REPORTS
IN
INFORMATICS

ISSN 0333-3590

On the Convergence of an Inexact Primal–Dual Interior Point Method for Linear Programming

Venansius Baryamureeba **Trond Steihaug**

REPORT NO 188 March 2000

Department of Informatics

UNIVERSITY OF BERGEN

Bergen, Norway

This report has URL http://www.ii.uib.no/publikasjoner/texrap/ps/2000-188.ps

Reports in Informatics from Department of Informatics, University of Bergen, Norway, is available at http://www.ii.uib.no/publikasjoner/texrap/.

Requests for paper copies of this report can be sent to:

Department of Informatics, University of Bergen, Høyteknologisenteret, P.O. Box 7800, N-5020 Bergen, Norway

On the Convergence of an Inexact Primal–Dual Interior Point Method for Linear Programming *

Venansius Baryamureeba
barya@ii.uib.no

Trond Steihaug
trond@ii.uib.no

March 31, 2000

Department of Informatics
University of Bergen
N–5020 Bergen Norway

Abstract

The inexact primal-dual interior point method which is discussed in this paper chooses a new iterate along an approximation to the Newton direction. The method is the Kojima, Megiddo, and Mizuno globally convergent infeasible interior point algorithm The inexact variation is shown to have the same convergence properties accepting a residual in both the primal and dual Newton step equation also for feasible iterates.

1 Introduction

Consider the primal linear programming problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to:} & Ax = b, \quad x \geq 0, \end{array} \tag{1a}$$

where A is an m-by-n matrix of full rank m, b an m-vector, and c an n-vector; and its dual problem

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to:} & A^T y + z = c, \quad z \geq 0. \end{array} \tag{1b}$$

The optimality conditions for the linear program pair (1a) and (1b) are the Karush-Kuhn-Tucker (KKT) conditions:

$$F(x, y, z) \equiv \begin{pmatrix} Ax - b \\ A^T y + z - c \\ XZe \end{pmatrix} = 0, \quad x \geq 0, \; z \geq 0, \tag{2}$$

where $X = \text{diag}(x)$, $Z = \text{diag}(z)$ and e is the vector of all ones in $\Re^n$. A point (x, y, x) is interior if $x > 0$ and $z > 0$ and an interior point is feasible if it satisfies all the equality constraints, otherwise it is infeasible. An interior point algorithm solves a linear programming problem by generating a sequence of interior points from an initial interior point.

Bellavia [2] proved global convergence of an inexact interior point method. Ito, Kelley, and Sachs [8] and Ito [7] discuss an inexact primal-dual interior point iteration for linear programs in functional spaces. Mizuno and Jarre [13] proved global and polynomial-time convergence of an infeasible interior point algorithm

*Technical report number 188, Department of Informatics, University of Bergen

using inexact computation. Portugal, Resende, Veiga, and Judice [15] presented a truncated primal-infeasible dual-feasible interior point algorithm for linear programming. Portugal, Fernandes, and Judice [14] presented a truncated primal-infeasible dual-feasible interior point algorithm for solving monotone linear complementarity problems. The methods suggested in [8, 14, 15] have a major drawback of remaining primal-feasible once they become primal-feasible. Thus if they happen to become primal-feasible before the complementarity gap is significantly reduced, satisfying the termination criterion in the iterative linear system solver is computationally expensive.

In this paper we discuss the convergence properties of an inexact implementation of the interior point method by Kojima, Megiddo, and Mizuno [9]. This method has been studied by many researchers [10, 11, 12] and is known to be practically efficient among the numerous variations and extensions of the primal-dual interior point algorithm.

The inexact computation in the interior point methods discussed in this paper is in solving the linear system called the Newton step equation. In Section 2 we formulate the linear equations and accuracy requirements in the linear equations on augmented and normal equations forms. In Section 3 we state the inexact interior point variation of the algorithm by Kojima, Megiddo, and Mizuno [9]. Section 4 discusses the global convergence results.

Throughout this paper we use the following notation. For any vector x, x^k denotes x at the k-th (interior point) iteration. Similarly for any matrix X or real number η. X^k and η^k denotes X and η at the k-th iteration. We have adopted the notation $(u, v, w) = (u^T, v^T, w^T)^T$ so for any vectors $x \in \Re^n, y \in \Re^m$, $(x, y) \in \Re^{m+n}$.

2 Inexact Computation

The perturbed KKT conditions for (2) with a positive μ is the nonlinear system

$$F_\mu(x, y, z) \equiv \begin{pmatrix} Ax - b \\ A^T y + z - c \\ XZe - \mu e \end{pmatrix} = 0, \quad x \geq 0,\ z \geq 0. \tag{3}$$

μ is referred to as the μ-complementarity parameter and the set of triples (x, y, z) that satisfy (3) for all $\mu > 0$ is called the (primal-dual) central path [17].

Let (x^k, y^k, z^k) with $(x^k, z^k) > 0$ be the iterate at interior point iteration k and consider the perturbed KKT condition (3). Newton's method defines the equation of directional change (the Newton step equation)

$$F'_{\mu^k}(x^k, y^k, z^k) \begin{pmatrix} \Delta x^k \\ \Delta y^k \\ \Delta z^k \end{pmatrix} = -F_{\mu^k}(x^k, y^k, z^k). \tag{4}$$

If the linear system of equations is solved approximately, then equation (4) has a residual error r^k

$$F'_{\mu^k}(x^k, y^k, z^k) \begin{pmatrix} \Delta x^k \\ \Delta y^k \\ \Delta z^k \end{pmatrix} = -F_{\mu^k}(x^k, y^k, z^k) + r^k. \tag{5}$$

The idea behind inexact interior point algorithms is to derive a stopping criterion of the iterative linear system solvers that minimizes the computational effort involved in computing the search directions and guarantee global convergence.

The residual will be partitioned into the primal infeasibility $\bar{r}^k$, dual infeasibility $\hat{r}^k$, and deviation in complementarity $\tilde{r}^k$, $r^k = (\bar{r}^k, \hat{r}^k, \tilde{r}^k)$. An inexact Newton step of (5) is an an approximate solution of the Newton step equation derived from (2)

$$F'(x^k, y^k, z^k) \begin{pmatrix} \Delta x^k \\ \Delta y^k \\ \Delta z^k \end{pmatrix} = -F(x^k, y^k, z^k) + r_g^k. \tag{6}$$

where

$$r_g^k = \mu^k \begin{pmatrix} 0 \\ 0 \\ e \end{pmatrix} + r^k.$$

since $F' = F'_\mu$ and $F_\mu = F - \mu^k\,(0, 0, e)$.

Consider a function $F : \Re^n \to \Re^n$ so that $F(x^*) = 0$ and F is continuously differentiable in a neighborhood of x^*. For a given $x^k \approx x^*$ the inexact Newton method generates a new iterate $x^{k+1} = x^k + s^k$ where s^k is an approximate solution of the Newton step equation

$$F'(x^k)s^k = -F(x^k). \tag{7}$$

Let

$$r^k = F'(x^k)s^k + F(x^k), \text{ where } \|r^k\| \le \eta^k \|F(x^k)\|. \tag{8}$$

If the iterates $\{x^k\}$ are in a neighborhood of x^*, $\{\|F'(x^k)\|\}$ and $\{\|F'(x^k)^{-1}\|\}$ are bounded, and

$$\eta^k \le \eta_{\max} < 1 \tag{9}$$

then $x^k \to x^*$ as $k \to \infty$. The sequence $\{\eta^k\}$ is referred to as a forcing sequence of inexact Newton methods [4]

A typical choice of the μ-complementarity parameter at iteration k is

$$\mu^k = \beta_1 \frac{(x^k)^T z^k}{n} \tag{10}$$

where $\beta_1 \in [0, 1]$ is the centering parameter. Bellavia [2] observed that

$$\|F(x^k, y^k, z^k)\|_2 \ge \frac{(x^k)^T z^k}{\sqrt{n}}.$$

So if

$$\|r^k\|_2 \le \eta^k (x^k)^T z^k$$

then

$$\begin{aligned} \|r_g^k\|_2 &\le \mu^k \sqrt{n} + \|r^k\|_2 \\ &\le \frac{(x^k)^T z^k}{\sqrt{n}} \left(\beta_1 + \eta^k \sqrt{n}\right) \\ &\le \left(\beta_1 + \eta^k \sqrt{n}\right) \|F(x^k, y^k, z^k)\|_2 \end{aligned}$$

Hence the sequence $\{\beta_1 + \eta^k \sqrt{n}\}$ can be regarded as a forcing sequence of inexact Newton methods applied to the nonlinear system (2).

Numerical testing [1] indicates that

$$\|r^k\|_2 \le \eta^k\,(x^k)^T z^k \text{ for } \eta^k < (1 - \beta_1)/\sqrt{n} \tag{11}$$

is overly restrictive and that we can tolerate a much larger error in (5). To see this, consider the $\|\cdot\|_1$, then

$$\|F(x^k, y^k, z^k)\|_1 \ge (x^k)^T z^k.$$

If

$$\|r^k\|_1 \le \eta^k (x^k)^T z^k$$

then

$$\begin{aligned} \|r_g^k\|_1 &\le \mu^k n + \|r^k\|_1 \\ &\le (x^k)^T z^k \left(\beta_1 + \eta^k\right) \\ &\le \left(\beta_1 + \eta^k\right) \|F(x^k, y^k, z^k)\|_1. \end{aligned}$$

This suggests solving the linear system (5) with an accuracy

$$\|r^k\|_1 \le \eta^k \, (x^k)^T z^k \text{ for } 0 \le \eta^k < 1 - \beta_1 \tag{12}$$

will be sufficient to achieve convergence.

The size of the problem (5) is $m + 2n$-by-$m + 2n$. However, by noticing the structure of the Jacobian matrix

$$F'_{\mu^k}(x^k, y^k, z^k) = \begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Z^k & 0 & X^k \end{bmatrix},$$

it is evident that (4) can be reduced to a (square) system with $m+n$ or m equations. However, in both cases $\tilde{r}^k = 0$.

Bellavia [2] establishes global convergence results for an inexact interior point method by interpreting it as an inexact Newton method using the global inexact Newton method of Eisenstat and Walker [5]. The convergence theory and implementation [2, 3] is based on solving the inexact Newton equation (6) with a forcing sequence where $\eta^k < (1 - \beta_1)/\sqrt{n}$.

In this paper, we prove global convergence results based on the reduced linear systems (the augmented systems or the normal equations), and we bound the residual $r^k = (\bar{r}^k, \hat{r}^k, 0)$ by

$$\|r^k\| \le \eta^k (x^k)^T z^k \quad \text{for} \quad 0 \le \eta^k < 1 \tag{13}$$

for suitable choices of the input parameters of the algorithm. The norm $\|\cdot\|$ can be any l_p norm or any norm in appropriate spaces that satisfies

$$\max\{\|\bar{r}^k\|, \|\hat{r}^k\|\} \le \|(\bar{r}^k, \hat{r}^k)\|. \tag{14}$$

It is evident that termination criterion (11) may be more demanding to satisfy than termination criterion (13).

To simplify the notation, we follow [8] and introduce the residuals

$$\begin{aligned} \xi^k &= b - Ax^k \\ \zeta^k &= c - A^T y^k - z^k. \end{aligned}$$

Let G^k be defined by $G^k = (Z^k)^{-1} X^k$. Eliminating Δz^k in (4) leads to an indefinite augmented system

$$\begin{bmatrix} O & A \\ A^T & -(G^k)^{-1} \end{bmatrix} \begin{pmatrix} \Delta y^k \\ \Delta x^k \end{pmatrix} = \begin{pmatrix} \xi^k \\ z^k + \zeta^k - (X^k)^{-1} \mu^k e \end{pmatrix} \tag{15a}$$

and

$$\Delta z^k = (X^k)^{-1} (\mu^k e - Z^k \Delta x^k) - z^k. \tag{15b}$$

An inexact solution of (15a) satisfies

$$\begin{bmatrix} O & A \\ A^T & -(G^k)^{-1} \end{bmatrix} \begin{pmatrix} \Delta y^k \\ \Delta x^k \end{pmatrix} = \begin{pmatrix} \xi^k \\ z^k + \zeta^k - (X^k)^{-1} \mu^k e \end{pmatrix} + \begin{pmatrix} \bar{r}^k \\ \hat{r}^k \end{pmatrix}. \tag{16}$$

After computing $(\Delta y^k, \Delta x^k)$, Δz^k can be found from (15b). The approximate step $(\Delta y^k, \Delta x^k, \Delta z^k)$ is thus an inexact Newton step of (5) with $r^k = (\bar{r}^k, \hat{r}^k, 0)$.

Eliminating Δx^k from (15a) gives the normal equations

$$AG^k A^T \Delta y^k = AG^k (z^k - (X^k)^{-1} \mu^k e + \zeta^k) + \xi^k. \tag{17}$$

An approximate solution of (17) satisfies

$$AG^k A^T \Delta y^k \quad = \quad AG^k (z^k - (X^k)^{-1} \mu^k e + \zeta^k) + \xi^k + \bar{r}^k. \tag{18a}$$

After computing Δy^k, Δx^k and Δz^k are computed from

$$\Delta x^k \quad = \quad -G^k (c - A^T (y^k + \Delta y^k) - \mu^k (X^k)^{-1} e) \tag{18b}$$

$$\Delta z^k \quad = \quad (X^k)^{-1} (\mu^k e - Z^k \Delta x^k) - z^k. \tag{18c}$$

The linear system (16) is equivalent to (18a) and (18b) when $\hat{r}^k = 0$ and the approximate step $(\Delta y^k, \Delta x^k, \Delta z^k)$ is an inexact Newton step of (5) with $r^k = (\bar{r}^k, 0, 0)$.

If the results in [8] are restricted to the finite dimensional problems (1a) and (1b) then the error in (18a) is required to satisfy the bound

$$\|\bar{r}^k\| \leq \eta^k \|\xi^k\|, \text{ where } \eta^k < 1. \tag{19}$$

Similar control of the error in the primal equation is used in [15]

$$\|\bar{r}^k\| \leq \beta_0 \|\xi^k\|, \text{ where } \beta_0 < \beta_1. \tag{20}$$

If the problems (1a) and (1b) are formulated as linear complementarity problem and the inexact (truncated) Newton interior point algorithm by Portugal, Fernandes, and Judice [14] is applied (16) then the error

$$\|(\bar{r}^k, \hat{r}^k)\| \leq \beta_0 \|(\xi^k, \zeta^k)\|, \text{ where } \beta_0 < \beta_1. \tag{21}$$

Satisfying the termination criterion used in [7, 8, 14, 15] for the iterative linear system solver is computationally expensive once the iterates become almost primal feasible or primal-dual feasible.

Freund, Jarre and Mizuno [6] suggest to control the errors

$$\|\bar{r}^k\| \leq \varepsilon_p^k \text{ and } \|\hat{r}^k\| \leq \varepsilon_d^k \tag{22}$$

which may avoid the problem with exact solution when the iterates are feasible. The variables ε_p^k and ε_d^k are updated at every iteration k, but are not known a priori. The variables go to zero faster than $\|\xi^k\|$ and $\|\zeta^k\|$ respectively and thus requiring high accuracy close to a solution.

Mizuno and Jarre [13] control the error $\|(\bar{r}^k, \hat{r}^k)\|$ using a semi norm. This norm is not computable, but accuracy requirement is shown to be slightly weaker than the relative error with respect to the right hand side in a modified linear system. This system require a QR factorization of the matrix A and the accuracy requirement is thus not suitable for computation.

3 The Kojima–Megiddo–Mizuno Algorithm

The algorithm we discuss in this paper is a variant of the infeasible primal-dual interior point algorithm by Kojima, Megiddo, and Mizuno [9]. For any given accuracy $\epsilon > 0$ required for the total complementarity, any tolerance $\epsilon_p > 0$ for the primal feasibility, any tolerance $\epsilon_d > 0$ for the dual feasibility define

$$\begin{aligned}
\mathcal{N} = \{(x,y,z) \in \mathcal{Q} :& \\
x_j z_j &\geq \gamma x^T z/n \text{ for } j = 1,2,\ldots,n, && (23a)\\
x^T z &\geq \gamma_p \|Ax-b\| \text{ or } \|Ax-b\| \leq \epsilon_p, && (23b)\\
x^T z &\geq \gamma_d \|A^T y + z - c\| \text{ or } \|A^T y + z - c\| \leq \epsilon_d\} && (23c)
\end{aligned}$$

where $\mathcal{Q} = \{(x,y,z) \in \Re^{n+m+n} : x > 0, z > 0\}$. The constants $0 < \gamma < 1$, $\gamma_p > 0$, $\gamma_d > 0$ will in a weak sense depend on the starting point, but will be chosen so that the neighborhood $\mathcal{N}$ is as large as possible.

Further, let ω^* be any large number and let $\epsilon^* = \min\{\epsilon, \gamma_p\epsilon_p, \gamma_d\epsilon_d\}$. Then the neighborhood [9]

$$\mathcal{N}^* = \{(x,y,z) \in \mathcal{N} : \epsilon^* \leq x^T z \leq \omega^*\} \tag{24}$$

is a compact set. We will show that either $(x^k, y^k, z^k) \in \mathcal{N}^*$ or satisfies the *termination criteria*

$$(x^k)^T z^k \leq \epsilon, \|Ax^k - b\| \leq \epsilon_p \text{ and } \|A^T y^k + z^k - c\| \leq \epsilon_d \tag{25a}$$

or

$$(x^k)^T z^k > \omega^* \tag{25b}$$

after a finite number of steps.

The inequalities $x_j^k z_j^k \geq \gamma (x^k)^T z^k/n$ for $j = 1,2,\ldots,n$, in (23a) prevent the generated sequence $\{(x^k, y^k, z^k)\}$ from reaching the boundary of $\mathcal{Q}$ before the total complementarity $(x^k)^T z^k$ becomes 0. On the other hand, the inequalities (23b) and (23c) prevent the possibility of the generated sequence $\{(x^k, y^k, z^k)\}$ converging to an infeasible complementary solution [9].

We choose any initial point $(x^1, y^1, z^1) \in \mathcal{Q}$ and parameters $\gamma, \gamma_p, \gamma_d$ and ω^* so that $(x^1, y^1, z^1) \in \mathcal{N}$ and $(x^1)^T z^1 \leq \omega^*$. Let $0 < \beta_1 < \beta_2 < \beta_3 < 1$. We state the algorithm below:

Algorithm 3.1 *Inexact Infeasible Primal-Dual Algorithm*
Step 1. *Set* $k = 1$. *Assume* $(x^1, y^1, z^1) \in \mathcal{N}$ *such that* $(x^1)^T z^1 \leq \omega^*$.
Step 2. *If (25a) or (25b) is satisfied then terminate.*
Step 3. *Let* $\mu^k = \beta_1 (x^k)^T z^k/n$.
Step 4. *Compute the inexact solution* $(\Delta x^k, \Delta y^k, \Delta z^k)$ *of (4).*
Step 5. *Let* $0 < \bar{\alpha}^k < 1$ *such that*

$$(x^k, y^k, z^k) + \alpha(\Delta x^k, \Delta y^k, \Delta z^k) \in \mathcal{N} \tag{26a}$$

$$(x^k + \alpha\Delta x^k)^T (z^k + \alpha\Delta z^k) \leq (1 - \alpha(1-\beta_2))(x^k)^T z^k, \tag{26b}$$

hold for every $\alpha \in (0, \bar{\alpha}^k]$.
Step 6. *Choose a primal step length* $\alpha_p^k \in [\bar{\alpha}^k, 1]$, *a dual step length* $\alpha_d^k \in [\bar{\alpha}^k, 1]$ *and a new iterate*

$$(x^{k+1}, y^{k+1}, z^{k+1}) = (x^k + \alpha_p^k \Delta x^k, y^k + \alpha_d^k \Delta y^k, z^k + \alpha_d^k \Delta z^k) \tag{27}$$

such that

$$\begin{aligned}
(x^{k+1}, y^{k+1}, z^{k+1}) &\in \mathcal{N}, && (28a)\\
(x^{k+1})^T z^{k+1} &\leq (1 - \bar{\alpha}^k(1-\beta_3))(x^k)^T z^k. && (28b)
\end{aligned}$$

Step 7. *Increase* k *by 1. Go to Step 2.*

In this algorithm we take relatively short steps when the search directions are computed to a relatively low accuracy. In Section 4 we will show the existence of an $\bar{\alpha}^k > 0$ for all k as long as (25a) is not satisfied.

4 Convergence

For $(x^k, y^k, z^k) \in \mathcal{N}$ the following inequalities will be used in the discussion of the algorithm.

$$(x^k)^T z^k \geq \gamma_p \|Ax^k - b\| \text{ or } \|Ax^k - b\| \leq \epsilon_p \tag{29a}$$
$$(x^k)^T z^k \geq \gamma_d \|A^T y^k + z^k - c\| \text{ or } \|A^T y^k + z^k - c\| \leq \epsilon_d \tag{29b}$$
$$(x^k)^T \Delta z^k + (\Delta x^k)^T z^k = -(1 - \beta_1)(x^k)^T z^k \tag{29c}$$
$$x_i^k \Delta z_i^k + \Delta x_i^k z_i^k = \beta_1 (x^k)^T z^k / n - x_i^k z_i^k \tag{29d}$$

Inequalities (29a) and (29b) follow from $(x^k, y^k, z^k) \in \mathcal{N}$ and (23b) and (23c). Equation (29c) and (29d) follow from simple manipulations of (15b) with $\mu^k = \beta_1 (x^k)^T z^k / n$.

Let

$$0 \leq \eta_{\max} < \min \left\{ \frac{\beta_1}{\max\{\gamma_p, \gamma_d\}}, 1 \right\}. \tag{30}$$

The neighborhood $\mathcal{N}$ is made large by making γ_p and γ_d small which implies that in most cases the only restriction on $\eta_{\max}$ is $0 \leq \eta_{\max} < 1$.

The inexact Newton direction $(\Delta x^k, \Delta y^k, \Delta z^k)$ computed in (16) and (15b) satisfies

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Z^k & 0 & X^k \end{bmatrix} \begin{pmatrix} \Delta x^k \\ \Delta y^k \\ \Delta z^k \end{pmatrix} = \begin{pmatrix} b - Ax^k \\ c - A^T y^k - z^k \\ \mu^k e - X^k Z^k e \end{pmatrix} + \begin{pmatrix} \bar{r}^k \\ \hat{r}^k \\ 0 \end{pmatrix} \tag{31}$$

where the residual satisfies

$$\|(\bar{r}^k, \hat{r}^k)\| \leq \eta^k \, (x^k)^T z^k \tag{32}$$

for

$$0 \leq \eta^k \leq \eta_{\max}. \tag{33}$$

The coefficient matrix F'_μ on the left hand side of (31) is nonsingular and continuous for $(x, y, z) \in \mathcal{N}^*$ defined in (24). Since $\|(\bar{r}^k, \hat{r}^k)\|$ in (32) is bounded for $(x^k, y^k, z^k) \in \mathcal{N}^*$ and $\eta^k \leq \eta_{\max}$ the inexact Newton direction $(\Delta x^k, \Delta y^k, \Delta z^k)$ in (31) is well defined (in the sense that for a given residual the solution of (31) is unique) and is bounded over the compact set $\mathcal{N}^*$. Hence there exists a positive constant τ such that the inexact Newton direction $(\Delta x^k, \Delta y^k, \Delta z^k)$ from (31) computed at Step 4 of every interior point iteration satisfies the inequalities

$$|\Delta x_i^k \Delta z_i^k - \gamma (\Delta x^k)^T \Delta z^k / n| \leq \tau \text{ and } |(\Delta x^k)^T \Delta z^k| \leq \tau. \tag{34}$$

We will now show that for an inexact Newton direction from Step 4 in Algorithm 3.1, we can find $0 < \bar{\alpha}^k < 1$ so that the conditions (26a) and (26b) in Step 5 are satisfied. It will be shown that for a given $\mathcal{N}^*$ there exists a $0 < \alpha^*$ so that $\alpha^* \leq \bar{\alpha}^k$. Let $\xi(\alpha)$ and $\zeta(\alpha)$ be defined by

$$\xi(\alpha) \equiv b - A(x^k + \alpha \Delta x^k) = \xi^k - \alpha A \Delta x^k \tag{35a}$$
$$\zeta(\alpha) \equiv c - A^T (y^k + \alpha \Delta y^k) - z^k = \zeta^k - \alpha (A^T \Delta y^k + \Delta z^k). \tag{35b}$$

For Δy^k and Δx^k given by (16), and Δz^k by (15b) it follows from from (31) that the expressions for $\xi(\alpha)$ and $\zeta(\alpha)$ simplify to

$$\xi(\alpha) = (1 - \alpha)\xi^k - \alpha \bar{r}^k \tag{36a}$$
$$\zeta(\alpha) = (1 - \alpha)\zeta^k - \alpha \hat{r}^k. \tag{36b}$$

Define the real-valued functions $f_i, i = 1, 2, \ldots, n$, g_p, g_d, and h as follows:

$$\begin{aligned}
f_i(\alpha) &= (x_i^k + \alpha\Delta x_i^k)(z_i^k + \alpha\Delta z_i^k) - \gamma(x^k + \alpha\Delta x^k)^T(z^k + \alpha\Delta z^k)/n \\
g_p(\alpha) &= (x^k + \alpha\Delta x^k)^T(z^k + \alpha\Delta z^k) - \gamma_p\|\xi(\alpha)\| \\
g_d(\alpha) &= (x^k + \alpha\Delta x^k)^T(z^k + \alpha\Delta z^k) - \gamma_d\|\zeta(\alpha)\| \\
h(\alpha) &= (1 - \alpha(1 - \beta_2))(x^k)^T z^k - (x^k + \alpha\Delta x^k)^T(z^k + \alpha\Delta z^k).
\end{aligned}$$

Consider Step 5 in the Inexact Infeasible Primal-Dual Algorithm. From the definition of the neighborhood $\mathcal{N}$ condition (26a) is equivalent to

$$\begin{gathered}
f_i(\alpha) \geq 0 \ (i = 1, 2, \ldots, n) \\
g_p(\alpha) \geq 0 \text{ or } \|\xi(\alpha)\| \leq \epsilon_p \\
g_d(\alpha) \geq 0 \text{ or } \|\zeta(\alpha)\| \leq \epsilon_d
\end{gathered}$$

for $0 < \alpha \leq \bar{\alpha}^k$. Similarly, (26b) is equivalent to

$$h(\alpha) \geq 0$$

for $0 < \alpha \leq \bar{\alpha}^k$.

Consider the function g_p and the simplified expression for ξ in (36a). Then

$$\begin{aligned}
g_p(\alpha) &\geq (x^k + \alpha\Delta x^k)^T(z^k + \alpha\Delta z^k) - \gamma_p(1 - \alpha)\|\xi^k\| - \gamma_p\alpha\|\bar{r}^k\| \\
&\geq (x^k)^T z^k + \alpha[(\Delta x^k)^T z^k + (x^k)^T\Delta z^k] + \alpha^2(\Delta x^k)^T\Delta z^k \\
&\quad -(1 - \alpha)(x^k)^T z^k - \gamma_p\alpha\|\bar{r}^k\| \\
&\geq \alpha\beta_1(x^k)^T z^k + \alpha^2(\Delta x^k)^T\Delta z^k - \gamma_p\alpha\|\bar{r}^k\| \qquad (37)
\end{aligned}$$

using (23b), and (29c). By (32) and the property of the norm (14)

$$\|\bar{r}^k\| \leq \|(\bar{r}^k, \hat{r}^k)\| \leq \eta^k\,(x^k)^T z^k,$$

and $\eta^k \leq \eta_{\max}$

$$g_p(\alpha) \geq \alpha\,(\beta_1 - \eta_{\max}\,\gamma_p)\,(x^k)^T z^k + \alpha^2(\Delta x^k)^T\Delta z^k. \qquad (38)$$

For $(x^k, y^k, z^k) \in \mathcal{N}^*$ and (34) $(x^k)^T z^k \geq \epsilon^*$ and $(\Delta x^k)^T\Delta z^k \geq -\tau$

$$g_p(\alpha) \geq \alpha\,[(\beta_1 - \eta_{\max}\,\gamma_p)\,\epsilon^* - \alpha\tau] \qquad (39)$$

Hence, if $0 \leq \alpha \leq \frac{(\beta_1 - \eta_{\max}\,\gamma_p)\epsilon^*}{\tau}$ and $g_p(0) \geq 0$ then $g_p(\alpha) \geq 0$.

Consider $g_p(0) < 0$. From the definition of g_p this is equivalent to $\gamma_p\|Ax^k - b\| > (x^k)^T z^k$ and from (36a)

$$\begin{aligned}
\|A(x^k + \alpha\Delta x^k) - b\| &\leq (1 - \alpha)\|Ax^k - b\| + \alpha\,\eta^k(x^k)^T z^k \\
&\leq (1 - \alpha)\|Ax^k - b\| + \alpha\,\eta^k\,\gamma_p\|Ax^k - b\| \\
&\leq (1 - \alpha(1 - \beta_1))\|Ax^k - b\| \\
&\leq \|Ax^k - b\| \qquad (40)
\end{aligned}$$

from the choice of η^k in (30). From the definition of $\mathcal{N}$ and Step 5 of Algorithm 3.1 we know the iterate $(x^k, y^k, z^k) \in \mathcal{N}$ for all k. Thus $g_p(0) < 0$ implies that $\|Ax^k - b\| \leq \epsilon_p$. From (40) it follows that $\|A(x^k + \alpha\Delta x^k) - b\| \leq \epsilon_p$ if $g_p(0) < 0$ for $\alpha \leq 1$.

Similarly, $\|\hat{r}^k\| \leq \|(\bar{r}^k, \hat{r}^k)\| \leq \eta^k\,(x^k)^T z^k$, and

$$\begin{aligned}
g_d(\alpha) &\geq (x^k + \alpha\Delta x^k)^T(z^k + \alpha\Delta z^k) - \gamma_d(1 - \alpha)\|\zeta^k\| - \alpha\gamma_d\|\hat{r}^k\| \\
&\geq \alpha\beta_1(x^k)^T z^k + \alpha^2(\Delta x^k)^T\Delta z^k - \alpha\gamma_d\|\hat{r}^k\| \\
&\geq \alpha\,(\beta_1 - \eta_{\max}\,\gamma_d)\,(x^k)^T z^k + \alpha^2(\Delta x^k)^T\Delta z^k \\
&\geq \alpha\,[(\beta_1 - \eta_{\max}\,\gamma_d)\,\epsilon^* - \alpha\tau]
\end{aligned}$$

Hence, if $0 \leq \alpha \leq \frac{(\beta_1 - \eta_{\max} \gamma_d)\epsilon^*}{\tau}$, and $g_d(0) \geq 0$ then $g_d(\alpha) \geq 0$.

Next consider

$$\begin{aligned} f_i(\alpha) &= (x_i^k + \alpha\Delta x_i^k)(z_i^k + \alpha\Delta z_i^k) - \gamma(x^k + \alpha\Delta x^k)^T(z^k + \alpha\Delta z^k)/n \\ &= (x_i^k z_i^k - \gamma(x^k)^T z^k/n)(1-\alpha) + \alpha\beta_1(1-\gamma)(x^k)^T z^k/n + \\ &\quad (\Delta x_i^k \Delta z_i^k - \gamma(\Delta x^k)^T \Delta z^k/n)\alpha^2 \\ &\geq \beta_1(1-\gamma)(\epsilon^*/n)\alpha - \tau\alpha^2 = \alpha[\beta_1(1-\gamma)(\epsilon^*/n) - \tau\alpha]. \end{aligned} \tag{41}$$

Inequality (41) follows from application of (29c), (29d) and (34) on the expression for $f_i(\alpha)$ above.

Finally, consider

$$\begin{aligned} h(\alpha) &= (1 - \alpha(1-\beta_2))(x^k)^T z^k - (x^k + \alpha\Delta x^k)^T(z^k + \alpha\Delta z^k) \\ &= \alpha(\beta_2 - \beta_1)(x^k)^T z^k + \alpha^2(\Delta x^k)^T \Delta z^k \\ &\geq \alpha[(\beta_2 - \beta_1)\epsilon^* - \alpha\tau]. \end{aligned} \tag{42}$$

The inequality for $h(\alpha)$ (42) follows from application of (29c) and (34) on $h(\alpha)$ above.

Let $\eta_{\max}$ be given by (30), $\bar{\gamma} = \max\{\gamma_p, \gamma_d\}$ and define

$$\alpha^* = \min\left\{1, \frac{(\beta_1 - \eta_{\max}\bar{\gamma})\epsilon^*}{\tau}, \frac{\beta_1(1-\gamma)\epsilon^*}{n\tau}, \frac{(\beta_2 - \beta_1)\epsilon^*}{\tau}\right\}. \tag{43}$$

For α^* defined in (43) we observe that

$$f_i(\alpha) \geq 0 \ (i = 1, 2, \ldots, n)$$

$$\begin{aligned} & g_p(\alpha) \geq 0 \text{ if } g_p(0) = (x^k)^T z^k - \gamma_p\|Ax^k - b\| \geq 0 \\ & \|A(x^k + \alpha\Delta x^k) - b\| \leq \epsilon_p \text{ if } g_p(0) < 0 \end{aligned} \tag{44a}$$

$$\begin{aligned} & g_d(\alpha) \geq 0 \text{ if } g_d(0) = (x^k)^T z^k - \gamma_d\|A^T y^k + z^k - c\| \geq 0 \\ & \|A^T(y^k + \alpha\Delta y^k) + (z^k + \alpha\Delta z^k) - c\| \leq \epsilon_d \text{ if } g_d(0) < 0 \end{aligned} \tag{44b}$$

$$h(\alpha) \geq 0$$

hold for every $\alpha \in [0, \alpha^*]$.

Note that $\eta^k \gamma_d < \beta_1$ for $\eta^k \leq \eta_{\max}$ and $\eta_{\max}$ defined by (30). By a similar argument as for (44a), observation (44b) follows.

Thus, we have shown that there exists an $\alpha^* > 0$ such that $\bar{\alpha}^k \geq \alpha^*$ in Algorithm 3.1 for all k. By the construction of the real-valued functions $f_i (i = 1, \ldots, n), g_p, g_d$, and h, this is equivalent to saying that (26a) and (26b) hold for every $\alpha \in (0, \bar{\alpha}^k]$.

Consider Step 6 in the algorithm. Conditions (28a) and (28b) are satisfied for a common primal and dual steplength $\alpha_p^k = \alpha_p^k = \bar{\alpha}^k$ and since $\beta_2 < \beta_3$ (28b) is satisfied with strict inequality. Further, note that the right hand side of (28b) is constant with respect to α_p^k and α_p^k.

Theorem 4.1 *Define* $\psi^k = \max\left\{\|Ax^k - b\|, \|A^T y^k + z^k - c\|, (x^k)^T z^k\right\}$. *Let the norm of the residual* $(\bar{r}^k, \hat{r}^k)$ *in (16) satisfy*

$$\|(\bar{r}^k, \hat{r}^k)\| \leq \eta^k (x^k)^T z^k, \ \textit{for } 0 \leq \eta^k \leq \eta_{\max} \tag{45}$$

where $\eta_{\max}$ *satisfies (30). Then there exists* $\tilde{\alpha} \in (0, 1)$ *so that*

$$\psi^{k+1} \leq (1 - \tilde{\alpha})\psi^k \tag{46}$$

where $\tilde{\alpha} = \alpha^*(1 - \max\{\beta_3, \eta_{\max}\})$ *and* α^* *is given by (43).*

Proof: From (36a) we have

$$\begin{aligned}\|\xi(\alpha_p^k)\| &\leq (1-\alpha_p^k)\|\xi^k\| + \alpha_p^k\|\bar{r}^k\| \\ &\leq (1-\alpha_p^k)\|\xi^k\| + \alpha_p^k\eta_{\max}(x^k)^T z^k \\ &\leq (1-\alpha_p^k(1-\eta_{\max}))\psi^k, \qquad (47)\end{aligned}$$

since $\max\left\{\|\xi^k\|, (x^k)^T z^k\right\} \leq \psi^k$. Similarly from (36b) we get

$$\|\zeta(\alpha_d^k)\| \leq (1-\alpha_d^k(1-\eta_{\max}))\psi^k. \qquad (48)$$

From (28b)

$$\begin{aligned}(x^{k+1})^T z^{k+1} &\leq (1-\bar{\alpha}^k(1-\beta_3))(x^k)^T z^k \qquad (49)\\ &\leq (1-\bar{\alpha}^k(1-\beta_3))\psi^k. \qquad (50)\end{aligned}$$

Using (47), (48) and (50) with ψ^{k+1} we get

$$\begin{aligned}\psi^{k+1} &\leq \max\left\{1-\alpha_p^k(1-\eta_{\max}),\ 1-\alpha_d^k(1-\eta_{\max}),\ 1-\bar{\alpha}^k(1-\beta_3)\right\}\psi^k \\ &\leq \max\left\{1-\bar{\alpha}^k(1-\eta_{\max}),\ 1-\bar{\alpha}^k(1-\beta_3)\right\}\psi^k \\ &\leq \left(1-\bar{\alpha}^k\left(1-\max\left\{\eta_{\max},\beta_3\right\}\right)\right)\psi^k \\ &\leq \left(1-\alpha^*\left(1-\max\left\{\eta_{\max},\beta_3\right\}\right)\right)\psi^k. \ \blacksquare\end{aligned}$$

Theorem 4.1 implies that $\|\xi^k\| \to 0$, $\|\zeta^k\| \to 0$ and $(x^{k+1})^T z^{k+1} \to 0$ as $k \to \infty$. This contradicts the assumption (24), i.e. $(x^k)^T z^k \geq \epsilon^*$ for all k. Thus as $k \to \infty$ there exists a k after which the algorithm generates a point $(x^{k+1}, y^{k+1}, z^{k+1})$ which is either an approximate optimal solution $(x^{k+1}, y^{k+1}, z^{k+1})$ satisfying (25a) or satisfies (25b). Therefore Algorithm 3.1 is globally convergent.

It follows from Theorem 4.1 and the proof that the sequences $\{\psi^k\}$, $\{(x^k)^T z^k\}$, and $\{\max\left\{\|Ax^k-b\|, \|A^Ty^k+z^k-c\|\right\}\}$ are q-linearly convergent.

When we compute an exact search direction $(\Delta x^k, \Delta y^k, \Delta z^k)$ at every Step 4 of Algorithm 3.1, i.e. $\eta^k = 0$, Algorithm 3.1 reduces to the globally convergent infeasible interior point algorithm of Kojima, Megiddo and Mizuno [9]. A mixed interior point algorithm where we alternate between a direct method and an iterative method to solve the linear system (16) at distinct interior point iterations is globally convergent if the norm of $(\bar{r}^k, \hat{r}^k)$ is bounded as in (32). Hence the class of preconditioners in [1, 16] can be applied in the implementation of mixed interior point algorithm, and still maintain global convergence.

For $u \in \Re^m$, $v \in \Re^n$, and $w \in \Re^n$ then

$$|||(u,v,w)||| = \max\{\|u\|, \|v\|, \|w\|_1\} \qquad (51)$$

is a norm in $u \in \Re^{m+n+n}$. Theorem 4.1 shows that for the function F (2) in the KKT condition inequality (46) is equivalent to

$$|||F(x^{k+1}, y^{k+1}, z^{k+1})||| \leq (1-\tilde{\alpha})|||F(x^k, y^k, z^k)|||$$

Classical stopping criterion [6] includes

$$\frac{\|Ax^k-b\|}{\max\{1,\|b\|\}} + \frac{\|A^Ty^k+z^k-c\|}{\max\{1,\|c\|\}} + \frac{(x^k)^T z^k}{\max\{1,\|b\|\ \|c\|\}} \leq \varepsilon \qquad (52)$$

Let $\lambda_1, \lambda_2, \lambda_3$ be positive numbers so that $\lambda_1+\lambda_2+\lambda_3 = 1$ and for $u \in \Re^m$, $v \in \Re^n$, and $w \in \Re^n$ define

$$|||(u,v,w)|||_+ = \lambda_1\|u\| + \lambda_2\|v\| + \lambda_3\|w\|_1. \qquad (53)$$

With appropriate choices of λ_i and ε_+ the stopping criteria (52) is equivalent to

$$|||F(x^k, y^k, z^k)|||_+ \leq \varepsilon_+$$

Corollary 4.1 *Under the assumptions of Theorem 4.1*

$$\|\!|F(x^{k+1}, y^{k+1}, z^{k+1})|\!\|_+ \leq (1-\tilde{\alpha})\|\!|F(x^k, y^k, z^k)|\!\|_+$$

Proof: This follows from the (47), (48), (50) and

$$\begin{aligned} \|\xi(\alpha_p^k)\| &\leq (1-\tilde{\alpha})\|\xi^k\| \\ \|\zeta(\alpha_d^k)\| &\leq (1-\tilde{\alpha})\|\zeta^k\| \\ (x^{k+1})^T z^{k+1} &\leq (1-\tilde{\alpha})(x^k)^T z^k. \end{aligned}$$

Taking a convex combination of the above inequalities yields the desired result. ■

In the inexact Newton method [4] the residual is required to satisfy $\|r^k\| \leq \eta^k\|F(x^k)\|$. The next result shows that the norms $\|\!|\cdot|\!\|$ and $\|\!|\cdot|\!\|_+$ can be used to control the accuracy in (16).

Corollary 4.2 *Choose* $\gamma_p = \gamma_d\ (= \bar{\gamma})$ *and* $\eta_{\max} < \min\{\bar{\gamma}, \beta_1, \beta_1/\bar{\gamma}\}$*. Assume that*

$$\|(\bar{r}^k, \hat{r}^k)\| \leq \eta^k\ \|\!|F(x^k, y^k, z^k)|\!\| \ \textit{where}\ \eta^k \leq \eta_{\max} \tag{54}$$

then there exists $\tilde{\alpha} \in (0,1)$ *so that*

$$\|\!|F(x^{k+1}, y^{k+1}, z^{k+1})|\!\| \leq (1-\tilde{\alpha})\|\!|F(x^k, y^k, z^k)|\!\|.$$

Proof: To guarantee the existence of $\bar{\alpha}^k$ we only need to consider the case $g_p(0) \geq 0$ and $g_d(0) \geq 0$. Then

$$\|\!|F(x^k, y^k, z^k)|\!\| \leq \max\{\frac{1}{\bar{\gamma}}, 1\}(x^k)^T z^k.$$

The existence of $\tilde{\alpha}$ follows from Theorem 4.1 provided

$$\max\{\frac{1}{\bar{\gamma}}, 1\}\eta^k < \min\left\{\frac{\beta_1}{\bar{\gamma}}, 1\right\}. \ ■$$

From the property of the norm (14) and (51) $\|\!|(u,v,0)|\!\| \leq \|(u,v)\|$, hence (54) implies

$$\|\!|(\bar{r}^k, \hat{r}^k, 0)|\!\| \leq \eta^k\ \|\!|F(x^k, y^k, z^k)|\!\| \ \text{where}\ \eta^k \leq \eta_{\max}$$

A similar result can be obtained using the norm $\|\!|\cdot|\!\|_+$.

Corollary 4.3 *Let*

$$\lambda = \frac{\lambda_1}{\gamma_p} + \frac{\lambda_2}{\gamma_d} + \lambda_3.$$

Choose $\eta_{\max} < \min\{\beta_1/\max\{\gamma_p, \gamma_d\}, 1\}/\lambda$*. Assume that*

$$\|(\bar{r}^k, \hat{r}^k)\| \leq \eta^k\ \|\!|F(x^k, y^k, z^k)|\!\|_+ \ \textit{where}\ \eta^k \leq \eta_{\max} \tag{55}$$

then there exists $\tilde{\alpha} \in (0,1)$ *so that*

$$\|\!|F(x^{k+1}, y^{k+1}, z^{k+1})|\!\|_+ \leq (1-\tilde{\alpha})\|\!|F(x^k, y^k, z^k)|\!\|_+.$$

Proof: To guarantee the existence of $\bar{\alpha}^k$ we only need to consider the case $g_p(0) \geq 0$ and $g_d(0) \geq 0$. Then

$$\|\!|F(x^k, y^k, z^k)|\!\|_+ \leq \lambda(x^k)^T z^k.$$

The existence of $\tilde{\alpha}$ follows from Theorem 4.1 provided

$$\lambda\eta^k < \min\left\{\frac{\beta_1}{\max\{\gamma_p, \gamma_d\}}, 1\right\}. \ ■$$

Since $\|\!|(u,v,0)|\!\|_+ \leq \|(u,v)\|$ (55) implies

$$\|\!|(\bar{r}^k, \hat{r}^k, 0)|\!\|_+ \leq \eta^k\ \|\!|F(x^k, y^k, z^k)|\!\|_+ \ \text{where}\ \eta^k \leq \eta_{\max}.$$

References

[1] V. Baryamureeba, T. Steihaug, and Y. Zhang, *Properties of a class of preconditioners for weighted least squares problems,* Technical Report No. 170, Department of Informatics, University of Bergen, N-5020 Bergen, Norway, April 30, 1999 (Revised July 6, 1999).

[2] S. Bellavia, *Inexact interior point method,* Journal of Optimization Theory and Applications, Vol. 96, pp. 109–121, 1998.

[3] S. Bellavia, *Numerical performance of an inexact interior point method,* R. De Leone, A. Murli, P.M. Pardalos, and G. Toraldo (Eds), High Performance Algorithms and Software in Nonlinear Optimization, Kluwer Academic Publishers, 1998, pp. 43–51.

[4] R.S. Dembo, S.C. Eisenstat, and T. Steihaug, *Inexact Newton Methods,* SIAM J. Numer. Anal., Vol. 19, pp. 400–408, 1982.

[5] S.C. Eisenstat, and H.F. Walker, *Globally convergent inexact Newton methods,* SIAM J. Optimization, Vol. 4, pp. 393–422, 1994.

[6] R.W. Freund, F. Jarre, and S. Mizuno, *Convergence of a class of inexact interior point algorithms for linear programs,* Mathematics of Operations Research, Vol. 24, pp. 50–71, 1999.

[7] S. Ito *Inexact implementation of interior point algorithms for optimal control problems*, Lecture Notes in Num. Appl. Anal., Vol 14, pp. 245–248, 1995.

[8] S. Ito, C.T. Kelley, and E.W. Sachs, *Inexact primal-dual interior-point iteration for linear programs in functional spaces*, Computational Optimization and Applications, Vol. 4, pp. 189–201, 1995.

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

[10] I.J. Lustig, *Feasibility issues in primal-dual interior point method for linear programming,* Mathematical Programming, Vol. 49, pp. 145–162, 1990/91.

[11] I.J. Lustig, R.E. Marsten, and D.F. Shanno, *Computational experience with a primal-dual interior point method for linear programming,* Linear Algebra and its Applications, Vol. 152, pp. 191–222, 1991.

[12] R. Marsten, R. Subramanian, M. Saltzman, I.J. Lustig, and D.F. Shanno, *Interior point methods for linear programming: Just call Newton, Lagrange, and Fiacco and McCormick!,* Interfaces, Vol. 20, pp. 105–116, 1990.

[13] S. Mizuno and F. Jarre, *Global and polynomial-time convergence of an infeasible–interior–point algorithm using inexact computation*, Mathematical Programming, Vol. 84, pp. 105–122, 1999.

[14] L. Portugal, L. Fernandes, and J. Judice, *A truncated Newton interior point algorithm for the solution of a multicommodity spatial equilibrium model,* Complementarity and Variational Problems, edited by J. S. Pang and M. Ferris, SIAM, 1997, pp. 315–344.

[15] L.F. Portugal, M.G.C. Resende, G. Veiga, and J. Judice, *A truncated primal–infeasible dual–feasible interior point network flow method,* Networks, Vol. 35, pp. 91–108, 2000.

[16] W. Wang and D.P. O'Leary, *Adaptive use of iterative methods in interior point methods for linear programming,* Computer Science Department Report CS–TR–3560, Institute for Advanced Computer Studies Report UMIACS–95–111, University of Maryland, November 1995.

[17] S. Wright, *Primal–dual interior–point methods.* SIAM, Philadelphia, 1997.