CENTER FOR PARALLEL OPTIMIZATION

CONVERGENCE OF INFEASIBLE INTERIOR-POINT ALGORITHMS FROM ARBITRARY STARTING POINTS

by

Stephen C. Billups and Michael C. Ferris

Computer Sciences Technical Report #1180

October 1993

Convergence of Infeasible Interior-Point Algorithms from Arbitrary Starting Points*

Stephen C. Billups and Michael C. Ferris†

October, 1993

Abstract

An important advantage of infeasible interior-point methods compared to feasible interior-point methods is their ability to be warm-started from approximate solutions. It is therefore important for the convergence theory of these algorithms not to depend on being able to alter the starting point. In two recent papers, Yin Zhang and Stephen Wright prove convergence results for some infeasible interior-point methods. Unfortunately, their analysis places a restriction on the starting point. It is easy to meet the restriction by altering the starting point, but this may take the point farther away from the solution, removing the advantage of warm-starting the algorithms. In this paper we extend Zhang and Wright's results to apply to arbitrary strictly positive starting points. We then present an algorithm for solving the Box-Constrained Linear Complementarity problem and prove its convergence.

1 Introduction

Quite often, in using an iterative method to solve a problem, it is possible to use a previously derived approximate solution as a starting point. Such a starting point may be available, for example, as the result of solving a "nearby" problem. When such an approximate solution is used as a starting point, we say that the algorithm has been *warm-started*. Warm-starting is particularly important in the context of sequential quadratic programming. Here, a difficult problem is tackled by solving a sequence of easier subproblems. Typically, the solution of each subproblem provides an excellent starting point for solving the next subproblem in the sequence. It is therefore important to consider how well an algorithm can be warm-started.

Until recently, interior-point methods were not amenable to warm-starting. The difficulty was that such methods required that the starting point be strictly feasible. This was an unfortunate constraint since a solution to a nearby problem would not, in general, be feasible. Thus, it would have to be modified to bring it into the feasible region–a process which would typically carry it farther away from the solution.

*This material is based on research supported by National Science Foundation Grant CCR-9157632 and the Air Force Office of Scientific Research Grant AFOSR-89-0410.

†Computer Sciences Department, University of Wisconsin, Madison, Wisconsin 53706

More recently, infeasible interior-point methods have been developed, which differ from (feasible) interior-point methods by allowing the iterates to violate the equality constraints while strictly satisfying the inequality constraints. Put differently, the iterates are infeasible with respect to the equality constraints, but are interior to the region defined by the inequality constraints. We call this region the *inequality-feasible region* or more simply the *i-feasible region.*

The relaxation of the feasibility requirements allows these infeasible interior-point methods to handle warm starts quite effectively. Typically, the problems can be formulated in such a way that the only difference between the current problem and the nearby problem is in the equality constraints. Thus, with only a slight modification, the solution to the nearby problem can be used as the starting point for the current problem. This slight modification is needed simply to move the point from the boundary of the i-feasible region to the interior of the i-feasible region, and can be arbitrarily small.

A significant amount of work has been devoted to the development of infeasible interior-point algorithms. This line of research first produced practical algorithms along with numerical tests and comparisons, which demonstrated superior practical performance of this class of algorithms. See, for example, [Ans89, Ans91, LMS91, LMS92, MMS89, Meh92, KMT92, Miz93], and [Pot92a, Pot92b].

More recently, a number of theoretical papers have been written which analyze convergence and complexity behavior of various algorithms. See, for example, [KMM91, Wri92a, Wri92b, Zha92], and [ZZ92]. Of particular importance is the paper by Yin Zhang, [Zha92]. In it, Zhang demonstrates global Q-linear convergence and polynomial complexity for a class of infeasible interior-point methods for a generalization of the linear complementarity problem called the horizontal linear complementarity problem. This work is particularly significant because the class of algorithms Zhang analyzes is closely related to already existing algorithms with proven effectiveness. More recently Stephen Wright [Wri92a, Wri92b] extended Zhang's algorithm to produce two algorithms that achieve local Q-quadratic convergence.

Unfortunately, both Zhang and Wright place a restriction on the starting point that will pose problems when warm-starting the algorithms. Their restriction is very easy to satisfy if we are allowed to vary the starting point. However, this completely defeats the purpose of warm-starting, since changing the starting point may take us farther away from the solution. Fortunately, the restriction on the starting points is unnecessary. Demonstrating that fact is the primary purpose of this paper. Indeed, we will show that both Zhang's and Wright's convergence results are valid starting from arbitrary i-feasible starting points.

The extended convergence results also make it possible to extend one of Wright's algorithms to derive an algorithm for the box-constrained linear complementarity problem and to prove global Q-linear and local Q-quadratic convergence for it.

The paper is organized as follows. In Section 2 we describe the horizontal linear complementarity problem and present Zhang's algorithms for solving it. We then extend Zhang's convergence results to apply to arbitrary strictly positive starting points. In Section 3 we state one of Wright's algorithms for solving the linear complementarity problem and extend his convergence results to apply to arbitrary strictly positive starting points. Finally, in Section 4 we discuss our algorithm for the box-constrained linear complementarity problem, and prove convergence results for it.

Some words about notation are in order. Unless otherwise specified, $\|\cdot\|$ denotes the Euclidean norm of a vector. Iteration numbers appear as superscripts on vectors and matrices and as subscripts on scalars. A subscript on a vector (or matrix) represents either a subvector (submatrix) or a component of the vector (matrix).

In expressions, vectors are assumed to be column vectors unless explicitly transposed. Commas are used to separate columns of matrices, semicolons are used to separate rows. For example, if we have the expressions $z = (x, y)$ and $w = (x; y)$, then z is a matrix with columns x and y, whereas w is a column vector formed by concatenating x and y.

We use the notation $(\cdot)_+$, $(\cdot)_-$, and $|\cdot|$ to represent the plus, minus, and absolute value operators, respectively, for vectors. That is, $x_+ := (\max(x_1, 0); \ldots; \max(x_n, 0))$, $x_- := (\max(-x_1, 0); \ldots; \max(-x_n, 0))$ and $|x| := (|x_1|; \ldots; |x_n|)$.

We also refer to the sets $\mathbb{R}_+$ and $\mathbb{R}_{++}$, which represent the nonnegative real numbers and the positive real numbers, respectively.

2 Zhang's Algorithm for the Horizontal Linear Complementarity Problem

The horizontal linear complementarity problem (HLCP) [CPS92] can be stated as follows:

(HLCP) Find $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ such that

$$F(x, y) := \begin{pmatrix} Mx + Ny - h \\ XYe \end{pmatrix} = 0, \quad (x, y) \geq 0,$$

where $M, N \in \mathbb{R}^{n \times n}, e, h \in \mathbb{R}^n, X = \operatorname{diag}(x), Y = \operatorname{diag}(y), e = (1, 1, \cdots, 1)^\top$. Note that if $N = -I$, this is exactly the linear complementarity problem (LCP), which we discuss in Section 3.

In his paper [Zha92], Yin Zhang presents two algorithms for solving the horizontal linear complementarity problem. The first algorithm is a very general algorithm about which a number of useful lemmas can be proved. The second algorithm is a special case of the first for which Zhang proves a global Q-linear convergence result.

For convenience of discussion, Zhang defines the following sets:

$$\begin{aligned}
\mathcal{S} &= \{(x, y) \in \mathbb{R}^{2n} : h = Mx + Ny, (x, y) \geq 0, x^\top y = 0\}, \text{i.e., the solution set,} \\
\mathcal{A} &= \{(x, y) \in \mathbb{R}^{2n} : h = Mx + Ny\}, \\
\mathcal{F} &= \{(x, y) \in \mathcal{A} : (x, y) \geq 0\}, \text{i.e., the set of feasible points,} \\
\mathcal{F}_+ &= \{(x, y) \in \mathcal{A} : (x, y) > 0\}, \text{i.e., the set of strictly feasible points.}
\end{aligned}$$

Zhang's algorithms can be described as centered and damped Newton methods that work as follows: given a starting point $(x^0, y^0) > 0$, both algorithms generate a sequence of strictly positive iterates $\{(x^k, y^k)\}$ that, under appropriate assumptions, converge to a solution (x^*, y^*) of (HLCP).

To prove his results, Zhang makes the following assumptions on the problem:

Assumption 2.1 *For any* $(x, y) \in \mathcal{A}$ *and* $(\hat{x}, \hat{y}) \in \mathcal{A}, (x - \hat{x})^\top (y - \hat{y}) \geq 0$, *i.e.* $\mathcal{A}$ *is the graph of a monotone operator.*

Assumption 2.2 $\mathcal{F} \neq \emptyset$, *i.e., a feasible point exists.*

It is known that Assumptions 2.1–2.2 imply the existence of a solution (x^*, y^*) to (HLCP)(see [Gül92, Theorem 3.1]). It is also well-known that Assumption 2.1 is satisfied by linear programs, convex quadratic programs, and monotone linear complementarity problems.

In addition to these two explicit assumptions, Zhang also makes an implicit assumption about the starting point. Given a point $(u^0, v^0) \in \mathcal{A}$ (such a point exists by Assumption 2.2), Zhang proves his convergence results by choosing a starting point $(x^0, y^0) > 0$ that satisfies $(x^0, y^0) \geq (u^0, v^0)$. It is easy to find such an (x^0, y^0) (simply choose $x^0 = \max(\zeta, u^0)$ and $y^0 = \max(\zeta, v^0)$ for some $\zeta > 0$). However, since we are interested in warm-starting the algorithm, we do not want to change the starting point. Thus, given a *fixed* starting point (x^0, y^0) Zhang's results are based on the following implicit assumption:

Assumption 2.3 *There exists* $(u^0, v^0) \in \mathcal{A}$ *such that* $(x^0, y^0) \geq (u^0, v^0)$.

We shall spend the remainder of this section proving Zhang's results without this implicit assumption. We start with Zhang's first algorithm.

Algorithm 1– Zhang's First Algorithm

Given $(x^0, y^0) > 0$, for $k = 0, 1, 2, \ldots$, do

1. Choose $\sigma_k \in [0, 1)$ and let $\mu_k = \frac{1}{n} {x^k}^\top y^k$. Solve the following linear system for $(\Delta x^k, \Delta y^k)$

$$\begin{pmatrix} M & N \\ Y & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = \begin{pmatrix} h - Mx^k - Ny^k \\ -X^k Y^k + \sigma_k \mu_k e \end{pmatrix}. \tag{1}$$

2. Choose a steplength $\alpha_k \in (0, 1]$ and

$$\alpha_k < \hat{\alpha}_k = \frac{-1}{\min((X^k)^{-1} \Delta x^k, (Y^k)^{-1} \Delta y^k, -1/2)}.$$

   Let $x^{k+1} = x^k + \alpha_k \Delta x^k$ and $y^{k+1} = y^k + \alpha_k \Delta y^k$.

Step 1 of this algorithm is derived by applying one step of Newton's method to the equation

$$F(x, y) = \begin{pmatrix} 0 \\ \sigma_k \mu_k e \end{pmatrix}.$$

The term $\sigma_k \mu_k e$ is used to bias the search direction toward the path of centers, and is thus called the *centering term.* In addition, the Newton step is damped (i.e., multiplied by $\alpha_k \leq 1$) to keep the next iterate positive.

We shall prove a number of technical lemmas about Algorithm 1, which we will then use to prove Q-linear convergence of Zhang's second algorithm, which is a special case of Algorithm 1. We shall also use these results in Section 3 to prove convergence results for one of Wright's algorithms for the linear complementarity problem.

We begin as Zhang does by constructing an auxiliary sequence $\{(u^k, v^k)\}$. Given a pair $(u^0, v^0) \in \mathcal{A}$, for $k = 0, 1, \dots$ we define

$$u^{k+1} := u^k + \alpha_k(\Delta x^k + x^k - u^k), \quad v^{k+1} := v^k + \alpha_k(\Delta y^k + y^k - v^k), \tag{2}$$

where x^k, y^k, Δx^k and Δy^k are defined in Algorithm 1. The sequence is strictly a tool for analysis and is not actually computed. The following lemma summarizes some of the properties of the auxiliary sequence.

Lemma 2.4 *Let $\{(x^k, y^k)\}$ and $\{\alpha_k\}$ be generated by Algorithm 1, $\{(u^k, v^k)\}$ be given by (2) and $\nu_k := \prod_{j=0}^{k-1}(1 - \alpha_j)$. Then for $k \geq 0$*

1. $(u^k, v^k) \in \mathcal{A}$, *i.e.*, $h = Mu^k + Nv^k$;
2. $x^k - u^k = \nu_k(x^0 - u^0)$ *and* $y^k - v^k = \nu_k(y^0 - v^0)$;
3. $|x^k - u^k| = \nu_k|x^0 - u^0| \leq |x^0 - u^0|$ *and* $|y^k - v^k| = \nu_k|y^0 - v^0| \leq |y^0 - v^0|$.

Proof Statements 1 and 2 are proven in [Zha92, Lemma 4.1]. Statement 3 follows immediately from Statement 2 and the fact that $0 \leq \nu_k < 1$. □

Lemma 2.5 *Let $(x^0, y^0) \in \mathbb{R}^n_{++} \times \mathbb{R}^n_{++}$, $(u^0, v^0) \in \mathcal{A}$, $(\hat{x}, \hat{y}) \in \mathcal{F}$, and let $\{(x^k, y^k)\}$ be generated by Algorithm 1. Then, under Assumptions 2.1–2.2, for all $k \geq 0$,*

$$\begin{aligned} &(\hat{x} - (x^0 - u^0)_-)^\top y^k + (\hat{y} - (y^0 - v^0)_-)^\top x^k \\ &\quad \leq \hat{x}^\top \hat{y} + {x^k}^\top y^k + \nu_k \left(|x^0 - u^0|^\top \hat{y} + |y^0 - v^0|^\top \hat{x} + \nu_k |x^0 - u^0|^\top |y^0 - v^0| \right). \end{aligned} \tag{3}$$

Proof Define $\{(u^k, v^k)\}$ according to (2). Then, by Lemma 2.4(1), $(u^k, v^k) \in \mathcal{A}$, so $(\hat{x} - u^k)^\top(\hat{y} - v^k) \geq 0$, by Assumption 2.1.

Using this fact,

$$\begin{aligned} &\hat{x}^\top y^k + \hat{y}^\top x^k + (x^k - u^k)^\top y^k + (y^k - v^k)^\top x^k \\ &\quad \leq \hat{x}^\top y^k + \hat{y}^\top x^k + (x^k - u^k)^\top y^k + (y^k - v^k)^\top x^k + (\hat{x} - u^k)^\top(\hat{y} - v^k) \\ &\quad = \hat{x}^\top \hat{y} + {x^k}^\top y^k + (x^k - u^k)^\top \hat{y} + (y^k - v^k)^\top \hat{x} + (x^k - u^k)^\top(y^k - v^k). \end{aligned} \tag{4}$$

Thus,

(5)

$$
\begin{aligned}
&(\hat{x} - (x^0 - u^0)_-)^\top y^k + (\hat{y} - (y^0 - v^0)_-)^\top x^k \\
&\quad = \hat{x}^\top y^k + \hat{y}^\top x^k - (x^0 - u^0)_-^\top y^k - (y^0 - v^0)_-^\top x^k \\
&\quad \le \hat{x}^\top y^k + \hat{y}^\top x^k - (x^k - u^k)_-^\top y^k - (y^k - v^k)_-^\top x^k \quad \text{(by Lemma 2.4(2))} \\
&\quad \le \hat{x}^\top y^k + \hat{y}^\top x^k + (x^k - u^k)^\top y^k + (y^k - v^k)^\top x^k \\
&\quad \le \hat{x}^\top \hat{y} + {x^k}^\top y^k + (x^k - u^k)^\top \hat{y} + (y^k - v^k)^\top \hat{x} + (x^k - u^k)^\top (y^k - v^k) \quad \text{(by (4))} \\
&\quad \le \hat{x}^\top \hat{y} + {x^k}^\top y^k + |x^k - u^k|^\top \hat{y} + |y^k - v^k|^\top \hat{x} + |x^k - u^k|^\top |y^k - v^k| \quad \text{(since } (\hat{x}, \hat{y}) \ge 0\text{)} \\
&\quad = \hat{x}^\top \hat{y} + {x^k}^\top y^k + \nu_k(|x^0 - u^0|^\top \hat{y} + |y^0 - v^0|^\top \hat{x} + \nu_k |x^0 - u^0|^\top |y^0 - v^0|) \\
&\qquad\qquad \text{(by Lemma 2.4(3))}.
\end{aligned}
$$

□

Lemma 2.6 *Let $\{(x^k, y^k)\}$ be generated by Algorithm 1 in such a way that $\phi_0 \ge {x^k}^\top y^k$ for some $\phi_0 > 0$ and let (x^*, y^*) be a solution to (HLCP). If for some i, $x_i^* > 0$, then the sequence $\{y_i^k\}$ is bounded. Similarly, if $y_i^* > 0$, then $\{x_i^k\}$ is bounded.*

Proof Define $\tilde{x} := x^* - (x^0 - x^*)_-$ and $\tilde{y} := y^* - (y^0 - y^*)_-$. By applying Lemma 2.5 with $(\hat{x}, \hat{y}) = (u^0, v^0) = (x^*, y^*)$, and noting that ${x^*}^\top y^* = 0$, we get

$$
\begin{aligned}
&\tilde{x}^\top y^k + \tilde{y}^\top x^k \\
&\qquad \le {x^k}^\top y^k + \nu_k(|x^0 - x^*|^\top y^* + |y^0 - y^*|^\top x^* + \nu_k |x^0 - x^*|^\top |y^0 - y^*|) \\
&\qquad \le \phi_0 + |x^0 - x^*|^\top y^* + |y^0 - y^*|^\top x^* + |x^0 - x^*||y^0 - y^*| =: C.
\end{aligned}
$$

Thus,

$$
\sum_{i=1}^{n} \left(\tilde{x}_i y_i^k + x_i^k \tilde{y}_i \right) \le C. \tag{6}
$$

Now, $\tilde{y}_i = \min(y_i^0, y_i^*) \ge 0$ and $\tilde{x}_i = \min(x_i^0, x_i^*) \ge 0$. So, each term on the left side of (6) is nonnegative. Therefore, for all i,

$$
x_i^k \tilde{y}_i \le C, \quad \text{and} \quad \tilde{x}_i y_i^k \le C.
$$

Thus, if $\tilde{y}_i > 0$, then $\{x_i^k\}$ is bounded. Similarly, if $\tilde{x}_i > 0$, then $\{y_i^k\}$ is bounded. □

The next lemma is the counterpart in our paper to [Zha92, Lemma 6.1].

Lemma 2.7 *Let $\{(x^k, y^k)\}$ be generated by Algorithm 1 in such a way that $\phi_0 \ge {x^k}^\top y^k \ge \beta \nu_k$ for some $\phi_0, \beta > 0$. For any $(x^*, y^*) \in \mathcal{S}$, let $(u^0, v^0) := (x^*, y^*)$ and generate $\{(u^k, v^k)\}$ according to (2). Then there exists $K > 0$ such that*

$$
\frac{|x^k - u^k|^\top y^k + |y^k - v^k|^\top x^k}{{x^k}^\top y^k} \le K.
$$

Proof Partition the indices $\{1, \dots, n\}$ as follows: $H_1 := \{i : x_i^0 \geq x_i^*, y_i^0 \geq y_i^*\}, H_2 := \{i : x_i^0 < x_i^*\}, H_3 := \{i : y_i^0 < y_i^*\}$. Note that $H_2 \bigcap H_3 = \emptyset$ since one of x_i^* and y_i^* is zero for each i and (x^0, y^0) is strictly positive.

By Lemma 2.4(2), $x_i^k - u_i^k = \nu_k(x_i^0 - x_i^*)$ and $y_i^k - v_i^k = \nu_k(y_i^0 - y_i^*)$, so

$$|x_i^k - u_i^k| = \begin{cases} (x_i^k - u_i^k), & \text{for } i \in H_1 \bigcup H_3, \\ (u_i^k - x_i^k), & \text{for } i \in H_2, \end{cases} \qquad |y_i^k - v_i^k| = \begin{cases} (y_i^k - v_i^k), & \text{for } i \in H_1 \bigcup H_2, \\ (v_i^k - y_i^k), & \text{for } i \in H_3. \end{cases}$$

By Lemma 2.6, $i \in H_2 \Rightarrow x_i^* > 0 \Rightarrow \{y_i^k\}$ is bounded $\Rightarrow \{v_i^k\}$ is bounded. Similarly, for $i \in H_3$, $\{x_i^k\}$ and $\{u_i^k\}$ are bounded. Thus, there exists $K_1 > 0$ such that $|y_i^k + v_i^k| < K_1$ for $i \in H_2$, and $|x_i^k + u_i^k| < K_1$ for $i \in H_3$. Thus,

$$\begin{aligned}
|x^k - u^k|^\top y^k &+ |y^k - v^k|^\top x^k \\
&\leq {x^*}^\top y^k + {y^*}^\top x^k + |x^k - u^k|^\top y^k + |y^k - v^k|^\top x^k + (x^* - u^k)^\top (y^* - v^k) \\
&\qquad \text{(By Assumption 2.1)} \\
&= {x^*}^\top y^* + (x^k - u^k)^\top y^* + (y^k - v^k)^\top x^* + {u^k}^\top v^k + \sum_{i \in H_1} \left\{ (x_i^k - u_i^k) y_i^k + (y_i^k - v_i^k) x_i^k \right\} \\
&\quad + \sum_{i \in H_2} \left\{ (u_i^k - x_i^k) y_i^k + (y_i^k - v_i^k) x_i^k \right\} + \sum_{i \in H_3} \left\{ (x_i^k - u_i^k) y_i^k + (v_i^k - y_i^k) x_i^k \right\} \\
&= (x^k - u^k)^\top y^* + (y^k - v^k)^\top x^* + \sum_{i \in H_1} \left\{ (x_i^k - u_i^k)(y_i^k - v_i^k) + x_i^k y_i^k \right\} \\
&\quad + \sum_{i \in H_2} \left\{ u_i^k y_i^k - v_i^k x_i^k + u_i^k v_i^k - x_i^k y_i^k + x_i^k y_i^k \right\} + \sum_{i \in H_3} \left\{ v_i^k x_i^k - u_i^k y_i^k + u_i^k v_i^k - x_i^k y_i^k + x_i^k y_i^k \right\} \\
&= {x^k}^\top y^k + (x^k - u^k)^\top y^* + (y^k - v^k)^\top x^* + \sum_{i \in H_1} (x_i^k - u_i^k)(y_i^k - v_i^k) \\
&\quad + \sum_{i \in H_2} (u_i^k - x_i^k)(y_i^k + v_i^k) + \sum_{i \in H_3} (x_i^k + u_i^k)(v_i^k - y_i^k) \\
&\leq {x^k}^\top y^k + |x^k - u^k|^\top y^* + |y^k - v^k|^\top x^* + \sum_{i \in H_1} |x_i^k - u_i^k||y_i^k - v_i^k| \\
&\quad + \sum_{i \in H_2} |x_i^k - u_i^k| K_1 + \sum_{i \in H_3} K_1 |y_i^k - v_i^k| \\
&\leq {x^k}^\top y^k + \nu_k \left(|x^0 - u^0|^\top y^* + |y^0 - v^0|^\top x^* + \sum_{i \in H_1} |x_i^0 - u_i^0||y_i^0 - v_i^0| \right. \\
&\quad \left. + \sum_{i \in H_2} |x_i^0 - u_i^0| K_1 + \sum_{i \in H_3} K_1 |y_i^0 - v_i^0| \right) \quad \text{(By Lemma 2.4(3))}.
\end{aligned}$$

So,

$$\frac{|x^k - u^k|^\top y^k + |y^k - v^k|^\top x^k}{{x^k}^\top y^k}$$
$$\leq 1 + \frac{1}{\beta}\Bigg(|x^0 - u^0|^\top y^* + |y^0 - u^0|^\top x^* + \sum_{i \in H_1} |x_i^0 - u_i^0||y_i^0 - v_i^0|$$
$$+ \sum_{i \in H_2} |x_i^0 - u_i^0| K_1 + \sum_{i \in H_3} K_1 |y_i^0 - v_i^0| \Bigg) =: K.$$

□

We are now ready to discuss Algorithm 2. This algorithm is identical to Algorithm 1 except that the steplength α_k is defined more precisely. We use the following merit function:

$$\phi(x, y) := x^\top y + \|Mx + Ny - h\|.$$

The first term in this merit function measures the complementarity gap and the second term measures the infeasibility. Clearly, a point (x^*, y^*) is a solution to (HLCP) if and only if $(x^*, y^*) \geq 0$ and $\phi(x^*, y^*) = 0$. For convenience we make several additional definitions:

$$x(\alpha) := x^k + \alpha \Delta x^k, \quad y(\alpha) := y^k + \alpha \Delta y^k,$$
$$\phi(\alpha) := \phi(x(\alpha), y(\alpha)), \quad \phi_k := \phi(x^k, y^k).$$

The steplength α_k is chosen so as to minimize the merit function $\phi(\alpha)$ subject to the following constraints:

(7a) $$\alpha \in [0, 1],$$

(7b) $$x(\alpha) > 0, \quad y(\alpha) > 0,$$

(7c) $$x(\alpha)^\top y(\alpha) \geq (1 - \alpha)\nu_k {x^0}^\top y^0,$$

(7d) $$x(\alpha)_i y(\alpha)_i \geq (\gamma/n) x(\alpha)^\top y(\alpha), \quad i = 1, \ldots, n.$$

where $\gamma \in (0, 1)$ is chosen so that $\gamma \leq \min(X^0 Y^0 e)/({x^0}^\top y^0/n)$.

Condition (7d) is used to prevent the iterates from prematurely getting too close to the boundary of the positive orthant. Condition (7c) gives a priority to feasibility over complementarity, and implies that

(8) $$\frac{{x^k}^\top y^k}{{x^0}^\top y^0} \geq \nu_k.$$

The algorithm can now be stated as Algorithm 2.

Note that Algorithm 2 is a special case of Algorithm 1, so all the lemmas proved for Algorithm 1 also apply for Algorithm 2. In particular, since ϕ_k is a decreasing sequence, it follows that ${x^k}^\top y^k \leq \phi_0$, for all k. We now show that Algorithm 2 has global Q-linear convergence.

Algorithm 2– Zhang's Second Algorithm

Given $(x^0, y^0) > 0$, for $k = 0, 1, 2, \ldots$, do

1. Choose $\sigma_k \in [0, 1)$ and let $\mu_k = \frac{1}{n} {x^k}^\top y^k$. Solve the linear system (1) for $(\Delta x^k, \Delta y^k)$
2. Set the steplength α_k by minimizing $\phi(\alpha)$ subject to the constraints (7).
   Let $x^{k+1} = x^k + \alpha_k \Delta x^k$, and $y^{k+1} = y^k + \alpha_k \Delta y^k$.

Let $(u^0, v^0) \in \mathcal{A}$ and let $\{(u^k, v^k)\}$ be defined by (2). Define

(9) $$\xi_k := \left(\frac{n}{\gamma}\right)^{1/2} \frac{|x^k - u^k|^\top y^k + |y^k - v^k|^\top x^k}{{x^k}^\top y^k},$$

(10) $$\eta_k := 1 - 2\sigma_k + \frac{{\sigma_k}^2}{\gamma} + \frac{2\nu_k |x^0 - u^0|^\top |y^0 - v^0|}{{x^0}^\top y^0},$$

(11) $$\omega_k := \left(\xi_k + \sqrt{{\xi_k}^2 + \eta_k}\right)^2,$$

where x^k, y^k arise from Algorithm 2. Note that if $(x^k, y^k) \geq (u^k, v^k)$ (Assumption 2.3), then these quantities are identical to the ones defined by Zhang. However, we now prove Zhang's main convergence theorem without Assumption 2.3.

Theorem 2.8 *Let $\{\phi^k\}$ be generated by Algorithm 2 with σ^k satisfying $0 < \sigma \leq \sigma^k \leq 1/2$. Then $\{\phi^k\}$ converges to zero at a global Q-linear rate, i.e., there exists $\delta \in (0, 1)$ such that*

$$\phi^{k+1} \leq (1 - \delta)\phi^k, \quad k = 0, 1, 2, \cdots.$$

The proof is identical to Zhang's proof except that we use the following lemma in place of [Zha92, Lemma 6.2]. Our proof follows the spirit of the proof of [Zha92, Lemma 6.2] closely.

Lemma 2.9 *Let $\{(x^k, y^k)\}$ and $\{(\Delta x^k, \Delta y^k)\}$ be generated by Algorithm 2 and let $D^k := (Y^k)^{1/2}(X^k)^{-1/2}$. Then*

$$\left\|D^k \Delta x^k\right\|^2 + \left\|(D^k)^{-1} \Delta y^k\right\|^2 \leq \omega_k {x^k}^\top y^k.$$

Moreover, the sequence $\{\omega_k\}$ is bounded, i.e. there is a constant $\omega > 0$ such that $\omega_k \leq \omega$, for all k.

Proof Define

$$t_k := \left(\left\|D^k \Delta x^k\right\|^2 + \left\|(D^k)^{-1} \Delta y^k\right\|^2\right)^{1/2}.$$

Notice that $\left\|D^k \Delta x^k\right\| \leq t_k$ and $\left\|(D^k)^{-1} \Delta y^k\right\| \leq t_k$.

By (1), $(x^k + \Delta x^k, y^k + \Delta y^k) \in \mathcal{A}$. Thus, by Lemma 2.4(1) $(\Delta x^k + x^k - u^k)^\top(\Delta y^k + y^k - v^k) \geq 0$. So,

$$
\begin{aligned}
(12) \qquad & {\Delta x^k}^\top \Delta y^k \\
&= [(\Delta x^k + x^k - u^k) - (x^k - u^k)]^\top [(\Delta y^k + y^k - v^k) - (y^k - v^k)] \\
&= (\Delta x^k + x^k - u^k)^\top (\Delta y^k + y^k - v^k) - (\Delta x^k + x^k - u^k)^\top (y^k - v^k) \\
&\quad - (x^k - u^k)^\top (\Delta y^k + y^k - v^k) + (x^k - u^k)^\top (y^k - v^k) \\
&\geq -{\Delta x^k}^\top (y^k - v^k) - (x^k - u^k)^\top \Delta y^k - (x^k - u^k)^\top (y^k - v^k) \\
&= -[(D^k)^{-1}(y^k - v^k)]^\top [D^k \Delta x^k] - [D^k (x^k - u^k)]^\top [(D^k)^{-1} \Delta y^k] \\
&\quad - (x^k - u^k)^\top (y^k - v^k) \\
&\geq -e^\top (D^k)^{-1} |y^k - v^k| t_k - e^\top D^k |x^k - u^k| t_k - (x^k - u^k)^\top (y^k - v^k).
\end{aligned}
$$

Consider now the two terms involving D^k in the above inequality. By the construction of $\{(x^k, y^k)\}$ (see (7d)),

$$
(13) \qquad \frac{{x^k}^\top y^k}{x_i^k y_i^k} \leq \frac{n}{\gamma}.
$$

This leads to

$$
({x^k}^\top y^k)^{1/2} D^k = [{x^k}^\top y^k (X^k Y^k)^{-1}]^{1/2} Y^k \leq \left(\frac{n}{\gamma}\right)^{1/2} Y^k,
$$

and

$$
\begin{aligned}
(14) \qquad e^\top D^k |x^k - u^k| &= \frac{({x^k}^\top y^k)^{1/2} e^\top [({x^k}^\top y^k)^{1/2} D^k] |x^k - u^k|}{{x^k}^\top y^k} \\
&\leq \left(\frac{n{x^k}^\top y^k}{\gamma}\right)^{1/2} \frac{e^\top Y^k |x^k - u^k|}{{x^k}^\top y^k} \\
&= \left(\frac{n{x^k}^\top y^k}{\gamma}\right)^{1/2} \frac{|x^k - u^k|^\top y^k}{{x^k}^\top y^k}.
\end{aligned}
$$

Similarly,

$$
e^\top (D^k)^{-1} |y^k - v^k| \leq \left(\frac{n{x^k}^\top y^k}{\gamma}\right)^{1/2} \frac{|y^k - v^k|^\top x^k}{{x^k}^\top y^k}.
$$

Thus,

$$
{\Delta x^k}^\top \Delta y^k \geq -\left(\frac{n{x^k}^\top y^k}{\gamma}\right)^{1/2} \frac{|x^k - u^k|^\top y^k + |y^k - v^k|^\top x^k}{{x^k}^\top y^k} t_k - (x^k - u^k)^\top (y^k - v^k),
$$

which by (9) is equivalent to

$$
(15) \qquad {\Delta x^k}^\top \Delta y^k \geq -({x^k}^\top y^k)^{1/2} \xi_k t_k - (x^k - u^k)^\top (y^k - v^k).
$$

Also from Lemma 2.4(2) and (8),

$$(16) \qquad (x^k - u^k)^\top (y^k - v^k) = (\nu_k)^2 (x^0 - u^0)^\top (y^0 - v^0)$$
$$\leq (\nu_k)^2 |x^0 - u^0|^\top |y^0 - v^0|$$
$$\leq \frac{\nu_k {x^k}^\top y^k |x^0 - u^0|^\top |y^0 - v^0|}{{x^0}^\top y^0}.$$

So,

$$(17) \qquad {x^k}^\top y^k [1 - 2\sigma_k + \sigma_k^2/\gamma]$$
$$= \sum_{i=1}^n \left(x_i^k y_i^k - 2\sigma_k \mu_k + \frac{\sigma_k^2 \mu_k}{\gamma} \right)$$
$$\geq \sum_{i=1}^n \left(x_i^k y_i^k - 2\sigma_k \mu_k + \frac{\sigma_k^2 \mu_k^2}{x_i y_i} \right) \quad \text{(by (13))}$$
$$= \left\| (X^k Y^k)^{-1/2} (\sigma_k \mu_k e - X^k Y^k e) \right\|^2$$
$$= \left\| (X^k Y^k)^{-1/2} (Y^k \Delta x^k + X^k \Delta y^k) \right\|^2 \quad \text{(by (1))}$$
$$= (t_k)^2 + 2\Delta {x^k}^\top \Delta y^k$$
$$\geq (t_k)^2 - 2({x^k}^\top y^k)^{1/2} \xi_k t_k - 2(x^k - u^k)^\top (y^k - v^k) \quad \text{(by (15))}$$
$$\geq (t_k)^2 - 2({x^k}^\top y^k)^{1/2} \xi_k t_k - 2\frac{\nu_k {x^k}^\top y^k |x^0 - u^0|^\top |y^0 - v^0|}{{x^0}^\top y^0} \quad \text{(by (16))}.$$

Thus, from (10),

$${t_k}^2 - 2({x^k}^\top y^k)^{1/2} \xi_k t_k - {x^k}^\top y^k \eta_k \leq 0.$$

The quadratic $t^2 - 2({x^k}^\top y^k)^{1/2} \xi_k t - {x^k}^\top y^k \eta_k$ is convex and has a unique positive root at

$$t = \left(\xi_k + \sqrt{{\xi_k}^2 + \eta_k} \right) ({x^k}^\top y^k)^{1/2}.$$

This implies that

$${t_k}^2 \leq \left(\xi_k + \sqrt{{\xi_k}^2 + \eta_k} \right)^2 {x^k}^\top y^k = \omega_k {x^k}^\top y^k.$$

Clearly, $\{\eta_k\}$ is bounded. Moreover, from Lemma 2.7 and (8), $\{\xi_k\}$ is also bounded. Hence, $\{\omega_k\}$ is bounded. □

We now turn our attention to a special case of (HLCP) and extend some global and local convergence results due to Wright using the technical lemmas proven above.

3 Wright's Algorithm for the Linear Complementarity Problem

The linear complementarity problem (LCP) [CPS92] is defined as follows:

$$\text{(LCP)} \qquad \text{Find } (x,y) \in \mathbb{R}^n \times \mathbb{R}^n \text{ such that}$$
$$y = Mx - h, \quad (x,y) \geq (0,0), \quad x^\top y = 0.$$

Note that this is a special case of (HLCP) with $N = -I$.

In a recent paper [Wri92a], Stephen Wright presents a locally Q-quadratic algorithm for solving (LCP) which is based on Zhang's algorithm for (HLCP). In fact, Wright's algorithm is a special case of Zhang's Algorithm 1. Unfortunately, Wright's convergence results, like Zhang's, suffer from the same restriction on the starting point (x^0, y^0) (see Assumption 2.3). In this section, we will remove this assumption so that Wright's results will apply to arbitrary strictly positive starting points.

The calculation of the search direction in Wright's algorithm is exactly the same as in Zhang's algorithms. By substituting $N = -I$ into (1), we get the following equation for calculating the search direction $(\Delta x^k, \Delta y^k)$.

$$\begin{pmatrix} M & -I \\ Y^k & X^k \end{pmatrix} \begin{pmatrix} \Delta x^k \\ \Delta y^k \end{pmatrix} = \begin{pmatrix} h - Mx^k + y^k \\ -X^k Y^k e + \sigma_k \mu_k e \end{pmatrix}. \tag{18}$$

The same substitution into the merit function gives us $\phi(x,y) := x^\top y + \|y - Mx + h\|$. The definitions of $\mu_k, X^k, Y^k, x(\alpha)$, and $y(\alpha)$ are unchanged.

The difference in Wright's algorithm is that the constraints placed on the steplength are relaxed in order to achieve local Q-quadratic convergence. Given the parameters $\gamma_k \in (0,1)$, and $\beta_k \in [0,1)$, the steplength α_k is calculated by minimizing the function $\phi(\alpha) := \phi(x(\alpha), y(\alpha))$ subject to the following constraints:

$$\alpha \in [0,1], \tag{19a}$$
$$x(\alpha) > 0, \quad y(\alpha) > 0, \tag{19b}$$
$$x(\alpha)^\top y(\alpha) \geq (1-\beta_k)(1-\alpha)\nu_k {x^0}^\top y^0, \tag{19c}$$
$$x(\alpha)_i y(\alpha)_i \geq (\gamma_k/n) x(\alpha)^\top y(\alpha), \quad i = 1, \ldots, n. \tag{19d}$$

The condition (19c) is a relaxation of the condition (7c) enforced by Zhang's Algorithm 2. Setting $\beta_k > 0$ allows the reduction in the complementarity gap to exceed the reduction in the feasibility, thereby allowing larger steps. Note that by setting $\beta_k = 0$ we get Zhang's algorithm.

Another notable difference is that Wright's algorithm can use a different γ_k at each iteration in condition (19d). In fact, the local Q-quadratic convergence is dependent on being able to choose successively smaller choices of γ_k at each iteration.

The complete algorithm is given in Algorithm 3.

Algorithm 3– Wright's Algorithm

Given $\gamma \in (0,1/2), \sigma \in (0,1/2), \rho \in (0,\gamma)$,
$\phi > 0$, and $(x^0, y^0) > (0,0)$, with $x_i^0 y_i^0 \geq 2\gamma\mu_0$;

$t_0 \leftarrow 1, \gamma_0 \leftarrow 2\gamma$;

for $k = 0, 1, 2, \ldots$
  if $\phi_k := \phi(x^k, y^k) \leq \phi$
  then Compute a "fast" step by setting $\sigma_k \leftarrow \mu_k, \beta_k \leftarrow \gamma^{t_k}$,
    and $\gamma_k \leftarrow \gamma(1+\gamma^{t_k})$ and solving (18)–(19)
    to calculate $(\Delta x^k, \Delta y^k)$ and α_k;
    if $\phi(x^k + \alpha_k \Delta x^k, y^k + \alpha_k \Delta y^k) \leq \rho\phi_k$
    then $(x^{k+1}, y^{k+1}) \leftarrow (x^k, y^k) + \alpha_k(\Delta x^k, \Delta y^k)$
      $t_{k+1} \leftarrow t_k + 1$;
      go to next k;
    end if
  end if
  Compute a "safe" step by setting $\sigma_k \in [\sigma, 1/2], \beta_k = 0$,
  and $\gamma_k = \gamma_{k-1}$, and solving (18)–(19)
  to calculate $(\Delta x^k, \Delta y^k)$ and α_k;
  $(x^{k+1}, y^{k+1}) \leftarrow (x^k, y^k) + \alpha_k(\Delta x^k, \Delta y^k)$
  $t_{k+1} \leftarrow t_k$;
  go to next k;
end for.

At each iteration, either a *safe step* or a *fast step* is taken. A safe step works exactly like Zhang's algorithm; we set $\beta_k = 0$ and hold γ_k constant for the next iteration. A fast step works by setting $\beta_k > 0$ and $\sigma_k = \mu_k$. It is these fast steps that allow the algorithm to attain local Q-quadratic convergence. Unfortunately, a fast step requires reducing the size of γ_k for subsequent iterations. Therefore, the fast step is only taken if it results in a significant decrease in ϕ. If it doesn't, the step is discarded and a "safe step" is taken instead.

We again will find it convenient to refer to the sets $\mathcal{S}, \mathcal{A}, \mathcal{F}$, and $\mathcal{F}_+$ defined earlier. For convenience we restate their definitions here with $N = -I$:

$$\mathcal{S} = \{(x,y) \in \mathbb{R}^{2n} : y = Mx - h, (x,y) \geq 0, x^\top y = 0\}, \text{i.e., the solution set,}$$
$$\mathcal{A} = \{(x,y) \in \mathbb{R}^{2n} : y = Mx - h\},$$
$$\mathcal{F} = \{(x,y) \in \mathcal{A} : (x,y) \geq 0\}, \text{i.e., the set of feasible points,}$$
$$\mathcal{F}_+ = \{(x,y) \in \mathcal{A} : (x,y) > 0\}, \text{i.e., the set of strictly feasible points.}$$

Wright proves two convergence results for his algorithm. First, he shows that the algorithm has global Q-linear convergence. Second, he shows that the algorithm attains local

Q-quadratic convergence. His results are based on the following explicit assumptions:

Assumption 3.1 *M is positive semidefinite.*

Assumption 3.2 *(LCP) has a strictly feasible point $(\bar{x}, \bar{y})$, that is $\mathcal{F}_+ \neq \emptyset$.*

Assumption 3.3 *The solution set for (LCP) is nonempty and, moreover, there is a strictly complementary solution (x^*, y^*).*

Wright's results, like Zhang's, are also dependent on Assumption 2.3. These assumptions are more restrictive than Zhang's assumptions. Assumption 3.1 is equivalent to Assumption 2.1 in the case of LCP, but Assumption 3.2 is stronger than Assumption 2.2. In fact, Zhang's assumptions are sufficient to prove the global Q-linear convergence. However, Wright's more restrictive assumptions are used to prove the local Q-quadratic convergence. We now proceed to prove global Q-linear convergence of Wright's algorithm using only Assumptions 2.1–2.2.

Note that since Wright's algorithm is a special case of Algorithm 1, Lemmas 2.4–2.7 are applicable for it. We shall also need the following result from [Wri92a]:

Lemma 3.4 *Let $\hat{\beta} := \prod_{k=0}^{\infty}(1-\beta_k)$ where β_k is defined in Algorithm 3, and let $\mu_k := {x^k}^\top y^k/n$. Then $\hat{\beta} > 0$ and*

$$\mu_k \geq \hat{\beta}\nu_k\mu_0, \quad \text{and}$$
$${x^k}^\top y^k \geq \hat{\beta}\nu_k {x^0}^\top y^0.$$

Proof [Wri92a], Lemmas 3.1 and 3.2. □

We now define the quantities $\hat{\xi}_k, \hat{\eta}_k$, and $\hat{\omega}_k$, which we use to establish the convergence rates:

$$\hat{\xi}_k := \left(\frac{n}{\gamma_k}\right)^{1/2} \frac{|x^k - u^k|^\top y^k + |y^k - v^k|^\top x^k}{{x^k}^\top y^k}, \tag{20}$$

$$\hat{\eta}_k := 1 - 2\sigma_k + \frac{(\sigma_k)^2}{\gamma_k} + \frac{2\nu_k |x^0 - u^0|^\top |y^0 - v^0|}{\hat{\beta}{x^0}^\top y^0}, \tag{21}$$

$$\hat{\omega}_k := \left(\hat{\xi}_k + \sqrt{(\hat{\xi}_k)^2 + \hat{\eta}_k}\right)^2. \tag{22}$$

Note the similarity to the definitions of ξ_k, η_k, and ω_k in (9), (10) and (11). $\hat{\xi}_k$ is identical to ξ_k except that it has γ_k in the denominator instead of γ. $\hat{\eta}_k$ differs from η_k only by dividing the last term by $\hat{\beta}$.

Lemma 3.5 *Let $\{(x^k, y^k)\}$ and $\{(\Delta x^k, \Delta y^k)\}$ be generated by Algorithm 3 and let $D^k := (Y^k)^{1/2}(X^k)^{-1/2}$. Then*

$$\left\|D^k \Delta x^k\right\|^2 + \left\|(D^k)^{-1}\Delta y^k\right\|^2 \leq \omega_k {x^k}^\top y^k.$$

Moreover, the sequence $\{\omega_k\}$ is bounded, i.e. there is a constant $\omega > 0$ such that $\omega_k \leq \omega$, for all k.

Proof The proof is similar to the proof of Lemma 2.9 with the following changes:

1. Replace ξ_k, η_k, and ω_k by $\hat{\xi}_k, \hat{\eta}_k$, and $\hat{\omega}_k$, respectively.

2. Replace (16) with the inequality

$$(x^k - u^k)^\top (y^k - v^k) \le \frac{\nu_k {x^k}^\top y^k |x^0 - u^0|^\top |y^0 - v^0|}{\hat{\beta} {x^0}^\top y^0}. \tag{23}$$

   which we justify by Lemmas 2.4(2) and 3.4.

3. Replace the last line of (17) with

$$\begin{aligned} &{x^k}^\top y^k [1 - 2\sigma_k + \sigma_k^2/\gamma_k] \\ &\quad \ge (t_k)^2 - 2({x^k}^\top y^k)^{1/2} \xi_k t_k - 2\frac{\nu_k {x^k}^\top y^k |x^0 - u^0|^\top |y^0 - v^0|}{\hat{\beta} {x^0}^\top y^0} \quad \text{(by (23))}. \end{aligned}$$

We then get the inequality

$$(t_k)^2 \le (\xi_k + \sqrt{(\xi_k)^2 + \eta_k})^2 {x^k}^\top y^k = \omega_k {x^k}^\top y^k.$$

Clearly, $\{\eta_k\}$ is bounded. Moreover, from Lemmas 2.7 and 3.4,

$$|\xi_k| = \left(\frac{n}{\gamma_k}\right)^{1/2} \frac{|x^k - u^k|^\top y^k + |y^k - v^k|^\top x^k}{{x^k}^\top y^k} \le \left(\frac{n}{\gamma}\right)^{1/2} K,$$

so $\{\xi_k\}$ is also bounded. Hence, $\{\omega_k\}$ is bounded. □

We can now state the global Q-linear convergence theorem.

Theorem 3.6 *Under Assumptions 2.1–2.2, there is a constant $\delta \in (0, 1)$ such that*

$$\phi_{k+1} \le (1 - \delta)\phi_k, \quad k = 0, \dots,$$

that is, Algorithm 3 converges globally and Q-linearly.

Proof The proof is identical to the proof of [Wri92a, Theorem 4.2] but using $\hat{\xi}_k, \hat{\eta}_k$, and $\hat{\omega}_k$ in place of ξ_k, η_k, and ω_k, and also using Lemma 3.5 in place of [Wri92a, Lemma 4.1]. □

We now turn our attention toward proving global Q-quadratic convergence. We shall need to use Wright's stronger assumptions 3.1–3.3.

We first prove two lemmas which place bounds on the iterates (x^k, y^k). We need the following definitions:

$$B = \{i | x_i^* > 0\}, \quad N = \{i | y_i^* > 0\}.$$

where (x^*, y^*) is the strictly complementary solution guaranteed by Assumption 3.3. Note that $N \bigcup B = \{1, 2, \dots, n\}$ and $N \bigcap B = \emptyset$.

Lemma 3.7 *Let $\{(x^k, y^k)\}$ be generated by Algorithm 3. There is a constant $C_1 > 0$ such that*

$$i \in N \Rightarrow x_i^k \le C_1\mu_k, \quad y_i^k \ge \gamma/C_1, \tag{24a}$$

$$i \in B \Rightarrow y_i^k \le C_1\mu_k, \quad x_i^k \ge \gamma/C_1. \tag{24b}$$

Proof Define $\tilde{x} := x^* - (x^0 - x^*)_-$ and $\tilde{y} := y^* - (y^0 - y^*)_-$, where (x^*, y^*) is the strictly complementary solution guaranteed by Assumption 3.3. By applying Lemma 2.5 with $(\hat{x}, \hat{y}) = (u^0, v^0) = (x^*, y^*)$, and noting that ${x^*}^\top y^* = 0$, we get

(25)

$$\begin{aligned}
&\tilde{x}^\top y^k + \tilde{y}^\top x^k \\
&\quad \le {x^k}^\top y^k + \nu_k(|x^0 - x^*|^\top y^* + |y^0 - y^*|^\top x^* + |x^0 - x^*|^\top |y^0 - y^*|) \\
&\quad = {x^k}^\top y^k \left(1 + \nu_k \left(\frac{|x^0 - x^*|^\top y^* + |y^0 - y^*|^\top x^* + |x^0 - x^*|^\top |y^0 - y^*|}{{x^k}^\top y^k}\right)\right) \\
&\quad \le \bar{C}_1 {x^k}^\top y^k, \quad \text{by Lemmas 2.4(3) and 3.4,}
\end{aligned}$$

where

$$\bar{C}_1 := \left(1 + \frac{|x^0 - x^*|^\top y^* + |y^0 - y^*|^\top x^* + |x^0 - x^*|^\top |y^0 - y^*|}{\hat{\beta} {x^0}^\top y^0}\right).$$

Thus,

$$\sum_{i=1}^{n} \left(\tilde{x}_i^\top y_i^k + {x_i^k}^\top \tilde{y}_i\right) \le n\bar{C}_1\mu_k. \tag{26}$$

Now, $\tilde{y}_i = \min(y_i^0, y_i^*) \ge 0$ and $\tilde{x}_i = \min(x_i^0, x_i^*) \ge 0$. So, each term on the left side of (26) is nonnegative. Therefore,

$$x_i^k \tilde{y}_i \le n\bar{C}_1\mu_k, \quad \text{and} \quad \tilde{x}_i y_i^k \le n\bar{C}_1\mu_k.$$

Note further that for $i \in N$, $\tilde{y}_i > 0$, so

$$x_i^k \le \frac{n\bar{C}_1}{\tilde{y}_i}\mu_k,$$

Similarly, for $i \in B, \tilde{x}_i > 0$ and

$$y_i^k \le \frac{n\bar{C}_1}{\tilde{x}_i}\mu_k,$$

Finally, we obtain our result by taking

$$C_1 := n\bar{C}_1 \max\left(\max_{i \in B} \frac{1}{\tilde{x}_i}, \max_{i \in N} \frac{1}{\tilde{y}_i}\right).$$

Then for $i \in B, x_i^k \le C_1\mu_k$, and by (19d)

$$x_i^k y_i^k \ge \gamma_k\mu_k \Rightarrow y_i^k \ge \frac{\gamma_k\mu_k}{x_i^k} \ge \frac{\gamma_k}{C_1} \ge \frac{\gamma}{C_1}.$$

Similarly, for $i \in N, y_i^k \le C_1\mu_k$ and $x_i^k \ge \gamma/C_1$. □

Lemma 3.8 *Let $\{(x^k, y^k)\}$ be generated by Algorithm 3. There is a constant $C_2 > 0$ such that*

$$0 < x_i^k \leq C_2, \quad 0 < y_i^k \leq C_2. \tag{27}$$

Proof Define $\check{x} := \bar{x} - (x^0 - \bar{x})_-$ and $\check{y} := \bar{y} - (y^0 - \bar{y})_-$, where $(\bar{x}, \bar{y})$ is the strictly feasible point guaranteed by Assumption 3.2. Note that $(\check{x}, \check{y}) > 0$. Now, by applying Lemma 2.5 with $(\hat{x}, \hat{y}) = (u^0, v^0) = (\bar{x}, \bar{y})$, we get

$$\begin{aligned}
&\check{x}^\top y^k + \check{y}^\top x^k \\
&\leq \bar{x}^\top \bar{y} + {x^k}^\top y^k + \nu_k(|x^0 - \bar{x}|^\top \bar{y} + |y^0 - \bar{y}|^\top \bar{x} + |x^0 - \bar{x}|^\top |y^0 - \bar{y}|) \\
&\leq \bar{x}^\top \bar{y} + {x^0}^\top y^0 + |x^0 - \bar{x}|^\top \bar{y} + |y^0 - \bar{y}|^\top \bar{x} + |x^0 - \bar{x}|^\top |y^0 - \bar{y}| \\
&=: \bar{C}_2.
\end{aligned} \tag{28}$$

Hence,

$$0 < y_i^k \leq \frac{\bar{C}_2}{\check{x}_i}, \quad 0 < x_i^k \leq \frac{\bar{C}_2}{\check{y}_i}, \quad i = 1, 2, \ldots, n.$$

The result is obtained by setting

$$C_2 := \bar{C}_2 \max\left(\max_{i=1,\ldots,n} \frac{1}{\check{x}_i}, \max_{i=1,\ldots,n} \frac{1}{\check{y}_i} \right).$$

□

The remainder of Wright's results can now be proved simply by replacing all references to [Wri92a, Lemmas 3.4 and 3.5] by references to our Lemmas 3.7 and 3.8.

Theorem 3.9 *Under Assumptions 3.1–3.3, there is an $\hat{\epsilon} > 0$ such that if K is the smallest integer such that $\phi_K \leq \hat{\epsilon}$, then*

1. *the algorithm will take fast steps at iteration K and at all subsequent iterations, and*
2. *the sequences $\{\mu_k\}$ and $\{\phi_k\}$ converge Q-quadratically to zero.*

Proof The proof is identical to the proof of [Wri92a, Theorems 6.3 and 6.4], but using Lemmas 3.7 and 3.8 in place of [Wri92a, Lemmas 3.4 and 3.5]. □

Corollary 3.10 *Algorithm 3 has local Q-quadratic convergence.*

Proof Follows immediately from Theorem 3.9. □

4 Algorithm for Box-Constrained LCP

We now turn our attention to linear complementarity problems with more general constraints. In particular, we consider the box-constrained linear complementarity problem (BLCP). Before we state this problem, we need to make some definitions. Let $\mathbb{B} := \prod_{i=1}^{n}[l_i, u_i]$, where $l_i \in [-\infty, \infty)$ and $u_i \in (-\infty, \infty]$ and for each i, at least one of l_i and u_i is finite. We also define a generalized inner product $\langle(\cdot,\cdot),\cdot\rangle : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, by

$$\langle (w,v), z\rangle := \sum_{i\in\{i:l_i>-\infty\}} w_i(z_i - l_i) + \sum_{i\in\{i:u_i<\infty\}} v_i(u_i - z_i).$$

The box-constrained linear complementarity problem (BLCP) can now be stated as follows:

(BLCP) Given $M \in \mathbb{R}^{n\times n}, q \in \mathbb{R}^n$, find a vector triple $(z, w, v) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ such that

$$\begin{aligned} & w - v = Mz + q, \\ & z \in \mathbb{B}, \quad w \geq 0, \quad v \geq 0 \\ & \langle (w,v), z\rangle = 0. \end{aligned}$$

Note that the linear complementarity problem is a special case of the box-constrained linear complementarity problem (simply set $l = 0$ and $u = \infty$).

We now make some observations about (BLCP). Let us partition the indices according to which bounds are finite.

$$H := \{i : -\infty < l_i, u_i < \infty\}, \quad J := \{i : u_i = \infty\}, \quad K := \{i : l_i = -\infty\}.$$

Note that H, J, and K are disjoint and further that $H \bigcup J \bigcup K = \{1, \ldots, n\}$. Without loss of generality, we can assume that the rows and columns of M and the vectors q, l, and u are ordered so that the indices in H occur first, those in J occur second, and those in K occur last. Let p, s, and t be the size of the sets H, J, and K respectively.

Note that if (z^*, w^*, v^*) is a solution to (BLCP), then $w^*_K = 0$, and $v^*_J = 0$. Thus, we can remove w_K and v_J from the problem. This motivates the definition of the set

$$\mathcal{G}_1 := \{(w,v) \in \mathbb{R}^n \times \mathbb{R}^n : w_K = 0, v_J = 0\}.$$

We now define an invertible linear map $L : \mathcal{G}_1 \to \mathbb{R}^{p+s} \times \mathbb{R}^{p+t}$ by

$$L(w,v) := (\hat{w}, \hat{v}),$$

$$\text{where} \quad \hat{w} := \begin{pmatrix} w_H \\ w_J \end{pmatrix} \quad \textit{and} \quad \hat{v} := \begin{pmatrix} v_H \\ v_K \end{pmatrix}.$$

Our plan now is to create an algorithm that generates iterates $\{(z^k, \hat{w}^k, \hat{v}^k)\}$ such that $\{(z^k, w^k, v^k)\} := \{(z^k, L^{-1}(\hat{w}^k, \hat{v}^k))\}$ converges to a solution (z^*, w^*, v^*) of (BLCP). To do this, we shall exploit the fact that (BLCP) can be reformulated as an LCP with higher dimension. We shall now discuss this reformulation.

Define the maps

$$\mathcal{X} : \mathbb{R}^n \times \mathcal{G}_1 \to \mathbb{R}^{n+p} := (z, w, v) \mapsto (z_H - l_H, z_J - l_J, u_K - z_K, v_H),$$
$$\mathcal{Y} : \mathbb{R}^n \times \mathcal{G}_1 \to \mathbb{R}^{n+p} := (z, w, v) \mapsto (w_H, w_J, v_K, u_H - z_H).$$

In order to refer to the last p components of $\mathcal{X}$ and $\mathcal{Y}$, we define the set of indices $\hat{H} = H+n$. For example, if $x = \mathcal{X}(z, w, v)$, then $x_{\hat{H}} = v_H$. Now, define the set

$$\mathcal{G}_2 := \{(x, y) \in \mathbb{R}^{n+p} \times \mathbb{R}^{n+p} : x_H + y_{\hat{H}} = u_H - l_H\}.$$

We can now define an invertible linear map $T : \mathbb{R}^n \times \mathcal{G}_1 \to \mathcal{G}_2$ by the relation

$$T(z, w, v) := (\mathcal{X}(z, w, v), \mathcal{Y}(z, w, v)).$$

We are now ready to state the reformulation of (BLCP). Let

$$(x, y) := T(z, w, v), \tag{29}$$

$$\hat{M} := \begin{pmatrix} M_{HH} & M_{HJ} & -M_{HK} & I \\ M_{JH} & M_{JJ} & -M_{JK} & 0 \\ -M_{KH} & -M_{KJ} & M_{KK} & 0 \\ -I & 0 & 0 & 0 \end{pmatrix}, \tag{30}$$

$$\hat{h} := \begin{pmatrix} -q_H - M_{HH} l_H - M_{HJ} l_J - M_{HK} u_K \\ -q_J - M_{JH} l_H - M_{JJ} l_J - M_{JK} u_K \\ q_K + M_{KH} l_H + M_{KJ} l_J + M_{KK} u_K \\ -u_H + l_H \end{pmatrix}, \qquad m := n + p. \tag{31}$$

With these definitions, (BLCP) is equivalent to the linear complementarity problem formed by replacing M, h, and n in (LCP) by $\hat{M}$, $\hat{h}$, and m, respectively. Thus, given a starting point $(z^0, w^0, v^0) \in (\mathbb{R}^n \times \mathcal{G}_1) \cap \mathbb{R}^{3n}_{++}$ we can solve (BLCP) simply by applying Algorithm 3 with the starting point $(x^0, y^0) := T(z^0, w^0, v^0)$. If the algorithm finds a solution (x^*, y^*) of (LCP) with $M = \hat{M}$, $h = \hat{h}$, then $(x^*, y^*) \in \mathcal{G}_2$ and $(z^*, w^*, v^*) := T^{-1}(x^*, y^*)$ is a solution of (BLCP).

Our plan now is to substitute (29)–(31) into Algorithm 3 and to simplify in order to produce an algorithm that generates iterates $\{(z^k, \hat{w}^k, \hat{v}^k)\}$, such that for all k,

$$T(z^k, L^{-1}(\hat{w}^k, \hat{v}^k)) = (x^k, y^k),$$

where $\{(x^k, y^k)\}$ are the iterates generated by Algorithm 3. Throughout our discussion, we will occasionally find it convenient to refer to (w^k, v^k). In such cases, we are implying the relationship $(w^k, v^k) = L^{-1}(\hat{w}^k, \hat{v}^k)$.

We look first at the equation used to calculate the search direction. Direct substitution into (18) yields

$$(32)\qquad \begin{pmatrix} M_{HH} & M_{HJ} & -M_{HK} & I_p & -I_p & 0 & 0 & 0 \\ M_{JH} & M_{JJ} & -M_{JK} & 0 & 0 & -I_s & 0 & 0 \\ -M_{KH} & -M_{KJ} & M_{KK} & 0 & 0 & 0 & -I_t & 0 \\ -I_p & 0 & 0 & 0 & 0 & 0 & 0 & -I_p \\ W_H^k & 0 & 0 & 0 & R_H^k & 0 & 0 & 0 \\ 0 & W_J^k & 0 & 0 & 0 & R_J^k & 0 & 0 \\ 0 & 0 & V_K^k & 0 & 0 & 0 & S_K^k & 0 \\ 0 & 0 & 0 & S_H^k & 0 & 0 & 0 & V_H^k \end{pmatrix} \begin{pmatrix} \Delta x_H^k \\ \Delta x_J^k \\ \Delta x_K^k \\ \Delta x_{\hat{H}}^k \\ \Delta y_H^k \\ \Delta y_J^k \\ \Delta y_K^k \\ \Delta y_{\hat{H}}^k \end{pmatrix} = \begin{pmatrix} w_H^k - v_H^k - M_{H\cdot}z^k - q_H \\ w_J^k - M_{J\cdot}z^k - q_J \\ v_K^k + M_{K\cdot}z^k + q_K \\ 0 \\ -W_H^k R_H^k e + \sigma_k\mu_k e \\ -W_J^k R_J^k e + \sigma_k\mu_k e \\ -V_K^k S_K^k e + \sigma_k\mu_k e \\ -V_H^k S_H^k e + \sigma_k\mu_k e \end{pmatrix}$$

where $R := \operatorname{diag}(z-l)$, and $S := \operatorname{diag}(u-z)$. By the fourth row of this system, $\Delta x_H^k = -\Delta y_{\hat{H}}^k$. We can thus replace the last equation of (32) with $-V^k \Delta x_H^k + S_H^k \Delta x_{\hat{H}}^k = -V_H^k S_H^k e + \sigma_k \mu_k e$. Removing, the fourth row and the last column, we get

$$(33)\qquad \begin{pmatrix} M_{HH} & M_{HJ} & M_{HK} & I_p & -I_p & 0 & 0 \\ M_{JH} & M_{JJ} & M_{JK} & 0 & 0 & -I_s & 0 \\ M_{KH} & M_{KJ} & M_{KK} & 0 & 0 & 0 & I_t \\ W_H^k & 0 & 0 & 0 & R_H^k & 0 & 0 \\ 0 & W_J^k & 0 & 0 & 0 & R_J^k & 0 \\ 0 & 0 & -V_K^k & 0 & 0 & 0 & S_K^k \\ -V_H^k & 0 & 0 & S_H^k & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta x_H^k \\ \Delta x_J^k \\ -\Delta x_K^k \\ \Delta x_{\hat{H}}^k \\ \Delta y_H^k \\ \Delta y_J^k \\ \Delta y_K^k \end{pmatrix} = \begin{pmatrix} w_H^k - v_H^k - M_{H\cdot}z^k - q_H \\ w_J^k - M_{H\cdot}z^k - q_J \\ -v_K^k - M_{H\cdot}z^k - q_K \\ -W_H^k R_H^k e + \sigma_k\mu_k e \\ -W_J^k R_J^k e + \sigma_k\mu_k e \\ -V_K^k S_K^k e + \sigma_k\mu_k e \\ -V_H^k S_H^k e + \sigma_k\mu_k e \end{pmatrix}.$$

Finally, moving column 4 to column 6, and switching rows 6 and 7, we obtain the equation

$$(34)\qquad \begin{pmatrix} M & -I_W & I_V \\ \hat{W}^k & \hat{R}^k & 0 \\ -\hat{V}^k & 0 & \hat{S}^k \end{pmatrix} \begin{pmatrix} \Delta z^k \\ \Delta \hat{w}^k \\ \Delta \hat{v}^k \end{pmatrix} = \begin{pmatrix} -q - Mz^k - v^k + w^k \\ -\hat{R}^k \hat{W}^k e + \sigma_k\mu_k e \\ -\hat{S}^k \hat{V}^k e + \sigma_k\mu_k e \end{pmatrix},$$

where

$$\Delta z^k := \begin{pmatrix} \Delta x_H^k \\ \Delta x_J^k \\ -\Delta x_K^k \end{pmatrix}, \quad \Delta \hat{w}^k := \begin{pmatrix} \Delta y_H^k \\ \Delta y_J^k \end{pmatrix}, \quad \Delta \hat{v}^k := \begin{pmatrix} \Delta x_{\hat{H}}^k \\ \Delta y_K^k \end{pmatrix},$$

$$I_W := \begin{pmatrix} I_p & 0 \\ 0 & I_s \\ 0 & 0 \end{pmatrix}, \quad I_V := \begin{pmatrix} I_p & 0 \\ 0 & 0 \\ 0 & I_t \end{pmatrix},$$

$$\hat{W}^k := \begin{pmatrix} W_H^k & 0 & 0 \\ 0 & W_J^k & 0 \end{pmatrix}, \quad \hat{V}^k := \begin{pmatrix} V_H^k & 0 & 0 \\ 0 & 0 & V_K^k \end{pmatrix},$$

$$\hat{R} := \begin{pmatrix} R_H & 0 \\ 0 & R_J \end{pmatrix}, \quad \hat{S} := \begin{pmatrix} S_H & 0 \\ 0 & S_k \end{pmatrix}.$$

We now turn our attention to the equations governing the calculation of the steplength. Let us define the merit function

$$\psi(z, \hat{w}, \hat{v}) := \phi(x, y), \qquad \text{where } (x, y) = T(z, w, v).$$

Then $\psi(z, \hat{w}, \hat{v}) = \langle (w, v), z \rangle + \|w - v - Mz - q\|$. Define

$$\begin{aligned} z(\alpha) &:= z^k + \alpha \Delta z^k, \\ \hat{w}(\alpha) &:= \hat{w}^k + \alpha \Delta \hat{w}^k, \\ \hat{v}(\alpha) &:= \hat{v}^k + \alpha \Delta \hat{v}^k. \end{aligned}$$

Direct substitution into (19) gives the following equations for calculating the steplength for the BLCP-algorithm:

$$\alpha_k = \arg\min_{\alpha} \psi(z(\alpha), \hat{w}(\alpha), \hat{v}(\alpha)) \tag{35}$$

subject to

$$\alpha \in [0, 1], \tag{36a}$$

$$l < z(\alpha) < u, \quad \text{and} \quad (\hat{w}(\alpha), \hat{v}(\alpha)) > 0, \tag{36b}$$

$$\langle (w(\alpha), v(\alpha)), z(\alpha) \rangle \geq (1 - \beta_k)(1 - \alpha) \langle (w^k, v^k), z^k \rangle, \tag{36c}$$

$$(z(\alpha)_i - l_i) w(\alpha)_i \geq (\gamma_k / 2n) \langle (w(\alpha), v(\alpha)), z(\alpha) \rangle, \quad i \in H \bigcup J, \tag{36d}$$

$$(u_i - z(\alpha)_i) v(\alpha)_i \geq (\gamma_k / 2n) \langle (w(\alpha), v(\alpha)), z(\alpha) \rangle, \quad i \in H \bigcup K. \tag{36e}$$

Finally, we note that from Algorithm 3, $\mu_k = {x^k}^\top y^k / m = \langle (w^k, v^k), z^k \rangle / (n + p)$. The complete algorithm is given in Algorithm 4.

By construction, there is a 1-1 correspondence between the iterates $\{(x^k, y^k)\}$ of Algorithm 3 and the iterates $\{(z^k, \hat{w}^k, \hat{v}^k)\}$ of Algorithm 4 given by $T(z^k, L^{-1}(\hat{w}^k, \hat{v}^k)) = (x^k, y^k)$. Thus, we can prove convergence results for the iterates $\{(z^k, \hat{w}^k, \hat{v}^k)\}$ of Algorithm 4 simply by analyzing the iterates $\{(x^k, y^k)\}$ of Algorithm 3.

Algorithm 4– BLCP Algorithm

Given $\gamma \in (0, 1/2), \sigma \in (0, 1/2), \rho \in (0, \gamma)$,
$\quad \psi > 0$, and $(z^0, \hat{w}^0, \hat{v}^0)$ with $l < z^0 < u$, $(\hat{w}^0, \hat{v}^0) > 0$,
$\quad (z_i - l_i)w_i \geq 2\gamma\mu_0$ for $i \in H \bigcup J$, and $v_i(u_i - z_i) \geq 2\gamma\mu_0$ for $i \in H \bigcup K$;
$t_0 \leftarrow 1, \gamma_0 \leftarrow 2\gamma$;
for $k = 0, 1, 2, \ldots$
 if $\quad \psi_k := \psi(z^k, \hat{w}^k, \hat{v}^k) \leq \psi$
 then Compute a "fast" step by setting $\sigma_k \leftarrow \mu_k, \beta_k \leftarrow \gamma^{t_k}$,
  and $\gamma_k \leftarrow \gamma(1 + \gamma^{t_k})$ and solving (34)–(36) to calculate
  $(\Delta z^k, \Delta \hat{w}^k, \Delta \hat{v}^k)$ and α_k;
  if $\quad \psi_k(z^k + \alpha_k \Delta z^k, \hat{w}^k + \alpha_k \Delta \hat{w}^k, \hat{v}^k + \alpha_k \Delta \hat{v}^k) \leq \rho\psi_k$
  then $(z^{k+1}, \hat{w}^{k+1}, \hat{v}^{k+1}) \leftarrow (z^k, \hat{w}^k, \hat{v}^k) + \alpha_k(\Delta z^k, \Delta \hat{w}^k, \Delta \hat{v}^k)$,
   $t_{k+1} \leftarrow t_k + 1$;
   go to next k;
  end if
 end if
 Compute the "safe" step by setting $\sigma_k \in [\sigma, 1/2], \beta_k \leftarrow 0, \gamma_k \leftarrow \gamma_{k-1}$
 and solving (34)–(36) to calculate $(\Delta z^k, \Delta \hat{w}^k, \Delta \hat{v}^k)$ and α_k;
 $(z^{k+1}, \hat{w}^{k+1}, \hat{v}^{k+1}) \leftarrow (z^k, \hat{w}^k, \hat{v}^k) + \alpha_k(\Delta z^k, \Delta \hat{w}^k, \Delta \hat{v}^k)$,
 $t_{k+1} \leftarrow t_k$;
 $(z^{k+1}, \hat{w}^{k+1}, \hat{v}^{k+1}) \leftarrow (z(\alpha_k), \hat{w}(\alpha_k), \hat{v}(\alpha_k))$;
 go to next k;
end for.

We now state several convergence theorems for Algorithm 4. These results are based on the following assumptions:

Assumption 4.1 *M is positive semidefinite.*

Assumption 4.2 *(BLCP) has a point $(\bar{z}, \bar{w}, \bar{v}) \in \mathcal{I}_+ := \{(z, w, v) : l < z < u, L(w, v) > 0, w_K = 0, v_J = 0\}$ and $\bar{w} - \bar{v} = M\bar{z} + q$.*

Assumption 4.3 *The solution set for (BLCP) is nonempty and, moreover, there is a strictly complementary solution (z^*, w^*, v^*), that is $z_i^* = l_i \implies w_i^* > 0$ and $z_i^* = u_i \implies v_i^* > 0$.*

The following lemma shows that the above assumptions guarantee the assumptions for the convergence of Algorithm 3.

Lemma 4.4 *Given the relationship between (BLCP) and (LCP) defined by equations (29) – (31), (i) Assumption 4.1 $\Rightarrow$ Assumption 3.1; (ii) Assumption 4.2 $\Rightarrow$ Assumption 3.2; (iii) Assumption 4.3 $\Rightarrow$ Assumption 3.3;*

Proof (i)

$$\begin{aligned}(x^\top, y^\top, z^\top, w^\top)\hat{M}(x; y; z; w) \\ &= x^\top M_{HH}x + x^\top M_{HJ}y + y^\top M_{JH}x - x^\top M_{HK}z - z^\top M_{KH}x + x^\top Iw \\ &\quad - w^\top Ix + y^\top M_{JJ}y - y^\top M_{JK}z - z^\top M_{KJ}y + z^\top M_{KK}z \\ &= (x^\top, y^\top, -z^\top)M(x; y; -z),\end{aligned}$$

so $\hat{M}$ is positive semidefinite whenever M is positive semidefinite.

(ii) If $(\bar{z}, \bar{w}, \bar{v}) \in \mathcal{I}_+$, then we can define $(\bar{x}, \bar{y}) := T(\bar{z}, \bar{w}, \bar{v})$. Clearly, $(\bar{x}, \bar{y}) > 0$. Moreover,

$$\hat{M}\bar{x} - \hat{h} = \hat{M}\begin{pmatrix} \bar{z}_H - l_H \\ \bar{z}_J - l_J \\ u_K - \bar{z}_K \\ \bar{v}_H \end{pmatrix} - \hat{h} = \begin{pmatrix} M_{H.}\bar{z} + q_H \\ M_{J.}\bar{z} + q_J \\ -M_{K.}\bar{z} - q_K \\ u_H - \bar{z}_H \end{pmatrix} = \begin{pmatrix} \bar{w}_H \\ \bar{w}_J \\ \bar{v}_K \\ u - \bar{z}_H \end{pmatrix} = \bar{y}.$$

Thus, $(\bar{x}, \bar{y})$ is a strictly feasible point for (LCP).

(iii) By a similar argument to (ii), if (z^*, w^*, v^*) is a strictly complementary solution of (BLCP), then $(x^*, y^*) := T(z^*, w^*, v^*)$ is a solution to (LCP). It is easy to check from Assumption 4.3 and the definition of T that $x_i^* = 0$ implies $y_i^* > 0$, so that (x^*, y^*) is strictly complementary. □

We can now state the following convergence theorems for Algorithm 4.

Theorem 4.5 *Under Assumptions 4.1–4.3, there is a constant $\delta \in (0, 1)$ such that*

$$\psi_{k+1} \leq (1 - \delta)\psi_k, \quad k = 0, 1, 2, \ldots,$$

that is, Algorithm 4 converges globally and Q-linearly to a solution of (BLCP).

Proof By construction, $\psi_k = \phi_k$, where ϕ_k is as defined in Algorithm 3. The result follows from Theorem 3.6. □

Theorem 4.6 *Under Assumptions 4.1–4.3, there is an $\hat{\epsilon} > 0$ such that if K is the smallest positive integer such that $\psi_K \leq \hat{\epsilon}$, then*

1. *Algorithm 4 will take fast steps at iteration K and at all subsequent iterations, and*

2. *the sequences $\{\mu_k\}$ and $\{\psi_k\}$ converge Q-quadratically to zero.*

Proof Follows directly from the definitions of μ_k and ψ_k and Theorem 3.9. □

5 Summary and Conclusions

In this paper, we have extended the convergence results of Zhang and Wright to apply to arbitrary strictly positive starting points. This extension is important because it allows the convergence theory to be applied to cases where the algorithms are warm-started from points not satisfying Zhang's and Wright's restrictions.

The extension also plays an important role in proving the convergence results of the BLCP algorithm discussed in Section 4. Recall that the restriction imposed by Zhang and Wright on the starting point can easily be satisfied simply by making the starting point (x^0, y^0) large in every component. However, in the analysis of the BLCP algorithm, we defined $(x^0, y^0) = T(z^0, w^0, v^0)$ so that (x^0, y^0) is required to lie in the range of T. In particular, $x_H + y_{\hat{H}} = u_H - l_H$, so increases in the components of x_H must be offset by decreases in the components of $y_{\hat{H}}$. Thus, for the (BLCP), it may not be possible to find a starting point that meets the restriction. By removing the restriction from the convergence results, this difficulty is eliminated.

References

[Ans89] K.M. Anstreicher, *A combined phase I–phase II projective algorithm for linear programming*, Mathematical Programming **43** (1989), 209–223.

[Ans91] K.M. Anstreicher, *A combined phase I–phase II scaled potential algorithm for linear programming*, Mathematical Programming **52** (1991), 429–440.

[CPS92] R. W. Cottle, J. S. Pang, and R. E. Stone, *The linear complementarity problem*, Academic Press, Boston, 1992.

[Gül92] O. Güler, *Generalized linear complementarity problems and interior point algorithms for their solutions.*, Tech. report, Technical Mathematics and Computer Science, Delft University of Technology, The Netherlands, April 1992.

[KMM91] M. Kojima, N. Meggido, and S. Mizuno, *A primal-dual exterior point algorithm for linear programming*, Tech. Report RJ 8500, IBM Alamaden Research Center, San Jose, California, 1991.

[KMT92] M. Kojima, S. Mizuno, and M. Todd, *Infeasible interior-point primal-dual potential-reduction algorithms for linear programming*, Tech. Report 1023, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, August 1992.

[LMS91] 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** (1991), 191–222.

[LMS92] I.J. Lustig, R.E. Marsten, and D.F. Shanno, *On implementting Mehrotra's predictor-corrector interior point method for linear programming*, SIAM Journal on Optimization **2** (1992), 435–449.

[Meh92] S. Mehrotra, *On the implementation of a primal-dual interior point method*, SIAM Journal on Optimization **2** (1992), 575–601.

[Miz93] S. Mizuno, *Polynomiality of Kojima-Meggido-Mizuno infeasible-interior-point algorithm for linear programming*, Tech. Report 1006, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, January 1993.

[MMS89] K.A. McShane, C.L. Monma, and D.F. Shanno, *An implementation of a primal-dual interior point method for linear programming*, ORSA Journal of Computing (1989), 70–83.

[Pot92a] F. A. Potra, *An infeasible interior-point predictor-corrector algorithm for linear programming*, Tech. Report 26, Department of Mathematics, University of Iowa, Iowa City, Iowa, June 1992.

[Pot92b] F. A. Potra, *A quadratically convergent infeasible interior-point algorithm for linear programming*, Tech. Report 28, Department of Mathematics, University of Iowa, Iowa City, Iowa, July 1992.

[Wri92a] Stephen Wright, *An infeasible-interior-point algorithm for linear complementarity problems*, Preprint MCS-P331-1092, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois, October 1992.

[Wri92b] Stephen Wright, *A path-following infeasible-interior-point algorithm for linear complementarity problems*, Preprint MCS-P334-1192, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois, November 1992.

[Zha92] Y. Zhang, *On the convergence of an infeasible interior-point algorithm for linear programming and other problems*, Tech. Report 92-07, Department of Mathematics and Statistics, University of Maryland, Baltimore County, Maryland, April 1992.

[ZZ92] Y. Zhang and D. Zhang, *Superlinear convergence of infeasible interior-point methods for linear programming*, Tech. Report 92-15, Department of Mathematics and Statistics, University of Maryland, Baltimore County, Maryland, September 1992.