

**ERROR BOUND AND CONVERGENCE ANALYSIS
OF MATRIX SPLITTING ALGORITHMS
FOR THE AFFINE VARIATIONAL INEQUALITY PROBLEM***

by

Paul Tseng[†] and Zhi-Quan Luo[‡]

ABSTRACT

Consider the affine variational inequality problem. We show that, near the solution set, the distance from a point to the solution set can be bounded by the norm of the natural residual at that point. We then use this bound to prove linear convergence of a matrix splitting algorithm for solving the symmetric case of the problem. This latter result improves upon a recent result of Luo and Tseng which further assumes the problem to be monotone.

KEY WORDS. Affine variational inequality, linear complementarity, error bound, matrix splitting, linear convergence.

* This research is partially supported by the U.S. Army Research Office, contract DAAL03-86-K-0171 (Center for Intelligent Control Systems), by the National Science Foundation, grant NSF-DDM-8903385, and by a grant from the Science and Engineering Research Board of McMaster University.

[†] Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

[‡] Room 225/CRL, Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, L8S 4L7, Canada.

1 Introduction

Let M be an $n \times n$ matrix and let q be a vector in \mathfrak{R}^n , the n -dimensional Euclidean space. Let X be a polyhedral set in \mathfrak{R}^n . We consider the following *affine* variational inequality problem associated with F and X :

$$\text{find an } x^* \in X \text{ satisfying } \langle x - x^*, Mx^* + q \rangle \geq 0, \quad \forall x \in X. \quad (1.1)$$

The problem (1.1) is well-known in optimization, containing as special cases linear (and quadratic) programming, bimatrix games, etc. (see Cottle and Dantzig [CoD68]). When X is the non-negative orthant in \mathfrak{R}^n , it is called the linear complementarity problem (LCP for short). We will not attempt to survey the literature on this problem, which is vast. Expository articles on the subject include [CoD68], [Eve71b], [CGL80], [Mur88]. For discussion of variational inequality problems in general, see [Aus76], [BeT89], [CGL80], [KiS80].

Let X^* denote the set of solutions of the affine variational inequality problem (1.1), which we assume hereon to be nonempty. It is well-known (and not difficult to see from the convexity of X) that X^* is precisely the set of fixed points of the nonlinear mapping $x \mapsto [x - Mx - q]^+$, where $[\cdot]^+$ denotes the orthogonal projection onto X , i.e., $[x]^+ = \arg \min_{z \in X} \|x - z\|$ and $\|\cdot\|$ denotes the Euclidean norm in \mathfrak{R}^n , so that

$$X^* = \{ x^* \in \mathfrak{R}^n \mid x^* = [x^* - Mx^* - q]^+ \}. \quad (1.2)$$

An important topic in the study of variational inequalities and complementarity problems concerns *error bounds* for estimating the closeness of a point to X^* (see [Pan85],[MaD88],[MaS86]). Such error bounds can serve as termination criteria for iterative algorithms and can be used to estimate the amount of error allowable in an inexact computation of the iterates (see [Pan86b]). Recently Luo and Tseng [LuT90] showed that one such bound, based on the norm of the natural residual function

$$\|x - [x - Mx - q]^+\|, \quad (1.3)$$

is also useful for analyzing the *rate* of convergence of iterative algorithms for solving (1.1). In particular, they showed that, for the problem of minimizing a certain convex essentially smooth function over a polyhedral set, a bound analogous to the above can be used as the basis for proving the linear convergence of a number of well-known iterative algorithms (applied to solve this problem).

The contribution of this paper is two fold: (i) we show that the error bound (1.3) holds locally for the affine variational inequality problem (1.1) for general M , thus extending a result of [LuT90, Sec. 2] for the case where M is symmetric positive semi-definite, (ii) we show, by using the above error bound, that if M is symmetric, then any matrix splitting algorithm using regular Q -splitting, applied to solve (1.1), is linearly convergent. This latter result extends that in [LuT90, Sec. 5],

which proved linear convergence for the same algorithm under the additional assumption that M is positive semi-definite. It also improves upon the results by Pang [Pan84, Sec. 4], [Pan86, Sec. 2], which showed convergence (respectively, weak convergence) for a special case of the algorithm, i.e., one that solves LCP, under the additional assumption that M is non-degenerate (respectively, strictly copositive). Matrix splitting algorithms using regular Q -splitting represent an important class of algorithms for solving affine variational inequality problems and LCP's (see [LiP87]), so the resolution of their convergence (and their rate of convergence) is of great interest. [See Section 3 for a more detailed discussion of the subject.]

This paper proceeds as follows. In Section 2, we prove that an error bound based on (1.3) holds for all points near X^* . In Section 3, we consider the special case of (1.1) where M symmetric and we use the bound of Section 2 to prove linear convergence for matrix splitting algorithms using regular Q -splitting, applied to solve this problem. Finally, in Section 4, we give our conclusion and discuss possible extensions.

We adopt the following notations throughout. For any $x \in \mathfrak{R}^n$ and $y \in \mathfrak{R}^n$, we denote by $\langle x, y \rangle$ the Euclidean inner product of x with y . For any $x \in \mathfrak{R}^n$, we denote by $\|x\|$ the Euclidean norm of x , i.e., $\|x\| = \sqrt{\langle x, x \rangle}$. For any two subsets C_1, C_2 of \mathfrak{R}^n , we denote by $d(C_1, C_2)$ the usual Euclidean distance between the sets C_1 and C_2 , that is

$$d(C_1, C_2) = \inf_{x \in C_1, y \in C_2} \|x - y\|.$$

For any $l \times k$ matrix A , we denote by A^T the transpose of A , by $\|A\|$ the matrix norm of A induced by the vector norm $\|\cdot\|$ (i.e., $\|A\| = \max_{\|x\|=1} \|Ax\|$), by A_i the i -th column of A and, for any nonempty $I \subseteq \{1, \dots, k\}$, by A_I the submatrix of A obtained by removing all columns $i \notin I$ of A . Analogously, for any k -vector x , we denote by x_i the i -th coordinate of x , and, for any nonempty subset $I \subseteq \{1, \dots, k\}$, by x_I the vector with components $x_i, i \in I$ (with the x_i 's arranged in the same order as in x).

2 An Error Bound

In this section we show that $d(x, X^*)$ can be upper bounded by the norm of $x - [x - Mx - q]^+$, the natural residual at x , whenever the latter quantity is small. Our proof, like the proof of Theorem 2.1 in [LuT90], exploits heavily the affine structure of the problem.

Since X is a polyhedral set, we can for convenience express it as

$$X = \{ x \in \mathfrak{R}^n \mid Ax \geq b \},$$

for some $m \times n$ matrix A and some $b \in \mathfrak{R}^m$. Then, for any $x \in X$, the vector $z = [x - Mx - q]^+$ is simply the unique vector which, together with some multiplier vector $\lambda \in \mathfrak{R}^m$, satisfies the Kuhn–Tucker conditions

$$z - x + Mx + q - A^T \lambda = 0, \quad Az \geq b, \quad \lambda \geq 0, \quad (2.1)$$

$$\lambda_i = 0, \quad \forall i \notin I(x), \quad A_i z = b_i, \quad \forall i \in I(x), \quad (2.2)$$

where we denote

$$I(x) = \{ i \in \{1, \dots, m\} \mid A_i z = b_i \}.$$

We say that an $I \subseteq \{1, \dots, m\}$ is *active* at a vector $x \in X$ if $z = [x - Mx - q]^+$ together with some $\lambda \in \mathfrak{R}^m$ satisfies

$$z - x + Mx + q - A^T \lambda = 0, \quad Az \geq b, \quad \lambda \geq 0, \quad (2.3)$$

$$\lambda_i = 0, \quad \forall i \notin I, \quad A_i z = b_i, \quad \forall i \in I. \quad (2.4)$$

[Clearly, $I(x)$ is active at x for all $x \in X$.]

The following lemma, due originally to Hoffman [Hof52] (also see [Rob73], [MaS87]), will be used extensively in the analysis to follow.

Lemma 2.1. Let B be some $k \times l$ matrix and let $S = \{ y \in \mathfrak{R}^l \mid Cy \geq d \}$, for some $h \times l$ matrix C and some $d \in \mathfrak{R}^h$. Then, there exists a scalar constant $\theta > 0$ depending on B and C only such that, for any $\bar{x} \in S$ and any $e \in \mathfrak{R}^k$ such that the linear system $B y = e$, $y \in S$ is consistent, there is a point $\bar{y} \in S$ satisfying $B\bar{y} = e$ and $\|\bar{x} - \bar{y}\| \leq \theta \|B\bar{x} - e\|$.

We next have the following lemma which roughly says that if $x \in X$ is sufficiently close to X^* , then those constraints that are active at x are also active at some element of X^* .

Lemma 2.2. There exists a scalar $\epsilon > 0$ such that, for any $x \in X$ with $\|x - [x - Mx - q]^+\| \leq \epsilon$, $I(x)$ is active at some $x^* \in X^*$.

Proof. We argue by contradiction. If the claim does not hold, then there would exist an $I \subseteq \{1, \dots, m\}$ and a sequence of vectors $\{x^1, x^2, \dots\}$ in X satisfying $I(x^r) = I$ for all r and $x^r - z^r \rightarrow 0$, where we let $z^r = [x^r - Mx^r - q]^+$ for all r , and yet there is no $x^* \in X^*$ for which I is active at x^* .

For each r , consider the following linear system in x , z , and λ :

$$z - x + Mx^r - A^T\lambda = -q, \quad Az \geq b, \quad \lambda \geq 0,$$

$$\lambda_i = 0, \quad \forall i \notin I, \quad A_i z = b_i, \quad \forall i \in I,$$

$$z - x = z^r - x^r.$$

The above system is consistent since, by $I(x^r) = I$ and (2.1)–(2.2), (x^r, z^r) together with some $\lambda^r \in \mathfrak{R}^m$ is a solution of it. Then, by Lemma 2.1, it has a solution $(\hat{x}^r, \hat{z}^r, \hat{\lambda}^r)$ whose size is bounded by some constant (depending on A and M only) times the size of the right hand side. Since the right hand side of the above system is clearly bounded as $r \rightarrow \infty$, we have that $\{(\hat{x}^r, \hat{z}^r, \hat{\lambda}^r)\}$ is bounded. Moreover, every one of its limit points, say $(x^\infty, z^\infty, \lambda^\infty)$, satisfies [cf. $z^r - x^r \rightarrow 0$]

$$z^\infty - x^\infty + Mx^\infty - A^T\lambda^\infty = -q, \quad Az^\infty \geq b, \quad \lambda^\infty \geq 0,$$

$$\lambda_i^\infty = 0, \quad \forall i \notin I, \quad A_i z^\infty = b_i, \quad \forall i \in I,$$

$$z^\infty - x^\infty = 0.$$

This shows $x^\infty = [x^\infty - Mx^\infty - q]^+$ [cf. (2.1), (2.2)] and I is active at x^∞ , a contradiction of our earlier hypothesis on I . **Q.E.D.**

By using Lemma 2.2, we can now establish the main result of this section:

Theorem 2.1. There exist scalars $\epsilon > 0$ and $\tau > 0$ such that

$$d(x, X^*) \leq \tau \|x - [x - Mx - q]^+\|,$$

for all $x \in X$ with $\|x - [x - Mx - q]^+\| \leq \epsilon$.

Proof. Let ϵ be the scalar given in Lemma 2.2. Consider any $x \in X$ satisfying the hypothesis of Lemma 2.2, and let $z = [x - Mx - q]^+$. Then, by (2.1) and (2.2), there exists some $\lambda \in \mathfrak{R}^m$ satisfying, together with x and z ,

$$Mz - A^T\lambda = (I - M)(x - z) - q, \quad Az \geq b, \quad \lambda \geq 0,$$

$$\lambda_i = 0, \quad \forall i \notin I(x), \quad A_i z = b_i, \quad \forall i \in I(x).$$

By Lemma 2.2, there exists an $x^* \in X^*$ such that $I(x)$ is active at x^* , so the following linear system in (x^*, λ^*)

$$Mx^* - A^T\lambda^* = -q, \quad Ax^* \geq b, \quad \lambda^* \geq 0,$$

$$\lambda_i^* = 0, \quad \forall i \notin I(x), \quad A_i x^* = b_i, \quad \forall i \in I(x),$$

is consistent. Moreover, every solution (x^*, λ^*) of this linear system satisfies $x^* \in X^*$. Upon comparing the above two systems, we see that, by Lemma 2.1, there exists a solution (x^*, λ^*) to the second system such that

$$\|(x^*, \lambda^*) - (z, \lambda)\| \leq \theta \|(I - M)(x - z)\|,$$

where θ is some scalar constant depending on A and M only. Hence,

$$\|x^* - z\| \leq \theta \|I - M\| \|x - z\|,$$

implying that

$$\|x^* - x\| \leq (\theta \|I - M\| + 1) \|x - z\|.$$

Since $x^* \in X^*$ so $d(x, X^*) \leq \|x^* - x\|$, this then completes the proof. **Q.E.D.**

Error bounds for estimating the distance from a point to the solution set, similar to that given in Theorem 2.1, have been fairly well studied. In fact, the same bound had been demonstrated by Pang [Pan85] for the special case of an LCP where M is positive *definite*, and by Luo and Tseng [LuT90] for the special case where M is symmetric and positive semi-definite. [This bound also extends to strongly monotone variational inequality problems [Pan85] and to problems of minimizing a certain convex essentially smooth function over a polyhedral set [LuT90].]

Alternative bounds have also been proposed, by Mangasarian and Shiau [MaS86] for the special case of an LCP where M is positive semi-definite, and for strongly convex programs [MaD88]. These alternative error bounds have the advantage that they hold *globally* (even for points outside of X), whereas the bound of Theorem 2.1 holds only locally. Might the latter bound hold globally also? For general matrices M , the answer unfortunately is “no”, as shown by an example of a non-symmetric LCP furnished in [MaS86] (see Example 2.10 therein). What if M is symmetric? The answer still is “no”, as shown by the following modification of Example 2.10 in [MaS86]:

Example 2.1. Let

$$M = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad q = \begin{pmatrix} -1 \\ -2 \end{pmatrix}, \quad X = [0, \infty)^2.$$

It is easily checked that $X^* = \{ (1, 1), (0, 2) \}$. Let $x(\theta) = (\theta, 1)$, where $\theta \in [0, \infty)$. Then, as $\theta \rightarrow \infty$, we have $d(x(\theta), X^*) \rightarrow \infty$ but $\|x(\theta) - [x(\theta) - Mx(\theta) - q]^+\|$ remains bounded.

3 Linear Convergence of Matrix Splitting Algorithm for the Symmetric Case

In this section we further assume that M is symmetric, in which case the variational inequality problem (1.1) may be casted as a quadratic program of the form

$$\begin{aligned} & \text{minimize} && f(x) \\ & \text{subject to} && x \in X, \end{aligned} \tag{3.1}$$

where f is the quadratic function in \mathfrak{R}^n given by

$$f(x) = \frac{1}{2}\langle x, Mx \rangle + \langle q, x \rangle. \tag{3.2}$$

It is easily seen that the set of stationary points for (3.1) is precisely X^* [cf. (1.2)] which, by assumption, is nonempty. Notice, however, that f may not be bounded from below on X .

Let (B, C) be a *regular splitting* of M (see, e.g., [OrR70], [Kel65], [LiP87]), i.e.,

$$M = B + C, \quad B - C \text{ is positive definite.} \tag{3.3}$$

Consider the following well-known iterative algorithm for solving (3.1), based on the splitting (B, C) :

Matrix Splitting Algorithm

At the r -th iteration we are given an $x^r \in X$ ($x^0 \in X$ is chosen arbitrarily), and we compute a new iterate x^{r+1} in X satisfying

$$x^{r+1} = [x^{r+1} - Bx^{r+1} - Cx^r - q + h^r]^+, \tag{3.4}$$

where h^r is some n -vector.

The problem of finding an x^{r+1} satisfying (3.4) may be viewed as an affine variational inequality problem, whereby x^{r+1} is the vector in X which satisfies the variational inequality

$$\langle Bx^{r+1} + Cx^r + q - h^r, z - x^{r+1} \rangle \geq 0, \quad \forall z \in X. \tag{3.5}$$

In general, such an x^{r+1} need not exist, in which case the above algorithm would break down. To ensure that this does not happen, we will, following [LiP87], assume that

$$(B, C) \text{ is a } Q\text{-splitting} \tag{3.6}$$

or, equivalently, an x^{r+1} satisfying (3.5) exists for all r . [For example, (B, C) is a Q -splitting if B is positive definite (see [BeT89], [KiS80]).]

The vector h^r can be thought of as an “error” vector arising as a result of an inexact computation of x^{r+1} . [This idea of introducing an error vector is adopted from Mangasarian [Man90].] Let γ denote the smallest eigenvalue of the symmetric part of $B - C$ (which by hypothesis is positive) and let ϵ be a fixed scalar in $(0, \gamma/2]$. We will consider the following restriction on h^r governing how fast h^r tends to zero:

$$\|h^r\| \leq (\gamma/2 - \epsilon)\|x^r - x^{r+1}\|, \quad \forall r. \quad (3.7)$$

[Notice that the above restriction on h^r is practically enforceable and can in fact be used as a termination criterion for any method that computes x^{r+1} .]

The above matrix splitting algorithm was first proposed by Pang [Pan82], based on the works of Hildreth [Hil57], Mangasarian [Man77] and others. [Actually, Pang considered the somewhat simpler case of an LCP with no error vector, i.e., X is the non-negative orthant in \mathfrak{R}^n and $h^r = 0$ for all r .] This algorithm has been studied extensively (see [LiP87], [LuT89], [LuT90], [Man77], [Man90], [Pan82], [Pan84], [Pan86a] and references therein), but, owing to the possible unboundedness of the set of stationary points, its convergence had been very difficult to establish and were typically shown under restrictive assumptions on the problem (such as that the stationary point is unique). It was shown only recently that, if M is positive semi-definite (in addition to being symmetric) and f given by (3.2) is bounded from below on X , then the iterates generated by this algorithm converge to a stationary point [LuT89] with a rate of convergence that is at least linear [LuT90, Sec. 5]. In this section we show that the same linear convergence result holds for *any* symmetric M , and thus resolve the issue of convergence (and rate of convergence) for this algorithm on symmetric problems. The convergence of this algorithm for the special case of a symmetric LCP has been studied by Pang (see [Pan84, Sec. 4] and [Pan86, Sec. 2]). However, Pang did not analyze the rate of convergence of the algorithm and his convergence results require restrictive assumptions on the problem such as that the set of stationary points be finite.

The line of our analysis follows that outlined in [LuT90] (also see [LuT89b] for a similar analysis) and is based on using the error bound of Theorem 2.1 to show that, asymptotically, the objective function value, evaluated at the new iterate x^{r+1} and at some stationary point, differ by only an order of $\|x^{r+1} - x^r\|^2$ [see (3.19)]. This then enables us to show that the objective function values converge at least linearly, from which one can deduce that the iterates converge at least linearly. [This is the main motivation for considering the symmetric case, so that an objective function exists and can be used to monitor the progress of the algorithm. The algorithm itself is well-defined whether M is symmetric or not.] On the other hand, because f is not convex so the set of stationary points X^* is not necessarily convex or even connected, a new analysis, different from that in [LuT90], is needed to show the above relation.

We begin our analysis by giving, in the lemma below, a characterization of the connected

components of X^* and the behaviour of f over these connected components.

Lemma 3.1. Suppose that M is symmetric. Let C_1, C_2, \dots, C_t denote the connected components of X^* , where t is some positive integer. Then,

$$X^* = \bigcup_{i=1}^t C_i,$$

and the following hold:

- (a) Each C_i is the union of a collection of polyhedral sets.
- (b) The C_i 's are properly separated from one another, that is, $d(C_i, C_j) > 0$ for all $i \neq j$.
- (c) f given by (3.2) is constant on each C_i .

Proof. Since X is polyhedral set, we can express it as

$$X = \{ x \in \mathfrak{R}^n \mid Ax \geq b \},$$

for some $m \times n$ matrix A and some $b \in \mathfrak{R}^m$. For each $I \subseteq \{1, 2, \dots, m\}$, let

$$X_I = \{ x \mid Ax \geq b, A_I x = b_I, Mx + q = A^T \lambda \text{ for some } \lambda \in [0, \infty)^m \text{ with } \lambda_i = 0 \ \forall i \notin I \}. \quad (3.8)$$

Then, each X_I simply comprises those elements of X^* at which I is active [see (2.1) and (2.2)], so it readily follows that

$$X^* = \bigcup_{I \subseteq \{1, \dots, m\}} X_I. \quad (3.9)$$

Moreover, each X_I , if nonempty, is a polyhedral set. We claim that f is *constant* on each nonempty X_I . To see this, fix any $I \subseteq \{1, \dots, m\}$ for which X_I is nonempty. Let x and y be any two elements of X_I (possibly equal). Since $x \in X_I$ and $y \in X_I$, we have from (3.8) that $A_I(x - y) = 0$ and there exists some $\lambda \in [0, \infty)^m$ with $Mx + q = (A_I)^T \lambda_I$. Then we have from (3.2) that

$$\begin{aligned} f(x) - f(y) &= \langle My + q, x - y \rangle + \frac{1}{2} \langle x - y, M(x - y) \rangle \\ &= \langle (A_I)^T \lambda_I, x - y \rangle + \frac{1}{2} \langle x - y, M(x - y) \rangle \\ &= \langle \lambda_I, A_I(x - y) \rangle + \frac{1}{2} \langle x - y, M(x - y) \rangle \\ &= \frac{1}{2} \langle x - y, M(x - y) \rangle. \end{aligned}$$

By symmetry, we also have

$$f(y) - f(x) = \frac{1}{2} \langle x - y, M(x - y) \rangle,$$

and thus $f(x) = f(y)$. Since the above choice of x and y was arbitrary, then $f(y) = f(x)$ for all $x \in X_I, y \in X_I$.

Since each X_I is connected, it follows from (3.9) that each C_i is the union of a collection of nonempty X_I 's. Since the nonempty X_I 's are polyhedral and the C_i 's are, by definition, mutually disjoint, this then proves parts (a) and (b). Since f is constant on each X_I , this also proves part (c). **Q.E.D.**

[Part (c) of Lemma 3.1 is quite remarkable since the gradient of f needs not be constant on each C_i , as can be seen from an example.]

By using Theorem 2.1 and Lemma 3.1, we can now prove the main result of this section. (The first third of our proof follows closely that of Theorem 5.1 in [LuT90].)

Theorem 3.1. Suppose that M is symmetric and that f given by (3.2) is bounded from below on X . Let $\{x^r\}$ be iterates generated by the matrix splitting algorithm (3.3), (3.4), (3.6), (3.7). Then $\{x^r\}$ converges at least linearly to an element of X^* .

Proof. First we claim that

$$f(x^{r+1}) - f(x^r) \leq -\epsilon \|x^{r+1} - x^r\|^2, \quad \forall r. \quad (3.10)$$

To see this, fix any r . Since the variational inequality (3.5) holds, then, by plugging in x^r for z in (3.5), we obtain

$$\langle Bx^{r+1} + Cx^r + q - h^r, x^{r+1} - x^r \rangle \leq 0.$$

Also, from $M = B + C$ [cf. (3.3)] and the definition of f [cf. (3.2)] we have that

$$f(x^{r+1}) - f(x^r) = \langle Bx^{r+1} + Cx^r + q, x^{r+1} - x^r \rangle + \langle x^{r+1} - x^r, (C - B)(x^{r+1} - x^r) \rangle / 2.$$

Combining the above two relations then gives

$$\begin{aligned} f(x^{r+1}) - f(x^r) &\leq \langle h^r, x^{r+1} - x^r \rangle + \langle x^{r+1} - x^r, (C - B)(x^{r+1} - x^r) \rangle / 2 \\ &\leq \|h^r\| \|x^{r+1} - x^r\| - \gamma \|x^{r+1} - x^r\|^2 / 2 \\ &\leq -\epsilon \|x^{r+1} - x^r\|^2, \end{aligned}$$

where the last inequality follows from (3.7). Thus, (3.10) holds.

Next we claim that there exists a scalar constant $\kappa_1 > 0$ for which

$$\|x^r - [x^r - Mx^r - q]^+\| \leq \kappa_1 \|x^{r+1} - x^r\|, \quad \forall r. \quad (3.11)$$

To see this, fix any r . From (3.4) we have that

$$\begin{aligned} \|x^r - [x^r - Mx^r - q]^+\| &= \|x^r - [x^r - Mx^r - q]^+ - x^{r+1} + [x^{r+1} - Bx^{r+1} - Cx^r - q + h^r]^+\| \\ &\leq \|x^r - x^{r+1}\| + \|[x^r - Mx^r - q]^+ - [x^{r+1} - Bx^{r+1} - Cx^r - q + h^r]^+\| \\ &\leq 2\|x^r - x^{r+1}\| + \|Mx^r - Bx^{r+1} - Cx^r + h^r\| \\ &\leq 2\|x^r - x^{r+1}\| + \|B(x^r - x^{r+1})\| + \|h^r\| \\ &\leq (2 + \|B\| + \gamma/2)\|x^r - x^{r+1}\|, \end{aligned}$$

where the second inequality follows from the nonexpansive property of the projection operator $[\cdot]^+$, the third inequality follows from $M = B + C$, and the last inequality follows from (3.7). This shows that (3.11) holds with $\kappa_1 = 2 + \|B\| + \gamma/2$.

Since f is bounded from below on X , (3.10) implies

$$\|x^{r+1} - x^r\| \rightarrow 0. \quad (3.12)$$

Then we have from (3.11) that $\|x^r - [x^r - Mx^r - q]^+\| \rightarrow 0$, so, by Theorem 2.1 (and using (3.11)), there exist a scalar constant $\kappa_2 > 0$ and an index r_1 such that

$$d(x^r, X^*) \leq \kappa_2 \|x^{r+1} - x^r\|, \quad \forall r \geq r_1. \quad (3.13)$$

Thus, we also have from (3.12) that

$$d(x^r, X^*) \rightarrow 0. \quad (3.14)$$

Let C_1, C_2, \dots, C_t denote the connected components of X^* , where t is some positive integer. By Lemma 3.1 (b), the C_i 's are properly separated from one another, so (3.12) and (3.14) imply that, for all r sufficiently large, it is the same connected component of X^* which is nearest to x^r . In other words, there exists a $k \in \{1, \dots, t\}$ and a scalar $r_2 \geq r_1$ such that

$$d(x^r, X^*) = d(x^r, C_k), \quad \forall r \geq r_2. \quad (3.15)$$

Since $d(x^r, X^*) \rightarrow 0$ [cf. (3.14)], it follows from (3.15) that $d(x^r, C_k) \rightarrow 0$. By Lemma 3.1 (c), f is constant on C_k . Let us denote this constant by f^∞ .

We now show that $f(x^r) \rightarrow f^\infty$ and estimate the speed at which this convergence takes place. Fix any $r \geq r_2$. Let y^r be any element of C_k nearest to x^r , so

$$f(y^r) = f^\infty \quad (3.16)$$

and, by (3.15), $\|y^r - x^r\| = d(x^r, X^*)$. Since $r \geq r_1$ [cf. $r_2 \geq r_1$] so that (3.13) holds, then the latter implies

$$\|y^r - x^r\| \leq \kappa_2 \|x^{r+1} - x^r\|. \quad (3.17)$$

This in turn implies

$$\begin{aligned} \langle My^r + q, x^{r+1} - y^r \rangle &\leq \langle My^r + q, x^{r+1} - y^r \rangle + \langle Bx^{r+1} + Cx^r + q - h^r, y^r - x^{r+1} \rangle \\ &= \langle B(x^{r+1} - x^r) + M(x^r - y^r) - h^r, y^r - x^{r+1} \rangle \\ &\leq \left(\|B\| \|x^{r+1} - x^r\| + \|M\| \|x^r - y^r\| + \|h^r\| \right) \|y^r - x^{r+1}\| \\ &\leq \left(\|B\| \|x^{r+1} - x^r\| + \|M\| \kappa_2 \|x^{r+1} - x^r\| + \|h^r\| \right) (\kappa_2 + 1) \|x^{r+1} - x^r\| \\ &\leq (\|B\| + \|M\| \kappa_2 + \gamma/2) (\kappa_2 + 1) \|x^{r+1} - x^r\|^2, \end{aligned} \quad (3.18)$$

where the first inequality follows from (3.5) with z set to y^r , the equality follows from $C = M - B$ [cf. (3.3)], the third inequality follows from (3.17), and the last inequality follows from (3.7). For convenience, let κ_3 denote the scalar constant on the right hand side of (3.18). Then we obtain from (3.16) that

$$\begin{aligned}
f(x^{r+1}) - f^\infty &= f(x^{r+1}) - f(y^r) \\
&= \langle My^r + q, x^{r+1} - y^r \rangle + \frac{1}{2} \langle x^{r+1} - y^r, M(x^{r+1} - y^r) \rangle \\
&\leq \kappa_3 \|x^{r+1} - x^r\|^2 + \frac{1}{2} \|M\| \|x^{r+1} - y^r\|^2 \\
&\leq \left(\kappa_3 + \frac{1}{2} \|M\| (\kappa_2 + 1) \right) \|x^{r+1} - x^r\|^2,
\end{aligned} \tag{3.19}$$

where the second equality follows from (3.2), the first inequality follows from (3.18), and the last inequality follows from (3.17).

Let κ_4 denote the scalar constant on the right hand side of (3.19). Then (3.10) and (3.19) yield

$$\begin{aligned}
f(x^{r+1}) - f^\infty &\leq \kappa_4 \|x^{r+1} - x^r\|^2 \\
&\leq \frac{\kappa_4}{\epsilon} (f(x^r) - f(x^{r+1})), \quad \forall r \geq r_2.
\end{aligned}$$

Upon rearranging terms, we find that

$$\left(1 + \frac{\kappa_4}{\epsilon}\right) (f(x^{r+1}) - f^\infty) \leq \frac{\kappa_4}{\epsilon} (f(x^r) - f^\infty), \quad \forall r \geq r_2.$$

Hence $f(x^r)$ converges at least linearly to f^∞ . By (3.10), $\{x^r\}$ also converges at least linearly. Since $d(x^r, X^*) \rightarrow 0$ [cf. (3.14)], the point to which $\{x^r\}$ converges is an element of X^* . **Q.E.D.**

Notice that we can allow the matrix splitting (B, C) to vary from iteration to iteration, provided that the eigenvalues of the symmetric part of $B - C$ are bounded away from zero and that $\|B\|$ is bounded.

Also notice that because f is not convex, the point to which the iterates converge needs not be an optimal solution of (3.1). [In some cases, finding such an optimal solution may be desirable.] On the other hand, it is easily seen from Lemma 3.1 (c) and the fact that the f value of the iterates are monotonically decreasing that *local* convergence to an optimal solution holds, that is, if the initial iterate (namely x^0) is sufficiently close to the optimal solution set of (3.1), then the point to which the iterates converge is an optimal solution of (3.1).

4 Concluding Remarks

In this paper, we have shown that a certain error bound holds locally for the affine variational inequality problem. By using this bound, we are able to prove the linear convergence of matrix splitting algorithms using regular Q -splitting for the symmetric case of the problem.

There are a number of open questions raised by our work. The first question concerns whether the error bound studied here holds globally. Example 2.1 shows that it does not hold globally even when M is symmetric. But what if M in addition is positive semi-definite? A “yes” answer to this question would allow us to show *global* linear convergence for the matrix splitting algorithm of Section 3 on symmetric monotone problems. Also, our convergence result (Theorem 3.1) asserts convergence only when f given by (3.2) is bounded from below on X . If this were not the case, can something meaningful about convergence still be said? Another question concerns whether other error bounds, such as those proposed in [MaD88] and [MaS86], can be used to analyze the convergence of iterative algorithm, like is done here. Also, can the analysis of Section 3 be extended to the non-symmetric case by finding an appropriate “objective function” to work with? Or to the simpler case of a non-symmetric LCP? [It is well-known that any LCP can be converted to a quadratic program. However, except under certain conditions (see [CPV89]), the set of solutions for the former need not coincide with the set of stationary points for the latter.]

It would also be worthwhile to find other problem classes for which the error bound studied here holds. Then, we can be hopeful of proving linear convergence results for these other problems.

Acknowledgement. We thank Professor O.L. Mangasarian for pointing out to us the references on error bounds and the examples therein.

References

- [Aus76] Auslender, A., *Optimisation: Méthodes Numériques*, Masson, Paris, New York, Barcelona, Milan (1976).
- [BeT89] Bertsekas, D. P. and Tsitsiklis, J. N., *Parallel and Distributed Computation: Numerical Methods*, Prentice-Hall, Englewood Cliffs, NJ (1989).
- [CoD68] Cottle, R. W. and Dantzig, G. B., “Complementary Pivot Theory of Mathematical Programming”, *Linear Algebra and Its Applications*, 1 (1968), 103–125.
- [CGL80] Cottle, R. W. , Giannessi, F. and Lions, J.-L., *Variational Inequalities and Complementarity Problems: Theory and Applications*, Wiley, New York, NY (1980).
- [CPV89] Cottle, R. W., Pang, J.-S. and Venkateswaran, V., “Sufficient Matrices and the Linear Complementarity Problem”, *Linear Algebra and Its Applications*, 114/115 (1989), 231–249.
- [Cry71] Cryer, C. W., “The Solution of a Quadratic Programming Problem Using Systematic Overrelaxation”, *SIAM J. Contr. & Optim.*, 9 (1971), 385–392.
- [Eve71a] Eaves, B. C., “On Quadratic Programming”, *Manag. Sci.*, 17 (1971), 698–711.
- [Eve71b] Eaves, B. C., “The Linear Complementarity Problem”, *Manag. Sci.*, 17 (1971), 612–635.
- [Hil57] Hildreth, C., “A Quadratic Programming Procedure”, *Naval Res. Logist. Quart.*, 4 (1957), 79–85; see also “Erratum”, *Naval Res. Logist. Quart.*, 4 (1957), 361.
- [Hof52] Hoffman, A. J., “On Approximate Solutions of Systems of Linear Inequalities,” *J. Res. Natl. Bur. Standards*, 49 (1952), 263–265.
- [Kel65] Keller, H. B., “On the Solution of Singular and Semidefinite Linear Systems by Iteration”, *SIAM Journal of Numerical Analysis*, 2 (1965), 281–290.
- [KiS80] Kinderlehrer, D. and Stampacchia, G., *An Introduction to Variational Inequalities and Applications*, Academic Press, New York, NY (1980).
- [LiP87] Lin, Y. Y. and Pang, J.-S., “Iterative Methods for Large Convex Quadratic Programs: A Survey”, *SIAM J. Contr. & Optim.*, 25 (1987), 383–411.
- [LuT89a] Luo, Z.-Q. and Tseng, P., “On the Convergence of a Matrix-Splitting Algorithm for the Symmetric Monotone Linear Complementarity Problem,” Laboratory for Information and Decision Systems Report No. P-1884, M. I. T., Cambridge, MA (July 1989; revised May 1990); to appear in *SIAM J. Contr. & Optim.*

- [LuT89b] Luo, Z.-Q. and Tseng, P., “On the Convergence of the Coordinate Descent Method for Convex Differentiable Minimization”, Laboratory for Information and Decision Systems Report No. P-1924, M. I. T., Cambridge, MA (December 1989; revised January 1990); to appear in *J. Optim. Theory & Appl.*
- [LuT90] Luo, Z.-Q. and Tseng, P., “On the Linear Convergence of Descent Methods for Convex Essentially Smooth Minimization”, Laboratory for Information and Decision Systems Report No. P-1979, M. I. T., Cambridge, MA (June 1990).
- [Man90] Mangasarian, O. L., “Convergence of Iterates of an Inexact Matrix Splitting Algorithm for the Symmetric Monotone Linear Complementarity Problem”, Computer Sciences Technical Report #917, University of Wisconsin, Madison, WI (March, 1990).
- [MaD88] Mangasarian, O. L. and De Leone, R., “Error Bounds for Strongly Convex Programs and (Super)Linearly Convergent Iterative Schemes for the Least 2-Norm Solution of Linear Programs”, *Applied Mathematics and Optimization*, 17 (1988), 1–14.
- [MaS86] Mangasarian, O. L. and Shiau, T.-H., “Error Bounds for Monotone Linear Complementarity Problems”, *Math. Prog.*, 36 (1986), 81–89.
- [MaS87] Mangasarian, O. L. and Shiau, T.-H., “Lipschitz Continuity of Solutions of Linear Inequalities, Programs and Complementarity Problems”, *SIAM J. Contr. & Optim.* 25 (1987), 583–595.
- [Mur88] Murty, K. G., *Linear Complementarity, Linear and Nonlinear Programming*, Helderman-Verlag, Berlin (1988).
- [OrR70] Ortega, J. M. and Rheinboldt, W. C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, NY (1970).
- [Pan82] Pang, J.-S., “On the Convergence of a Basic Iterative Method for the Implicit Complementarity Problem”, *J. Optim. Theory & Appl.*, 37 (1982), 149–162.
- [Pan84] Pang, J.-S., “Necessary and Sufficient Conditions for the Convergence of Iterative Methods for the Linear Complementarity Problem”, *J. Optim. Theory & Appl.*, 42 (1984), 1–17.
- [Pan85] Pang, J.-S., “A Posteriori Error Bounds for the Linearly-Constrained Variational Inequality Problem”, School of Management Report, University of Texas, Dallas, TX (June 1985).
- [Pan86a] Pang, J.-S., “More Results on the Convergence of Iterative Methods for the Symmetric Linear Complementarity Problem”, *J. Optim. Theory & Appl.*, 49 (1986), 107–134.
- [Pan86b] Pang, J.-S., “Inexact Newton Methods for the Nonlinear Complementarity Problem”, *Math. Prog.*, 36 (1986), 54–71.

[Rob73] Robinson, S. M., "Bounds for Error in the Solution Set of a Perturbed Linear Program",
Lin. Algeb. & Appl., 6 (1973), 69–81.

[Roc70] Rockafellar, R. T., *Convex Analysis*, Princeton University Press, Princeton, NJ (1970).