Central European Journal of Operations Research (2023) 31:691–711
https://doi.org/10.1007/s10100-022-00812-6

A new long-step interior point algorithm for linear programming based on the algebraic equivalent transformation

Marianna E.-Nagy[1,2] · **Anita Varga**[1]

Accepted: 15 July 2022 / Published online: 13 August 2022
© The Author(s) 2022

Abstract
In this paper, we investigate a new primal-dual long-step interior point algorithm for linear optimization. Based on the step size, interior point algorithms can be divided into two main groups, short-step, and long-step methods. In practice, long-step variants perform better, but usually, a better theoretical complexity can be achieved for the short-step methods. One of the exceptions is the large-update algorithm of Ai and Zhang. The new wide neighborhood and the main characteristics of the presented algorithm are based on their approach. In addition, we use the algebraic equivalent transformation technique of Darvay to determine new modified search directions for our method. We show that the new long-step algorithm is convergent and has the best known iteration complexity of short-step variants. We present our numerical results and compare the performance of our algorithm with two previously introduced Ai-Zhang type interior point algorithms on a set of linear programming test problems from the Netlib library.

Keywords Mathematical programming · Linear optimization · Interior point algorithms · Algebraic equivalent transformation technique

✉ Anita Varga
vanita@math.bme.hu

Marianna E.-Nagy
marianna.eisenberg-nagy@uni-corvinus.hu

[1] Department of Differential Equations, Institute of Mathematics, Budapest University of Technology and Economics, Müegyetem rkp. 3., Budapest 1111, Hungary

[2] Corvinus Centre for Operations Research, Corvinus University of Budapest, Fővám tér 8., Budapest 1093, Hungary

1 Introduction

In this paper, we propose a new long-step interior point algorithm (IPA) for linear optimization. We consider the primal-dual linear programming (LP) problem pair in the following standard form:

$$\left.\begin{aligned} \min \mathbf{c}^T\mathbf{x} \\ A\mathbf{x} = \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \end{aligned}\right\} \qquad \left.\begin{aligned} \max \mathbf{b}^T\mathbf{y} \\ A^T\mathbf{y} + \mathbf{s} = \mathbf{c} \\ \mathbf{s} \geq \mathbf{0} \end{aligned}\right\} \tag{1}$$

where $A \in \mathbb{R}^{m\times n}$ with full row rank, $\mathbf{b} \in \mathbb{R}^m$ and $\mathbf{c} \in \mathbb{R}^n$ are given.

The simplex method for solving linear optimization problems was developed by Dantzig (1951). Although there were different attempts to propose new methods, this was the only numerically efficient algorithm to solve LP problems for many years. Khachian (1979) proved that the ellipsoid method can solve the linear optimization problem in polynomial time. This result received much attention because of its theoretical importance, but it turned out that in practice, its performance is significantly worse than that of the simplex algorithm. Karmarkar (1984) proposed a new polynomial algorithm for LP, and this result started a new era in operations research. This algorithm generates a sequence of points in the interior of the feasible polyhedron (i.e., it is an IPA) and therefore follows an entirely different approach from the simplex method, which gives a sequence of vertices of the feasible set. Since then, this approach has received much attention, and numerous new IPAs have been introduced not just for linear optimization but also for many other problem classes, such as linear complementarity problems (LCPs), convex optimization, symmetric optimization, second-order cone optimization, etc.

Based on the step length, IPAs can be divided into two main groups, short-step and long-step methods. Long-step methods perform better in practice, but in general short-step variants have better theoretical complexity $O(\sqrt{n}L)$. Here, n denotes the dimension of the problem and $L = \log \frac{\mathbf{x}_0^T\mathbf{s}_0}{\varepsilon}$, where $(\mathbf{x}_0, \mathbf{y}_0, \mathbf{s}_0)$ is the given starting point and ε is the required precision. This discrepancy was pointed out by Renegar (2001) as the "irony of IPAs". In the last twenty years, different attempts have been made to overcome this issue (e.g., Bai et al. 2008; Peng et al. 2002; Potra 2004).

The wide neighbourhood $\mathcal{N}_\infty^-$ (to be defined in Sect. 3) has been proposed by Kojima et al. (1989). Their algorithm turned out to be efficient in practice, and its complexity was $O(nL)$. Ai and Zhang (2005) introduced an IPA that works in a new wide neighbourhood of the central path. They proved that the method has the same theoretical complexity as the short-step variants.

Using the wide neighborhood applied by Ai and Zhang, several authors proposed new long-step methods with the best known theoretical complexity. There are related results for linear programming (Darvay and Takács 2018; Liu et al. 2011; Yang et al. 2016), for horizontal linear complementarity problems (Potra 2014), and also for semidefinite optimization (Feng and Fang 2014; Li and Terlaky 2010; Pirhaji et al. 2017).

To be able to determine new search directions in IPAs, Darvay (2003) introduced the method of algebraic equivalent transformation. His main idea was to apply a strictly increasing, continuously differentiable function φ to the centering equation of the central path system, then apply Newton's method to determine the new search directions. In his paper, Darvay applied the function $\varphi(t) = \sqrt{t}$ and introduced a new, short-step algorithm for linear optimization. Most algorithms in the literature can be considered as a special case of this technique, where $\varphi(t) = t$, i.e., the identity map. The function $\varphi(t) = t - \sqrt{t}$ has been introduced by Darvay et al. (2016), also in the context of linear optimization, and has recently been investigated in several papers of Darvay and his coauthors. They presented a corrector-predictor IPA for linear optimization (Darvay et al. 2020a), and proposed another corrector-predictor IPA for sufficient LCPs (Darvay et al. 2020b), and also introduced a short-step IPA for sufficient LCPs (Darvay et al. 2021). Furthermore, the function $\varphi(t) = \frac{\sqrt{t}}{2(1+\sqrt{t})}$ has been proposed by Kheirfam and Haghighi (2016), to solve $\mathcal{P}^*(\kappa)$ linear complementarity problems. In this paper, we investigate a new long-step IPA for linear optimization, based on the function $\varphi(t) = t - \sqrt{t}$.

Most of the algorithms based on the algebraic equivalent transformation technique are short-step variants, except for the method of Darvay and Takács (2018), which is based on the function $\varphi(t) = \sqrt{t}$ and applies an Ai-Zhang type wide neighborhood.

Throughout this paper, we use the following notations. Scalars and indices are denoted by lowercase Latin letters. Vectors are denoted by bold lowercase Latin letters, and we use uppercase Latin letters to denote matrices. Sets are denoted by capital calligraphic letters. Let $\mathbf{x}, \mathbf{s} \in \mathbb{R}^n$ be two vectors; then $\mathbf{xs}$ is componentwise, namely, the Hadamard product of $\mathbf{x}$ and $\mathbf{s}$. $\mathbf{x}^+$ and $\mathbf{x}^-$ represent the positive and negative part of the vector $\mathbf{x}$, i.e.,

$$\mathbf{x}^+ = \max\{\mathbf{x}, \mathbf{0}\} \in \mathbb{R}^n \quad \text{and} \quad \mathbf{x}^- = \min\{\mathbf{x}, \mathbf{0}\} \in \mathbb{R}^n,$$

where the maximum and minimum are taken componentwise.

If $\alpha \in \mathbb{R}$, then $\mathbf{x}^\alpha = [x_1^\alpha, x_2^\alpha, \ldots, x_n^\alpha]^T$. If $s_i \neq 0$ holds for all $i \in \{1, \ldots, n\}$, then the fraction of $\mathbf{x}$ and $\mathbf{s}$ is the vector $\mathbf{x}/\mathbf{s} = [x_1/s_1, x_2/s_2 \ldots, x_n/s_n]^T$. The vector of ones is denoted by $\mathbf{e}$. $\|\mathbf{x}\|$ is the Euclidean norm of $\mathbf{x}$, $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$ denotes the L^1 (Manhattan) norm of $\mathbf{x}$, and $\|\mathbf{x}\|_\infty = \max_{i=1}^n |x_i|$ is the infinity norm of $\mathbf{x}$. $\operatorname{diag}(\mathbf{x})$ is the diagonal matrix with the elements of the vector $\mathbf{x}$ in its diagonal. Finally, $\mathcal{I}$ denotes the index set $\mathcal{I} = \{1, \ldots, n\}$.

The paper is organized as follows. In Sect. 2 we give an overview of Darvay's algebraic equivalent transformation technique. In Sect. 3 we define a new wide neighborhood, introduce a large-update IPA, and examine its correctness. In the last subsection, we prove that the complexity of the new method is $O(\sqrt{n}L)$. In Sect. 6 we present our preliminary numerical results. Section 7 summarizes our conclusions.

2 The algebraic equivalent transformation technique

The optimality criteria of the primal-dual pair (1) can be formulated as:

$$\left.\begin{aligned} A\mathbf{x} &= \mathbf{b}, & \mathbf{x} &\geq \mathbf{0} \\ A^T\mathbf{y} + \mathbf{s} &= \mathbf{c}, & \mathbf{s} &\geq \mathbf{0} \\ \mathbf{xs} &= \mathbf{0}. \end{aligned}\right\}$$

In the case of IPAs, instead of the third equation of the optimality criteria (the complementarity condition), we consider a perturbed version

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

where ν is a given positive parameter.

Let $\mathcal{F} = \{(\mathbf{x}, \mathbf{y}, \mathbf{s}) : \ A\mathbf{x} = \mathbf{b},\ A^T\mathbf{y} + \mathbf{s} = \mathbf{c},\ \mathbf{x} \geq \mathbf{0},\ \mathbf{s} \geq \mathbf{0}\}$ denote the set of primal-dual feasible solutions and $\mathcal{F}_+ = \{(\mathbf{x}, \mathbf{y}, \mathbf{s}) \in \mathcal{F} :\ \mathbf{x} > \mathbf{0},\ \mathbf{s} > \mathbf{0}\}$ the set of strictly feasible solutions.

If $\mathcal{F}_+ \neq \emptyset$, then for each $\nu > 0$ system (2) has a unique solution (Sonnevend 1986), the ν-center. The set of ν-centers form a path that is called the central path, and system (2) is called the central path problem. Furthermore, as ν tends to 0, the ν-centers converge to a solution of the linear programming problem (1).

To be able to find new search directions, Darvay (2003) introduced the algebraic equivalent transformation technique (AET). His main idea was to transform the central path problem (2) to an equivalent form:

$$\left.\begin{aligned} A\mathbf{x} &= \mathbf{b}, & \mathbf{x} &\geq \mathbf{0} \\ A^T\mathbf{y} + \mathbf{s} &= \mathbf{c}, & \mathbf{s} &\geq \mathbf{0} \\ \varphi\left(\frac{\mathbf{xs}}{\nu}\right) &= \varphi(\mathbf{e}), \end{aligned}\right\} \tag{3}$$

where $\varphi : (\xi, \infty) \to \mathbb{R}$ is a continuously differentiable function with $\varphi'(t) > 0$ for all $t \in (\xi, \infty)$, $\xi \in [0, 1)$. It is important to note that the transformed system (3) does not modify the central path; it determines different search directions depending on the function φ. More precisely, if we are at the point $(\mathbf{x}, \mathbf{y}, \mathbf{s}) \in \mathcal{F}_+ \subset \mathbb{R}^{n+m+n}$ and take a step toward the $\nu = \tau\mu$-center, where $\mu = \mathbf{x}^T\mathbf{s}/n$ and $\tau \in (0, 1)$ is a given update parameter, then applying Newton's method to (3), the search direction $(\Delta\mathbf{x}, \Delta\mathbf{y}, \Delta\mathbf{s})$

is the solution of the following system:

$$\left.\begin{aligned} A\Delta\mathbf{x} &= \mathbf{0} \\ A^T\Delta\mathbf{y} + \Delta\mathbf{s} &= \mathbf{0} \\ \mathbf{s}\Delta\mathbf{x} + \mathbf{x}\Delta\mathbf{s} &= \tau\mu\frac{\varphi(\mathbf{e}) - \varphi\left(\frac{\mathbf{xs}}{\tau\mu}\right)}{\varphi'\left(\frac{\mathbf{xs}}{\tau\mu}\right)}. \end{aligned}\right\} \tag{4}$$

Traditionally, in the analysis of Ai-Zhang type methods, the value of the update parameter τ is included in the formulation of the Newton-system; this is the main reason why we chose the value of ν as $\tau\mu$. The value of τ does not depend on the dimension of the problem; i.e., we propose a large-update IPA.

Since we assumed that A has full row rank and $\mathbf{x}$ and $\mathbf{s}$ are strictly positive vectors, the Newton-directions are uniquely determined by the system (4).

To facilitate the analysis of IPAs, we consider a scaled version of (4). Let

$$\mathbf{v} = \sqrt{\frac{\mathbf{xs}}{\tau\mu}}, \quad \mathbf{dx} = \frac{\mathbf{v}\Delta\mathbf{x}}{\mathbf{x}}, \quad \mathbf{ds} = \frac{\mathbf{v}\Delta\mathbf{s}}{\mathbf{s}}, \text{ and } \bar{A} = A\,\mathrm{diag}\left(\frac{\mathbf{v}}{\mathbf{s}}\right).$$

With these notations, the scaled Newton-system can be written as:

$$\left.\begin{aligned} \bar{A}\mathbf{dx} &= \mathbf{0} \\ \bar{A}^T\Delta\mathbf{y} + \mathbf{ds} &= \mathbf{0} \\ \mathbf{dx} + \mathbf{ds} &= \mathbf{p}_\varphi, \end{aligned}\right\}$$

where

$$\mathbf{p}_\varphi = \frac{\varphi(\mathbf{e}) - \varphi(\mathbf{v}^2)}{\mathbf{v}\varphi'(\mathbf{v}^2)}.$$

In this paper, we investigate the function $\varphi(t) = t - \sqrt{t}$, $t > 1/2$ (i.e., $\xi = 1/2$) introduced by Darvay et al. (2016). Since we fixed the function φ, from now on, we omit the subscript φ and simply write

$$\mathbf{p} = \frac{2(\mathbf{v} - \mathbf{v}^2)}{2\mathbf{v} - \mathbf{e}}.$$

Our goal is to introduce a new long-step IPA based on this function. To be able to prove the correctness of this method, we need to ensure that $\mathbf{p}$ is well-defined. Therefore, we assume that $v_i > 1/2$ is satisfied for all $i \in \mathcal{I}$.

Let p be the function for which $p(v_i) = p_i$ holds for all $v_i \in (1/2, \infty)$, i.e.,

$$p : \left(\frac{1}{2}, \infty\right) \to \mathbb{R}, \quad p(t) = \frac{2(t - t^2)}{2t - 1}.$$

Throughout the analysis, we will also investigate different estimations of the function $p(t)$.

3 The new algorithm

The main idea of Ai and Zhang (2005) was to decompose the Newton-directions into positive and negative parts and use different step lengths with the two components. If we apply this approach to the system (4), we get the following two systems:

$$\left.\begin{aligned} A\Delta\mathbf{x}_- &= \mathbf{0} \\ A^T\Delta\mathbf{y}_- + \Delta\mathbf{s}_- &= \mathbf{0} \\ \mathbf{s}\Delta\mathbf{x}_- + \mathbf{x}\Delta\mathbf{s}_- &= \tau\mu\mathbf{v}\mathbf{p}^- \end{aligned}\right\} \qquad \left.\begin{aligned} A\Delta\mathbf{x}_+ &= \mathbf{0} \\ A^T\Delta\mathbf{y}_+ + \Delta\mathbf{s}_+ &= \mathbf{0} \\ \mathbf{s}\Delta\mathbf{x}_+ + \mathbf{x}\Delta\mathbf{s}_+ &= \tau\mu\mathbf{v}\mathbf{p}^+, \end{aligned}\right\} \tag{5}$$

and the new point with step length $\alpha = (\alpha_1, \alpha_2)$ will be $\mathbf{x}(\alpha) = \mathbf{x} + \alpha_1\Delta\mathbf{x}_- + \alpha_2\Delta\mathbf{x}_+$, $\mathbf{y}(\alpha) = \mathbf{y} + \alpha_1\Delta\mathbf{y}_- + \alpha_2\Delta\mathbf{y}_+$ and $\mathbf{s}(\alpha) = \mathbf{s} + \alpha_1\Delta\mathbf{s}_- + \alpha_2\Delta\mathbf{s}_+$. For both systems, the coefficient matrix is exactly the same as in the system (4); therefore, using the same reasoning, it is easy to see that both systems have unique solutions.

It is important to notice that $\Delta\mathbf{x}_+$ is not the positive part of $\Delta\mathbf{x}$ (in this case the sign $+$ is a subscript instead of a superscript), it is the solution of the system with $\mathbf{p}^+$ on its right-hand side. The notation is similar for the other solutions of these systems.

We introduce the index sets $\mathcal{I}_+ = \{i \in \mathcal{I} : x_i s_i \leq \tau\mu\} = \{i \in \mathcal{I} : v_i \leq 1\}$, and $\mathcal{I}_- = \mathcal{I} \setminus \mathcal{I}_+$. Under the technical assumption $v_i > \frac{1}{2}$, the nonnegativity of a coordinate p_i is equivalent to $i \in \mathcal{I}_+$.

To facilitate the analysis of the algorithm, we introduce the scaled search directions

$$\mathbf{d}\mathbf{x}_- = \frac{\mathbf{v}\Delta\mathbf{x}_-}{\mathbf{x}}, \ \mathbf{d}\mathbf{s}_- = \frac{\mathbf{v}\Delta\mathbf{s}_-}{\mathbf{s}}, \ \mathbf{d}\mathbf{x}_+ = \frac{\mathbf{v}\Delta\mathbf{x}_+}{\mathbf{x}}, \ \mathbf{d}\mathbf{s}_+ = \frac{\mathbf{v}\Delta\mathbf{s}_+}{\mathbf{s}}.$$

The systems (5) then transform to the following systems

$$\left.\begin{aligned} \bar{A}\mathbf{d}\mathbf{x}_- &= \mathbf{0} \\ \bar{A}^T\Delta\mathbf{y}_- + \mathbf{d}\mathbf{s}_- &= \mathbf{0} \\ \mathbf{d}\mathbf{x}_- + \mathbf{d}\mathbf{s}_- &= \mathbf{p}^-, \end{aligned}\right\} \qquad \left.\begin{aligned} \bar{A}\mathbf{d}\mathbf{x}_+ &= \mathbf{0} \\ \bar{A}^T\Delta\mathbf{y}_+ + \mathbf{d}\mathbf{s}_+ &= \mathbf{0} \\ \mathbf{d}\mathbf{x}_+ + \mathbf{d}\mathbf{s}_+ &= \mathbf{p}^+. \end{aligned}\right\} \tag{6}$$

The wide neighborhood $\mathcal{N}_\infty^-$ has been introduced by Kojima et al. (1989). It is defined as follows:

$$\mathcal{N}_\infty^-(1-\tau) = \{(\mathbf{x}, \mathbf{y}, \mathbf{s}) \in \mathcal{F}_+ : \mathbf{x}\mathbf{s} \geq \tau\mu\mathbf{e}\} = \{(\mathbf{x}, \mathbf{y}, \mathbf{s}) \in \mathcal{F}_+ : \mathbf{v} \geq \mathbf{e}\}.$$

Notice that this means that a point is in the neighborhood $\mathcal{N}_\infty^-(1-\tau)$ if and only if the corresponding index set $\mathcal{I}_+$ is empty, namely $\mathbf{p}^+ = \mathbf{0}$. In the analysis, we are going

to use a new neighborhood that depends only on the positive part of the vector $\mathbf{p}$:

$$\mathcal{W}(\tau,\beta)=\left\{(\mathbf{x},\mathbf{y},\mathbf{s})\in\mathcal{F}_+ : \|\mathbf{p}^+\|\le\beta \text{ and } \mathbf{v}>\frac{1}{2}\mathbf{e}\right\},$$

where $0<\beta<1/2$ is a given parameter value. The role of the technical condition $\mathbf{v}>\mathbf{e}/2$ has been discussed at the end of Sect. 2. This neighborhood is a modification of the one introduced by Ai and Zhang (2005) (since they require $\|\mathbf{vp}^+\|\le\beta$) and it is equivalent to the one used by Darvay and Takács (2018) for the function $\varphi(t)=\sqrt{t}$.

Following the idea of Ai and Zhang (2005), the next lemma verifies that $\mathcal{W}(\tau,\beta)$ is indeed a wide neighborhood:

Lemma 1 *Let $0<\beta<1/2$ and $0<\tau<1$ be given parameters, and let $\gamma=1/4\,(1+\sqrt{1-2\beta})^2\tau$. Then*

$$\mathcal{N}_\infty^-(1-\tau)\subseteq\mathcal{W}(\tau,\beta)\subseteq\mathcal{N}_\infty^-(1-\gamma).$$

Proof If $(\mathbf{x},\mathbf{y},\mathbf{s})\in\mathcal{N}_\infty^-(1-\tau)$, then $\|\mathbf{p}^+\|=0<\beta$ and $\mathbf{v}\ge\mathbf{e}>1/2\mathbf{e}$.

For the second inclusion, let $(\mathbf{x},\mathbf{y},\mathbf{s})\in\mathcal{W}(\tau,\beta)$ and assume indirectly that there exists an index $i\in\mathcal{I}$ for which $x_is_i<\gamma\mu$, i.e.,

$$v_i^2<\gamma/\tau=1/4\,(1+\sqrt{1-2\beta})^2.$$

Since $p(t)$ is a strictly decreasing function,

$$p_i=p(v_i)>\frac{2\left(\sqrt{\frac{\gamma}{\tau}}-\frac{\gamma}{\tau}\right)}{2\sqrt{\frac{\gamma}{\tau}}-1}=\frac{\beta}{\sqrt{1-2\beta}}>\beta,$$

which is a contradiction. □

The following lower and upper bounds on the coordinates of the vector $\mathbf{v}$ will be useful for different estimations during the analysis.

Corollary 1 *Let $(\mathbf{x},\mathbf{y},\mathbf{s})\in\mathcal{W}(\tau,\beta)$, then*

$$\begin{aligned}\frac{1+\sqrt{1-2\beta}}{2}\le v_i\le 1 &\qquad \forall i\in\mathcal{I}_+,\\ 1<v_i\le\sqrt{n/\tau} &\qquad \forall i\in\mathcal{I}_-.\end{aligned}$$

Proof The first statement follows directly from Lemma 1. The upper bound $v_i\le\sqrt{n/\tau}$ holds for all $i\in\mathcal{I}$ since

$$\sum_{i\in\mathcal{I}}v_i^2=\sum_{i\in\mathcal{I}}\frac{x_is_i}{\tau\mu}=\frac{1}{\tau\mu}\mathbf{x}^T\mathbf{s}=\frac{n}{\tau}. \tag{7}$$

□

Before presenting the analysis, we give the pseudocode of the IPA.

Input: $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{c} \in \mathbb{R}^n$
the update parameter $0 < \tau < 1$,
the neighborhood parameter $0 < \beta < 1$,
the accuracy parameter $\varepsilon > 0$,
an initial point $(\mathbf{x}_0, \mathbf{y}_0, \mathbf{s}_0) \in \mathcal{W}(\tau, \beta)$.

$\mathbf{x} := \mathbf{x}_0$, $\mathbf{y} := \mathbf{y}_0$, $\mathbf{s} := \mathbf{s}_0$ and $\mu := \mu_0 = \mathbf{x}_0^T \mathbf{s}_0/n$
while $\mathbf{x}^T\mathbf{s} > \epsilon$ **do**
  Determine $\Delta\mathbf{x}_+$, $\Delta\mathbf{s}_+$, $\Delta\mathbf{y}_+$ and $\Delta\mathbf{x}_-$, $\Delta\mathbf{s}_-$, $\Delta\mathbf{y}_-$ according to (5);
  Set $\alpha_2 = 1$ and $\alpha_1 = \max\{\alpha_1 \in [0, 1] : (\mathbf{x}(\alpha), \mathbf{y}(\alpha), \mathbf{s}(\alpha)) \in \mathcal{W}(\tau, \beta)\}$,
  where $\mathbf{x}(\alpha) = \mathbf{x} + \alpha_1\Delta\mathbf{x}_- + \alpha_2\Delta\mathbf{x}_+$, $\mathbf{y}(\alpha) = \mathbf{y} + \alpha_1\Delta\mathbf{y}_- + \alpha_2\Delta\mathbf{y}_+$ and
  $\mathbf{s}(\alpha) = \mathbf{s} + \alpha_1\Delta\mathbf{s}_- + \alpha_2\Delta\mathbf{s}_+$;
  $(\mathbf{x}, \mathbf{y}, \mathbf{s}) := (\mathbf{x}(\alpha), \mathbf{y}(\alpha), \mathbf{s}(\alpha))$;
  $\mu := \mathbf{x}^T\mathbf{s}/n$;
end

Algorithm 1: New IPA for LP

During the analysis, we consider the case of $\alpha_2 = 1$, i.e., we take a full Newton-step in the direction $(\Delta\mathbf{x}_+, \Delta\mathbf{y}_+, \Delta\mathbf{s}_+)$, and determine a value of α_1 so that the desired complexity of the algorithm can be achieved.

From now on, we assume that a point $(\mathbf{x}, \mathbf{y}, \mathbf{s}) \in \mathcal{W}(\tau, \beta)$ is given, and in the next section, we prove the correctness of the algorithm.

4 Analysis of the algorithm

Let us introduce the following notations:

$$\mathbf{dx}(\alpha) = \alpha_1\mathbf{dx}_- + \alpha_2\mathbf{dx}_+, \quad \mathbf{ds}(\alpha) = \alpha_1\mathbf{ds}_- + \alpha_2\mathbf{ds}_+,$$
$$\mathbf{h}(\alpha) = \tau\mu\mathbf{v}^2 + \alpha_1\tau\mu\mathbf{vp}^- + \alpha_2\tau\mu\mathbf{vp}^+,$$

where $\alpha_1, \alpha_2 \in [0, 1]$ are given step lengths, whose values will be specified later. With these notations, the equation

$$\mathbf{x}(\alpha)\mathbf{s}(\alpha) = (\mathbf{x} + \alpha_1\Delta\mathbf{x}_- + \alpha_2\Delta\mathbf{x}_+)(\mathbf{s} + \alpha_1\Delta\mathbf{s}_- + \alpha_2\Delta\mathbf{s}_+)$$

can be written as

$$\mathbf{x}(\alpha)\mathbf{s}(\alpha) = \mathbf{h}(\alpha) + \tau\mu\mathbf{dx}(\alpha)\mathbf{ds}(\alpha).$$

It is important to note that the search directions are orthogonal, as usually in the case of LP problems, since

$$\mathbf{dx}(\alpha)^T\mathbf{ds}(\alpha) = \alpha_1^2\mathbf{dx}_-^T\mathbf{ds}_- + \alpha_1\alpha_2(\mathbf{dx}_-^T\mathbf{ds}_+ + \mathbf{dx}_+^T\mathbf{ds}_-) + \alpha_2^2\mathbf{dx}_+^T\mathbf{ds}_+.$$

Furthermore, $\mathbf{dx}_+$ and $\mathbf{dx}_-$ are in the kernel of the matrix $\bar{A}$, while $\mathbf{ds}_+$ and $\mathbf{ds}_-$ are in the rowspace of $\bar{A}$ (see system (6)). Therefore, all four scalar products are 0 in the previous expression.

The next two lemmas give lower bounds on the value of $\mathbf{h}(\alpha)$.

Lemma 2 *Let $\alpha \in [0,1]^2$. Then $h_i(\alpha) \geq \tau\mu$ for all $i \in \mathcal{I}_-$.*

Proof In the case of $i \in \mathcal{I}_-$, $v_i > 1$ and $h_i(\alpha) = \tau\mu v_i(v_i + \alpha_1 p_i)$. We need to prove that $v_i(v_i + \alpha_1 p_i) \geq 1$, i.e., $\alpha_1 \leq \frac{1-v_i^2}{v_i p_i}$ holds.

Let us examine the expression $\frac{1-t^2}{tp(t)}$ over the interval $(1, \infty)$:

$$\frac{1-t^2}{tp(t)} = \frac{1-t^2}{t}\frac{2t-1}{2t(1-t)} = \frac{2t^2+t-1}{2t^2} = 1 + \frac{t-1}{2t^2} > 1.$$

On the other hand, $\alpha_1 \leq 1$ by definition. Thus, $h_i(\alpha) \geq \tau\mu$ holds for all $i \in \mathcal{I}_-$. □

We show that $\mathbf{h}(\alpha)$ is a componentwise strictly positive vector.

Lemma 3 *Let $(\mathbf{x}, \mathbf{y}, \mathbf{s}) \in \mathcal{W}(\tau, \beta)$ and $\alpha \in [0,1]^2$. Then $\mathbf{h}(\alpha) \geq \gamma\mu\mathbf{e}$, and consequently $\mathbf{h}(\alpha) > \mathbf{0}$.*

Proof By Lemma 1, $\tau\mu v_i^2 = \mathbf{x}_i\mathbf{s}_i \geq \gamma\mu$ for all $i \in \mathcal{I}$. Furthermore, if $i \in \mathcal{I}_+$, then $v_i p_i > 0$, so $h_i(\alpha) \geq \tau\mu v_i^2 \geq \gamma\mu$.

In the case of $i \in \mathcal{I}_-$, the statement is a consequence of Lemma 2, since $h_i(\alpha) \geq \tau\mu \geq \gamma\mu$. □

To be able to prove the feasibility of the new iterates and ensure that they stay in the neighborhood $\mathcal{W}(\tau, \beta)$, we need the following technical lemma:

Lemma 4 *Let $(\mathbf{x}, \mathbf{y}, \mathbf{s}) \in \mathcal{W}(\tau, \beta)$, $\alpha_1 = \sqrt{\frac{\beta\tau}{n}}$ and $\alpha_2 = 1$. Then*

$$\| [\mathbf{dx}(\alpha)\mathbf{ds}(\alpha)]^- \|_1 = \| [\mathbf{dx}(\alpha)\mathbf{ds}(\alpha)]^+ \|_1 \leq \frac{1}{2}\beta.$$

Proof According to Lemma 3.5 of Ai and Zhang (2005) and using the orthogonality of $\mathbf{dx}(\alpha)$ and $\mathbf{ds}(\alpha)$, we have

$$\begin{aligned} \|[\mathbf{dx}(\alpha)\mathbf{ds}(\alpha)]^-\|_1 = \|[\mathbf{dx}(\alpha)\mathbf{ds}(\alpha)]^+\|_1 &\leq \tfrac{1}{4}\|\mathbf{dx}(\alpha) + \mathbf{ds}(\alpha)\|^2 \\ = \tfrac{1}{4}\|\alpha_1(\mathbf{dx}_- + \mathbf{ds}_-) + \alpha_2(\mathbf{dx}_+ + \mathbf{ds}_+)\|^2 &= \tfrac{1}{4}\left(\alpha_1^2\|\mathbf{p}^-\|^2 + \alpha_2^2\|\mathbf{p}^+\|^2\right). \end{aligned}$$

By the definition of $\mathcal{W}(\tau, \beta)$, we have $\|\mathbf{p}^+\| \leq \beta$. We need to estimate the term $\|\mathbf{p}^-\|^2$. According to (7), we have

$$\|\mathbf{p}^-\|^2 = \sum_{i\in\mathcal{I}_-}\left(v_i - \frac{v_i}{2v_i - 1}\right)^2 \leq \sum_{i\in\mathcal{I}_-} v_i^2 \leq \sum_{i\in\mathcal{I}} v_i^2 = \frac{n}{\tau}.$$

Using these two estimations and substituting the values of α_1 and α_2, we can write

$$\frac{1}{4}\left(\alpha_1^2\|\mathbf{p}^-\|^2+\alpha_2^2\|\mathbf{p}^+\|^2\right)\le\frac{1}{4}\frac{\beta\tau}{n}\frac{n}{\tau}+\frac{1}{4}\beta^2=\frac{1}{4}\beta+\frac{1}{4}\beta^2\le\frac{1}{2}\beta.$$

□

The next lemma gives a positive lower bound on the vector $\mathbf{x}(\alpha)\mathbf{s}(\alpha)$, which is the first step to prove the strict feasibility of the new point.

Lemma 5 *Let* $(\mathbf{x},\mathbf{y},\mathbf{s})\in\mathcal{W}(\tau,\beta)$, $\alpha_1=\sqrt{\frac{\beta\tau}{n}}$ *and* $\alpha_2=1$. *Then*

$$\mathbf{x}(\alpha)\mathbf{s}(\alpha)\ge\frac{1-2\beta+\sqrt{1-2\beta}}{2}\tau\mu\mathbf{e}$$

holds.

Proof By Lemma 3, we have $\mathbf{h}(\alpha)\ge\gamma\mu\mathbf{e}$. Using Lemma 4 and substituting the value of γ, we get

$$\begin{aligned}\mathbf{x}(\alpha)\mathbf{s}(\alpha)&=\mathbf{h}(\alpha)+\tau\mu\mathbf{dx}(\alpha)\mathbf{ds}(\alpha)\ge\gamma\mu\mathbf{e}-\tau\mu\|[\mathbf{dx}(\alpha)\mathbf{ds}(\alpha)]^-\|_1\mathbf{e}\\&\ge\gamma\mu\mathbf{e}-\tau\mu\frac{1}{2}\beta\mathbf{e}=\tau\mu\left(\frac{\gamma}{\tau}-\frac{\beta}{2}\right)\mathbf{e}=\frac{1-2\beta+\sqrt{1-2\beta}}{2}\tau\mu\mathbf{e}.\end{aligned}$$

□

The following statement is the linear programming analogue of Proposition 3.2 by Ai and Zhang (2005) (they proposed it for monotone linear complementarity problems). The proof remains the same.

Lemma 6 *Let* $(\mathbf{x},\mathbf{y},\mathbf{s})\in\mathcal{F}^+$ *and* $(\Delta\mathbf{x},\Delta\mathbf{y},\Delta\mathbf{s})$ *be the solution of the system*

$$\begin{aligned}A\Delta\mathbf{x}&=\mathbf{0}\\A^T\Delta\mathbf{y}+\Delta\mathbf{s}&=\mathbf{0}\\\mathbf{s}\Delta\mathbf{x}+\mathbf{x}\Delta\mathbf{s}&=\mathbf{z}.\end{aligned}$$

If $\mathbf{z}+\mathbf{xs}>0$ *and* $(\mathbf{x}+t_0\Delta\mathbf{x})(\mathbf{s}+t_0\Delta\mathbf{s})>0$ *holds for some* $t_0\in(0,1]$, *then* $\mathbf{x}+t\Delta\mathbf{x}>0$ *and* $\mathbf{s}+t\Delta\mathbf{s}>0$ *for all* $t\in(0,t_0]$.

We have already proved that $\mathbf{h}(\alpha)>\mathbf{0}$ for all $\alpha\in[0,1]$ (see Lemma 3), and $\mathbf{x}(\alpha)\mathbf{s}(\alpha)>0$ for $\alpha_1=\sqrt{\beta\tau/n}$ and $\alpha_2=1$ (see Lemma 5), therefore by Lemma 6, we have that the new points are also strictly positive, namely $\mathbf{x}(\alpha)>\mathbf{0}$ and $\mathbf{s}(\alpha)>\mathbf{0}$.

The following two statements propose bounds on the duality gap of the new point: $\mu(\alpha)=\mathbf{x}(\alpha)^T\mathbf{s}(\alpha)/n$.

Lemma 7 *Let* $\alpha_1=\sqrt{\frac{\beta\tau}{n}}$ *and* $\alpha_2=1$. *Then* $\mu(\alpha)\ge(1-\alpha_1)\,\mu$.

Proof Since $\mathbf{v}^T\mathbf{p}^+ \geq 0$, and $v_i p_i = v_i^2 - \frac{v_i^2}{2v_i - 1} \leq v_i^2$ for all $i \in \mathcal{I}_-$ then by (7) we have

$$\mu(\alpha) = \frac{\mathbf{x}(\alpha)^T\mathbf{s}(\alpha)}{n} = \mu + \frac{\alpha_1\tau\mu}{n}\mathbf{v}^T\mathbf{p}^- + \frac{\alpha_2\tau\mu}{n}\mathbf{v}^T\mathbf{p}^+ \geq \mu + \frac{\alpha_1\tau\mu}{n}\mathbf{v}^T\mathbf{p}^-$$
$$= \mu - \frac{\alpha_1\tau\mu}{n}\sum_{i\in\mathcal{I}_-}\frac{2v_i(v_i^2 - v_i)}{2v_i - 1} \geq \mu - \frac{\alpha_1\tau\mu}{n}\sum_{i\in\mathcal{I}} v_i^2 = (1 - \alpha_1)\,\mu.$$

□

The following theorem guarantees the proper reduction of the duality gap after an iteration:

Lemma 8 *Assume that* $(\mathbf{x}, \mathbf{y}, \mathbf{s}) \in \mathcal{W}(\tau, \beta)$, $\alpha_1 = \sqrt{\frac{\beta\tau}{n}}$ *and* $\alpha_2 = 1$. *Then*

$$\mu(\alpha) \leq \left(1 - \sqrt{\frac{\beta\tau}{n}}\left[\frac{8}{9}(1 - \tau) - \sqrt{\beta\tau}\right]\right)\mu. \tag{8}$$

Proof Observe that

$$\mu(\alpha) = \frac{\mathbf{x}(\alpha)^T\mathbf{s}(\alpha)}{n} = \mu + \frac{\alpha_1\tau\mu}{n}\mathbf{v}^T\mathbf{p}^- + \frac{\alpha_2\tau\mu}{n}\mathbf{v}^T\mathbf{p}^+.$$

First, let us estimate the term $\mathbf{v}^T\mathbf{p}^+$:

$$\mathbf{e}^T\left(\mathbf{v}\mathbf{p}^+\right) = \left\|\mathbf{v}\mathbf{p}^+\right\|_1 \leq \sqrt{n}\left\|\mathbf{v}\mathbf{p}^+\right\| \leq \sqrt{n}\beta. \tag{9}$$

The first equality holds since $\mathbf{v}$ is positive, and we consider only the positive part of $\mathbf{p}$. By applying the Cauchy-Schwarz inequality, we get the first estimation. Using the property $v_i \leq 1$ when $i \in \mathcal{I}_+$ and the definition of the neighborhood $\mathcal{W}(\tau, \beta)$, the last inequality can also be verified.

To obtain an upper bound on the expression $\mathbf{v}^T\mathbf{p}^-$, consider the inequalities $2\mathbf{v} - \mathbf{e} > \mathbf{0}$ and $v_i > 1$ for all $i \in \mathcal{I}_-$:

$$\mathbf{v}^T\mathbf{p}^- = \mathbf{e}^T\left(\mathbf{v}\frac{2(\mathbf{v} - \mathbf{v}^2)^-}{2\mathbf{v} - \mathbf{e}}\right) = \sum_{i\in\mathcal{I}_-}\frac{2v_i(v_i - v_i^2)}{2v_i - 1}$$
$$= \sum_{i\in\mathcal{I}_-}\frac{2v_i^2}{(1 + v_i)(2v_i - 1)}(1 - v_i^2)$$
$$\leq \sum_{i\in\mathcal{I}_-}\frac{8}{9}(1 - v_i^2) \leq \sum_{i\in\mathcal{I}}\frac{8}{9}(1 - v_i^2) = \frac{8}{9}n\left(1 - \frac{1}{\tau}\right). \tag{10}$$

Using (9) and (10) we obtain

$$\mu(\alpha) \leq \mu + \frac{\alpha_1 \tau \mu}{n} \frac{8}{9} n \left(1 - \frac{1}{\tau}\right) + \frac{\alpha_2 \tau \mu}{n} \sqrt{n} \beta = \left(1 - \alpha_1 \left[\frac{8}{9}(1-\tau) - \sqrt{\beta\tau}\right]\right) \mu.$$

□

Notice, that the upper bound on $\mu(\alpha)$ in (8) is positive for all $\beta, \tau \in (0, 1)$. Indeed,

$$1 - \sqrt{\frac{\beta\tau}{n}} \left[\frac{8}{9}(1-\tau) - \sqrt{\beta\tau}\right] \geq 1 - \frac{8}{9}(1-\tau) > \frac{1}{9}.$$

With a suitable parameter setting, we can ensure that the duality gap decreases strictly monotonically, i.e., $\mu(\alpha) < \mu$.

Corollary 2 *Let $\tau \leq 1/2$ and $\beta \leq 1/4$. If $(\mathbf{x}, \mathbf{y}, \mathbf{s}) \in \mathcal{W}(\tau, \beta)$, $\alpha_1 = \sqrt{\frac{\beta\tau}{n}}$ and $\alpha_2 = 1$, then $\mu(\alpha) < \mu$ holds.*

Proof We need to check whether the multiplier of μ in inequality (8) is less than 1. This means that $8/9(1-\tau) - \sqrt{\beta\tau} > 0$ and this holds when $\beta < 64/81(1-\tau)^2/\tau$, which is satisfied for our choice of parameter values. □

In addition to strict feasibility, we also need to prove the fulfilment of the technical condition $\mathbf{v}(\alpha) = \sqrt{\frac{\mathbf{x}(\alpha)\mathbf{s}(\alpha)}{\tau\mu(\alpha)}} > \frac{1}{2}\mathbf{e}$.

Lemma 9 *Let $(\mathbf{x}, \mathbf{y}, \mathbf{s}) \in \mathcal{W}(\tau, \beta)$, $\alpha_1 = \sqrt{\frac{\beta\tau}{n}}$ and $\alpha_2 = 1$. If $\beta < \frac{\sqrt{3}}{4}$, then $\mathbf{v}(\alpha) > \frac{1}{2}\mathbf{e}$ holds.*

Proof From Lemma 5 and Corollary 2, we have

$$\mathbf{v}^2(\alpha) = \frac{\mathbf{x}(\alpha)\mathbf{s}(\alpha)}{\tau\mu(\alpha)} \geq \frac{1 - 2\beta + \sqrt{1-2\beta}}{2}\mathbf{e}. \tag{11}$$

Since $\frac{1-2\beta+\sqrt{1-2\beta}}{2} > \frac{1}{4}$ if $\beta < \frac{\sqrt{3}}{4}$, we have proved the statement. □

To show that the new iterates remain in the neighborhood $\mathcal{W}(\tau, \beta)$, we need another technical lemma:

Lemma 10 *Let $(\mathbf{x}, \mathbf{y}, \mathbf{s}) \in \mathcal{W}(\tau, \beta)$, $\alpha_1 = \sqrt{\frac{\beta\tau}{n}}$ and $\alpha_2 = 1$. Then*

$$\|[\tau\mu(\alpha)\mathbf{e} - \mathbf{h}(\alpha)]^+\| \leq \beta\tau\mu(\alpha)\left(1 - \frac{1 + \sqrt{1-2\beta}}{2}\right).$$

Proof Based on Lemma 2, $\tau\mu(\alpha) - h_i(\alpha) \leq 0$ for all $i \in \mathcal{I}_-$. Therefore we need to examine indices only from the set $\mathcal{I}_+$.

Since $1/2 < v_i \leq 1$ for all $i \in \mathcal{I}_+$, we have

$$\frac{1-v_i^2}{p_i} = \frac{(1-v_i^2)(2v_i-1)}{2v_i(1-v_i)} = \frac{2v_i^2+v_i-1}{2v_i} = v_i + \frac{1}{2} - \frac{1}{2v_i} \leq v_i \leq 1 \quad \forall i \in \mathcal{I}_+. \tag{12}$$

Using Corollary 2 and (12), we obtain that

$$\begin{aligned}\tau\mu(\alpha) - h_i(\alpha) &= \tau\mu(\alpha) - \tau\mu\left(v_i^2 + v_i p_i\right) \leq \tau\mu(\alpha)\left(1 - v_i^2 - v_i p_i\right) \\ &= \tau\mu(\alpha) p_i(1 - v_i) \leq \tau\mu(\alpha) p_i \left(1 - \frac{1+\sqrt{1-2\beta}}{2}\right) \ \forall\, i \in \mathcal{I}_+,\end{aligned}$$

where in the last estimation, we used the first statement of Corollary 1.

Using the definition of $\mathcal{W}(\tau, \beta)$, we obtain

$$\|(\tau\mu(\alpha)\mathbf{e} - \mathbf{h}(\alpha))^+\| \leq \tau\mu(\alpha)\|\mathbf{p}^+\| \left(1 - \tfrac{1+\sqrt{1-2\beta}}{2}\right) \leq \beta\tau\mu(\alpha)\left(1 - \tfrac{1+\sqrt{1-2\beta}}{2}\right),$$

which concludes the proof. □

Now we are ready to prove that after an iteration, if the right-hand side of the third equation in the Newton system (6) is denoted by $\mathbf{p}(\alpha)$, then $\|\mathbf{p}(\alpha)^+\| \leq \beta$ holds. Together with Lemma 9, this means that the new iterates after the Newton-step remain in the neighborhood $\mathcal{W}(\tau, \beta)$.

Lemma 11 *Let $\beta \leq \frac{1}{8}$, $\tau \leq \frac{1}{8}$. If $(\mathbf{x}, \mathbf{y}, \mathbf{s}) \in \mathcal{W}(\tau, \beta)$, $\alpha_1 = \sqrt{\frac{\beta\tau}{n}}$ and $\alpha_2 = 1$, then the new iterate stays in the same neighborhood, namely $(\mathbf{x}(\alpha), \mathbf{y}(\alpha), \mathbf{s}(\alpha)) \in \mathcal{W}(\tau, \beta)$.*

Proof By the definition of $\mathcal{W}(\tau, \beta)$ and Lemma 9, we need to prove

$$\left\|\mathbf{p}(\alpha)^+\right\| = \left\| \left[\frac{2\mathbf{v}(\alpha)(\mathbf{e} - \mathbf{v}(\alpha))}{2\mathbf{v}(\alpha) - \mathbf{e}} \right]^+ \right\| \leq \beta.$$

Since $\frac{2\mathbf{v}(\alpha)}{2\mathbf{v}^2(\alpha)+\mathbf{v}(\alpha)-\mathbf{e}} > \mathbf{0}$ when $\mathbf{v}(\alpha) > 1/2\mathbf{e}$, we have

$$\begin{aligned}\left\|\mathbf{p}(\alpha)^+\right\| &= \left\| \frac{2\mathbf{v}(\alpha)}{(2\mathbf{v}(\alpha) - \mathbf{e})\,(\mathbf{e} + \mathbf{v}(\alpha))} \left[\mathbf{e} - \mathbf{v}^2(\alpha)\right]^+ \right\| \\ &\leq \left\| \frac{2\mathbf{v}(\alpha)}{2\mathbf{v}^2(\alpha) + \mathbf{v}(\alpha) - \mathbf{e}} \right\|_\infty \left\| \left[\mathbf{e} - \mathbf{v}^2(\alpha)\right]^+ \right\|. \end{aligned} \tag{13}$$

Let $q : \left(\frac{1}{2}, \infty\right) \to \mathbb{R}$ defined by $q(t) = \frac{2t}{2t^2+t-1}$. This function is strictly decreasing on its domain, therefore using (11), the first term in (13) can be estimated as

$$\left\| \frac{2\mathbf{v}(\alpha)}{2\mathbf{v}^2(\alpha) + \mathbf{v}(\alpha) - \mathbf{e}} \right\|_\infty \leq q\left(\sqrt{\frac{1 - 2\beta + \sqrt{1-2\beta}}{2}} \right), \tag{14}$$

where the expression $\sqrt{(1-2\beta+\sqrt{1-2\beta})/2}$ is strictly decreasing in β, implying that the upper bound is strictly increasing in β.

To give an upper bound on $\left\|\left[\mathbf{e}-\mathbf{v}^2(\alpha)\right]^+\right\|$, we use Lemmas 10, 4 and then 7:

$$\begin{aligned}\left\|\left[\mathbf{e}-\mathbf{v}^2(\alpha)\right]^+\right\| &= \frac{1}{\tau\mu(\alpha)}\left\|[\tau\mu(\alpha)\mathbf{e}-\mathbf{x}(\alpha)\mathbf{s}(\alpha)]^+\right\| \\ &\le \frac{1}{\tau\mu(\alpha)}\left(\left\|[\tau\mu(\alpha)\mathbf{e}-\mathbf{h}(\alpha)]^+\right\|+\tau\mu\left\|[\mathbf{dx}(\alpha)\mathbf{ds}(\alpha)]^-\right\|\right) \\ &\le \frac{1}{\tau\mu(\alpha)}\left(\beta\tau\mu(\alpha)\left(1-\frac{1+\sqrt{1-2\beta}}{2}\right)+\tau\mu\frac{\beta}{2}\right) \\ &\le \beta\left(\frac{1-\sqrt{1-2\beta}}{2}+\frac{1}{2-2\sqrt{\beta\tau}}\right), \end{aligned} \tag{15}$$

where the last term is strictly increasing in both β and τ.

Using the just proved inequalities (13), (14) and (15), we obtain

$$\left\|\left(\frac{2\mathbf{v}(\alpha)(\mathbf{e}-\mathbf{v}(\alpha))}{2\mathbf{v}(\alpha)-\mathbf{e}}\right)^+\right\| \le \beta\left[q\left(\sqrt{\frac{1-2\beta+\sqrt{1-2\beta}}{2}}\right)\left(\frac{1-\sqrt{1-2\beta}}{2}+\frac{1}{2-2\sqrt{\beta\tau}}\right)\right]. \tag{16}$$

To prove that this expression is less than or equal to β, we need to ensure that the value of the term in square brackets is at most 1. Notice that by the monotonicity of the estimations (14) and (15), their product is also strictly increasing both in β and τ. Moreover, substituting $\beta=\tau=1/8$, the coefficient of β on the right-hand side of (16) is less than 0.77, which concludes the proof. □

5 Iteration bound of the algorithm

Theorem 1 *Let $\beta=\tau=\frac{1}{8}$, $\alpha_1=\sqrt{\frac{\beta\tau}{n}}$, $\alpha_2=1$, and suppose that a starting point $(\mathbf{x}_0,\mathbf{y}_0,\mathbf{s}_0)\in\mathcal{W}(\tau,\beta)$ is given. The algorithm then provides an ε-optimal solution of the primal-dual pair of LPs in*

$$O\left(\sqrt{n}\log\frac{\mathbf{x}_0^T\mathbf{s}_0}{\varepsilon}\right)$$

iterations.

Proof Let $(\mathbf{x}_k,\mathbf{y}_k,\mathbf{s}_k)$ denote the point given by the algorithm in the k^{th} iteration. According to Lemma 8, the following inequality holds for the duality gap in the k^{th} iteration:

$$\frac{\mathbf{x}_k^T\mathbf{s}_k}{n}=\mu_k\le\mu_{k-1}\left(1-\sqrt{\frac{\beta\tau}{n}}\left[\tfrac{8}{9}(1-\tau)-\sqrt{\tau\beta}\right]\right)\le\mu_0\left(1-\sqrt{\frac{\beta\tau}{n}}\left[\tfrac{8}{9}(1-\tau)-\sqrt{\tau\beta}\right]\right)^k.$$

From the above inequalities, we get that $\mathbf{x}_k^T \mathbf{s}_k \leq \varepsilon$ holds if

$$\left(1 - \sqrt{\frac{\beta\tau}{n}}\left[\frac{8}{9}(1-\tau) - \sqrt{\tau\beta}\right]\right)^k \mu_0 n \leq \varepsilon$$

is satisfied. Taking the logarithm of both sides, we obtain

$$k \log\left[1 - \sqrt{\frac{\beta\tau}{n}}\left(\frac{8}{9}(1-\tau) - \sqrt{\tau\beta}\right)\right] + \log(\mu_0 n) \leq \log \varepsilon.$$

Using the inequality $-\log(1-\vartheta) \geq \vartheta$, we can require the fulfillment of the stronger inequality

$$-k\sqrt{\frac{\beta\tau}{n}}\left(\frac{8}{9}(1-\tau) - \sqrt{\tau\beta}\right) + \log(\mu_0 n) \leq \log \varepsilon.$$

The last inequality is satisfied when

$$k \geq \sqrt{\frac{n}{\beta\tau}} \frac{1}{\frac{8}{9}(1-\tau) - \sqrt{\tau\beta}} \log\left(\frac{\mathbf{x}_0^T \mathbf{s}_0}{\varepsilon}\right),$$

and this proves the statement. □

In the analysis, we applied the fixed step lengths $\alpha_1 = \sqrt{\frac{\beta\tau}{n}}, \alpha_2 = 1$. When describing the IPA in Algorithm 1, we chose α_1 as the largest value so that the new iterate remains in the neighborhood $\mathcal{W}(\tau, \beta)$. Since the duality gap is strictly decreasing in α_1 and $\sqrt{\frac{\beta\tau}{n}}$ is a lower bound on the value of α_1 in Algorithm 1, its complexity is at least as good as the analyzed case, i.e., the derived complexity result holds for Algorithm 1 as well.

As can be seen from Theorem 1, the investigated method can produce an ε-optimal solution to LP problems in polynomial time. There are different results in the literature on rounding the solutions provided by an IPA to an exact solution in polynomial time, see, e.g., Mehrotra and Ye (1993), Roos et al. (1997).

6 Numerical results

To illustrate that the method can be applied to solve LP problems in practice, we implemented it in Matlab and solved selected linear programming problem instances from the Netlib library (Gay 1985). The numerical experiments were carried out on a Dell laptop with an Intel i7 processor and 16 GB RAM.

First, we transformed the problems to the standard form, then eliminated the redundant constraints using the procedure `eliminateRedundantRows.m` by Ploskas

Table 1 Size of the selected LP instances before and after preprocessing and the time required for the preprocessing and postsolving procedures

Name	m_0	n_0	m	n	Prep.(s)	Posts.(s)
25fv47	821	1876	1856	3712	0.2866	0.0137
adlittle	56	138	139	278	0.0167	0.0032
afiro	27	51	53	106	0.0043	0.0024
agg	488	615	591	1182	0.0345	0.0021
agg2	516	758	758	1516	0.0476	0.0033
agg3	516	758	758	1516	0.0474	0.0015
bandm	305	472	418	836	0.0341	0.0018
beaconfd	173	295	222	444	0.0138	0.0008
blend	74	114	116	232	0.0051	0.0003
bnl1	643	1586	1550	3100	0.1157	0.0060
bore3d	245	346	260	520	0.0195	0.0008
brandy	220	303	247	494	0.0175	0.0011
degen2	444	757	759	1518	0.0427	0.0029
e226	223	472	443	886	0.0177	0.0009
etamacro	617	1033	869	1738	0.0684	0.0034
fffff800	524	1028	1007	2014	0.0577	0.0016
finnis	578	1145	1038	2076	0.0670	0.0014
fit1d	1050	2075	2077	4154	0.4593	0.0012
fit1p	1026	2076	2078	4156	0.3102	0.0004
ganges	1706	2103	1933	3866	0.9625	0.0081
gfrd_pnc	874	1418	1404	2808	0.1534	0.0020
grow15	900	1245	1247	2494	0.2266	0.0049
grow7	420	581	583	1166	0.0458	0.0011
israel	174	316	318	636	0.0074	0.0002
kb2	52	77	79	158	0.0016	0.0003
lotfi	153	366	366	732	0.0078	0.0007
nug05	210	225	227	454	0.0314	0.0023
nug06	372	486	488	976	0.0463	0.0092
nug07	602	931	933	1866	0.1290	0.0244
nug08	912	1632	1634	3268	0.4180	0.0836
recipe	186	299	249	498	0.0118	0.0005
sc105	105	163	164	328	0.0044	0.0011
sc205	205	317	317	634	0.0095	0.0005
sc50a	50	78	79	158	0.0015	0.0002
sc50b	50	78	78	156	0.0014	0.0003
scagr25	471	671	672	1344	0.0367	0.0006
scagr7	129	185	186	372	0.0053	0.0003
scfxm1	330	600	583	1166	0.0276	0.0018
scfxm2	660	1200	1164	2328	0.1050	0.0048

Table 1 continued

Name	m_0	n_0	m	n	Prep.(s)	Posts.(s)
scfxm3	990	1800	1745	3490	0.2532	0.0073
scrs8	490	1275	1270	2540	0.0698	0.0016
sctap1	300	660	662	1324	0.0284	0.0004
scsd1	77	760	762	1524	0.0168	0.0007
scsd6	147	1350	1352	2704	0.0444	0.0013
scsd8	397	2750	2752	5504	0.1928	0.0069
share2b	96	162	164	328	0.0032	0.0002
ship04l	402	2166	1965	3930	0.1103	0.0034
ship08s	778	2467	1714	3428	0.2446	0.0060
standata	479	1394	1350	2700	0.0785	0.0018
standgub	481	1503	1351	2702	0.0779	0.0020
standmps	587	1394	1350	2700	0.1004	0.0026
stocfor1	117	165	159	318	0.0137	0.0006

Table 2 Numerical results for the selected Netlib LP instances

	$\varphi(t) = t$		$\varphi(t) = \sqrt{t}$		$\varphi(t) = t - \sqrt{t}$	
Name	Iter.	Time (s)	Iter.	Time (s)	Iter.	Time (s)
25fv47	89	1008.5749	83	912.0627	89	981.3782
adlittle	22	0.3030	22	0.2779	22	0.3256
afiro	16	0.0392	16	0.0422	16	0.0437
agg	39	17.7995	35	15.4779	39	17.2854
agg2	36	31.2053	35	29.9788	36	31.4254
agg3	45	38.9589	40	34.6220	44	38.4465
bandm	43	7.9849	39	7.0379	42	7.4236
beaconfd	19	0.8271	18	0.6353	19	0.6934
blend	19	0.1517	19	0.1713	20	0.1716
bnl1	105	689.9010	94	619.1983	104	683.0412
bore3d	50	2.6550	46	2.5350	50	2.5969
brandy	44	2.1004	41	1.9017	44	2.0936
degen2	25	20.3961	24	19.6907	25	20.8438
e226	45	8.6461	42	7.9522	45	8.6376
etamacro	96	113.5285	84	105.7878	94	117.6574
fffff800	83	155.3398	75	140.5098	81	150.2565
finnis	72	145.8691	65	132.8389	72	148.3899
fit1d	37	552.6591	35	519.2029	37	545.2137
fit1p	39	576.3481	36	532.4356	38	559.8121
ganges	30	363.0230	29	352.3398	30	368.3899
gfrd_pnc	36	172.6343	32	154.3308	36	184.8429
grow15	31	132.5574	28	120.7086	32	136.6458

Table 2 continued

	$\varphi(t) = t$		$\varphi(t) = \sqrt{t}$		$\varphi(t) = t - \sqrt{t}$	
Name	Iter.	Time (s)	Iter.	Time (s)	Iter.	Time (s)
grow7	30	15.9793	28	15.3579	29	15.6023
israel	49	5.3376	45	4.7431	48	5.1576
kb2	29	0.1273	26	0.1113	29	0.1305
lotfi	31	4.7294	30	4.6403	31	4.7940
nug05	12	0.5650	13	0.7227	12	0.5866
nug06	14	4.8167	14	4.6753	14	4.8035
nug07	21	40.7705	21	40.8739	21	40.4843
nug08	17	162.6864	17	162.1375	16	153.6536
recipe	20	1.0631	20	0.9909	20	1.0163
sc105	16	0.3212	15	0.2657	16	0.3036
sc205	18	1.7174	18	1.6634	18	1.8461
sc50a	15	0.0692	15	0.0646	15	0.0637
sc50b	13	0.0515	13	0.0558	13	0.0575
scagr25	30	18.3103	29	17.5205	30	18.4399
scagr7	22	0.5123	21	0.5026	22	0.5148
scfxm1	61	26.1691	56	23.0922	61	26.1172
scfxm2	69	196.5458	64	200.2992	70	242.9177
scfxm3	72	790.3487	65	720.7563	71	785.0079
scrs8	69	305.1052	63	279.3418	69	308.3720
sctap1	54	40.1256	52	38.5267	55	40.8623
scsd1	17	18.6181	17	19.3082	17	18.7609
scsd6	22	120.3443	23	125.8426	23	125.9444
scsd8	19	817.8832	20	858.1465	20	854.7997
share2b	22	0.4630	22	0.5933	22	0.4665
ship04l	30	473.9215	30	473.8443	30	474.7306
ship08s	41	437.4367	38	401.7406	40	421.5452
standata	45	238.3908	44	232.4935	45	238.2968
standgub	45	239.3328	44	234.3439	45	239.0084
standmps	53	280.6806	51	268.9258	50	280.6102
stocfor1	33	0.6242	31	0.6198	32	0.6079
Average:	38.6538	159.3183	36.2115	150.8065	38.4423	159.8292

and Samaras (2017). After these reformulations, we applied a similar method to procedure CLEAN from Adler et al. (1989) to eliminate fix-valued variables from the linear programming problems.

To be able to give strictly feasible initial points in the neighborhood $\mathcal{W}(\tau, \beta)$, we first transformed the problems into symmetric form and then applied the self-dual embedding technique (Ye et al. 1994). To avoid doubling the number of constraints in

the first case, we carried out this reformulation according to the last Remark of Jansen et al. (1994, p. 232).

The numbers of rows and columns of the original LP problems (in standard form) are denoted by m_0 and n_0, respectively, while the sizes after the reformulations and the embedding procedure are denoted by m and n. These are shown in the second to fifth columns of Table 1. We note that m_0 and n_0 differ from the number of rows and columns given on the Netlib site, since the original formulation of the Netlib LP problems possibly contains lower and upper bounds on the variables. In these cases, the sizes were modified when we reformulated the problems in standard form. The times required to clean and embed the problem (preprocessing) and retrieve the solution of the original optimization problem (postsolve) are also shown in Table 1, in the columns "Prep. (s)" and "Posts. (s)".

For the embedded problem, we may choose $\mathbf{x} = \mathbf{e}$ and $\mathbf{s} = \mathbf{e}$ as proper initial points, since they are strictly feasible and are included in the neighborhood $\mathcal{W}(\tau, \beta)$. The step lengths α_1 and α_2 were calculated in the following greedy way. We fixed the value of α_2 as 1 and determined the largest value α_1 so that the new point $(\mathbf{x}(\alpha), \mathbf{y}(\alpha), \mathbf{s}(\alpha))$ remains in the neighborhood $\mathcal{W}(\tau, \beta)$.

We compared three variants of Algorithm 1, based on the functions $\varphi(t) = t$, $\varphi(t) = \sqrt{t}$ and $\varphi(t) = t - \sqrt{t}$. The first IPA is a moderately modified version of the original method of Ai and Zhang (2005) (we used a slightly different neighborhood definition $\mathcal{W}(\tau, \beta)$). The second case is the IPA proposed by Darvay and Takács (2018). The third IPA is the algorithm introduced in this paper.

The value of the precision parameter ε was 10^{-6}. The number of iterations and the running time (in seconds) required to achieve this precision (i.e., to find a point for which the duality gap is less than ε) for the different algorithm variants are shown in Table 2.

According to our numerical results, there is no significant difference in the performance of the three algorithms for linear programming problems; however, the second variant is moderately better on this test problem set, both in terms of average number of iterations and average running time. It can also be observed that the new variant performs slightly better than the algorithm based on the function $\varphi(t) = t$.

7 Conclusion

We investigated a new long-step IPA based on the algebraic equivalent transformation technique, using the function $\varphi(t) = t - \sqrt{t}$ and a new Ai-Zhang-type wide neighborhood $\mathcal{W}(\tau, \beta)$.

We proved that the algorithm is well-defined and provides an ε-optimal solution in at most $O\left(\sqrt{n}\log\left(\frac{\mathbf{x}_0^T\mathbf{s}_0}{\varepsilon}\right)\right)$ steps, therefore, it has the same theoretical complexity as the best short-step variants. According to our preliminary numerical results, the new algorithm performs well in practice.

To extend our results, we would like to propose a similar long-step algorithm for $\mathcal{P}_*(\kappa)$ linear complementarity problems, based on the function $\varphi(t) = t - \sqrt{t}$. We

expect that the choice of function φ will cause significant difference in the performance of the different variants.

Another interesting question for further research is investigating an infeasible variant of the proposed IPA to avoid applying the self-dual embedding technique when determining the starting points.

Acknowledgements This research was supported by the Hungarian Research Fund, OTKA (Grant No. NKFIH 125700) and the NRDI Fund (TKP2020 NC, Grant No. BME-NC) based on the charter of bolster issued by the NRDI Office under the auspices of the Ministry for Innovation and Technology.
Furthermore, the work of Anita Varga was supported by the ÚNKP-20-3 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund.

Funding Open access funding provided by Budapest University of Technology and Economics.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Adler I, Karmarkar N, Resende MG, Veiga G (1989) Data structures and programming techniques for the implementation of Karmarkar's algorithm. ORSA J Comput 1(2):84–106. https://doi.org/10.1287/ijoc.1.2.84

Ai W, Zhang S (2005) An $O(\sqrt{n}L)$ iteration primal-dual path-following method, based on wide neighborhoods and large updates, for monotone LCP. SIAM J Optim 16(2):400–417. https://doi.org/10.1137/040604492

Bai Y, Lesaja G, Roos C, Wang G, El Ghami M (2008) A class of large-update and small-update primal-dual interior-point algorithms for linear optimization. J Optim Theory Appl 138(3):341–359. https://doi.org/10.1007/s10957-008-9389-z

Dantzig GB (1951) Maximization of a linear function of variables subject to linear inequalities. Activity Anal Prod Alloc 13:339–347

Darvay ZS (2003) New interior point algorithms in linear programming. Adv Model Optim 5(1):51–92

Darvay ZS, Takács PR (2018) Large-step interior-point algorithm for linear optimization based on a new wide neighbourhood. Cent Eur J Oper Res 26(3):551–563. https://doi.org/10.1007/s10100-018-0524-0

Darvay ZS, Papp IM, Takács PR (2016) Complexity analysis of a full-Newton step interior-point method for linear optimization. Period Math Hung 73(1):27–42. https://doi.org/10.1007/s10998-016-0119-2

Darvay ZS, Illés T, Kheirfam B, Rigó PR (2020) A corrector-predictor interior-point method with new search direction for linear optimization. Cent Eur J Oper Res 28(3):1123–1140. https://doi.org/10.1007/s10100-019-00622-3

Darvay ZS, Illés T, Povh J, Rigó PR (2020) Feasible corrector–predictor interior-point algorithm for $\mathcal{P}_*(\kappa)$-linear complementarity problems based on a new search direction. SIAM J Optim 30(3):2628–2658. https://doi.org/10.1137/19M1248972

Darvay ZS, Illés T, Majoros CS (2021) Interior-point algorithm for sufficient LCPs based on the technique of algebraically equivalent transformation. Optim Lett 15(2):357–376. https://doi.org/10.1007/s11590-020-01612-0

Feng Z, Fang L (2014) A new $O(nL)$-iteration predictor–corrector algorithm with wide neighborhood for semidefinite programming. J Comput Appl Math 256:65–76. https://doi.org/10.1016/j.cam.2013.07.011

Gay DM (1985) Electronic mail distribution of linear programming test problems. Math Program Soc COAL Newsl 13:10–12

Jansen B, Roos C, Terlaky T (1994) The theory of linear programming: skew symmetric self-dual problems and the central path. Optimization 29(3):225–233. https://doi.org/10.1080/02331939408843952

Karmarkar N (1984) A new polynomial-time algorithm for linear programming. In: Proceedings of the sixteenth annual ACM symposium on theory of computing, pp 302–311. https://doi.org/10.1145/800057.808695

Khachian LG (1979) A polynomial algorithm in linear programming. Sov Math Dokl 20:191–194 (**English traslation**)

Kheirfam B, Haghighi M (2016) A full-Newton step feasible interior-point algorithm for $\mathcal{P}_*(\kappa)$-LCP based on a new search direction. Croat Oper Res Rev 277–290

Kojima M, Mizuno S, Yoshise A (1989) A primal-dual interior point algorithm for linear programming. In: Progress in mathematical programming. Springer, pp 29–47. https://doi.org/10.1007/978-1-4613-9617-8_2

Li Y, Terlaky T (2010) A new class of large neighborhood path-following interior point algorithms for semidefinite optimization with $O\left(n \log \frac{Tr(X^0 S^0)}{\varepsilon}\right)$ iteration complexity. SIAM J Optim 20(6):2853–2875. https://doi.org/10.1137/080729311

Liu C, Liu H, Cong W (2011) An $O(\sqrt{n}L)$ iteration primal–dual second-order corrector algorithm for linear programming. Optim Lett 5(4):729–743. https://doi.org/10.1007/s11590-010-0242-6

Mehrotra S, Ye Y (1993) Finding an interior point in the optimal face of linear programs. Math Program 62(1):497–515. https://doi.org/10.1007/BF01585180

Peng J, Roos C, Terlaky T (2002) Self-regularity: a new paradigm for primal–dual interior-point algorithms. Princeton series in applied mathematics. Princeton University Press, Princeton. https://doi.org/10.2307/j.ctt7sf0f

Pirhaji M, Mansouri H, Zangiabadi M (2017) An $O\left(\sqrt{n}L\right)$ wide neighborhood interior-point algorithm for semidefinite optimization. Comput Appl Math 36(1):145–157. https://doi.org/10.1007/s40314-015-0220-9

Ploskas N, Samaras N (2017) Linear programming using MATLAB®, vol 127. Springer, Berlin. https://doi.org/10.1007/978-3-319-65919-0

Potra FA (2004) A superlinearly convergent predictor–corrector method for degenerate LCP in a wide neighborhood of the central path with $O(\sqrt{n}L)$ iteration complexity. Math Program 100(2):317–337. https://doi.org/10.1007/s10107-003-0472-9

Potra FA (2014) Interior point methods for sufficient horizontal LCP in a wide neighborhood of the central path with best known iteration complexity. SIAM J Optim 24(1):1–28. https://doi.org/10.1137/120884341

Renegar J (2001) A mathematical view of interior-point methods in convex optimization. SIAM, Philadelphia

Roos C, Terlaky T, Vial JP (1997) Theory and algorithms for linear optimization: an interior point approach. Wiley, Chichester

Sonnevend Gy (1986) An "analytical centre" for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. In: System modelling and optimization. Springer, pp 866–875. https://doi.org/10.1007/BFb0043914

Yang X, Zhang Y, Liu H (2016) A wide neighborhood infeasible-interior-point method with arc-search for linear programming. J Appl Math Comput 51(1–2):209–225. https://doi.org/10.1007/s12190-015-0900-z

Ye Y, Todd MJ, Mizuno S (1994) An $O\left(\sqrt{n}L\right)$-iteration homogeneous and self-dual linear programming algorithm. Math Oper Res 19(1):53–67. https://doi.org/10.1287/moor.19.1.53

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.