delta x, s \rangle. \end{aligned}$$

From the third block equation in (3.10), we obtain

$$(4.15) \quad s + \Delta s = -\tau\mu F_1(x) - \mu F_2(x)\Delta x,$$

giving

$$(4.16) \quad \begin{aligned} \langle x, s + \Delta s \rangle &= \langle -x, \tau\mu F_1(x) + \mu F_2(x)\Delta x \rangle \\ &= -\tau\mu \langle x, F'(x) + E_1(x) \rangle - \mu \langle x, (F''(x) + E_2(x))\Delta x \rangle \\ &= \tau\mu\nu - \tau\mu \langle x, E_1(x) \rangle + \mu \langle F'(x), \Delta x \rangle - \mu \langle x, E_2(x)\Delta x \rangle, \end{aligned}$$

where we have used Lemma 2.5(b),(c). Since $\langle \Delta x, \Delta s \rangle = 0$, it also follows from (4.15) that

$$(4.17) \quad \begin{aligned} \langle \Delta x, s \rangle &= \langle \Delta x, -\mu F_2(x)\Delta x - \tau \mu F_1(x) \rangle \\ &= -\mu \langle \Delta x, F_2(x)\Delta x \rangle - \tau \mu \langle F'(x) + E_1(x), \Delta x \rangle. \end{aligned}$$

Combining (4.14), (4.16), and (4.17) we have

$$\begin{aligned} \nu \mu_+ &= \tau \mu \nu - \tau \mu \langle x, E_1(x) \rangle + (1 - \tau) \mu \langle F'(x), \Delta x \rangle - \tau \mu \langle E_1(x), \Delta x \rangle \\ &\quad - \mu \langle x, E_2(x)\Delta x \rangle - \mu \langle \Delta x, F_2(x)\Delta x \rangle; \end{aligned}$$

i.e.,

$$(4.18) \quad \begin{aligned} \frac{\mu_+}{\mu} &= \tau - \frac{\tau}{\nu} \langle x, E_1(x) \rangle + \frac{1 - \tau}{\nu} \langle F'(x), \Delta x \rangle - \frac{\tau}{\nu} \langle E_1(x), \Delta x \rangle \\ &\quad - \frac{1}{\nu} \langle x, E_2(x)\Delta x \rangle - \frac{1}{\nu} \langle \Delta x, F_2(x)\Delta x \rangle. \end{aligned}$$

To reduce clutter, let $t := \|\Delta x\|_{x,F}$. In view of Lemma 4.1, we have $\langle \Delta x, F_2(x)\Delta x \rangle \geq (1 - \epsilon_2)t^2$. Using this inequality and appealing to Lemma 4.2, we obtain the following upper bound on μ_+/μ :

$$\begin{aligned} \frac{\mu_+}{\mu} &\leq \tau + \frac{\tau}{\nu} |\langle x, E_1(x) \rangle| + \frac{1 - \tau}{\nu} |\langle F'(x), \Delta x \rangle| + \frac{\tau}{\nu} |\langle E_1(x), \Delta x \rangle| \\ &\quad + \frac{1}{\nu} |\langle x, E_2(x)\Delta x \rangle| - \frac{1 - \epsilon_2}{\nu} t^2 \\ &\leq \tau + \tau \epsilon_1 + \frac{1 - \tau}{\nu} \nu^{1/2} t + \frac{\tau}{\nu} \epsilon_1 \nu^{1/2} t + \frac{1}{\nu} \epsilon_2 \nu^{1/2} t - \frac{1 - \epsilon_2}{\nu} t^2 \\ &=: u(t). \end{aligned}$$

It follows from Corollary 4.4 that $t \leq \beta_1$, so a uniform upper bound on μ_+/μ is $\max\{u(t) : 0 \leq t \leq \beta_1\}$. The unconstrained maximizer is given by

$$t^* = \frac{(1 - \tau + \tau \epsilon_1 + \epsilon_2) \nu^{1/2}}{2(1 - \epsilon_2)}.$$

If this nonnegative solution satisfies $t^* \leq \beta_1$, then the constrained maximum of $u(t)$ is $u(t^*)$. Otherwise the maximum is $u(\beta_1)$. Hence

$$\frac{\mu_+}{\mu} \leq u(\min\{\beta_1, t^*\}),$$

and the latter bound is $\bar{\delta}$ as given in (4.12) and (4.13). This proves the second inequality in (4.10).

In view of Corollary 4.4, $\langle \Delta x, F_2(x)\Delta x \rangle \leq \beta_0^2 = (1 - \epsilon_2)\beta_1^2$, so from (4.18) we can also obtain a lower bound on μ_+/μ :

$$\begin{aligned} \frac{\mu_+}{\mu} &\geq \tau - \tau \epsilon_1 - \frac{1 - \tau}{\nu} \nu^{1/2} t - \frac{\tau}{\nu} \epsilon_1 \nu^{1/2} t - \frac{1}{\nu} \epsilon_2 \nu^{1/2} t - \frac{1 - \epsilon_2}{\nu} \beta_1^2 \\ &= \tau - \tau \epsilon_1 - \frac{1 - \epsilon_2}{\nu} \beta_1^2 - \left(\frac{1 - \tau + \tau \epsilon_1 + \epsilon_2}{\nu^{1/2}} \right) t \\ &\geq \tau - \tau \epsilon_1 - \frac{1 - \epsilon_2}{\nu} \beta_1^2 - \left(\frac{1 - \tau + \tau \epsilon_1 + \epsilon_2}{\nu^{1/2}} \right) \beta_1 \\ &= \underline{\delta}, \end{aligned}$$

proving the first inequality in (4.10). \square

LEMMA 4.7. *The quantities $\underline{\delta}$ and $\bar{\delta}$ defined in Lemma 4.6 satisfy $\underline{\delta} < \tau < \bar{\delta}$.*

Proof. It is clear from (4.11) that $\underline{\delta} < \tau$. The inequality $\tau < \bar{\delta}$ is easily seen after rewriting $\bar{\delta}$ as

$$\bar{\delta} = \tau + \tau\epsilon_1 + \frac{\phi(1 - \epsilon_2)}{\nu} \left[\frac{(1 - \tau + \tau\epsilon_1 + \epsilon_2)\nu^{1/2}}{1 - \epsilon_2} - \phi \right],$$

and noting that the term in the square brackets must be positive, due to (4.13). □

We will now show that condition (4.9) together with the following two conditions, where dependence on $\theta, \tau, \epsilon_1, \epsilon_2,$ and ν is again emphasized, insures that the next iterate lies in $\mathcal{N}(\theta)$:

$$(4.19) \quad \underline{\delta}(\theta, \tau, \epsilon_1, \epsilon_2, \nu) > 0,$$

$$(4.20) \quad \theta_+(\theta, \tau, \epsilon_1, \epsilon_2, \nu) \text{ is well defined, and } \theta_+(\theta, \tau, \epsilon_1, \epsilon_2, \nu) \leq \theta,$$

where

$$(4.21) \quad \begin{aligned} \theta_+ := & \frac{\tau\epsilon_1\nu^{1/2} + \epsilon_2\beta_1}{\underline{\delta}(1 - \beta_1)} + \left(\frac{\tau}{\underline{\delta}} - 1\right)\nu^{1/2} + \frac{(1 - \tau)\beta_1}{\underline{\delta}(1 - \beta_1)} \\ & + \frac{2\tau}{\underline{\delta}} \left(\log(1 - \beta_1) + \frac{\beta_1}{1 - \beta_1} \right). \end{aligned}$$

The next lemma will prove useful in the sequel.

LEMMA 4.8. *If $\theta, \tau, \epsilon_1, \epsilon_2,$ and ν satisfy conditions (4.19) and (4.20), then $\beta_1, \beta_2 < 1$.*

Proof. If θ_+ is well defined, then clearly $\beta_1 < 1$. Moreover, for all $\beta_1 \in (0, 1)$, $\log(1 - \beta_1) + \frac{\beta_1}{1 - \beta_1} > 0$. It follows from Lemma 4.7 and (4.19) that $\tau/\underline{\delta} > 1$. Hence

$$(4.22) \quad 2 \left(\log(1 - \beta_1) + \frac{\beta_1}{1 - \beta_1} \right) < \frac{2\tau}{\underline{\delta}} \left(\log(1 - \beta_1) + \frac{\beta_1}{1 - \beta_1} \right) < \theta_+ \leq \theta < \beta_1.$$

The inequality $2 \log(1 - \beta_1) + \frac{2\beta_1}{1 - \beta_1} < \beta_1$ implies $\beta_1 < 0.46$. Now from (4.5) we have

$$\begin{aligned} \beta_2 &= \max \left\{ \beta_1, \beta_0 \frac{(1 + \epsilon_2)^{1/2}}{1 - \theta} \right\} \\ &= \beta_1 \max \left\{ 1, \frac{(1 - \epsilon_2^2)^{1/2}}{1 - \theta} \right\} \\ &\leq \frac{\beta_1}{1 - \theta} \\ &< \frac{\beta_1}{1 - \beta_1} \\ &< \frac{0.46}{1 - 0.46} < 1. \quad \square \end{aligned}$$

LEMMA 4.9. *Let $\theta, \tau, \epsilon_1, \epsilon_2,$ and ν satisfy conditions (4.19) and (4.20). If $(x, w, s) \in \mathcal{N}(\theta)$, then Iteration Short_step produces a triple $(x_+, w_+, s_+) \in \mathcal{N}(\theta)$.*

Proof. By Lemma 4.8, $\beta_1, \beta_2 < 1$, so (4.9) holds. Applying Lemma 4.5, we conclude that (x_+, w_+, s_+) is strongly feasible. It remains to show that

$$(4.23) \quad \|s + \mu F'(x)\|_{x,F}^* \leq \theta\mu \implies \|s_+ + \mu_+ F'(x_+)\|_{x_+,F}^* \leq \theta\mu_+ .$$

From the third block equation in the linear system (3.10), we have

$$(4.24) \quad s_+ + \mu_+ F'(x_+) = -\tau\mu F_1(x) - \mu F_2(x)\Delta x + \mu_+ F'(x_+) = D_1 + D_2 + D_3,$$

where

$$\begin{aligned} D_1 &= -\tau\mu E_1(x) - \mu E_2(x)\Delta x, \\ D_2 &= (\mu_+ - \tau\mu)F'(x_+) + \mu(\tau - 1)F''(x)\Delta x, \\ D_3 &= \tau\mu(F'(x_+) - F'(x) - F''(x)\Delta x). \end{aligned}$$

We now appropriately bound the norms of D_1 , D_2 , and D_3 . First we bound $\|D_1\|_{x_+,F}^*$. Since $\|x_+ - x\|_{x,F} = \|\Delta x\|_{x,F} \leq \beta_1$ (Corollary 4.4) and $\beta_1 < 1$, we can use (2.6) to bound $\|D_1\|_{x_+,F}^*$ in terms of $\|D_1\|_{x,F}^*$; the result is (4.25a). The inequality (4.25b) follows from the definition of D_1 , and (4.25c) follows from (4.1). The inequality (4.25d) follows from (4.2) and Corollary 4.4.

$$(4.25a) \quad \|D_1\|_{x_+,F}^* \leq \frac{1}{1 - \beta_1} \|D_1\|_{x,F}^*$$

$$(4.25b) \quad \leq \frac{1}{1 - \beta_1} (\tau\mu \|E_1(x)\|_{x,F}^* + \mu \|E_2(x)\Delta x\|_{x,F}^*)$$

$$(4.25c) \quad \leq \frac{1}{1 - \beta_1} (\tau\mu\epsilon_1\nu^{1/2} + \mu \|F''(x)^{-1/2} E_2(x) F''(x)^{-1/2} F''(x)^{1/2} \Delta x\|)$$

$$\leq \frac{1}{1 - \beta_1} (\tau\mu\epsilon_1\nu^{1/2} + \mu \|F''(x)^{-1/2} E_2(x) F''(x)^{-1/2}\| \|F''(x)^{1/2} \Delta x\|)$$

$$(4.25d) \quad \leq \frac{1}{1 - \beta_1} (\tau\mu\epsilon_1\nu^{1/2} + \mu\epsilon_2\beta_1).$$

Now in Lemma 4.6 we established that $\underline{\delta}\mu \leq \mu_+$, where $\underline{\delta}$ is given in (4.11). Moreover, by assumption, $\underline{\delta} > 0$. It follows that

$$(4.26) \quad \begin{aligned} \|D_1\|_{x_+,F}^* &\leq \frac{\tau\epsilon_1\nu^{1/2} + \epsilon_2\beta_1}{\underline{\delta}(1 - \beta_1)} \mu_+ \\ &=: d_1\mu_+. \end{aligned}$$

Next we bound $\|D_2\|_{x_+,F}^*$. First note that for $\underline{\delta}$ and $\bar{\delta}$ given in (4.11) and (4.12),

$$\begin{aligned} \underline{\delta} + \bar{\delta} &= 2\tau + (\phi - \beta_1) \frac{1 - \tau + \tau\epsilon_1 + \epsilon_2}{\nu^{1/2}} - \frac{1 - \epsilon_2}{\nu} (\phi^2 + \beta_1^2) \\ &< 2\tau, \end{aligned}$$

where the inequality follows from $\phi \leq \beta_1$; see (4.13). It follows that $\bar{\delta} - \tau < \tau - \underline{\delta}$. Using this, the fact that $\underline{\delta} > 0$, and the result $\tau < \bar{\delta}$ from Lemma 4.7, we have

$$\frac{\tau}{\underline{\delta}} - 1 = \frac{\tau - \underline{\delta}}{\underline{\delta}} \geq \frac{\tau - \underline{\delta}}{\bar{\delta}} > \frac{\bar{\delta} - \tau}{\bar{\delta}} = 1 - \frac{\tau}{\bar{\delta}} > 0.$$

Hence

$$(4.27) \quad \max_{\underline{\delta}\mu \leq \mu_+ \leq \bar{\delta}\mu} \left| 1 - \tau \frac{\mu}{\mu_+} \right| = \max \left\{ \frac{\tau}{\underline{\delta}} - 1, 1 - \frac{\tau}{\bar{\delta}} \right\} = \frac{\tau}{\underline{\delta}} - 1.$$

In the following, (4.28a) follows from the definition of D_2 , and (4.28b) follows from Lemma 2.5(e) and (2.6). Furthermore, (4.28c) follows from (4.27) and Corollary 4.4.

$$(4.28a) \quad \|D_2\|_{x_+,F}^* \leq |\mu_+ - \tau\mu| \|F'(x_+)\|_{x_+,F}^* + \mu(1 - \tau) \|F''(x)\Delta x\|_{x_+,F}^*$$

$$(4.28b) \quad \leq \mu_+ \left| 1 - \tau \frac{\mu}{\mu_+} \right| \nu^{1/2} + \mu(1 - \tau) \frac{\|F''(x)\Delta x\|_{x_+,F}^*}{1 - \|\Delta x\|_{x_+,F}}$$

$$(4.28c) \quad \leq \mu_+ \left(\frac{\tau}{\underline{\delta}} - 1 \right) \nu^{1/2} + \frac{\mu_+}{\underline{\delta}} (1 - \tau) \frac{\beta_1}{1 - \beta_1}$$

$$(4.28d) \quad =: d_2\mu_+.$$

Finally, we bound $\|D_3\|_{x_+,F}^*$. In what follows, we will be working with integrals of vectors and matrices. All such integrals are meant componentwise. Applying the fundamental theorem of calculus to F' , then to F'' , yields

$$D_3 = \tau\mu \int_0^1 (F''(x + t\Delta x) - F''(x))\Delta x \, dt = \tau\mu \int_0^1 \int_0^t F'''(x + u\Delta x)[\Delta x, \Delta x] \, du \, dt,$$

giving

$$\|D_3\|_{x_+,F}^* \leq \tau\mu \int_0^1 \int_0^t \|F'''(x + u\Delta x)[\Delta x, \Delta x]\|_{x_+,F}^* \, du \, dt.$$

Since $\beta_1 < 1$, Corollary 2.4 can be applied, yielding

$$\|F'''(x + u\Delta x)[\Delta x, \Delta x]\|_{x_+,F}^* \leq \frac{2\beta_1^2}{(1 - \beta_1)(1 - u\beta_1)},$$

and⁷

$$(4.29) \quad \begin{aligned} \|D_3\|_{x_+,F}^* &\leq \tau\mu \int_0^1 \int_0^t \frac{2\beta_1^2}{(1 - \beta_1)(1 - u\beta_1)} \, du \, dt \\ &= \tau\mu \left(2 \log(1 - \beta_1) + \frac{2\beta_1}{1 - \beta_1} \right) \end{aligned}$$

$$(4.30) \quad \leq \frac{2\tau}{\underline{\delta}} \left(\log(1 - \beta_1) + \frac{\beta_1}{1 - \beta_1} \right) \mu_+$$

$$=: d_3\mu_+.$$

Combining the bounds in (4.26), (4.28d), and (4.30) with (4.24), we obtain

$$\begin{aligned} \|s_+ + \mu_+ F'(x_+)\|_{x_+,F}^* &\leq \|D_1\|_{x_+,F}^* + \|D_2\|_{x_+,F}^* + \|D_3\|_{x_+,F}^* \\ &\leq \mu_+(d_1(\nu) + d_2(\nu) + d_3(\nu)), \end{aligned}$$

where

$$(4.31a) \quad d_1(\nu) := \frac{\tau\epsilon_1\nu^{1/2} + \epsilon_2\beta_1}{\underline{\delta}(1 - \beta_1)},$$

$$(4.31b) \quad d_2(\nu) := \left(\frac{\tau}{\underline{\delta}} - 1 \right) \nu^{1/2} + \frac{(1 - \tau)\beta_1}{\underline{\delta}(1 - \beta_1)},$$

$$(4.31c) \quad d_3(\nu) := \frac{2\tau}{\underline{\delta}} \left(\log(1 - \beta_1) + \frac{\beta_1}{1 - \beta_1} \right).$$

⁷Our bound on $\frac{1}{\tau\mu}\|D_3\|_{x_+,F}^*$ of $2 \log(1 - \beta_1) + \frac{2\beta_1}{1 - \beta_1} = \beta_1^2 + \frac{4}{3}\beta_1^3 + \dots$ in (4.29) below can be improved in the case that K is a self-scaled cone: In [22, Theorem 4.3] a bound of no greater than β_1^2 is derived using special properties of self-scaled barriers and self-scaled cones.

From (4.21), we see that $\theta_+ \equiv d_1(\nu) + d_2(\nu) + d_3(\nu)$. Since $\theta_+ \leq \theta$ (4.20), the proof is complete. \square

The significance of θ_+ is that it specifies a neighborhood $\mathcal{N}(\theta_+)$ of the central path such that $(x_+, w_+, s_+) \in \mathcal{N}(\theta_+) \subset \mathcal{N}(\theta)$. An important consequence of this is that Iteration Short_step may be repeated. In what follows we will refer to “Algorithm Short_step”, which is simply an iterative algorithm in which at iteration k , Iteration Short_step is invoked with μ_k as the duality measure of the current iterate $(x, w, s) = (x^k, w^k, s^k)$. The algorithm terminates when an ε -optimal solution is obtained, i.e., a feasible solution with $\mu_k \leq \varepsilon$, for some prescribed ε .

4.2. Polynomial complexity of an inexact IPM. For convenience let us gather the conditions previously imposed on the algorithm parameters. The conditions (4.32a) and (4.32b) in the box below are copied from (4.19) and (4.20). (Due to Lemma 4.8 it is unnecessary to include the condition (4.9).) We also impose a new condition (4.32c) that will ensure the duality measure decreases at a fast enough rate. Since we are interested in a complexity result that holds regardless of the complexity parameter ν , we ask that these conditions be satisfied for all $\nu \geq 1$.

Algorithm parameters $\theta, \tau \in (0, 1)$ and $\epsilon_1, \epsilon_2 \in [0, 1)$ are selected as functions of ν with the property that there exists a constant $\alpha \in (0, 1)$ independent of ν such that for every $\nu \geq 1$:

$$(4.32a) \quad \underline{\delta}(\theta, \tau, \epsilon_1, \epsilon_2, \nu) > 0,$$

$$(4.32b) \quad \theta_+(\theta, \tau, \epsilon_1, \epsilon_2, \nu) \text{ is well defined, and } \theta_+(\theta, \tau, \epsilon_1, \epsilon_2, \nu) \leq \theta,$$

$$(4.32c) \quad \bar{\delta}(\theta, \tau, \epsilon_1, \epsilon_2, \nu) \leq 1 - \frac{\alpha}{\nu^{1/2}}.$$

We now give the main result regarding the worst-case complexity of Algorithm Short_step.

THEOREM 4.10. *Let F be a ν -self-concordant barrier for the cone K in (3.1). Let $\theta, \tau \in (0, 1)$ and $\epsilon_1, \epsilon_2 \in [0, 1)$ vary with ν in such a way that for some constant $\alpha > 0$, they satisfy conditions (4.32a)–(4.32c) for all $\nu \geq 1$. Suppose that the initial primal-dual iterate lies in $\mathcal{N}(\theta)$, and that for all primal iterates $x^k, F_1(x^k)$ and $F_2(x^k)$ satisfy the inequalities in (4.1) and (4.2). Then for any $\varepsilon > 0$, an ε -optimal solution to (3.1)–(3.2) is obtained in a polynomial number (in ν and $\log(\mu_0/\varepsilon)$) of iterations. Specifically, given $\varepsilon > 0$, there exists a number $k^* = \mathcal{O}(\nu^{1/2} \log(\mu_0/\varepsilon))$ such that $k \geq k^*$ implies $\mu_k \leq \varepsilon$.*

Proof. Since (4.32a) and (4.32b) hold, and the starting iterate lies in $\mathcal{N}(\theta)$, it follows from Lemma 4.9 that all iterates generated by Iteration Short_step lie in $\mathcal{N}(\theta)$. In fact, for each k , they remain in $\mathcal{N}(\theta)$ restricted to the set $S(\mu_k) := \{(x, w, s) : 0 \leq \mu \leq \mu_k\}$. Now from Lemma 4.6 the ratio of successive duality measures μ_{k+1}/μ_k is bounded above by the constant $\bar{\delta}$, which by (4.32c) is bounded away from 1. Hence $\mu_k \rightarrow 0^+$ and all limit points of the sequence $\{(x^k, w^k, s^k)\}$ lie in $\mathcal{N}(\theta)$ restricted to the set $S(\lim_k \mu_k) = S(0)$, which is the primal-dual optimal solution set. Now the particular bound on $\bar{\delta}$ in (4.32c) implies that $\mu_{k+1} \leq (1 - \alpha/\nu^{1/2})\mu_k$ for $k = 0, 1, \dots$. Our polynomial iteration bound is readily obtained from this inequality. Specifically, if $k \geq \lceil \frac{1}{\alpha} \nu^{1/2} \log(\mu_0/\varepsilon) \rceil$, then $\mu_k \leq \varepsilon$. \square

Remark 4.11. A tradeoff exists between the distance from a given quadruple $(\theta, \tau, \epsilon_1, \epsilon_2)$ to the boundary of the set of quadruples satisfying (4.32a)–(4.32b), and the largest possible α in (4.32c). Given θ and τ , it is natural to want to choose ϵ_1 and

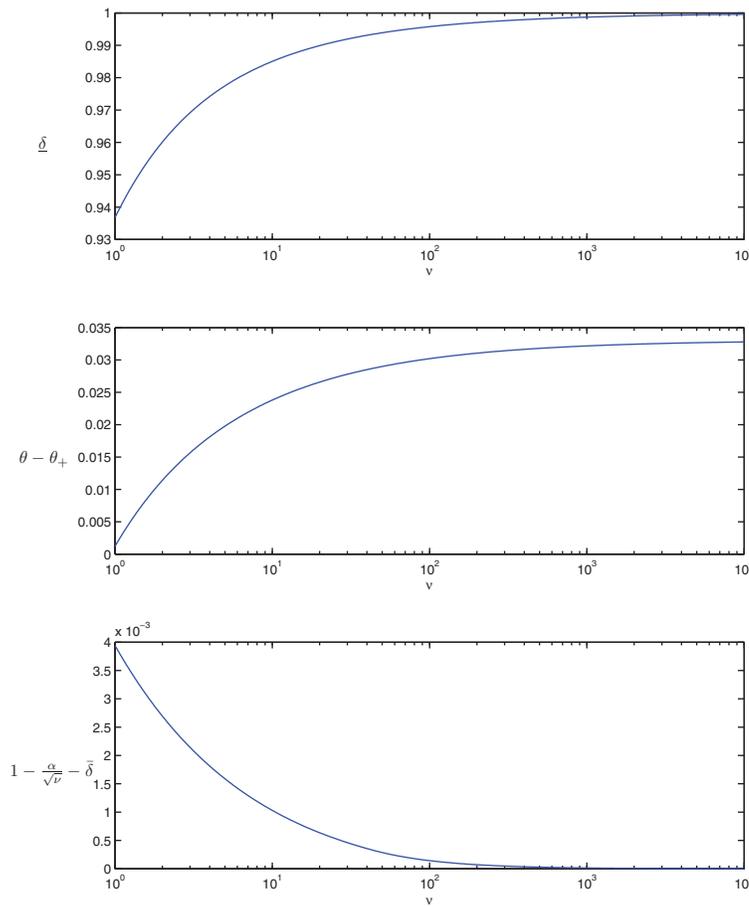


FIG. 4.1. Plot of the slacks in conditions (4.32a), (4.32b), and (4.32c), for $\alpha = 10^{-4}$, $\theta = 0.1$, $\tau = 1 - 0.02\nu^{-1/2}$, $\epsilon_1 = 0.015\nu^{-1/2}$, and $\epsilon_2 = 0.035$.

ϵ_2 as large as possible. This reduces α , which according to the proof of Theorem 4.10, increases the order constant in the complexity of the algorithm.

It remains to verify that Theorem 4.10 is not vacuous, i.e., that there exists some $\alpha \in (0, 1)$ and (possibly ν -dependent) values $\theta, \tau \in (0, 1)$ and $\epsilon_1, \epsilon_2 \in [0, 1)$, such that inequalities (4.32a)–(4.32c) hold for all $\nu \geq 1$. To that effect, let $\alpha = 10^{-4}$, $\theta = 0.1$, $\tau = 1 - 0.02\nu^{-1/2}$, $\epsilon_1 = 0.015\nu^{-1/2}$, and $\epsilon_2 = 0.035$. The slack in the three inequalities (4.32a)–(4.32c) is plotted in Figure 4.1, suggesting that the conditions hold indeed. (That they hold for arbitrarily large ν can be formally verified with little effort.)

Remark 4.12. The worst-case complexity bound obtained in Theorem 4.10, $\mathcal{O}(\nu^{1/2} \log(\mu_0/\varepsilon))$, matches the best worst-case bound currently available for convex optimization algorithms, even for linear optimization algorithms that use exact evaluations of the barrier gradient and Hessian.

COROLLARY 4.13. *Given a family $\{K_n\}$ of cones, where K_n has dimension n , and associated barrier functions $F^n : \text{int}(K_n) \rightarrow \mathcal{R}$, if the complexity parameter $\nu(n)$ of F^n is polynomial in n , and if the gradient and Hessian estimates $F_1^n(\cdot)$ and $F_2^n(\cdot)$ can each be computed in a polynomial (in n) number of arithmetic operations, then Algorithm Short_step with (possibly ν -dependent) values of θ , τ , ϵ_1 , and ϵ_2 chosen to*

satisfy inequalities (4.32a)–(4.32c), generates an ε -optimal solution in a polynomial (in n) number of arithmetic operations.

As a special case, Theorem 4.10 provides sufficient conditions on the parameters θ and τ for polynomiality of Algorithm ShortStep under *exact* evaluation of the gradient and Hessian of F . Letting $\alpha = 10^{-4}$ again, it is readily verified that, when $\theta = 0.1$ is again selected, τ can now be as small as $1 - 0.069\nu^{-1/2}$, and when the larger value $\tau = 1 - 0.02\nu^{-1/2}$ is maintained, θ can be as large as 0.28, i.e., the size of the neighborhood of the central path can be significantly increased.

5. Applications. Our results have potential applications in several areas, some of which we now discuss briefly.

1. **Inexactness in computation of the Newton step.** Inexactness in the solution of (3.10) can be reinterpreted as inexactness in the computation of $F'(x)$ or $F''(x)$. In order to do this, the increment Δx should be defined as $P_A \Delta p$, where the columns of P_A form a basis for the nullspace of A . Then the first block equation is automatically satisfied. The value of Δw can be computed by approximately solving the normal equations using an iterative method, and Δs can then be determined to satisfy the second block equation. As a consequence, the only residual is in the third block equation, and this can be analyzed as a perturbation in $F'(x)$ or $F''(x)$.

2. **Optimization over cones with intractable barrier functions.** The well-studied convex optimization problems (linear programming, SDP, SOCP, etc.) all have barrier functions that can be computed in polynomial time. For obvious reasons, there is a dearth of literature on optimization over cones with intractable barrier functions. One notable exception is the universal barrier function for a cone generated by a Chebyshev system over an interval. This function has been proven to be the limit of the sequence of barrier functions for discretizations of the interval [7], and such discretization provides a natural way to approximate such barrier functions. Our results shed light on the effects of such approximations on convergence of the inexact algorithm.

An explicit formula for the universal barrier function—which is defined for any proper cone—is given in [20], but this function and its derivatives are defined as multidimensional integrals that are intractable except for a few cones K . Different characterizations of the universal barrier function and its derivatives, and a possible procedure for estimating these quantities, are discussed in [27, Chapter 5], and our results in section 4 quantify the effects of the estimation errors.

For those conic optimization problems in which the cone's barrier function F and derivatives F' and F'' cannot be computed exactly, the use of automatic differentiation on the approximate function evaluator (defined, for example, using quadrature), provides another application for our results. Similarly, our results provide insight into the effects of using finite difference approximations to the gradient and Hessian of a barrier function.

3. **Exploiting structure in the problem.** Even for barrier functions where computation of derivatives is tractable, exact computation is sometimes not desirable. For example, in SDP there is increasing interest in problems involving large matrices, and forming the Newton system (3.10) can be too expensive. If in these problems the data matrices have special structure (for example, block-diagonal matrices plus matrices of small norm), then the small-norm pieces might be ignored in forming the normal equations. The

resulting inexactness in the Newton step can again be interpreted as resulting from a perturbation in $F'(x)$ or $F''(x)$.

6. Conclusion. A simple interior-point method for *general* conic optimization was proposed. It allows for *inexact* evaluation of the gradient and Hessian of the primal barrier function, and does not involve any evaluation of the dual barrier function or its derivatives. It was shown that as long as the relative evaluation errors remain within certain fixed, prescribed bounds, under standard assumptions, an ε -optimal solution is obtained in a polynomial number of iterations.

Acknowledgments. We are grateful to the referees for helpful comments.

REFERENCES

- [1] A. BEN-TAL AND A. NEMIROVSKII, *Lectures on Modern Convex Optimization. Analysis, Algorithms, and Engineering Applications*, MPS/SIAM Ser. Optim., SIAM, Philadelphia, 2001.
- [2] J. F. BONNANS, C. POLA, AND R. RÉBAI, *Perturbed path following predictor-corrector interior point algorithms*, Optim. Methods Softw., 11/12 (1999), pp. 183–210.
- [3] S. BOYD AND L. VANDENBERGHE, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.
- [4] M. A. BRANCH, T. F. COLEMAN, AND Y. LI, *A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems*, SIAM J. Sci. Comput., 21 (1999), pp. 1–23.
- [5] R. H. BYRD, J. C. GILBERT, AND J. NOCEDAL, *A trust region method based on interior point techniques for nonlinear programming*, Math. Program., 89 (2000), pp. 149–185.
- [6] T. J. CARPENTER AND D. F. SHANNO, *An interior point method for quadratic programs based on conjugate projected gradients*, Comput. Optim. Appl., 2 (1993), pp. 5–28.
- [7] L. FAYBUSOVICH AND M. GEKHTMAN, *Calculation of universal barrier functions for cones generated by Chebyshev systems over finite sets*, SIAM J. Optim., 14 (2004), pp. 965–979.
- [8] R. W. FREUND, F. JARRE, AND S. MIZUNO, *Convergence of a class of inexact interior-point algorithms for linear programs*, Math. Oper. Res., 24 (1999), pp. 50–71.
- [9] P. E. GILL, W. MURRAY, M. A. SAUNDERS, J. A. TOMLIN, AND M. H. WRIGHT, *On projected Newton barrier methods for linear programming and an equivalence to Karmarkar's projective method*, Math. Program., 36 (1986), pp. 183–209.
- [10] D. GOLDFARB AND S. MEHROTRA, *A relaxed version of Karmarkar's method*, Math. Program., 40 (1988), pp. 289–315.
- [11] N. K. KARMARKAR AND K. G. RAMAKRISHNAN, *Computational results of an interior point algorithm for large scale linear programming*, Math. Program., 52 (1991), pp. 555–586.
- [12] J. KORZAK, *Convergence analysis of inexact infeasible-interior-point algorithms for solving linear programming problems*, SIAM J. Optim., 11 (2000), pp. 133–148.
- [13] Z. LU, R. D. C. MONTEIRO, AND J. W. O'NEAL, *An iterative solver-based infeasible primal-dual path-following algorithm for convex quadratic programming*, SIAM J. Optim., 17 (2006), pp. 287–310.
- [14] S. MEHROTRA, *Implementation of affine scaling methods: Approximate solutions of systems of linear equations using preconditioned conjugate gradient methods*, ORSA J. Comput., 4 (1992), pp. 103–118.
- [15] S. MIZUNO AND F. JARRE, *Global and polynomial-time convergence of an infeasible-interior-point algorithm using inexact computation*, Math. Program., 84 (1999), pp. 105–122.
- [16] K. NAKATA, K. FUJISAWA, AND M. KOJIMA, *Using the conjugate gradient method in interior-point methods for semidefinite programs*, Proc. Inst. Statist. Math., 46 (1998), pp. 297–316.
- [17] Y. NESTEROV, *Long-step strategies in interior-point primal-dual methods*, Math. Program., 76 (1997), pp. 47–94.
- [18] Y. NESTEROV, *Introductory Lectures on Convex Optimization. A Basic Course*, Kluwer Academic Publishers, Boston/Dordrecht/London, 2004.
- [19] Y. NESTEROV, *Towards nonsymmetric conic optimization*, Technical Report: CORE Discussion Paper 2006/28, Center for Operations Research and Econometrics, Catholic University of Louvain, March 2006.
- [20] Y. NESTEROV AND A. NEMIROVSKII, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM Stud. Appl. Math. 13, SIAM, Philadelphia, 1994.

- [21] Y. NESTEROV, M. J. TODD, AND Y. YE, *Infeasible-start primal-dual methods and infeasibility detectors for nonlinear programming problems*, Math. Program., 84 (1999), pp. 227–267.
- [22] Y. E. NESTEROV AND M. J. TODD, *Self-scaled barriers and interior-point methods for convex programming*, Math. Oper. Res., 22 (1997), pp. 1–42.
- [23] Y. E. NESTEROV AND M. J. TODD, *Primal-dual interior-point methods for self-scaled cones*, SIAM J. Optim., 8 (1998), pp. 324–364.
- [24] J. RENEGAR, *A polynomial-time algorithm, based on Newton's method, for linear programming*, Math. Program., 40 (1988), pp. 59–93.
- [25] J. RENEGAR, *A Mathematical View of Interior-Point Methods in Convex Optimization*, MPS/SIAM Ser. Optim., SIAM, Philadelphia, 2001.
- [26] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original.
- [27] S. P. SCHURR, *An inexact interior-point algorithm for conic convex optimization problems*, Ph.D. thesis, University of Maryland, College Park, MD 20472, USA, 2006.
- [28] K.-C. TOH, R. H. TÜTÜNCÜ, AND M. J. TODD, *Inexact primal-dual path-following algorithms for a special class of convex quadratic SDP and related problems*, Pac. J. Optim., 3 (2007), pp. 135–164.
- [29] L. TUNCEL, *Generalization of primal-dual interior-point methods to convex optimization problems in conic form*, Found. Comput. Math., 1 (2001), pp. 229–254.
- [30] W. WANG AND D. P. O'LEARY, *Adaptive use of iterative methods in predictor-corrector interior point methods for linear programming*, Numer. Algorithms, 25 (2000), pp. 387–406.
- [31] S. J. WRIGHT, *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, 1997.
- [32] G. ZHOU AND K.-C. TOH, *Polynomiality of an inexact infeasible interior point algorithm for semidefinite programming*, Math. Program., 99 (2004), pp. 261–282.