

A NEW AND EFFICIENT LARGE-UPDATE INTERIOR-POINT METHOD FOR LINEAR OPTIMIZATION

J. PENG, C. ROOS

Delft University of Technology, Netherlands

e-mail: J.Peng@its.tudelft.nl, c.roos@its.tudelft.nl

T. TERLAKY

McMaster University, Hamilton, Ontario, Canada

e-mail: terlaky@cas.mcmaster.ca

В данной статье представлен уточненный анализ нового метода внутренних точек для линейной оптимизации, использующий некоторые новые математические инструменты, в частности, разработанные в [11], где рассмотрено семейство аналогичных методов.

1. Introduction

Interior-point methods (IPMs) are among the most effective methods for solving wide classes of Linear Optimization (LO) problems. Since the seminal work of Karmarkar [5], many researchers have proposed and analyzed various IPMs for LO and a large amount of results have been reported. For a survey we refer to recent books on the subject [14, 16, 17, 19]. An interesting fact is that almost all known polynomial-time variants of IPMs use the so-called *central path* [15] as a guideline to the optimal set, and some variant of Newton's method is used to follow the central path approximately. Therefore, analyzing the behavior of Newton's method has a prominent role in the theoretical investigation of IPMs.

1.1. Notation

We use the following notational conventions. Throughout the paper, $\|\cdot\|$ denotes the 2-norm of a vector, whereas $\|\cdot\|_\infty$ denotes the infinity norm. For any $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$, x_{\min} denotes the smallest and x_{\max} the largest value of the components of x . If also $s \in \mathbb{R}^n$, then xs denotes the coordinatewise product of the vectors x and s . Furthermore, e denotes the all-one vector of length n . Finally, if $z \in \mathbb{R}_+^n$ and $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, then $f(z)$ denotes the vector in \mathbb{R}_+^n whose i -th component is $f(z_i)$, with $1 \leq i \leq n$.

1.2. The gap between theory and practice

At present there is still a gap between the practical behavior of the algorithms and the theoretical performance results, in favor of the practical behavior. This is especially true for so-called primal-dual large-update methods, which are the most efficient methods in practice (see, e.g. Andersen et al. [1]). The aim of this paper is to present a new analysis of the primal-dual Newton algorithm proposed in [10]. The new iteration bound further improves the bound in [10], and almost closes the above gap, at least from a theoretical point of view. The method of [10] belongs to a wider family of methods, introduced and thoroughly studied in [11]. By applying some of the new analysis tools developed in [11] to the method of [10], a relatively simple analysis leads to the presently best possible iteration bound for large-update methods.

To be more concrete we need to go into more detail at this stage. We deal with the following LO-problem:

$$(P) \quad \min\{c^T x : Ax = b, x \geq 0\},$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, and its dual problem

$$(D) \quad \max\{b^T y : A^T y + s = c, s \geq 0\}.$$

We assume that both (P) and (D) satisfy the interior-point condition (IPC), i. e., there exists (x^0, s^0, y^0) such that

$$Ax^0 = b, x^0 > 0, \quad A^T y^0 + s^0 = c, s^0 > 0.$$

It is well known that the IPC can be assumed without loss of generality. In fact we may, and will assume that $x^0 = s^0 = e$. For this and some other properties mentioned below, see, e.g., [14]. Finding an optimal solution of (P) and (D) is equivalent to solving the following system

$$\begin{aligned} Ax &= b, & x &\geq 0, \\ A^T y + s &= c, & s &\geq 0, \\ xs &= 0. \end{aligned} \tag{1}$$

The basic idea of primal-dual IPMs is to replace the third equation in (1), the so-called *complementarity condition* for (P) and (D), by the parametrized equation $xs = \mu e$, where e denotes the all-one vector and $\mu > 0$. Thus we consider the system

$$\begin{aligned} Ax &= b, & x &\geq 0, \\ A^T y + s &= c, & s &\geq 0, \\ xs &= \mu e. \end{aligned} \tag{2}$$

If $\text{rank}(A) = m$ and the IPC holds, then for each $\mu > 0$, the parameterized system (2) has a unique solution. This solution is denoted as $(x(\mu), y(\mu), s(\mu))$ and we call $x(\mu)$ the μ -center of (P) and $(y(\mu), s(\mu))$ the μ -center of (D). The set of μ -centers (with μ running through all positive real numbers) gives a homotopy path, which is called *the central path* of (P) and (D). The relevance of the central path for LO was recognized first by Sonnevend [15] and Megiddo [6]. If $\mu \rightarrow 0$ then the limit of the central path exists and since the limit points satisfy the complementarity condition, the limit yields optimal solutions for (P) and (D).

IPMs follow the central path approximately. Let us briefly indicate how this goes. Without loss of generality we assume that $(x(\mu), y(\mu), s(\mu))$ is known for some positive μ . For example,

due to the above assumption we may assume this for $\mu = 1$, with $x(1) = s(1) = e$. We then decrease μ to $\mu_+ = (1 - \theta)\mu$, for some $\theta \in (0, 1)$ and we solve the following Newton system

$$\begin{aligned} A\Delta x &= 0, \\ A^T\Delta y + \Delta s &= 0, \\ s\Delta x + x\Delta s &= \mu_+e - xs. \end{aligned} \tag{3}$$

This system defines a search direction $(\Delta x, \Delta s, \Delta y)$. By taking a step along the search direction, with the step size defined by some line search rules, one constructs a new triple (x, y, s) . If necessary, we repeat the procedure until we find iterates that are ‘close’ to $(x(\mu_+), y(\mu_+), s(\mu_+))$. Then $\mu := \mu_+$ is again reduced by the factor $1 - \theta$ and we apply Newton’s method targeting at the new μ -centers, and so on. This process is repeated until μ is small enough, say until $n\mu \leq \varepsilon$; at this stage we have found an ε -solution of the problems (P) and (D) .

The ‘closeness’ of some iterates (x, y, s) to the μ -center will be quantified by using a proximity measure $\delta(xs, \mu)$ that will be defined later on. More formally, the algorithm can now be described as follows.

Primal-Dual Newton Algorithm for LO

Input:

A proximity parameter τ ;
 an accuracy parameter $\varepsilon > 0$;
 a fixed barrier update parameter θ , $0 < \theta < 1$;

begin

$x := e$; $s := e$; $\mu := 1$;

while $n\mu \geq \varepsilon$ **do**

begin

$\mu := (1 - \theta)\mu$;

while $\delta(xs; \mu) \geq \tau$ **do**

begin

$x := x + \alpha\Delta x$;

$s := s + \alpha\Delta s$;

$y := y + \alpha\Delta y$

end

end

end

As said before, we take $x = s = e$ and $\mu = 1$ to initialize the algorithm. The parameter α ($0 < \alpha \leq 1$) is the so called *step size* (or *damping factor*) and will be specified later on. If the proximity measure $\delta(xs; \mu)$ exceeds some threshold value τ then we use one or more damped Newton steps to recenter; otherwise μ is reduced by the factor $1 - \theta$. This is repeated until $n\mu < \varepsilon$. The step size (or damping factor) α has to be taken such that the proximity measure function δ decreases by a sufficient amount. In the theoretical analysis the step size α is usually given some default value.

Let us mention that most practical algorithms use the ε -solution to construct a basic solution and then produce an optimal basic solution by *crossing-over* to the Simplex method. An alternative way is to apply a rounding procedure as described by Ye [18] (see also Mehrotra and Ye [7], and Roos et al. [14]).

The choice of the parameter θ plays an important role both in theory and practice of IPMs. Usually, if θ is a constant independent of the dimension order n of the problem, for instance $\theta = \frac{1}{2}$, then we call the algorithm a *large-update* (or *long-step*) method. If θ depends on the problem, such as $\theta = \frac{1}{\sqrt{n}}$, then the algorithm is named a *small-update* (or *short-step*) method.

It is now well known that small-update methods have the best iteration bound; they require $O\left(\sqrt{n} \log \frac{n}{\varepsilon}\right)$ iterations to produce an ε -solution. On the other hand, large-update methods, in practice much more efficient than small-update methods [1], have a worse iteration bound, namely $O\left(n \log \frac{n}{\varepsilon}\right)$ [14, 17, 19]. This phenomenon is what we called “the gap between theory and practice”.

1.3. Aim of this paper

Several authors have suggested to use the so called higher-order methods to improve the complexity of large-update IPMs [3, 4, 8, 20, 21]. In these methods, in each iteration one solves some additional equations based on the higher-order approximations to the system (2).

In this paper, as in [10] and [11], we propose a different way to narrow the gap between theory and practice for large-update IPMs. In the next section we briefly discuss how the above mentioned iteration bounds are obtained. Then we indicate some inconsistency that occurs in the analysis and which suggests how to improve the iteration bound by slightly changing the search direction. The new search direction is defined in Section 2.4 and the rest of the paper is devoted to the analysis of this new direction. The result is an $O\left(\sqrt{n} \log n \log \frac{n}{\varepsilon}\right)$ iteration bound. Clearly, this bound almost closes the gap between the iteration bounds for small-update and large-update methods.

The result of this paper improves the result in [10], where the same authors used the same search direction as in this paper but could obtain only a $O\left(n^{\frac{2}{3}} \log \frac{n}{\varepsilon}\right)$ iteration bound. The improvement in this paper is due to the use of some new analytical tools, that were developed in [11]. In fact, in [11] a whole class of IPMs is considered, based on so called *self-regular* proximity functions, and for some IPMs in this class complexity results are derived that are the same as the result of this paper. The aim of this paper is to concentrate on one member of this class, thus showing that the currently best possible complexity result for large-update methods can be obtained with a relatively simple analysis. In addition, we devote an extra section to explain which ideas have led to the new approach. This will be done in Section 2.

2. Preliminaries

2.1. The classical approach

As noted before, for the analysis of the Newton process we need to measure the “distance” from a primal-dual pair (x, s) to the μ -center $(x(\mu), s(\mu))$. The most popular tool for this is

the so-called primal-dual logarithmic barrier function. This function is, up to the “constant” $-n + n \log \mu$, given by

$$\Phi_c(xs, \mu) = \frac{x^T s}{\mu} - \sum_{i=1}^n \log x_i s_i - n + n \log \mu. \quad (4)$$

Its usefulness can easily be understood by introducing the vector

$$v := \sqrt{\frac{xs}{\mu}}. \quad (5)$$

Note that the pair (x, s) coincides with the μ -center $(x(\mu), s(\mu))$ if and only if $v = e$. Now defining

$$\psi_c(t) := t^2 - 1 - 2 \log t, \quad t > 0,$$

one may easily verify that $\Phi_c(xs, \mu)$ can be expressed in terms of the vector v as follows

$$\Phi_c(xs, \mu) = \Psi_c(v) := \sum_{i=1}^n \psi_c(v_i). \quad (6)$$

Since $\psi_c(t)$ is strictly convex, and attains its minimal value at $t = 1$, with $\psi_c(1) = 0$, it follows that $\Phi_c(xs, \mu)$ is nonnegative and vanishes if and only if $v = e$, i. e., if and only if $xs = \mu e$.

Below we need a second proximity measure, a so-called norm-based measure, namely

$$\delta(xs, \mu) := \frac{1}{2} \|v - v^{-1}\| = \frac{1}{2} \left\| \sqrt{\frac{xs}{\mu}} - \sqrt{\frac{\mu e}{xs}} \right\|. \quad (7)$$

Note that

$$\delta(xs, \mu) = 0 \Leftrightarrow \Phi_c(xs, \mu) = 0 \Leftrightarrow v = e \Leftrightarrow xs = \mu e,$$

and otherwise $\delta(xs, \mu) > 0$ and $\Phi_c(xs, \mu) > 0$.

The classical analysis of large-update methods depends on the following two results. The first result estimates the increase in $\Phi_c(xs, \mu)$ if, at the start of an outer iteration, μ is multiplied by the factor $1 - \theta$.

Proposition 2.1 (Lemma II.72 in [14]). *Let (x, y, s) be strictly feasible and $\mu > 0$. If $\mu_+ = (1 - \theta)\mu$ then*

$$\Phi_c(xs, \mu^+) \leq \Phi_c(xs, \mu) + \frac{2\delta\rho(\delta)\theta\sqrt{n}}{1 - \theta} + n\psi_c\left(\sqrt{\frac{\theta}{1 - \theta}}\right),$$

where $\delta = \delta(xs, \mu)$ and $\rho(\delta) = \delta + \sqrt{1 + \delta^2} = O(\delta)$.

The second result estimates the decrease in $\Phi_c(xs, \mu)$ during a damped Newton step, with an appropriate damping factor α .

Proposition 2.2 (Section 7.8.2 in [14]). *If $\delta = \delta(xs, \mu) \geq 1$ there exists a step size α such that*

$$\Phi_c(x_+s_+, \mu) \leq \Phi_c(xs, \mu) - \psi_c\left(\sqrt{\frac{2\delta}{\rho(\delta)}}\right) = \Phi_c(xs, \mu) - \Theta(1).$$

With these two results we can derive an iteration bound for a large-update method. Taking $\tau = 1$ in the algorithm, at the start of each outer iteration one has $\delta = \delta(xs, \mu) \leq 1$. From this it follows that then $\Phi_c(xs, \mu) = O(1)$, and also $\rho(\delta) = O(1)$. Taking $\theta = \frac{1}{2}$, or some other fixed constant in the interval $(0, 1)$, it follows from Proposition 2.1 that after the update of μ we have

$$\Phi_c(xs, \mu) \leq O(1) + O(\sqrt{n}) + O(n) = O(n).$$

As long as $\delta \geq 1$, each inner iteration decreases the barrier function with at least a constant, by Proposition 2.2. As a consequence, no more than $O(n)$ inner iterations are necessary to reach the situation where $\delta(xs, \mu) \leq 1$. This proves that each outer iteration requires no more than $O(n)$ inner iterations. The number of outer iterations is bounded by (cf. Lemma II.17 in [14])

$$\frac{1}{\theta} \log \frac{n}{\varepsilon}. \quad (8)$$

Multiplication of this number by the number of inner iterations per outer iteration yields a bound for the total number of iterations. If $\theta = \Theta(1)$, the iteration bound becomes

$$O\left(n \log \frac{n}{\varepsilon}\right). \quad (9)$$

This is the currently best bound for large-update methods, derived in the past in many ways and by many authors.

2.2. An alternative approach

To make clear what is the underlying idea of this paper, we proceed with a second derivation of the iteration bound (9). In this derivation we do not use the logarithmic barrier function, but we only use the norm-based proximity measure $\delta(xs, \mu)$. The analysis was carried out first in [9], and is based on the following two results.

Proposition 2.3 (Lemma IV.36 in [14]). *Using the above notation, we have*

$$\delta(xs, \mu_+) \leq \frac{2\delta(xs, \mu) + \theta\sqrt{n}}{2\sqrt{1-\theta}}.$$

Proposition 2.4 (Theorem 3.6 in [9]). *If $\delta = \delta(xs, \mu) \geq 1$ and $\alpha = \frac{1}{2\delta^2}$ then*

$$\delta(x_{+s_+}, \mu)^2 \leq \delta^2 - \frac{5}{12} = \delta^2 - \Theta(1).$$

Just as in the previous section, with these two results the iteration bound easily follows. If $\tau = 1$, at the start of an outer iteration we have $\delta = \delta(xs, \mu) \leq 1$. With $\theta = \frac{1}{2}$, or some other fixed constant in the interval $(0, 1)$, Proposition 2.3 implies that after the update of μ we have

$$\delta(xs, \mu)^2 \leq (O(1) + O(\sqrt{n}))^2 = O(n).$$

By Proposition 2.4, as long as $\delta \geq 1$, each inner iteration decreases the squared proximity with at least a constant. As a consequence, no more than $O(n)$ inner iterations are necessary to reach the situation where $\delta(xs, \mu) \leq 1$. Hence, each outer iteration requires no more than $O(n)$ inner iterations. Using (8) again, we obtain (9) as a bound for the total number of iterations.

2.3. Two observations

2.3.1. First observation

In [9] we found another interesting result that suggests a direction where a better complexity bound for the algorithm might be obtained.

Proposition 2.5 (Corollary 4.4 in [9]). *Let $\delta = \delta(xs, \mu) \geq 1$ and $v_{\min} \geq 1$. Then the step size $\alpha = \frac{1}{1+\delta}$ gives*

$$\delta(x_{+s_+}, \mu)^2 \leq \delta^2 - \frac{3\delta^2}{4(1+\delta)} \leq \delta^2 - \frac{3}{8}\delta.$$

If we could prove the inequality in Proposition 2.5 without the assumption $v_{\min} \geq 1$, this would give the desired iteration bound

$$O\left(\sqrt{n} \log \frac{n}{\varepsilon}\right) \quad (10)$$

for large-update methods, as we will show below. We need two simple technical results for this purpose; for completeness' sake we include their (short) proofs.

Lemma 2.6 (Lemma 2.1 in [10]). *If $\alpha \in [0, 1]$, then*

$$(1-t)^\alpha \leq 1 - \alpha t, \quad \forall t \in [0, 1]. \quad (11)$$

Proof: For any fixed $\alpha \leq 1$ the function $(1-t)^\alpha - 1 + \alpha t$ is decreasing for $t \in [0, 1]$ and zero if $t = 0$. This implies (11).

Lemma 2.7 (Proposition 2.2 in [10]). *Let t_0, t_1, \dots, t_K be a sequence of positive numbers such that*

$$t_{k+1} \leq t_k - \kappa t_k^{1-\gamma}, \quad k = 0, 1, \dots, K-1, \quad (12)$$

where $\kappa > 0$ and $0 < \gamma \leq 1$. Then $K \leq \left\lfloor \frac{t_0^\gamma}{\kappa\gamma} \right\rfloor$.

Proof: Using (12), we have

$$0 < t_{k+1}^\gamma \leq (t_k - \kappa t_k^{1-\gamma})^\gamma = t_k^\gamma (1 - \kappa t_k^{-\gamma})^\gamma \leq t_k^\gamma (1 - \kappa\gamma t_k^{-\gamma}) = t_k^\gamma - \kappa\gamma,$$

where the second inequality follows from (11). Hence, for each k , $t_k^\gamma \leq t_0^\gamma - k\gamma\kappa$. Taking $k = K$ we obtain $0 < t_0^\gamma - K\gamma\kappa$, which implies the lemma. \square

By taking in $\gamma = \frac{1}{2}$, $\kappa = \frac{3}{8}$, and $t_k = \delta(x^k s^k, \mu^k)^2$, where x^k , s^k , and μ^k denote the values of x , s , and μ at the k -th iteration, Lemma 2.7 yields the iteration bound (10).

Comparison of the results in Proposition 2.4 and Proposition 2.5 indicates that a larger decrease of the proximity occurs when the smallest coordinate of the vector v is large. Hence, **the question becomes how to keep v_{\min} large during the course of the algorithm.** In the next section we make another observation that points in the same direction.

2.3.2. Second observation

In the sequel we will frequently use scaled versions of the displacements Δx and Δs . These are defined as follows:

$$d_x := \frac{v\Delta x}{x}, \quad d_s := \frac{v\Delta s}{s}. \quad (13)$$

Now the system (3) defining the Newton search directions can be rewritten as

$$\begin{aligned} \bar{A}d_x &= 0, \\ \bar{A}^T \Delta y + d_s &= 0, \\ d_x + d_s &= v^{-1} - v, \end{aligned}$$

where $\bar{A} = AV^{-1}X$, with $V = \text{diag}(v)$, $X = \text{diag}(x)$. The last equation in the above system is called the scaled Newton equation. Yet we observe that this equation can be written as

$$d_x + d_s = v^{-1} - v = -\frac{1}{2}\nabla\Psi_c(v),$$

with $\Psi_c(v)$ denoting the ‘scaled’ logarithmic barrier function, as defined in (6). This shows that the scaled search directions depend on the steepest descent direction of the scaled logarithmic barrier function. This seems to be natural when dealing with the analysis of Section 2.1. However, when dealing with the analysis of Section 2.2, we are trying to minimize the squared proximity, i. e., $\|v^{-1} - v\|^2$, and then it seems to be more attractive to relate the search direction to its gradient, as follows:

$$d_x + d_s = -\frac{1}{2}\nabla_v \|v^{-1} - v\|^2 = v^{-3} - v. \quad (14)$$

Changing the definition of the search directions in this way, the authors could obtain a better iteration bound for large-update methods in [10], namely

$$O\left(n^{\frac{2}{3}} \log \frac{n}{\varepsilon}\right).$$

The analysis in [10] used as a proximity measure, albeit implicitly, $\|v^{-3} - v\|$. Note that small coordinates (smaller than 1) in v may blow up this measure. Hence, this measure is in favour of large values of v_{\min} . This strengthens our conclusion of the previous section, namely that **the theoretical complexity may improve if we keep v_{\min} large during the course of the algorithm.**

We conclude this section by considering the direction (14) in the original space. One has

$$s\Delta x + x\Delta s = \mu v (d_x + d_s) = \mu (v^{-2} - v^2),$$

or, equivalently,

$$s\Delta x + x\Delta s = \mu \frac{\mu e}{xs} - xs.$$

Thus the new direction is the Newton direction targeting at $\mu \frac{\mu e}{xs}$ instead of μe .

2.4. A class of new directions

Inspired by the result discussed in the previous section, we modify the Newton equation in this paper as follows:

$$d_x + d_s = -\nabla\Psi(v) = v^{-q} - v, \quad (15)$$

with $q > 1$, and where the functions ψ and Ψ are defined according to

$$\Psi(v) := \sum_{i=1}^n \psi(v_i), \quad \psi(v_i) := \left(\frac{v_i^2 - 1}{2} + \frac{v_i^{1-q} - 1}{q-1} \right). \quad (16)$$

Note that since $q > 1$, the function $\Psi(v)$ has the *barrier property*, i.e., the property that it goes to infinity if the vector v approaches the boundary of the feasible region, i.e., if one of the coordinates of v goes to zero. We call it a *polynomial barrier function*. Only in the limit, when q approaches 1, the barrier function becomes logarithmic:

$$\lim_{q \downarrow 1} \Psi(v) = \frac{1}{2} \sum_{i=1}^n (v_i^2 - 1 - \log v_i^2) = \frac{1}{2} \left(\frac{x^T s}{\mu} - \sum_{i=1}^n x_i s_i + n \log \mu - n \right).$$

This is, up to the factor $\frac{1}{2}$, precisely the logarithmic barrier function (4).

In the analysis of the algorithm with the modified Newton direction, we also use a norm-based proximity measure $\sigma(v)$, naturally related to our new barrier function, according to

$$\sigma(v) := \|\nabla\Psi(v)\| = \|d_x + d_s\| = \|v^{-q} - v\|. \quad (17)$$

Letting $\Delta x, \Delta y, \Delta s$ denote the displacements in the original space, the result of a damped Newton step with damping factor α is denoted as

$$x_+ = x + \alpha\Delta x, \quad y_+ = y + \alpha\Delta y, \quad s_+ = s + \alpha\Delta s. \quad (18)$$

The rest of the paper is devoted to the complexity analysis of the large-update algorithm with the modified Newton direction (15). Let us mention that basically the same algorithm, with the same search direction, was proposed and analyzed in [10]. In that paper a parameter ρ occurs in the definition of the search direction (cf. [10, page 4]); by taking $\rho = q - 1$, the definition coincides with (15). By using some new tools in the analysis we obtain in this paper a better complexity result. This is due to the fact that in [10] we had to restrict ourselves to the case $\rho \leq 2$, which corresponds to $q \leq 3$, whereas we can now deal also with larger values of q in the analysis.

The rest of the paper is organized as follows. In the next section we derive bounds for the coordinates of the vector v in terms of $\sigma(v)$ and we derive a lower bound for the maximum value of a feasible step size. In Section 4 we estimate the decrease of the barrier function during a damped step, and in Section 5 the increase due to the update of the barrier parameter μ . When having these results, the announced iteration bound easily follows, in Section 6.

3. Bounds for v and the step size

Our first lemma in this section provides a lower and an upper bound for the coordinates of v in terms of $\sigma(v)$.

Lemma 3.1. *Let $\sigma := \sigma(v)$, as defined by (17). Then*

$$v_{\min} \geq (1 + \sigma)^{-\frac{1}{q}}, \quad v_{\max} \leq 1 + \sigma.$$

Proof: The lemma is trivial if $v_{\min} \geq 1$ and $v_{\max} \leq 1$. Consider the case that $v_{\min} < 1$. From (17) we derive

$$\sigma = \|v - v^{-q}\| \geq |v_i^{-q} - v_i|, \quad 1 \leq i \leq n,$$

which implies

$$\sigma \geq \frac{1}{v_{\min}^q} - v_{\min} \geq \frac{1}{v_{\min}^q} - 1.$$

On the other hand, if $v_{\max} > 1$ then we find

$$\sigma \geq v_{\max} - \frac{1}{v_{\max}^q} \geq v_{\max} - 1.$$

The lemma follows directly from the above inequalities. \square

Note that if q goes to infinity, then the lower bound for v_{\min} approaches 1, for all values of σ .

In the sequel we use the following notations:

$$\Delta_x = \frac{\Delta x}{x} = \frac{d_x}{v}, \quad \Delta_s = \frac{\Delta s}{s} = \frac{d_s}{v}. \quad (19)$$

We then may write

$$x_+ = x(e + \alpha \Delta_x), \quad s_+ = s(e + \alpha \Delta_s). \quad (20)$$

Hence the maximum step size is determined by the vector (Δ_x, Δ_s) : the step size α is feasible if and only if $e + \alpha \Delta_x \geq 0$ and $e + \alpha \Delta_s \geq 0$, and this will certainly hold if

$$1 - \alpha \|(\Delta_x, \Delta_s)\| \geq 0. \quad (21)$$

Since the displacements Δx and Δs are orthogonal, the scaled displacements d_x and d_s are orthogonal as well, i. e., $d_x^T d_s = 0$. This enables us, using (17), to write

$$\sigma(v) = \|d_x + d_s\| = \|(d_x, d_s)\|. \quad (22)$$

Now we are ready to state our next result.

Lemma 3.2. *One has*

$$\|(\Delta_x, \Delta_s)\| \leq \sigma(1 + \sigma)^{\frac{1}{q}}.$$

Consequently, the maximal feasible step size, α_{\max} , satisfies

$$\alpha_{\max} \geq \frac{1}{\sigma(1 + \sigma)^{\frac{1}{q}}}.$$

Proof: Using Lemma 3.1 and (22), we may write

$$\|(\Delta_x, \Delta_s)\| = \left\| \left(\frac{d_x}{v}, \frac{d_s}{v} \right) \right\| \leq \frac{\|(d_x, d_s)\|}{v_{\min}} = \frac{\sigma}{v_{\min}} \leq \sigma(1 + \sigma)^{\frac{1}{q}}.$$

From (21) we derive that

$$\alpha_{\max} \geq \frac{1}{\|(\Delta_x, \Delta_s)\|}.$$

Hence the lemma follows. \square

In the next section it will become clear that the default step size in our analysis is only a small fraction of the maximal step size in Lemma 3.2.

4. Estimate of the decrease of the barrier function

Our aim in this section is to estimate the decrease of the new barrier function

$$\Phi(xs, \mu) := \Psi(v) = \sum_{i=1}^n \psi(v_i) \quad (23)$$

during one modified Newton step, for a suitable (feasible) step size α . It will be convenient to introduce

$$\Psi_b(v) := \sum_{i=1}^n \psi_b(v_i), \quad \psi_b(v_i) := \frac{v_i^{1-q} - 1}{q - 1}. \quad (24)$$

Then we may write

$$\Phi(xs, \mu) = \sum_{i=1}^n \left(\frac{v_i^2 - 1}{2} + \frac{v_i^{1-q} - 1}{q - 1} \right) = \frac{e^T v^2 - n}{2} + \Psi_b(v).$$

Hence, the value of the barrier function after one step is given by

$$\Phi(x_{+s_+}, \mu) = \frac{e^T v_+^2 - n}{2} + \Psi_b(v_+), \quad v_+ = \sqrt{\frac{x_{+s_+}}{\mu}}. \quad (25)$$

We proceed by deriving a simple expression for v_+ . Using (20) and (19) we find

$$v_+^2 = \frac{x_{+s_+}}{\mu} = \frac{xs(e + \alpha\Delta_x)(e + \alpha\Delta_s)}{\mu} = v^2(e + \alpha\Delta_x)(e + \alpha\Delta_s) = (v + \alpha d_x)(v + \alpha d_s).$$

Hence, using the orthogonality of d_x and d_s we obtain

$$e^T v_+^2 = e^T (v^2 + \alpha v(d_x + d_s) + \alpha^2 d_x d_s) = e^T v^2 + \alpha v^T (v^{-q} - v).$$

Furthermore,

$$\Psi_b(v_+) = \Psi_b \left(\sqrt{(v + \alpha d_x)(v + \alpha d_s)} \right).$$

At this stage we need a simple, but important lemma. As we will see later, it provides a new and convenient tool in the analysis of our large-update method.

Lemma 4.1 (cf. Lemma 2.2 in [11]). *Let $t_1 > 0$ and $t_2 > 0$. Then*

$$\psi_b(\sqrt{t_1 t_2}) \leq \frac{1}{2} (\psi_b(t_1) + \psi_b(t_2)). \quad (26)$$

Proof: One may easily verify that the property in the lemma holds if and only if the function $\psi_b(e^z)$ is convex in z , and this holds if and only if $\psi_b'(t) + t\psi_b''(t) \geq 0$, whenever $t > 0$. Since

$$\psi_b'(t) = -t^{-q}, \quad \psi_b''(t) = qt^{-q-1},$$

one has $\psi_b'(t) + t\psi_b''(t) = (q - 1)t^{-q} > 0$. Hence, the lemma follows. \square

The above proof makes clear that the property (26) is equivalent to convexity (with respect to ξ) of the composed function $\psi_b(e^\xi)$. We therefore say that ψ_b is *right exponentially convex*, or shortly, right e-convex.¹

Due to the definition of the function Ψ_b , the property in Lemma 4.1 extends to Ψ_b , and hence we have²

$$\Psi_b(v_+) = \Psi_b\left(\sqrt{(v + \alpha d_x)(v + \alpha d_s)}\right) \leq \frac{1}{2} \sum_{i=1}^n (\psi_b(v_i + \alpha d_{x_i}) + \psi_b(v_i + \alpha d_{s_i})).$$

Substitution of the above results in (25) gives

$$\Phi(x_+s_+, \mu) \leq \frac{e^T v^2 + \alpha v^T (v^{-q} - v) - n}{2} + \frac{1}{2} \sum_{i=1}^n (\psi_b(v_i + \alpha d_{x_i}) + \psi_b(v_i + \alpha d_{s_i})). \quad (27)$$

Let us denote the change in Φ during the step as $f(\alpha)$. Then

$$f(\alpha) = \Phi(x_+s_+, \mu) - \Phi(xs, \mu) \leq f_1(\alpha),$$

where

$$f_1(\alpha) := -\Phi(xs, \mu) + \frac{e^T v^2 + \alpha v^T (v^{-q} - v) - n}{2} + \frac{1}{2} \sum_{i=1}^n (\psi_b(v_i + \alpha d_{x_i}) + \psi_b(v_i + \alpha d_{s_i})).$$

Note that $f(0) = f_1(0) = 0$. We next consider the first and second derivative of $f_1(\alpha)$ to α .

$$f'_1(\alpha) = \frac{1}{2} v^T (v^{-q} - v) + \frac{1}{2} \sum_{i=1}^n (\psi'_b(v_i + \alpha d_{x_i}) d_{x_i} + \psi'_b(v_i + \alpha d_{s_i}) d_{s_i}).$$

One easily verifies, using (15) and (17), that

$$f'_1(0) = -\frac{1}{2} \sigma^2. \quad (28)$$

Furthermore, using that $\psi_b''(t) = qt^{-q-1}$ is monotonically decreasing,

$$\begin{aligned} f''_1(\alpha) &= \frac{1}{2} \sum_{i=1}^n (\psi_b''(v_i + \alpha d_{x_i}) d_{x_i}^2 + \psi_b''(v_i + \alpha d_{s_i}) d_{s_i}^2) = \\ &= \frac{1}{2} \sum_{i=1}^n ((q(v_i + \alpha d_{x_i})^{-q-1}) d_{x_i}^2 + (q(v_i + \alpha d_{s_i})^{-q-1}) d_{s_i}^2) \leq \frac{q}{2} \sum_{i=1}^n (v_{\min} - \alpha \sigma)^{-q-1} (d_{x_i}^2 + d_{s_i}^2). \end{aligned}$$

The last inequality is due to

$$\begin{aligned} v_i + \alpha d_{x_i} &\geq v_{\min} - \alpha \|d_x\| \geq v_{\min} - \alpha \sigma, \\ v_i + \alpha d_{s_i} &\geq v_{\min} - \alpha \|d_s\| \geq v_{\min} - \alpha \sigma. \end{aligned}$$

¹Recall that $f : \mathbb{R} \rightarrow (0, \infty)$ is called *logarithmically convex* (or log-convex) if the composite function $\log \circ f$ is convex [2]. A natural way to extend this well known notion would be to define a function f to be left g -convex (or shortly g -convex) if $g \circ f$ is convex; in case $f \circ g$ is convex we use the adjective ‘right’ to refer to the reversed order of the factors, and call the function f right g -convex. This terminology is consistent with the above definition of right e-convexity: f is right e-convex if and only if $f \circ \exp$ is convex.

²It may be noted that in [11] it is assumed that the barrier function Ψ itself is right e-convex, whereas in this paper we use right e-convexity of its ‘barrier part’ Ψ_b .

Thus, using also (22), we obtain

$$f_1''(\alpha) \leq h(\alpha) := \frac{1}{2}q\sigma^2 (v_{\min} - \alpha\sigma)^{-q-1}. \quad (29)$$

By integration, this implies, for any α such that $v_{\min} - \alpha\sigma \geq 0$,

$$f_1'(\alpha) = f_1'(0) + \int_0^\alpha f_1''(\xi) d\xi \leq f_1'(0) + \int_0^\alpha h(\xi) d\xi.$$

Since $f(\alpha) \leq f_1(\alpha)$, by integrating once more we get, using $f_1(0) = 0$,

$$f(\alpha) \leq f_1(\alpha) = \int_0^\alpha f_1'(\zeta) d\zeta \leq f_2(\alpha) := f_1'(0)\alpha + \int_0^\alpha \int_0^\zeta h(\xi) d\xi d\zeta.$$

Obviously, $f_2''(\alpha) = h(\alpha) > 0$. Hence $f_2(\alpha)$ is convex in α , and twice differentiable. Since $f_2(0) = 0$, $f_2'(0) = f_1'(0) < 0$, and $f_2''(\alpha) = h(\alpha)$ goes to infinity if α approaches v_{\min}/σ , the function $f_2(\alpha)$ attains its minimal value at some positive value $\tilde{\alpha}$ of α , and $\tilde{\alpha}$ is the stationary point of $f_2(\alpha)$. Hence $\tilde{\alpha}$ is the solution of the equation

$$f_2'(\alpha) = f_1'(0) + \int_0^\alpha h(\xi) d\xi = 0.$$

Using (29), this equation is equivalent to

$$f_1'(0) + \frac{1}{2}q\sigma^2 \int_0^\alpha (v_{\min} - \xi\sigma)^{-q-1} d\xi = 0.$$

Also using (28), and by evaluating the integral, we find that the stationary point $\tilde{\alpha}$ satisfies

$$-\frac{\sigma^2}{2} + \frac{\sigma}{2} ((v_{\min} - \alpha\sigma)^{-q} - v_{\min}^{-q}) = 0.$$

By solving $\bar{\alpha}$ from this equation, we find

$$\bar{\alpha} = \frac{v_{\min}}{\sigma} \left(1 - (1 + \sigma v_{\min}^q)^{-\frac{1}{q}} \right). \quad (30)$$

By Lemma 3.1,

$$v_{\min}^q \geq \frac{1}{\sigma + 1}.$$

Substitution in (30) gives

$$\bar{\alpha} \geq \frac{v_{\min}}{\sigma} \left(1 - \left(1 + \frac{\sigma}{\sigma + 1} \right)^{-\frac{1}{q}} \right).$$

Using (11), in Lemma 2.6, we derive that

$$\left(1 + \frac{\sigma}{\sigma + 1} \right)^{-\frac{1}{q}} = \left(1 - \frac{\sigma}{2\sigma + 1} \right)^{\frac{1}{q}} \leq 1 - \frac{\sigma}{q(2\sigma + 1)},$$

and hence we obtain, using Lemma 3.1 once more,

$$\bar{\alpha} \geq \frac{v_{\min}}{\sigma} \frac{\sigma}{q(2\sigma + 1)} = \frac{v_{\min}}{q(2\sigma + 1)} \geq \frac{1}{q(2\sigma + 1)(\sigma + 1)^{\frac{1}{q}}}. \quad (31)$$

In the sequel we assume $\sigma \geq 1$, and we use the step size

$$\alpha^* = \frac{1}{3q\sigma(\sigma+1)^{\frac{1}{q}}}. \quad (32)$$

Note that $\alpha^* \leq \bar{\alpha}$. To get an estimate for the value of $f_2(\alpha^*)$ we use another useful lemma from [11]. For completeness' sake we include its elementary proof.

Lemma 4.2 (Lemma 3.12 in [11]). *Let $h(t)$ be a twice differentiable convex function with $h(0) = 0$, $h'(0) < 0$ and let $h(t)$ attain its (global) minimum at $t^* > 0$. If $h''(t)$ is increasing for $t \in [0, t^*]$ then*

$$h(t) \leq \frac{th'(0)}{2}, \quad 0 \leq t \leq t^*.$$

Proof: Using the hypothesis of the lemma we may write

$$\begin{aligned} h(t) &= \int_0^t h'(\xi)d\xi = h'(0)t + \int_0^t \int_0^\xi h''(\zeta)d\zeta d\xi \leq h'(0)t + \int_0^t \xi h''(\xi)d\xi = \\ &= h'(0)t + \int_0^t \xi dh'(\xi) = h'(0)t + (\xi h'(\xi))|_0^t - \int_0^t h'(\xi)d\xi \leq h'(0)t - \int_0^t dh'(\xi) = h'(0)t - h(t). \end{aligned}$$

This implies the lemma. □

We apply this lemma with $h = f_2$ and $t = \alpha$. First we verify that the hypotheses of the lemma are satisfied: we have $f_2(0) = 0$ and $f_2'(0) < 0$. Furthermore, $f_2''(\alpha) = h(\alpha) > 0$, and

$$f_2'''(\alpha) = h'(\alpha) = \frac{q(q+1)\sigma^3}{2} (v_{\min} - \alpha\sigma)^{-q-2} > 0.$$

Hence the lemma applies, and using also (28) we obtain

$$f(\alpha^*) \leq f_2(\alpha^*) \leq \frac{\alpha^* f_2'(0)}{2} = \frac{\alpha^* f_1'(0)}{2} = -\frac{\alpha^* \sigma^2}{4}.$$

Finally, by substitution of (32) and using $\sigma \geq 1$, we get

$$f(\alpha^*) \leq \frac{-\sigma}{12q(\sigma+1)^{\frac{1}{q}}} \leq \frac{-\sigma}{12q(2\sigma)^{\frac{1}{q}}} \leq \frac{-\sigma^{\frac{q-1}{q}}}{24q}. \quad (33)$$

Remark 4.3. *If $v_{\min} \geq 1$ then one may easily verify that the same chain of arguments leads to*

$$f(\alpha^*) \leq \frac{-\sigma^2}{4q(\sigma+1)} \leq \frac{-\sigma}{8q},$$

with

$$\alpha^* = \frac{1}{q(\sigma+1)}.$$

5. Effect of updating the barrier parameter

We start this section with some technical results. As before, we write

$$\psi(t) = \frac{t^2 - 1}{2} + \psi_b(t), \quad \psi_b(t) = \frac{t^{1-q} - 1}{q - 1}, \quad q > 1. \tag{34}$$

Since

$$\psi'(t) = t - t^{-q}, \quad \psi''(t) = 1 + qt^{-q-1},$$

we have $\psi''(t) \geq 1$ for all t . This will be used in the proof of the next lemma.

Lemma 5.1. *One has*

$$\frac{1}{2}(t - 1)^2 \leq \psi(t) \leq \frac{1}{2}\psi'(t)^2, \quad t > 0.$$

Proof: Using that $\psi(1) = \psi'(1) = 0$, we may write

$$\psi(t) = \int_1^t \int_1^\xi \psi''(\zeta) d\zeta d\xi \geq \int_1^t \int_1^\xi d\zeta d\xi = \frac{1}{2}(t - 1)^2,$$

which proves the first inequality. The second inequality is obtained as follows:

$$\psi(t) = \int_1^t \int_1^\xi \psi''(\zeta) d\zeta d\xi \leq \int_1^t \int_1^\xi \psi''(\xi) \psi''(\zeta) d\zeta d\xi = \int_1^t \psi''(\xi) \psi'(\xi) d\xi = \int_1^t \psi'(\xi) d\psi'(\xi) = \frac{1}{2}\psi'(t)^2.$$

This completes the proof.

Corollary 5.2. *One has $\|v\| \leq \sqrt{n} + \sqrt{2\Psi(v)}$.*

Proof: Using the first inequality in Lemma 5.1 and the inequality of Cauchy – Schwarz we find

$$2\Psi(v) = 2 \sum_{i=1}^n \psi(v_i) \geq \sum_{i=1}^n (v_i - 1)^2 = \|v\|^2 - 2e^T v + n \geq \|v\|^2 - 2\|e\| \|v\| + \|e\|^2 = (\|v\| - \|e\|)^2.$$

This implies

$$\|v\| \leq \|e\| + \sqrt{2\Psi(v)} = \sqrt{n} + \sqrt{2\Psi(v)},$$

proving the corollary.

Corollary 5.3. *One has $\Psi(v) \leq \frac{1}{2}\sigma(v)^2$.*

Proof: Using the second inequality in Lemma 5.1 and (17) we may write

$$\Psi(v) = \sum_{i=1}^n \psi(v_i) \leq \frac{1}{2} \sum_{i=1}^n \psi'(v_i)^2 = \frac{1}{2} \|\nabla\Psi(v)\|^2 = \frac{1}{2}\sigma(v)^2.$$

Hence the proof is complete. □

We proceed by estimating the increase of $\psi(t)$ when t increases to βt , with $\beta \geq 1$.

Lemma 5.4. *Let $\beta \geq 1$. Then*

$$\psi(\beta t) \leq \psi(t) + \frac{1}{2} (\beta^2 - 1) t^2.$$

Proof: Using the notation of (34), we may write

$$\psi(\beta t) = \frac{\beta^2 t^2 - 1}{2} + \psi_b(\beta t) = \psi(t) + \frac{1}{2} (\beta^2 t^2 - t^2) + \psi_b(\beta t) - \psi_b(t).$$

Since $\psi_b(t)$ is monotonically decreasing, $\psi_b(\beta t) - \psi_b(t) \leq 0$. Therefore, the lemma follows. \square

With the last lemma we are able to estimate the effect of an update of the barrier parameter to the value of the barrier function.

Lemma 5.5. *Let $\theta \in (0, 1)$ and $\mu_+ = (1 - \theta)\mu$. Then*

$$\Phi(xs, \mu_+) \leq \Phi(xs, \mu) + \frac{\theta}{2(1 - \theta)} \left(2\Psi(v) + \sqrt{2n\Psi(v)} + n \right).$$

Proof: Recall that $\Phi(xs, \mu) := \Psi(v)$. When changing μ to μ_+ , the vector v becomes $v_+ = \beta v$, with $\beta = 1/\sqrt{1 - \theta}$. Using Lemma 5.4 with this β , we write

$$\Psi(\beta v) = \sum_{i=1}^n \psi(\beta v_i) \leq \sum_{i=1}^n \left(\psi(v_i) + \frac{1}{2} (\beta^2 - 1) v_i^2 \right) = \Psi(v) + \frac{\theta \|v\|^2}{2(1 - \theta)}.$$

This implies, by Corollary 5.2,

$$\Psi(\beta v) \leq \Psi(v) + \frac{\theta \left(\sqrt{n} + \sqrt{2\Psi(v)} \right)^2}{2(1 - \theta)} = \Psi(v) + \frac{\theta}{2(1 - \theta)} \left(2\Psi(v) + \sqrt{2n\Psi(v)} + n \right).$$

The proof is completed.

6. New algorithm and its complexity analysis

The algorithm we analyse in this section is a slight modification of the ‘classical’ primal-dual algorithm considered in Section 1. In the new algorithm we monitor the progress of inner iterations by the new barrier function $\Phi(xs, \mu)$, as defined in (23). Thus the algorithm goes as follows.

New Primal-Dual Newton Algorithm for LO

Input:

A threshold parameter $\tau \geq 1$;
 an accuracy parameter $\varepsilon > 0$;
 a fixed barrier update parameter θ , $0 < \theta < 1$;

begin

$x := e$; $s := e$; $\mu := 1$;

while $n\mu \geq \varepsilon$ **do**

begin

$\mu := (1 - \theta)\mu$;

while $\Phi(xs, \mu) \geq \tau$ **do**

begin

$x := x + \alpha\Delta x$;

$s := s + \alpha\Delta s$;

$y := y + \alpha\Delta y$

end**end****end**

At the start of an outer iteration, just before the update of the barrier parameter, we have $\Phi(xs, \mu) \leq \tau$. By Lemma 5.5, after the update of the barrier parameter,

$$\Phi(xs, \mu_+) \leq \tau + \frac{\theta}{2(1-\theta)} \left(2\tau + \sqrt{2n\tau} + n \right).$$

Using the default step size α^* , as given by (32), each inner iteration decreases the barrier function with at least

$$-\frac{\sigma^{\frac{q-1}{q}}}{24q}, \quad (35)$$

where $\sigma = \sigma(v)$, by (33). Recall that this result is only valid if $\sigma \geq 1$. This certainly holds if $\tau \geq 1$, because then, at the start of each inner iteration one has $\Phi(xs, \mu) \geq 1$. By Corollary 5.3 this implies $\sigma \geq 1$. As a consequence of (35) and Corollary 5.3 we find that each inner iteration decreases the barrier function value with at least

$$-\frac{\Phi^{\frac{q-1}{2q}}}{24q},$$

where Φ denotes the barrier function value before the inner iteration. This brings us in a situation where we can apply Lemma 2.7. With $\kappa = \frac{1}{24q}$ and $\gamma = \frac{q+1}{2q}$ we then obtain that at most

$$\left\lceil \frac{48q^2 \left(\tau + \frac{\theta}{2(1-\theta)} (2\tau + \sqrt{2n\tau} + n) \right)^{\frac{q+1}{2q}}}{q+1} \right\rceil \quad (36)$$

inner iterations are needed to recenter. We are now ready for the main complexity result of this paper.

Theorem 6.1. *The total number of iterations is at most*

$$\left\lceil \frac{48q^2 \left(\tau + \frac{\theta}{2(1-\theta)} (2\tau + \sqrt{2n\tau} + n) \right)^{\frac{q+1}{2q}}}{q+1} \right\rceil \left\lceil \frac{1}{\theta} \log \frac{n}{\varepsilon} \right\rceil.$$

Proof: According to (8) the number of outer iterations is at most

$$\left\lceil \frac{1}{\theta} \log \frac{n}{\varepsilon} \right\rceil.$$

Multiplying this number by the bound (36) for the number of inner iterations per outer iteration we obtain the theorem. \square

Removing the integer brackets the iteration bound of Theorem 6.1 becomes

$$\frac{48q^2 \left(\tau + \frac{\theta}{2(1-\theta)} (2\tau + \sqrt{2n\tau} + n) \right)^{\frac{q+1}{2q}}}{q+1} \frac{1}{\theta} \log \frac{n}{\varepsilon}.$$

Thus, if $\theta = \frac{2}{3}$ and $\tau = n$ the bound becomes

$$\frac{72q^2 (4n + n\sqrt{2})^{\frac{q+1}{2q}}}{q+1} \log \frac{n}{\varepsilon} \leq 72 (4 + \sqrt{2}) \frac{q^2 n^{\frac{q+1}{2q}}}{q+1} \log \frac{n}{\varepsilon} \leq 390 q n^{\frac{q+1}{2q}} \log \frac{n}{\varepsilon}.$$

Fixing $q > 1$, this bound improves the best bound known for large-update methods, which is $O(n \log \frac{n}{\varepsilon})$. Note that the order of magnitude does not change if we take $\tau = 1$, or $\tau = O(n)$.

Another interesting choice is $q = \frac{1}{2} \log n$, which minimizes the iteration bound. Then the iteration bound becomes

$$O\left(\sqrt{n} \log n \log \frac{n}{\varepsilon}\right),$$

which is quite close to the currently best iteration bound known for interior-point methods, namely $O\left(\sqrt{n} \log \frac{n}{\varepsilon}\right)$.

7. Concluding remarks

We have shown that by slightly modifying the classical primal-dual Newton search direction, the complexity of large-update methods can be improved. From a theoretical point of view, the existing gap between small-update and large-update methods has thus been almost closed.

It has become clear that the obtained complexity result highly depends on the choice of a suitable barrier function $\Phi(xs, \mu)$. As we have shown the gradient of the scaled version $\Psi(v)$ of $\Phi(xs, \mu)$ naturally induces both a search direction and a norm-based proximity measure. The barrier function used in this paper is a member of a wider class of barrier functions, each

of which can be used to derive the same complexity bound as obtained in this paper. This is the subject of the recent paper [11], where we introduced the class of *self-regular barrier functions*. It should be no surprise that the analysis in this paper is much simpler than in [11], since when dealing with one specific barrier function one can explore all nice properties of this function that are not necessarily common to all self-regular barrier functions. It might be mentioned that we extended the theory of IPMs based on self-regular barrier functions in [11] to Semidefinite Optimization, in [12] to Smooth Nonlinear Complementarity problems and in [13] also to optimization over the Second Order Cone.

Finally, let us discuss briefly how the ideas presented in this paper can be incorporated into practical implementations of IPMs. So far only very limited computational tests have been done. Based on these experiments we intend to believe that it might be a good idea to use a dynamic scheme such that when v_{\min} is relatively large, then the classic Newton direction is used, while if it goes too close to zero, then we use our new search direction, with an appropriate value of the parameter q . It requires further investigation to work out the details of such a strategy.

References

- [1] Andersen E. D., Gondzio J., Mészáros Cs., Xu X. Implementation of interior-point methods for large scale linear programming // Interior-Point Methods of Math. T. Terlaky (ed.). Programming. Dordrecht: Kluwer Acad. Publ. 1996. P. 189–252.
- [2] Hiriart-Urruty J. B., Lemarechal C. Convex Analysis and Minimization Algorithms I. Vol. 305 of A series in Comprehensive Studies in Mathematics. Berlin: Springer Verlag, 1996.
- [3] Hung P., Ye Y. An asymptotically $O(\sqrt{n}L)$ -iteration path-following linear programming algorithm that uses long steps // SIAM J. on Optimization. 1996. Vol. 6. P. 570–586.
- [4] Jansen B., Roos C., Terlaky T., Ye Y. Improved complexity using higher-order correctors for primal-dual Dikin affine scaling // Math. Programming. 1997. Vol. 76. P. 117–130.
- [5] Karmarkar N. K. A new polynomial-time algorithm for linear programming // Combinatorica. 1984. Vol. 4. P. 373–395.
- [6] Megiddo N. Pathways to the optimal set in linear programming // Progress in Math. Programming: Interior-Point and Related Methods. N. Megiddo (ed.). N. Y.: Springer Verlag, 1989. P. 131–158. [Proc. of the 6th Math. Programming Symp. of Japan, Nagoya. 1986. P. 1–35.]
- [7] Mehrotra S., Ye Y. On finding the optimal facet of linear programs // Math. Programming. 1993. Vol. 62. P. 497–515.
- [8] Monteiro R. D. C., Adler I., Resende M. G. C. A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series extension // Mathematics of Operations Research. 1990. Vol. 15. P. 191–214.
- [9] Peng J., Roos C., Terlaky T. New complexity analysis of the primal-dual Newton method for linear optimization // Annals of Operations Research. To appear.
- [10] Peng J., Roos C., Terlaky T. A new class of polynomial primal-dual methods for linear and semidefinite optimization // Europ. J. of Operations Research. Submitted.

- [11] Peng J., Roos C., Terlaky T. Self-regular proximities and new search directions for linear and semidefinite optimization // *Math. Programming.* 2000. Accepted.
- [12] Peng J., Roos C., Terlaky T., Yoshise A. Self-regular proximities and new directions for nonlinear $P_*(\kappa)$ complementarity problems // *Math. of Operations Research.* 2000. Submitted.
- [13] Peng J., Roos C., Terlaky T. New primal-dual algorithms for second-order cone optimization based on self-regular proximities // *SIAM J. on Optimization.* 2000. Submitted.
- [14] Roos C., Terlaky T., Vial J.-Ph. *Theory and Algorithms for Linear Optimization. An Interior-Point Approach.* Chichester: John Wiley & Sons, 1997.
- [15] Sonnevend G. An “Analytic center” for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming / *System Modelling and Optimization* // *Proc. of the 12th IFIP-Conf.* A. Prékopa, J. Szelezsán, B. Strazicky (Eds). Budapest, Hungary, Sept. 1985. Vol. 84 [Lecture Notes in Control and Information Sci. Berlin: Springer Verlag, 1986. P. 866–876.]
- [16] Vanderbei R. J. *Linear Programming: Foundations and Extensions.* Boston: Kluwer Acad. Publ., 1996.
- [17] Wright S. J. *Primal-Dual Interior-Point Methods.* SIAM, Philadelphia, USA, 1997.
- [18] Ye Y. On the finite convergence of interior-point algorithms for linear programming // *Math. Programming.* 1992. Vol. 57. P. 325–335.
- [19] Ye Y. *Interior-Point Algorithms, Theory and Analysis.* Chichester: John Wiley & Sons, 1997.
- [20] Zhang Y., Zhang D. On polynomiality of the Mehrotra-type predictor-corrector interior-point algorithms // *Math. Programming.* 1995. Vol. 68. P. 303–318.
- [21] Zhao G. Interior-point algorithms for linear complementarity problems based on large neighborhoods of the central path // *SIAM J. on Optimization.* 1998. Vol. 8. P. 397–413.

Received for publication February 1, 2001