Improved Full-Newton Step $O(nL)$ Infeasible Interior-Point Method for Linear Optimization

G. Gu · H. Mansouri · M. Zangiabadi · Y.Q. Bai · C. Roos

Abstract We present several improvements of the full-Newton step infeasible interior-point method for linear optimization introduced by Roos (SIAM J. Optim. 16(4):1110–1136, 2006). Each main step of the method consists of a feasibility step and several centering steps. We use a more natural feasibility step, which targets the μ^+-center of the next pair of perturbed problems. As for the centering steps, we apply a sharper quadratic convergence result, which leads to a slightly wider neighborhood for the feasibility steps. Moreover, the analysis is much simplified and the iteration bound is slightly better.

G. Gu (✉) · C. Roos
Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University
of Technology, P.O. Box 5031, 2600 GA Delft, Netherlands
e-mail: G.Gu@tudelft.nl

C. Roos
e-mail: C.Roos@tudelft.nl

H. Mansouri · M. Zangiabadi
Department of Mathematical Science, Shahrekord University, P.O. Box 115, Shahrekord, Iran

H. Mansouri
e-mail: H.Mansouri@tudelft.nl

M. Zangiabadi
e-mail: M.Zangiabadi@tudelft.nl

Y.Q. Bai
Department of Mathematics, Shanghai University, Shanghai, 200444, China
e-mail: yqbai@shu.edu.cn

1 Introduction

We consider the linear optimization (LO) problem in the standard form

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

with its dual problem

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

Here $A \in \mathbf{R}^{m \times n}$, b, $y \in \mathbf{R}^m$, and $c, x, s \in \mathbf{R}^n$. Without loss of generality, we assume that $\operatorname{rank}(A) = m$. The vectors x, y, s are the vectors of variables.

In [1], a new infeasible interior-point method (IIPM) is proposed to solve the above LO problems. It differs from the classical IIPMs (e.g. [2–9]) in that the new method uses only full steps (instead of damped steps), which has the advantage that no line searches are needed. Our motivation for the use of full-Newton steps is that, though such methods are less greedy, the best complexity results for interior-point methods are obtained for such methods. In our approach, as in [6], the size of the residual vectors reduces with the same speed as the duality gap.

This paper is organized as follows. First, we present some results for feasible interior-point methods (IPMs), as these will be used to analyze the centering steps of our IIPM. Then, we present our improved full-Newton step IIPM. Each main step of the method consists of a feasibility step and several centering steps. We use a more natural feasibility step than in [1, 10], which targets the μ^+-center (see Remark 3.1). Moreover, for the centering steps, we apply a sharper quadratic convergence result, which results in a slightly wider neighborhood for the feasibility steps. Besides these, the analysis is much simplified and the iteration bound is slightly better. Finally, we give some concluding remarks.

2 Full-Newton Step for Feasible IPMs

In preparation for dealing with our IIPM, we recall briefly the notions of central path and of a feasible full-Newton step, as well as some of their properties. We refer to [11, 12] for more details, where it is also described how to obtain a polynomial-time algorithm by using such steps. To solve problems (P) and (D), one needs to find a solution of the following system of equations:[1]

$$Ax = b, \quad x \geq 0,$$
$$A^{\mathrm{T}}y + s = c, \quad s \geq 0,$$
$$xs = 0.$$

In these optimality conditions, the first two constraints represent primal and dual feasibility, whereas the last equation is the so-called complementarity condition. The

---

[1]We denote by 0 and e (used later) the zero vector and the all-one vector, respectively, of appropriate size. Moreover, if $x, s \in \mathbf{R}^n$, then xs denotes the componentwise or Hadamard product of the vectors x and s.

nonnegativity constraints in the feasibility conditions make the problem already non-trivial: only iterative methods can find solutions of linear systems involving inequality constraints. The complementarity condition is nonlinear, which makes it extra hard to solve this system.

2.1 Central Path

IPMs replace the complementarity condition with the so-called centering condition $xs = \mu e$, where μ may be any positive number. This yields the system

$$Ax = b, \quad x \geq 0,$$
$$A^{\mathrm{T}}y + s = c, \quad s \geq 0,$$
$$xs = \mu e. \tag{1}$$

Surprisingly enough, if this system has a solution for some $\mu > 0$, then a solution exists for every $\mu > 0$, and this solution is unique. This happens if and only if problems (P) and (D) satisfy the interior-point condition (IPC); i.e., if (P) has a feasible solution $x > 0$ and (D) has a solution (y, s) with $s > 0$ (see, e.g., [11]). If the IPC is satisfied, then the solution of (1) is denoted by $(x(\mu), y(\mu), s(\mu))$ and is called the μ-center of (P) and (D). The set of all μ-centers forms a path, which is called the central path. As μ goes to zero, $x(\mu)$, $y(\mu)$, $s(\mu)$ converge to optimal solutions of problems (P) and (D). Of course, the system (1) is still hard to solve, but by applying Newton's method one can easily find approximate solutions.

2.2 Newton Step

We proceed to describe the Newton method for solving (1) with μ fixed. Given any x and (y, s), we want to find displacements Δx, Δy, Δs such that

$$A(x + \Delta x) = b,$$
$$A^{\mathrm{T}}(y + \Delta y) + s + \Delta s = c,$$
$$(x + \Delta x)(s + \Delta s) = \mu e.$$

Neglecting the quadratic term $\Delta x \Delta s$ in the left-hand side of the third equation, we obtain the following linear system of equations in the search directions Δx, Δy, Δs:

$$A\Delta x = b - Ax,$$
$$A^{\mathrm{T}}\Delta y + \Delta s = c - A^{\mathrm{T}}y - s,$$
$$s\Delta x + x\Delta s = \mu e - xs. \tag{2}$$

Since A has full row rank, and since the vectors x and s are positive, one may easily verify that the coefficient matrix in the linear system (2) is nonsingular. Hence, this system uniquely defines the search directions Δx, Δy, Δs. These search directions are used in all existing primal-dual (feasible and infeasible) IPMs.

If x is primal feasible and (y, s) is dual feasible, then

$$b - Ax = 0,$$
$$c - A^{T}y - s = 0,$$

whence the above system reduces to

$$A\Delta x = 0,$$
$$A^{T}\Delta y + \Delta s = 0,$$
$$s\Delta x + x\Delta s = \mu e - xs, \tag{3}$$

which gives the usual search directions for feasible primal-dual IPMs. Then the new iterates are given by

$$x^{+} = x + \Delta x,$$
$$y^{+} = y + \Delta y,$$
$$s^{+} = s + \Delta s.$$

An important observation is that Δx lies in the null space of A, whereas Δs belongs to the row space of A. This implies that Δx and Δs are orthogonal, i.e., $\Delta x^{T}\Delta s = 0$. As a consequence, we have the important property that, after a full-Newton step, the duality gap assumes the same value as at the μ-centers, namely $n\mu$.

Lemma 2.1 (See [11], Lemma II.47) *After a primal-dual Newton step, one has* $(x^{+})^{T}s^{+} = n\mu$.

We use the quantity $\delta(x, s; \mu)$ to measure the proximity of a feasible triple (x, y, s) to the μ-center $(x(\mu), y(\mu), s(\mu))$. Following [1, 11], this quantity is defined as follows:[2]

$$\delta(x, s; \mu) := \delta(v) := \frac{1}{2}\|v - v^{-1}\|, \quad \text{where } v := \sqrt{\frac{xs}{\mu}}. \tag{4}$$

It is crucial for us to know the effect on $\delta(x, s; \mu)$ of a full-Newton step targeting the μ-center of (P) and (D). For that purpose, Theorem II.50 of [11] was used in [1]. This theorem states that, if $\delta := \delta(x, s; \mu) \leq 1$, then the primal-dual Newton step is feasible, i.e., x^{+} and s^{+} are nonnegative; moreover, if $\delta < 1$, then x^{+} and s^{+} are positive and

$$\delta(x^{+}, s^{+}; \mu) \leq \frac{\delta^{2}}{\sqrt{2(1 - \delta^{2})}}.$$

This result implies that the Newton process is locally quadratically convergent, and has been crucial in the analysis in [1]. We use a tighter upper bound for $\delta(x^{+}, s^{+}; \mu)$,

---

[2]The short-hand notation in the definition of v means that v is the vector obtained by taking square roots of the elements of the vector xs/μ.

which provides a slightly wider neighborhood for the feasibility step of our IIPM. As the previous lemma, we recall it without proof.

Theorem 2.1 ([11], Theorem II.52) *If* $\delta := \delta(x, s; \mu) < 1$, *then*

$$\delta(x^+, s^+; \mu) \leq \frac{\delta^2}{\sqrt{2(1 - \delta^4)}}.$$

As a result, the following corollary follows trivially.

Corollary 2.1 *If* $\delta := \delta(x, s; \mu) \leq \frac{1}{\sqrt[4]{2}}$, *then* $\delta(x^+, s^+; \mu) \leq \delta^2$.

3 Full-Newton Step IIPM

In the case of an infeasible method, we call the triple (x, y, s) an ε-optimal solution of (P) and (D) if the 2-norms of the residual vectors $b - Ax$ and $c - A^\mathrm{T}y - s$ do not exceed ε, and if the duality gap satisfies $x^\mathrm{T}s \leq \varepsilon$. In this section, we present an infeasible-start algorithm that generates an ε-optimal solution of (P) and (D), if it exists, or establishes that no such solution exists.

3.1 Perturbed Problems

We start with choosing arbitrarily $x^0 > 0$ and (y^0, s^0), with $s^0 > 0$ such that $x^0 s^0 = \mu^0 e$ for some (positive) number μ^0. We denote the initial values of the primal and dual residuals r_b^0 and r_c^0 respectively as

$$r_b^0 = b - Ax^0,$$
$$r_c^0 = c - A^\mathrm{T}y^0 - s^0.$$

For any ν with $0 < \nu \leq 1$ we consider the perturbed problem (P_ν), defined by

$$(\mathrm{P}_\nu) \quad \min\{(c - \nu r_c^0)^\mathrm{T}x : Ax = b - \nu r_b^0, \ x \geq 0\},$$

and its dual problem (D_ν), which is given by

$$(\mathrm{D}_\nu) \quad \max\{(b - \nu r_b^0)^\mathrm{T}y : A^\mathrm{T}y + s = c - \nu r_c^0, \ s \geq 0\}.$$

We note that, if $\nu = 1$, then $x = x^0$ yields a strictly feasible solution of (P_ν) and $(y, s) = (y^0, s^0)$ yields a strictly feasible solution of (D_ν). We conclude that, if $\nu = 1$, then (P_ν) and (D_ν) satisfy the IPC.

Theorem 3.1 ([12, Theorem 5.13]) *The original problems* (P) *and* (D) *are feasible if and only if, for each* ν *satisfying* $0 < \nu \leq 1$, *the perturbed problems* (P_ν) *and* (D_ν) *satisfy the IPC.*

In the sections to follow, we assume that (P) and (D) are feasible. Only in Sect. 4.5 will we discuss how our algorithm can be used to detect infeasibility or unboundedness of (P) and (D). It may be worth noting that, if x^0 and (y^0, s^0) are feasible for (P) and (D), then $(P_\nu) \equiv (P)$ and $(D_\nu) \equiv (D)$ for each $\nu \in (0, 1]$.

3.2 Central Path of the Perturbed Problems

Let (P) and (D) be feasible and $0 < \nu \le 1$. Then, Theorem 3.1 implies that the perturbed problems (P_ν) and (D_ν) satisfy the IPC; hence, their central paths exist. This means that the system

$$Ax = b - \nu r_b^0, \quad x \ge 0, \tag{5}$$

$$A^\mathrm{T} y + s = c - \nu r_c^0, \quad s \ge 0, \tag{6}$$

$$xs = \mu e$$

has a unique solution for every $\mu > 0$. This unique solution is denoted by $(x(\mu, \nu), y(\mu, \nu), s(\mu, \nu))$ and is the μ-center of the perturbed problems (P_ν) and (D_ν). In what follows, the parameters μ and ν always satisfy the relation $\mu = \nu \mu^0$. Thus, we may denote the μ-centers of the perturbed problems (P_ν) and (D_ν) simply as $(x(\nu), y(\nu), s(\nu))$.

Note that, since $x^0 s^0 = \mu^0 e$, x^0 is the μ^0-center of the perturbed problem (P_1) and (y^0, s^0) is the μ^0-center of the perturbed problem (D_1). In other words, $(x(1), y(1), s(1)) = (x^0, y^0, s^0)$.

3.3 An Iteration of the Algorithm

We just established that, if $\nu = 1$ and $\mu = \mu^0$, then $x = x^0$ is the μ-center of the perturbed problem (P_ν) and $(y, s) = (y^0, s^0)$ is the μ-center of (D_ν). This pair is our initial iterate.

We measure proximity to the μ-center of the perturbed problems by the quantity $\delta(x, s; \mu)$ as defined in (4). Thus, initially we have $\delta(x, s; \mu) = 0$. In what follows, we assume that, at the start of each iteration, just before the feasibility step, $\delta(x, s; \mu)$ is smaller than or equal to a (small) threshold value $\tau > 0$. So, this is certainly true at the start of the first iteration.

Now, we describe one (main) iteration of our algorithm. Suppose that, for some $\mu \in (0, \mu^0]$, we have (x, y, s) satisfying the feasibility conditions (5) and (6) with $\nu = \mu/\mu^0$ and such that $x^\mathrm{T} s = n\mu$ and $\delta(x, s; \mu) \le \tau$. We reduce μ to $\mu^+ = (1 - \theta)\mu$, with $\theta \in (0, 1)$, and find a new iterate (x^+, y^+, s^+) that satisfies (5) and (6), with ν replaced by $\nu^+ = (1 - \theta)\nu = \mu^+/\mu^0$, and such that $(x^+)^\mathrm{T} s^+ = n\mu^+$ and $\delta(x^+, s^+; \mu^+) \le \tau$.

To be more precise, this is achieved as follows. Each main iteration consists of a feasibility step and a few centering steps. The feasibility step serves to get an iterate (x^f, y^f, s^f) that is strictly feasible for (P_{ν^+}) and (D_{ν^+}) and close to their μ^+-center $(x(\nu^+), y(\nu^+), s(\nu^+))$. In fact, the feasibility step is designed in such a way that $\delta(x^f, s^f; \mu^+) \le 1/\sqrt[4]{2}$, i.e., (x^f, y^f, s^f) lies in the quadratic convergence neighborhood with respect to the μ^+-center of (P_{ν^+}) and (D_{ν^+}). Then we can easily get

an iterate (x^+, y^+, s^+) that is strictly feasible for $(P_{\nu+})$ and $(D_{\nu+})$ and such that $(x^+)^\mathsf{T} s^+ = n\mu^+$ and $\delta(x^+, s^+; \mu^+) \le \tau$, just by performing a few centering steps starting from (x^f, y^f, s^f) and targeting the μ^+-center of $(P_{\nu+})$ and $(D_{\nu+})$.

In the following, we describe the feasibility step in detail. Suppose that we have a strictly feasible iterate (x, y, s) for (P_ν) and (D_ν). This means that (x, y, s) satisfies (5) and (6), with $\nu = \mu/\mu^0$. We need displacements $\Delta^f x$, $\Delta^f y$, $\Delta^f s$ such that

$$
\begin{aligned}
x^f &= x + \Delta^f x, \\
y^f &= y + \Delta^f y, \\
s^f &= s + \Delta^f s,
\end{aligned}
$$

are feasible for $(P_{\nu+})$ and $(D_{\nu+})$. One may verify easily that (x^f, y^f, s^f) satisfies (5) and (6), with ν replaced by $\nu^+ = (1 - \theta)\nu$, only if the first two equations in the following system are satisfied:

$$
A\Delta^f x = \theta \nu r_b^0, \tag{7}
$$

$$
A^\mathsf{T}\Delta^f y + \Delta^f s = \theta \nu r_c^0, \tag{8}
$$

$$
s\Delta^f x + x\Delta^f s = (1 - \theta)\mu e - xs. \tag{9}
$$

The third equation is inspired by the third equation in the system (3) that we used to define the search directions for the feasible case, except that we target the μ^+-centers of $(P_{\nu+})$ and $(D_{\nu+})$.

Remark 3.1 For (9), we use the linearization of $x^f s^f = (1 - \theta)\mu e$, which means that we are targeting the μ^+-center of $(P_{\nu+})$ and $(D_{\nu+})$. While in [1], the linearization of $x^f s^f = \mu e$ is used (targeting the μ-center), and in [10], the linearization of $x^f s^f = xs$ is used (targeting the old xs). As our aim is to calculate a feasible solution to $(P_{\nu+})$ and $(D_{\nu+})$, which should also lie in the quadratic convergence neighborhood to its μ^+-center, the direction used here is more natural and better intuitively.

We conclude that, after the feasibility step, the iterate satisfies the affine equations (5) and (6), with $\nu = \nu^+$. The hard part in the analysis is to guarantee that x^f and s^f are positive and satisfy $\delta(x^f, s^f; \mu^+) \le 1/\sqrt[4]{2}$.

After the feasibility step, we perform a few centering steps in order to get iterate (x^+, y^+, s^+) which satisfies $(x^+)^\mathsf{T} s^+ = n\mu^+$ and $\delta(x^+, s^+; \mu^+) \le \tau$. By using Corollary 2.1, the required number of centering steps can be obtained easily. Indeed, assuming $\delta = \delta(x^f, s^f; \mu^+) \le 1/\sqrt[4]{2}$, after k centering steps we will have iterates (x^+, y^+, s^+) that are still feasible for $(P_{\nu+})$ and $(D_{\nu+})$ and satisfy

$$
\delta(x^+, s^+; \mu^+) \le \left(\frac{1}{\sqrt[4]{2}} \right)^{2^k}.
$$

From this, one deduces easily that $\delta(x^+, s^+; \mu^+) \le \tau$ holds after at most

$$
2 + \left\lceil \log_2 \left(\log_2 \frac{1}{\tau} \right) \right\rceil \tag{10}
$$

centering steps.

3.4 Algorithm

A more formal description of the algorithm is given in Algorithm 1.

Algorithm 1 A Full-Newton Step Infeasible IPM for LO

Step 0: Accuracy parameter $\varepsilon > 0$;
 update parameter θ, $0 < \theta < 1$;
 threshold parameter $\tau > 0$;
 starting point (x^0, y^0, s^0) with $x^0 > 0$, $s^0 > 0$ and $x^0 s^0 = \mu^0 e$;
 $x := x^0$; $y := y^0$; $s := s^0$; $\mu := \mu^0$; $\nu := 1$.
Step 1: Feasibility step: $(x, y, s) := (x, y, s) + (\Delta^f x, \Delta^f y, \Delta^f s)$.
Step 2: Update of μ and ν: $\mu := (1 - \theta)\mu$; $\nu := (1 - \theta)\nu$.
Step 3: Centering step: $(x, y, s) := (x, y, s) + (\Delta x, \Delta y, \Delta s)$.
Step 4: If $\delta(x, s; \mu) > \tau$, go to Step 3.
Step 5: If $\max(x^T s, \|b - Ax\|, \|c - A^T y - s\|) \leq \varepsilon$, stop; else, go to Step 1.

Note that, after each iteration, the residuals and the duality gap are reduced the factor $1 - \theta$. The algorithm stops if the norms of the residuals and the duality gap are less than the accuracy parameter ε.

4 Analysis of the Algorithm

Let x, y, s denote the iterates at the start of a main iteration; i.e., for some $\mu \in (0, \mu^0]$, we have x, y, s satisfying the feasibility conditions (5) and (6) with $\mu = \nu\mu^0$ and such that $x^T s = n\mu$ and $\delta(x, s; \mu) \leq \tau$. Recall that, at the start of the first iteration, this is certainly true.

4.1 Feasibility Step

As we established in Sect. 3.3, the feasibility step generates new iterate (x^f, y^f, s^f) that satisfies the feasibility conditions for $(P_{\nu+})$ and $(D_{\nu+})$, except possibly the non-negativity constraints. A crucial element in the analysis is to show that, after the feasibility step $\delta(x^f, s^f; \mu^+) \leq 1/\sqrt[4]{2}$, i.e., the iterate (x^f, y^f, s^f) is within the neighborhood where the Newton process targeting the μ^+-center of $(P_{\nu+})$ and $(D_{\nu+})$ is quadratically convergent.

Define

$$d_x^f := \frac{\nu \Delta^f x}{x}, \qquad d_s^f := \frac{\nu \Delta^f s}{s}, \tag{11}$$

where ν is defined in (4). We have, using (9) and (11),

$$x^f s^f = xs + (s\Delta^f x + x\Delta^f s) + \Delta^f x \Delta^f s$$

$$= (1 - \theta)\mu e + \Delta^f x \Delta^f s$$

$$= \mu\left[(1 - \theta)e + d_x^f d_s^f\right]. \tag{12}$$

Lemma 4.1 (See [11], Lemma II.46) *The iterates* (x^f, y^f, s^f) *are strictly feasible if and only if* $(1 - \theta)e + d_x^f d_s^f > 0$.

Proof Note that, if x^f and s^f are positive, then (12) makes clear that $(1 - \theta)e + d_x^f d_s^f > 0$, proving the only if part of the lemma. For the proof of the converse implication, we introduce a steplength $\alpha \in [0, 1]$ and we define

$$x^\alpha = x + \alpha \Delta^f x, \qquad s^\alpha = s + \alpha \Delta^f s.$$

We then have $x^0 = x$, $x^1 = x^f$ and a similar convention for s. Hence, we have $x^0 s^0 = xs > 0$. We write

$$x^\alpha s^\alpha = (x + \alpha \Delta^f x)(s + \alpha \Delta^f s) = xs + \alpha(s\Delta^f x + x\Delta^f s) + \alpha^2 \Delta^f x \Delta^f s.$$

Using (4), (9), (11) gives

$$x^\alpha s^\alpha = xs + \alpha[(1 - \theta)\mu e - xs] + \alpha^2 \Delta^f x \Delta^f s$$
$$= \mu[(1 - \alpha)v^2 + \alpha(1 - \theta)e + \alpha^2 d_x^f d_s^f].$$

Now, suppose that

$$(1 - \theta)e + d_x^f d_s^f > 0.$$

Then,

$$d_x^f d_s^f > -(1 - \theta)e.$$

Substitution gives

$$x^\alpha s^\alpha > \mu\big[(1 - \alpha)v^2 + \alpha(1 - \theta)e - \alpha^2(1 - \theta)e\big]$$
$$= \mu(1 - \alpha)\big[v^2 + \alpha(1 - \theta)e\big], \quad \alpha \in [0, 1].$$

Since

$$\mu(1 - \alpha)\big[v^2 + \alpha(1 - \theta)e\big] \geq 0,$$

it follows that $x^\alpha s^\alpha > 0$ for $0 \leq \alpha \leq 1$. Hence, none of the entries of x^α and s^α vanishes for $0 \leq \alpha \leq 1$. Since x^0 and s^0 are positive, and since x^α and s^α depend linearly on α, this implies that $x^\alpha > 0$ and $s^\alpha > 0$ for $0 \leq \alpha \leq 1$. Hence, x^1 and s^1 must be positive, proving the if part of Lemma 4.1.   □

We proceed by deriving an upper bound for $\delta(x^f, s^f; \mu^+)$. According to definition (4), one has

$$\delta(x^f, s^f; \mu^+) = \frac{1}{2}\left\| v^f - \frac{e}{v^f} \right\|, \quad \text{where } v^f = \sqrt{\frac{x^f s^f}{\mu^+}}.$$

In the sequel, we denote $\delta(x^f, s^f; \mu^+)$ by $\delta(v^f)$ and we have the following result.

Lemma 4.2 *If $\|d_x^f d_s^f\|_\infty < 1 - \theta$, then*

$$4\delta(v^f)^2 \leq \frac{\left\|\frac{d_x^f d_s^f}{1-\theta}\right\|^2}{1 - \left\|\frac{d_x^f d_s^f}{1-\theta}\right\|_\infty}.$$

Proof To simplify the notation in this proof, let $z := \frac{d_x^f d_s^f}{1-\theta}$. After division of both sides in (12) by μ^+, we get

$$(v^f)^2 = \frac{\mu[(1-\theta)e + d_x^f d_s^f]}{\mu^+} = \frac{\mu[(1-\theta)e + (1-\theta)z]}{(1-\theta)\mu} = e + z.$$

Hence, we have

$$
\begin{aligned}
4\delta(v^f)^2 &= \sum_{i=1}^n \left((v_i^f)^2 + (v_i^f)^{-2} - 2\right) = \sum_{i=1}^n \left(1 + z_i + \frac{1}{1+z_i} - 2\right) \\
&= \sum_{i=1}^n \frac{z_i^2}{1+z_i} \leq \sum_{i=1}^n \frac{z_i^2}{1-|z_i|} \leq \sum_{i=1}^n \frac{z_i^2}{1-\|z\|_\infty} = \frac{\|z\|^2}{1-\|z\|_\infty},
\end{aligned}
$$

where the inequalities are due to $\|z\|_\infty < 1$. This proves Lemma 4.2.   □

4.2 First Upper Bound for θ

Because we need to have $\delta(v^f) \leq 1/\sqrt[4]{2}$, it follows from Lemma 4.2 that it suffices to have

$$\frac{\left\|\frac{d_x^f d_s^f}{1-\theta}\right\|^2}{1 - \left\|\frac{d_x^f d_s^f}{1-\theta}\right\|_\infty} \leq 2\sqrt{2}. \tag{13}$$

We may easily verify that

$$\|d_x^f d_s^f\|^2 \leq \left(\|d_x^f\|\|d_s^f\|\right)^2 \leq \frac{1}{4}\left(\|d_x^f\|^2 + \|d_s^f\|^2\right)^2$$

and

$$\|d_x^f d_s^f\|_\infty \leq \frac{1}{2}\left(\|d_x^f\|_\infty^2 + \|d_s^f\|_\infty^2\right) \leq \frac{1}{2}\left(\|d_x^f\|^2 + \|d_s^f\|^2\right). \tag{14}$$

For the moment, we assume that

$$\frac{\|d_x^f\|^2 + \|d_s^f\|^2}{1-\theta} < 2.$$

Then,

$$\left\|\frac{d_x^f d_s^f}{1-\theta}\right\|_\infty < 1,$$

whence inequality (13) holds if

$$\frac{\frac{1}{4}\left(\frac{\|d_x^f\|^2+\|d_s^f\|^2}{1-\theta}\right)^2}{1-\frac{1}{2}\frac{\|d_x^f\|^2+\|d_s^f\|^2}{1-\theta}} \leq 2\sqrt{2}.$$

Considering $\frac{\|d_x^f\|^2+\|d_s^f\|^2}{1-\theta}$ as a single term, and by some elementary calculations, we obtain that (13) holds if

$$\frac{\|d_x^f\|^2 + \|d_s^f\|^2}{1-\theta} \leq 2\sqrt{2}\left(\sqrt{1+\sqrt{2}}-1\right) \approx 1.566. \tag{15}$$

Also by Lemma 4.1 and inequality (14), the strict feasibility of (x^f, y^f, s^f) can be derived from (15). In other words, the inequality (15) implies that, after the feasibility step, (x^f, y^f, s^f) is strictly feasible and lies in the quadratic convergence neighborhood with respect to the μ^+-center of $(P_{\nu+})$ and $(D_{\nu+})$.

4.3 Upper Bound for $\|d_x^f\|^2 + \|d_s^f\|^2$

Obtaining an upper bound for $\|d_x^f\|^2 + \|d_s^f\|^2$ is the subject of this subsection. In subsequent subsections, this will enable us to find a default value for θ.

One may easily check that the system (7)–(9), which defines the search directions $\Delta^f x$, $\Delta^f y$, $\Delta^f s$, can be expressed in terms of the scaled search directions d_x^f and d_s^f as follows:

$$\bar{A}d_x^f = \theta v r_b^0, \tag{16}$$

$$\bar{A}^{\mathrm{T}}\frac{\Delta^f y}{\mu} + d_s^f = \theta v v s^{-1} r_c^0, \tag{17}$$

$$d_x^f + d_s^f = (1-\theta)v^{-1} - v, \tag{18}$$

where

$$\bar{A} = AV^{-1}X, \quad V = \mathrm{diag}(v), \quad X = \mathrm{diag}(x).$$

From the above definition of $\bar{A}$ we deduce that $\bar{A} = \sqrt{\mu}AD$, where

$$D = \mathrm{diag}\left(\frac{xv^{-1}}{\sqrt{\mu}}\right) = \mathrm{diag}\left(\sqrt{\frac{x}{s}}\right) = \mathrm{diag}\left(\sqrt{\mu}vs^{-1}\right).$$

For the moment, let us define

$$r_b := \theta v r_b^0, \qquad r_c := \theta v r_c^0, \qquad r := (1-\theta)v^{-1} - v. \tag{19}$$

With $\xi = -\frac{\Delta^f y}{\mu}$, by eliminating d_s^f from (16)–(18), we then have

$$\sqrt{\mu}ADd_x^f = r_b, \tag{20}$$

$$\sqrt{\mu}DA^\mathrm{T}\xi + d_x^f = r - \frac{1}{\sqrt{\mu}}Dr_c. \tag{21}$$

By multiplying both sides of (21) from the left with $\sqrt{\mu}AD$ and using (20), it follows that

$$\mu AD^2A^\mathrm{T}\xi + r_b = \sqrt{\mu}AD\left(r - \frac{1}{\sqrt{\mu}}Dr_c\right).$$

Therefore,

$$\xi = \frac{1}{\mu}(AD^2A^\mathrm{T})^{-1}\left[\sqrt{\mu}AD\left(r - \frac{1}{\sqrt{\mu}}Dr_c\right) - r_b\right].$$

Substitution into (21) gives

$$d_x^f = r - \frac{1}{\sqrt{\mu}}Dr_c - \frac{1}{\sqrt{\mu}}DA^\mathrm{T}(AD^2A^\mathrm{T})^{-1}\left[\sqrt{\mu}AD\left(r - \frac{1}{\sqrt{\mu}}Dr_c\right) - r_b\right]$$

$$= \left[I - DA^\mathrm{T}(AD^2A^\mathrm{T})^{-1}AD\right]\left(r - \frac{1}{\sqrt{\mu}}Dr_c\right) + \frac{1}{\sqrt{\mu}}DA^\mathrm{T}(AD^2A^\mathrm{T})^{-1}r_b.$$

To simplify notation, we denote

$$P = DA^\mathrm{T}(AD^2A^\mathrm{T})^{-1}AD.$$

Note that P is the matrix of the orthogonal projection to the row space of the matrix AD. Now, we may write

$$d_x^f = [I - P]\left(r - \frac{1}{\sqrt{\mu}}Dr_c\right) + \frac{1}{\sqrt{\mu}}DA^\mathrm{T}(AD^2A^\mathrm{T})^{-1}r_b.$$

Let $(\bar{x}, \bar{y}, \bar{s})$ be such that $A\bar{x} = b$ and $A^\mathrm{T}\bar{y} + \bar{s} = c$. Then, we may write

$$r_b = \theta v r_b^0 = \theta v(b - Ax^0) = \theta v A(\bar{x} - x^0),$$

$$r_c = \theta v r_c^0 = \theta v(c - A^\mathrm{T}y^0 - s^0) = \theta v\left(A^\mathrm{T}(\bar{y} - y^0) + \bar{s} - s^0\right).$$

Thus, we obtain

$$d_x^f = [I - P]\left(r - \frac{\theta v}{\sqrt{\mu}}D\left(A^\mathrm{T}(\bar{y} - y^0) + \bar{s} - s^0\right)\right) + \frac{\theta v}{\sqrt{\mu}}PD^{-1}(\bar{x} - x^0).$$

Since $I - P$ is the orthogonal projection to the null space of AD, we have

$$[I - P]DA^\mathrm{T}(\bar{y} - y^0) = 0$$

and the expression for d_x^f reduces to

$$d_x^f = [I - P]\left(r - \frac{\theta v}{\sqrt{\mu}}D\left(\bar{s} - s^0\right)\right) + \frac{\theta v}{\sqrt{\mu}}PD^{-1}(\bar{x} - x^0).$$

To proceed, we further simplify the notation by defining

$$u^x = \frac{\theta v}{\sqrt{\mu}} D^{-1}(\bar{x} - x^0), \qquad u^s = \frac{\theta v}{\sqrt{\mu}} D(\bar{s} - s^0). \tag{22}$$

Then, we may write

$$d_x^f = [I - P](r - u^s) + P u^x.$$

For d_s^f, by using (18) and the definition (19) of r, we obtain

$$d_s^f = r - d_x^f = r - [I - P]r + [I - P]u^s - P u^x$$
$$= [I - P]u^s + P(r - u^x).$$

We denote $[I - P]r = r_1$ and $Pr = r_2$, and use similar notations for the projection of u^x and u^s. Then, from the above expressions for d_x^f and d_s^f, we derive that

$$d_x^f = r_1 - u_1^s + u_2^x, \qquad d_s^f = u_1^s + r_2 - u_2^x.$$

Therefore, using the orthogonality of vectors with different subscripts, we may write

$$\|d_x^f\|^2 + \|d_s^f\|^2$$
$$= \|r_1 - u_1^s\|^2 + \|u_2^x\|^2 + \|u_1^s\|^2 + \|r_2 - u_2^x\|^2$$
$$= \|r_1\|^2 + \|u_1^s\|^2 - 2r_1^{\mathrm{T}} u_1^s + \|u_2^x\|^2 + \|u_1^s\|^2 + \|r_2\|^2 + \|u_2^x\|^2 - 2r_2^{\mathrm{T}} u_2^x$$
$$= \|r\|^2 + 2\|u_2^x\|^2 + 2\|u_1^s\|^2 - 2r_1^{\mathrm{T}} u_1^s - 2r_2^{\mathrm{T}} u_2^x.$$

Further by the Cauchy-Schwartz inequality, the inequality $2ab \le a^2 + b^2$ and the properties of the orthogonal projection, we obtain

$$\|d_x^f\|^2 + \|d_s^f\|^2 \le \|r\|^2 + 2\|u_2^x\|^2 + 2\|u_1^s\|^2 + 2\|r_1\|\|u_1^s\| + 2\|r_2\|\|u_2^x\|$$
$$\le \|r\|^2 + 2\|u_2^x\|^2 + 2\|u_1^s\|^2 + \|r_1\|^2 + \|u_1^s\|^2 + \|r_2\|^2 + \|u_2^x\|^2$$
$$\le 2\|r\|^2 + 3\left(\|u^x\|^2 + \|u^s\|^2\right). \tag{23}$$

Since v and $v^{-1} - v$ are orthogonal and since $\|v\|^2 = n$, we have

$$\|r\|^2 = \left\|(1-\theta)v^{-1} - v\right\|^2 = \left\|(1-\theta)(v^{-1} - v) - \theta v\right\|^2$$
$$= (1-\theta)^2 \left\|v^{-1} - v\right\|^2 + \theta^2 \|v\|^2 = 4(1-\theta)^2 \delta^2 + \theta^2 n. \tag{24}$$

Due to (22), we have

$$\|u^x\|^2 + \|u^s\|^2 = \frac{\theta^2 v^2}{\mu} \left(\|D^{-1}(\bar{x} - x^0)\|^2 + \|D(\bar{s} - s^0)\|^2\right). \tag{25}$$

To proceed, we have to specify our initial iterate (x^0, y^0, s^0). We assume that $\zeta > 0$ is such that

$$\|x^* + s^*\|_\infty \leq \zeta, \qquad (26)$$

for some optimal solutions x^* of (P) and (y^*, s^*) of (D); as usual, we start the algorithm with

$$x^0 = s^0 = \zeta e, \qquad y^0 = 0, \qquad \mu^0 = \zeta^2. \qquad (27)$$

We are still free to choose $\bar{x}$ and $\bar{s}$, subject to the constraints $A\bar{x} = b$ and $A^T\bar{y} + \bar{s} = c$. Taking $\bar{x} = x^*$ and $\bar{s} = s^*$, then the entries of the vectors $x^0 - \bar{x}$ and $s^0 - \bar{s}$ satisfy

$$0 \leq x^0 - \bar{x} \leq \zeta e, \qquad 0 \leq s^0 - \bar{s} \leq \zeta e.$$

Thus, it follows that

$$
\begin{aligned}
&\|D^{-1}(\bar{x} - x^0)\|^2 + \|D(\bar{s} - s^0)\|^2 \\
&\quad \leq \zeta^2 \left(\|De\|^2 + \|D^{-1}e\|^2 \right) = \zeta^2 e^T \left(\frac{x}{s} + \frac{s}{x} \right) \\
&\quad = \zeta^2 e^T \left(\frac{x^2 + s^2}{xs} \right) \leq \frac{\zeta^2 e^T (x^2 + s^2)}{\min_i |x_i s_i|} \leq \frac{\zeta^2 \left[e^T (x + s) \right]^2}{\mu \min_i v_i^2}. \qquad (28)
\end{aligned}
$$

Summarizing, while using (23), (24), (25), (28) and $\mu = \nu\zeta^2$, we obtain

$$\|d_x^f\|^2 + \|d_s^f\|^2 \leq 8(1 - \theta)^2\delta^2 + 2\theta^2 n + \frac{3\theta^2 \left[e^T(x+s) \right]^2}{\zeta^2 \min_i v_i^2}. \qquad (29)$$

Recall that x is feasible for (P_ν) and (y, s) is feasible for (D_ν), with $x^T s = n\mu$ and, moreover $\delta(x, s; \mu) \leq \tau$; i.e., these iterates are close to the μ-centers of (P_ν) and (D_ν). Based on this information, we present the following two lemmas to estimate an upper bound for $e^T(x + s)$ and a lower bound for $\min_i v_i^2$.

Lemma 4.3 *Let x and (y, s) be feasible for the perturbed problems (P_ν) and (D_ν), respectively, with $x^T s = n\mu$, ζ as defined in (26), and (x^0, y^0, s^0) as in (27). We then have*

$$e^T(x + s) \leq 2n\zeta.$$

Proof Let (x^*, y^*, s^*) be optimal solutions satisfying (26). Then, from the feasibility conditions (5) and (6) of the perturbed problems (P_ν) and (D_ν), it is easily seen that

$$
\begin{aligned}
A\left[x - \nu x^0 - (1 - \nu)x^* \right] &= 0, \\
A^T\left[y - \nu y^0 - (1 - \nu)y^* \right] + \left[s - \nu s^0 - (1 - \nu)s^* \right] &= 0.
\end{aligned}
$$

This implies that

$$\left[x - \nu x^0 - (1 - \nu)x^* \right]^T \left[s - \nu s^0 - (1 - \nu)s^* \right] = 0.$$

By expanding the above equality and using the fact that $(x^*)^\mathrm{T} s^* = 0$, we obtain

$$v\big((x^0)^\mathrm{T} s + (s^0)^\mathrm{T} x\big) = x^\mathrm{T} s + v^2 (x^0)^\mathrm{T} s^0 - (1-v)(x^\mathrm{T} s^* + s^\mathrm{T} x^*)$$
$$+ v(1-v)\big((x^0)^\mathrm{T} s^* + (s^0)^\mathrm{T} x^*\big).$$

Since (x^0, y^0, s^0) are as defined in (27), we have

$$(x^0)^\mathrm{T} s + (s^0)^\mathrm{T} x = \zeta e^\mathrm{T}(x+s),$$
$$(x^0)^\mathrm{T} s^0 = n\zeta^2,$$
$$(x^0)^\mathrm{T} s^* + (s^0)^\mathrm{T} x^* = \zeta e^\mathrm{T}(x^* + s^*).$$

Due to (26), we have

$$e^\mathrm{T}(x^* + s^*) \le n\zeta.$$

Furthermore,

$$x^\mathrm{T} s = n\mu = v\zeta^2 n \quad \text{and} \quad x^\mathrm{T} s^* + s^\mathrm{T} x^* \ge 0.$$

Substitution of these relations gives

$$v\zeta e^\mathrm{T}(x+s) \le v\zeta^2 n + nv^2\zeta^2 + v(1-v)n\zeta^2 = 2nv\zeta^2.$$

This implies the lemma.   □

Lemma 4.4 (See [11], Theorem II.62) *Let $\rho(\delta) = \delta + \sqrt{1+\delta^2}$. Then,*

$$\frac{1}{\rho(\delta)} \le v_i \le \rho(\delta), \quad 1 \le i \le n.$$

Substituting the results of the above two lemmas into (29), we obtain

$$\|d_x^f\|^2 + \|d_s^f\|^2 \le 8(1-\theta)^2\delta^2 + 2\theta^2 n + 12\theta^2 n^2 \rho(\delta)^2$$
$$\le 8\delta^2 + 2\theta^2 n + 12\theta^2 n^2 \rho(\delta)^2. \tag{30}$$

4.4 Value for θ

We have found so far that $\delta(v^f) \le 1/\sqrt[4]{2}$ certainly holds if the inequality (15) is satisfied. Then, by (30), inequality (15) holds if

$$8\delta^2 + 2\theta^2 n + 12\theta^2 n^2 \rho(\delta)^2 \le 2\sqrt{2}\left(\sqrt{1+\sqrt{2}} - 1\right)(1-\theta).$$

We set $\tau = 1/16$. Obviously, the left-hand side of the above inequality is increasing in δ, due to the definition

$$\rho(\delta) = \delta + \sqrt{1+\delta^2}.$$

Using this, one may easily verify that, if

$$\tau = \frac{1}{16}, \qquad \theta = \frac{1}{4n}, \tag{31}$$

then the above inequality is satisfied.

Then, according to (10), with τ as given, after the feasibility step at most

$$2 + \left\lceil \log_2 \left(\log_2 \frac{1}{\tau} \right) \right\rceil = 4 \tag{32}$$

centering steps suffice to get an iterate (x^+, y^+, s^+) that satisfies $\delta(x^+, s^+; \mu^+) \leq \tau$.

4.5 Complexity Analysis

In the previous sections, we have found that, if at the start of an iteration the iterate satisfies $\delta(x, s; \mu) \leq \tau$, with $\tau = 1/16$, then after the feasibility step, with θ as defined in (31), the iterate satisfies $\delta(x^f, s^f; \mu^+) \leq 1/\sqrt[4]{2}$.

According to (32), at most 4 centering steps suffice to get the iterate (x^+, y^+, s^+) that satisfies $\delta(x^+, s^+; \mu^+) \leq \tau$ again. So, each main iteration consists of at most 5 so-called inner iterations, in each of which we need to compute a search direction (for either a feasibility step or a centering step).

It has become customary to measure the complexity of an IPM by the required number of inner iterations. In each main iteration, both the duality gap and the norms of the residual vectors are reduced by the factor $1 - \theta$. Hence, using $(x^0)^{\mathsf{T}} s^0 = n\zeta^2$, the total number of main iterations is bounded above by

$$\frac{1}{\theta} \log \frac{\max\{n\zeta^2, \|r_b^0\|, \|r_c^0\|\}}{\varepsilon}.$$

Taking the value of θ as in (31), the total number of inner iterations is bounded above by

$$20n \log \frac{\max\{n\zeta^2, \|r_b^0\|, \|r_c^0\|\}}{\varepsilon}.$$

Thus, we may state without further proof the main result of the paper.

Theorem 4.1 *If* (P) *and* (D) *are feasible and* $\zeta > 0$ *is such that* $\|x^* + s^*\|_\infty \leq \zeta$ *for some optimal solutions* x^* *of* (P) *and* (y^*, s^*) *of* (D), *then after at most*

$$20n \log \frac{\max\{n\zeta^2, \|r_b^0\|, \|r_c^0\|\}}{\varepsilon}$$

inner iterations, the algorithm finds an ε*-optimal solution of* (P) *and* (D).

Note that this bound is slightly better than that in [1, Theorem 4.8].

Remark 4.1 The above iteration bound is derived under the assumption that there exists some optimal solutions of (P) and (D) with $\|x^* + s^*\|_\infty \leq \zeta$. One might ask what happens if this condition is not satisfied. In that case, during the course of the algorithm, it may happen that, after some main steps, the proximity measure δ (after the feasibility step) exceeds $1/\sqrt[4]{2}$, because otherwise there is no reason why the algorithm would not generate an ε-optimal solution. So, if this happens, it tells us that the problems (P) and (D) do not have any optimal solution that satisfies $\|x^* + s^*\|_\infty \leq \zeta$. Recall that our starting point is defined in (27), which depends on ζ.

5 Concluding Remarks

We presented an improved full-Newton step IIPM for LO, which is motivated by [1]. The new method can be viewed as a homotopy method, which turns out to have many nice properties. First, as the name suggests, it uses full steps (instead of damped steps), so there is no need to calculate the step length (which is always 1). Second, the iterates always lie in the quadratic convergence neighborhood with respect to some perturbed problems, which makes the algorithm more stable. Third, during the solution process, both "feasibility" and "optimality" are improved at the same rate, which is also credited by Potra [6]. Finally, the iteration bound coincides with the currently best-known bound for IIPMs.

Each main step of our method consists of a feasibility step and at most 4 centering steps. The new feasibility step is more natural than in [1, 10], as it targets the μ^+-center with respect to the next pair of perturbed problems. For the centering steps, a sharper quadratic convergence result is used, which results in a wider neighborhood for the feasibility steps. Moreover, the analysis is much simplified and the iteration bound is slightly better. For our iteration bound, we mention that it does not depend on the big O notation.

Our new method admits the best known iteration bound, but from a practical perspective a severe shortcoming is its worst-case-oriented nature: it will always perform according to its worst-case theoretical complexity bound. A topic of further research is the use of *adaptive updates*, as described in [11]. This may enhance significantly the practical performance of the algorithm; thereby, we may take profit of the wider neighborhood for the feasibility steps that is proposed in this paper. Another topic for further research is the generalization to other classes of optimization problems, as second-order cone optimization, semidefinite optimization, and also P_*-matrix LCP.

References

1. Roos, C.: A full-Newton step $O(n)$ infeasible interior-point algorithm for linear optimization. SIAM J. Optim. **16**(4), 1110–1136 (2006) (electronic)
2. Ji, J., Potra, F.A., Sheng, R.: A predictor-corrector method for solving the $P * (\kappa)$-matrix LCP from infeasible starting points. Optim. Methods Softw. **6**(2), 109–126 (1996)

3. Kojima, M., Megiddo, N., Mizuno, S.: A primal-dual infeasible-interior-point algorithm for linear programming. Math. Program. Ser. A **61**(3), 263–280 (1993)
4. Lustig, I.J.: Feasibility issues in a primal-dual interior-point method for linear programming. Math. Program. Ser. A **49**(2), 145–162 (1990/91)
5. Mizuno, S.: Polynomiality of infeasible-interior-point algorithms for linear programming. Math. Program. Ser. A **67**(1), 109–119 (1994)
6. Potra, F.A.: An infeasible-interior-point predictor-corrector algorithm for linear programming. SIAM J. Optim. **6**(1), 19–32 (1996)
7. Potra, F.A., Sheng, R.: A large-step infeasible-interior-point method for the P_*-matrix LCP. SIAM J. Optim. **7**(2), 318–335 (1997)
8. Potra, F.A., Sheng, R.: Predictor-corrector algorithm for solving $P_*(\kappa)$-matrix LCP from arbitrary positive starting points. Math. Program. Ser. B **76**(1), 223–244 (1997). Interior point methods in theory and practice (Iowa City, IA, 1994)
9. Zhang, Y.: On the convergence of a class of infeasible interior-point methods for the horizontal linear complementarity problem. SIAM J. Optim. **4**(1), 208–227 (1994)
10. Mansouri, H., Roos, C.: Simplified $O(nL)$ infeasible interior-point algorithm for linear optimization using full-Newton steps. Optim. Methods Softw. **22**(3), 519–530 (2007)
11. Roos, C., Terlaky, T., Vial, J.P.: Interior Point Methods for Linear Optimization. Springer, New York (2006) (2nd edn. of Theory and Algorithms for Linear Optimization. Wiley, Chichester (1997))
12. Ye, Y.: Interior Point Algorithms. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley, New York (1997)