Asymptotic convergence in a generalized predictor–corrector method

Sanjay Mehrotra *

Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208-3119, USA

Received November 1992; revised September 1994

Abstract

The asymptotic convergence properties of a generalized predictor–corrector method are analyzed. This method is based on making a sequence of corrections to the primal–dual affine scaling (predictor) direction. It is shown that a method making r corrections to a predictor direction has the Q-order convergence of order $r + 2$. It is also shown that asymptotically the problem can be solved by only computing corrections to the predictor direction.

Keywords: Linear programming; Primal–dual methods; Predictor–corrector methods; Asymptotic convergence analysis

1. Introduction

In this paper we analyze the asymptotic convergence properties of a variation of the generalized (multiple-correction) predictor–corrector method for linear programming described and implemented in Mehrotra [6,8], Lustig, Marsten and Shanno [4] and Carpenter, Lustig, Mulvey and Shanno [2]. We show that in this method, if the predictor step is followed by r corrector steps, its Q-order convergence is at least $r+2$. The results in this paper partially explain the superior performance of the practical implementations of this method.

The results proved in this paper can be viewed as a generalization of a similar result for solving a system of nonlinear equations (see Ortega and Rheinboldt [13, pp. 315–316]). In our case the Jacobian matrix of the primal–dual system can be singular and the "simplified Newton steps" are restricted to the positive orthant.

* Supported in part by a grant from the GTE Laboratories and by the grant CCR-9019469 from the National Science Foundation.

SSDI 0025-5610(95)00028-3

It was recently shown by Mehrotra [9] and Ye, Güler, Tapia and Zhang [19] that the Mizuno–Todd–Ye predictor-corrector primal–dual interior point method converges Q-quadratically. The results proved in these papers are derived from the properties of the primal–dual affine scaling direction near the central trajectory. The properties of this direction combined with the centering step in Mizuno–Todd–Ye $O(\sqrt{n}L)$ iteration method lead to the Q-quadratic convergence result.

The idea of a generalized predictor-corrector method in Mehrotra [6, 8] is to make a sequence of corrections to the predictor step. Tapia, Zhang, Saltzman and Weiser [15] observed that the correction steps used in the computational studies of Mehrotra [6, 8], Lustig, Marsten and Shanno [4], Carpenter, Lustig, Mulvey and Shanno [2] can be viewed as "simplified Newton steps." In this step the Jacobian matrix of the primal–dual system is kept fixed for some "sub-iterations." Under some assumptions on the performance of the method, Tapia et al. [15] also showed that the predictor–corrector method with one correction has cubic convergence. The results proved here do not make any such assumptions.

This paper is organized as follows. Section 2 is divided into three subsections. Section 2.1 gives some desirable properties of the optimal solution set of a linear program. Section 2.2 describes a generic generalized predictor–corrector method which is analyzed in this paper. Some results that have been proved in earlier papers are restated in Section 2.3. Sections 3 and 4 contain the analysis for the results proved in this paper.

2. Background material for analysis

2.1. Properties of optimal solutions of a linear program

We consider the linear program (LP):

$$\begin{array}{ll} \min & c^{\mathrm{T}}x, \\ \text{s.t.} & Ax = b, \quad x \geqslant 0, \end{array}$$

and its dual (LD):

$$\begin{array}{ll} \max & b^{\mathrm{T}}z, \\ \text{s.t.} & A^{\mathrm{T}}z + y = c, \quad y \geqslant 0, \end{array}$$

where $A \in \mathbb{R}^{m\times n}$, $c, x, y \in \mathbb{R}^n$, and $b, z \in \mathbb{R}^m$. We assume that A has full row rank. The optimality conditions for linear programming ensure that feasible solutions x^* and (y^*, z^*) respectively for (LP) and (LD) are optimal if and only if,

$$x_i^* y_i^* = 0 \quad \text{for } i = 1, 2, \ldots, n.$$

Let $\sigma(x)$ represent the index set of positive components in $x \geqslant 0$, that is,

$$\sigma(x) = \{i : x_i > 0\}.$$

Among all the optimal solutions for (LP) and (LD), there exists at least one optimal solution pair (x^*, y^*) which is strictly complementary (see Schrijver [14, pp. 95–96]), that is,

$$\sigma(x^*) \cap \sigma(y^*) = \emptyset \quad \text{and} \quad \sigma(x^*) \cup \sigma(y^*) = \{1, 2, \ldots, n\}$$

for every complementary solution (x^*, y^*). Moreover, $\sigma(x^*)$ and $\sigma(y^*)$ remain invariant for every strictly complementary solution (x^*, y^*). Hence, we can denote $\sigma(x^*)$ by σ^* for (LP) and let $\bar{\sigma}^* = \{1, \ldots, n\} \setminus \sigma^*$. One can further show that

$$\sigma(x^*) \subset \sigma^* \quad \text{and} \quad \sigma(y^*) \subset \bar{\sigma}^*$$

for every complementary solution (x^*, y^*).

2.2. *A generic generalized predictor–corrector primal–dual method*

We first give a generic generalized predictor–corrector primal–dual method. For details on the primal–dual methods we refer the reader to Kojima, Mizuno and Yoshise [3], Lustig, Marsten and Shanno [5], Mizuno, Todd and Ye [11], and Monteiro and Adler [12].

2.2.1. *Computation of the predictor direction*

Let us assume that at the beginning of iteration l a feasible solution $x^l > 0$, $y^l > 0$, z^l is available and define $X^l = \operatorname{diag}(x_i^l)$. Similar notation is followed to define diagonal matrices associated with other vectors. We first compute a direction by solving the following system of linear equations:

$$J(x^l, y^l) \begin{pmatrix} \bar{d}_x^0 \\ \bar{d}_y^0 \\ \bar{d}_z^0 \end{pmatrix} \equiv \begin{bmatrix} Y^l & X^l & 0 \\ 0 & I & A^{\mathrm{T}} \\ A & 0 & 0 \end{bmatrix} \begin{pmatrix} \bar{d}_x^0 \\ \bar{d}_y^0 \\ \bar{d}_z^0 \end{pmatrix} = \begin{pmatrix} -X^l y^l \\ 0 \\ 0 \end{pmatrix}. \tag{1}$$

The direction given by $(\bar{d}_x^0, \bar{d}_y^0, \bar{d}_z^0)$ is called the primal–dual affine scaling direction. The primal–dual affine scaling direction computed in this way is called the "predictor" direction. The first intermediate point is computed as follows.

$$\begin{aligned} u^1 &\leftarrow x^l + \tau_0 \bar{d}_x^0, \\ v^1 &\leftarrow y^l + \tau_0 \bar{d}_y^0, \\ w^1 &\leftarrow z^l + \tau_0 \bar{d}_z^0. \end{aligned}$$

Here τ_0 is the largest step satisfying (4) below. The direction $-(\bar{d}_x^0, \bar{d}_y^0, \bar{d}_z^0)$ can be viewed as the Newton direction at (x^l, y^l, z^l) for the following system of nonlinear equations:

$$\begin{aligned} Xy &= 0, \\ A^{\mathrm{T}} z + y &= 0, \\ Ax &= 0. \end{aligned}$$

2.2.2. Computation of corrections

Next we describe how corrections are computed in the generic method. The kth correction (d_u^k, d_v^k, d_w^k) is computed by solving:

$$J(x^l, y^l)\begin{pmatrix} d_u^k \\ d_v^k \\ d_w^k \end{pmatrix} = \begin{pmatrix} -U^k v^k \\ 0 \\ 0 \end{pmatrix}. \tag{2}$$

Note that the computations in (2) use the same Jacobian matrix used for computing the predictor direction. Therefore, it can be viewed as a simplified Newton direction. Let

$$\begin{aligned} d_x^{k-1} &= \tau_{k-1}\bar{d}_x^{k-1}, \\ d_y^{k-1} &= \tau_{k-1}\bar{d}_y^{k-1}, \\ d_z^{k-1} &= \tau_{k-1}\bar{d}_z^{k-1}. \end{aligned}$$

Now, a corrected direction is computed as

$$\begin{aligned} \bar{d}_x^k &= d_u^k + d_x^{k-1}, \\ \bar{d}_y^k &= d_v^k + d_y^{k-1}, \\ \bar{d}_z^k &= d_w^k + d_z^{k-1}, \end{aligned} \tag{3}$$

and a new intermediate solution is generated from:

$$\begin{aligned} u^{k+1} &\leftarrow x^l + \tau_k \bar{d}_x^k, \\ v^{k+1} &\leftarrow y^l + \tau_k \bar{d}_y^k, \\ w^{k+1} &\leftarrow z^l + \tau_k \bar{d}_z^k, \end{aligned}$$

where τ_k takes the largest value for which

$$\left(1 - \frac{0.2}{\sqrt{n}}\right)\frac{x_i^l y_i^l}{x^{l\mathrm{T}} y^l} \leqslant \frac{u_i^{k+1} v_i^{k+1}}{(u^{k+1})^{\mathrm{T}} v^{k+1}} \leqslant \left(1 + \frac{0.2}{\sqrt{n}}\right)\frac{x_i^l y_i^l}{x^{l\mathrm{T}} y^l} \tag{4}$$

is satisfied for all i. The constants $(1 - 0.2/\sqrt{n})$ and $(1 + 0.2/\sqrt{n})$ are used in the analysis of the next section.

2.2.3. Centering

The solution available after performing a fixed number (say r) of corrections is taken as an intermediate solution, i.e.,

$$\begin{aligned} \hat{x}^l &\leftarrow u^{r+1}, \\ \hat{y}^l &\leftarrow v^{r+1}, \\ \hat{z}^l &\leftarrow w^{r+1}. \end{aligned}$$

Let $\hat{\mu}_l = (\hat{x}^l)^{\mathrm{T}}\hat{s}^l/n$ and e be a vector of all ones. At this stage, we perform a centering step by solving the system of linear equations:

$$\begin{bmatrix} \hat{Y}^l & \hat{X}^l & 0 \\ 0 & I & A^{\mathrm{T}} \\ A & 0 & 0 \end{bmatrix} \begin{pmatrix} d^l_{\hat{x}} \\ d^l_{\hat{y}} \\ d^l_{\hat{z}} \end{pmatrix} = \begin{pmatrix} -\hat{X}^l \hat{y}^l + \hat{\mu}_l e \\ 0 \\ 0 \end{pmatrix}. \tag{5}$$

The new solution is obtained from

$$\begin{aligned} x^{l+1} &\leftarrow \hat{x}^l + d^l_{\hat{x}}, \\ y^{l+1} &\leftarrow \hat{y}^l + d^l_{\hat{y}}, \\ z^{l+1} &\leftarrow \hat{z}^l + d^l_{\hat{z}}. \end{aligned}$$

We make a few comments about the generalized predictor–corrector method described above.

Remark 1. In this paper we are only interested in the asymptotic analysis of this method. It is sufficient to define the corrected directions using (3) for our analysis. To ensure global convergence of the method a two dimensional line search using $(d_x^{k-1}, d_y^{k-1}, d_z^{k-1})$ and (d_u^k, d_v^k, d_w^k) can be performed. The line search can be based on ensuring condition (4) or similar conditions on the proximity to the central path (see for example Mizuno, Todd and Ye [11]). Alternatively, the Tanabe–Todd–Ye potential function (for example see [16]) can also be used for this line search. A variety of algorithms can be constructed by using combinations of predictor, corrector and centering directions and forcing some path following measure or sufficient potential reduction. The asymptotic properties of predictor and corrector directions depend only on the current iterate, and they are independent of how that iterate was generated.

Remark 2. The analysis in the next section uses a theoretical value of τ_k. Also, the above description performs centering at the end of making all corrections. Excellent computational results were reported in Mehrotra [8] for the method $PCM(2,2,k)$ described in that paper, which uses a more practical (but heuristic) approach to combine the directions $(d_x^{k-1}, d_y^{k-1}, d_z^{k-1})$ and (d_u^k, d_v^k, d_w^k), and a centering direction. Another difference between the method here and the method that has been implemented in [2,8] is that here the simplified Newton step is computed at a feasible interior point.

Remark 3. If no corrections are performed and the predictor step is followed by the centering step, then the generalized predictor–corrector method is same as the Mizuno–Todd–Ye predictor–corrector method.

2.3. *Results from the previous work*

Let $\mathcal{N}(\mu, \epsilon) \equiv \{(x,y) \mid \|Xy/\mu - e\|_2 \leqslant \epsilon\}$. Then, for $(x,y) \in \mathcal{N}(\mu,\epsilon)$,

$$(1-\epsilon) \leqslant \frac{x_i y_i}{\mu} \leqslant (1+\epsilon). \tag{6}$$

The following Lemma is a restatement of Lemma 3.6 in Mehrotra [9].

Lemma 4. *Let (x, y) be a feasible solution and $\mu = x^{\mathrm{T}}y/n$. Assume that (x, y) satisfies (6) for $\epsilon = 1/2$. Then it is possible to generate directions p_x and p_y such that $x^* = x + p_x$ and $y^* = y + p_y$, are strictly complementary optimal solutions. Furthermore, p_x and p_y satisfy*

$$\max_{i \in \sigma^*}(|(p_x)_i/x_i|) \leqslant C^* x^{\mathrm{T}}y/n,$$

$$\max_{i \in \bar{\sigma}^*}(|(p_y)_i/y_i|) \leqslant C^* x^{\mathrm{T}}y/n,$$

for some constant C^*.

The next theorem shows that the predictor step is close to an "optimal direction" constructed for Lemma 4.

Theorem 5 (Theorem 3.7 in [9]). *Let x, y, p_x, p_y, x^*, y^* be as in Lemma 4. Let $\bar{d}_x^0$, $\bar{d}_y^0, \bar{d}_z^0$ be the predictor step computed at (x, y, z). Then, for some constant C,*

$$\left\| \begin{matrix} X^{-1}(p_x - \bar{d}_x^0) \\ Y^{-1}(p_y - \bar{d}_y^0) \end{matrix} \right\| \leqslant C\mu.$$

The following theorem is about the quality of the centering step. This theorem is a consequence of Q-quadratic convergence of Newton's method for the centering problem. It ensures that the centering step would bring the solution closer to the central path, provided that the starting solution is sufficiently close.

Theorem 6 (Monteiro and Adler [12], Mizuno, Todd and Ye [11])). *Let $(\hat{x}, \hat{y})$ be an interior feasible solution, and $\hat{\mu} = \hat{x}^{\mathrm{T}}\hat{y}/n$. Let $(\hat{x}, \hat{y}) \in \mathcal{N}(\hat{\mu}, 2\beta)$, $0 < \beta \leqslant 1/4$ (say $\beta = 1/4$). Let $x^+ = \hat{x} + d_{\hat{x}}$ and $y^+ = \hat{y} + d_{\hat{y}}$ where $(d_{\hat{x}}, d_{\hat{y}})$ is computed from (5) at $(\hat{x}, \hat{y})$. Then, $(x^+, y^+) \in \mathcal{N}(\hat{\mu}, \beta)$.*

3. Properties of corrector directions

This section establishes some basic results on the properties of corrections generated in the generalized predictor–corrector method described in the previous section.

The notation of the previous section is simplified by assuming that the current point is given by (x, y, z). We take $\mu = x^{\mathrm{T}}y/n$. Theorem 7 is the main result of this section. This theorem provides the induction argument for the results proved in the next section. It gives a bound on the quality of the corrected direction in the generalized predictor–corrector method after one additional correction step is performed.

Theorem 7. *Let $(\bar{d}_x^k, \bar{d}_y^k, \bar{d}_z^k)$ be the corrected direction after k corrections in the predictor–corrector method have been performed. Assume that we have constants C_k^1, C_k^2 such that $C_k^1\mu \leqslant 0.01/n$, $C_k^2\mu^{k+1} < 1$, and $(\bar{d}_x^k, \bar{d}_y^k, \bar{d}_z^k)$ satisfies*

$$\frac{|(\tilde{d}_x^k)_i|}{x_i} \leqslant C_k^1 \mu, \quad i \in \sigma^*, \tag{7}$$

$$-C_k^2 \mu^{k+1} \leqslant 1 + \frac{(\tilde{d}_y^k)_i}{y_i} \leqslant C_k^2 \mu^{k+1}, \quad i \in \sigma^*, \tag{8}$$

$$-C_k^2 \mu^{k+1} \leqslant 1 + \frac{(\tilde{d}_x^k)_i}{x_i} \leqslant C_k^2 \mu^{k+1}, \quad i \in \bar{\sigma}^*, \tag{9}$$

$$\frac{|(\tilde{d}_y^k)_i|}{y_i} \leqslant C_k^1 \mu, \quad i \in \bar{\sigma}^*. \tag{10}$$

Let $(x, y) \in \mathcal{N}(\mu, 1/4)$. *Then for*

$$1 \geqslant \tau_k = 1 - \rho_k \mu^{k+1} \geqslant 0; \quad \rho_k = (1 + 10\sqrt{n}) C_k^2, \tag{11}$$

$(u^{k+1}, v^{k+1}) \in \mathcal{N}(u^{k+1\,\mathrm{T}} v^{k+1}/n, 1/2)$. *Now let* $(d_u^{k+1}, d_v^{k+1}, d_w^{k+1})$ *be the* $(k+1)$*st correction, and* $(\tilde{d}_x^{k+1}, \tilde{d}_y^{k+1}, \tilde{d}_z^{k+1})$ *be the* $(k+1)$*st corrected direction. Then we have constants* C_{k+1}^1, C_{k+1}^2 *such that*

$$\frac{|(\tilde{d}_x^{k+1})_i|}{x_i} \leqslant C_{k+1}^1 \mu, \quad i \in \sigma^*, \tag{12}$$

$$-C_{k+1}^2 \mu^{k+2} \leqslant 1 + \frac{(\tilde{d}_y^{k+1})_i}{y_i} \leqslant C_{k+1}^2 \mu^{k+2}, \quad i \in \sigma^*, \tag{13}$$

$$-C_{k+1}^2 \mu^{k+2} \leqslant 1 + \frac{(\tilde{d}_x^{k+1})_i}{x_i} \leqslant C_{k+1}^2 \mu^{k+2}, \quad i \in \bar{\sigma}^*, \tag{14}$$

$$\frac{|(\tilde{d}_y^{k+1})_i|}{y_i} \leqslant C_{k+1}^1 \mu, \quad i \in \bar{\sigma}^*, \tag{15}$$

where $C_{k+1}^1 = C_k^1 + \beta_k \mu^{k+1}$ *and*

$$\begin{aligned} C_{k+1}^2 = \sqrt{4n/3}\,(2 \max\{&C_k^1 (C_k^2 + 2\rho_k), \rho_k C^* (1 + C_k^1 \mu)(1 + (C_k^2 + 2\rho_k)\mu^{k+1})\} \\ &+ (1 + C_k^1 \mu)^{1/2} (C_k^2 + 2\rho_k)^{1/2} \rho_k^{3/2} C^* \mu^{k+1}), \end{aligned}$$

$$\beta_k = (1 + C_k^1 \mu) \rho_k C^* + C_{k+1}^2.$$

The following lemmas and propositions are used in the proof of this Theorem 7. Proposition 8 obtains bounds on quantities that are closely related to $(\tilde{d}_x^k, \tilde{d}_y^k, \tilde{d}_z^k)$. Lemma 9 shows that the duality gap reduces linearly by moving in the corrected direction. Lemma 10 ensures that $(u^{k+1}, v^{k+1}) \in \mathcal{N}(u^{k+1\,\mathrm{T}} v^{k+1}/n, 1/2)$. Lemma 11 provides a bound on the relative distance of a strictly complementary optimal solution from (u^{k+1}, v^{k+1}). Propositions 12, 13 and 14 are preparatory towards Lemma 15. Lemma 15 bounds deviation of $k+1$st correction direction from a direction that generates an optimal solution. The bounds proved in Lemma 15 are central to the proof of Theorem 7.

Proposition 8. *Under the hypothesis of Theorem 7, we have*

$$\frac{|(d_x^k)_i|}{x_i} \leqslant C_k^1 \mu, \quad i \in \sigma^*, \tag{16}$$

$$-(C_k^2 + 2\rho_k)\mu^{k+1} \leqslant \frac{v_i^{k+1}}{y_i} = 1 + \frac{(d_y^k)_i}{y_i} \leqslant (C_k^2 + 2\rho_k)\mu^{k+1}, \quad i \in \sigma^*, \tag{17}$$

$$-(C_k^2 + 2\rho_k)\mu^{k+1} \leqslant \frac{u_i^{k+1}}{x_i} = 1 + \frac{(d_x^k)_i}{x_i} \leqslant (C_k^2 + 2\rho_k)\mu^{k+1}, \quad i \in \bar{\sigma}^*, \tag{18}$$

$$\frac{|(d_y^k)_i|}{y_i} \leqslant C_k^1 \mu, \quad i \in \bar{\sigma}^*. \tag{19}$$

Also,

$$1 - C_k^1\mu \leqslant \frac{u_i^{k+1}}{x_i} \leqslant (1 + C_k^1\mu), \quad i \in \sigma^*, \tag{20}$$

$$1 - C_k^1\mu \leqslant \frac{v_i^{k+1}}{y_i} \leqslant (1 + C_k^1\mu), \quad i \in \bar{\sigma}^*. \tag{21}$$

Proof. The inequality (16) follows since $|(d_x^k)_i| = |\tau_k(\bar{d}_x^k)_i| \leqslant |(\bar{d}_x^k)_i|$ because $\tau_k \leqslant 1$. (19) follows similarly. To prove (17) note that

$$\begin{aligned} 1 + \frac{(d_y^k)_i}{y_i} &= 1 + \tau_k\frac{(\bar{d}_y^k)_i}{y_i} = 1 + \frac{(\bar{d}_y^k)_i}{y_i} + (\tau_k - 1)\frac{(\bar{d}_y^k)_i}{y_i} \\ &\leqslant C_k^2\mu^{k+1} + (1 + C_k^2\mu^{k+1})\rho_k\mu^{k+1} \leqslant (C_k^2 + 2\rho_k)\mu^{k+1}. \end{aligned}$$

Here the first inequality above uses (8) and (11), and the second inequality uses $C_k^2\mu^{k+1} < 1$. The proof for the other side of the inequality in (17) is similar. The proof for (18) is similar to the proof of (17). To prove (20) once again note that

$$\frac{u_i^{k+1}}{x_i} = 1 + \tau_k\frac{(\bar{d}_x^k)_i}{x_i} \leqslant 1 + \tau_k C_k^1\mu \leqslant 1 + C_k^1\mu.$$

The other side of the inequality is proved similarly, and finally, the proof of (21) is similar to the proof of (20).

Lemma 9. *Let $(\bar{d}_x^k, \bar{d}_y^k, \bar{d}_z^k)$ be the direction obtained after k corrections. Then,*

$$(x + \tau_k\bar{d}_x^k)^{\mathrm{T}}(y + \tau_k\bar{d}_y^k) = (1 - \tau_k)x^{\mathrm{T}}y.$$

Proof. Since $(\bar{d}_x^0)^{\mathrm{T}}(\bar{d}_y^0) = 0$ and $x^{\mathrm{T}}\bar{d}_y^0 + y^{\mathrm{T}}\bar{d}_x^0 = -x^{\mathrm{T}}y$, we have

$$(x + \tau_0\bar{d}_x^0)^{\mathrm{T}}(y + \tau_0\bar{d}_y^0) = (1 - \tau_0)x^{\mathrm{T}}y.$$

Now for $k \geqslant 1$, since $(\bar{d}_x^k)^{\mathrm{T}}(\bar{d}_y^k) = 0$ we have

$$(x + \tau_k\bar{d}_x^k)^{\mathrm{T}}(y + \tau_k\bar{d}_y^k) = x^{\mathrm{T}}y + \tau_k(x^{\mathrm{T}}\bar{d}_y^k + y^{\mathrm{T}}\bar{d}_x^k).$$

From equation (2) we know that $Yd_u^k + Xd_v^k = -U^k v^k = -(X + D_x^{k-1})(y + d_y^{k-1})$. Now using (3) we have $Y\tilde{d}_x^k + X\tilde{d}_y^k = -Xy - D_x^{k-1} d_y^{k-1}$, which gives the desired results since ${d_x^{k-1}}^{\mathrm{T}} d_y^{k-1} = 0$. □

The next lemma shows that by moving in directions satisfying the assumptions of Theorem 7 we slowly leave the neighborhood of the central path.

Lemma 10. *Let $x(\tau) = x + \tau\tilde{d}_x^k$, $y(\tau) = y + \tau\tilde{d}_y^k$, $\mu = x^{\mathrm{T}}y/n$, and $\mu(\tau) = x(\tau)^{\mathrm{T}}y(\tau)/n$. Assume that (7–8) is satisfied and $C_k^1\mu \leqslant 0.01/n$. Then for $\tau = 1 - \rho_k\mu^{k+1}$, $\rho_k = (1 + 10\sqrt{n})C_k^2$, we have*

$$(1 - 0.2/\sqrt{n})\frac{x_i y_i}{\mu} \leqslant \frac{x_i(\tau)y_i(\tau)}{\mu(\tau)} \leqslant (1 + 0.2/\sqrt{n})\frac{x_i y_i}{\mu}. \tag{22}$$

Furthermore, if $(x, y) \in \mathcal{N}(\mu, 1/4)$, then $(x(\tau), y(\tau)) \in \mathcal{N}(\mu(\tau), 1/2)$.

Proof. From Lemma 9 we have,

$$\frac{x_i(\tau)y_i(\tau)}{\mu(\tau)} = \frac{x_i y_i}{\mu}\frac{(1 + \tau(\tilde{d}_x^k)_i/x_i)(1 + \tau(\tilde{d}_y^k)_i/y_i)}{1 - \tau},$$

therefore, to show (22) it is sufficient to show that for $\tau = 1 - (1 + 10\sqrt{n})C_k^2\mu^{k+1}$

$$1 - 0.2/\sqrt{n} \leqslant \frac{(1 + \tau(\tilde{d}_x^k)_i/x_i)(1 + \tau(\tilde{d}_y^k)_i/y_i)}{1 - \tau} \leqslant 1 + 0.2/\sqrt{n}. \tag{23}$$

First we show the upper bound in (23). Given any index i, if $i \in \sigma^*$, then using (7) and (8), otherwise $i \in \bar{\sigma}^*$ and using (9) and (10) we have

$$\begin{aligned}
&\frac{(1 + \tau(\tilde{d}_x^k)_i/x_i)(1 + \tau(\tilde{d}_y^k)_i/y_i)}{1 - \tau} \\
&\quad \leqslant \frac{(1 + \tau C_k^1\mu)(1 + \tau(C_k^2\mu^{k+1} - 1))}{1 - \tau} \\
&\quad \leqslant \frac{(1 + C_k^1\mu)(1 + \tau(C_k^2\mu^{k+1} - 1))}{1 - \tau} \\
&\quad = \frac{(1 + C_k^1\mu)\left(1 + [1 - (1 + 10\sqrt{n})C_k^2\mu^{k+1}][C_k^2\mu^{k+1} - 1]\right)}{(1 + 10\sqrt{n})C_k^2\mu^{k+1}} \\
&\quad \leqslant \frac{(1 + C_k^1\mu)\left((2 + 10\sqrt{n})C_k^2\mu^{k+1}\right)}{(1 + 10\sqrt{n})C_k^2\mu^{k+1}} \\
&\quad = \frac{(1 + C_k^1\mu)(1 + 0.2/\sqrt{n})}{1 + 0.1/\sqrt{n}} \\
&\quad \leqslant 1 + 0.2/\sqrt{n}.
\end{aligned}$$

The last inequality above follows by using $C_k^1\mu \leqslant 0.01/n < 0.1/\sqrt{n}$. Now the lower bound in (23) is proved. Once again for any given index i, if $i \in \sigma^*$, then using (7) and (8), otherwise $i \in \bar{\sigma}^*$ and using (9) and (10) we have

$$\begin{aligned}
&\frac{(1+\tau(\tilde{d}_x)_i/x_i)(1+\tau(\tilde{d}_y)_i/y_i)}{1-\tau} \\
&\quad \geqslant \frac{(1-\tau C_k^1\mu)(1-\tau(C_k^2\mu^{k+1}+1))}{1-\tau} \\
&\quad \geqslant \frac{\left(1-C_k^1\mu)\right)\left(1-[1-(1+10\sqrt{n})C_k^2\mu^{k+1}][C_k^2\mu^{k+1}+1]\right)}{(1+10\sqrt{n})C_k^2\mu^{k+1}} \\
&\quad \geqslant \frac{\left(1-C_k^1\mu\right)\left(10\sqrt{n}C_k^2\mu^{k+1}+(1+10\sqrt{n})(C_k^2\mu^{k+1})^2\right)}{(1+10\sqrt{n})C_k^2\mu^{k+1}} \\
&\quad \geqslant \frac{\left(1-C_k^1\mu\right)}{1+0.1/\sqrt{n}} \\
&\quad \geqslant \frac{\left(1-C_k^1\mu\right)(1-0.1/\sqrt{n})}{1-0.01/n} \\
&\quad \geqslant 1-0.2/\sqrt{n}.
\end{aligned}$$

The last inequality above follows by using $C_k^1\mu \leqslant 0.01/n$. This completes the proof of (23). We now show that $(x(\tau), y(\tau)) \in \mathcal{N}(\mu(\tau), 1/2)$. First note that for $(x,y) \in \mathcal{N}(\mu, 1/4)$ from (6) we have

$$3/4 \leqslant x_iy_i/\mu \leqslant 5/4. \tag{24}$$

Also we can write

$$\frac{x_i(\tau)y_i(\tau)}{\mu(\tau)} = \left(1+\frac{\nu_i}{\sqrt{n}}\right)\frac{x_iy_i}{\mu},$$

and from (22) we know that $|\nu_i| \leqslant 0.2$. Now by letting $\bar{y}$ represent a vector whose components are $\nu_iy_i/\sqrt{n}$, we have

$$\begin{aligned}
\left\|\frac{1}{\mu(\tau)}X(\tau)y(\tau)-e\right\| = \left\|\frac{1}{\mu}Xy-e+\frac{1}{\mu}X\bar{y}\right\| &\leqslant \left\|\frac{1}{\mu}Xy-e\right\| + \left\|\frac{1}{\mu}X\bar{y}\right\| \\
&\leqslant 0.25 + \left(\max_i\{|\nu_i|\}/\sqrt{n}\right)\left\|\frac{1}{\mu}Xy\right\| \\
&\leqslant 0.25 + 0.2(1.25) = 1/2.
\end{aligned}$$

Here the last inequality above follows by using (24). □

Lemma 11. *Let the hypothesis of Lemma 10 hold. Let* $\tau_k = 1-\rho_k\mu^{k+1}$, *and* $u^{k+1} = x(\tau_k), v^{k+1} = y(\tau_k)$. *Then*

$$u^{k+1^{\mathrm{T}}}v^{k+1}/n \leqslant \rho_k\mu^{k+2}. \tag{25}$$

Furthermore, at (u^{k+1}, v^{k+1}) we have a direction p_u^{k+1}, p_v^{k+1} satisfying

$$\max_{i\in\sigma^*}(|(p_u^{k+1})_i/u_i^{k+1}|) \leqslant \rho_k C^* \mu^{k+2}, \tag{26}$$

$$\max_{i\in\bar\sigma^*}(|(p_v^{k+1})_i/v_i^{k+1}|) \leqslant \rho_k C^* \mu^{k+2}, \tag{27}$$

and $u^{k+1} + p_u^{k+1} = x^*$, $v^{k+1} + p_v^{k+1} = y^*$, is a strictly complementary optimal solution pair.

Proof. The inequality (25) follows immediately from Lemma 9. From Lemma 4 we know that at (u^{k+1}, v^{k+1}) we can construct a direction (p_u^{k+1}, p_v^{k+1}) satisfying

$$\max_{i\in\sigma^*}(|(p_u^{k+1})_i/u_i^{k+1}|) \leqslant C^*({u^{k+1}}^{\mathrm{T}} v^{k+1})/n,$$

$$\max_{i\in\bar\sigma^*}(|(p_v^{k+1})_i/v_i^{k+1}|) \leqslant C^*({u^{k+1}}^{\mathrm{T}} v^{k+1})/n.$$

and $u^{k+1} + p_u^{k+1} = x^*$, $v^{k+1} + p_v^{k+1} = y^*$ are strictly complementary solutions. The proof then follows by using (25). □

The next three propositions are preparatory for Lemma 15.

Proposition 12. *Let $(d_u^{k+1}, d_v^{k+1}, d_w^{k+1})$ be defined by (2) and (p_u^{k+1}, p_v^{k+1}) be defined as in Lemma 11. Let $D_x^k = \mathrm{diag}(d_x^k)$ and follow similar notation for other vectors. Let $(\xi_u^{k+1}, \xi_v^{k+1}, \xi_w^{k+1})$ be the solution of*

$$J(x,y)\begin{pmatrix}\xi_u^{k+1}\\ \xi_v^{k+1}\\ \xi_w^{k+1}\end{pmatrix} = \begin{pmatrix}-(D_x^k p_v^{k+1} + D_y^k p_u^{k+1} + P_u^{k+1} p_v^{k+1})\\ 0\\ 0\end{pmatrix}. \tag{28}$$

Then,

$$d_u^{k+1} + \xi_u^{k+1} = p_u^{k+1} \quad \textit{and} \quad d_v^{k+1} + \xi_v^{k+1} = p_v^{k+1}. \tag{29}$$

Proof. Since $u^{k+1} + p_u^{k+1} = x^*$ and $v^{k+1} + p_v^{k+1} = y^*$, from $(U^{k+1} + P_u^{k+1})(v^{k+1} + p_v^{k+1}) = X^* y^* = 0$ we have

$$U^{k+1} p_v^{k+1} + V^{k+1} p_u^{k+1} = -U^{k+1} v^{k+1} - P_u^{k+1} p_v^{k+1},$$

therefore, from the definition of (u^{k+1}, v^{k+1}) we have

$$X p_v^{k+1} + Y p_u^{k+1} = -(U^{k+1} v^{k+1} + P_u^{k+1} p_v^{k+1} + D_x^k p_v^{k+1} + D_y^k p_u^{k+1}).$$

Hence, (p_u^{k+1}, p_v^{k+1}) is a solution of

$$J(x,y)\begin{pmatrix}p_u^{k+1}\\ p_v^{k+1}\\ p_w^{k+1}\end{pmatrix} = \begin{pmatrix}-(U^{k+1} v^{k+1} + P_u^{k+1} p_v^{k+1} + D_x^k p_v^{k+1} + D_y^k p_u^{k+1})\\ 0\\ 0\end{pmatrix}. \tag{30}$$

Since $J(x,y)$ is an invertible matrix, (2), (28) and (30), have unique solutions. Now, the right hand side of (30) is the sum of the right hand side of (2) (for $k+1$) and the right hand side of (28), therefore (29) follows. □

Proposition 13. *Let* $\Theta = \left(XY^{-1}\right)^{1/2}$, *and use the notation of Proposition 12. Then,*

$$\left\| \begin{matrix} \Theta^{-1}\xi_u^{k+1} \\ \Theta\xi_v^{k+1} \end{matrix} \right\| \leqslant \left\|(XY)^{-1/2}D_x^k p_v^{k+1}\right\| + \left\|(XY)^{-1/2}D_y^k p_u^{k+1}\right\| + \left\|(XY)^{-1/2}P_u^{k+1}p_v^{k+1}\right\|.$$

Proof. From (28) we have $(\Theta^{-1}\xi_u^{k+1})^{\mathrm{T}}(\Theta\xi_v^{k+1}) = (\xi_u^{k+1})^{\mathrm{T}}(\xi_v^{k+1}) = 0$, and

$$\Theta^{-1}\xi_u^{k+1} + \Theta\xi_v^{k+1} = -(XY)^{-1/2}\left(D_x^k p_v^{k+1} + D_y^k p_u^{k+1} + P_u^{k+1}p_v^{k+1}\right).$$

Hence,

$$\left\| \begin{matrix} \Theta^{-1}\xi_u^{k+1} \\ \Theta\xi_v^{k+1} \end{matrix} \right\|^2 = \|\Theta^{-1}\xi_u^{k+1} + \Theta\xi_v^{k+1}\|^2 = \|(XY)^{-1/2}\left(D_x^k p_v^{k+1} + D_y^k p_u^{k+1} + P_u^{k+1}p_v^{k+1}\right)\|^2.$$

The result follows from using the triangular inequality. □

Proposition 14. *Following the notation of Proposition 12, we have*

$$\|(XY)^{-1/2}D_x^k p_v^{k+1}\| \leqslant \max\{C_k^1(C_k^2+2\rho_k), \rho_k C^*(1+C_k^1\mu)(1+(C_k^2+2\rho_k)\mu^{k+1})\}\mu^{k+2}(x^{\mathrm{T}}y)^{1/2}, \tag{31}$$

$$\|(XY)^{-1/2}D_y^k p_u^{k+1}\| \leqslant \max\{C_k^1(C_k^2+2\rho_k), \rho_k C^*(1+C_k^1\mu)(1+(C_k^2+2\rho_k)\mu^{k+1})\}\mu^{k+2}(x^{\mathrm{T}}y)^{1/2}, \tag{32}$$

$$\|(XY)^{-1/2}P_u^{k+1}p_v^{k+1}\| \leqslant \left((1+C_k^1\mu)^{1/2}(C_k^2+2\rho_k)^{1/2}\rho^{3/2}C^*\mu^{k+1}\right)\mu^{k+2}(x^{\mathrm{T}}y)^{1/2}. \tag{33}$$

Proof. We first show (31). (34) below uses $(p_v^{k+1})_i/v_i^{k+1} = -1$, $i \in \sigma^*$ because $y_i^* = 0$, $i \in \sigma^*$. Inequality (35) uses (17), (16), (21), (18) and (27).

$$\|(XY)^{-1/2}D_x^k p_v^{k+1}\| = \left(\sum_{i=1}^{n} \frac{\left((D_x^k p_v^{k+1})_i\right)^2}{x_i y_i}\right)^{1/2}$$

$$= \left(\sum_{i\in\sigma^*}(x_i y_i)\left(\frac{v_i^{k+1}}{y_i}\frac{(d_x^k)_i}{x_i}\right)^2 + \sum_{i\in\bar\sigma^*}(x_i y_i)\left(\frac{v_i^{k+1}}{y_i}\frac{(d_x^k)_i}{x_i}\frac{(p_v^{k+1})_i}{v_i^{k+1}}\right)^2\right)^{1/2} \tag{34}$$

$$\leqslant \left(\sum_{i\in\sigma^*} (x_i y_i) \left([(C_k^2+2\rho_k)\mu^{k+1}][C_k^1\mu]\right)^2 \right.$$

$$\left. + \sum_{i\in\bar\sigma^*} (x_i y_i) \left([1+C_k^1\mu][1+(C_k^2+2\rho_k)\mu^{k+1}][\rho_k C^* \mu^{k+2}]\right)^2 \right)^{1/2} \tag{35}$$

$$\leqslant \max\{C_k^1(C_k^2+2\rho_k), \rho_k C^*(1+C_k^1\mu)(1+(C_k^2+2\rho_k)\mu^{k+1})\}\mu^{k+2}(x^{\mathrm{T}}y)^{1/2}.$$

This completes the proof of (31). The proof for (32) is similar to the proof of (31). Next we prove (33). In the following proof (36) follows from using (20), (17), (21), (18), (26) and (27). (37) uses (25).

$$\|(XY)^{-1/2}p_u^{k+1}p_v^{k+1}\| = \left(\sum_{i=1}^n \frac{((P_u^{k+1}p_v^{k+1})_i)^2}{x_i y_i}\right)^{1/2}$$

$$= \left(\sum_{i=1}^n (u_i^{k+1}v_i^{k+1}) \frac{u_i^{k+1}v_i^{k+1}}{x_i y_i} \left(\frac{(p_u^{k+1})_i(p_v^{k+1})_i}{u_i^{k+1}v_i^{k+1}}\right)^2\right)^{1/2}$$

$$\leqslant \left(\sum_{i=1}^n (u_i^{k+1}v_i^{k+1})[1+C_k^1\mu][(C_k^2+2\rho_k)\mu^{k+1}][(\rho_k C^*\mu^{k+2})^2]\right)^{1/2} \tag{36}$$

$$\leqslant \left((1+C_k^1\mu)(C_k^2+2\rho_k)\mu^{k+1}(\rho_k C^*\mu^{k+2})^2(\rho_k\mu^{k+1}x^{\mathrm{T}}y)\right)^{1/2} \tag{37}$$

$$= \left((1+C_k^1\mu)^{1/2}(C_k^2+2\rho_k)^{1/2}\rho_k^{3/2}C^*\mu^{k+1}\right)\mu^{k+2}(x^{\mathrm{T}}y)^{1/2}.$$

This completes the proof of Proposition 14. □

Lemma 15. *Following the notations of Propositions 12–14, we have*

$$\left\| \begin{matrix} \Theta^{-1}\xi_u^{k+1} \\ \Theta\xi_v^{k+1} \end{matrix} \right\| \leqslant \sqrt{3/4n}\,C_{k+1}^2\mu^{k+2}(x^{\mathrm{T}}y)^{1/2}, \tag{38}$$

where C_{k+1}^2 is defined as in Theorem 7. Furthermore,

$$\frac{|(\xi_u^{k+1})_i|}{x_i} \leqslant C_{k+1}^2\mu^{k+2}, \tag{39}$$

$$\frac{|(\xi_v^{k+1})_i|}{y_i} \leqslant C_{k+1}^2\mu^{k+2}. \tag{40}$$

Proof. The proof of (38) follows directly from Propositions 13, 14 and the definition of C_{k+1}^2. To prove (39) note that

$$\begin{aligned}
\frac{|(\xi_u^{k+1})_i|}{x_i} &= (x_iy_i)^{-1/2}|x_i^{-1/2}y_i^{1/2}(\xi_u^{k+1})_i| \leqslant (x_iy_i)^{-1/2}\left\| \begin{matrix} \Theta^{-1}\xi_u^{k+1} \\ \Theta\xi_v^{k+1} \end{matrix} \right\| \\
&\leqslant \sqrt{3/4n}\left(\frac{x^{\mathrm{T}}y}{x_iy_i}\right)^{1/2} C_{k+1}^2\mu^{k+2} \quad \text{(using (38))} \\
&\leqslant C_{k+1}^2\mu^{k+2} \quad \text{(using (6))}.
\end{aligned}$$

The proof for (40) is similar to the proof of (39). □

Proof of Theorem 7. The first statement of the theorem has been proved in Lemma 10. Now we prove (12). For $i \in \sigma^*$ we have

$$\begin{aligned}
\frac{|(\tilde{d}_x^{k+1})_i|}{x_i} &= \frac{|(d_x^k)_i + (d_u^{k+1})_i|}{x_i} \\
&= \frac{|(d_x^k)_i + (p_u^{k+1})_i - (\xi_u^{k+1})_i|}{x_i} \quad \text{(using (29))} \\
&\leqslant \frac{|(d_x^k)_i| + |(p_u^{k+1})_i| + |(\xi_u^{k+1})_i|}{x_i} \\
&\leqslant C_k^1\mu + \frac{u_i^{k+1}}{x_i}\rho_k C^*\mu^{k+2} + C_{k+1}^2\mu^{k+2} \quad \text{(using (27) and (39))} \\
&\leqslant C_k^1\mu + \beta_k\mu^{k+2} \quad \text{(using (20))}.
\end{aligned}$$

Next we prove (14). For $i \in \bar{\sigma}^*$,

$$\begin{aligned}
\left|1 + \frac{(\tilde{d}_x^{k+1})_i}{x_i}\right| = \left|\frac{x_i + (\tilde{d}_x^{k+1})_i}{x_i}\right| &= \left|\frac{x_i + (d_x^k)_i + (p_u^{k+1})_i - (\xi_u^{k+1})_i}{x_i}\right| \\
&= |(\xi_u^{k+1})_i/x_i| \qquad \text{(because } (d_x^k)_i + (p_u^{k+1})_i = -x_i) \\
&\leqslant C_{k+1}^2\mu^{k+2}.
\end{aligned}$$

The proof of (13) is similar to the proof of (14), and the proof of (15) is similar to the proof of (12). □

4. Asymptotic convergence analysis

In this section we establish the asymptotic convergence results for the generalized predictor–corrector method. It shows that the duality gap converges with Q-order $r+2$. Furthermore, the complementarity violation x_iy_i also converges with the same order.

Theorem 16. *Let (x^l, y^l, z^l) be the solution available after l major iterations of a primal-dual interior point method. Let $(\tilde{d}_x^r, \tilde{d}_y^r, \tilde{d}_z^r)$ be the corrected direction after r corrections in the generalized predictor–corrector method have been performed. Let*

$\mu_l = x^{l^T} y^l / n$, and for simplicity assume that $C = \max\{C^*, C\}$. Then for sufficiently small μ_l we have constants C_r^1, C_r^2 such that $(\tilde{d}_x^r, \tilde{d}_y^r)$ satisfies

$$\frac{|(\tilde{d}_x^r)_i|}{x_i^l} \leqslant C_r^1 \mu_l, \quad i \in \sigma^*, \tag{41}$$

$$-C_r^2 \mu_l^{r+1} \leqslant 1 + \frac{(\tilde{d}_y^r)_i}{y_i^l} \leqslant C_r^2 \mu_l^{r+1}, \quad i \in \sigma^*, \tag{42}$$

$$-C_r^2 \mu_l^{r+1} \leqslant 1 + \frac{(\tilde{d}_x^r)_i}{x_i^l} \leqslant C_r^2 \mu_l^{r+1}, \quad i \in \bar{\sigma}^*, \tag{43}$$

$$\frac{|(\tilde{d}_y^k)_i|}{y_i^l} \leqslant C_r^1 \mu_l, \quad i \in \bar{\sigma}^*. \tag{44}$$

Furthermore,

$$C_r^2 \leqslant (\hat{\beta} C)^r C, \tag{45}$$

$$C_r^1 \leqslant C \left[1 + (\tilde{\beta} C \mu_l) \sum_{i=1}^{r-1} (\hat{\beta} C \mu_l)^i \right] \leqslant 2C, \tag{46}$$

for all values of r. Here $\hat{\beta} = 3\sqrt{12n}(1 + 10\sqrt{n})$ and $\tilde{\beta} = 2(1 + 10\sqrt{n}) + \hat{\beta}$ and we assume that $\tilde{\beta} C \mu_l < 1/2$.

Proof. From Theorem 5 (for $k = 0$) we have

$$-C\mu_l \leqslant 1 + \frac{(\tilde{d}_y^0)_i}{y_i^l} \leqslant C\mu_l, \quad i \in \sigma^*, \tag{47}$$

$$-C\mu_l \leqslant 1 + \frac{(\tilde{d}_x^0)_i}{x_i^l} \leqslant C\mu_l, \quad i \in \bar{\sigma}^*. \tag{48}$$

Now from (1) we have

$$\frac{(\tilde{d}_x^0)_i}{x_i^l} + \frac{(\tilde{d}_y^0)_i}{y_i^l} = -1,$$

therefore, we have

$$\frac{|(\tilde{d}_x^0)_i|}{x_i^l} \leqslant C\mu_l, \quad i \in \sigma^*, \tag{49}$$

$$\frac{|(\tilde{d}_y^0)_i|}{y_i^l} \leqslant C\mu_l, \quad i \in \bar{\sigma}^*. \tag{50}$$

Clearly (41)–(46) hold for $r = 0$. Now assume that they are true for $r = k$. We show that if that is the case, then (45) and (46) is also true for $r = k + 1$. First note that $\tilde{\beta} C \mu_l < 1/2$ and the induction hypothesis implies

$$\hat{\beta} C \mu_l \leqslant 1/2, \qquad C_k^1 \mu_l \leqslant 1, \qquad (3 + 20\sqrt{n}) C_k^2 \mu_l^{k+1} \leqslant 1.$$

each correction requires only $O(n^2)$ arithmetic operations, from Theorem 7 we can state results such as the following.

Corollary 18. *Let $\{(x^l, y^l), (\hat{x}^l, \hat{y}^l)\}$ be generated by the generalized predictor–corrector algorithm with n corrections, each requiring $O(n^2)$ work. Also assume that the efforts in solving a system of linear equations is $O(n^3)$. Then:*
(i) $x^{l^{\mathrm{T}}} y^l \to 0$ *with Q-order* $(n+2)^{1/3}$.
(ii) $X_i^l y_i^l \to 0$, *with Q-order* $(n+2)^{1/3}$.

The following corollary of Theorem 16 shows that once μ has become sufficiently small, it is enough to perform corrector steps to solve the problem, i.e., no further computations of the predictor step is needed.

Corollary 19. *Asymptotically it is sufficient to perform corrections to a predictor direction to solve a linear program.*

Proof. From Theorem 16 once the value of μ_l becomes sufficiently small we have

$$\lim_{r\to\infty} \frac{|(\tilde{d}_x^r)_i|}{x_i^l} \leqslant 2C\mu_l, \quad i \in \sigma^*,$$

$$\lim_{r\to\infty} (\tilde{d}_y^r)_i = -y_i^l, \quad i \in \sigma^*,$$

$$\lim_{r\to\infty} (\tilde{d}_x^r)_i = -x_i^l, \quad i \in \bar{\sigma}^*,$$

$$\lim_{r\to\infty} \frac{|(\tilde{d}_y^k)_i|}{y_i^l} \leqslant 2C\mu_l, \quad i \in \bar{\sigma}^*.$$

Hence, as $r \to \infty$, we can take $\tau_r \to 1$, without violating feasibility. The result then follows from using Lemma 9. □

References

[1] A.J. Goldman and A.W. Tucker, Theory of linear programming, in: H.W. Kuhn and A.W. Tucker, eds., *Linear Inequalities and Related Systems* (Princeton University Press, Princeton, NJ, 1956) 53–97.
[2] T.J. Carpenter, I.J. Lustig, J.M. Mulvey and D.F. Shanno, Higher order predictor–corrector interior point methods with applications to quadratic programming, Technical Report SOR 90-09, Department of Civil Engineering and Operations Research, Princeton University (Princeton, NJ, 1990).
[3] M. Kojima, S. Mizuno and A. Yoshise, A polynomial-time algorithm for a class of linear complementarity problems, *Mathematical Programming* 44 (1989) 1–26.
[4] 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 (3) (1992) 435–449.
[5] 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* 152 (1989).
[6] S. Mehrotra, On finding a vertex solution using interior point method, *Linear Algebra and Its Applications* 152 (1989) 233–253.
[7] S. Mehrotra, On the implementation of a primal–dual interior point method, *SIAM Journal on Optimization* 2 (4) (1992) 575–601.

[8] S. Mehrotra, Generalized predictor–corrector methods and their performance, Technical Report, Dept. of IE/MS, Northwestern University (Evanston, IL, 1991).
[9] S. Mehrotra, Quadratic convergence in a primal-dual method, *Mathematics of Operations Research* 18 (3) (1991) 741–751.
[10] S. Mehrotra and Y. Ye, Finding an interior point in the optimal face of linear programs, *Mathematical Programming* 62 (1993) 497–515.
[11] 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.
[12] R.C. Monteiro and I. Adler, Interior path following primal-dual algorithms. Part I: Linear programming, *Mathematical Programming* 44 (1989) 27–41.
[13] J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables* (Academic Press, New York, 1990).
[14] A. Schrijver, *Theory of Linear and Integer Programming* (Wiley, New York, 1986).
[15] R. Tapia, Y. Zhang, M. Saltzman and A. Weiser, The predictor–corrector interior point method as a composite Newton method, Technical Report 90-06, Department of Mathematical Sciences, Rice University (Houston, TX, 1990).
[16] M.J. Todd and Y. Ye, A centered projective algorithm for linear programming, *Mathematics of Operations Research* 15 (3) (1990) 508–529.
[17] Y. Ye, An $O(n^3L)$ potential reduction algorithm for linear programming, *Mathematical Programming* 50 (2) (1988) 239–258.
[18] Y. Ye, On the finite convergence of interior-point algorithms for linear programming, *Mathematical Programming* 57 (2) (1993) 325–336.
[19] 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 (2) (1993) 151–162.