McMaster University

Advanced Optimization Laboratory

Title: Postponing the Choice of the Barrier Parameter in Mehrotra-Type Predictor-Corrector Algorithms

Authors:

Maziar Salahi and Tamás Terlaky

AdvOl-Report No. 2005/16

September 2005, Hamilton, Ontario, Canada

Postponing the Choice of the Barrier Parameter in Mehrotra-Type Predictor-Corrector Algorithms

Maziar Salahi*‡ and Tamás Terlaky†‡

September 1, 2005

Abstract

In [8] the authors considered a variant of Mehrotra's predictor-corrector algorithm that has been widely used in several IPMs based optimization packages. By an example they showed that this variant might make very small steps in order to keep the iterate in a certain neighborhood of the central path, that itself implies the inefficiency of the algorithm. This observation motivated them to incorporate a safeguard in their algorithmic scheme that gives a warranted lower bound for the maximum step size at each iteration. In this paper we propose a different approach that enables us to have control on the iterates. Our new approach is based on postponing the choice of the barrier parameter and does not require any safeguard strategy like the one in [8]. To do so, first we fix a step size in the corrector step, then by solving a one dimensional optimization problem we estimate the barrier parameter. Finally, using the estimated barrier parameter it computes the maximum step size that can be taken and makes the next iterate. We proved that for the feasible case in the worst case, our new algorithm stops after at most $O\left(n^2 \log \frac{n}{\epsilon}\right)$ iterations without any safeguard strategy. We further modified the proposed algorithm by slightly modifying the Newton system that has to be solved in the corrector step. This modified variant enjoys better iteration complexity i.e., $O\left(n \log \frac{n}{\epsilon}\right)$. The superlinear convergence of both algorithms are established. Finally, we report some limited encouraging numerical results.

Keywords: Linear Optimization, Predictor-Corrector Method, Mehrotra-Type Algorithm, Polynomial Complexity, Superlinear Convergence.

AMS Subject Classification: 90C05, 90C51

---

*Corresponding author: Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada, L8S 4K1, salahim@math.mcmaster.ca.

†Department of Computing and Software, McMaster University, Hamilton, Ontario, Canada, L8S 4K1, terlaky@mcmaster.ca.

‡The research is supported by the NSERC Discovery Grant DG:5-48923, the Canada Research Chair program and by MITACS.

1 Introduction

In this paper we consider primal-dual *interior-point methods* (IPMs) for solving the following standard *Linear Optimization* (LO) problem:

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

where $A \in R^{m\times n}$ satisfies $\operatorname{rank}(A) = m,\ b \in R^m,\ c \in R^n$, and its dual problem

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

Before getting to the main theme of the paper we first give a brief introduction to IPMs. We may assume without loss of generality [7] that both (P) and (D) satisfy the interior point condition (IPC), i.e., there exists an (x^0, y^0, s^0) such that

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

which is essential in order to be able to apply IPMs. Finding optimal solutions of (P) and (D) is equivalent to solving the following system:

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

where xs denotes the componentwise product of the vectors x and s. The basic idea of primal-dual IPMs is to replace the third equation in (1) by the parameterized equation $xs = \mu e$, where e is the all one vector. This leads to the following system:

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

If the IPC holds, then for each $\mu > 0$, system (2) has a unique solution. This solution, denoted by $(x(\mu), y(\mu), s(\mu))$, is called the μ-center of the primal-dual pair (P) and (D). The set of μ-centers for all $\mu > 0$ gives *the central path* of (P) and (D) [3, 10]. It has been shown that the limit of the central path (as μ goes to zero) exists. Because the limit point satisfies the complementarity condition, it naturally yields optimal solutions for both (P) and (D), respectively [7].

Applying Newton's method to (2) from a given interior point, gives the following linear system of equations

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

where $(\Delta x, \Delta y, \Delta s)$ give the Newton step.

Predictor-corrector algorithms use (3) with different values of μ in the predictor and corrector steps. In the predictor step they use the so-called affine scaling step with $\mu = 0$, and in the corrector step e.g., the Mzuno-Tod-Ye's algorithm [5] uses $\mu = \mu_g := \frac{x^T s}{n}$.

It is worth mentioning that various predictor-corrector algorithms have been proposed in the last decades, see e.g., [1, 2, 5, 7, 9] for more details.

Most IPMs based software packages [13, 14] are using a variant of the Mehortra' algorithm [4]. The authors of [8] realized that in the worst case it may suffer from some drawbacks that motivated them to slightly modify the algorithm in order to guarantee both theoretical and practical efficiency. In what follows we briefly review this variant of the original algorithm.

In the predictor step it solves the following system of equation, which is defined from (3) by letting $\mu = 0$, and it is called the *affine scaling* system of equations:

$$\begin{aligned} A\Delta x^a &= 0, \\ A^T \Delta y^a + \Delta s^a &= 0, \\ s\Delta x^a + x\Delta s^a &= -xs. \end{aligned} \tag{4}$$

Then one computes the maximum feasible step size α_{a} in this direction such that

$$(x + \alpha_{\mathrm{a}}\Delta x^a, s + \alpha_{\mathrm{a}}\Delta s^a) \geq 0.$$

However, the algorithm does not make such a step. It uses the information from the predictor step to compute the centering direction as follows:

$$\begin{aligned} A\Delta x &= 0, \\ A^T \Delta y + \Delta s &= 0, \\ s\Delta x + x\Delta s &= \mu e - xs - \Delta x^a \Delta s^a, \end{aligned} \tag{5}$$

where μ is defined adaptively as

$$\mu_t = \left(\frac{g_a}{g}\right)^2 \frac{g_a}{n},$$

where $g_a = (x + \alpha_{\mathrm{a}}\Delta x^a)^T(s + \alpha_{\mathrm{a}}\Delta s^a)$ and $g = x^T s$. Since $(\Delta x^a)^T \Delta s^a = 0$, the previous relation can be simplified to

$$\mu_t = (1 - \alpha_{\mathrm{a}})^3 \mu_g. \tag{6}$$

As it has been discussed in [8], this variant might make very small steps in order to keep the iterates in a certain neighborhood of the central path. In [8] the authors proposed a new variant that performs excellent in practice, while it enjoys polynomial worst case iteration complexity as well. However, in the worst case this variant requires an extra backsolve at each iteration. In this paper we propose a new approach that is different than the existing ideas in the literature of IPMs. In this approach we fix a required step size for the corrector step, then by solving a one dimensional optimization problem (see Section 2) we estimate the best possible value of the parameter μ, that replaces Mehrotra's μ in the centering step computed by (6). As we will see in the sequel this strategy is applicable without incorporating a safeguard strategy like the one in [8]. The rest of the paper is organized as it follows.

In Section 2, we describe our algorithmic scheme and give the worst case iteration complexity analysis of our new algorithm in details. A slightly modified variant of the propose algorithm is presented in Section 3 with its worst case iteration complexity analysis. The superlinear convergence of both algorithm are discussed in Section 4. Some preliminary numerical results are given in Section 5. Finally the paper is concluded by some remarks in Section 6. For ease of understanding we moved some technical lemmas to the Appendix. It is worth mentioning that for self containdness we give the detailed proofs for some lemmas which appeared in [8].

Some Notations:

- X and S denote $n \times n$ matrices whose diagonal elements are the elements of x and s, respectively.
- s^{-1} denotes the componentwise inverse of the vector s.
- $\mathcal{I} = \{1, \cdots, n\}$.
- $\mathcal{I}_+ = \{i \in \mathcal{I} \mid \Delta x_i^a \Delta s_i^a > 0\}$.
- $\mathcal{F} = \{(x, y, s) \mid Ax = b,\ A^T y + s = c,\ x \geq 0,\ s \geq 0\}$.
- $\mathcal{F}^0 = \{(x, y, s) \mid Ax = b,\ A^T y + s = c,\ x > 0,\ s > 0\}$.

2 Algorithmic Scheme

This section is devoted to our new algorithmic scheme and its worst case iteration complexity analysis.

In the rest of the paper we deal with algorithms that operate in the so called negative infinity neighborhood, defined by

$$\mathcal{N}_\infty^-(\gamma) := \{(x, y, s) \in \mathcal{F}^0 :\ x_i s_i \geq \gamma \mu_g\ \forall i \in \mathcal{I}\}, \tag{7}$$

where $\gamma \in (0, 1)$ is a constant independent of n. This neighborhood is the one which is widely used in the implementations too.

The following lemma will be used later in this section.

Lemma 2.1 (Lemma 3.2 in [8]) *Suppose that the current iterate* $(x, y, s) \in \mathcal{N}_\infty^-(\gamma)$ *and* $(\Delta x, \Delta y, \Delta s)$ *be the solution of (5). Then we have*

$$\|\Delta x \Delta s\| \le 2^{\frac{-3}{2}} \left(\frac{n\mu^2}{\gamma \mu_g} + \frac{17 n \mu_g}{16} + \frac{n^2 \mu_g}{16\gamma} - 2n\mu + \frac{n\mu}{2\gamma} \right).$$

Proof: If we multiply the third equation of (5) by $(XS)^{-\frac{1}{2}}$, then by Lemma 5.3 of [11] we have

$$\begin{aligned}
\|\Delta x \Delta s\| &\le 2^{\frac{-3}{2}} \left\| \mu(XSe)^{-\frac{1}{2}} - (XSe)^{\frac{1}{2}} - (XS)^{-\frac{1}{2}} \Delta x^a \Delta s^a \right\|^2 \\
&= 2^{\frac{-3}{2}} \left(\mu^2 \sum_{i \in \mathcal{I}} \frac{1}{x_i s_i} + \sum_{i \in \mathcal{I}} x_i s_i + \sum_{i \in \mathcal{I}} \frac{(\Delta x_i^a \Delta s_i^a)^2}{x_i s_i} - 2n\mu - 2\mu \sum_{i \in \mathcal{I}} \frac{\Delta x_i^a \Delta s_i^a}{x_i s_i} \right) \\
&\le 2^{\frac{-3}{2}} \left(\frac{n\mu^2}{\gamma \mu_g} + n\mu_g + \frac{n\mu_g}{16} + \frac{n^2 \mu_g}{16\gamma} - 2n\mu + \frac{n\mu}{2\gamma} \right),
\end{aligned}$$

where the last inequality follows from Lemma 7.2 and the fact that the previous iterate is in $\mathcal{N}_\infty^-(\gamma)$. □

The predictor step of the algorithm is the primal-dual affine scaling step, analogous to the predictor step of the algorithm presented in the introduction, except that here we do not compute the maximum step size in this direction. In the corrector step, in contrast to the algorithm in the introduction, first we fix a target step size, then by solving a one dimensional optimization problem we find an appropriate μ that allows the iterates to stay in $\mathcal{N}_\infty^-(\gamma)$ while sufficiently reducing the complementarity gap at the same time.

Now let us recast the system of equations that have to be solved in the corrector step of our algorithm:

$$\begin{aligned}
A\Delta x &= 0, \\
A^T \Delta y + \Delta s &= 0, \\
s\Delta x + x\Delta s &= \mu e - xs - \Delta x^a \Delta s^a,
\end{aligned} \tag{8}$$

where μ is now a variable, rather than being a constant. The following lemma gives information about the complementarity gap after one step. The proof is a direct consequence of the third equation of (8).

Lemma 2.2 *Let* $(\Delta x, \Delta y, \Delta s)$ *be the solution of (8). Then for the complementarity gap after one step with step size* α *one has*

$$(x + \alpha \Delta x)^T (s + \alpha \Delta s) = (1-\alpha) x^T s + \alpha n \mu = \left(1 - \alpha + \alpha \frac{\mu}{\mu_g} \right) x^T s.$$

Now, let us start to derive the explicit solution of (8) as a function of μ. From the third equation of (8) one has

$$\Delta s = X^{-1}(\mu e - xs - \Delta x^a \Delta s^a - s\Delta x).$$

Then, from the second equation of (8) one has

$$\Delta y = -(ADA^T)^{-1}AS^{-1}(\mu e - xs - \Delta x^a \Delta s^a - s\Delta x),$$

where $D = \text{diag}\,(xs^{-1})$. Let us define $P = (ADA^T)^{-1}AS^{-1}$, then

$$\Delta y = -\mu Pe - PXs - P\text{diag}\,(\Delta x^a)\Delta s^a.$$

Using again the second equation of (8) one has

$$\Delta s = \mu A^T Pe + A^T PXs + A^T P\text{diag}\,(\Delta x^a)\Delta s^a.$$

Finally, from the third equation of (8) one can get

$$\Delta x = \mu S^{-1}(e - XA^T Pe) - x - S^{-1}\left(\text{diag}\,(\Delta x^a)\Delta s^a - XA^T PXs - XA^T P\text{diag}\,(\Delta x^a)\Delta s^a\right).$$

Therefore, one has the following representation for $(\Delta x, \Delta y, \Delta s)$ as a function of μ:

$$\begin{aligned} \Delta x &= \mu p_x^1 + p_x^2, \\ \Delta y &= \mu p_y^1 + p_y^2, \\ \Delta s &= \mu p_s^1 + p_s^2, \end{aligned} \tag{9}$$

where

$$\begin{aligned} p_x^1 &= S^{-1}(e - XA^T Pe), \\ p_x^2 &= -x - S^{-1}\left(\text{diag}\,(\Delta x^a)\Delta s^a - XA^T PXs - XA^T P\text{diag}\,(\Delta x^a)\Delta s^a\right), \\ p_y^1 &= Pe, \\ p_y^2 &= -PXs - P\text{diag}\,(\Delta x^a)\Delta s^a, \\ p_s^1 &= A^T Pe, \\ p_s^2 &= A^T PXs + A^T P\text{diag}\,(\Delta x^a)\Delta s^a. \end{aligned}$$

Now the goal is to solve the following one dimensional optimization problem in order to estimate the smallest value of parameter μ for which the already chosen target[1] step size α_t gives an iterate in $\mathcal{N}_\infty^-(\gamma)$.

$$\begin{aligned} \min \;\; & (1-\alpha_t)g + \alpha_t n \mu_t \\ & (x(\alpha_t), s(\alpha_t)) \in \mathcal{N}_\infty^-(\gamma). \end{aligned} \tag{10}$$

We will show in the sequel that problem (10) is solvable for properly chosen α values and we give an upper bound for the optimal value of (10). To do so, after each corrector step one has

$$\begin{aligned} x(\alpha_t)s(\alpha_t) &= (x + \alpha_t \Delta x)(s + \alpha_t \Delta s) \\ &= (1-\alpha_t)xs + \alpha_t \mu_t e - \alpha_t \Delta x^a \Delta s^a + \alpha_t^2 \Delta x \Delta s \\ &= (1-\alpha_t)xs + \alpha_t \mu_t e - \alpha_t \Delta x^a \Delta s^a + \alpha_t^2 \mu_t^2 p_x^1 p_s^1 + \alpha_t^2 \mu_t (p_x^1 p_s^2 + p_x^2 p_s^1) + \alpha_t^2 p_x^2 p_s^2 \\ &= \alpha_t^2 p_x^1 p_s^1 \mu_t^2 + (\alpha_t e + \alpha_t^2 p_x^1 p_s^2 + \alpha_t^2 p_x^2 p_s^1)\mu_t + (1-\alpha_t)xs + \alpha_t^2 p_x^2 p_s^2 \end{aligned}$$

[1]The index 't' for α_t and μ_t denotes the target step size and barrier parameter.

and

$$\mu_g(\alpha_t) = (1-\alpha_t)\mu_g + \alpha_t\mu_t. \tag{11}$$

Therefore, the constraint in (10) is equivalent to

$$\alpha_t^2 p_x^1 p_s^1 \mu_t^2 + (\alpha_t e + \alpha_t^2 p_x^1 p_s^2 + \alpha_t^2 p_x^2 p_s^1 - \gamma\alpha_t e)\mu_t + (1-\alpha_t)xs + \alpha_t^2 p_x^2 p_s^2 \\ -\gamma(1-\alpha_t)\mu_g e \geq 0, \tag{12}$$

namely n quadratic inequalities with μ_t as a variable. It is worth mentioning that in advance we do not have any information about the convexity or concavity of these polynomials.

In what follows we give an upper bound for the optimal solution μ_t of problem (10) by fixing the target step size α_t. This estimation enables us to prove the polynomial iteration complexity of our new algorithm.

Let us fix $\alpha_t = \frac{\gamma^2}{n^2}$, then the ith constraint of (12) holds if

$$x_i(\alpha)s_i(\alpha) \geq \gamma\mu_g(\alpha).$$

After expanding this inequality using (8) and (11) one has

$$(1-\alpha)x_i s_i + \alpha(1-\gamma)\mu - \alpha\Delta x_i^a \Delta s_i^a + \alpha^2 \Delta x_i \Delta s_i \geq \gamma(1-\alpha)\mu_g. \tag{13}$$

We get the worst bound if $\Delta x_i^a \Delta s_i^a > 0$ and $\Delta x_i \Delta s_i < 0$. By Lemma 7.1 we also know that $\Delta x_i^a \Delta s_i^a \leq \frac{x_i s_i}{4}$. Therefore, using this inequality and the fact that the previous iterate is in $\mathcal{N}_\infty^-(\gamma)$, inequality (13) holds if

$$-\frac{\gamma\mu_g}{4} + (1-\gamma)\mu - \frac{\gamma^2 |\Delta x_i \Delta s_i|}{n^2} \geq 0.$$

Using Lemma 2.1, the previous inequality is true if

$$-\frac{\gamma\mu_g}{4} + (1-\gamma)\mu - \frac{2^{\frac{-3}{2}}\gamma\mu^2}{\mu_g n} - \frac{2^{\frac{-3}{2}}17\gamma^2\mu_g}{16n} - \frac{2^{\frac{-3}{2}}\gamma\mu_g}{16} + \frac{2^{\frac{-1}{2}}\gamma^2\mu}{n} - \frac{2^{\frac{-3}{2}}\gamma\mu}{2n} \geq 0.$$

After reordering and simplifying one has

$$-\frac{\gamma}{2\sqrt{2}\mu_g n}\mu^2 + \left(\frac{\gamma^2}{\sqrt{2}n} - \frac{\gamma}{4\sqrt{2}n} + 1 - \gamma\right)\mu - \frac{(8\sqrt{2}+1)\gamma\mu_g}{32\sqrt{2}} - \frac{17\gamma^2\mu_g}{32\sqrt{2}n} \geq 0. \tag{14}$$

The following lemma gives the interval for which inequality (14) holds.

Lemma 2.3 *For $\gamma \in (0, 0.68)$ and $n \geq 2$ the step size $\alpha_t = \frac{\gamma^2}{n^2}$ is feasible for any*

$$\mu_t \in [\beta_1\mu_g \ , \ \beta_2\mu_g],$$

where

$$\beta_1 = \frac{-4\sqrt{2}\gamma n + 4\sqrt{2}n + 4\gamma^2 - \gamma - \sqrt{\Delta}}{4\gamma}$$

and

$$\beta_2 = \frac{-4\sqrt{2}\gamma n + 4\sqrt{2}n + 4\gamma^2 - \gamma + \sqrt{\Delta}}{4\gamma}$$

and

$$\Delta = 32\gamma^2 n^2 - 64\gamma n^2 - 32\sqrt{2}\gamma^3 n + (32\sqrt{2} - 1)\gamma^2 n + 32n^2 - 8\sqrt{2}\gamma n + 16\gamma^4 - 25\gamma^3 + \gamma^2.$$

Moreover $0 < \beta_1 < \frac{7}{8}$.

Proof: For the worst case analysis it suffices to find the interval for which inequality (14) holds. By simple calculus one may verify that $\beta_1\mu_g$ and $\beta_2\mu_g$ are two real roots of the second order polynomial in (14) if and only if $\Delta \geq 0$, which holds for all $\gamma \in (0, 0.73)$ and $n \geq 2$. Finally we have to show that $0 < \beta_1 < \frac{7}{8}$. To this end, using the definition of β_1, $\beta_1 > 0$ is equivalent to $(8\sqrt{2} + 1)\gamma^2 n + 17\gamma^3 > 0$, which is definitely true. Analogously, $\beta_1 < \frac{7}{8}$ is equivalent to $(144\sqrt{2} + 4)\gamma^2 n - 112\sqrt{2}\gamma n - 44\gamma^3 + 77\gamma^2 < 0$. This inequality holds for all $\gamma \in (0, 0.68)$ and $n \geq 2$ that completes the proof of the lemma. □

Corollary 2.4 *The optimal value* μ_t *of problem (10) satisfies* $\mu_t \leq \beta_1\mu_g$.

Remark 2.5 *Because* $\mu_t \leq \beta_1\mu_g \leq \frac{7}{8}\mu_g$, *our new algorithm features a large update of parameter* μ *at each step.*

Now, let us outline the scheme of our new algorithm.

Algorithm 1

Input:

A neighborhood parameter $\gamma \in (0, 0.68)$;
$\alpha_t \in (0, 1]$;
an accuracy parameter $\epsilon > 0$;
$(x^0, y^0, s^0) \in \mathcal{N}_\infty^-(\gamma)$.

begin
while $x^T s > \epsilon$ **do**
begin
Predictor Step
Solve (4) for $(\Delta x^a, \Delta y^a, \Delta s^a)$.
end
begin
Corrector step
Solve (10) with α_t to determine μ_t;
Compute $(\Delta x, \Delta y, \Delta s)$ by (9);
Compute the maximum step size α_{c} such that $(x(\alpha_{\mathrm{c}}), y(\alpha_{\mathrm{c}}), s(\alpha_{\mathrm{c}})) \in \mathcal{N}_\infty^-(\gamma)$;
Set $(x(\alpha_{\mathrm{c}}), y(\alpha_{\mathrm{c}}), s(\alpha_{\mathrm{c}})) = (x + \alpha_{\mathrm{c}}\Delta x, y + \alpha_{\mathrm{c}}\Delta y, s + \alpha_{\mathrm{c}}\Delta s)$.
end
end

The following theorem gives an upper bound for the maximum number of iterations of Algorithm 1.

Theorem 2.6 *Let $\alpha_t = \frac{\gamma^2}{n^2}$, then Algorithm 1 stops after at most*

$$O\left(n^2 \log \frac{n}{\epsilon}\right)$$

iterations with a solution (x, y, s) that satisfies $x^T s \le \epsilon$.

Proof: After each iteration one has

$$x(\alpha_{\mathrm{c}})^T s(\alpha_{\mathrm{c}}) = \left(1 - \alpha_{\mathrm{c}} + \alpha_{\mathrm{c}} \frac{\mu_t}{\mu_g}\right) x^T s.$$

In the corrector step of the algorithm by construction $\alpha_{\mathrm{c}} \ge \frac{\gamma^2}{n^2}$ and $\mu_t \le \mu_g$. This implies

$$x(\alpha_c)^T s(\alpha_c) \le \left(1 - \frac{\gamma^2}{n^2} + \frac{\gamma^2 \mu_t}{\mu_g n^2}\right) x^T s.$$

By Lemma 2.3 we also have $\mu_t \leq \beta_1 \mu_g$, where $0 < \beta_1 < \frac{7}{8}$. Therefore

$$x(\alpha_c)^T s(\alpha_c) \leq \left(1 - \frac{\gamma^2}{8n^2}\right) x^T s$$

that completes the proof by Theorem 3.2 of [11]. □

Remark 2.7 *We see that the iteration complexity of Algorithm 1 in the worst case is in the same order as the complexity bound for the one presented in [8], but Algorithm 1 is not using any safeguard strategy while the one in [8] does.*

Remark 2.8 *By changing the right hand side of (4) and (8) one can propose an infeasible variant of the proposed algorithm. The complexity analysis of this algorithm is left for the interested reader.*

3 A Modified Version of Algorithm 1

In this section we slightly modify the predictor and corrector steps of Algorithm 1 that improves the iteration complexity significantly. The motivation for this modification is the bound for $\|\Delta x \Delta s\|$ in Lemma 2.1. In the proof of Lemma 2.1 one can see that the bound for the negative components of $\Delta x^a \Delta s^a$ bring a factor of n^2 that itself leads to an upper bound of order of n^2. To avoid this effect we need to modify the Newton system that is solved in the corrector step of the algorithm. This itself 'unlike Algorithm 1' requires to compute the maximum step size in the predictor step. Therefore, let us first derive a lower bound for the maximum step size in the predictor step.

Theorem 3.1 (Theorem 3.1 in [8]) *Suppose the current iterate* $(x, y, s) \in \mathcal{N}_\infty^-(\gamma)$ *and* $(\Delta x^a, \Delta y^a, \Delta s^a)$ *be the solution of (4). Then the maximum feasible step size* α_a*, that warrantees* $(x(\alpha_\mathrm{a}), s(\alpha_\mathrm{a})) \geq 0$, *satisfies*

$$\alpha_\mathrm{a} \geq \frac{2\sqrt{\gamma^2 + \gamma n} - 2\gamma}{n}. \tag{15}$$

Proof: For all $i \in \mathcal{I}_+$ there is no upper bound for a feasible step size. Since $(\Delta x^a)^T \Delta s^a = 0$, there is at least one $i \in \mathcal{I}_-$. Furthermore, $(x, s) \in \mathcal{N}_\infty^-(\gamma)$ implies by Lemma 7.2 that

$$x_i(\alpha) s_i(\alpha) = (1-\alpha) x_i s_i + \alpha^2 \Delta x^a{}_i \Delta s^a{}_i \geq \gamma \left(1 - \alpha - \frac{n\alpha^2}{4\gamma}\right) \mu_g. \tag{16}$$

Our aim is to prove that $x_i(\alpha) s_i(\alpha) \geq 0$. For this it suffices to prove that

$$\gamma \left(1 - \alpha - \frac{n\alpha^2}{4\gamma}\right) \mu_g \geq 0 \tag{17}$$

that is equivalent to

$$n\alpha^2 + 4\gamma\alpha - 4\gamma \leq 0.$$

The previous inequality holds when $\alpha \in \left[\frac{-2\sqrt{\gamma^2+\gamma n}-2\gamma}{n},\ \frac{2\sqrt{\gamma^2+\gamma n}-2\gamma}{n}\right]$ that completes the proof of the lemma. □

For the negative components of $\Delta x^a \Delta s^a$ using (16) and (17) one has

$$-\Delta x_i^a \Delta s_i^a \leq \frac{1}{\alpha_{\mathrm{a}}}\left(\frac{1}{\alpha_{\mathrm{a}}} - 1\right) x_i s_i \quad \forall i \in \mathcal{I}_-. \tag{18}$$

Therefore, if we use this upper bound, by slightly modifying the Newton system in the corrector step the bound in Lemma 2.1 can be improved which we prove in the sequel. Compared to the corrector step of (8) we damp the effect of the second order corrector term $\Delta x^a \Delta s^a$ by multiplying it with the maximum feasible step size of the affine scaling step. The new system is given by

$$\begin{aligned} A\Delta x &= 0, \\ A^T \Delta y + \Delta s &= 0, \\ s\Delta x + x\Delta s &= \mu e - xs - \alpha_{\mathrm{a}} \Delta x^a \Delta s^a. \end{aligned} \tag{19}$$

Analogous to the previous section, one can derive a representation of $(\Delta x, \Delta y, \Delta s)$ as a function of the barrier parameter μ. Therefore, problem (10) has to be solved in order to estimate the parameter μ for this modified version.

The following result is analogous to Lemma 2.1. One has to notice that the bound in the following lemma is much stronger than the one in Lemma 2.1.

Lemma 3.2 (Lemma 4.2 in [8]) *Suppose that the current iterate $(x, y, s) \in \mathcal{N}_\infty^-(\gamma)$ and let $(\Delta x, \Delta y, \Delta s)$ be the solution of (19). Then by Lemma 5.3 of [11] we have*

$$\|\Delta x \Delta s\| \leq 2^{-\frac{3}{2}} \left(\frac{n\mu^2}{\gamma\mu_g} + n\mu_g + \frac{\alpha_{\mathrm{a}}^2 n\mu_g}{16} + (1-\alpha_{\mathrm{a}})\frac{n\mu_g}{4} - 2n\mu + \frac{\alpha_{\mathrm{a}} n\mu}{2\gamma}\right).$$

Proof: Since $(\Delta x^a)^T \Delta s^a = 0$, both $\mathcal{I}_+$ and $\mathcal{I}_-$ are nonempty. If we multiply the third equation of (19) by $(XS)^{-\frac{1}{2}}$, then by Lemma 5.3 of [11] we have

$$\begin{aligned} \|\Delta x \Delta s\| &\leq 2^{-\frac{3}{2}} \left\| \mu(XSe)^{-\frac{1}{2}} - (XSe)^{\frac{1}{2}} - \alpha_{\mathrm{a}}(XS)^{-\frac{1}{2}} \Delta x^a \Delta s^a \right\|^2 \\ &= 2^{-\frac{3}{2}} \left(\mu^2 \sum_{i\in\mathcal{I}} \frac{1}{x_i s_i} + x^T s + \alpha_{\mathrm{a}}^2 \sum_{i\in\mathcal{I}} \frac{(\Delta x_i^a \Delta s_i^a)^2}{x_i s_i} - 2n\mu - 2\alpha_{\mathrm{a}}\mu \sum_{i\in\mathcal{I}} \frac{\Delta x_i^a \Delta s_i^a}{x_i s_i}\right) \\ &\leq 2^{-\frac{3}{2}} \left(\frac{n\mu^2}{\gamma\mu_g} + n\mu_g + \frac{\alpha_{\mathrm{a}}^2 n\mu_g}{16} - (1-\alpha_{\mathrm{a}}) \sum_{i\in\mathcal{I}_-} \Delta x_i^a \Delta s_i^a - 2n\mu + \frac{\alpha_{\mathrm{a}} n\mu}{2\gamma}\right) \\ &\leq 2^{-\frac{3}{2}} \left(\frac{n\mu^2}{\gamma\mu_g} + n\mu_g + \frac{\alpha_{\mathrm{a}}^2 n\mu_g}{16} + (1-\alpha_{\mathrm{a}})\frac{n\mu_g}{4} - 2n\mu + \frac{\alpha_{\mathrm{a}} n\mu}{2\gamma}\right), \end{aligned}$$

where the first inequality follows from (18), Lemma 7.2, and the assumption that the previous iterate is in $\mathcal{N}_\infty^-(\gamma)$. The second inequality also follows from Lemma 7.2. $\square$

Now, analogous to the analysis of Algorithm 1, we have to solve the one dimensional optimization problem (10) for a given α_t in order to estimate the parameter μ_t. In this case we may choose a large value $\alpha_t = \frac{\gamma^2}{n}$. For the worst case analysis the ith constraint of (10) holds, if for $\alpha_t = \frac{\gamma^2}{n}$ one has

$$-\frac{\gamma\mu_g}{4} + (1-\gamma)\mu - \frac{\gamma^2\,|\Delta x_i \Delta s_i|}{n} \geq 0.$$

By Lemma 3.2 the previous inequality holds if

$$-\frac{\gamma\mu_g}{4} + (1-\gamma)\mu - \frac{2^{\frac{-3}{2}}\gamma\mu^2}{\mu_g} - 2^{\frac{-3}{2}}\gamma^2\mu_g - \frac{2^{\frac{-3}{2}}\gamma^2\alpha_{\mathrm{a}}^2\mu_g}{16} - \frac{2^{\frac{-3}{2}}\gamma^2(1-\alpha_{\mathrm{a}})\mu_g}{4} + 2^{\frac{-1}{2}}\gamma^2\mu - \frac{2^{\frac{-3}{2}}\alpha_{\mathrm{a}}\gamma\mu}{2} \geq 0.$$

After reordering we have

$$-\frac{\gamma}{2\sqrt{2}\mu_g}\mu^2 + \left(\frac{\gamma^2}{\sqrt{2}} - \frac{\gamma\alpha_{\mathrm{a}}}{4\sqrt{2}} + 1 - \gamma\right)\mu - \frac{\gamma^2\mu_g}{2\sqrt{2}} - \frac{\gamma\mu_g}{4} - \frac{\gamma^2(1-\alpha_{\mathrm{a}})\mu_g}{8\sqrt{2}} - \frac{\gamma^2\alpha_{\mathrm{a}}^2\mu_g}{32\sqrt{2}} \geq 0. \tag{20}$$

In the following lemma we give an interval in which the previous inequality holds.

Lemma 3.3 *For $\gamma \in (0, 0.62)$ and $n \geq 2$ the step size $\alpha_t = \frac{\gamma^2}{n}$ is feasible for any*

$$\mu \in [\eta_1\mu_g \ , \ \eta_2\mu_g],$$

where

$$\eta_1 = \frac{-4\sqrt{2}\gamma + 4\gamma^2 - \gamma\alpha_{\mathrm{a}} + 4\sqrt{2} - \sqrt{\Delta}}{4\gamma}$$

and

$$\eta_2 = \frac{-4\sqrt{2}\gamma + 4\gamma^2 - \gamma\alpha_{\mathrm{a}} + 4\sqrt{2} + \sqrt{\Delta}}{4\gamma}$$

with

$$\Delta = (32 + 24\sqrt{2})\gamma^2 - (32\sqrt{2} + 20)\gamma^3 + 8\sqrt{2}\gamma^2\alpha_{\mathrm{a}} - 64\gamma + 16\gamma^4 - 4\gamma^3\alpha_{\mathrm{a}} + \gamma^2\alpha_{\mathrm{a}}^2 - 8\sqrt{2}\gamma\alpha_{\mathrm{a}} + 32 - \gamma^3\alpha_{\mathrm{a}}^2.$$

Moreover, $0 < \eta_1 < 0.9$.

Proof: For the worst case analysis it suffices to find the interval for which inequality (20) holds. By simple calculus one can show that $\eta_1\mu_g$ and $\eta_2\mu_g$ are two real roots of inequality (20) if and only if $\Delta \geq 0$, which holds for all $\gamma \in (0, 0.64)$ and $n \geq 2$. By using the definition of η_1, $\eta_1 > 0$ is equivalent to $8\sqrt{2}\gamma^2 + 20\gamma^3 - 4\gamma^3\alpha_{\mathrm{a}} + \gamma^3\alpha_{\mathrm{a}}^2 > 0$, which is definitely true for all $\gamma \in (0, 1)$. Finally, $\eta_1 < 0.9$ is equivalent to $(920\sqrt{2} + 324)\gamma^2 - 220\gamma^3 + 180\gamma^2\alpha_{\mathrm{a}} - 770\sqrt{2}\gamma - 100\gamma^3\alpha_{\mathrm{a}} + 25\gamma^3\alpha_{\mathrm{a}}^2 < 0$, which holds for all $\gamma \in (0, 0.62)$. This completes the proof of the lemma. $\square$

Corollary 3.4 *The optimal value μ_t of problem (10) satisfies $\mu_t \leq \eta_1 \mu_g$.*

Remark 3.5 *Because $\mu_t \leq \eta_1 \mu_g \leq 0.9\mu_g$, Algorithm 2 features a large update of parameter μ at each step.*

The scheme of the modified version of Algorithm 1 is presented as Algorithm 2:

Algorithm 2

Input:

A neighborhood parameter $\gamma \in (0, 0.62)$;
$\alpha_t \in (0, 1]$;
an accuracy parameter $\epsilon > 0$;
$(x^0, y^0, s^0) \in \mathcal{N}_\infty^-(\gamma)$.

begin
while $x^T s > \epsilon$ **do**
begin
Predictor Step
Solve (4) for $(\Delta x, \Delta y, \Delta s)$;
Compute the maximum step size $\alpha_{a\max}$ such that $(x(\alpha_{\mathrm{a}}), y(\alpha_{\mathrm{a}}), s(\alpha_{\mathrm{a}})) \in \mathcal{F}$ (algorithm does not make this step).
end
begin
Corrector step
Solve (10) with α_t to determine the target μ_t;
Solve (19) for $(\Delta x, \Delta y, \Delta s)$;
Compute the maximum step size α_{c} such that $(x(\alpha_{\mathrm{c}}), y(\alpha_{\mathrm{c}}), s(\alpha_{\mathrm{c}})) \in \mathcal{N}_\infty^-(\gamma)$;
Set $(x(\alpha_{\mathrm{c}}), y(\alpha_{\mathrm{c}}), s(\alpha_{\mathrm{c}})) = (x + \alpha_{\mathrm{c}}\Delta x, y + \alpha_{\mathrm{c}}\Delta y, s + \alpha_{\mathrm{c}}\Delta s)$.
end
end

The following theorem gives an upper bound for the maximum number of iterations that Algorithm 2 (the modified version of Algorithm 1) needs to find an $\epsilon-$approximation solution.

Theorem 3.6 *Let $\alpha_t = \frac{\gamma^2}{n}$ in the corrector step, then Algorithm 2 stops after at most*

$$O\left(n \log \frac{n}{\epsilon}\right)$$

iterations with a solution (x, y, s) satisfying $x^T s \leq \epsilon$.

Proof: After each iteration one has

$$x(\alpha_c)^T s(\alpha_c) = \left(1 - \alpha_c + \alpha_c \frac{\mu_t}{\mu_g}\right) x^T s.$$

In the corrector step of Algorithm 2, by construction $\alpha_c \geq \frac{\gamma^2}{n}$ and $\mu_t \leq \mu_g$. This implies

$$x(\alpha_c)^T s(\alpha_c) \leq \left(1 - \frac{\gamma^2}{n} + \frac{\gamma^2 \mu_t}{\mu_g n}\right) x^T s.$$

By Corollary 3.4 $\mu_t \leq \eta_1 \mu_g$ and by Lemma 3.3 we have $0 < \eta_1 < 0.9$, therefore

$$x(\alpha_c)^T s(\alpha_c) \leq \left(1 - \frac{\gamma^2}{10n}\right) x^T s$$

that completes the proof by Theorem 3.2 of [11]. □

Remark 3.7 *We see that the iteration complexity bound of Algorithm 2 is in the same order as for the one in [8], but Algorithm 2 is not using any safeguard strategy while the one in [8] does.*

Remark 3.8 *By changing the right hand side of (4) and (19) one can propose an infeasible variant of the Algorithm 2. The complexity analysis of this algorithm is left for the interested reader.*

4 Superlinear Convergence

In this section we prove the superlinear convergence of both algorithms. As we know, for sufficiently small μ_g it has been proved in [12] that $|\Delta x_i^a \Delta s_i^a| = \mathcal{O}(\mu_g^2)$, that imply

$$\begin{aligned} x_i(\alpha)s_i(\alpha) &= (1-\alpha)x_i s_i + \alpha^2 \Delta {x^a}_i \Delta {s^a}_i \\ &\geq (1-\alpha)x_i s_i - \mathcal{O}(\mu_g^2)\alpha^2 \geq 0. \end{aligned} \tag{21}$$

Using Lemmas II.64 and II.65 of [7], one can prove that $\alpha_a \geq 1 - \mathcal{O}(\mu_g)$. In the following theorem we prove the superlinear convergence of both Algorithms 1 and 2 .

Theorem 4.1 *Let the iterate (x^k, y^k, s^k) be generated by Algorithm 1 or Algorithm 2. When μ_g is sufficiently small, by postponing the choice of the barrier parameter μ, Algorithm 1 and it's modified version Algorithm 2 are superlinearly convergent in the sense that $\mu_g^{k+1} = \mathcal{O}\left((\mu_g^k)^{1+r}\right)$, where $r \in (0, 1)$.*

Proof: Let us consider first Algorithm 1. Let $\alpha = 1 - \mu_g^r$, where $r \in (0,1)$. Then, analogous to the proof of Theorem 7.4 in [11] one can prove that there exist a $\rho \in (0,1)$, such that for $\mu = \frac{2C}{1-\gamma}\mu_g^{1+\rho}$ the inequality

$$2C\mu_g^{1+\rho} - C\mu_g^2 - (1-\mu_g^r)\mathcal{O}(\mu_g^2) \geq 0.$$

holds. This implies the superlinear convergence of Algorithm 1. The superlinear convergence of Algorithm 2 can be proved analogously. □

5 Numerical Results

Our implementation is based on an infeasible variant of Algorithm1. The predictor step of the implemented algorithm is the same as in other Mehrotra-type predictor-corrector algorithms, namely it solves the affine scaling system of equations (3), without making this step (one do not computes the maximum step size in this direction in Algorithm 1, while in Algorithm 2 one do). Then, by using the search directions computed in the predictor step, one computes the target parameter μ_t by solving problem (10) (for the infeasible case this problem is slightly different). Since there is no information regarding the convexity or concavity of the quadratic polynomials in (10), we use a simple line search technique to solve problem (10). This is the major difference between our new approach and usual IPMs. In Table 1 we report some limited computational results using the LIPSOL software package [13]. We solve the problems from the NETLIB test set that are given in standard format. We have modified some of LIPSOL's subroutines to implement our new algorithm. For all test problems we choose $\alpha = 0.99$ and $\gamma = 0.001$. If there is no feasible solution of problem (10) for this fix step size, then we reduce α_t by a constant factor of 0.9. Finally, it is worth mentioning that in our implementation we do not use any heuristic to improve the results so one can see how competitive is our new algorithm compare to a state of the art software package like LIPSOL.

6 Final Remarks

In this paper we have proposed a new technique for IPMs in solving linear optimization problems. We proved that the new algorithms have the same order of polynomial complexity as the algorithms in [8], but the new algorithms do not employ any safeguard strategy akin the ones presented in [8]. The superlinear convergence of both algorithms are established. Finally, we reported some limited encouraging computational results that shows that our new approach is competitive with the traditional Mehrotra-type predictor-corrector algorithm. Further investigation is needed to discover more features of this novel strategy, for example how to tune the target step size α_t and the neighborhood parameter γ for a set of problems.

Table 1: Comparison of Iteration Numbers

Problem	MLIPSOL	LIPSOL	Problem	MLIPSOL	LIPSOL
25fv47	24	25	sc205	11	10
afiro	9	8	scagr7	14	14
adlittle	14	13	scagr25	19	17
agg2	19	18	scfxm1	19	19
agg3	18	17	scfxm2	22	21
bandm	17	18	scfxm3	21	21
beaconfd	14	13	ship04l	15	14
blend	12	12	ship04s	15	14
brandy	16	17	ship08l	17	16
d2q06c	32	32	ship08s	17	15
d6cube	26	23	scorpion	15	15
lotfi	20	18	stocfor1	22	17
sc50a	10	10	stocfor2	28	21
sc50b	9	7	wood1p	20	19
sc105	10	10	woodw	26	28

References

[1] C.C. Gonzaga. Complexity of predictor-corrector algorithms for LCP based on a large neighborhood of the central path. *SIAM Journal on Optimization,* **10**, 183-194, 1999.

[2] P. Hung and Y. Ye. An asymptotically $O(\sqrt{n}L)$-iteration path-following linear programming algorithm that uses long steps. *SIAM Journal on Optimization*, **6**, 570-586, 1996.

[3] N. Megiddo. Pathways to the optimal set in linear programming. In: N. Megiddo, editor, *Progress in Mathematical Programming: Interior Point and Related Methods*, 131-158, 1989. Springer Verlag, New York. Identical version in: *Proceedings of the 6th Mathematical Programming Symposium of Japan, Nagoya, Japan*, 1-35, 1986.

[4] S. Mehrotra. On the implementation of a (primal-dual) interior point method. *SIAM Journal on Optimization*, **2**, 575-601, 1992.

[5] S. Mizuno. M.J. Todd, and Y. Ye. On adaptive step primal-dual interior-point algorithms for linear programming. *Mathematics of Operations Research*, **18**, 964-981, 1993.

[6] J. Peng, T. Terlaky, and Y.B. Zhao. A predictor-corrector algorithm for linear optimization based on a specific self-regular proximity function. *SIAM Journal on Optimization,* **15**, 1105-1127, 2005.

[7] C. Roos, T. Terlaky, and J.-Ph. Vial. *Theory and Algorithms for Linear Optimization: An Interior Point Approach.* John Wiley & Sons, Chichester, UK, 1997, Springer Science, 2005 (2nd edition).

[8] M. Salahi, J. Peng, and T. Terlaky. On Mehrotra-Type Predictor-Corrector Algorithms. *Technical Report* 2005/4*, Advanced Optimization Lab. Department of Computing and Software, McMaster University.*`http://www.cas.mcmaster.ca/∼oplab/publication.`

[9] M. Salahi and T. Terlaky. Adaptive large neighborhood self-regular predictor-corrector IPMs for LO. *Technical Report* 2004/7*, Advanced Optimization Lab. Department of Computing and Software, McMaster University,* `http://www.cas.mcmaster.ca/∼oplab/publication,` to appear in *Journal of Optimization Theory and Applications.*

[10] G. Sonnevend. An "analytic center" for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. In: A. Prékopa, J. Szelezsán, and B. Strazicky, editors, *System Modeling and Optimization: Proceeding of the 12th IFIP Conference*, Budapest, Hungary, September 1985, Volume 84, Lecture Notes in Control and Information Sciences, pp. 866-876. Springer Verlag, Berlin, 1986.

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

[12] Y. Ye and K.M. Anstreicher. On quadratic and $O(\sqrt{n}L)$ convergence of a predictor-corrector algorithm for LCP. *Mathematical Programming,* **62**, 537-551, 1993.

[13] Y. Zhang, Solving large-scale linear programms by inteior point methods under the MATLAB enviroment. *Optimization Methods and Software,* **10**, 1-31, 1999.

[14] X. Zhu, J. Peng, T. Terlaky, and G. Zhang. On implementing self-regular proximity based feasible IPMs. *Technical Report* 2003/2, *Advanced Optimization Lab. Department of Computing and Software, McMaster University,* http://www.cas.mcmaster.ca/∼oplab/publication.

7 Appendix

In this section we present some technical lemmas that are used frequently throughout the paper. These lemmas are quoted from [8].

Lemma 7.1 *Let* $(\Delta x^a, \Delta y^a, \Delta s^a)$ *be the solution of (4), then*

$$\Delta x_i^a \Delta s_i^a \leq \frac{x_i s_i}{4}, \ \forall i \in \mathcal{I}_+.$$

Proof: By equation (4) for $i \in \mathcal{I}_+$ we have

$$s_i \Delta x_i^a + x_i \Delta s_i^a = -x_i s_i.$$

If we divide this equation by $x_i s_i$ we get

$$\frac{\Delta x_i^a}{x_i} + \frac{\Delta s_i^a}{s_i} = -1.$$

Since $\Delta x_i^a \Delta s_i^a > 0$, this equality implies that both $\Delta x_i^a < 0$ and $\Delta s_i^a < 0$. Then one can easily prove that

$$\Delta x_i^a \Delta s_i^a \leq \frac{x_i s_i}{4}.$$

□

Lemma 7.2 *Let* $(\Delta x^a, \Delta y^a, \Delta s^a)$ *be the solution of (4), then we have*

$$\sum_{i \in \mathcal{I}_+} \Delta x_i^a \Delta s_i^a = \sum_{i \in \mathcal{I}_-} |\Delta x_i^a \Delta s_i^a| \leq \frac{x^T s}{4}.$$

Proof: The proof is a direct consequence of the previous lemma. □