A ROBUST INTERIOR POINT METHOD FOR COMPUTING THE ANALYTIC CENTER OF AN ILL-CONDITIONED POLYTOPE WITH ERRORS[*]

Zhouhong Wang

School of Science, Beijing Jiaotong University, Beijing 100044, China
Email: wangzhh@bjtu.edu.cn

Yuhong Dai

LESC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences,
Beijing 100190, and School of Mathematical Sciences, University of Chinese Academy of Sciences,
Beijing 100049, China
Email: dyh@lsec.cc.ac.cn

Fengmin Xu

School of Economics and Finance, Xi'an Jiaotong University, Xi'an 710061, China
Email: fengminxu@mail.xjtu.edu.cn

Abstract

In this paper we propose an efficient and robust method for computing the analytic center of the polyhedral set $P = \{x \in R^n \mid Ax = b, x \geq 0\}$, where the matrix $A \in R^{m \times n}$ is ill-conditioned, and there are errors in A and b. Besides overcoming the difficulties caused by ill-conditioning of the matrix A and errors in A and b, our method can also detect the infeasibility and the unboundedness of the polyhedral set P automatically during the computation. Detailed mathematical analyses for our method are presented and the worst case complexity of the algorithm is also given. Finally some numerical results are presented to show the robustness and effectiveness of the new method.

Mathematics subject classification: 65K05, 90C51.
Key words: Analytic center, Ill-conditioning, Unboundedness, Primal-dual interior point algorithm, Convergence, Polynomial complexity.

1. Introduction

The computation of the analytic center of a polytope has been discussed in many papers and books as the development of interior point methods (IPMs) for linear programming (LP), see, e. g., [1,9,30] and the references therein. In this paper, we will consider the following problem:

$$
\begin{aligned}
\min \quad & -\sum_{j=1}^{n} \ln x_j \\
\text{s.t.} \quad & Ax = b, \\
& x_j > 0, \quad j = 1, \ldots, n,
\end{aligned}
\tag{1.1}
$$

where the matrix $A \in R^{m \times n}$ is ill-conditioned, $b \in R^m$, and there are errors in A and b which may lead to an infeasible problem. Problem (1.1) comes from the study of determining the neutron energy spectrum from multiple activation foils in nuclear physics which is based on the

maximum entropy principle of thermodynamics and Boltzmann's entropy formula, see [6, 35] for more details.

In fact, problem (1.1) is equivalent to the computation of the analytic center of the polyhedral set $P = \{x \in R^n \mid Ax = b, x \geq 0\}$ when P is solid (i.e., the set has at least one positive feasible point) and bounded. Theory of IPMs for LP has established abundant results on the computation of analytic center (see, e.g., [5, 24, 28, 30] and the references therein). If we solve the problem (1.1) directly by some kind of interior point method, a linear system $AD^2A^Ty = h$ or some equivalent linear systems will be formed with ill-conditioned matrices A and D. Although M. H. Wright [27] and S. J. Wright [29] have shown that the ill-conditioning of matrix D does not noticeably impair the accuracy of the computed primal-dual steps in certain cases, the ill-conditioning of matrix A can cause serious numerical difficulties during computation, let alone the errors in A and b. We have used some optimization tools in MATLAB which are based on LIPSOL [34] and IPOPT [25] to solve some instances of problem (1.1), but their numerical performances are unsatisfactory (which will be shown by numerical experiments in Section 4). Despite the fact that some precondition techniques (see, e.g., [19, 20] and references therein) can effectively reduce the ill-conditioning of matrix A, the errors in A and b will still prevent problem (1.1) from being solved effectively.

Based on the primal-dual infeasible-interior-point (IIP) method, we will propose an efficient and robust method for problem (1.1) which can overcome the difficulties caused by ill-conditioning of the matrix A and errors in A and b. The paper is organized as follows. In Section 2 we reformulate problem (1.1) into a problem which is equivalent to finding the analytic center of the optimal solution set of a usually well-conditioned linear programming (LP) problem. The reformulation not only can reduce the impact of errors in A and b but also can improve the condition number and detect the infeasibility of the set P. Then in Section 3 we will consider a hybrid primal-dual IIP algorithm for computing the analytic center of the optimal solution set of an LP problem. Moreover, we will study a method in detail for detecting unboundedness or infeasibility of the set P which may cause computational difficulty for general interior point algorithms. The convergent properties and the worst case complexity of the algorithm are also analysed. Finally some numerical results are presented to show the effectiveness and robustness of our algorithm in Section 4, and some conclusions and remarks are given in Section 5.

2. Reformualtion

If the polyhedral set $P = \{x \in R^n \mid Ax = b, x \geq 0\}$ is unbounded or not solid, problem (1.1) will have no optimal solution. Even when the set P is solid and bounded, the ill-conditioning of matrix A will cause serious numerical difficulties in computation, and errors in A and b will prevent us from using precondition techniques directly. Hence some reformulation of problem (1.1) is needed.

Various reformulations for ill-posed problems have been proposed in literature, such as Tikhonov regularization method and trust region method (see, e.g., [26]). However, these methods can not be used directly to problem (1.1) because of the errors. Hence we consider the following first phase problem for LP:

$$
\begin{aligned}
\min \quad & e^Ty \equiv \sum_{i=1}^{m} y_i \\
\text{s.t.} \quad & Ax + y = b, \\
& x \geq 0, \ y \geq 0.
\end{aligned}
\tag{2.1}
$$

where $A \in R^{m \times n}, e = (1, \dots, 1)^T \in R^m, x = (x_1, \dots, x_n)^T \in R^n, y = (y_1, \dots, y_m)^T \in R^m$. If we balance the matrix A in advance, the coefficients matrix $[A, I]$ is usually well conditioned when $m \ll n$. In fact the following result has been proved in [6].

Lemma 2.1. *Suppose $A = (a_{ij})_{m \times n}$ and $\sum_{j=1}^{n} |a_{ij}| \leq 1$ for every $i = 1, \dots, m$. Denote $B = [A, I]$. Then the condition $\kappa_2(B)$ of the matrix B satisfies*

$$\kappa_2(B) \leq \sqrt{m+1}.$$

Hence problem (2.1) is usually well-conditioned when $m \ll n$ by Lemma 2.1. Furthermore, problem (2.1) can also reduce the impact of the errors in A and b by introducing the vector variable y. In practical computation of determining neutron energy spectrum, the vector b is usually much larger than desired. Hence we require that $y \geq 0$. The vector y can be seen as a correction for the errors in A and b. In fact, when the polyhedral set $P = \{x \in R^n \mid Ax = b, x \geq 0\}$ is bounded and solid, the solution of problem (1.1) is just the analytic center of the optimal solution set of problem (2.1). For completeness, we present the definition of analytic center here, which can be found in many books on IPMs, see, e. g., [5, 24, 30].

Definition 2.1 (Analytic center). *Let the nonempty and bounded set $\mathcal{T}$ be the intersection of an affine space in R^n with the nonnegative orthant of R^n. Define the support $\sigma(\mathcal{T})$ of $\mathcal{T}$ as the subset of the full index set $\{1, 2, \dots, n\}$ given by $\sigma(\mathcal{T}) = \{i : \exists x \in \mathcal{T}$ such that $x_i > 0\}$. The analytic center of $\mathcal{T}$ is defined as the zero vector if $\sigma(\mathcal{T})$ is empty; otherwise it is the vector in $\mathcal{T}$ that maximizes the product*

$$\prod_{i \in \sigma(\mathcal{T})} x_i, \ x \in \mathcal{T}.$$

When the set $\mathcal{T}$ is *unbounded* or *empty*, for simplicity we will say that the analytic center of $\mathcal{T}$ *does not exist*, which is not in accordance with the *generalized analytic center* defined in [18].

Denote

$$
\begin{aligned}
P &= \{x \in R^n \mid Ax = b, \ x \geq 0\}, \\
Q &= \{(x, y) \in R^n \times R^m \mid Ax + y = b, \ x \geq 0, \ y \geq 0\}.
\end{aligned}
\tag{2.2}
$$

It is obvious that $P \subset Q$ in x-part. Without loss of generality, we will assume that $\mathbf{b} > \mathbf{0}$ afterwards[1].

Lemma 2.2. 1. *The polyhedral set Q is solid, i.e., it has a positive feasible point.*

2. *Denote the optimal objective value and the optimal solution set of problem (2.1) by g^* and $\mathcal{P}^*$ respectively. Then $\mathcal{P}^* \neq \emptyset$. Moreover, we have $P \neq \emptyset$ if and only if $g^* = 0$.*

3. *If P is nonempty and bounded, then the set P is the same as the set $\mathcal{P}^*$ in x-part, and so do their the analytic centers.*

Proof. 1. By $b > 0$ we can find a positive integer p large enough such that $A\left(\frac{e}{p}\right) < b$, where $e = (1, 1, \dots, 1)^T \in R^n$. Then $x_0 = \frac{e}{p}, y_0 = b - Ax_0$ is a positive feasible point of Q, and we get the result.

---

[1] First we may assume that $b \neq 0$ (Otherwise the set P is a cone and the analytic center does not exist for nontrival case). If $b_i < 0$ for some i, multiply the equation by -1; if $b_i = 0$ for some i, add some equation with $b_k > 0$ to this equation and we get $0 < b_i \leftarrow b_i + b_k$.

2. Since problem (2.1) is feasible and bounded below, we have $\mathcal{P}^* \neq \emptyset$. By the two-phase simplex method for LP, we know that $P \neq \emptyset$ if and only if $g^* = 0$.

3. It is obvious by item 2 and the definition of analytic center. $\square$

In fact, when we do not know whether P is solid, problem (1.1) should be described as

$$
\begin{aligned}
\min \quad & -\sum_{j \in B} \ln x_j \\
\text{s.t.} \quad & Ax = b, \\
& x_j \geq 0, \quad j = 1, \ldots, n.
\end{aligned}
\tag{2.3}
$$

where $B = \{i \mid \text{there exists some } x \in P \text{ such that } x_i > 0\}$. When P is nonempty and bounded, problem (2.3) is equivalent to finding the analytic center of the optimal solution set $\mathcal{P}^*$ of problem (2.1) by Lemma 2.2. By Lemma 2.1 it should be more advisable to compute the analytic center of the optimal solution set $\mathcal{P}^*$ of problem (2.1) by some kind of primal-dual interior point method instead of solving problem (1.1) directly for numerical stability. The dual problem of the problem (2.1) is

$$
\begin{aligned}
\max_{w,r,s} \quad & b^T w \\
\text{s.t.} \quad & A^T w + r = 0, \\
& w + s = e, \\
& r \geq 0, \ s \geq 0.
\end{aligned}
\tag{2.4}
$$

where $w \in R^m, r \in R^n, s \in R^m$ are the corresponding dual variables of the problem (2.1).

Lemma 2.3. *Denote the optimal solution sets of problems* (2.1) *and* (2.4) *by* $\mathcal{P}^*$ *and* $\mathcal{D}^*$ *respectively. Then* $\mathcal{D}^*$ *is a nonempty convex compact set, and* $\mathcal{P}^*$ *is bounded if and only if the dual problem* (2.4) *has a positive feasible point.*

Proof. By Lemma 2.2 and the strong duality theorem for LP we know that $\mathcal{D}^*$ is nonempty. Since both problems (2.1) and (2.4) are feasible, and problem (2.1) has a positive feasible point, by Corollaries II.11 and II.12 in [24] we get the results. $\square$

Corollary 2.1. *If P is unbounded, the dual problem* (2.4) *has no positive feasible point.*

Proof. Since P is unbounded, there exists some $d \in R^n$ such that $d \neq 0, Ad = 0$, and $d \geq 0$. For any $(\bar{x}, \bar{y}) \in \mathcal{P}^*$ and $\lambda \geq 0$, we have $(\bar{x} + \lambda d, \bar{y}) \in \mathcal{P}^*$. Hence $\mathcal{P}^*$ is unbounded. Then by Lemma 2.3 we get the result. $\square$

According to Corollary 2.1, the dual problem (2.4) will have no positive feasible point when P is unbounded, which may cause numerical difficulties for primal-dual interior-point methods.

Lemma 2.4. *If problem* (2.1) *has a feasible solution* (w_0, r_0, s_0) *with* $b^T w_0 > 0$, *then the set P is empty, and problem* (1.1) *is infeasible.*

Proof. Let $(\bar{x}, \bar{y})$ be an optimal solution of problem (2.1), and g^* be the optimal value. Then we have $g^* = e^T \bar{y} \geq b^T w_0 > 0$, and by item 2 of Lemma 2.2 we get the result. $\square$

By Lemma 2.4, if we find a feasible point $(\bar{w}, \bar{r}, \bar{s})$ of problem (2.4) with $b^T \bar{w} > 0$ (which can be detected during the computation), the set P is empty. However, when the set P is empty, the dual problem (2.4) might have a positive feasible point or not, and we have the following result.

Lemma 2.5. *The dual problem* (2.4) *have no positive feasible point if and only if there exists a vector $p \in R^n$ such that*

$$Ap = 0, \ p \geq 0, \ p \neq 0, \tag{2.5}$$

and the set P is either empty or unbounded in this case.

Proof. The dual problem (2.4) have no positive feasible point is equivalent to the inequality $A^T w > 0$ has no solution. Then by Gordan's theorem (see, e. g., Corollary 1 of Theorem 2.4.5 in [2]) we get (2.5).

By Lemmas 2.2 and 2.3, the set P cannot be "nonempty and bounded" when the dual problem (2.4) have no positive feasible point. If P is nonempty, there exists some $\bar{x} \in P$. Then by (2.5) we have $\bar{x} + \lambda p \in P$ for any $\lambda \geq 0 (p \neq 0)$. Hence P is unbounded. $\square$

By the above results, we know that problem (1.1) is equivalent to finding the analytic center of the optimal solution set of the LP problem (2.1), which is usually well-conditioned when $m \ll n$. We will discuss how to compute the analytic center of the optimal solution set of (2.1) and detect the unboundedness of P during computation in the next section.

3. The Algorithm and Its Analysis

By the theory of primal-dual IPMs for LP, we know that the central path converges to the analytic center of the optimal solution set. The basic primal-dual interior algorithm was proposed by Kojima, Mizuno and Yoshise [11]. The most efficient implementations of IPMs seem to be some variants of Mehrotra's infeasible predictor-corrector primal-dual interior point method, see, e. g., [12, 14, 34]. However, it is not guaranteed in theory that the iteration sequences generated by these algorithms converge to the analytic center of the optimal solution set of LP. Mizuno, Todd, and Ye [17] proposed simplified feasible version of Mehrotra's method with $O(\sqrt{n}L)$ iteration complexity (where the parameter L means that we need to reduce the residuals in feasibility and complementarity by a factor of 2^{-L}), which we will call it the "MTY algorithm" in the following text. The analysis of the MTY algorithm theoretically illuminates why Mehrotra's predictor-corrector method is so efficient. Then Ye et al. [31] and Mehrotra [15] independently proved the sequence generated by the MTY algorithm is quadratically convergent in duality gap. Bonnans and Gonzaga [3], Gonzaga and Tapia [8] proved that the iteration sequence generated by the MTY algorithm converges to the analytic center of the optimal solution set. Hence the MTY algorithm should be the best choice for computing the analytic center of the optimal solution set of problem (2.1). The standard method to generate a proper strictly interior point for the MTY algorithm to start is to transform the problem into a self-dual form (see, e.g., [32]). However, in this way the optimal solution obtained by the MTY algorithm would be the analytic center of the optimal solution set of the transformed problem instead of the original problem itself. Hence we need to find a proper starting interior point of problem (2.1) instead of transforming it into some self-dual form in order to solve problem (1.1) by the MTY algorithm. However, in practical computations of problem (2.1) we found out that it was very difficult to get an initial interior point for the MTY algorithm to start, and the techniques proposed in [9] failed in our computations due to the ill-conditioning and errors in A and b. Hence we would use the infeasible-interior-point (IIP) techniques proposed in [12, 14, 34] to design and implement an algorithm to compute such a starting point and overcome the numerical difficulties.

Many papers have analysed the properties of IIP methods. It is Kojima, Megiddo and Mizuno [10] who first proved the global convergence of a primal-dual IIP algorithm for LP. It is Zhang [33] who first proved the worst case iteration complexity $O(n^2L)$ of a primal-dual IIP algorithm. Soon after that, Mizuno [16] also proved that a modification of the algorithm in [10] has $O(n^2L)$ iteration complexity, and proposed a predictor-corrector IIP algorithm with $O(nL)$ iteration complexity which is based on the MTY predictor-corrector algorithm. In the meantime Potra [21] proposed a new IIP predictor-corrector method where the residuals in feasibility and optimality are improved at the same rate, and proved that the algorithm is globally convergent and has $O(nL)$ iteration complexity. Bonnans and Potra [4] presented sufficient conditions for the iteration sequence of infeasible path following algorithms for linear complementarity problems converging to the shifted analytic center (but not the analytic center) of the optimal face, and proposed an IIP algorithm which is quadratically convergent and has $O(nL)$ iteration complexity based on the feasible algorithm proposed in [7]. Roos [23] proposes a path-following full-Newton step infeasible interior-point algorithm for LP and proves the algorithm also has $O(nL)$ iteration complexity.

It is worth pointing out that at the moment the best known complexity bounds $O(nL)$ for IIP methods are all obtained by following a narrow neighborhood of some infeasible central path, and the iteration sequence generated by them usually will not converge to the analytic center of the optimal solution set. It is indicated in [17] that the worst-case number of iterations usually grows as a wider neighborhood of central path is used, and it is one of the ironies of the IPM literature that algorithms which are more efficient in practice often have somewhat worse complexity bounds, as pointed out in [22]. For computational efficiency and robustness, we will use wide neighborhoods of an infeasible central path to design an algorithm, and the convergent properties and the worst case iteration complexity $O((m+n)^2L)$ of the algorithm will be given.

We first present some basic concepts in IPM (for more details, see, e.g., [24,30] and references therein). The optimality condition for problem (2.1) and its dual (2.4) is (which is also called as the *KKT condition*):

$$
\begin{aligned}
&Ax + y = b, \\
&A^T w + r = 0, \\
&w + s = e, \\
&rx = 0, \ r \geq 0, \ x \geq 0, \\
&sy = 0, \ s \geq 0, \ y \geq 0.
\end{aligned}
\tag{3.1}
$$

The perturbed KKT system of the problem (2.1) and its dual (2.4) is:

$$
\begin{aligned}
&Ax + y = b, \\
&A^T w + r = 0, \\
&w + s = e, \\
&rx = \mu e, \ r \geq 0, \ x \geq 0, \\
&sy = \mu e, \ s \geq 0, \ y \geq 0,
\end{aligned}
\tag{3.2}
$$

where the multiplications rx and sy are componentwise, and we will use this notation without indication in the following paragraphs. When the set $P = \{x \in R^n \mid Ax = b, x \geq 0\}$ is nonempty and bounded, by Lemmas 2.2 and 2.3 we know both problems (2.1) and (2.4) have positive feasible points, which are also called as feasible *interior points*. Then by the well-known *interior-point condition* of IPM, i.e., both the primal and the dual problems have positive feasible points, the system (3.2) has a unique solution $z(\mu) = (x(\mu), y(\mu), w(\mu), r(\mu), s(\mu))$ for

every $\mu > 0$, and the curve $z(\mu)$ is smooth for $\mu > 0$ by the implicit function theorem. The curve $z(\mu)$ $(\mu > 0)$ is called as *the central path* of problem (2.1) and its dual (2.4). By the IPM theory (see, e.g., [13]), it converges to the analytic center of the primal-dual optimal solution set, which is the same as that of set P in x-part by Lemma 2.2.

Given a primal-dual interior point $z = (x, y, w, r, s)$ (i.e., we have $x > 0, y > 0, r > 0, s > 0$, which may be feasible or infeasible), the Newton direction of problem (3.2) at the point z is determined by the following linear systems:

$$
\begin{aligned}
A\Delta x + \Delta y &= -r_b, \\
A^T \Delta w + \Delta r &= -r_c, \\
\Delta w + \Delta s &= -r_e, \\
R\Delta x + X\Delta r &= \sigma\hat{\mu}e - rx, \\
Y\Delta s + S\Delta y &= \sigma\hat{\mu}e - sy,
\end{aligned}
\tag{3.3}
$$

where $r_b = Ax + y - b, r_c = A^T w + r, r_e = w + s - e$ are residuals of linear equations, and $\sigma \in [0, 1]$ is a factor to decrease the duality gap, and R, X, Y, S are diagonal matrices of vectors r, x, y, s respectively. Denote $v = \begin{bmatrix} xr \\ ys \end{bmatrix}$, define

$$
\delta(z, \mu) = \delta(x, y, w, r, s, \mu) \equiv \left\| \frac{v}{\mu} - e \right\|,
\tag{3.4}
$$

$$
\mu(z) = \mu(x, y, w, r, s) \equiv \frac{x^T r + y^T s}{n + m},
\tag{3.5}
$$

$$
\delta(z) = \delta(x, y, w, r, s) \equiv \delta(z, \mu(z)),
\tag{3.6}
$$

$$
\mathcal{N}_2(\beta) \equiv \left\{ z = (x, y, w, r, s) \mid \delta(z) \leq \beta \right\},
\tag{3.7}
$$

$$
\mathcal{N}_{-\infty}(\gamma) \equiv \left\{ z = (x, y, w, r, s) \mid (x, y, r, s) > 0, xr \geq \gamma\mu e, ys \geq \gamma\mu e \right\},
\tag{3.8}
$$

where $\mu(z)$ is the average duality gap at point z, the function $\delta(z)$ in (3.6) is used to measure *the proximity* of interior point z to the central path, parameter $\beta \in (0, 1), \gamma \in (0, 1)$, and $\mathcal{N}_2(\beta)$ and $\mathcal{N}_{-\infty}(\gamma)$ in (3.7)-(3.8) are usually called as the narrow neighborhood and the wide neighborhood of the central path respectively.

In primal-dual interior point methods, system (3.3) is first solved to get the Newton direction, then we do line search to get a new point which is in some neighborhood of the central path and to decrease the duality gap. The classical IIP algorithm for problem (2.1) and its dual (2.4) (or equivalently problem (3.1)) can be described as follows:

Algorithm 3.1. The Classical IIP Algorithm for Problem (2.1)

Step 1. Give a point $z^{(0)} = (x^{(0)}, y^{(0)}, w^{(0)}, r^{(0)}, s^{(0)})$ with $x^{(0)} > 0, y^{(0)} > 0, r^{(0)} > 0, s^{(0)} > 0$, which may be feasible or infeasible. Set $\sigma_0 = 1, k = 0$.

Step 2. Let $z = z^{(k)}, \sigma = \sigma_k, \mu = \mu(z^{(k)})$ and solve linear system (3.3) to get the Newton search direction $\Delta^{(k)} = (\Delta x^{(k)}, \Delta y^{(k)}, \Delta w^{(k)}, \Delta r^{(k)}, \Delta s^{(k)})$.

Step 3. Compute a step $\alpha_k > 0$ such that $z^{(k+1)} = z^{(k)} + \alpha_k \Delta^{(k)} \in N_2(\beta)$(or $\in N_{-\infty}(\gamma)$). Determine the parameter σ_{k+1}. Set $k \leftarrow k + 1$ and go to Step 2.

The basic idea of our method is first using IIP method to compute a feasible initial interior point for MTY algorithm to start, then using MTY algorithm to compute the analytic center. However, when primal-dual problems (2.1) and (2.4) have no feasible interior point (which is equivalent to that the dual problem (2.4) has no feasible interior point by Lemma 2.2), system (3.2) will have no solution for any fixed parameter $\mu > 0$ according to the theory of IPM, which may cause numerical difficulties for primal-dual IPMs. By Lemma 2.5 we know that P is either empty or unbounded in this case. Hence we need a method to detect the unboundedness or emptiness of set $P = \{x \in R^n \mid Ax = b, x \geq 0\}$ during computation.

Denote residuals $r_b^{(k)} = Ax^{(k)} + y^{(k)} - b, r_c^{(k)} = A^T w^{(k)} + r^{(k)}, r_e^{(k)} = w^{(k)} + s^{(k)} - e$. Since $z^{(k+1)} = z^{(k)} + \alpha_k \Delta^{(k)}$, by (3.3) we get

$$r_b^{(k+1)} = (1 - \alpha_k) r_b^{(k)} = \nu_{k+1} r_b^{(0)}, \tag{3.9}$$

$$r_c^{(k+1)} = (1 - \alpha_k) r_c^{(k)} = \nu_{k+1} r_c^{(0)}, \tag{3.10}$$

$$r_e^{(k+1)} = (1 - \alpha_k) r_e^{(k)} = \nu_{k+1} r_e^{(0)}, \tag{3.11}$$

where

$$\nu_k = (1 - \alpha_{k-1}) \nu_{k-1} = \prod_{i=0}^{k-1} (1 - \alpha_i), \ \ k = 1, 2, \ldots, \tag{3.12}$$

and $\nu_0 = 1$.

According to Lemma 2.2, let $(x^{(0)}, y^{(0)})$ be a positive feasible point of Q. Let $w^{(0)} = 0, r^{(0)} = e, s^{(0)} = e$. Then $z^{(0)} = (x^{(0)}, y^{(0)}, w^{(0)}, r^{(0)}, s^{(0)})$ is an initial interior point of system (3.2) and we have:

$$
\begin{aligned}
r_b^{(0)} &= Ax^{(0)} + y^{(0)} - b = 0, \\
r_c^{(0)} &= A^T w^{(0)} + r^{(0)} = e, \\
r_e^{(0)} &= w^{(0)} + s^{(0)} - e = 0.
\end{aligned}
\tag{3.13}
$$

By (3.9)–(3.13), at the k-th iteration of Algorithm 3.1 we have

$$r_b^{(k)} = 0, \ r_c^{(k)} = \nu_k e, \ r_e^{(k)} = 0, \quad k = 1, 2, \ldots \tag{3.14}$$

which means that $z^{(k)} = (x^{(k)}, y^{(k)}, w^{(k)}, r^{(k)}, s^{(k)})$ is a feasible solution of system

$$Ax + y = b, \ A^T w + r = \nu_k e, \ w + s = e.$$

Hence we consider the following perturbed system of (3.2):

$$
\begin{aligned}
& Ax + y = b, \\
& A^T w + r = \tau a, \\
& w + s = e, \\
& rx = \mu e, \ r \geq 0, \ x \geq 0, \\
& sy = \mu e, \ s \geq 0, \ y \geq 0,
\end{aligned}
\tag{3.15}
$$

where $a > 0, \tau > 0$, and $\mu > 0$ is fixed. Since the primal-dual linear system "$Ax + y = b, A^T w + r = \tau a, w + s = e, x \geq 0, y \geq 0, r \geq 0, s \geq 0$" has a primal-dual interior point "$x_0 = x^{(0)} > 0, y_0 = y^{(0)} > 0, w_0 = 0, r_0 = \tau a > 0, s_0 = e > 0$", system (3.15) has a unique solution for every $\tau > 0$ by the interior point condition. In fact, it is easy to check that system (3.15) is equivalent to the KKT optimal condition of the following problem:

$$
\begin{aligned}
\min \quad & \tau a^T x + e^T y - \mu \left(\sum_{i=1}^n \ln x_i + \sum_{i=1}^m \ln y_i \right) \\
\text{s.t.} \quad & Ax + y = b, \\
& x \geq 0, \ y \geq 0,
\end{aligned}
\tag{3.16}
$$

whose objective is strongly convex for any $\mu > 0$.

Theorem 3.1. *Let $a \in R^n, a > 0$, and $\mu > 0$ be fixed. Then for every $\tau > 0$, system (3.15) has a unique solution $(x(\tau), y(\tau), w(\tau), r(\tau), s(\tau))$. Denote $F(\tau, x, y) = \tau a^T x + e^T y - \mu(\sum_{i=1}^n \ln x_i + \sum_{i=1}^m \ln y_i)$, $F^*(\tau) = F(\tau, x(\tau), y(\tau))$ and $p(\tau) = a^T x(\tau)$.*

1. *For any $\tau_1 > \tau_2 > 0$, we have $p(\tau_1) < p(\tau_2)$ and $F^*(\tau_1) > F^*(\tau_2)$.*

2. *If the dual problem (2.4) has a positive feasible point (in this case $P = \{x \in R^n \mid Ax = b, x \geq 0\}$ is either empty or bounded), system (3.2) has a unique solution $(\bar{x}, \bar{y}, \bar{w}, \bar{r}, \bar{s})$. Moreover, we have $x(\tau) \to \bar{x}, y(\tau) \to \bar{y}$ as $\tau \downarrow 0$, and $p(\tau) \uparrow a^T \bar{x}, F^*(\tau) \downarrow [e^T \bar{y} - \mu(\sum_{i=1}^n \ln \bar{x}_i + \sum_{i=1}^m \ln \bar{y}_i)]$ as $\tau \downarrow 0$.*

3. *If the dual problem (2.4) has no positive feasible points (in this case $P = \{x \in R^n \mid Ax = b, x \geq 0\}$ is either empty or unbounded), we have $p(\tau) \uparrow +\infty$ and $F^*(\tau) \downarrow -\infty$.*

Proof. As noted above, the linear part of system (3.15) has an interior feasible point. Hence system (3.15) has a unique solution for every $\tau > 0$ by the well known interior-point condition.

1. By the equivalence of system of (3.15) and problem (3.16), and the strong convexity of $F(\tau, x, y)$ on variables (x, y), we have

$$F^*(\tau_1) = \tau_1 a^T x(\tau_1) + e^T y(\tau_1) - \mu\left(\sum_{i=1}^n \ln x_i(\tau_1) + \sum_{i=1}^n \ln y_i(\tau_1)\right)$$

$$< \tau_1 a^T x(\tau_2) + e^T y(\tau_2) - \mu\left(\sum_{i=1}^n \ln x_i(\tau_2) + \sum_{i=1}^m \ln y_i(\tau_2)\right), \qquad (3.17)$$

$$F^*(\tau_2) = \tau_2 a^T x(\tau_2) + e^T y(\tau_2) - \mu\left(\sum_{i=1}^n \ln x_i(\tau_2) + \sum_{i=1}^n \ln y_i(\tau_2)\right)$$

$$< \tau_2 a^T x(\tau_1) + e^T y(\tau_1) - \mu\left(\sum_{i=1}^n \ln x_i(\tau_1) + \sum_{i=1}^m \ln y_i(\tau_1)\right). \qquad (3.18)$$

By adding the above two inequalities we get

$$(\tau_1 - \tau_2)(a^T x(\tau_1) - a^T x(\tau_2)) < 0.$$

Since $\tau_1 > \tau_2$, we get $p(\tau_1) = a^T x(\tau_1) < a^T x(\tau_2) = p(\tau_2)$.

By $\tau_1 > \tau_2$, $a^T x(\tau) > 0$ for any $\tau > 0$ and formula (3.18) we get

$$F^*(\tau_2) < \tau_1 a^T x(\tau_1) + e^T y(\tau_1) - \mu\left(\sum_{i=1}^n \ln x_i(\tau_1) + \sum_{i=1}^m \ln y_i(\tau_1)\right) = F^*(\tau_1).$$

2. By Lemmas 2.2 and 2.3 we know that both problems (2.1) and (2.4) have nonempty interiors in this case. Hence system (3.2) has a unique solution $(\bar{x}, \bar{y}, \bar{w}, \bar{r}, \bar{s})$ for any $\mu > 0$ with $\bar{x} > 0, \bar{y} > 0, \bar{r} > 0, \bar{s} > 0$. Since $(\bar{x}, \bar{y}, \bar{w}, \bar{r} + \tau a, \bar{s})$ is a feasible interior point of system "$Ax + y = b, A^T w + r = \tau a, w + s = e, x \geq 0, y \geq 0, r \geq 0, s \geq 0$" for every $\tau \geq 0$, the perturbed system (3.15) has a unique solution $z(\tau) = (x(\tau), y(\tau), w(\tau), r(\tau), s(\tau))$ for every $\tau \geq 0$ by the interior-point condition, and $z(\tau) = (x(\tau), y(\tau), w(\tau), r(\tau), s(\tau))$ is a continuously differentiable function of τ on $[0, +\infty)$. Hence we have $z(\tau) \to z(0) = (\bar{x}, \bar{y}, \bar{w}, \bar{r}, \bar{s})$ as $\tau \downarrow 0$,

and $x(\tau) \to \bar{x}, y(\tau) \to \bar{y}$ as $\tau \downarrow 0$. Then by the monotonicity of $p(\tau)$ and $F^*(\tau)$ we obtain $p(\tau) \uparrow a^T \bar{x}$ and $F^*(\tau) \downarrow [e^T \bar{y} - \mu(\sum_{i=1}^n \ln \bar{x}_i + \sum_{i=1}^m \ln \bar{y}_i)]$ as $\tau \downarrow 0$.

3. We first prove that $p(\tau) \uparrow +\infty$ by contradiction. If $p(\tau)$ is bounded on $(0, 1]$, by monotonicity of $p(\tau)$ on $(0, 1]$ we have $p(\tau) = a^T x(\tau) \uparrow \bar{p}$ as $\tau \downarrow 0$ for some $\bar{p} > 0$. Hence $0 < \underline{p} = a^T x(1) \le a^T x(\tau) \le \bar{p}$ on $(0, 1]$. Since $a > 0$, there exists a constant $\xi_1 > 0$ such that $0 < x(\tau) \le \xi_1$ on $(0, 1]$. Then we know that $y(\tau)$ is bounded by $Ax(\tau) + y(\tau) = b$ in (3.15). Denote $\psi(t) = t - \mu \ln t$. By monotonicity of $F^*(\tau)$, for any $\tau \in (0, 1]$ we have

$$F^*(1) \ge F^*(\tau) = \tau a^T x(\tau) - \mu \sum_{i=1}^n \ln x_i(\tau) + \sum_{i=1}^m \psi(y_i(\tau)). \tag{3.19}$$

Since $\psi(t) = t - \mu \ln t \ge \mu(1 - \ln \mu)$, $0 < x(\tau) \le \xi_1$ and $0 < \tau a^T x(\tau) \le \bar{p}$ for $\tau \in (0, 1]$, by (3.19) we know there exist constants $\xi 0$ and M such that $0 < \xi_2 \le x(\tau)$ and $\psi(y_i(\tau)) \le M$ for any $\tau \in (0, 1]$ and every $i = 1, \ldots, m$. By the fact that $\psi(t)$ is strictly decreasing in $(0, \mu)$, and strictly increasing in $(\mu, +\infty)$, and $\lim_{t \to 0+} \psi(t) = \lim_{t \to +\infty} \psi(t) = +\infty$, there exist constants $t_2^M > t_1^M > 0$ such that $t_1^M \le y_i(\tau) \le t_2^M$ for any $\tau \in (0, 1]$ and every $i = 1, \ldots, m$. Then by $sy = \mu e$ in (3.15) we know $s(\tau)$ $(\tau \in (0, 1])$ is bounded, and $w(\tau)$ $(\tau \in (0, 1])$ is bounded by $w + s = e$ in (3.15), and $r(\tau)$ $(\tau \in (0, 1])$ is bounded by $A^T w + r = \tau a$ in (3.15). Hence there exists a sequence $\{\tau_k\}_{k=0}^{+\infty} \subset (0, 1]$ with $\tau_k \to 0$ such that $z(\tau_k) = (x(\tau_k), y(\tau_k), w(\tau_k), r(\tau_k), s(\tau_k))$ converges to some point $\hat{z} = (\hat{x}, \hat{y}, \hat{w}, \hat{r}, \hat{s})$ as $k \to +\infty$. By (3.15) and let $k \to +\infty$ we know $\hat{z} = (\hat{x}, \hat{y}, \hat{w}, \hat{r}, \hat{s})$ is a solution of (3.2), and $(\hat{w}, \hat{r}, \hat{s})$ is an interior point of the dual problem (2.4), which is a contradiction. Hence we have $p(\tau) = a^T x(\tau) \uparrow +\infty$ as $\tau \downarrow 0$.

Now let $(x^{(0)}, y^{(0)})$ be an interior point of the set Q described in Lemma 2.2, and denote $x^{(0)} = (x_1^0, \ldots, x_n^0)^T, y^{(0)} = (y_1^0, \ldots, y_m^0)^T$. By Lemma 2.5 we have formula (2.5), and $(x^{(0)} + \lambda p, y^{(0)})$ is feasible for problem (3.16) for any $\lambda \ge 0$. Define $\phi(\tau, \lambda) = F(\tau, x^{(0)} + \lambda p, y^{(0)})$ and $\phi^*(\tau) = \min_{\lambda \ge 0} \phi(\tau, \lambda)$. Then $\phi^*(\tau) \ge F^*(\tau)$, and we only need show $\phi^*(\tau) \to -\infty$ as $\tau \downarrow 0$. Let $J = \{i : p_i > 0\}$, then

$$\phi(\tau, \lambda) = \tau a^T (x^{(0)} + \lambda p) + e^T y^{(0)} - \mu \left[\sum_{i=1}^n \ln(x_i^0 + \lambda p_i) + \sum_{i=1}^m \ln y_i^0 \right]$$

$$= \lambda \tau \sum_{i \in J} a_i p_i - \mu \sum_{i \in J} \ln(x_i^0 + \lambda p_i) + \tau a^T x^{(0)} + C, \tag{3.20}$$

where $C = e^T y^{(0)} - \mu(\sum_{i \notin J} \ln x_i^0 + \sum_{i=1}^m \ln y_i^0)$. It is obvious that $\phi(\tau, \lambda)$ is a strong convex function of λ. Let $\lambda(\tau) = \arg \min_{\lambda \ge 0} \phi(\tau, \lambda)$. Then $\lambda(\tau)$ satisfies:

$$\phi_\lambda'(\tau, \lambda) = \tau \sum_{i \in J} a_i p_i - \mu \sum_{i \in J} \frac{p_i}{x_i^0 + \lambda p_i} = 0. \tag{3.21}$$

Since $\sum_{i \in J} a_i p_i > 0, \mu > 0$, and $p_i > 0$ for every $i \in J$, we know that equation (3.21) has a unique solution $\lambda(\tau) > 0$ for every $\tau > 0$, and $\lambda(\tau) \uparrow +\infty$ as $\tau \downarrow 0$. By (3.20) and (3.21) we get

$$\phi^*(\tau) = \mu \left(\sum_{i \in J} \frac{\lambda(\tau) p_i}{x_i^0 + \lambda(\tau) p_i} - \sum_{i \in J} \ln(x_i^0 + \lambda(\tau) p_i) \right) + \tau a^T x^{(0)} + C,$$

which diverges to $-\infty$ as $\lambda(\tau) \uparrow +\infty$ and $\tau \downarrow 0$. Hence we have $F^*(\tau) \downarrow -\infty$ as $\tau \downarrow 0$. $\qquad \square$

By Theorem 3.1, we could detect the unboundedness or the emptiness of P if we keep an eye on the variation of $F^*(\nu_k)$ and $p(\nu_k)$ when ν_k is small enough and μ is fixed during computation.

Hence we propose the following IIP algorithm for computing an initial feasible interior point where the MTY algorithm can start. In the following algorithm we denote

$$\mu_k \equiv \mu(z^{(k)}) = \frac{(x^{(k)})^T r^{(k)} + (y^{(k)})^T s^{(k)}}{n + m}, \quad \delta_k \equiv \delta(z^{(0)}) = \left\| \begin{bmatrix} \frac{x^{(k)} r^{(k)}}{\mu_k} \\ \frac{y^{(k)} s^{(k)}}{\mu_k} \end{bmatrix} - e \right\| \qquad (3.22)$$

for $k = 0, 1, \ldots$. Comparing with the algorithm proposed in [6], the new algorithm can detect the unboundedness or emptiness of P.

Algorithm 3.2. The IIP Algorithm for an Initial Point of the MTY Algorithm

Step 0a. Balance matrix $A = (a_{ij})_{m \times n}$, i.e., for $i = 1, \ldots, m$, compute $t_i = \sum_{j=1}^n |a_{ij}|$. If $t_i > 1$, set $a_{ij} \leftarrow a_{ij}/t_i$ for $j = 1, \ldots, n$, and $b_i \leftarrow b_i/t_i$.

Step 0b. Compute an interior point $(x^{(0)}, y^{(0)})$ of the set Q according to the method described in the proof of Lemma 2.2, i.e., find a proper positive number h_0 such that $x^{(0)} = h_0 e$ and $y^{(0)} = b - Ax^{(0)} > 0$. Set $w^{(0)} = 0, r^{(0)} = e, s^{(0)} = e$. Set $r_b^{(0)} = 0, r_c^{(0)} = A^T w^{(0)} + r^{(0)} = e, r_e^{(0)} = 0$. Set a convergence tolerance parameter $\varepsilon > 0$. Compute μ_0, δ_0 and $\xi = \min\{\frac{x_j^{(0)} r_j^{(0)}}{\mu_0}, j = 1, \ldots, n; \frac{y_i^{(0)} s_i^{(0)}}{\mu_0}, i = 1, \ldots, m\}$. Set $\gamma \in [10^{-8}, 10^{-3}], \beta > 0, \tau_0 \in [2\varepsilon, 100\varepsilon], M \in [3, 10]$. Set $\sigma_0 = 1, \hat{\mu} = \mu_0, \bar{\mu} = \min\{0.1\mu_0, 0.5\}, \nu_0 = 1, k = 0$.

Step 1. If $\|r_c^{(k)}\|_\infty = \nu_k \leq \varepsilon$ and $b^T w^{(k)} > \varepsilon$, the set P is empty and stop.

If $\delta_k \leq \beta$ with $\|r_c^{(k)}\|_\infty = \nu_k \leq \varepsilon$, stop and return the results, we find the initial point for MTY algorithm to start.

Otherwise, if $\nu_k \leq \tau_0$, and $p_k = e^T x^{(k)}$ still increases a lot, and $F_k = \nu_k p_k + e^T y^{(k)} - \hat{\mu}(\sum_{i=1}^n \ln x_i^{(k)} + \sum_{i=1}^m \ln y_i^{(k)})$ still decreases a lot for M consecutive iterations, the polyhedral set P should be empty or unbounded by Theorem 3.1, stop.

Otherwise go to next step.

Step 2. Let $z = z^{(k)}, \sigma = \sigma_k, r_b = 0, r_c = r_c^{(k)}, r_e = 0$ and solve (3.3) to get the Newton search direction
$$\Delta_k = (\Delta x^{(k)}, \Delta y^{(k)}, \Delta w^{(k)}, \Delta r^{(k)}, \Delta s^{(k)}).$$

Compute
$$t = \min\left\{ (x^{(k)})^{-1} \Delta x^{(k)}, (y^{(k)})^{-1} \Delta y^{(k)}, (r^{(k)})^{-1} \Delta r^{(k)}, (s^{(k)})^{-1} \Delta s^{(k)} \right\},$$
$$\zeta = \frac{-1}{\min\{t, -0.5\}}, \quad \alpha_{max} = \min\{1, 0.9995 * \zeta\}.$$

Step 3. If $\alpha_{max} = 1$, set $z^{(k+1)} = z^{(k)} + \Delta_k$. Otherwise use backtracking technique to compute the smallest nonnegative integer p with $\alpha_k = \theta^p \alpha_{max}$ (where $\theta < 1$) such that
$$z^{(k+1)} = z^{(k)} + \alpha_k \Delta_k \in \mathcal{N}_{-\infty}(\gamma),$$
$$\bar{\mu} \leq \mu(z^{(k+1)}). \qquad (3.23)$$

Moreover, if $\delta_k > \beta$, the new point $z^{(k+1)}$ should also satisfy

$$G(z^{(k+1)}) \leq (1 - 0.01\alpha_k)G(z^{(k)}), \tag{3.24}$$

where $G(z) = \|v - \mu(z)e\|$ with $v = \begin{bmatrix} xr \\ ys \end{bmatrix}$.

Step 4. Compute $\nu_{k+1} = (1 - \alpha_k)\nu_k, r_c^{(k+1)} = \nu_{k+1}e, \mu_{k+1} = \mu(z^{(k+1)})$ and $\delta_{k+1} = \delta(z^{(k+1)})$ according to (3.22).

If $\nu_k \geq \tau_0$ and $\nu_{k+1} < \tau_0$, set $\tilde{\mu} = \max\{\mu_{k+1}, 2\bar{\mu}\}$.

If $\nu_{k+1} < \tau_0$, set $\hat{\mu} = \tilde{\mu}, \sigma_{k+1} = 1$. Otherwise, if $\mu_{k+1} > 10\bar{\mu}$ and $\alpha_k > 0.5$ with $\nu_{k+1} \geq \tau_0$, set $\sigma_{k+1} = 0.2, \hat{\mu} = \mu_{k+1}$; otherwise $\mu_{k+1} \leq 10\bar{\mu}$ or $\alpha_k \leq 0.5$ with $\nu_{k+1} \geq \tau_0$, set $\sigma_{k+1} = 1, \hat{\mu} = \max\{\mu_{k+1}, 2\bar{\mu}\}$. Set $k \leftarrow k + 1$ and go to step 1.

In the above algorithm, we first balance the ill-conditioned matrix A in Step 0a according to Lemma 2.1. In Step 0b we compute an initial infeasible interior point and set parameters, where the parameter ε is the convergence tolerance, and the parameter γ is used to determine the wide neighborhood of the central path $\mathcal{N}_{-\infty}(\gamma)$, and ν_k is the decreasing factor of infeasibility in k-th iteration which is determined by formula (3.12). Since the purpose of Algorithm 3.2 is to compute an initial feasible interior point for MTY algorithm to start, we use the parameter $\bar{\mu}$ to prevent the duality gap μ_k from becoming too small, which may cause numerical difficulties due to the ill-conditioning. In fact the parameters σ_k and $\hat{\mu}$ which are used in system (3.3) will be fixed when ν_k or μ_k is small enough according to Step 4 in the above algorithm. Parameters τ_0 and M are used to determine whether the set P is unbounded or empty according to Theorem 3.1. If we find a feasible point $z^{(k)}$ with $b^T w^{(k)} > \varepsilon$, the set P is empty by the duality theory of LP. By Theorem 3.1, when the dual problem (2.4) has a positive feasible point, the sequences $p_k = e^T x^{(k)}$ and $F_k = \nu_k p_k + e^T y^{(k)} - \sigma_k \hat{\mu}(\sum_{i=1}^n \ln x_i^{(k)} + \sum_{i=1}^m \ln y_i^{(k)})$ will converge as $k \to +\infty$. On the contrary, when the dual problem (2.4) has no positive feasible point, the set P is either empty or unbounded by Lemma 2.5, and we will have $p_k \uparrow +\infty, F_k \downarrow -\infty$ as $k \to +\infty$. Therefor in Step 1 of Algorithm 3.2 we will keep an eye on the sequences $\{p_k\}$ and $\{F_k\}$, and stop immediately when we find the variations of the sequences $\{p_k\}$ and $\{F_k\}$ are still very big for M consecutive iterations when $\nu_k < \tau_0$. We usually set $M = 3$ in the implementation of the algorithm. We could have more certainty about the emptiness or unboundedness of the set P when we use bigger M. However, this parameter should not be too big to affect the performance of the algorithm.

Lemma 3.1. *In Algorithm* 3.2, *for any* $\varepsilon > 0$, *there exists a constant* $C > 0$ *such that* $\alpha_k \geq \frac{C}{(n+m)^2}$ *for all* $k = 0, 1, \ldots$, *where* $\frac{C}{(n+m)^2} < 1$.

The proof of Lemma 3.1 is a little long, but the techniques used are standard in IIP analyses which have been proposed in $[10, 16, 28, 33]$, etc. For brevity and clarity, we put the proof in the appendix. It is worth pointing out that the parameter C is a constant relative to the problem, and it does not depend on m and n (for more details, please read the the proof in the appendix). With the help of Lemma 3.1, we get the following result.

Theorem 3.2. *For any* $\varepsilon > 0$, *Algorithm* 3.2 *will stop in finite steps, and its worst case complexity is* $O((n+m)^2(\ln \frac{1}{\varepsilon}))$.

Proof. By Lemma 3.1, (3.12) and (3.14), we have

$$\nu_k = (1 - \alpha_{k-1})\nu_{k-1} = \prod_{i=0}^{k-1}(1 - \alpha_i) \leq \left[1 - \frac{C}{(n+m)^2}\right]^k,$$

$$\|r_c^{(k)}\|_\infty = \|\nu_k e\|_\infty = \nu_k \leq \left[1 - \frac{C}{(n+m)^2}\right]^k. \tag{3.25}$$

By the fact $\ln(1 + x) \leq x$ for $x > -1$, we have $\left[1 - \frac{C}{(n+m)^2}\right]^k \leq \varepsilon$ when $k \geq \frac{(n+m)^2}{C}\ln\frac{1}{\varepsilon}$. Let $k_1 = \lceil \frac{(n+m)^2}{C}\ln\frac{1}{\varepsilon}\rceil$. Then by (3.25) we have $\nu_k = \|r_c^{(k)}\|_\infty \leq \varepsilon \leq \tau_0$ when $k \geq k_1$. If P is empty or unbounded, the algorithm will stop after another M consecutive iterations according to Step 1 of Algorithm 3.2. Hence the algorithm will stop in $O((n+m)^2(\ln\frac{1}{\varepsilon}))$ iterations. Otherwise the set P is bounded and nonempty, and the generated iteration sequence $\{z^{(k)}\}$ lies in a compact set. We have $\delta_k > \beta$ when $k \geq k_1$ according to Step 1 of Algorithm 3.2 in this case. Then by Step 3 of Algorithm 3.2 for $k > k_1$ we have

$$G(z^{(k)}) \leq (1 - 0.01\alpha_{k-1})G(z^{(k-1)}) \leq \cdots \leq \prod_{j=k_1}^{k-1}(1 - 0.01\alpha_j)G(z^{(k_1)}).$$

By Lemma 3.1 we get

$$G(z^{(k)}) \leq \prod_{j=k_1}^{k-1}(1 - 0.01\alpha_j)G(z^{(k_1)}) \leq \left[1 - \frac{0.01C}{(n+m)^2}\right]^{k-k_1}G(z^{(k_1)}). \tag{3.26}$$

In the same way as above we get $G(z^{(k)}) \leq \beta\bar\mu$ when $k - k_1 \geq \frac{100(n+m)^2}{C}\ln\frac{G(z^{(k_1)})}{\beta\bar\mu}$. Since $\beta, \bar\mu, \varepsilon$ are positive constants, and $G(z^{(k)})$ is bounded on a compact set, we have $\frac{100(n+m)^2}{C}\ln\frac{G(z^{(k_1)})}{\beta\bar\mu} = O((n+m)^2(\ln\frac{1}{\varepsilon}))$. Hence after $O((n+m)^2(\ln\frac{1}{\varepsilon}))$ iterations, by (3.23) and the above formulae we have

$$\delta(z^{(k)}) = \frac{G(z^{(k)})}{\mu(z^{(k)})} \leq \frac{G(z^{(k)})}{\bar\mu} \leq \beta,$$

and the algorithm will stop according to Step 1.   □

After getting an initial point for the MTY algorithm to start, we will implement the MTY algorithm until we get the analytic center of the set P, or find the set P is infeasible. For completeness, we present the full algorithm as follows.

Algorithm 3.3. The Algorithm for Computing the Analytic Center

Step 1. Call Algorithm 3.2 with $\beta = 0.25$. If Algorithm 3.2 get an initial point $z^{(i)} = (x^{(i)}, y^{(i)}, w^{(i)}, r^{(i)}, s^{(i)})$ satisfying $\|r_c^{(i)}\| \leq \varepsilon$ and $\delta(z^{(i)}) \leq \beta = 0.25$ at i-th iteration, set $k = i, \beta = 0.25$ and go to Step 2. Otherwise, stop.

Step 2. If $b^T w^{(k)} > \varepsilon$, by Lemma 2.2 we know $P = \emptyset$, stop.

Step 3. (Predictor) Set $\sigma = 0, z = z^{(k)} = (x^{(k)}, y^{(k)}, w^{(k)}, r^{(k)}, s^{(k)}), \hat\mu = \mu(z)$ and solve (3.3) to get the direction $\overline{\Delta z} = (\overline{\Delta x}, \overline{\Delta y}, \overline{\Delta w}, \overline{\Delta r}, \overline{\Delta s})$.

Step 4. Compute the largest $\hat{\alpha} \in (0, 1]$ such that $\hat{z} = z + \hat{\alpha}\overline{\Delta z} \in \mathcal{N}(2\beta)$, which can be done by solving two quadratic equations in one variable.

Step 5. (Corrector) Set $\sigma = 1, \hat{\mu} = (1 - \hat{\alpha})\mu_k, z = \hat{z}$ and solve (3.3) to get the search direction $\Delta z = (\Delta x, \Delta y, \Delta w, \Delta r, \Delta s)$, and set $\bar{z} = \hat{z} + \Delta z$.

Step 6. If $(n + m)\mu < \varepsilon$, stop and return the solution; Otherwise set $k \leftarrow k + 1, z^{(k)} = \bar{z}$ and go to Step 3.

According to the results gotten in [17], [31], [3] and [8], we have the following results.

Theorem 3.3. *Suppose the set $P = \{x \in R^n \mid Ax = b, x \geq 0\}$ is nonempty and bounded. Then the sequence $z^{(k)} = (x^{(k)}, y^{(k)}, w^{(k)}, r^{(k)}, s^{(k)})$ generated by the above MTY algorithm converges to $z^* = (x^*, y^*, w^*, r^*, s^*)$ which is the analytic center of the primal-dual optimal solution set of problem (2.1) and its dual (2.4) with $y^* = 0$, and x^* is the analytic center of the polytope P. The local convergence rate in duality gap is quadratic, and the worst case complexity of Step 3–Step 6 in Algorithm 3.3 is $O(\sqrt{(n + m)} \ln \frac{1}{\varepsilon})$.*

4. Some Numerical Results

In this section we will present some numerical results for academic purpose. All test problems are constructed by using the well-known Hilbert matrices, which are seriously ill-condtioned as the dimension increases. All the numerical tests are done under the MATLAB environment in Windows 7 system on a laptop computer. The MATLAB version we used is R2018b. The parameters of Algorithm 3.3 are set as $\varepsilon = 10^{-6}, \tau_0 = 10^{-5}, M = 3, \gamma = 10^{-8}$ for all test problems, and its corresponding numerical results will be denoted by "IIPMTY".

Letting H_m be an $m \times m$ Hilbert matrix, i.e., we have $H = (\frac{1}{i+j-1})_{m \times m}$. First let us consider the following ill-conditioned polytope

$$P_m = \left\{ x \in R^n \mid Ax = b, x \geq 0 \right\},$$

where $A = [H_m, H_m], b = H_m e$, and $n = 2m$. It is easy to check that the analytic center of P_m is $\bar{x} = (0.5, \ldots, 0.5)$.

For comparison, we will use the interior point method used in MATLAB's optimization tool *linprog* to solve the following problem:

$$\begin{aligned} \min \quad & 0^T x \\ \text{s.t.} \quad & Ax = b, \\ & x \geq 0. \end{aligned} \tag{4.1}$$

The corresponding numerical results is denoted by "linprog-1". We also use the interior point method in *linprog* to solve problem (2.1) and denote corresponding numerical results by "linprog-2". We can first use the function "optimoptions" in MATLAB to set the options of the tool *linprog*:

```
options = optimoptions('linprog','Algorithm', 'interior-point',...
'Display', 'iter', 'OptimalityTolerance', 1e-6,'maxiter', 1000);
```

Then we call the optimization tool *linprog* in MATLAB by using the parameter *options*.

We will also use the interior methods in MATLAB's optimization tool *fmincon* to compute the following problem

$$
\begin{aligned}
\min \quad & -\sum_{j=1}^{n} \ln x_j \\
\text{s.t.} \quad & Ax = b, \\
& x \geq 0.
\end{aligned}
\tag{4.2}
$$

The corresponding numerical results will be denoted by "fmincon". The options for *fmincon* is set as follows:

```
options = optimoptions(@fmincon,'SpecifyObjectiveGradient',true, ...
    'Display', 'iter', 'Algorithm', 'interior-point',...
    'OptimalityTolerance', 1e-6, 'maxiter', 1000);
```

which means that *fmincon* will use its interior-point algorithm to compute.

Numerical results for computing the analytic center of P_m are collected in Table 4.1. The column "m" in the table is the dimension of the Hilbert matrix H. The column "Algo." indicates which method is used. The columns f and g are the function values of

$$
f(x) = -\sum_{j=1}^{n} \ln x_j \quad \text{and} \quad g(x) = \|Ax - b\|
$$

at the numerical solution $\bar{x}$ obtained by the corresponding algorithm respectively. The function value $f(\bar{x})$ indicates whether the solution $\bar{x}$ is the analytic center, and $g(\bar{x})$ indicates whether the solution $\bar{x}$ is feasible. In this two columns, the symbol "—" means that the corresponding method failed to obtain a solution, and in column "f" the symbol "$+\infty$" means that there exists at least one x_j whose value is 0 in the solution. The column "No." represents *the iteration number* of the corresponding algorithm needed to obtain the solution $\bar{x}$. The column "Opt." indicates whether the corresponding method stops with its optimality measure being satisfied, with the symbol "Y" meaning "Yes", the symbol "N" meaning "No", and the symbol "P" meaning "possible optimal". The column "Time" is the CPU time (in seconds) needed by the corresponding algorithm to compute the problem.

By the results in Table 4.1, we can see that only for $m = 10$, the method "linprog-1" succeeded in obtaining the optimal solution. For other problems it failed, and at the end of its computation it displayed either

> *Solver stopped prematurely. Linprog stopped because it exceeded the iteration limit, options.MaxIterations = 1000.*

when the iteration No. = 1000, or

> *Linprog stopped because it was unable to find a point that satisfies the constraints within the default value of the constraint tolerance.*

when the iteration No. < 1000.

The method "linprog-2" succeeded in obtaining optimal solutions according to its optimization measure for $m = 10, 20, 50, 100$. But it failed for $m = 300, 500$, and displayed similar messages as the method "linprog-1". Comparing numerical results obtained by "linprog-1" with "linprog-2", we can see that problem (2.1) is much more stable than problem (4.1), as

Table 4.1: Numerical Results for $A_m = [H_m, H_m], b = H_m e$.

m	Algo.	No.	f	g	Opt.	Time(s)
10	linprog-1	6	$+\infty$	4.110e-07	Y	0.235
	linprog-2	5	$+\infty$	1.421e-09	Y	0.243
	fmincon	39	13.863	1.286e-13	Y	1.392
	IIPMTY	11	13.863	8.674e-09	Y	0.071
20	linprog-1	1001	—	—	N	0.285
	linprog-2	5	$+\infty$	1.101e-08	Y	0.266
	fmincon	15	27.726	1.528e-13	Y	0.763
	IIPMTY	13	27.726	3.589e-11	Y	0.092
50	linprog-1	1001	—	—	N	0.820
	linprog-2	6	$+\infty$	6.305e-08	Y	0.241
	fmincon	14	69.315	2.336e-13	Y	1.021
	IIPMTY	16	69.315	1.780e-10	Y	0.127
100	linprog-1	337	—	—	N	1.062
	linprog-2	5	$+\infty$	1.299e-07	Y	0.246
	fmincon	13	138.629	1.483e-13	Y	1.117
	IIPMTY	18	138.629	1.096e-10	Y	0.390
300	linprog-1	1001	—	—	N	41.905
	linprog-2	19	—	—	N	1.401
	fmincon	156	415.888	3.855e-13	N	166.530
	IIPMTY	22	415.888	2.379e-11	Y	3.600
500	linprog-1	676	—	—	N	149.613
	linprog-2	8	—	—	N	3.548
	fmincon	77	693.147	6.652e-13	P	385.377
	IIPMTY	24	693.147	6.288e-11	Y	11.828

indicated by Lemma 2.1. By numerical results obtained by 'linprog-2" in column "f", we can see that the interior method used in *linprog*, which is based on LIPSOL [34], can not guarantee the iteration sequence generated converges to the analytic center of the optimal solution set.

When we used the tool *fmincon* to solve problem (4.2), it also met some numerical difficulties during computation and displayed the following message

> Warning: Matrix is close to singular or badly scaled. Results may be inaccurate.
> RCOND = 1.013015e-016.

or something alike. However, the method *fmincon* still obtained optimal solutions for $m = 10, 20, 50, 100, 300$ according to its optimization measure, i.e., stopped with the "exitflag = 1". For $m = 500$, the method *fmincon* stopped with the "exitflag = 2", and displayed the following message:

> Local minimum possible. Constraints satisfied.

> Change in x was less than options.StepTolerance and maximum constraint violation was less than options.ConstraintTolerance.

For all test problems, the method *fmincon* got the approximate analytic center according to the column f in Table 4.1. However, its CPU time is much more than our method "IIPMTY"

for all test problems. For every test problem, our method "IIPMTY" terminated with the optimal rule being satisfied, and got the approximate analytic center with the least CPU time (in seconds). From the numerical results in Table 4.1, we can see that the performance of our method "IIPMTY" is the most effective and robust one.

Next we consider computing the analytic center of the following polyhedral set

$$\tilde{P}_m = \left\{ x \in R^n \mid Ax = b, x \geq 0 \right\},$$

where $A = [H_m, -H_m], b = H_m e,\ n = 2m$, and $m = 10, 20, 50, 100, 300, 500$. It is easy to check that the polyhedral set $\tilde{P}_m$ is unbounded, hence the analytic center of $\tilde{P}_m$ does not exist. For this kind of problems, we only compare our method "IIPMTY" with "fmincon", since there is no sense to compare it with the methods "linprog-1" and "lingprog-2" for this kind of problems. The method "fmincon" was not able to find the unboundedness of the set $\tilde{P}_m$ for all test problems with $m = 10, 20, 50, 100, 300, 500$, and did a lot of unnecessary computations until they cannot move any further. They also gave some wrong judgements about the solution they got. For example, sometimes the method "fmincon" reported that it find a local minimum, but in fact there is no local minimum by the strong convex properties of problem (4.2). However, for all this kind of test problems with $m = 10, 20, 50, 100, 300, 500$ our method "IIPMTY" succeeded in detecting the unboundedness of the set $\tilde{P}_m$ only by 13 iterations.

5. Conclusions

In this paper we have discussed how to compute the analytic center of the polyhedral set $P = \{x \in R^n \mid Ax = b, x \geq 0\}$, where the matrix $A \in R^{m \times n}$ may be badly ill-conditioned with errors in A and b. After some studies, it is surprising to find that we only need to transform the problem into the well-known first-phase problem in the simplex method for LP, and compute the analytic center of the optimal solution set of the transformed problem. The new method can effectively reduce the ill-conditioning of the coefficient matrix A, and can detect the infeasibility or the unboundedness of the set P during computation. In this new method, we proposed a robust and efficient algorithm for computing the initial point for the MTY algorithm to start. Then we combined it with the well-known MTY algorithm the compute the analytic center of the set P. The mathematical correctness of our method is rigorously proved, and its convergent properties and worst case complexity analysis are presented. Finally, some numerical tests only for academic purposes are given, which show the robustness and effectiveness of our method.

In our computations, we find out that the parameter β in the MTY algorithm sometimes is too restrictive which would slow the convergence rate before the quadratic convergence property of the algorithm is shown. When the analytic center of the optimal solution set is needed, one simple way to improve the efficiency of the MTY algorithm is to combine it with the Mehrotras predictor-corrector method proposed in [14] while ensuring that the generated iteration sequence converges to the analytic center. We will do further studies in this direction.

Appendix A. Proof of Lemma 3.1

Proof. According to Step 0b, Step 3, and Step 4 of Algorithm 3.2, we always have:

$$\bar{\mu} \le \mu_k \le \hat{\mu} \le 2\mu_k, \quad \sigma_k\mu_k \ge 2\bar{\mu}, \quad \sigma_k\hat{\mu} \ge 2\bar{\mu}. \tag{A.1}$$

For simplicity, denote $\tilde{A} = [A, I], \tilde{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \tilde{s} = \begin{bmatrix} r \\ s \end{bmatrix}, \tilde{w} = w, c = \begin{bmatrix} 0 \\ e \end{bmatrix}$. Then the system (3.2) can be described as

$$
\begin{aligned}
\tilde{A}\tilde{x} &= b, \\
\tilde{A}^T\tilde{w} + \tilde{s} &= c, \\
\tilde{x}\tilde{s} &= \mu e, \\
\tilde{x} \ge 0, \tilde{s} &\ge 0,
\end{aligned}
\tag{A.2}
$$

and the corresponding Newton system (3.3) can be described as

$$
\begin{aligned}
\tilde{A}\Delta\tilde{x} &= -\tilde{r}_b, \\
\tilde{A}^T\Delta\tilde{w} + \Delta\tilde{s} &= -\tilde{r}_c, \\
\tilde{s}\Delta\tilde{x} + \tilde{x}\Delta\tilde{s} &= \sigma\hat{\mu}e - \tilde{x}\tilde{s}, \\
\tilde{x} > 0, \tilde{s} &> 0,
\end{aligned}
\tag{A.3}
$$

where $\tilde{r}_b = \tilde{A}\tilde{x} - b, \tilde{r}_c = \tilde{A}^T\tilde{w} + \tilde{s} - c$. Let $(\tilde{x}^{(k)}, \tilde{w}^{(k)}, \tilde{s}^{(k)})$ be the iteration sequence generated by Algorithm 3.2. Denote $\tilde{r}_b^{(k)} = \tilde{A}\tilde{x}^{(k)} - b, \tilde{r}_c^{(k)} = \tilde{A}^T\tilde{w}^{(k)} + \tilde{s}^{(k)} - c$, by Step 3 of Algorithm 3.2 and (A.3) we have

$$
\begin{aligned}
(\tilde{x}^{(k+1)}, \tilde{w}^{(k+1)}, \tilde{s}^{(k+1)}) &= (\tilde{x}^{(k)}, \tilde{w}^{(k)}, \tilde{s}^{(k)}) + \alpha_k(\Delta\tilde{x}^{(k)}, \Delta\tilde{w}^{(k)}, \Delta\tilde{s}^{(k)}), \\
\tilde{r}_b^{(k+1)} &= \tilde{A}\tilde{x}^{(k+1)} - b = \tilde{r}_b^{(k)} - \alpha_k\tilde{r}_b^{(k)} = (1 - \alpha_k)\tilde{r}_b^{(k)}, \\
\tilde{r}_c^{(k+1)} &= \tilde{A}^T\tilde{w}^{(k+1)} + \tilde{s}^{(k+1)} - c = \tilde{r}_c^{(k)} - \alpha_k\tilde{r}_c^{(k)} = (1 - \alpha_k)\tilde{r}_c^{(k)}.
\end{aligned}
\tag{A.4}
$$

Denote $\mathcal{A} = \{(x, w, s) \mid \tilde{A}x = b, \tilde{A}^Tw + s = c\}$. With $(\hat{x}^{(0)}, \hat{w}^{(0)}, \hat{s}^{(0)}) \in \mathcal{A}$, we construct the following feasible auxiliary sequences according to [33]:

$$
\begin{aligned}
\hat{x}^{(k+1)} &= \hat{x}^{(k)} + \alpha_k(\tilde{x}^{(k)} + \Delta\tilde{x}^{(k)} - \hat{x}^{(k)}), \\
\hat{w}^{(k+1)} &= \hat{w}^{(k)} + \alpha_k(\tilde{w}^{(k)} + \Delta\tilde{w}^{(k)} - \hat{w}^{(k)}), \\
\hat{s}^{(k+1)} &= \hat{s}^{(k)} + \alpha_k(\tilde{s}^{(k)} + \Delta\tilde{s}^{(k)} - \hat{s}^{(k)}).
\end{aligned}
\tag{A.5}
$$

It is easy to check that $(\hat{x}^{(k)}, \hat{w}^{(k)}, \hat{s}^{(k)}) \in \mathcal{A}$ for $k = 0, 1, \cdots$. Denote

$$\eta_k^x = \tilde{x}^{(k)} - \hat{x}^{(k)}, \ \eta_k^w = \tilde{w}^{(k)} - \hat{w}^{(k)}, \ \eta_k^s = \tilde{s}^{(k)} - \hat{s}^{(k)},$$

then by (A.4) and (A.5) we have

$$
\begin{aligned}
\eta_k^x &= \tilde{x}^{(k)} - \hat{x}^{(k)} = (1 - \alpha_{k-1})(\tilde{x}^{(k-1)} - \hat{x}^{(k-1)}) = \nu_k\eta_0^x, \\
\eta_k^w &= \tilde{w}^{(k)} - \hat{w}^{(k)} = (1 - \alpha_{k-1})(\tilde{w}^{(k-1)} - \hat{w}^{(k-1)}) = \nu_k\eta_0^w, \\
\eta_k^s &= \tilde{s}^{(k)} - \hat{s}^{(k)} = (1 - \alpha_{k-1})(\tilde{s}^{(k-1)} - \hat{s}^{(k-1)}) = \nu_k\eta_0^s,
\end{aligned}
\tag{A.6}
$$

where ν_k is defined by (3.12).

Since $(\hat{x}^{(k)}, \hat{w}^{(k)}, \hat{s}^{(k)}) \in \mathcal{A}$, for any $(\bar{x}, \bar{w}, \bar{s}) \in \mathcal{A}$, we have

$$(\hat{x}^{(k)} - \bar{x})^T(\hat{s}^{(k)} - \bar{s}) = 0.$$

By $\hat{x}^{(k)} = \tilde{x}^{(k)} - \eta_k^x, \hat{s}^{(k)} = \tilde{s}^{(k)} - \eta_k^s$ and the above formula we get

$$
\begin{aligned}
&(\tilde{x}^{(k)})^T(\eta_k^s + \bar{s}) + (\tilde{s}^{(k)})^T(\bar{x} + \eta_k^x) \\
&= \bar{x}^T \bar{s} + (\tilde{x}^{(k)})^T \tilde{s}^{(k)} + (\eta_k^x)^T \eta_k^s + \bar{x}^T \eta_k^s + \bar{s}^T \eta_k^x.
\end{aligned} \tag{A.7}
$$

By Lemmas 2.2 and 2.3 the primal-dual optimal solution set $\mathcal{S}$ of (2.1) and its dual (2.4) (which is equal to the solution set of system (3.1)) is nonempty. Let $(\tilde{x}^*, \tilde{w}^*, \tilde{s}^*) \in \mathcal{S}$ and in (A.7) set

$$
(\bar{x}, \bar{w}, \bar{s}) = (\hat{x}^{(0)}, \hat{w}^{(0)}, \hat{s}^{(0)}) = (\tilde{x}^*, \tilde{w}^*, \tilde{s}^*),
$$

then by (A.6), $\bar{x}\bar{s} = 0$ in (3.1) and $\nu_k < 1$ we get

$$
\nu_k\left((\tilde{s}^{(0)})^T \tilde{x}^{(k)} + (\tilde{x}^{(0)})^T \tilde{s}^{(k)}\right) \leq (\tilde{x}^{(k)})^T \tilde{s}^{(k)} + \nu_k^2(\bar{\eta}_0^x)^T \bar{\eta}_0^s + \nu_k(\bar{x}^T \bar{\eta}_0^s + \bar{s}^T \bar{\eta}_0^x),
$$

where $\bar{\eta}_0^x = \tilde{x}^{(0)} - \tilde{x}^*, \bar{\eta}_0^s = \tilde{s}^{(0)} - \tilde{s}^*$. Denote $\tilde{n} = n + m, \xi_0 = \min\{\tilde{x}_i^{(0)}, \tilde{s}_i^{(0)}, i = 1, \ldots, \tilde{n}\}$, by $\mu_k \geq \bar{\mu}, 1 \geq \nu_k$ and $(\tilde{x}^{(k)})^T \tilde{s}^{(k)} = \tilde{n}\mu_k$ we get

$$
\nu_k(\|\tilde{x}^{(k)}\|_1 + \|\tilde{s}^{(k)}\|_1) \leq C_1 \mu_k, \tag{A.8}
$$

where

$$
C_1 = \frac{1}{\xi_0}\left(\tilde{n} + \frac{\nu_k((\bar{\eta}_0^x)^T \bar{\eta}_0^s + \bar{x}^T \bar{\eta}_0^s + \bar{s}^T \bar{\eta}_0^x)}{\bar{\mu}}\right) = O(\tilde{n}).
$$

Denote $d^{(k)} = \sqrt{\tilde{x}^{(k)}/\tilde{s}^{(k)}}$, where the operations are componentwise. By the third equation in (A.3) we have

$$
(d^{(k)})^{-1}\Delta\tilde{x}^{(k)} + d^{(k)}\Delta\tilde{s}^{(k)} = h_k,
$$

where $h_k = \frac{\sigma_k\hat{\mu}}{\sqrt{\tilde{x}^{(k)}\tilde{s}^{(k)}}} - \sqrt{\tilde{x}^{(k)}\tilde{s}^{(k)}}$. Hence

$$
(d^{(k)})^{-1}(\Delta\tilde{x}^{(k)} + \eta_k^x) + d^{(k)}(\Delta\tilde{s}^{(k)} + \eta_k^s) = h_k + (d^{(k)})^{-1}\eta_k^x + d^{(k)}\eta_k^s. \tag{A.9}
$$

Since $\tilde{A}\eta_k^x = \tilde{A}\tilde{x}^{(k)} - b = \tilde{r}_b^{(k)}$ and $\tilde{A}^T\eta_k^w + \eta_k^s = \tilde{A}^T\tilde{w}^{(k)} + \tilde{s}^{(k)} - c = \tilde{r}_c^{(k)}$, by (A.3) we have

$$
\begin{aligned}
\tilde{A}(\Delta\tilde{x}^{(k)} + \eta_k^x) &= -\tilde{r}_b^{(k)} + \tilde{r}_b^{(k)} = 0, \\
\tilde{A}^T(\Delta\tilde{w}^{(k)} + \eta_k^w) + (\Delta\tilde{s}^{(k)} + \eta_k^s) &= -\tilde{r}_c^{(k)} + \tilde{r}_c^{(k)} = 0.
\end{aligned} \tag{A.10}
$$

Hence $(\Delta\tilde{x}^{(k)} + \eta_k^x) \in N(\tilde{A}), (\Delta\tilde{s}^{(k)} + \eta_k^s) \in R(\tilde{A}^T)$, where $N(\tilde{A})$ is the null space of $\tilde{A}$ and $R(\tilde{A}^T)$ is the range space of $\tilde{A}^T$, and we have $(\Delta\tilde{x}^{(k)} + \eta_k^x)^T(\Delta\tilde{s}^{(k)} + \eta_k^s) = 0$. Denote by $\tilde{P}$ the orthogonal projection onto the subspace $N(\tilde{A}D_k)$, where $D_k = \text{diag}(d^{(k)})$, by (A.9) and (A.10) we get

$$
\begin{aligned}
(D^{(k)})^{-1}\Delta\tilde{x}^{(k)} &= \tilde{P}h_k - (I - \tilde{P})(D^{(k)})^{-1}\eta_k^x + \tilde{P}D^{(k)}\eta_k^s, \\
D^{(k)}\Delta\tilde{s}^{(k)} &= (I - \tilde{P})h_k + (I - \tilde{P})(D^{(k)})^{-1}\eta_k^x - \tilde{P}D^{(k)}\eta_k^s.
\end{aligned} \tag{A.11}
$$

For convenience, we will denote $(D^{(k)})^{-1}\Delta\tilde{x}^{(k)}, (D^{(k)})^{-1}\eta_k^x$ and $D^{(k)}\eta_k^s$ as $(d^{(k)})^{-1}\Delta\tilde{x}^{(k)}, (d^{(k)})^{-1}\eta_k^x$ and $d^{(k)}\eta_k^s$ respectively, where the multiplications between vectors are all componentwise. Since $\tilde{P}$ and $I - \tilde{P}$ are orthogonal projections, we have

$$
\begin{aligned}
\|(d^{(k)})^{-1}\Delta\tilde{x}^{(k)}\| &\leq \|h_k\| + \|(d^{(k)})^{-1}\eta_k^x\| + \|d^{(k)}\eta_k^s\|, \\
\|d^{(k)}\Delta\tilde{s}^{(k)}\| &\leq \|h_k\| + \|(d^{(k)})^{-1}\eta_k^x\| + \|d^{(k)}\eta_k^s\|.
\end{aligned} \tag{A.12}
$$

Since $\tilde{x}_i^{(k)}\tilde{s}_i^{(k)} \geq \gamma\mu_k$ and $0.2 = \sigma_{\min} \leq \sigma_k \leq \sigma_{\max} = 1$ according to Algorithm 3.2, by (A.1) we have

$$\|h_k\|^2 = \|\frac{\sigma_k\hat{\mu}}{\sqrt{\tilde{x}^{(k)}\tilde{s}^{(k)}}} - \sqrt{\tilde{x}^{(k)}\tilde{s}^{(k)}}\|^2$$
$$= \sum_{i=1}^{\tilde{n}}\left(\frac{(\sigma_k\hat{\mu})^2}{\tilde{x}_i^{(k)}\tilde{s}_i^{(k)}} - 2\sigma_k\hat{\mu} + \tilde{x}_i^{(k)}\tilde{s}_i^{(k)}\right) \leq \tilde{n}\left(\frac{4\sigma_{\max}^2}{\gamma} - 2\sigma_{\min} + 1\right)\mu_k.$$

By (A.8) we get

$$\|(d^{(k)})^{-1}\eta_k^x\| = \|\sqrt{\frac{1}{\tilde{x}^{(k)}\tilde{s}^{(k)}}}\tilde{s}^{(k)}\nu_k\eta_0^x\|$$
$$\leq \frac{1}{\gamma\mu_k}\nu_k\|\tilde{s}^{(k)}\|_1\|\eta_0^x\| \leq \frac{C_1\|\eta_0^x\|}{\sqrt{\gamma}}\sqrt{\mu_k}.$$

Similarly we have

$$\|d^{(k)}\eta_k^s\| = \|\sqrt{\frac{1}{\tilde{x}^{(k)}\tilde{s}^{(k)}}}\tilde{x}^{(k)}\nu_k\eta_0^s\| \leq \frac{C_1\|\eta_0^s\|}{\sqrt{\gamma}}\sqrt{\mu_k}.$$

Then by (A.12) we get

$$\|(d^{(k)})^{-1}\Delta\tilde{x}^{(k)}\| \leq C_2\sqrt{\mu_k}, \quad \|d^{(k)}\Delta\tilde{s}^{(k)}\| \leq C_2\sqrt{\mu_k}, \tag{A.13}$$

where

$$C_2 = \sqrt{\tilde{n}(\frac{4\sigma_{\max}^2}{\gamma} - 2\sigma_{\min} + 1)} + \frac{C_1(\|\eta_0^s\| + \|\eta_0^x\|)}{\sqrt{\gamma}} = O(\tilde{n}).$$

By (A.13) we get

$$|(\Delta\tilde{x}^{(k)})^T\Delta\tilde{s}^{(k)}| \leq C_2^2\mu_k, \quad \|\Delta\tilde{x}^{(k)}\Delta\tilde{s}^{(k)}\| \leq C_2^2\mu_k. \tag{A.14}$$

Denote $t_1 = \frac{(\Delta\tilde{x}^{(k)})^T\Delta\tilde{s}^{(k)}}{\tilde{n}}$. By $0 < \gamma < 0.1, \tilde{n} > 1$ and (A.14) we get

$$\|\Delta\tilde{x}^{(k)}\Delta\tilde{s}^{(k)} - \gamma t_1 e\| \leq C_2^2\mu_k + \gamma\sqrt{\tilde{n}}\frac{C_2^2\mu_k}{\tilde{n}} \leq (1+\gamma)C_2^2\mu_k, \tag{A.15a}$$

$$\|\Delta\tilde{x}^{(k)}\Delta\tilde{s}^{(k)} - t_1 e\| \leq C_2^2\mu_k + \sqrt{\tilde{n}}\frac{C_2^2\mu_k}{\tilde{n}} \leq 2C_2^2\mu_k. \tag{A.15b}$$

Denote $x(\alpha) = \tilde{x}^{(k)} + \alpha\Delta\tilde{x}^{(k)}, s(\alpha) = \tilde{s}^{(k)} + \alpha\Delta\tilde{s}^{(k)}$, by (A.3) we have

$$x(\alpha)s(\alpha) = (1-\alpha)\tilde{x}^{(k)}\tilde{s}^{(k)} + \alpha\sigma_k\hat{\mu}e + \alpha^2\Delta\tilde{x}^{(k)}\Delta\tilde{s}^{(k)}, \tag{A.16}$$

$$\mu(\alpha) = \frac{(x(\alpha))^Ts(\alpha)}{\tilde{n}} = (1-\alpha)\mu_k + \alpha\sigma_k\hat{\mu} + t_1\alpha^2. \tag{A.17}$$

Define

$$\bar{\alpha} = \min\left\{\frac{\sigma_{\min}}{2C_2^2}, \frac{0.99\beta}{2C_2^2}\right\}. \tag{A.18}$$

Then for any $\alpha \in [0, \bar{\alpha}]$, by (A.1), (A.14)–(A.18) and $0 < \gamma < 0.1$, we get

$$
\begin{aligned}
g_1(\alpha) &= x(\alpha)s(\alpha) - \gamma\mu(\alpha)e = (1-\alpha)(\tilde{x}^{(k)}\tilde{s}^{(k)} - \gamma\mu_k e) \\
&\quad + \alpha\sigma_k\hat{\mu}(1-\gamma)e + \alpha^2(\Delta\tilde{x}^{(k)}\Delta\tilde{s}^{(k)} - \gamma t_1 e) \\
&\geq \alpha[\sigma_k\hat{\mu}(1-\gamma) - \alpha\|\Delta\tilde{x}^{(k)}\Delta\tilde{s}^{(k)} - \gamma t_1 e\|]e \\
&\geq \alpha[\sigma_{\min}\mu_k(1-\gamma) - \bar{\alpha}(1+\gamma)C_2^2\mu_k]e \\
&\geq \alpha\left(0.9\sigma_{min} - 1.1 \cdot \frac{\sigma_{\min}}{2C_2^2} \cdot C_2^2\right)\mu_k e > 0,
\end{aligned}
\tag{A.19}
$$

$$
\begin{aligned}
\mu(\alpha) &= (1-\alpha)\mu_k + \alpha\sigma_k\hat{\mu} + t_1\alpha^2 \geq (1-\alpha)\bar{\mu} + \alpha\bar{\mu} + \alpha\left(0.5\sigma_k\mu_k - \frac{C_2^2\mu_k}{\tilde{n}}\alpha\right) \\
&\geq \bar{\mu} + \alpha\mu_k\left(0.5\sigma_{\min} - C_2^2 \cdot \frac{\sigma_{\min}}{2C_2^2}\right) \geq \bar{\mu}.
\end{aligned}
\tag{A.20}
$$

When $\delta_k > \beta$, which means $\|\tilde{x}^{(k)}\tilde{s}^{(k)} - \mu_k e\| > \beta\mu_k$, then for any α, $0 \leq \alpha \leq \bar{\alpha}$ we have

$$
\begin{aligned}
g_2(\alpha) &= \|x(\alpha)s(\alpha) - \mu(\alpha)e\| \\
&= \|(1-\alpha)(\tilde{x}^{(k)}\tilde{s}^{(k)} - \mu_k e) + \alpha^2(\Delta\tilde{x}^{(k)}\Delta\tilde{s}^{(k)} - t_1 e)\| \\
&\leq (1-0.01\alpha)\|\tilde{x}^{(k)}\tilde{s}^{(k)} - \mu_k e\| \\
&\quad - \alpha(0.99\|\tilde{x}^{(k)}\tilde{s}^{(k)} - \mu_k e\| - \alpha\|\Delta\tilde{x}^{(k)}\Delta\tilde{s}^{(k)} - t_1 e)\|) \\
&\leq (1-0.01\alpha)\|\tilde{x}^{(k)}\tilde{s}^{(k)} - \mu_k e\| - \alpha[0.99\beta\mu_k - \alpha(2C_2^2\mu_k)] \\
&\leq (1-0.01\alpha)\|\tilde{x}^{(k)}\tilde{s}^{(k)} - \mu_k e\| - \alpha\mu_k[0.99\beta - \frac{0.99\beta}{2C_2^2}(2C_2^2)] \\
&= (1-0.01\alpha)\|\tilde{x}^{(k)}\tilde{s}^{(k)} - \mu_k e\|.
\end{aligned}
\tag{A.21}
$$

By (A.19)–(A.21) and Step 3 in the algorithm, we have $\alpha_k \geq \theta\bar{\alpha} =$ for each $k = 0, 1, \cdots$. By the definition of C_2 below (A.13), and by the definition of $\bar{\alpha}$ in (A.18), there exists a constant $\tilde{C} > 0$ such that $\bar{\alpha} \geq \frac{\tilde{C}}{\theta\tilde{n}^2}$. Therefore we have $\alpha_k \geq \frac{C}{\tilde{n}^2}$ for each $k = 0, 1, \cdots$, where $C = \frac{\tilde{C}}{\theta}$.

References

[1] D.S. Atkinson and P.M. Vaidya, A scaling technique for finding the weighted analytic center of a polytope, *Mathematical Programming*, **57** (1992), 163–192.

[2] M.S. Bazaraa, H.D. Sherali and C.M. Shetty, *Nonlinear Programming: Theory and Algorithms*, 3rd edition, John Wiley & Sons, New Jersey, 2006.

[3] J.F. Bonnans and C.C. Gonzaga, Convergence of interior point algorithms for the monotone linear complementarity problem, *Mathematics of Operations Research*, **21** (1996), 1–25.

[4] J.F. Bonnans and F.A. Potra, On the convergence of the iteration sequence of infeasible path following algorithms for linear complementarity problems, *Mathematics of Operations Research*, **22** (1997), 378–407.

[5] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.

[6] Y.H. Dai, Z.H. Wang and F.M. Xu, A primal-dual method for unfolding neutron energy spectrum from multiple activation foils, Research report, AMSS, Chinese Academy of Sciences, 2018.

[7] C.C. Gonzaga, The largest step path following algorithm for monotone linear complementarity problems, *Mathematical Programming*, **76** (1997), 309–332.

[8] C.C. Gonzaga and R.A. Tapia, On the convergence of the Mizuno-Todd-Ye algorithm to the analytic center of the solution set, *SIAM Journal on Optimization*, **7** (1997), 47–65.

[9] M.D. González-Lima, R.A. Tapia and F.A. Potra, On effectively computing the analytic center of the solution set by primal-dual interior-point methods, *SIAM Journal on Optimization*, **8** (1998), 1–25.

[10] M. Kojima, N. Megiddo and S. Mizuno, A primal-dual infeasible-interior-point algorithm for linear programming, *Mathematical Programming*, **61** (1993), 263–280.

[11] M. Kojima, S. Mizuno and A. Yoshise, A primal-dual interior point algorithm for linear programming, in *Progress in Mathematical Programming: Interior-Point and Related Methods*, N. Megiddo (ed.), Springer, New York, 1989, 29–47.

[12] I.J. Lustig, R.E. Marsten and D.F. Shanno, On implementing Mehrotra's predictor-corrector interior-point method for linear programming, *SIAM Journal on Optimization*, **2** (1992), 435–449.

[13] N. Megiddo, Pathways to the optimal set in linear programming, in *Progress in Mathematical Programming: Interior-Point and Related Methods*, N. Megiddo (ed.), Springer, New York, 1989, 131–158.

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

[15] S. Mehrotra, Quadratic convergence in a primal-dual method, *Mathematics of Operations Research*, **18** (1993), 741–751.

[16] S. Mizuno, Polynomiality of infeasible-interior-point algorithms for linear programming, *Mathematical Programming*, **67** (1994), 109–119.

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

[18] S. Mizuno, M.J. Todd and Y. Ye, A Surface of Analytic Centers and Primal-Dual Infeasible-Interior-Point Algorithms for Linear Programming, *Mathematics of Operations Research*, **20** (1995), 135–162.

[19] R. Monteiro, J. O'Neal and T. Tsuchiya, Uniform boundedness of a preconditioned normal matrix used in interior-point methods, *SIAM Journal on Optimization*, **15** (2004), 96–100.

[20] A. Oliveira and D. Sorensen, A new class of preconditioners for large-scale linear systems from interior point methods for linear programming, *Linear Algebra and its Applications*, **394** (2005), 1–24.

[21] F. Potra, An infeasible-interior-point predictor-corrector algorithm for linear programming, *SIAM Journal on Optimization*, **6** (1996), 19-32.

[22] J. Renegar, *A Mathematical View of Interior-Point Methods in Convex Optimization*, SIAM, Philadelphia, 2001.

[23] C. Roos, A full-newton step o(n) infeasible interior-point algorithm for linear optimization, *SIAM Journal on Optimization*, **16** (2006), 1110–1136.

[24] C. Roos, T. Terlaky and J.P. Vial, *Interior Point Methods for Linear Optimization*, 2nd edition, Springer, Berlin, 2005.

[25] A. Wächter and L.T. Biegler, On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming, *Mathematical Programming*, **106** (2006), 25–57.

[26] Y. Wang, Y. Yuan and H. Zhang, A trust region-CG algorithm for deblurring problem in atmospheric image reconstruction, *Science in China Series A: Mathematics*, **45** (2002), 731–740.

[27] M. Wright, Ill-conditioning and computational error in interior methods for nonlinear programming, *SIAM Journal on Optimization*, **9** (1998), 84–111.

[28] S.J. Wright, *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, 1996.

[29] S.J. Wright, Effects of finite-precision arithmetic on interior-point methods for nonlinear programming, *SIAM Journal on Optimization*, **12** (2001), 36–78.

[30] Y. Ye, *Interior Point Algorithms: Theory and Analysis*, John Wiley & Sons, New Jersey, 1997.

[31] Y. Ye, O. Güler, R.A. Tapia and Y. Zhang, A quadratically convergent $O(\sqrt{n}L)$-iteration algorithm for linear programming, *Mathematical Programming*, **59** (1993), 151–162.

[32] Y. Ye, M.J. Todd and S. Mizuno, An $O(\sqrt{n}L)$-iteration homogeneous and self-dual linear programming algorithm, *Mathematics of Operations Research*, **19** (1994), 53–67.

[33] Y. Zhang, On the convergence of a class of infeasible interior-point methods for the horizontal linear complementarity problem, *SIAM Journal on Optimization*, **4** (1994), 208–227.

[34] Y. Zhang, Solving large-scale linear programs by interior-point methods under the MATLAB environment, *Optimization Methods and Software*, **10** (1998), 1–31.

[35] Y. Zhangsun, *Unfolding Method Based on Entropy Theory for the Determination of Neutron Spectrum (in Chinese)*, Master's thesis, Northwest Institute of Nuclear Technology, Xi'an, Shanxi, P. R. China, 2015.