Convergence Analysis of Inexact Infeasible Interior Point Method for Linear Optimization

Ghussoun Al-Jeiroudi* Jacek Gondzio†

School of Mathematics

The University of Edinburgh

Mayfield Road, Edinburgh EH9 3JZ

United Kingdom.

MS-2007-003

August 30th, 2007

*Email: G.Al-Jeiroudi@sms.ed.ac.uk

†Email: J.Gondzio@ed.ac.uk, URL: http://maths.ed.ac.uk/~gondzio/

Convergence Analysis of Inexact Infeasible Interior Point Method for Linear Optimization

Abstract

In this paper we present the convergence analysis of the inexact infeasible path-following (IIPF) interior point algorithm. In this algorithm the preconditioned conjugate gradient method is used to solve the reduced KKT system (the augmented system). The augmented system is preconditioned by using a block triangular matrix.

The KKT system is solved approximately. Therefore, it becomes necessary to study the convergence of interior point method for this specific inexact case. We present the convergence analysis of the inexact infeasible path-following (IIPF) algorithm, prove the global convergence of this method and provide complexity analysis.

Keywords: Inexact Interior Point Methods, Linear Programming, Preconditioned Conjugate Gradients, Indefinite System.

1 Introduction

In this paper we are concerned with the use of primal-dual interior point method (IPM for short) to solve large-scale linear programming problems. The primal-dual method is applied to the primal-dual formulation of the linear program

$$
\begin{array}{llcll}
\min & c^T x & \qquad & \max & b^T y \\
\text{s.t.} & Ax = b, & & \text{s.t.} & A^T y + s = c, \\
 & x \geq 0; & & & y \text{ free}, \;\; s \geq 0,
\end{array}
$$

where $A \in \mathcal{R}^{m\times n}$, $x, s, c \in \mathcal{R}^n$ and $y, b \in \mathcal{R}^m$. We assume that $m \leq n$. The primal-dual algorithm is usually faster and more reliable than the pure primal or pure dual method [1, 17]. The main computational effort of this algorithm consists in the computation of the primal-dual Newton direction of the Karush-Kuhn-Tucker (KKT) conditions:

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

where $X = \text{diag}(x)$, $S = \text{diag}(s)$ and $e \in \mathcal{R}^n$ is a vector of ones.

In this paper, we focus on the use of the Infeasible Path-Following algorithm [17] chapter 6. This algorithm does not require the initial point to be strictly feasible but requires only that its x and s components be strictly positive. At each iteration the following nonlinear system need to be solved

$$F(t) = \begin{bmatrix} Ax - b \\ A^T y + s - c \\ XSe - \sigma\mu e \end{bmatrix} = 0, \tag{2}$$

where $t = (x, y, s)$, $\mu = x^T s/n$ is the average complementarity gap and $\sigma \in (0, 1)$. We use Newton's method to solve this nonlinear system, where the direction at each iteration k is computed according to

$$F^{'}(t^k)\Delta t^k = -F(t^k), \tag{3}$$

which yields

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x^k \\ \Delta y^k \\ \Delta s^k \end{bmatrix} = - \begin{bmatrix} Ax^k - b \\ A^T y^k + s^k - c \\ X^k S^k e - \sigma_k \mu_k e \end{bmatrix}. \tag{4}$$

Solving the linear system (4) with a direct method becomes sometimes very expensive for large problems. In these situations it is reasonable to use an iterative method. Therefore, instead of solving (3) exactly, we solve it with the inexact Newton method:

$$F^{'}(t^k)\Delta t^k = -F(t^k) + r^k, \tag{5}$$

where r^k is the residual of the inexact Newton method. Any approximate step is accepted provided that the residual r^k is small such as

$$\|r^k\| \leq \eta_k \|F(t^k)\|, \tag{6}$$

as required by the theory [7, 9]. We refer to the term η_k as the forcing term.

In the computational practice (4) is reduced: after substituting

$$\Delta s = -X^{-1}S\Delta x - s + \sigma\mu X^{-1}e, \tag{7}$$

in the second row we get the following symmetric indefinite system of linear equations, usually called the augmented system

$$\begin{bmatrix} -\Theta^{-1} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}, \tag{8}$$

where $\Theta = XS^{-1}$, $f = A^T y - c + \sigma\mu X^{-1}e$ and $g = Ax - b$.

We use the preconditioned conjugate gradient (PCG) method to solve the augmented system (8) preconditioned by a block triangular matrix P. We have in mind a particular class of preconditioners in this paper, the ones which try to guess a "basis", a nonsingular submatrix of A. There has been recently a growing interest in such preconditioners see for example [5, 6, 8, 13].

In [8] the following reasoning has been used to find the preconditioner for KKT system. From the complementarity condition we know that at the optimum $x_j s_j = 0, \forall j \in \{1, 2, \ldots, n\}$. Primal-dual interior point methods usually identify a strong optimal partition near the optimal solution. If at the optimal solution $x_j \to 0$ and $s_j \to \hat{s}_j > 0$, then the corresponding element $\Theta_j \to 0$. If, on the other hand, $x_j \to \hat{x}_j > 0$ and $s_j \to 0$, then the corresponding element $\Theta_j \to \infty$.

We partition the matrices:

$$A = [B, N], \quad \Theta^{-1} = \begin{bmatrix} \Theta_B^{-1} & 0 \\ 0 & \Theta_N^{-1} \end{bmatrix},$$

where B is $m \times m$ non-singular matrix. As done in [8], we permute the columns of A and Θ^{-1} such that $\theta_1^{-1} \leq \theta_2^{-1} \leq ... \leq \theta_n^{-1}$, and we pick the first m linearly independent columns of A in this order to construct B. The indefinite matrix in (8) can then be rewritten in the following form

$$K = \begin{bmatrix} -\Theta_B^{-1} & & B^T \\ & -\Theta_N^{-1} & N^T \\ B & N & \end{bmatrix}. \tag{9}$$

The preconditioner P is constructed as follows:

$$P = \begin{bmatrix} & & B^T \\ & -\Theta_N^{-1} & N^T \\ B & N & \end{bmatrix}. \tag{10}$$

P is easily invertible because it is a block-triangular matrix with nonsingular diagonal blocks B, Θ_N^{-1} and B^T.

The PCG method is used to solve the augmented system preconditioned by the block triangular matrix P. A consequence of using PCG method is that the search direction is computed approximately; this yields a specific inexact interior point method. This causes a major difference to the interior point algorithm, whose convergence is proved under the assumption that the search directions are calculated exactly. In this paper we present the convergence analysis of an infeasible path-following algorithm in which the search directions are computed inexactly. We call this method an inexact infeasible path-following algorithm (IIPF).

The use of inexact Newton methods in interior point methods for LP was investigated in [2, 3, 12]. In [2] the convergence of the infeasible interior point algorithm of Kojima, Megiddo, and Mizuno is proved under the assumption that the iterates are bounded. Monteiro and O'Neal [12] propose the convergence analysis of inexact infeasible long-step primal-dual algorithm and give complexity results for this method. In [12] the PCG method is used to solve the normal equations preconditioned with a sparse preconditioner. This preconditioner was proposed originally in [14] and determined according to the Maximum Weight Basis Algorithm. In [4] an inexact interior point method for semidefinite programming is presented. It allows the linear system to be solved to a low accuracy when the current iterate is far from the solution. In [11] the convergence analysis of inexact infeasible primal-dual path-following algorithm for convex quadratic programming is presented. In these papers the search directions are inexact as the PCG method is used to solve the normal equations. Korzak [10] proves the convergence of the infeasible interior point algorithm of Kojima, Megiddo and Mizuno as the search directions are computed inexactly, under the assumption that the iterates are bounded. Korzak's analysis [10] works for general case of inexact search directions.

In this paper we study the convergence analysis of inexact infeasible path following algorithm for linear programming as the PCG method is used to solve the augmented system preconditioned with block triangular sparse preconditioner. We prove the global convergence and the complexity result for this method without having to assume the boundedness of the iterates. We design a suitable stopping criteria for the PCG method. This plays an important role in the whole convergence of IIPF algorithm. This stopping criteria allows a low accuracy when the current iterate is far from the solution. We state conditions on the forcing term of inexact Newton method in order to prove the convergence of IIPF algorithm.

The approach in this paper can be used in the cases where the augmented system is solved iteratively, providing that the residual of this iterative method has a zero block $r = [r_1, 0]$. So we can carry out the approach in this paper to cases like [16] for example.

In order to prove the convergence of an inexact infeasible interior point method, we should prove first that the PCG method, when applied to an indefinite system, converges. Then, we prove the convergence of the IIPF Algorithm.

The paper is organized as follows. In Section 2 we study the behaviour of the residual of the PCG method when applied to an indefinite system (augmented system) preconditioned with (10). In Section 3 we compute the residual of the inexact Newton method and choose suitable stopping criteria to the PCG method which makes sense for the convergence of the inexact Newton method. In Section 4 we perform the convergence analysis and provide the complexity result for the IIPF Algorithm. In Section 5 we draw some conclusions.

2 Convergence of the PCG method

Following the theory developed by Rozložník and Simoncini [15], in [8] we studied the behaviour of the preconditioned conjugate gradient method on the indefinite system (8) preconditioned with (10). We can apply PCG method to an indefinite system because the following properties are satisfied. The first property is the preconditioned matrix KP^{-1} is J-symmetric where $J = P^{-1}$, and the second is that the residuals of the PCG method have zero block in the form $r^k_{PCG} = [r^k_1, 0]$ which results from the use of a specific starting point. See [8]. We gave explicit formulae describing the convergence of the error term. The following theorem, which is proved in [8], shows that the convergence of the error term is similar to that in the case of symmetric positive definite matrices.

Let e^j and r^j_{PCG} be the error term and the residual term on j-th PCG iteration respectively. The matrix of the augmented system K is indefinite. Therefore, the K-norm is not defined, but the K-inner product $(e^j)^T K e^j$ is always positive (Lemma 4 in [8]). So we allow ourselves to write $\|e^j\|_K = \sqrt{(e^j)^T K e^j}$. Accordingly to the partitioning of K in (9) we will partition the error $e^j = [e_1^j, e_2^j]$, where $e_1^j = [e_B^j, e_N^j]$. Later in this section we will use the same partitioning for the residual vector $r_{PCG}^j = [r_1^j, r_2^j]$, where $r_1^j = [r_B^j, r_N^j]$.

Theorem 2.1. *Let e^0 be the initial error of PCG. Then*

$$\|e^j\|_K^2 \leq \min_{\phi \in P_j, \phi(0)=1} \max_{\lambda \in \Lambda(I_m + WW^T)} [\phi(\lambda)]^2 \|e_B^0\|_{\Theta_B^{-1}}^2 + \min_{\phi \in P_j, \phi(0)=1} \max_{\lambda \in \Lambda(I_{n-m} + W^T W)} [\phi(\lambda)]^2 \|e_N^0\|_{\Theta_N^{-1}}^2, \tag{11}$$

where P_j is a polynomial of degree j, $\Lambda(G)$ is the set of eigenvalues of the matrix G and $W = \Theta_B^{-1/2} B^{-1} N \Theta_N^{1/2}$. $I_m + WW^T$ and $I_{n-m} + W^T W$ are symmetric positive definite matrices.

Theorem 2.1 states that the K-norm of the *error* e^j is minimized over the eigenvalues of the symmetric positive definite matrices $I_m + WW^T$ and $I_{n-m} + W^T W$. Consequently, the error term displays asymptotic convergence similar to that observed when PCG is applied to positive definite system.

In the rest of this section we show that the convergence of the *residual* of the PCG method applied to (8) with the preconditioner (10) is similar to the convergence observed in the case of PCG method applied to symmetric positive definite matrix.

2.1 The residual of the PCG method

The Euclidean norm of the residual is minimized over the eigenvalues of the symmetric positive definite matrix $I_m + WW^T$. The following Theorem shows that the residual term displays asymptotic convergence similar to that observed when PCG is applied to positive definite system.

Theorem 2.2. *The residual of the PCG method which is used to solve the augmented system (8) preconditioned by P satisfies*

$$\|r_{PCG}^j\| \leq \min_{\phi \in P_j, \phi(0)=1} \max_{\lambda \in \Lambda(I_m + WW^T)} |\phi(\lambda)| \|r_B^0\|. \tag{12}$$

Proof. The residual satisfies

$$r^j_{PCG} = -Ke^j,$$

and the error can be written as

$$e^j = \phi_j(P^{-1}K)e^0.$$

So we can write the residual as

$$r^j_{PCG} = -K\phi_j(P^{-1}K)e^0 = -\phi_j(KP^{-1})Ke^0 = \phi_j(KP^{-1})r^0_{PCG}.$$

Furthermore,

$$KP^{-1}r^0_{PCG} = \begin{bmatrix} (I + \Theta_B^{-1}B^{-1}N\Theta_N N^T B^{-T})r^0_B - \Theta_B^{-1}B^{-1}N\Theta_N r^0_N + \Theta_B^{-1}B^{-1}r^0_2 \\ r^0_N \\ r^0_2 \end{bmatrix},$$

where $r^j_{PCG} = [r^j_B, r^j_N, r^j_2]$. The initial residual has the form $r^0_{PCG} = [r^0_B, 0, 0]$ because of using the following starting point

$$t^{(0)} = \begin{bmatrix} B^{-1}(g - N\Theta_N f_N) \\ \Theta_N f_N \\ 0 \end{bmatrix},$$

see [8], so the previous equation becomes

$$KP^{-1}r^0_{PCG} = \begin{bmatrix} \Theta_B^{-1}(\Theta_B + B^{-1}N\Theta_N N^T B^{-T})r^0_B \\ 0 \\ 0 \end{bmatrix}. \tag{13}$$

Let us define $C = \Theta_B + B^{-1}N\Theta_N N^T B^{-T}$. It is easy to prove that C is a symmetric positive definite matrix. By an argument similar to the one used to derive (13) we obtain

$$r^j_{PCG} = \phi_j(KP^{-1})r^0_{PCG} = \begin{bmatrix} \phi_j(\Theta_B^{-1}C)r^0_B \\ 0 \\ 0 \end{bmatrix}, \tag{14}$$

and so

$$\|r^j_{PCG}\| = \|\phi_j(\Theta_B^{-1}C)r^0_B\|. \tag{15}$$

Let us observe that $(\Theta_B^{-1}C)^k = \Theta_B^{-1/2}(\Theta_B^{-1/2}C\Theta_B^{-1/2})^k\Theta_B^{1/2} = \Theta_B^{-1/2}(I_m + WW^T)^k\Theta_B^{1/2}$, where $I_m + WW^T$ is a symmetric positive definite matrix.

Using these definitions, (15) can be written as

$$\|r^j_{PCG}\| = \|\Theta_B^{-1/2}\phi_j(I_m + WW^T)\Theta_B^{1/2}r^0_B\| = \|\phi_j(I_m + WW^T)\Theta_B^{1/2}r^0_B\|_{\Theta_B^{-1}}.$$

Therefore,

$$\|r^j_{PCG}\| \le \min_{\phi\in P_j, \phi(0)=1} \max_{\lambda\in\Lambda(I_m+WW^T)} |\phi(\lambda)| \|\Theta_B^{1/2}r^0_B\|_{\Theta_B^{-1}},$$

and the claim is proved after substituting $\|\Theta_B^{1/2}r^0_B\|_{\Theta_B^{-1}} = \|r^0_B\|$. □

In this section we proved that the PCG method applied to the indefinite system (8) preconditioned with (10) and initialized with an appropriate starting point, converges similar to the case of when PCG is applied to positive definite system. In the next section we show that applying PCG to solve (8) with the preconditioner (10) can be analysed using the classical framework of the inexact Newton method (5).

3 The residual of inexact Newton method

Using an iterative method to solve the augmented system (8) produces a specific value of the residual of the inexact Newton method (5). So we shall find the value of the residual r in (5) in order to satisfy (6) and prove the convergence of inexact infeasible path following algorithm.

Solving (8) approximately gives

$$\begin{bmatrix} -\Theta^{-1} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix} + \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}, \tag{16}$$

where $r_1 = [r_B, r_N]$.

That gives the following equations:

$$-X^{-1}S\Delta x + A^T\Delta y = f + r_1 = c - A^T y - \sigma\mu X^{-1}e + r_1, \tag{17}$$

$$A\Delta x = g + r_2 = b - Ax + r_2. \tag{18}$$

Then we find Δs by substituting Δx in (7). However, we can shift the residual from (17) to (7) by assuming there is a residual h while computing Δs. Then (7) is replaced by

$$\Delta s = -X^{-1}S\Delta x - s + \sigma\mu X^{-1}e + h,$$

which we can rewrite as

$$-X^{-1}S\Delta x = \Delta s + s - \sigma\mu X^{-1}e - h.$$

Substituting it in (17) gives

$$A^T\Delta y + \Delta s = c - A^T y - s + h + r_1.$$

To satisfy the second equation of (4) we choose $h = -r_1$. This gives

$$A^T\Delta y + \Delta s = c - A^T y - s, \tag{19}$$

and

$$\Delta s = -X^{-1}S\Delta x - s + \sigma\mu X^{-1}e - r_1,$$

which implies

$$S\Delta x + X\Delta s = -XSe + \sigma\mu e - Xr_1. \tag{20}$$

Equations (18), (19) and (20) give

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} \xi_p \\ \xi_d \\ \xi_\mu \end{bmatrix} + \begin{bmatrix} r_2 \\ 0 \\ -Xr_1 \end{bmatrix},$$

where $\xi_p = b - Ax$, $\xi_d = c - A^T y - s$, $\xi_\mu = -XSe + \sigma\mu e$ and $\sigma \in [0, 1]$.

In the setting in which we apply the PCG method to solve (8) preconditioned with (10) we have $r_2 = 0$ and $r_1 = [r_B, 0]$, see equation (14) in the proof of Theorem 2.2. Therefore, the inexact Newton method residual r is

$$r = \begin{bmatrix} 0 \\ 0 \\ -Xr_1 \end{bmatrix}$$

with $Xr_1 = \begin{bmatrix} X_B r_B \\ X_N r_N \end{bmatrix} = \begin{bmatrix} X_B r_B \\ 0 \end{bmatrix}$.

Shifting the residual from (17) to (7) is an essential step to prove the convergence of the IIPF algorithm. It results in moving the residual from the second row to the last row of the inexact Newton system, which makes the proof of the convergence of the IIPF Algorithm much easier, as we will see in Section 4.

The issue of choosing the stopping criteria of inexact Newton method to satisfy the condition (6) has been discussed in many papers. See for example [2, 3, 4, 10]. In [3] the residual of inexact Newton method is chosen such that

$$\|r^k\| \le \eta_k \mu_k,$$

while in [4] the choice satisfies

$$\|r^k\| \leq \eta_k(n\mu_k).$$

Let the residual $r = [r_p, r_d, r_\mu]$. According to Korzak [10], the residual is chosen such that

$$\begin{aligned} \|r_p^k\|_2 &\leq (1-\tau_1)\|Ax^k - b\|_2, \\ \|r_d^k\|_2 &\leq (1-\tau_2)\|A^T y^k + s^k - c\|_2, \\ \|r_\mu^k\|_\infty &\leq \tau_3 \mu_k. \end{aligned}$$

where $\tau_1, \tau_2 \in (0, 1]$ and $\tau_3 \in [0, 1)$ are some appropriately chosen constants.

We use Korzak's stopping criteria of inexact Newton method. However, as in our case $r_p = r_d = 0$, we will stop the PCG algorithm when

$$\|r_\mu^k\|_\infty \leq \bar{\eta}_k \mu_k.$$

As $r_\mu^k = -X^k r_1^k$ and $r_1 = [r_B, 0]$, the stopping criteria becomes

$$\|X_B^k r_B^k\|_\infty \leq \bar{\eta}_k \mu_k. \tag{21}$$

Satisfying condition (21) guarantees that condition (6) is satisfied too.

We terminate the PCG algorithm when the stopping criteria (21) is satisfied. This stopping criteria allows a low accuracy when the current iterate is far from the solution. In the later iterations the accuracy increases because the average complementarity gap μ reduces from one iteration to another.

4 Convergence of the IIPF Algorithm

In this section we carry out the proof of the convergence of the IIPF algorithm and derive a complexity result. In the previous section we used the shifting residual strategy, which makes the proof of the convergence of this inexact algorithm similar to that of the exact case.

This section is organised as follows. First we describe the IIPF algorithm. Then in Lemmas 4.1, 4.2 and 4.3 we derive useful bounds on the iterates. In Theorems 4.4 and 4.5 we prove that there is a step length α such that the new iteration generated by IIPF algorithm belongs to the neighbourhood $\mathcal{N}_{-\infty}(\gamma, \beta)$ and the average complementarily gap decreases. In order to prove that we supply conditions on the forcing term $\bar{\eta}_k$. In Theorem 4.6 we show that the sequence $\{\mu_k\}$ converges Q-linearly to zero and the normal residual sequence $\{\|(\xi_p^k, \xi_d^k)\|\}$ converges R-linearly to zero. Finally in Theorem 4.7, we provide the complexity result for this algorithm.

Definition: The central path neighbourhood $\mathcal{N}_{-\infty}(\gamma, \beta)$ is defined by

$$\mathcal{N}_{-\infty}(\gamma, \beta) = \{(x, y, s) : \|(\xi_p, \xi_d)\|/\mu \leq \beta\|(\xi_p^0, \xi_d^0)\|/\mu_0,\ (x, s) > 0,\ x_i s_i \geq \gamma\mu,\ i = 1, 2, ..., n\}, \tag{22}$$

where $\gamma \in (0, 1)$ and $\beta \geq 1$ [17].

4.1 Inexact Infeasible Path-Following Algorithm (IIPF Algorithm):

1. Given $\gamma, \beta, \sigma_{min}, \sigma_{max}$ with $\gamma \in (0, 1)$, $\beta \geq 1$, $0 < \sigma_{min} < \sigma_{max} < 0.5$, and $0 < \eta_{min} < \eta_{max} < 1$; choose (x^0, y^0, s^0) with $(x^0, s^0) > 0$;

2. For $k = 0, 1, 2, ...$

- choose $\sigma_k \in [\sigma_{min}, \sigma_{max}]$ and $\bar{\eta}_k \in [\eta_{min}, \eta_{max}]$; and solve

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x^k \\ \Delta y^k \\ \Delta s^k \end{bmatrix} = \begin{bmatrix} \xi_p^k \\ \xi_d^k \\ -X^k S^k e + \sigma_k \mu_k e \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -X^k r_1^k \end{bmatrix}, \tag{23}$$

such that $x_N^k = 0$ and

$$\|X_B^k r_B^k\|_\infty \leq \bar{\eta}_k \mu_k, \tag{24}$$

- choose α_k as the largest value of an α in $[0, 1]$ such that

$$(x^k(\alpha), y^k(\alpha), s^k(\alpha)) \in \mathcal{N}_{-\infty}(\gamma, \beta) \tag{25}$$

and the following Armijo condition holds:

$$\mu_k(\alpha) \leq (1 - .01\alpha)\mu_k; \tag{26}$$

- set $(x^{k+1}, y^{k+1}, s^{k+1}) = (x^k(\alpha_k), y^k(\alpha_k), s^k(\alpha_k))$;
- stop when $\mu_k < \epsilon$, for a small positive constant ϵ.

In this section we will follow the convergence analysis of the infeasible path-following algorithm proposed in the book of Wright [17].

Firstly, let us introduce the quantity

$$\nu_k = \prod_{j=0}^{k-1} (1 - \alpha_j),\ \nu_0 = 1$$

Note that $\xi_p^{k+1} = b - Ax^{k+1} = b - A(x^k + \alpha_k \Delta x^k) = b - Ax^k - \alpha_k A\Delta x^k = \xi_p^k - \alpha_k A\Delta x^k$, from the first row of (23) we get

$$\xi_p^{k+1} = (1 - \alpha_k)\xi_p^k, \tag{27}$$

which implies

$$\xi_p^k = \nu_k \xi_p^0.$$

Note also $\xi_d^{k+1} = c - A^T y^{k+1} - s^{k+1} = c - A^T(y^k + \alpha_k \Delta y^k) - (s^k + \alpha_k \Delta s^k) = (c - A^T y^k - s^k) - \alpha_k(A^T \Delta y^k + \Delta s^k) = \xi_d^k - \alpha_k(A^T \Delta y^k + \Delta s^k)$. From the second row of (23) we get

$$\xi_d^{k+1} = (1 - \alpha_k)\xi_d^k, \tag{28}$$

which implies

$$\xi_d^k = \nu_k \xi_d^0,$$

Consequently, the quantity ν_k satisfies

$$\nu_k \leq \beta \frac{\mu_k}{\mu_0}.$$

More details can be found in [17].

Let (x^*, y^*, s^*) be any primal-dual solution.

Lemma 4.1. *Assume that $(x^k, y^k, s^k) \in \mathcal{N}_{-\infty}(\gamma, \beta)$, $(\Delta x^k, \Delta y^k, \Delta s^k)$ satisfies (23) and (24) for all $k \geq 0$, and $\mu_k \leq (1 - .01\alpha_{k-1})\mu_{k-1}$ for all $k \geq 1$. Then there is a positive constant C_1 such* *that for all $k \geq 0$*

$$\nu_k \|(x^k, s^k)\| \leq C_1 \mu_k, \tag{29}$$

where C_1 is given as

$$C_1 = \zeta^{-1}(n\beta + n + \beta\|(x^0, s^0)\|_\infty \|(x^*, s^*)\|_1/\mu_0),$$

where

$$\zeta = \min_{i=1,\ldots,n} \min(x_i^0, s_i^0).$$

The proof of this Lemma is similar to the proof of Lemma 6.3 in [17]. Moreover, we follow the same logic as in [17] to prove the following lemma.

Lemma 4.2. *Assume that $(x^k, y^k, s^k) \in \mathcal{N}_{-\infty}(\gamma, \beta)$, $(\Delta x^k, \Delta y^k, \Delta s^k)$ satisfies (23) and (24) for all $k \geq 0$, and $\mu_k \leq (1 - .01\alpha_{k-1})\mu_{k-1}$ for all $k \geq 1$. Then there is a positive constant C_2 such that*

$$\|D^{-1}\Delta x^k\| \leq C_2 \mu_k^{1/2}, \tag{30}$$

$$\|D\Delta s^k\| \leq C_2 \mu_k^{1/2}, \tag{31}$$

where $D = X^{1/2}S^{-1/2}$. For all $k \geq 0$.

Proof. For simplicity we omit the iteration index k in the proof.

Let

$$(\bar{x}, \bar{y}, \bar{s}) = (\Delta x, \Delta y, \Delta s) + \nu_k(x^0, y^0, s^0) - \nu_k(x^*, y^*, s^*).$$

Then $A\bar{x} = 0$ and $A^T\bar{y} + \bar{s} = 0$, which implies $\bar{x}^T\bar{s} = 0$.

$A\bar{x} = 0$ because

$$A\bar{x} = A\Delta x + \nu_k Ax^0 - \nu_k Ax^* = \xi_p + \nu_k Ax^0 - \nu_k b = \xi_p - \nu_k \xi_0 = 0.$$

Similarly one can show that $A^T\bar{y} + \bar{s} = 0$. Hence

$$0 = \bar{x}^T\bar{s} = (\Delta x + \nu_k x^0 - \nu_k x^*)^T(\Delta s + \nu_k s^0 - \nu_k s^*). \tag{32}$$

Using the last row of (23) implies

$$\begin{aligned} S(\Delta x + \nu_k x^0 - \nu_k x^*) + X(\Delta s + \nu_k s^0 - \nu_k s^*) &= S\Delta x + X\Delta s + \nu_k S(x^0 - x^*) + \nu_k X(s^0 - s^*) \\ &= -XSe + \sigma\mu e - Xr_1 + \nu_k S(x^0 - x^*) + \nu_k X(s^0 - s^*). \end{aligned}$$

By multiplying this system by $(XS)^{-1/2}$, we get

$$\begin{aligned} &D^{-1}(\Delta x + \nu_k x^0 - \nu_k x^*) + D(\Delta s + \nu_k s^0 - \nu_k s^*) \\ &= (XS)^{-1/2}(-XSe + \sigma\mu e - Xr_1) + \nu_k D^{-1}(x^0 - x^*) + \nu_k D(s^0 - s^*). \end{aligned}$$

The equality (32) gives

$$\|D^{-1}(\Delta x + \nu_k x^0 - \nu_k x^*) + D(\Delta s + \nu_k s^0 - \nu_k s^*)\|^2 = \|D^{-1}(\Delta x + \nu_k x^0 - \nu_k x^*)\|^2 + \|D(\Delta s + \nu_k s^0 - \nu_k s^*)\|^2.$$

Consequently,

$$\|D^{-1}(\Delta x + \nu_k x^0 - \nu_k x^*)\|^2 + \|D(\Delta s + \nu_k s^0 - \nu_k s^*)\|^2$$
$$= \|(XS)^{-1/2}(-XSe + \sigma\mu e - Xr_1) + \nu_k D^{-1}(x^0 - x^*) + \nu_k D(s^0 - s^*)\|^2,$$

which leads to

$$\|D^{-1}(\Delta x + \nu_k x^0 - \nu_k x^*)\| \leq \|(XS)^{-1/2}(-XSe + \sigma\mu e - Xr_1) + \nu_k D^{-1}(x^0 - x^*) + \nu_k D(s^0 - s^*)\|$$
$$\leq \|(XS)^{-1/2}(-XSe + \sigma\mu e - Xr_1)\| + \nu_k\|D^{-1}(x^0 - x^*)\| + \nu_k\|D(s^0 - s^*)\|.$$

The triangle inequality and addition of an extra term $\nu_k\|D(s^0 - s^*)\|$ to the right hand side give

$$\|D^{-1}\Delta x\| \leq \|(XS)^{-1/2}[-XSe + \sigma\mu e - Xr_1]\| + 2\nu_k\|D^{-1}(x^0 - x^*)\| + 2\nu_k\|D(s^0 - s^*)\|. \quad (33)$$

We can write

$$\|(XS)^{-1/2}(-XSe + \sigma\mu e - Xr_1)\|^2 = \sum_{i=1}^{n} \frac{(-x_i s_i + \sigma\mu - x_i r_{1,i})^2}{x_i s_i}$$
$$\leq \frac{\| - XSe + \sigma\mu e - Xr_1\|^2}{\min_i x_i s_i} \leq \frac{1}{\gamma\mu}\| - XSe + \sigma\mu e - Xr_1\|^2.$$

because $(x, y, s) \in \mathcal{N}_{-\infty}(\gamma, \beta)$ which implies $x_i s_i \geq \gamma\mu$ for $i = 1, ..., n$.

On the other hand,

$$\begin{aligned}
\| - XSe + \sigma\mu e\|^2 &= \|XSe\|^2 + \|\sigma\mu e\|^2 - 2\sigma\mu e^T XSe = \|XSe\|^2 + n\sigma^2\mu^2 - 2n\sigma\mu^2 \\
&\leq \|XSe\|_1^2 + n\sigma^2\mu^2 - 2n\sigma\mu^2 = (x^T s)^2 + n\sigma^2\mu^2 - 2n\sigma\mu^2 \\
&\leq n^2\mu^2 + n\sigma^2\mu^2 - 2n\sigma\mu^2 \leq n^2\mu^2,
\end{aligned}$$

as $\sigma \in (0,1)$. This leads to

$$\begin{aligned}
\|-XSe+\sigma\mu e - Xr_1\| &\leq \|-XSe+\sigma\mu e\| + \|Xr_1\| \\
&\leq n\mu + \sqrt{n}\|X_B r_B\|_\infty \leq n\mu + \sqrt{n}\bar{\eta}\mu \leq n\mu + \sqrt{n}\eta_{max}\mu,
\end{aligned}$$

which implies the following

$$\|(XS)^{-1/2}(-XSe+\sigma\mu e - Xr_1)\| \leq \gamma^{-1/2}(n+\sqrt{n}\eta_{max})\mu^{1/2}. \tag{34}$$

On the other hand

$$\nu_k\|D^{-1}(x^0-x^*)\| + \nu_k\|D(s^0-s^*)\| \leq \nu_k(\|D^{-1}\| + \|D\|)\max(\|x^0-x^*\|, \|s^0-s^*\|). \tag{35}$$

For the matrix norm $\|D^{-1}\|$, we have

$$\|D^{-1}\| \leq \max_i \|D_{ii}^{-1}\| = \|D^{-1}e\|_\infty = \|(XS)^{-1/2}Se\|_\infty \leq \|(XS)^{-1/2}\|\|s\|_1,$$

and similarly

$$\|D\| \leq \|(XS)^{-1/2}\|\|x\|_1.$$

Using Lemma 4.1 and (35) we get

$$\begin{aligned}
\nu_k\|D^{-1}(x^0-x^*)\| + \nu_k\|D(s^0-s^*)\| &\leq \nu_k\|(x,s)\|_1\|(XS)^{-1/2}\|\max(\|x^0-x^*\|, \|s^0-s^*\|) \\
&\leq C_1\gamma^{-1/2}\mu^{1/2}\max(\|x^0-x^*\|, \|s^0-s^*\|).
\end{aligned}$$

By substituting the previous inequality and (34) in (33) we get

$$\|D^{-1}\Delta x\| \leq (\gamma^{-1/2}(n + \sqrt{n}\eta_{max}) + 2C_1\gamma^{-1/2}\max(\|x^0 - x^*\|, \|s^0 - s^*\|))\mu^{1/2}.$$

Let us define C_2 as

$$C_2 = \gamma^{-1/2}(n + \sqrt{n}\eta_{max}) + 2C_1\gamma^{-1/2}\max(\|x^0 - x^*\|, \|s^0 - s^*\|).$$

In a similar way we get the bound on $\|D\Delta s\|$, and the proof is completed. □

Lemma 4.3. *Assume that $(x^k, y^k, s^k) \in \mathcal{N}_{-\infty}(\gamma, \beta)$, $(\Delta x^k, \Delta y^k, \Delta s^k)$ satisfies (23) and (24) for all $k \geq 0$, and $\mu_k \leq (1 - .01\alpha_{k-1})\mu_{k-1}$ for all $k \geq 1$. Then there is a positive constant C_3 such that*

$$|(\Delta x^k)^T \Delta s^k| \leq C_3\mu_k, \tag{36}$$

$$|\Delta x_i^k \Delta s_i^k| \leq C_3\mu_k \tag{37}$$

for all $k \geq 0$.

Proof. For simplicity we omit the iteration index k in the proof. From Lemma 4.2 we have

$$|\Delta x^T \Delta s| = |(D^{-1}\Delta x)^T(D\Delta s)| \leq \|D^{-1}\Delta x\|\|D\Delta s\| \leq C_2^2\mu.$$

Moreover, using Lemma 4.2 again we obtain

$$|\Delta x_i \Delta s_i| = |D_{ii}^{-1}\Delta x_i D_{ii}\Delta s_i| = |D_{ii}^{-1}\Delta x_i||D_{ii}\Delta s_i| \leq \|D^{-1}\Delta x\|\|D\Delta s\| \leq C_2^2\mu.$$

Let us denote $C_3 = C_2^2$, and the proof is complete. $\square$

Theorem 4.4. *Assume that* $(x^k, y^k, s^k) \in \mathcal{N}_{-\infty}(\gamma, \beta)$, $(\Delta x^k, \Delta y^k, \Delta s^k)$ *satisfies (23) and (24) for all* $k \geq 0$, *and* $\mu_k \leq (1 - .01\alpha_{k-1})\mu_{k-1}$ *for all* $k \geq 1$. *Then there is a value* $\bar{\alpha} \in (0, 1)$ *such that the following three conditions are satisfied for all* $\alpha \in [0, \bar{\alpha}]$ *for all* $k \geq 0$

$$(x^k + \alpha\Delta x^k)^T(s^k + \alpha\Delta s^k) \geq (1 - \alpha)(x^k)^T s^k \tag{38}$$

$$(x_i^k + \alpha\Delta x_i^k)(s_i^k + \alpha\Delta s_i^k) \geq \frac{\gamma}{n}(x^k + \alpha\Delta x^k)^T(s^k + \alpha\Delta s^k) \tag{39}$$

$$(x^k + \alpha\Delta x^k)^T(s^k + \alpha\Delta s^k) \leq (1 - .01\alpha)(x^k)^T s^k. \tag{40}$$

Proof. For simplicity we omit the iteration index k in the proof.

The last row of the system (23) implies

$$s^T\Delta x + x^T\Delta s = -x^T s + n\sigma\mu - x_B^T r_B,$$

and

$$s_i\Delta x_i + x_i\Delta s_i = -x_i s_i + \sigma\mu - x_i r_{1,i}$$

which leads to

$$\begin{aligned} (x + \alpha\Delta x)^T(s + \alpha\Delta s) &= x^T s + \alpha(x^T\Delta s + s^T\Delta x) + \alpha^2(\Delta x)^T\Delta s \\ &= x^T s + \alpha(-x^T s + n\sigma\mu - x_B^T r_B) + \alpha^2(\Delta x)^T\Delta s \\ &= (1 - \alpha)x^T s + n\alpha\sigma\mu - \alpha x_B^T r_B + \alpha^2(\Delta x)^T\Delta s. \end{aligned}$$

Similarly

$$\begin{aligned}(x_i+\alpha\Delta x_i)(s_i+\alpha\Delta s_i) &= x_is_i+\alpha(s_i\Delta x_i+x_i\Delta s_i)+\alpha^2\Delta x_i\Delta s_i\\ &= x_is_i+\alpha(-x_is_i+\sigma\mu-x_ir_{1,i})+\alpha^2\Delta x_i\Delta s_i\\ &= (1-\alpha)x_is_i+\alpha\sigma\mu-\alpha x_ir_{1,i}+\alpha^2\Delta x_i\Delta s_i.\end{aligned}$$

For (38) we have

$$\begin{aligned}(x+\alpha\Delta x)^T(s+\alpha\Delta s)-(1-\alpha)x^Ts &= (1-\alpha)x^Ts+n\alpha\sigma\mu-\alpha x_B^Tr_B+\alpha^2(\Delta x)^T\Delta s-(1-\alpha)x^Ts\\ &= n\alpha\sigma\mu-\alpha x_B^Tr_B+\alpha^2(\Delta x)^T\Delta s\\ &\geq n\alpha\sigma\mu-\alpha|x_B^Tr_B|-\alpha^2|(\Delta x)^T\Delta s|\\ &\geq n\alpha\sigma\mu-n\alpha\bar{\eta}\mu-\alpha^2C_3\mu\end{aligned}$$

where we used the fact that from (24) we have

$$|x_B^Tr_B|\leq n\|X_Br_B\|_\infty\leq n\bar{\eta}\mu.$$

Therefore, the condition (38) holds for all $\alpha\in[0,\alpha_1]$, where α_1 is given by

$$\alpha_1=\frac{n(\sigma-\bar{\eta})}{C_3}, \tag{41}$$

and we choose $\bar{\eta}<\sigma-\varepsilon_1$ to guarantee α_1 to be strictly positive, where ε_1 is a constant strictly greater than zero.

Let us consider (39)

$$(x_i+\alpha\Delta x_i)(s_i+\alpha\Delta s_i)-\frac{\gamma}{n}(x+\alpha\Delta x)^T(s+\alpha\Delta s)=$$
$$(1-\alpha)x_is_i+\alpha\sigma\mu-\alpha x_ir_{1,i}+\alpha^2\Delta x_i\Delta s_i-\frac{\gamma}{n}((1-\alpha)x^Ts+n\alpha\sigma\mu-\alpha x_B^Tr_B+\alpha^2(\Delta x)^T\Delta s)$$

because $(x, y, s) \in N_{-\infty}(\gamma, \beta)$, so $x_i s_i \geq \gamma\mu$, $\forall i = 1, ..., n$, that gives

$$(x_i + \alpha\Delta x_i)(s_i + \alpha\Delta s_i) - \frac{\gamma}{n}(x + \alpha\Delta x)^T(s + \alpha\Delta s) \geq$$

$$(1-\alpha)\gamma\mu + \alpha\sigma\mu - \alpha \max_i x_i r_{1,i} - \alpha^2|\Delta x_i \Delta s_i| - \gamma(1-\alpha)\mu - \gamma\alpha\sigma\mu + \frac{\gamma}{n}\alpha x_B^T r_B - \frac{\gamma}{n}\alpha^2(\Delta x)^T\Delta s$$

$$\geq \alpha\sigma\mu - \alpha\|X_B r_B\|_\infty - \alpha^2 C_3\mu - \alpha\sigma\gamma\mu - \frac{\gamma}{n}\alpha|x_B^T r_B| - \frac{\gamma}{n}\alpha^2 C_3\mu$$

$$\geq \alpha\sigma\mu - \alpha\bar{\eta}\mu - \alpha^2 C_3\mu - \alpha\sigma\gamma\mu - \gamma\alpha\bar{\eta}\mu - \frac{\gamma}{n}\alpha^2 C_3\mu = \alpha((1-\gamma)\sigma - \bar{\eta}(1+\gamma))\mu - \alpha^2(1 + \frac{\gamma}{n})C_3\mu$$

Condition (39) holds for all $\alpha \in [0, \alpha_2]$, where α_2 is given by:

$$\alpha_2 = \frac{\sigma(1-\gamma) - (1+\gamma)\bar{\eta}}{(1 + \frac{\gamma}{n})C_3}. \tag{42}$$

We choose $\bar{\eta} < \frac{\sigma(1-\gamma)}{(1+\gamma)} - \varepsilon_2$ to guarantee α_2 to be strictly positive, where ε_2 is a constant strictly greater than zero.

Finally, let us consider condition (40)

$$\frac{1}{n}[(x + \alpha\Delta x)^T(s + \alpha\Delta s) - (1 - .01\alpha)x^T s] =$$

$$= \frac{1}{n}[(1-\alpha)x^T s + n\alpha\sigma\mu - \alpha x_B^T r_B + \alpha^2(\Delta x)^T\Delta s - (1 - .01\alpha)x^T s]$$

$$= \frac{1}{n}[-.99\alpha x^T s + n\alpha\sigma\mu - \alpha x_B^T r_B + \alpha^2(\Delta x)^T\Delta s]$$

$$\leq -.99\alpha\mu + \alpha\sigma\mu + \frac{\alpha}{n}|x_B^T r_B| + \frac{\alpha^2}{n}C_3\mu \leq -.99\alpha\mu + \alpha\sigma\mu + \alpha\bar{\eta}\mu + \frac{\alpha^2}{n}C_3\mu.$$

We can conclude that condition (40) holds for all $\alpha \in [0, \alpha_3]$, where α_3 is given by:

$$\alpha_3 = \frac{n(0.99 - \sigma - \bar{\eta})}{C_3}. \tag{43}$$

We choose $\bar{\eta}$ and σ such that $\bar{\eta} + \sigma < 0.99 - \varepsilon_3$ to guarantee α_3 to be strictly positive, where ε_3 is a constant strictly greater than zero.

Combining the bounds (41), (42) and (43), we conclude that conditions (38), (39) and (40) hold for $\alpha \in [0, \bar{\alpha}]$, where

$$\bar{\alpha} = \min\left\{1, \frac{n(\sigma - \bar{\eta})}{C_3}, \frac{\sigma(1-\gamma) - (1+\gamma)\bar{\eta}}{(1+\frac{\gamma}{n})C_3}, \frac{n(0.99 - \sigma - \bar{\eta})}{C_3}\right\}. \tag{44}$$

□

We introduce the constants ε_1, ε_2 and ε_3 to guarantee that the limit of the step length $\bar{\alpha}$ is strictly greater than zero and to make it flexible to choose the parameters $\bar{\eta}_k$ and σ_k.

Note that if $\bar{\eta} < \frac{\sigma(1-\gamma)}{(1+\gamma)}$ then $\bar{\eta} < \sigma$ because $\frac{(1-\gamma)}{(1+\gamma)} < 1$ for any $\gamma \in (0, 1)$.

From this theorem we observe that the forcing term $\bar{\eta}_k$ should be chosen such that the following two conditions $\bar{\eta}_k < \frac{\sigma_k(1-\gamma)}{(1+\gamma)} - \varepsilon_2$ and $\bar{\eta}_k + \sigma_k < 0.99 - \varepsilon_3$ are satisfied. Under these assumption the following theorem guarantees that there is a step length α such that the new point belongs to the neighbourhood $\mathcal{N}_{-\infty}(\gamma, \beta)$ and its average complementarity gap decreases according to condition (26).

Below we prove two theorems using standard techniques which follow from Wright [17].

Theorem 4.5. *Assume that* $\bar{\eta}_k < \frac{\sigma_k(1-\gamma)}{(1+\gamma)} - \varepsilon_2$*,* $\bar{\eta}_k + \sigma_k < 0.99 - \varepsilon_3$ *for* $\varepsilon_2, \varepsilon_3 > 0$*,* $(x^k, y^k, s^k) \in \mathcal{N}_{-\infty}(\gamma, \beta)$ *and* $(\Delta x^k, \Delta y^k, \Delta s^k)$ *satisfies (23) and (24) for all* $k \geq 0$*,* $\mu_k \leq (1 - .01\alpha_{k-1})\mu_{k-1}$ *for all* $k \geq 1$*. Then* $(x^k(\alpha), y^k(\alpha), s^k(\alpha)) \in \mathcal{N}_{-\infty}(\gamma, \beta)$ *and* $\mu_k(\alpha) \leq (1 - .01\alpha)\mu_k$ *for all* $\alpha \in [0, \bar{\alpha}]$*, where* $\bar{\alpha}$ *is given by (44).*

Proof. Theorem 4.4 ensures that the conditions (38), (39) and (40) are satisfied. Note that (40) implies that the condition $\mu_k(\alpha) \leq (1 - .01\alpha)\mu_k$ is satisfied, while (39) guarantees that $x_i^k(\alpha)s_i^k(\alpha) \geq \gamma\mu_k(\alpha)$.

To prove that $(x^k(\alpha), y^k(\alpha), s^k(\alpha)) \in \mathcal{N}_{-\infty}(\gamma, \beta)$, we have to prove that $\|(\xi_p^k(\alpha), \xi_d^k(\alpha))\|/\mu_k(\alpha) \leq$

$\beta\|(\xi_p^0, \xi_d^0)\|/\mu_0$. From (27), (28) and (38) we have

$$\frac{\|(\xi_p^k(\alpha), \xi_d^k(\alpha))\|}{\mu_k(\alpha)} = \frac{(1-\alpha)\|(\xi_p^k, \xi_d^k)\|}{\mu_k(\alpha)} \leq \frac{(1-\alpha)\|(\xi_p^k, \xi_d^k)\|}{(1-\alpha)\mu_k} \leq \frac{\|(\xi_p^k, \xi_d^k)\|}{\mu_k} \leq \beta\frac{\|(\xi_p^0, \xi_d^0)\|}{\mu_0},$$

since $(x^k, y^k, s^k) \in \mathcal{N}_{-\infty}(\gamma, \beta)$. □

Theorem 4.6. *The sequence $\{\mu_k\}$ generated by the IIPF Algorithm converges Q-linearly to zero, and the sequence of residual norms $\{\|(\xi_p^k, \xi_d^k)\|\}$ converges R-linearly to zero.*

Proof. Q-linear convergence of $\{\mu_k\}$ follows directly from condition (26) and Theorem 4.4. Indeed, there exists a constant $\bar{\alpha} > 0$ such that $\alpha_k \geq \bar{\alpha}$ for every k such that

$$\mu_{k+1} \leq (1 - .01\alpha_k)\mu_k \leq (1 - .01\bar{\alpha})\mu_k, \text{ for all } k \geq 0.$$

From (27) and (28) we also have

$$\|(\xi_p^{k+1}, \xi_d^{k+1})\| \leq (1 - \alpha_k)\|(\xi_p^k, \xi_d^k)\|.$$

Therefore,

$$\|(\xi_p^{k+1}, \xi_d^{k+1})\| \leq (1 - \bar{\alpha})\|(\xi_p^k, \xi_d^k)\|.$$

Also from Theorem 4.5 we know that

$$\|(\xi_p^{k+1}, \xi_d^{k+1})\| \leq \mu_k\beta\frac{\|(\xi_p^0, \xi_d^0)\|}{\mu_0}.$$

Therefore, the sequence of residual norms is bounded above by another sequence that converges Q-linearly, so $\{\|(\xi_p^k, \xi_d^k)\|\}$ converges R-linearly. □

Theorem 4.7. *Let $\epsilon > 0$ and the starting point $(x^0, y^0, s^0) \in \mathcal{N}_{-\infty}(\gamma, \beta)$ in the Algorithm IIPF be given. Then there is an index K with*

$$K = O(n^2 |log \epsilon|)$$

such that the iterates $\{(x^k, y^k, s^k)\}$ generated by IIPF Algorithm satisfy

$$\mu_k \leq \epsilon, \ \text{for all } \epsilon \geq K.$$

Proof. If the conditions of Theorem 4.5 are satisfied, then the conditions (25) and (26) are satisfied for all $\alpha \in [0, \bar{\alpha}]$ for all $k \geq 0$. By Theorem 4.4, the quantity $\bar{\alpha}$ satisfies

$$\bar{\alpha} \geq \min \left\{ 1, \frac{n(\sigma - \bar{\eta})}{C_3}, \frac{\sigma(1-\gamma) - (1+\gamma)\bar{\eta}}{(1+\frac{\gamma}{n})C_3}, \frac{n(0.99 - \sigma - \bar{\eta})}{C_3} \right\}.$$

Furthermore, from Lemmas 4.1, 4.2 and 4.3 we have $C_3 = O(n^2)$, therefore

$$\bar{\alpha} \geq \frac{\delta}{n^2}$$

for some positive scalar δ independent of n. That implies

$$\mu_{k+1} \leq (1 - .01\bar{\alpha})\mu_k \leq (1 - \frac{.01\delta}{n^2})\mu_k, \ \text{for } k \geq 0.$$

The complexity result is an immediate consequence of Theorem 3.2 of [17]. □

5 Conclusions

In this paper we have considered the convergence analysis of the inexact infeasible path-following algorithm, where the augmented system is solved iteratively. We have analysed the behaviour of the residual term in the PCG method which is used to solve the augmented system (indefinite system). This analysis reveals that the residual converges to zero and, asymptotically, behaves in a similar way to the classical case when PCG is applied to positive definite system. We have chosen a suitable stopping criteria of the PCG method and have provided a condition on the forcing term. Furthermore, we have proved the global convergence of the IIPF algorithm and have provided a complexity result for this method. The technique to control accuracy in the inexact Newton method proposed and analysed in this paper has been implemented in HOPDM [8].

References

[1] E. D. Andersen, J. Gondzio, C. Mészáros, and X. Xu, *Implementation of interior point methods for large scale linear programming*, in Interior Point Methods in Mathematical Programming, T. Terlaky, ed., Kluwer Academic Publishers, 1996, pp. 189–252.

[2] V. Baryamureeba and T. Steihaug, *On the convergence of an inexact primal-dual interior point method for linear programming*, Technical report, University of Bergen, (2000).

[3] S. Bellavia, *An inexact interior point method*, Journal of Optimization Theory and Applications, 96 (1998), pp. 109–121.

[4] S. Bellavia and S. Pieraccini, *Convergence analysis of an inexact infeasible interior point method for semidefinite programming*, Computational Optimization and Applications, 29 (2004), pp. 289–313.

[5] S. Bocanegra, F. Campos, and A. Oliveira, *Using a hybrid preconditioner for solving large-scale linear systems arising from interior point methods*, Computational Optimization and Applications, 36 (2007), pp. 149–164.

[6] J. S. Chai and K. C. Toh, *Preconditioning and iterative solution of symmetric indefinite linear systems arising from interior point methods for linear programming*, Computational Optimization and Applications, 36 (2007), pp. 221–247.

[7] R. S. Dembo, S. C. Eisenstat, and T. Steihaug, *Inexact Newton methods*, SIAM Journal on Numerical Analysis, 19 (1982), pp. 400–408.

[8] G. A. Jeiroudi, J. Gondzio, and J. Hall, *Preconditioning indefinite systems in interior point methods for large scale linear optimization*, Optimization Methods and Software, published in electronic form: http://www.informaworld.com/smpp/title content=g769587169 db=all, (2007).

[9] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, vol. 16 of Frontiers in Applied Mathematics, SIAM, Philadelphia, 1995.

[10] J. Korzak, *Convergence analysis of inexact infeasible-interior-point-algorithm for solving linear progamming problems*, SIAM Journal on Optimization, 11 (2000), pp. 133–148.

[11] Z. Lu, R. D. S. Monteiro, and J. W. O'Neal, *An iterative solver-based infeasible primal-dual path-following algorithm for convex QP*, Georgia Institute of Technology, (2004).

[12] R. D. S. Monteiro and J. W. O'Neal, *Convergence analysis of long-step primal-dual infeasible interior point lp algorithm based on iterative linear solvers*, Georgia Institute of Technology, (2003).

[13] A. R. L. Oliveira and D. C. 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), pp. 1–24.

[14] M. G. C. Resende and G. Veiga, *An implementation of the dual affine scaling algorithm for minimum cost flow on bipartite uncapacitated networks*, SIAM Journal on Optimization, 3 (1993), pp. 516–537.

[15] M. Rozložník and V. Simoncini, *Krylov subspace methods for saddle point problems with indefinite preconditioning*, SIAM Journal of Matrix Analysis and Applications, 24 (2002), pp. 368–391.

[16] M. Rozloznik and V. Simoncini, *The conjugate gradients method with indefinite preconditioning*, MAT- TRIAD, (2005).

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