Some Disadvantages of a Mehrotra-Type Primal-Dual Corrector Interior Point Algorithm for Linear Programming

Coralia Cartis*

January, 2005

Abstract

The Primal-Dual Corrector (PDC) algorithm that we propose computes on each iteration a *corrector* direction in addition to the direction of the standard primal-dual path-following interior point method [8, 22] for Linear Programming (LP), in an attempt to improve performance. The new iterate is chosen by moving along the sum of these directions, from the current iterate. This technique is similar to the construction of Mehrotra's highly popular predictor-corrector algorithm [14]. We present examples, however, that show that the PDC algorithm may fail to converge to a solution of the LP problem, in both exact and finite arithmetic, regardless of the choice of stepsize that is employed. The cause of this bad behaviour is that the correctors exert too much influence on the direction in which the iterates move.

1 Introduction

In the past fifteen years, Interior Point Methods (IPMs) have become highly successful in solving Linear Programming (LP) problems, especially large-scale ones, while enjoying good theoretical convergence and complexity properties (see [3, 5, 18, 22, 20, 21] for comprehensive reviews of the field of IPMs for LP). Examples of IPMs that are reliable both in theory and in practice include the Primal-Dual (PD) path-following method of Kojima et al. [8] with some *long-step* linesearch procedure [22], and an *infeasible* formulation of this algorithm [7, 22]. The majority of commercial and public IPM codes implement a variant of the latter, Mehrotra's Predictor-Corrector (MPC) algorithm [14], and some of them employ in addition, Gondzio's higher-order corrections [4]. For descriptions of the MPC algorithm, see also [10, 24] and Chapter 10 of [22]. Since its first implementations [10, 14] and testing on the standard set of LP test problems (the Netlib test set) [10, 14], the MPC algorithm proved to be, especially on large-scale problems, much faster than the infeasible PD algorithm, in terms of both the number of iterations and the computational time [10]. Its past and present practical successes, however, have not been enhanced by equally praiseworthy theoretical guarantees of good performance: no global convergence or polynomial complexity results are known for this method. A local convergence property [19] has been derived from the view that this method is a perturbed-composite Newton's method, but assumptions are required (non-degeneracy, strict complementarity, etc.) that are much stronger than those of the corresponding analysis for standard IPMs. It is, in fact, acknowledged among practitioners that there are examples on which the MPC algorithm fails to converge (see [17], page 407). To our knowledge, no such examples have been published or analysed in the literature. Moreover, most implementations of the MPC algorithm do not include any safeguards to monitor convergence of the algorithm or to help the algorithm move away from troublesome situations since the generally very good performance of the MPC algorithm seems to render them unnecessary (see [17], page 407). Presently, we construct a Mehrotra-type method, the Primal-Dual Corrector (PDC), whose behaviour we can understand and explain.

The PDC algorithm computes on each iteration, an additional direction, a corrector, to augment the direction of the PD algorithm. In this paper, we find, however, that employing these correctors may have an adverse effect on the performance of the algorithm. In particular, we show that the PDC algorithm may fail to converge to the solution of an LP example in both exact and finite arithmetic. If certain starting points are chosen for the algorithm, and the centring parameters are set to be equal to the same value in $(0, 1)$ or if they are increasing, then we prove that the failure of the algorithm on the example problem occurs in exact arithmetic regardless of the stepsize procedure that is employed (see Section 4.1). We describe

two numerical calculations that exhibit this failure (see Section 4.2). In the first numerical example, the centring parameters are all equal to the same value, and in the second one, they are chosen automatically by the procedure employed in the MPC algorithm [14, 15, 22] and their numerical values are increasing. Though the example that we present does not apply to the MPC algorithm, it throws doubt nevertheless on its convergence properties in general, due to the similarities between the MPC and PDC algorithms and the cause of failure of the latter algorithm on the example (see Subsection 5.2).

The failure of the PDC algorithm to converge is due to the corrector exerting too much influence in the construction of the iterates, and determining the direction in which the iterates move. We attempt to reduce the impact of the correctors by multiplying them by the square of the stepsize in the expression of the new iterates (see Subsection 5.1). The resulting algorithm, the Primal-Dual Second-Order Corrector (PDSOC) [2, 25], has very good convergence, and even complexity, properties for practical choices of the centring parameters and the stepsize [2, 25], which imply that the PDSOC algorithm can overcome the failure that the PDC experienced on the aforementioned example [2]. Further, a substantive interpretation of its construction is given in [2].

The structure of the paper is as follows. Section 2 summarizes some LP duality and interior point method theory that is needed for the remainder of this article. In Section 3, we describe the construction of the PDC algorithm and discuss suitable stepsize procedures for this algorithm. In particular, we show that a standard long-step linesearch procedure, commonly employed in the PD algorithm [22], may not be well-defined for the PDC method. Section 4 presents the above-mentioned example of failure of the PDC algorithm to converge. Subsection 4.1 gives the promised theoretical analysis, and Subsection 4.2, the numerical evidence. Subsection 5.1 describes the PDSOC algorithm briefly, while Subsection 5.2 concludes on the relevance of the failure example to the behaviour of the MPC algorithm.

2   Some LP theory and terminology

Let the LP problem we are solving be given in the standard form

$$\min_{x \in \mathbb{R}^n} \quad c^\top x \quad \text{subject to} \quad Ax = b, \quad x \geq 0, \tag{P}$$

where $m < n$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, and A is a real matrix of dimension $m \times n$. The dual problem corresponding to the primal problem (P) is

$$\max_{(y,s) \in \mathbb{R}^m \times \mathbb{R}^n} \quad b^\top y \quad \text{subject to} \quad A^\top y + s = c, \quad s \geq 0. \tag{D}$$

Let $\mathcal{F}_{PD}$ denote the set of primal-dual feasible points, i.e.,

$$\mathcal{F}_{PD} := \{w = (x, y, s) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n : Ax = b, \quad A^\top y + s = c, \quad x \geq 0, \quad s \geq 0\}, \quad (2.1)$$

and $\mathcal{S}_{PD}$, the primal-dual solution set, containing all triplets $w^* = (x^*, y^*, s^*) \in \mathcal{F}_{PD}$ such that x^* is a solution of (P) and (y^*, s^*), a solution of (D).

A triplet $(x^*, y^*, s^*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ belongs to $\mathcal{S}_{PD}$ if and only if it satisfies the optimality conditions

$$Ax^* = b, \qquad x^* \geq 0, \qquad\qquad\qquad (2.2\text{a})$$
$$A^\top y^* + s^* = c, \quad s^* \geq 0, \qquad\qquad\qquad (2.2\text{b})$$
$$x_i^* s_i^* = 0, \qquad i = 1, 2, \ldots, n. \qquad\qquad (2.2\text{c})$$

Equivalently, a triplet $(x, y, s) \in \mathcal{F}_{PD}$ belongs to $\mathcal{S}_{PD}$ if and only if the duality gap $c^\top x - b^\top y = x^\top s$ is zero. The points in the relative interior of $\mathcal{S}_{PD}$ are called *strictly complementary solutions* of (P) and (D). Such primal-dual solutions $w^\dagger = (x^\dagger, y^\dagger, s^\dagger)$ exist whenever the set $\mathcal{S}_{PD}$ is nonempty, and they are characterized by the property $x^\dagger + s^\dagger > 0$ (see for example, [17]).

We assume that there exists a point $w^0 = (x^0, y^0, s^0)$ that satisfies

$$Ax^0 = b, \quad A^\top y^0 + s^0 = c, \quad x^0 > 0 \quad \text{and} \quad s^0 > 0, \qquad\qquad (2.3)$$

and that the matrix A has full row rank. We refer to these assumptions as the **IPM conditions**. They are equivalent to requiring the sets $\mathcal{F}_{PD}$ and $\mathcal{S}_{PD}$ to be nonempty and bounded, respectively (see for example, Corollary 2.8 in [2]). The first condition implies that $\mathcal{S}_{PD}$ is nonempty.

Any point $w = (x, y, s)$ that satisfies (2.3) is called a *primal-dual strictly feasible point*. These points form the relative interior of the set $\mathcal{F}_{PD}$.

Subject to the IPM conditions, the perturbed system of optimality conditions [22] associated to (P) and (D)

$$F_\mu(w) := \begin{pmatrix} Ax - b \\ A^\top y + s - c \\ XSe - \mu e \end{pmatrix} = 0, \quad x > 0, \quad s > 0, \qquad\qquad (2.4)$$

has a unique solution $w(\mu) = (x(\mu), y(\mu), s(\mu))$, for each $\mu > 0$ [11, 22], where in (2.4), XS is the diagonal matrix with diagonal elements $x_i s_i$, $i = \overline{1, n}$, and $e := (1, 1, \ldots, 1) \in \mathbb{R}^n$. As μ tends to zero, the points $w(\mu)$, $\mu > 0$, which form the *primal-dual central path*, converge to the *analytic centre* of the primal-dual solution set, which is a strictly complementary solution of problems (P) and (D) [23].

3 The Primal-Dual Corrector (PDC) algorithm

3.1 Description of the algorithm

Let problems (P) and (D) satisfy the IPM conditions, and assume that a point $w^0 = (x^0, y^0, s^0)$ satisfying (2.3) is available as a starting point of the algorithm.

The PDC algorithm attempts to follow the primal-dual central path of (P) and (D) approximately to a solution of these problems, in a similar fashion to long-step primal-dual path-following IPMs.

At the current iterate $w^k = (x^k, y^k, s^k)$, $k \geq 0$, of the PDC algorithm, a parameter $\mu > 0$ is picked

$$\mu := \sigma^k \mu^k, \tag{3.1}$$

where $\mu^k := (x^k)^\top s^k / n$, and $\sigma^k \in (0, 1)$ is a *centring parameter* that can be fixed at the start of the algorithm or computed on each iteration by some automatic procedure. Then we compute the Newton direction $dw^k = (dx^k, dy^k, ds^k)$ from w^k for the system $F_\mu(w) = 0$ in (2.4), i.e., dw^k is the solution of the linear system

$$F'_\mu(w^k) \, dw^k = -F_\mu(w^k), \tag{3.2}$$

where $F'_\mu(w^k)$ is the Jacobian of F_μ at w^k. The system (3.2) is equivalent to

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^\top & I \\ S^k & 0 & X^k \end{pmatrix} \begin{pmatrix} dx^k \\ dy^k \\ ds^k \end{pmatrix} = - \begin{pmatrix} Ax^k - b \\ A^\top y^k + s^k - c \\ X^k S^k e - \sigma^k \mu^k e \end{pmatrix}. \tag{3.3}$$

Next, a *corrector* direction $dw^{k,c} = (dx^{k,c}, dy^{k,c}, ds^{k,c})$ is computed by solving the linear system

$$F'_\mu(w^k) dw^{k,c} = -F_\mu(w^k + dw^k). \tag{3.4}$$

The right-hand side of the system (3.4) represents the error that is introduced in the system $F_\mu(w) = 0$ of (2.4) by its linearization around w^k, and it has the explicit expression

$$F_\mu(w^k + dw^k) = \begin{pmatrix} A(x^k + dx^k) - b \\ A^\top(y^k + dy^k) + (s^k + ds^k) - c \\ (X^k + dX^k)(S^k + dS^k)e - \sigma^k \mu^k e \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ dX^k dS^k e \end{pmatrix}, \tag{3.5}$$

where the last equation depends on (3.3), and where dX^k and dS^k are the diagonal matrices with diagonal elements dx_i^k, $i = \overline{1, n}$, and ds_i^k, $i = \overline{1, n}$, respectively. It follows from (3.4) that

the corrector direction attempts to correct this error, in order to position the new iterate closer to the primal-dual central path.

The resulting search direction $dw^{k,r} = (dx^{k,r}, dy^{k,r}, ds^{k,r})$ of the PDC algorithm is the sum

$$dw^{k,r} := dw^k + dw^{k,c}, \tag{3.6}$$

and the new iterate has the form

$$x^{k+1} := x^k + \theta_p^k dx^{k,r}, \quad y^{k+1} := y^k + \theta_d^k dy^{k,r}, \quad \text{and} \quad s^{k+1} := s^k + \theta_d^k ds^{k,r}, \tag{3.7}$$

where $\theta_p^k \in (0, 1]$ and $\theta_d^k \in (0, 1]$ are possibly different primal and dual stepsizes that provide the conditions

$$x^{k+1} > 0 \quad \text{and} \quad s^{k+1} > 0. \tag{3.8}$$

The strict inequalities (3.8), and those in (2.3), together with A having full row rank, imply that the Jacobian $F_\mu'(w^k)$ is nonsingular [22], and thus, the directions dw^k and $dw^{k,c}$ are well-defined, for every $k \geq 0$.

In the context of variants of Newton's method for solving nonlinear systems of equations, the construction of the search direction (3.6) and of the new iterate (3.7) when $\theta_p^k = \theta_d^k = \theta^k$ coincides with the *level-1 composite Newton direction and iterate*, respectively, for the nonlinear system $F_\mu(w) = 0$, starting at w^k, where $\mu := \sigma^k \mu^k$. For a description of this variant of Newton's method applied to general nonlinear systems of equations, see, for example, [19], page 48, and the references therein.

If $dw^{k,c} := 0$, for each $k \geq 0$, the PDC algorithm coincides with the PD algorithm (see pages 8–9 of [22]).

The PDC algorithm applied to problems (P) and (D) can be summarized as follows.

The PDC algorithm:
A point $w^0 = (x^0, y^0, s^0)$ is required that satisfies (2.3). Let $\epsilon > 0$ be a tolerance parameter. At the current iterate $w^k = (x^k, y^k, s^k)$, where $k \geq 0$, do:

Step 1: If $(x^k)^\top s^k \leq \epsilon$, STOP.

Step 2: Let $\mu^k := \frac{(x^k)^\top s^k}{n}$ and choose $\sigma^k \in (0, 1)$.
Compute the direction $dw^k = (dx^k, dy^k, ds^k)$ from the linear system (3.2).
Compute the corrector direction $dw^{k,c} = (dx^{k,c}, dy^{k,c}, ds^{k,c})$ from the system (3.4).
Compute the search direction $dw^{k,r} = (dx^{k,r}, dy^{k,r}, ds^{k,r})$ from (3.6).

Step 3: Choose the stepsizes $\theta_p^k \in (0, 1]$ and $\theta_d^k \in (0, 1]$ along $dx^{k,r}$ and $(dy^{k,r}, ds^{k,r})$, respectively, such that the new iterate $w^{k+1} = (x^{k+1}, y^{k+1}, s^{k+1})$ defined by (3.7) satisfies (3.8).

Step 4: Let $k := k + 1$. Go to Step 1. $\diamond$

It is easy to check that all the iterates w^k, $k \geq 0$, are primal-dual strictly feasible. Thus the only optimality condition that remains to be satisfied (asymptotically) by the iterates is the zero duality gap, i.e., $(x^k)^\top s^k = c^\top x^k - b^\top y^k \to 0$ as $k \to \infty$, which explains the termination criterion in Step 1. We remark that an *infeasible* variant of the PDC algorithm can be developed as for the PD algorithm [22], by abolishing the requirement that w^0 satisfies the primal-dual equality constraints. This, however, is not necessary for our present purpose.

3.2   Stepsize procedures for the PDC algorithm

In this subsection, we specify how to perform Step 3 of the PDC algorithm. We first investigate the possibility of employing a common and practical long-step linesearch procedure [22]. For this purpose, we define

$$w^k(\theta) := w^k + \theta dw^{k,r} \quad \text{and} \quad \mu^k(\theta) := \frac{1}{n}(x^k(\theta))^\top s^k(\theta), \quad \text{for} \quad \theta \geq 0 \quad \text{and} \quad k \geq 0. \quad (3.9)$$

The stepsize technique in question is described for the PD algorithm on pages 84 and 96 of [22], and we refer to it as the $\boldsymbol{\gamma}$ **stepsize procedure**. In the context of the PDC algorithm, this translates into choosing the stepsize $\theta_p^k = \theta_d^k = \theta^k$ to be the largest $\overline{\theta} \in (0,1]$ that is allowed by the inequalities

$$x_i^k(\theta)s_i^k(\theta) \geq \gamma\mu^k(\theta), \quad 0 \leq \theta \leq \overline{\theta}, \quad i = 1, \ldots, n, \quad (3.10)$$

where the positive constant γ is chosen at the start of the algorithm such that the constraints (3.10) are satisfied at the starting point w^0, i.e.,

$$0 < \gamma \leq \frac{1}{\mu^0}\min\left(x_i^0 s_i^0 : \; i = 1, \ldots, n\right). \quad (3.11)$$

Since the iterates satisfy the primal-dual equality constraints, the first $n + m$ equations of the systems (3.3) and (3.4) provide the orthogonality properties

$$(dx^k)^\top ds^k = 0, \; (dx^k)^\top ds^{k,c} = 0, \; (ds^k)^\top dx^{k,c} = 0, \; (dx^{k,c})^\top ds^{k,c} = 0, \; k \geq 0, \quad (3.12)$$

which together with (3.6) imply $(dx^{k,r})^\top ds^{k,r} = 0$. This, (3.12), (3.6) and (3.9) give the recurrence

$$\mu^k(\theta) = [1 - \theta(1 - \sigma^k)]\mu^k, \quad \theta \geq 0, \quad k \geq 0. \quad (3.13)$$

It follows from the $(m+n+i)$th equation of the systems (3.3), (3.4) and (3.5), that computing θ^k according to the above stepsize procedure is equivalent to finding the largest $\overline{\theta} \in (0,1]$ such that

$$\phi_i(\theta) := dx_i^{k,r}ds_i^{k,r}\theta^2 + [-x_i^k s_i^k + \gamma\mu^k + \sigma^k(1-\gamma)\mu^k - dx_i^k ds_i^k]\theta + x_i^k s_i^k - \gamma\mu^k \geq 0, \quad (3.14)$$

7

holds, for all $\theta \in [0, \overline{\theta}]$ and all $i \in \{1, \ldots, n\}$.

Condition (3.11) or the inequalities (3.10) on the previous iteration imply $\phi_i(0) \geq 0$, $i = \overline{1, n}$. We expect $\phi_i(0) = 0$ for some $i \in \{1, \ldots, n\}$ because of the previous choice of stepsize. In this case we require $\phi_i'(0) \geq 0$, so that the steplength of the current iteration can be positive. Thus, for $i \in \{1, \ldots, n\}$, we want the inequality

$$\sigma^k (1 - \gamma) \mu^k - dx_i^k ds_i^k \geq 0, \tag{3.15}$$

whenever i satisfies

$$x_i^k s_i^k = \gamma \mu^k. \tag{3.16}$$

The next example shows, however, that there are primal-dual strictly feasible points w and parameters $\sigma \in (0, 1)$ and $\gamma \in (0, 1)$ with $\gamma \leq \min \left(x_j s_j / \mu : \; j = \overline{1, n} \right)$, where $\mu := x^\top s / n$, such that (3.16) holds for some i, but (3.15) is violated for the same i. Thus if we make such choices of parameters at the start of the PDC algorithm when applied to the LP problem in question, and let $w^0 := w$, then the stepsize procedure (3.10) fails to be well-defined for the algorithm.

Example. Consider the LP problem

$$\min_{x \in \mathbb{R}^3} \; x_1 \quad \text{subject to} \quad x_2 + x_3 = 2, \; x = (x_1, x_2, x_3) \geq 0, \tag{3.17}$$

and its dual

$$\max_{(y,s) \in \mathbb{R} \times \mathbb{R}^3} \; 2y \quad \text{subject to} \quad s_1 = 1, \; y + s_2 = 0, \; y + s_3 = 0, \; s = (s_1, s_2, s_3) \geq 0, \tag{3.18}$$

which trivially satisfy the IPM conditions. Let the parameters satisfy

$$\sigma \in (0, 1) \quad \text{and} \quad 0 < \gamma \leq \frac{3}{20} \sigma, \tag{3.19}$$

and let $w = (x, y, s)$ be

$$x_1 = \frac{8 s_2}{\sigma}, \quad x_2 = \frac{\gamma}{3} \left[\frac{8}{\sigma} + 2 \right], \quad x_3 = 2 - x_2, \quad s_1 = 1, \quad s_2 = s_3 = -y > 0. \tag{3.20}$$

From (3.20), w satisfies the primal-dual equality constraints. Thus we have

$$\mu := \frac{1}{3} (x_1 s_1 + x_2 s_2 + x_3 s_3) = \frac{1}{3} (x_1 + 2 s_2). \tag{3.21}$$

Moreover, the choice of x_2 in (3.20), $\sigma < 1$, and the second inequality in (3.19) imply

$$x_2 < \frac{10 \gamma}{3 \sigma} \leq \frac{1}{2}. \tag{3.22}$$

Thus $x_3 > 0$ and any vector w satisfying (3.19) and (3.20) is a primal-dual strictly feasible point of (3.17) and (3.18). From (3.21) and $x_1 = 8s_2/\sigma$, we have

$$M := \frac{\sigma\mu}{s_2} = \frac{8}{3} + \frac{2\sigma}{3}, \tag{3.23}$$

and it follows from $\sigma \in (0,1)$ that

$$\frac{8}{3} < M < \frac{10}{3}. \tag{3.24}$$

The choice of x_2 in (3.20) provides $x_2 s_2 = \gamma\mu$. We are going to show that

$$x_1 s_1 > x_2 s_2 = \gamma\mu, \quad \text{and} \quad x_3 s_3 > x_2 s_2 = \gamma\mu, \tag{3.25}$$

and that (3.15) does not hold for $i = 2$, i.e.,

$$\sigma(1-\gamma)\mu - dx_2 ds_2 < 0. \tag{3.26}$$

Firstly, from $x_2 < 1$ and the choices of x_1, s_1 and σ, we obtain

$$x_1 s_1 = \frac{8s_2}{\sigma} > 8s_2 > 8x_2 s_2 > x_2 s_2. \tag{3.27}$$

Moreover, from $x_2 < 1$, $x_2 + x_3 = 2$ and $s_3 = s_2$, we have $x_3 s_3 = x_3 s_2 > x_2 s_2$.

To show (3.26), we compute from (3.3) the explicit expression of dw and deduce from the primal-dual equality constraints, that it has the components $dx_2 = \sigma\mu(1-x_2)/s_2$ and $ds_2 = \sigma\mu - s_2$. Thus we obtain

$$\sigma(1-\gamma)\mu - dx_2 ds_2 = \sigma\mu[1 - \gamma - (1-x_2)(M-1)]. \tag{3.28}$$

Substituting $x_2 = \gamma M/\sigma$, it remains to show that $1 - \gamma - (1 - M\gamma/\sigma)(M-1) < 0$, or equivalently,

$$\gamma M^2 - (\gamma + \sigma)M + \sigma(2-\gamma) < 0. \tag{3.29}$$

It follows from $\sigma/\gamma \geq 20/3$ that the expression on the left-hand side of (3.29) is strictly decreasing in M for $M < 10/3$. This and (3.24) imply that it is enough to show that (3.29) holds at $M := 8/3$. Then, (3.29) is equivalent to the inequality $(2/9) \cdot (20\gamma - 3\sigma) - \sigma\gamma < 0$, which is satisfied due to (3.19). $\diamond$

Decreasing the parameter γ on each iteration when the inequalities (3.10) do not provide a positive stepsize generates a linesearch strategy that overcomes the kind of difficulties that have been mentioned. We refer to this technique as the **variable γ stepsize procedure**. In particular, at the start of the algorithm, we choose $\gamma^0 \in (0,1)$ such that the inequalities (3.11) with $\gamma := \gamma^0$ are strictly satisfied. On each iteration, we let $\gamma^k := \gamma^{k-1}$, $k \geq 1$, unless (3.16) holds with $\gamma := \gamma^{k-1}$, for some $i \in \{1, \ldots, n\}$, and (3.15) fails to hold for the same

9

value of i and γ. Then, we choose $\gamma^k \in (0, 1)$ such that $\gamma^k < \gamma^{k-1}$. Next, the stepsize $\theta_p^k = \theta_d^k := \theta^k$ is computed as the largest $\overline{\theta} \in (0, 1]$ such that the inequalities (3.10) with $\gamma := \gamma^k \in (0, 1)$ are satisfied for all $\theta \in [0, \overline{\theta}]$. The sequence γ^k, $k \geq 0$, thus obtained is monotonically decreasing. See pages 136–145 of [22] for a similar stepsize technique for a different primal-dual algorithm.

The following standard stepsize procedure offers another way of avoiding the above-mentioned troubles. In order to ensure condition (3.8) on each iteration, the stepsizes θ_p^k and θ_d^k need to satisfy the inequalities

$$\theta_p^k < \overline{\theta}_p^k \quad \text{and} \quad \theta_d^k < \overline{\theta}_d^k, \quad k \geq 0, \tag{3.30}$$

where $\overline{\theta}_p^k$ and $\overline{\theta}_d^k$ are the steps from x^k and s^k along $dx^{k,r}$ and $ds^{k,r}$ to the boundaries of the primal and dual nonnegative bound constraints, respectively. In other words, we have

$$\overline{\theta}_p^k := 1/\max\left(0, -dx_i^{k,r}/x_i^k, i = \overline{1, n}\right), \quad \text{and} \quad \overline{\theta}_d^k := 1/\max\left(0, -ds_i^{k,r}/s_i^k, i = \overline{1, n}\right). \tag{3.31}$$

The $\boldsymbol{\tau_{pd}}$ **stepsize procedure** [17] chooses $\tau \in (0, 1)$ at the start of the PDC algorithm. Then, on every iteration $k \geq 0$ of the algorithm, at Step 3, a parameter $\tau^k \in [\tau, 1)$ is chosen, and the bounds (3.31) are computed. The stepsizes are given the values

$$\theta_p^k := \min\left(1, \tau^k \overline{\theta}_p^k\right) \quad \text{and} \quad \theta_d^k := \min\left(1, \tau^k \overline{\theta}_d^k\right). \tag{3.32}$$

Positive steps can always be taken from x^k and s^k, since condition (2.3) initially or condition (3.8) on the previous iteration implies $\overline{\theta}_p^k > 0$ and $\overline{\theta}_d^k > 0$. The parameter τ^k in (3.32) may be fixed at the start of the algorithm, or chosen on every iteration, possibly to tend to 1 as $k \to \infty$ [22]. From a numerical point of view, a small value of τ^k is usually inefficient, typical practical values of τ^k being 0.995, 0.9995, or even 0.99995. The stepsize procedure τ_{pd} is common in the practical implementations of various versions of the PD method, including those of the MPC algorithm [22].

The same stepsize in the primal and the dual space can be chosen by letting

$$\theta^k := \min\left(1, \theta_p^k, \theta_d^k\right), \quad k \geq 0, \tag{3.33}$$

where θ_p^k and θ_d^k are defined in (3.32) [22]. We refer to this choice of stepsize as the $\boldsymbol{\tau}$ **procedure**.

4   An example of failure of the PDC algorithm

We find in this section that some interesting features of the PDC algorithm are exposed by the LP problem

$$\min_{x \in \mathbb{R}^3} \quad x_1 + \alpha x_2 \quad \text{subject to} \quad x_2 + x_3 = 2, \quad x = (x_1, x_2, x_3) \geq 0, \qquad (4.1)$$

which depends on a positive parameter α. Its dual problem is

$$\max_{(y,s) \in \mathbb{R} \times \mathbb{R}^3} \quad 2y \quad \text{subject to} \quad s_1 = 1, \ y + s_2 = \alpha, \ y + s_3 = 0, \quad s = (s_1, s_2, s_3) \geq 0. \qquad (4.2)$$

For any $\alpha > 0$, problems (4.1) and (4.2) have the unique solution $w^* = (x^*, y^*, s^*)$, where

$$x^* = (0, 0, 2), \quad y^* = 0, \quad \text{and} \quad s^* = (1, \alpha, 0), \qquad (4.3)$$

and the IPM conditions are satisfied.

We apply the PDC algorithm to problems (4.1) and (4.2). Then every iterate $w^k = (x^k, y^k, s^k)$ satisfies the equations

$$x_2^k + x_3^k = 2, \quad s_1^k = 1, \quad y^k + s_2^k = \alpha, \quad y^k + s_3^k = 0, \quad s_2^k = s_3^k + \alpha, \qquad (4.4)$$

and the inequalities

$$0 < x_1^k, \quad 0 < x_2^k < 2, \quad 0 < x_3^k < 2, \quad s_2^k > \alpha, \quad s_3^k > 0, \quad k \geq 0. \qquad (4.5)$$

The direction $dw^k = (dx^k, dy^k, ds^k)$ defined by (3.3) has the following explicit expression

$$dx_1^k = -x_1^k + \sigma^k \mu^k, \quad dx_2^k = \frac{2\sigma^k \mu^k (1 - x_2^k) - \alpha x_2^k x_3^k}{2 s_2^k - \alpha x_2^k}, \quad dx_3^k = -dx_2^k, \qquad (4.6a)$$

$$ds_1^k = 0, \quad ds_2^k = \frac{\sigma^k \mu^k (2 s_2^k - \alpha) - 2 s_2^k s_3^k}{2 s_2^k - \alpha x_2^k}, \quad ds_3^k = ds_2^k, \quad dy^k = -ds_2^k. \qquad (4.6b)$$

The expression of the corrector direction $dw^{k,c} = (dx^{k,c}, dy^{k,c}, ds^{k,c})$ follows from the systems (3.4) and (3.5) and it is

$$dx_1^{k,c} = 0, \quad dx_2^{k,c} = \frac{-2 ds_2^k}{2 s_2^k - \alpha x_2^k} dx_2^k, \quad dx_3^{k,c} = -dx_2^{k,c}, \qquad (4.7a)$$

$$ds_1^{k,c} = 0, \quad ds_2^{k,c} = \frac{\alpha dx_2^k}{2 s_2^k - \alpha x_2^k} ds_2^k, \quad ds_3^{k,c} = ds_2^{k,c}, \quad dy^{k,c} = -ds_2^{k,c}. \qquad (4.7b)$$

From (4.6) and (4.7), we deduce that the resulting search direction $dw^{k,r} = dw^k + dw^{k,c}$ has the components

$$dx_1^{k,r} = -x_1^k + \sigma^k \mu^k, \quad dx_2^{k,r} = -dx_3^{k,r} = dx_2^k \left(1 - \frac{2 ds_2^k}{2 s_2^k - \alpha x_2^k} \right), \qquad (4.8a)$$

$$ds_1^{k,r} = 0, \quad ds_2^{k,r} = ds_3^{k,r} = ds_2^k \left(1 + \frac{\alpha dx_2^k}{2 s_2^k - \alpha x_2^k} \right). \qquad (4.8b)$$

4.1 Theoretical analysis of the example

Let $\alpha > 0$ in problems (4.1) and (4.2). In this subsection, we investigate the behaviour in exact arithmetic of the PDC algorithm when applied to these problems. Our only assumptions on the choice of stepsizes θ_p^k and θ_d^k are given in Step 3 of the PDC algorithm, so they can be chosen arbitrarily in $(0, 1]$, provided that they also satisfy the inequalities (3.30). The primal and dual stepsizes can be equal to each other, or distinct. We will show in what follows that, if the centring parameter σ^k is set to the same value $\sigma \in (0, 1)$ on each iteration, then there exist starting points w^0 for the PDC algorithm such that the sequence of duality gaps of the generated iterates does not converge to zero, which implies that the iterates do not converge to the solution of problems (4.1) and (4.2) (see our remarks following equation (2.2)).

The first lemma is an intermediate result, in which we identify conditions on the current iterate w^k of the PDC algorithm such that the second and third components of $dx^{k,c}$ and $ds^{k,c}$ are greater in absolute value than their dx^k and ds^k counterparts, yielding a search direction that prevents the progress of w^k, in particular of x^k, towards the optimum.

Lemma 4.1 *Consider problems (4.1) and (4.2), for some $\alpha > 0$. Let $w^k = (x^k, y^k, s^k)$, $k \geq 0$, be the sequence of iterates generated by the PDC algorithm when applied to these problems. If*

$$x_2^k \geq \xi^k := 2 - \frac{1}{2}\sigma^k \quad and \quad s_3^k \leq \nu^k := \frac{1}{8}\alpha\,\sigma^k, \tag{4.9}$$

then

$$dx_2^{k,c} > -dx_2^k > 0 \quad and \quad -ds_2^{k,c} > ds_2^k > 0, \tag{4.10}$$

which imply

$$dx_2^{k,r} = -dx_3^{k,r} > 0 \quad and \quad ds_2^{k,r} = ds_3^{k,r} < 0. \tag{4.11}$$

Proof. Throughout the proof, we drop the iteration superscript k. Firstly, we remark that $\xi \in (1.5, 2)$ and $\nu > 0$, since $\sigma \in (0, 1)$ and $\alpha > 0$. Thus from (4.9), (4.4) and (4.5), we have

$$x_2 = 2 - x_3 \in [\xi, 2) \quad and \quad s_3 = s_2 - \alpha \in (0, \nu]. \tag{4.12}$$

Since $x_2 < 2$ and $s_2 > \alpha$, the denominator $2s_2 - \alpha x_2$ of expressions (4.8a) and (4.8b) is positive. Therefore it is sufficient to establish the relations

$$\frac{2ds_2}{2s_2 - \alpha x_2} > 1 \quad and \quad \frac{\alpha dx_2}{2s_2 - \alpha x_2} < -1. \tag{4.13}$$

Indeed, they imply $ds_2 > 0$ and $dx_2 < 0$. Further, (4.7a) and (4.7b) give $|dx_2^c| > |dx_2|$ and $|ds_2^c| > |ds_2|$ with the sign changes of expression (4.10). The inequalities in (4.11) follow from (3.6) and (4.10), while the equalities are the expressions (4.8).

The mean value μ of the complementarity products can be written

$$\mu = \frac{1}{3}(x_1 s_1 + x_2 s_2 + x_3 s_3) = \frac{1}{3}(x_1 + \alpha x_2 + 2s_3), \qquad (4.14)$$

where (4.4) gives the second equality. We substitute (4.14) and the expression (4.6b) for ds_2 into the first part of (4.13). Then, using the feasibility relation $s_2 - s_3 = \alpha$, we obtain the following equivalent expression for the first inequality in (4.13), in terms of x_2, s_3 and x_1

$$8(3-\sigma)s_3^2 + 4\alpha[9 - \sigma - (3+\sigma)x_2]s_3 + \alpha^2[3x_2^2 - 2(6+\sigma)x_2 + 12] - 2\sigma(\alpha + 2s_3)x_1 < 0. \quad (4.15)$$

It follows from $x_1 > 0$, $s_3 > 0$, $\sigma > 0$ and $\alpha > 0$, that it is sufficient to show

$$8(3-\sigma)s_3^2 + 4\alpha[9 - \sigma - (3+\sigma)x_2]s_3 + \alpha^2[3x_2^2 - 2(6+\sigma)x_2 + 12] < 0, \qquad (4.16)$$

for $x_2 \in [\xi, 2)$ and $s_3 \in (0, \nu]$. The left-hand side of (4.16) is a convex function in s_3, and therefore its supremum occurs at one of the end points of the interval $(0, \nu]$. It remains to verify that (4.16) holds for $s_3 = 0$ and for $s_3 = \nu = \alpha\sigma/8$. At $s_3 = 0$, condition (4.16) becomes

$$3x_2^2 - 2(6+\sigma)x_2 + 12 < 0. \qquad (4.17)$$

In order to check that (4.17) is achieved for any $x_2 \in [\xi, 2)$, it is enough to verify that it holds at $x_2 = \xi$, since the the left-hand side of (4.17) is a decreasing function of x_2. In the case $x_2 = \xi = 2 - \sigma/2$, the left-hand side of (4.17) has the value $-4\sigma + 1.75\sigma^2$, which is negative as required due to $\sigma \in (0, 1)$. For $s_3 = \alpha\sigma/8$, condition (4.16) becomes

$$24x_2^2 - 4(\sigma^2 + 7\sigma + 24)x_2 - \sigma^3 - \sigma^2 + 36\sigma + 96 < 0, \qquad (4.18)$$

whose left-hand side is also decreasing in x_2. At $x_2 = \xi$, the above condition becomes $\sigma^3 + 11\sigma^2 - 20\sigma < 0$, which holds for any $\sigma \in (0, 1)$. Thus the first inequality in (4.13) is achieved.

Similarly, substituting (4.14) and the expression of dx_2 from (4.6a) into the second inequality in (4.13), and employing the feasibility relations $x_3 = 2 - x_2$ and $s_2 = \alpha + s_3$, we deduce the following form of this inequality

$$6s_3^2 + 2\alpha[6 + \sigma - (3+\sigma)x_2]s_3 + \alpha^2[(3-\sigma)x_2^2 - (9-\sigma)x_2 + 6] - \alpha\sigma(x_2 - 1)x_1 < 0. \quad (4.19)$$

Since $x_1 > 0$, $\alpha > 0$, $\sigma \in (0, 1)$, $x_2 \geq \xi > 1$, it is sufficient to establish

$$6s_3^2 + 2\alpha[6 + \sigma - (3+\sigma)x_2]s_3 + \alpha^2[(3-\sigma)x_2^2 - (9-\sigma)x_2 + 6] < 0, \qquad (4.20)$$

for $x_2 \in [\xi, 2)$ and $s_3 \in (0, \nu]$. As before, the left-hand side of (4.20) is convex in s_3. It is thus enough to show that (4.20) holds at $s_3 = 0$ and $s_3 = \nu = \alpha\sigma/8$. At $s_3 = 0$, condition (4.20) becomes

$$3(x_2^2 - 3x_2 + 2) + \sigma x_2(1 - x_2) < 0, \qquad (4.21)$$

13

which holds for any $x_2 \in (1, 2)$ and $\sigma \in (0, 1)$. For $s_3 = \alpha\sigma/8$, condition (4.20) becomes

$$32(3 - \sigma)x_2^2 - 8(\sigma^2 - \sigma + 36)x_2 + 11\sigma^2 + 48\sigma + 192 < 0. \qquad (4.22)$$

The left-hand side of (4.22) is convex in x_2. Substituting $x_2 = 2$ in expression (4.22) yields $-5\sigma^2 - 64\sigma < 0$. At $x_2 = \xi = 2 - \sigma/2$, the left-hand side of (4.22) is $-4\sigma^3 + 79\sigma^2 - 112\sigma$ which is negative for any $\sigma \in (0, 1)$. This proves that the second inequality in (4.13) also holds. $\qquad \square$

The next lemma gives some cases when the PDC algorithm fails to converge to the solution of problems (4.1) and (4.2).

Lemma 4.2 *Let the* PDC *algorithm be applied to problems* (4.1) *and* (4.2)*, for some $\alpha > 0$, and let the sequence of centring parameters $\sigma^k \in (0, 1)$, $k \geq 0$, be monotonically increasing. Let the starting point $w^0 = (x^0, y^0, s^0)$ of the algorithm be any primal-dual strictly feasible point of* (4.1) *and* (4.2) *with*

$$x_2^0 \geq \xi^0 \quad and \quad s_3^0 \leq \nu^0, \qquad (4.23)$$

where $\xi^0 := 2 - \sigma^0/2$ and $\nu^0 := \alpha\,\sigma^0/8$. Then the sequence of duality gaps of the iterates generated by the algorithm is bounded away from zero, and the following bound holds

$$(x^k)^\top s^k > \xi^0 \alpha, \quad k \geq 0. \qquad (4.24)$$

Proof. Let $\xi^k := 2 - \sigma^k/2$ and $\nu^k := \alpha\,\sigma^k/8$, $k \geq 0$, which also occur in (4.9). Since σ^k, $k \geq 0$, is monotonically increasing, we have

$$\xi^k \geq \xi^{k+1} \quad \text{and} \quad \nu^k \leq \nu^{k+1}, \quad k \geq 0. \qquad (4.25)$$

Due to condition (4.23), Lemma 4.1 applies with $k = 0$. Thus relations (4.11) provide the inequalities $dx_2^{0,r} = -dx_3^{0,r} > 0$ and $ds_2^{0,r} = ds_3^{0,r} < 0$. This, the positive stepsizes, (4.23) and (4.25) imply

$$x_2^1 > x_2^0 \geq \xi^0 \geq \xi^1 \quad \text{and} \quad s_3^1 < s_3^0 \leq \nu^0 \leq \nu^1. \qquad (4.26)$$

Thus Lemma 4.1 applies again, this time for $k = 1$, and by the same argument as for $k = 0$, we deduce the analogue of the relations (4.26), where each index is increased by one. Inductively, relations (4.26) hold for $k \geq 0$. They provide the bound $x_2^k \geq \xi^0 > 0$, $k \geq 0$, which implies, together with the feasibility condition $s_2^k > \alpha$, that the complementarity products $x_2^k s_2^k$ are bounded below by the positive constant $\xi^0 \alpha$ for $k \geq 0$. The required result now follows from $(x^k)^\top s^k > x_2^k s_2^k$, $k \geq 0$. $\qquad \square$

The result we promised at the beginning of this subsection is given next. It is the highlight of the subsection.

Corollary 4.3 *Let the* PDC *algorithm be applied to problems* (4.1) *and* (4.2)*, for some* $\alpha > 0$*, and let the centring parameters* σ^k *satisfy*

$$\sigma^k := \sigma \in (0, 1), \quad k \geq 0. \tag{4.27}$$

Let the starting point $w^0 = (x^0, y^0, s^0)$ *of the algorithm be any primal-dual strictly feasible point of* (4.1) *and* (4.2) *with*

$$x_2^0 \geq \xi \quad and \quad s_3^0 \leq \nu, \tag{4.28}$$

where $\xi := 2 - \sigma/2$ *and* $\nu := \alpha\sigma/8$*. Then the sequence of duality gaps of the iterates generated by the algorithm is bounded away from zero, and the following bound holds*

$$(x^k)^\top s^k > \xi\alpha, \quad k \geq 0. \tag{4.29}$$

Proof. Let $\sigma^k := \sigma$, $\xi^k := \xi$ and $\nu^k := \nu$, $k \geq 0$, in Lemma 4.2. □

The next corollary is concerned with some additional properties of the stepsize and the iterates generated by the PDC algorithm when the conditions of the above corollary are satisfied.

Corollary 4.4 *Under the conditions of Corollary 4.3, let* $\theta_p^k = \theta_d^k := \theta^k$*, for* $k \geq 0$*. Then we have*

$$\theta^k \to 0, \quad k \to \infty, \tag{4.30}$$

and

$$w^k \to \tilde{w} \neq w^*, \quad k \to \infty, \tag{4.31}$$

where $\tilde{w}$ *is a primal-dual feasible point of* (4.1) *and* (4.2)*, and* w^* *is defined in* (4.3)*.*

In particular, if the τ *stepsize procedure is employed in the* PDC *algorithm, then at least one of the following two sets of limits holds*

$$(x_2^k, x_3^k) \to (2, 0), \quad k \to \infty, \tag{4.32a}$$
$$s^k = (s_1^k, s_2^k, s_3^k) \to (1, \alpha, 0) = s^*, \quad k \to \infty. \tag{4.32b}$$

Proof. From (3.13), we deduce

$$\frac{\mu^{k+1}}{\mu^k} = 1 - \theta^k(1 - \sigma), \quad k \geq 0, \tag{4.33}$$

which further gives that the sequence μ^k, $k \geq 0$, is strictly decreasing. This, together with (4.29), implies that $\mu^{k+1}/\mu^k \to 1$, as $k \to \infty$. The limit (4.30) now follows from (4.33) and $\sigma \in (0, 1)$.

Lemma 4.1 implies

$$-dx_2^{k,r} = dx_3^{k,r} < 0, \quad \text{and} \quad ds_2^{k,r} = ds_3^{k,r} < 0, \quad k \geq 0. \tag{4.34}$$

It follows from relations (3.7) and $\theta^k > 0$ that the sequences x_3^k, s_2^k and s_3^k, $k \geq 0$, are strictly decreasing, while x_2^k, $k \geq 0$, is strictly increasing. Since, conforming to (4.5), they are also bounded, they are convergent. It follows from $\mu^k := (x_1^k + x_2^k s_2^k + x_3^k s_3^k)/3$, $k \geq 0$, being convergent that x_1^k converges as well. Thus w^k converges, and its limit $\tilde{w}$ is primal-dual feasible since every iterate is. Moreover, Corollary 4.3 implies $\tilde{w} \neq w^*$.

Now let θ^k, $k \geq 0$, be computed by the τ procedure. From (4.4), $\alpha > 0$ and (4.34), we have $-s_2^k/ds_2^{k,r} > -s_3^k/ds_3^{k,r}$, $k \geq 0$. Moreover, if $dx_1^{k,r} := -x_1^k + \sigma^k \mu^k < 0$, then $-x_1^k/dx_1^{k,r} > 1$. Thus the expressions (3.31) and (3.33) become

$$\overline{\theta}_p^k = -\frac{x_3^k}{dx_3^{k,r}}, \quad \overline{\theta}_d^k = -\frac{s_3^k}{ds_3^{k,r}}, \quad \text{and} \quad \theta^k = \min\left(1, \tau^k \overline{\theta}_p^k, \tau^k \overline{\theta}_d^k\right), \quad k \geq 0. \tag{4.35}$$

Since the limit (4.30) holds, θ^k cannot be equal to 1 asymptotically. It follows from (4.35) that there exists a subsequence k_i, $i \geq 0$, such that $\theta^{k_i} = \tau^{k_i} \overline{\theta}_j^{k_i}$, $i \geq 0$, where $j = p$ or $j = d$. If $j = p$, from (3.7), (4.35) and $\tau^k \geq \tau$, we deduce

$$x_3^{k_i+1} = x_3^{k_i} + \theta^{k_i} dx_3^{k_i,r} = (1 - \tau^{k_i})x_3^{k_i} \leq (1 - \tau)x_3^{k_i}, \quad i \geq 0. \tag{4.36}$$

Since x_3^k, $k \geq 0$, is strictly decreasing, it follows from (4.36) and $\tau \in (0,1)$ that $x_3^{k_i+1} \to 0$, as $i \to \infty$, which further gives $x_3^k \to 0$, as $k \to \infty$, since the sequence x_3^k, $k \geq 0$, is convergent. Moreover, from (4.4), we have $x_2^k \to 2$, $k \to \infty$.

Similarly, if $j = d$, we deduce that $s_3^k \to 0$, as $k \to \infty$, and the remaining limits in (4.32b) follow from (4.4). $\qquad \square$

Relations (4.29) and (4.32) imply that the iterates w^k generated by the PDC algorithm with the τ stepsize procedure satisfy the limit $x_3^k s_3^k/\mu^k \to 0$, $k \to \infty$, regardless of how the parameters τ^k and τ are chosen in the stepsize procedure. We remark that existing global convergence results for most long-step primal-dual path-following IPMs (which show that $(x^k)^\top s^k \to 0$, $k \to \infty$) require that the sequences $x_i^k s_i^k/\mu^k$, $k \geq 0$, are bounded away from zero, for $i = \overline{1,n}$.

Corollary 4.3 and the first part of Corollary 4.4 apply also to the case when the variable γ stepsize procedure is employed in the PDC algorithm. This linesearch technique provides the condition that $x_i^k s_i^k/\mu^k$, $k \geq 0$, are bounded away from zero, $i = \overline{1,n}$, provided the parameters γ^k, $k \geq 0$, are chosen to be bounded away from zero. Our results show, however, that for the PDC algorithm, this condition is not enough to ensure global convergence of the algorithm, the role of the corrector directions being decisive.

We remark that subject to the conditions of Corollary 4.3, the PD algorithm with the γ stepsize procedure converges to the solution of the problems (4.1) and (4.2), for any $\gamma \in (0, 1)$ that satisfies (3.11) (see [22] for a general result).

4.2 Numerical calculations

In this subsection, we illustrate the numerical performance of the PDC algorithm with the τ_{pd} stepsize procedure when applied to problems (4.1) and (4.2), for certain values of the parameters. We implemented this algorithm in MATLAB (version 6.0, R12).

Example 1. We set the parameters of the algorithm to the values

$$\sigma^k := 0.1 \quad \text{and} \quad \tau^k := 0.995, \quad \text{for} \quad k \geq 0, \quad \text{and} \quad \epsilon := 10^{-8}, \tag{4.37}$$

and we applied the algorithm to (4.1) and (4.2) with $\alpha := 8$, starting from $w^0 = (x^0, y^0, s^0)$,

$$x^0 := (8, 1.95, 0.05), \quad y^0 := -0.1, \quad s^0 := (1, 8.1, 0.1), \tag{4.38}$$

which is a primal-dual strictly feasible point of these problems.

The conditions of Corollary 4.3 are satisfied in this case, and the lower bound on the duality gap $(x^k)^\top s^k$ of the iterates in relation (4.29) has the value $\xi\alpha = 1.95 \cdot 8 = 15.6$. Since $(x^0)^\top s^0 = 23.8$ and $\epsilon = 10^{-8}$, the Corollary implies that the duality gap of the iterates cannot be decreased to a value lower than ϵ. Our implementation, indeed, does not terminate at Step 1 of the algorithm. Loss of accuracy occurs as k increases, due to growing ill-conditioning in the computation of the search directions (see (4.43) below), leading to the violation of the accuracy tests in our code. Then, the algorithm halts, 6–8 iterations being the usual, depending on the way in which the directions are calculated (*normal equations approach* [22], explicit expressions, etc.).

Tables 1 and 2 report the dual and the primal results computed during the first four iterations of the algorithm, respectively. For $k \geq 4$, the behaviour of the algorithm is similar. All the entries in the tables are rounded to 4 digits times a power of 10, though we do not specify the power when it is 1. The relations (4.4) are satisfied to machine precision. Analysing the data in these tables provides us with significant insight into the behaviour of the algorithm.

Firstly, we inspect the data in Table 1. The values in the third and fourth columns show that the lengths of both ds^k and $ds^{k,c}$ increase very rapidly (see also (4.43)). Moreover, the corrector components $ds_2^{k,c}$ and $ds_3^{k,c}$ are much longer in absolute value than ds_2^k and ds_3^k, respectively, and the relations (4.10) and (4.11) are verified for $k \in \{0, \ldots, 4\}$.

k	$(s_1^k, s_2^k, s_3^k)^\top$	$(ds_1^k, ds_2^k, ds_3^k)^\top$	$(ds_1^{k,c}, ds_2^{k,c}, ds_3^{k,c})^\top$	$\overline{\theta}_d^k$	$2y^k$	$(x^k)^\top s^k$
0	1.0000 8.1000 $1.0000 \cdot 10^{-1}$	0 8.1422 8.1422	0 $-4.1387 \cdot 10^2$ $-4.1387 \cdot 10^2$	$2.4647 \cdot 10^{-4}$	$-2.0000 \cdot 10^{-1}$	$2.3800 \cdot 10$
1	1.0000 8.0005 $5.0000 \cdot 10^{-4}$	0 $2.1305 \cdot 10^3$ $2.1305 \cdot 10^3$	0 $-3.0363 \cdot 10^9$ $-3.0363 \cdot 10^9$	$1.6467 \cdot 10^{-13}$	$-1.0000 \cdot 10^{-3}$	$2.3995 \cdot 10$
2	1.0000 8.0000 $2.5000 \cdot 10^{-6}$	0 $4.2660 \cdot 10^5$ $4.2660 \cdot 10^5$	0 $-2.4265 \cdot 10^{16}$ $-2.4265 \cdot 10^{16}$	$1.0303 \cdot 10^{-22}$	$-5.0000 \cdot 10^{-6}$	$2.3996 \cdot 10$
3	1.0000 8.0000 $1.2500 \cdot 10^{-8}$	0 $8.5321 \cdot 10^7$ $8.5321 \cdot 10^7$	0 $-1.9412 \cdot 10^{23}$ $-1.9412 \cdot 10^{23}$	$6.4392 \cdot 10^{-32}$	$-2.5000 \cdot 10^{-8}$	$2.3996 \cdot 10$
4	1.0000 8.0000 $6.2500 \cdot 10^{-11}$	0 $1.7064 \cdot 10^{10}$ $1.7064 \cdot 10^{10}$	0 $-1.5530 \cdot 10^{30}$ $-1.5530 \cdot 10^{30}$	$4.0245 \cdot 10^{-41}$	$-1.2500 \cdot 10^{-10}$	$2.3996 \cdot 10$

Table 1: The first five dual iterates of the PDC algorithm when applied to problems (4.1) and (4.2).

From (3.31) and the data in Table 1, we deduce

$$\overline{\theta}_d^k = -\frac{s_3^k}{ds_3^{k,r}}, \quad k = \overline{0,4}. \tag{4.39}$$

It follows from (3.7), (3.32) and (4.37) that the relations

$$s_3^{k+1} = s_3^k + \tau^k \overline{\theta}_d^k ds_3^{k,r} = (1 - \tau^k)s_3^k = 5 \cdot 10^{-3} s_3^k, \quad k = \overline{0,4}, \tag{4.40}$$

hold, which together with the dual equality constraints, explain the change to (y^k, s^k) in the second and the sixth columns of Table 1.

Thus after four iterations, the dual iterates and the dual objective function are within $\epsilon = 10^{-8}$ of their optimal values. The numbers in the last column of Table 1 show that the duality gap of the iterates increases, implying that the primal iterate x^4 is far from the optimum. We remark that the increase in the value of the duality gap is due to the fact that we allow different stepsizes to be taken in the primal and dual space (else, conform (3.13), the duality gap of the iterates is strictly decreasing).

Now we address the data in Table 2. As before, this data and the definition (3.31) imply

$$\overline{\theta}_p^k = -\frac{x_3^k}{dx_3^{k,r}}, \quad k = \overline{0,4}, \tag{4.41}$$

which, together with (3.7), (3.32) and (4.37), provides the recurrence

$$x_3^{k+1} = x_3^k + \tau^k \overline{\theta}_d^k dx_3^{k,r} = (1 - \tau^k)x_3^k = 5 \cdot 10^{-3} x_3^k, \quad k = \overline{0,4}. \tag{4.42}$$

k	$(x_1^k, x_2^k, x_3^k)^\top$	$(dx_1^k, dx_2^k, dx_3^k)^\top$	$(dx_1^{k,c}, dx_2^{k,c}, dx_3^{k,c})^\top$	$\overline{\theta}_p^k$	$x_1^k + 8x_2^k$	$(x^k)^\top s^k$
	8.0000	-7.2067	0			
0	1.9500	-3.8122	$1.0347 \cdot 10^2$	$5.0173 \cdot 10^{-4}$	$2.3600 \cdot 10$	$2.3800 \cdot 10$
	$5.0000 \cdot 10^{-2}$	3.8122	$-1.0347 \cdot 10^2$			
	7.9964	-7.1966	0			
1	1.9997	$-5.3443 \cdot 10^2$	$7.5908 \cdot 10^8$	$3.2935 \cdot 10^{-13}$	$2.3994 \cdot 10$	$2.3995 \cdot 10$
	$2.5000 \cdot 10^{-4}$	$5.3443 \cdot 10^2$	$-7.5908 \cdot 10^8$			
	7.9964	-7.1965	0			
2	2.0000	$-1.0665 \cdot 10^5$	$6.0664 \cdot 10^{15}$	$2.0605 \cdot 10^{-22}$	$2.3996 \cdot 10$	$2.3996 \cdot 10$
	$1.25 \cdot 10^{-6}$	$1.0665 \cdot 10^5$	$-6.0664 \cdot 10^{15}$			
	7.9964	-7.1965	0			
3	2.0000	$-2.1330 \cdot 10^7$	$4.8531 \cdot 10^{22}$	$1.2878 \cdot 10^{-31}$	$2.3996 \cdot 10$	$2.3996 \cdot 10$
	$6.2500 \cdot 10^{-9}$	$2.1330 \cdot 10^7$	$-4.8531 \cdot 10^{22}$			
	7.9964	-7.1965	0			
4	2.0000	$-4.2660 \cdot 10^9$	$3.8825 \cdot 10^{29}$	$8.0490 \cdot 10^{-41}$	$2.3996 \cdot 10$	$2.3996 \cdot 10$
	$3.1250 \cdot 10^{-11}$	$4.2660 \cdot 10^9$	$-3.8825 \cdot 10^{29}$			

Table 2: The first five primal iterates of the PDC algorithm when applied to problems (4.1) and (4.2).

It follows from $x_2^k = 2 - x_3^k$, that the pair (x_2^k, x_3^k) approach the point $(2, 0)$ very fast, and x^4 is within ϵ distance to the nonoptimal boundaries determined by the constraints $x_2 = 2$ and $x_3 = 0$. The values in the third and fourth column of Table 2 show that the lengths of dx^k and $dx^{k,c}$ increase very rapidly with k, due to the length of their second and third components (see also (4.43)). Moreover, $dx_2^{k,c}$ and $dx_3^{k,c}$ are much longer in absolute value than, and have opposite signs to, dx_2^k and dx_3^k, confirming that relations (4.10) and (4.11) hold. The direction dx^k 'points towards' the optimum $x^* = (0, 0, 2)$, while the second and third components of the corrector 'point away' from it. Thus these primal components of the resulting search direction $dw^{k,r}$ point away from the optimum.

If $x_2^k \nearrow 2$ and $s_2^k \searrow \alpha = 8$, as k increases, then the denominator $2s_2^k - \alpha x_2^k$ in the expressions (4.6) and (4.7) of dw^k and $dw^{k,c}$ tends to 0, or equivalently, the matrix of the systems (3.2) and (3.4) converges to a singular matrix. In our case, the values of $2s_2^k - \alpha x_2^k$, $k = \overline{0, 4}$, rounded to 4 digits times a power of 10 are

$$2s_2^0 - \alpha x_2^0 = 6.0000 \cdot 10^{-1}, \quad 2s_2^1 - \alpha x_2^1 = 3.0000 \cdot 10^{-3}, \quad 2s_2^2 - \alpha x_2^2 = 1.5000 \cdot 10^{-5},$$
$$2s_2^3 - \alpha x_2^3 = 7.5000 \cdot 10^{-8}, \quad 2s_2^4 - \alpha x_2^4 = 3.7500 \cdot 10^{-10}.$$

$$(4.43)$$

As we already remarked, the increasing ill-conditioning ultimately stops the algorithm. $\diamond$

Corollary 4.3 gives sufficient conditions for the PDC algorithm to fail to converge to the solution of our example, but our numerical evidence and the way we established the quantities

ξ and ν suggest that these conditions are not necessary. For example, the PDC algorithm with the τ_{pd} stepsize procedure also fails to converge numerically if in the previous example we let $x^0 := (8, 1.75, 0.25)$, where $x_2^0 = 1.75 < \xi = 1.95$. Nevertheless if the starting point w^0 of the algorithm is "sufficiently close" to the primal-dual central path of these problems, then we observed numerical convergence of the algorithm to the solution of the problems. We proved a more general result: a *short-step* PDC algorithm converges when applied to problems (P) and (D). For conciseness, this result is omitted.

Next we describe succintly a numerical calculation with a popular choice of σ^k.

Example 2. For $k \geq 0$, we compute the centring parameters $\sigma^k > 0$ in the PDC algorithm by the procedure employed in the MPC algorithm [15, 22]. Thus we let

$$\sigma^k := \left(\frac{(x^k + \overline{\theta}_p^k dx^{k,a})^\top (s^k + \overline{\theta}_d^k ds^{k,a})}{(x^k)^\top s^k} \right)^i, \tag{4.44}$$

where $dw^{k,a} = (dx^{k,a}, dy^{k,a}, ds^{k,a})$ is defined by (3.3) with $\sigma^k := 0$. The stepsizes $\overline{\theta}_p^k$ and $\overline{\theta}_d^k$ are the maximum steps from x^k and s^k, along $dx^{k,a}$ and $ds^{k,a}$, to the primal and dual constraint boundaries, respectively. Thus they are defined by (3.31) with $dw^{k,r} := dw^{k,a}$. The index $i \in \{1, 2, 3, 4\}$ is a constant that we fix at the start of the PDC algorithm. See Chapter 10 of [22] for more explanations about this choice of σ^k.

Let $\alpha := 8$ in (4.1) and (4.2), and let the starting point $w^0 = (x^0, y^0, s^0)$ of the PDC algorithm with the τ_{pd} stepsize procedure be

$$x^0 := (8, 1.99, 0.01), \quad y^0 := -0.1, \quad s^0 := (1, 8.1, 0.1). \tag{4.45}$$

Let $\epsilon := 10^{-8}$ and, for $k \geq 0$, let $\tau^k := 0.995$ and σ^k be computed from (4.44) with $i = 3$. Then the algorithm does not terminate at Step 1, but fails after 5 iterations due to loss of accuracy. The iterates generated by the algorithm are very similar to the ones in Tables 1 and 2. For example, the (x_2^4, x_3^4) components are within 10^{-11} distance to the point $(2, 0)$, and the dual iterates (y^4, s^4) are within ϵ distance to the optimum. The numerical values of σ^k, $k = \overline{0, 6}$, are all of order 10^{-1} and are strictly increasing. The values of the difference $\sigma^{k+1} - \sigma^k$, $k = 1, \ldots, 5$, rounded to 4 digits times a power of 10 are

$$\sigma^1 - \sigma^0 = 3.1762 \cdot 10^{-3}, \qquad \sigma^2 - \sigma^1 = 1.6115 \cdot 10^{-5}, \qquad \sigma^3 - \sigma^2 = 8.0581 \cdot 10^{-8},$$
$$\sigma^4 - \sigma^3 = 4.0290 \cdot 10^{-10}, \qquad \sigma^5 - \sigma^4 = 2.0102 \cdot 10^{-12}, \qquad \sigma^6 - \sigma^5 = 1.0880 \cdot 10^{-14}. \tag{4.46}$$

The condition (4.23) is satisfied numerically. Thus Lemma 4.2 applies, and the value of the lower bound in (4.24) is $\xi^0 \alpha \approx 1.9476 \cdot 8 \approx 15.58$. It follows from $(x^0)^\top s^0 = 24.12$ and (4.24)

that the duality gaps $(x^k)^\top s^k$ cannot be decreased below ϵ. Therefore the algorithm cannot terminate at Step 1.

The behaviour of the algorithm is similar for any $i \in \{1, 2, 4\}$ in (4.44). $\diamond$

Our numerical experience with the PDC algorithm is not restricted to the example problems (4.1) and (4.2). The algorithm does terminate at Step 1 on most LP instances that we tested.

5 Conclusions

5.1 Overcoming the failure: the Primal-Dual Second-Order Corrector (PDSOC) algorithm

The Primal-Dual Corrector (PDC) algorithm that we presented computes on each iteration an additional direction, a corrector, to augment the direction of the standard primal-dual path-following interior-point method for LP problems, in an attempt to improve performance. We found, however, that the new algorithm, the PDC, may fail to converge to the solution of problems (4.1) and (4.2) in both exact and finite arithmetic, regardless of the choice of stepsize that is employed. The cause of the bad performance of the algorithm on these problems is that the corrector direction had too much influence on the resulting search direction. Therefore in the PDSOC algorithm the contribution from the corrector is the quadratic function of the steplength $\theta^k = \theta_p^k = \theta_d^k$

$$w^{k+1} := w^k + \theta^k dw^k + (\theta^k)^2 dw^{k,c}, \tag{5.1}$$

where $w^k = (x^k, y^k, s^k)$ is the current iterate of the PDSOC algorithm applied to problems (P) and (D), and the directions dw^k and $dw^{k,c}$ are computed as before, for $k \geq 0$. The γ linesearch procedure can be shown to be well-defined for the PDSOC algorithm [2], and thus it can be employed to compute the stepsize θ^k.

The quadratic features of the linesearch (5.1) are supported by the interpretation that the new iterate w^{k+1} is chosen along the second-order Taylor approximation around w^k of a local nonlinear path that starts at w^k and ends at the point $w(\sigma^k \mu^k)$ of the primal-dual central path of the problems (see Section 5.1 of [2]).

Some convergence properties of the PDSOC algorithm are given in [25] (see Algorithm 2). It can be shown (see Appendix C of [2]) that these results ensure that, subject to the conditions of Corollary 4.3, the PDSOC algorithm with the γ stepsize procedure converges in exact arithmetic to the solution of problems (4.1) and (4.2), for any $\gamma \in (0, 1)$ that satisfies (3.11).

Additional convergence and complexity properties for the PDSOC algorithm with practical choices of the centring parameters and the stepsize are given in Section 5.2 and 5.3 of [2].

Our preliminary numerical experiments with the PDSOC algorithm are very encouraging.

5.2   The relevance of the example to the MPC algorithm

Relating the construction of the PDC to that of the MPC algorithm, we find that, when applied to problems (P) and (D), the search direction generated in the MPC algorithm is also the sum of dw^k and a corrector direction. The MPC corrector, however, attempts to adjust the error generated in the system of optimality conditions of problems (P) and (D) (i.e., the system $F_\mu(w) = 0$ in (2.4) with $\mu := 0$) by its Newton direction, $dw^{k,a}$, from w^k. Thus the MPC corrector direction is defined by the system (3.4) with $\sigma^k := 0$ and $dw^k := dw^{k,a}$. We remark that the centring parameters in the MPC algorithm are computed as in (4.44).

Concerning the use of the inequalities (3.10) to determine the stepsize for the MPC algorithm, a similar example to the one in Section 3 can be given to show that the γ stepsize procedure may be ill-defined for this algorithm (see Appendix B of [2]).

As we already mentioned in the introductory section, the example of failure of the PDC algorithm to converge that we presented in Section 4 does not apply to the MPC algorithm. We did not show theoretically that the MPC algorithm converges on this example, but found no numerical evidence that would imply otherwise. Our MATLAB implementation of the MPC algorithm with the τ_{pd} stepsize procedure was successful in solving problems (4.1) and (4.2), for various starting points. For example, for the starting point defined in (4.38) or in (4.45), the MPC algorithm similarly generates long correctors that move the primal iterate away from the optimum on early iterations. It "recovers" quickly, however, and converges fast to the solution. The example throws doubt nevertheless, on the convergence properties of the MPC algorithm in general, due to the above-mentioned similarities between the two algorithms and the cause of failure of the PDC on the example. It seems highly unlikely that the occurrence of long corrector directions in the performance of the MPC algorithm would always have a beneficial or harmless effect. A theoretical understanding of the numerical behaviour of the MPC algorithm constitutes potential future work.

Besides its strong connection to the MPC algorithm, we find the PDC algorithm to be interesting in itself, since the examples we presented emphasize the disadvantages of this particular way of constructing corrector directions and new iterates.

References

[1] T. J. Carpenter, I. J. Lustig, J. M. Mulvey, and D. F. Shanno. Higher-order predictor-corrector interior point methods with application to quadratic objectives. *SIAM J. Optim.*, 3(4):696–725, 1993.

[2] C. Cartis. On interior point methods for linear programming, 2004. PhD thesis, Department of Applied Mathematics and Theoretical Physics, University of Cambridge.

[3] R. M. Freund and S. Mizuno. Interior point methods: current status and future directions. In *High performance optimization*, volume 33 of *Appl. Optim.*, pages 441–466. Kluwer Acad. Publ., Dordrecht, 2000.

[4] J. Gondzio. Multiple centrality corrections in a primal-dual method for linear programming. *Comput. Optim. Appl.*, 6(2):137–156, 1996.

[5] C. C. Gonzaga. Path-following methods for linear programming. *SIAM Rev.*, 34(2), 1992.

[6] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4(4):373–395, 1984.

[7] M. Kojima, N. Megiddo, and S. Mizuno. A primal-dual infeasible-interior-point algorithm for linear programming. *Math. Programming*, 61(3, Ser. A):263–280, 1993.

[8] M. Kojima, S. Mizuno, and A. Yoshise. A primal-dual interior point algorithm for linear programming. In *Progress in mathematical programming (Pacific Grove, CA, 1987)*, pages 29–47. Springer, New York, 1989.

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

[10] I. J. Lustig, R. E. Marsten, and D. F. Shanno. On implementing Mehrotra's predictor-corrector interior-point method for linear programming. *SIAM J. Optim.*, 2(3):435–449, 1992.

[11] N. Megiddo. Pathways to the optimal set in linear programming. In *Progress in mathematical programming (Pacific Grove, CA, 1987)*, pages 131–158. Springer, New York, 1989.

[12] S. Mehrotra. Higher order methods and their performance. Technical Report 90–16R1, Dept. of IE/MS, Northwestern University, 1990.

[13] S. Mehrotra. Generalized predictor-corrector methods and their performance. Technical Report 91–17, Dept. of IE/MS, Northwestern University, 1991.

[14] S. Mehrotra. On finding a vertex solution using interior point methods. *Linear Algebra Appl.*, 152:233–253, 1991.

[15] S. Mehrotra. On the implementation of a primal-dual interior point method. *SIAM J. Optim.*, 2(4):575–601, 1992.

[16] S. Mehrotra. Asymptotic convergence in a generalized predictor-corrector method. *Math. Programming*, 74(1, Ser. A):11–28, 1996.

[17] J. Nocedal and S. J. Wright. *Numerical optimization.* Springer-Verlag, New York, 1999.

[18] F. A. Potra and S. J. Wright. Interior-point methods. *J. Comput. Appl. Math.*, 124(1-2):281–302, 2000.

[19] R. Tapia, Y. Zhang, M. Saltzman, and A. Weiser. The Mehrotra predictor-corrector interior-point method as a perturbed composite Newton method. *SIAM J. Optim.*, 6(1):47–56, 1996.

[20] M. H. Wright. Interior methods for constrained optimization. *Acta Numer.*, pages 341–407, 1992.

[21] M. H. Wright. The interior-point revolution in constrained optimization. In *High performance algorithms and software in nonlinear optimization (Ischia, 1997)*, volume 24 of *Appl. Optim.*, pages 359–381. Kluwer Acad. Publ., Dordrecht, 1998.

[22] S. J. Wright. *Primal-dual Interior-Point Methods.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1997.

[23] Y. Ye. *Interior Point Algorithms:Theory and Analysis.* John Wiley and Sons, New York, 1997.

[24] Y. Zhang. User's guide to LIPSOL: linear-programming interior point solvers V0.4. *Optim. Methods Softw.*, 11/12(1-4):385–396, 1999.

[25] Y. Zhang and D. T. Zhang. On polynomiality of the Mehrotra-type predictor-corrector interior-point algorithms. *Math. Programming*, 68(3, Ser. A):303–318, 1995.