July 28, 1989 LIDS-P-1884

On the Convergence of a Matrix Splitting Algorithm for the Symmetric Linear Complementarity Problem*

by

Zhi-Quan Luo†
Paul Tseng†

Abstract

We consider a matrix splitting algorithm for the linear complementarity problem where the matrix is symmetric positive semi-definite. We show that if the splitting is regular, then the iterates generated by the algorithm are well defined and converge to a solution. This result resolves in the affirmative a long standing question about the convergence of the point SOR method for solving this problem. We also extend this result to related iterative methods. As direct consequences, we obtain convergence of the methods of, respectively, Aganagic, Cottle et al., Mangasarian, Pang, and others, without making any additional assumption on the problem.

KEY WORDS: Convergence, iterative methods, convex quadratic program, linear complementarity, matrix splitting, gradient projection, SOR.

---

* This research is partially supported by the U.S. Army Research Office, contract DAAL03-86-K-0171 (Center for Intelligent Control Systems), by the Office of Naval Research, contract N00014-84-K-0519 (NR64-003), and by the National Science Foundation, grant NSF-ECS-8519058.

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

1. Introduction

Let M be an n×n symmetric positive semi-definite matrix, q be an element of $\Re^n$ (the n-dimensional Euclidean space), and c be an element of $[0,\infty]^n$. Consider the following problem of minimizing a convex quadratic function over a box:

$$\begin{array}{lll} \text{Minimize} & f(x) = \langle x, Mx\rangle/2 + \langle q, x\rangle & \\ \text{subject to} & 0 \le x_i \le c_i, & i = 1, \ldots, n. \end{array} \tag{P}$$

In our notation, all vectors are column vectors, $\langle\cdot,\cdot\rangle$ denotes the usual Euclidean inner product, and, for any vector x, x_i denotes its i-th coordinate. Notice that we allow for the possibility $c_i = \infty$ for some i. [Our results in fact hold for general box constraints (not restricted to the non-negative orthant), but for simplicity we will not consider this general case here.]

The problem (P) is an important optimization problem, with numerous applications to linear/quadratic programming [BeT89], [Man77], [MaD87], [LiP87] and to boundary value problems [CoG78], [CGS78], [DeT84]. In the special case where $c_i = \infty$ for all i, it reduces to the well-known symmetric linear complementarity problem.

We make the following standing assumption on (P):

Assumption A. f is bounded from below on the feasible set $X = [0,c_1] \times \ldots \times [0,c_n]$.

Since f is convex quadratic and X is a polyhedral set, it follows from a standard result in quadratic programming (e.g. [Eav71], [FrW57]) that (P) has a finite optimal value and the set of optimal solutions for (P), denoted by X^*, is nonempty. However, because M is only positive semi-definite, X^* may be unbounded.

From the Kuhn-Tucker conditions for (P) it is easily seen that an x belongs to X^* if and only if the orthogonal projection of $x - \nabla f(x)$ onto the feasible set X is x itself, i.e.

$$x = [\, x - (Mx + q) \,]^+, \tag{1.1}$$

where $[y]^+$ denotes the orthogonal projection of y onto X. Now, let us write M as

$$M = B + C, \tag{1.2}$$

for some $n \times n$ matrices B and C. In the terminology of numerical analysis [OrR70], such a pair (B,C) is called a splitting of M. If in addition B–C is positive definite (not necessarily symmetric), then (B,C) is called a regular splitting of M (cf. [LiP87]).

Suppose that, instead of solving the nonlinear equation (1.1) directly, we fix a solution estimate $x \in X$ and solve the following approximation to (1.1)

$$y = [\, y - (By + Cx + q) \,]^+, \tag{1.3}$$

to obtain a solution y. Then we set x to y and repeat the procedure. We formalize this procedure with the following iterative scheme: Let (B,C) be a regular splitting of M. Define a corresponding point-to-point mapping $\mathcal{A}_B : X \to X$ by (cf. (1.3))

$$\mathcal{A}_B(x) = \{\, y \in \Re^n \mid y = [\, y - (By + Cx + q) \,]^+ \,\}, \quad \forall\, x \in X. \tag{1.4}$$

We shall show in Section 2 that $\mathcal{A}_B$ is well-defined (see Lemma 2 (a)). Notice that an x satisfies (1.1) if and only if $x = \mathcal{A}_B(x)$. Consider the following algorithm for solving (P):

Matrix Splitting Algorithm: Choose an $x^0 \in X$. Generate a sequence of vectors $\{x^0, x^1, \ldots\}$ in X by the formula

$$x^{r+1} = \mathcal{A}_B(x^r), \qquad r = 0, 1, \ldots. \tag{1.5}$$

In order for the algorithm (1.4)-(1.5) to be practical, the splitting (B,C) is chosen such that Eq. (1.3) is easily solvable. We will discuss such choices in Section 5.

Consider the special case of the symmetric linear complementarity problem. The first matrix splitting method for solving this problem is the cyclic coordinate descent method of Hildreth [Hil57]. This method is simple, uses little storage, can exploit problem sparsity, and is practical for solving problems on a large scale. The method of Hildreth was subsequently extended by Cryer [Cry71] to a (point) SOR method, which in turn was extended by Cottle, Golub and Sacher [CGS78] and Cottle and Pang [CoP82] to block SOR methods. Cottle and Gohee [CoG78] further extended the Cottle-Golub-Sacher method to box constraints. An extension of Cryer's method along a different direction was proposed by Mangasarian [Man77], which is also closely related to a gradient projection algorithm of Aganagic [Aga78]. [Applications of Mangasarian's method to solving strictly convex quadratic programs and linear programs are discussed in [Man84] and [MaD87]. Parallel implementation of the method is discussed in [MaD87].] Pang [Pan82] showed that the above methods (with the possible exception of the block SOR methods) can be viewed as special cases of the matrix splitting algorithm (1.4)-(1.5). Pang then proceeded to give an extensive analysis of this algorithm [Pan82], [Pan84], [Pan86]. Yet, despite their long history and practical advantages, convergence of these iterative methods remained largely unresolved. [A summary of the current knowledge is given in [LiP87; §2-3]. See [BeT89; Chap. 3] for discussions on gradient projection algorithms.] In particular, none of the above methods has been shown to be convergent (in the sense that the iterates converge to an optimal solution) if the optimal solution is not unique. Convergence typically requires additional assumptions on the problem, all of which lead to the compactness of the solution set X^*, in which case the proof becomes rather routine (i.e., checking that each limit point is an optimal solution). In the absence of any such assumption, it was only known that the gradient of the iterates converge and that each limit point of the iterates, if it exists, is an optimal solution. The method of Cottle and Pang [CoP82] does generate a limit point, but this method includes, in addition to the standard block SOR iteration, a projection step which ensures the iterates to stay bounded and, moreover, it is applicable only to problems with a network structure. It is the aim of this paper to resolve this fundamental question of convergence by showing that the above methods are indeed convergent without making any additional assumption on the problem. In fact, we prove a more general result that, if the splitting is regular, then the corresponding matrix splitting algorithm (1.4)-(1.5) is well-defined and convergent, and the same conclusion holds for certain SOR extensions of the algorithm. [To the best of our knowledge, the only other matrix splitting algorithm that is known to be convergent in the same strong sense is one considered in Tseng [Tse89].] Our proof is of some interest in itself as it uses a number of (new) contraction properties of regular splitting and gives a detailed analysis of the trajectory of the iterates near the boundary of the feasible set X.

We remark that even for the simplest instance of the matrix splitting algorithm (1.4)-(1.5), such as the cyclic coordinate descent method, convergence is very difficult to establish when the cost function has unbounded level sets. The only other nontrivial problem having unbounded level sets in the cost function, and for which the cyclic coordinate descent method is known to be convergent in our strong sense, is a certain dual problem arising in nonlinear network optimization [BHT87].

This paper proceeds as follows: In Section 2 we derive a number of properties of the solutions of (P) and of regular splitting. In Section 3 we use these properties to prove that, when the splitting is regular, the iterates generated by the algorithm (1.4)-(1.5) converge to an optimal solution of (P). In Section 4 we propose SOR extensions of this algorithm. In Section 5 we apply the above results to a number of known methods.

In our notation, superscript T will denote transpose and $\|\cdot\|$, $\|\cdot\|_\infty$ will denote, respectively, the L_2-norm and the L_∞-norm in some Euclidean space. If A is a square matrix, $\|A\|$ will denote the matrix norm of A induced by the vector norm $\|\cdot\|$, i.e. $\|A\| = \max_{\|x\|=1} \|Ax\|$. For any k×m matrix A, we will denote by A_i the i-th row of A and, for any nonempty $I \subseteq \{1,\ldots,k\}$ and $J \subseteq \{1,\ldots,m\}$, by A_I the submatrix of A obtained by removing all rows i of A such that $i \notin I$, and by A_{IJ} the submatrix of A_I obtained by removing all columns j of A such that $j \notin J$. We will also denote by Span(A) the space spanned by the columns of A. Analogously, for any k-vector x and any nonempty subset $J \subseteq \{1,\ldots,k\}$, we denote by x_J the vector with components x_i, $i \in J$. For any finite set J, we denote by Card(J) the cardinality of J. Finally, for any $J \subseteq \{1,\ldots,n\}$, we denote by $\tilde{J}$ the complement of J with respect to $\{1,\ldots,n\}$.

2. Characterization of Optimal Solutions and Regular Splittings

In this section we derive various properties of the elements of X^* and the mapping $\mathcal{A}_B$ given by regular splittings (B,C) of M. These properties will be used in the following section to prove convergence of the algorithm (1.4)-(1.5).

The first result states that ∇f is invariant over the solution set X^*.

Lemma 1. There exists a $d^* \in \Re^n$ such that $Mx^* + q = d^*$ for all $x^* \in X^*$.

Proof. It is simple algebra to verify that, for any $x \in \Re^n$ and $y \in \Re^n$,

$$f(y) - f(x) \quad = \quad \|M^{1/2}(y-x)\|^2/2 + \langle y-x, Mx + q\rangle.$$

Hence if both x and y belong to X^*, so that $f(y) = f(x)$ and $\langle y-x, Mx + q\rangle = 0$, then $M^{1/2}(y-x) = 0$, or equivalently, $My = Mx$. **Q.E.D.**

The next result shows that, if (B,C) is a regular splitting of M, then $\mathcal{A}_B$ is a well-defined point-to-point mapping and possesses a certain descent property.

Lemma 2. Let (B,C) be a regular splitting of M. Then the following hold:

(a) $\mathcal{A}_B: X \to X$ is a well-defined point-to-point mapping.
(b) For any $x \in X$,

$$f(y) - f(x) \quad \leq \quad \langle y - x, (C - B)(y - x)\rangle/2,$$

where $y = \mathcal{A}_B(x)$.

Proof. We first prove part (a). Since $B - C$ is positive definite, it follows from $2B = M + (B - C)$ (cf. $M = B + C$) and the positive semi-definite property of M that B is positive definite. Hence, by a well-known result from variational inequality [BeT89, pp. 271], [KiS80, §2], we have that, for any $x \in X$, the nonlinear equation

$$y \quad = \quad [\, y - (By + Cx + q) \,]^+,$$

has a unique solution y. This proves part (a).

Now we prove part (b). It can be seen by using $M = B + C$ that, for any x and y in $\Re^n$,

$$f(y) - f(x) \quad = \langle y - x, By + Cx + q \rangle + \langle y - x, (C - B)(y - x) \rangle / 2.$$

On the other hand, we see from (1.4) that $y = \mathcal{A}_B(x)$ if and only if $y \in X$ and

$$\begin{aligned} B_i y + C_i x + q_i &> 0 \quad \Rightarrow \quad y_i = 0, \\ B_i y + C_i x + q_i &< 0 \quad \Rightarrow \quad y_i = c_i. \end{aligned}$$

Hence if in addition $x \in X$ (so that $0 \le x \le c$), then $\langle y - x, By + Cx + q \rangle \le 0$. **Q.E.D.**

[The results of Lemma 2 are quite well-known (e.g. [LiP87]). The proof of part (b) is based on one given in Lemma 4.1 of [Pan84].]

It can be seen that if the box constraints $x \in X$ are removed, then, for any splitting (B,C) of M, $y = \mathcal{A}_B(x)$ if and only if $By + Cx + q = 0$, or equivalently (assuming that B is invertible)

$$y \quad = -B^{-1}(Cx + q) \quad = (I - B^{-1}M)x - B^{-1}q.$$

A key property of regular splitting is that the corresponding iteration matrix $I - B^{-1}M$ has its spectral radius strictly less than one over a certain subspace:

Lemma 3. Let Q be an $m \times m$ symmetric positive semi-definite matrix and let (B,C) be a regular splitting of Q. Then B is positive definite and the following hold:

(a) The spectral radius of $I - QB^{-1}$ restricted to Span(Q) is strictly less than 1, i.e. there exist $\rho \in (0,1)$ and $\tau > 0$ such that

$$\|(I - QB^{-1})^k z\| \quad \le \tau (\rho)^k \|z\|, \quad \forall\, k \ge 1, \qquad \forall\, z \in \mathrm{Span}(Q).$$

(b) The spectral radius of $I - B^{-1}Q$ is less than or equal to 1, i.e. there exists $\Delta \geq 1$ such that

$$\|(I - B^{-1}Q)^k z\| \leq \Delta\|z\|, \qquad \forall\, k \geq 1, \qquad \forall\, z \in \Re^n.$$

Proof. Since $B - C$ is positive definite, it follows from $2B = Q + (B - C)$ (cf. $Q = B + C$) and the positive semi-definite property of Q that B is positive definite. For any $y^0 \in \Re^n$, consider the sequence of vectors $\{y^0, y^1, \ldots\}$ given by the recursion

$$By^{r+1} + Cy^r = 0. \tag{2.1}$$

Since B is positive definite, this sequence is well defined. By using (2.1) and an argument analogous to the proof of Lemma 2, we obtain that

$$g(y^{r+1}) = g(y^r) + \langle y^{r+1} - y^r, (C - B)(y^{r+1} - y^r)\rangle/2,$$

where we define $g:\Re^n \to \Re$ to be the function $g(y) = \langle y, Qy\rangle/2$. Since Q is positive semi-definite, $g(y)$ is non-negative for all y. Hence the above equation (and using the positive definite property of $B - C$) implies $y^{r+1} - y^r \to 0$. Let $d^r = Qy^r$. Then from (2.1) and $C = Q - B$ we have

$$d^r = B(y^r - y^{r+1}), \qquad \forall\, r. \tag{2.2}$$

Therefore $d^r \to 0$ and (multiplying both sides of (2.2) by QB^{-1})

$$d^{r+1} = (I - QB^{-1})d^r, \qquad \forall\, r.$$

This in turn implies that, for any $d^0 \in \mathrm{Span}(Q)$,

$$(I - QB^{-1})^r d^0 \to 0 \qquad \text{as } r \to \infty,$$

and part (a) then follows from the following fact proven in Appendix A:

Fact 1. For any m×m matrix A and any linear subspace $\mathcal{V}$ of $\Re^m$, if $(A)^k z \to 0$ as $k \to \infty$ for all $z \in \mathcal{V}$, then there exist $\rho \in (0,1)$ and $\tau > 0$ such that $\|(A)^k z\| \le \tau (\rho)^k \|z\|$ for all $k \ge 1$ and all $z \in \mathcal{V}$.

Since $\{d^r\}$ converges to zero at a geometric rate. By (2.2) and the fact that B is invertible, $\{y^{r+1} - y^r\}$ also converges to zero at a geometric rate. Hence $\{y^r\}$ satisfies the Cauchy criterion for convergence and therefore is convergent. Since (cf.(2.2)) $y^{r+1} = y^r - B^{-1}d^r$ for all r, it follows that

$$y^{r+1} \quad = \quad (I - B^{-1}Q)y^r, \qquad \forall\ r.$$

Since $\{y^r\}$ is convergent for any $y^0 \in \Re^n$, this implies that all eigenvalues of $I - B^{-1}Q$ lie either inside or on the unit circle. This proves part (b). **Q.E.D.**

Remark 1. Since $I - B^{-1}Q = B^{-1}(I - QB^{-1})B$, the two matrices $I - B^{-1}Q$ and $I - QB^{-1}$ are similar and therefore have identical eigenvalues. Hence part (b) of Lemma 3 implies that the eigenvalues of $I - QB^{-1}$ are also within the unit circle.

Remark 2. The matrix $I - \omega QB^{-1}$ also has the contraction properties described in part (a) of Lemma 3, provided that $0 < \omega \le 1$ (see Appendix A). This fact will be used in Section 4 where we introduce under/over-relaxation to the mapping $\mathcal{A}_B$.

As an immediate consequence of Lemma 3, we have that the coordinate descent method for solving the unconstrained version of (P) (i.e. find an x satisfying $Mx + q = 0$)

$$y^{r+1} \quad = \quad (I - (E + L)^{-1}M)y^r - (E + L)^{-1}q,$$

where E and L denote respectively the diagonal and the strictly lower triangular part of M, converges at a geometric rate (assuming that M has positive diagonal entries and that the problem has a solution). This result improves on one given in [Lue73; pp. 159] for the special case where M is symmetric positive definite.

Lemma 3 in turn implies the following facts:

Lemma 4. Let (B,C) be a regular splitting of M. Then the following hold:

(a) For any nonempty $J \subseteq \{1,\ldots,n\}$, there exist $\rho_J \in (0,1)$ and $\tau_J > 0$ such that

$$\|(I - M_{JJ}(B_{JJ})^{-1})^k z\| \le \tau_J (\rho_J)^k \|z\|, \qquad \forall\, k \ge 1, \qquad \forall\, z \in \mathrm{Span}(M_{JJ}).$$

(b) There exists a $\Delta \ge 1$ such that, for any nonempty $J \subseteq \{1,\ldots,n\}$,

$$\|(I - (B_{JJ})^{-1} M_{JJ})^k z\| \le \Delta \|z\|, \qquad \forall\, k \ge 1, \qquad \forall\, z.$$

Proof. Since $B - C$ is positive definite, $B_{JJ} - C_{JJ}$ is positive definite. Parts (a) and (b) then follow immediately from, respectively, parts (a) and (b) of Lemma 3. **Q.E.D.**

Let I^* denote the set of indices i for which the i-th coordinate of any element of X^* is not necessarilly fixed at either 0 or its upper bound c_i, i.e.,

$$I^* = \{\, i \mid d_i^* = 0 \,\}.$$

Then, for each $x^* \in X^*$, we have $M_{I^*}x^* + q_{I^*} = 0$ (cf. Lemma 1), so that

$$q_{I^*} \in \mathrm{Span}(M_{I^*}).$$

The submatrix of M indexed by I^* has a number of interesting properties which we show below:

Lemma 5. For any $J \subseteq I^*$, $\mathrm{Span}(M_{J\bar{J}}) \subseteq \mathrm{Span}(M_{JJ})$ and $q_J \in \mathrm{Span}(M_{JJ})$.

Proof. For each $i \notin J$, consider the restricted problem

$$\begin{array}{lll} \text{Minimize} & \langle x, Mx\rangle & \\ \text{subject to} & x_i = 1, & x_j = 0,\ \forall\, j \notin J \text{ such that } j \neq i. \end{array}$$

This problem is clearly feasible and, since M is positive semi-definite, its optimal value is finite. It follows that this problem has an optimal solution, and from its Kuhn-Tucker conditions we have $M_{Ji} \in \mathrm{Span}(M_{JJ})$. Since the choice of $i \notin J$ was arbitrary, we have $\mathrm{Span}(M_{J\bar{J}}) \subseteq \mathrm{Span}(M_{JJ})$. This, together with the fact $q_J \in \mathrm{Span}(M_J)$ (since $q_{I^*} \in \mathrm{Span}(M_{I^*})$ and $J \subseteq I^*$), implies $q_J \in \mathrm{Span}(M_{JJ})$. **Q.E.D.**

3. A General Convergence Theorem

Let $\{x^r\}$ be a sequence of iterates generated by the algorithm (1.4)-(1.5), i.e.,

$$x^{r+1} = \mathcal{A}_B(x^r), \qquad r = 0, 1, \ldots,$$

where (B,C) is some regular splitting of M. By Lemma 2 (a), $\{x^r\}$ is well-defined. We will show that $\{x^r\}$ converges to an element of X^*.

To motivate our proof, note from Lemma 2 (b) that, for all r,

$$\begin{aligned} f(x^{r+1}) &\le f(x^r) - \langle x^{r+1} - x^r, (B - C)(x^{r+1} - x^r)\rangle/2 \\ &\le f(x^r) - \gamma\|x^{r+1} - x^r\|^2/2, \end{aligned} \tag{3.1}$$

where γ denotes the modulus of the smallest eigenvalue of $B - C$. Upon summing this inequality over all r and using the fact that $f(x^r)$ is bounded from below for all r (cf. Assumption A), we obtain

$$\sum_{r=0}^{\infty} \|x^{r+1} - x^r\|^2 < \infty. \tag{3.2}$$

Hence $x^{r+1} - x^r \to 0$, which together with

$$x^{r+1} = [\, x^{r+1} - (Bx^{r+1} + Cx^r + q)\,]^+ \tag{3.3}$$

(cf. $x^{r+1} = \mathcal{A}_B(x^r)$), the Lipschitz continuity of $[\cdot]^+$, and the fact $B + C = M$, establishes the following:

Lemma 6.

(a) $x^{r+1} - x^r \to 0$.

(b) $x^r - [\, x^r - Mx^r - q\,]^+ \to 0$.

Hence any limit point x^∞ of $\{x^r\}$ satisfies $x^\infty = [\, x^\infty - Mx^\infty - q\,]^+$ and is therefore in X^*. This result is quite well-known (e.g., [Pan86], [LiP87]) and, as we just saw, is relatively easy to prove. The difficulty lies in showing that $\{x^r\}$ indeed has a limit point. This is a highly nontrivial task to which the remainder of this section will be devoted.

Remark 3. Eq. (3.2) gives an estimate of the rate at which $x^{r+1} - x^r \to 0$, but technically speaking, it is not enough for us to claim the convergence of $\{x^r\}$ since it does not prevent $x^{r+1} - x^r$ to decrease like $1/r$, in which case $x^r \to \infty$. Intuitively, it seems unlikely that such a sequence of iterates can be generated by the matrix splitting algorithm, but to show this rigorously is very difficult, as indicated by the complexity of the proof given below.

For each $x \in \Re^n$, let $\phi(x)$ denote the distance from x to X^*, i.e.

$$\phi(x) \quad = \min_{x^* \in X^*} \|x - x^*\|.$$

The next lemma, which shows that $\{Mx^r\}$ converges and that $\{x^r\}$ approaches X^*, follows as a direct consequence of Lemma 6:

Lemma 7.

(a) $Mx^r + q \to d^*$.

(b) $\phi(x^r) \to 0$.

Proof. We first prove part (a). [This result is known when $c_i = \infty$ for all i [Pan86; Theorem 3.1]. The following proof is simpler than that given in [Pan86] and holds for the general case where some of the c_i's may be finite.] Since M is symmetric positive semi-definite, $M^{1/2}$ exists. Let x^* be an element of X^*. By a direct calculation we have that, for all $r \geq 0$,

$$\begin{aligned} f(x^r) - f(x^*) &= \|M^{1/2}(x^r - x^*)\|^2/2 + \langle x^r - x^*, Mx^* + q\rangle \qquad (3.4)\\ &\geq \|M^{1/2}(x^r - x^*)\|^2/2, \end{aligned}$$

where the inequality follows from the optimality conditions for x^*. Since $f(x^r)$ is monotonically decreasing with r (cf. (3.1)) we see that $\{M^{1/2}x^r\}$ is bounded. Let z^∞ be any limit point of $\{M^{1/2}x^r\}$, let $d^\infty = M^{1/2}z^\infty + q$, and let $\{M^{1/2}x^r\}_{r\in R}$ be any subsequence of $\{M^{1/2}x^r\}$ converging to z^∞. Then $\{Mx^r + q\}_{r\in R} \to d^\infty$, so that Lemma 6 (b) yields

$$\begin{aligned} \{x_i^r\}_{r\in R} &\to 0, && \text{if } d_i^\infty > 0, \\ \{x_i^r\}_{r\in R} &\to c_i, && \text{if } d_i^\infty < 0. \end{aligned}$$

Since

$$\begin{aligned} \langle x^* - x^r, Mx^r + q\rangle &= \langle x^* - x^r, Mx^r - M^{1/2}z^\infty\rangle + \langle x^* - x^r, d^\infty\rangle \\ &= \langle M^{1/2}x^* - M^{1/2}x^r, M^{1/2}x^r - z^\infty\rangle \\ &+ \sum_{d_i^\infty>0} (x_i^* - x_i^r)d_i^\infty + \sum_{d_i^\infty<0} (x_i^* - x_i^r)d_i^\infty, \qquad \forall\ r, \end{aligned}$$

we obtain, upon passing into the limit as $r \to \infty$, $r\in R$, that

$$\begin{aligned} \lim_{r\to\infty; r\in R} \inf\{\langle x^* - x^r, Mx^r + q\rangle\} &\geq \sum_{d_i^\infty>0} x_i^* d_i^\infty + \sum_{d_i^\infty<0} (x_i^* - c_i)d_i^\infty, \\ &\geq 0. \qquad (3.5) \end{aligned}$$

where the second inequality follows from the fact $0 \leq x^* \leq c$. On the other hand, we have from (3.4) that

$$f(x^*) - f(x^r) = \|M^{1/2}x^r - M^{1/2}x^*\|^2/2 + \langle x^* - x^r, Mx^r + q\rangle, \qquad \forall\ r.$$

Since $0 \geq f(x^*) - f(x^r)$ for all r, by passing into the limit as $r \to \infty$, $r\in R$, and using (3.5), we obtain that

$$0 \geq \|z^\infty - M^{1/2}x^*\|^2/2.$$

Hence $z^\infty = M^{1/2}x^*$, so that $\{Mx^r\}_{r\in R} \to Mx^*$. Since the choice of the limit point z^∞ was arbitrary, this holds for all convergent subsequences of $\{Mx^r\}$ and part (a) is proven.

We now prove part (b). Since $Mx^r + q \to Mx^* + q = d^*$ (cf. part (a) and Lemma 1), we have from Lemma 6 (b) that

$$x_i^r \quad \to 0, \qquad \text{if } d_i^* > 0, \tag{3.6a}$$
$$x_i^r \quad \to c_i, \qquad \text{if } d_i^* < 0. \tag{3.6b}$$

Let $d^r = Mx^r + q$. Then we also have

$$d^r \quad \to d^*. \tag{3.6c}$$

Now, each x^r is a solution of the linear system

$$Mx + q \;=\; d^r, \quad 0 \le x \le c, \quad x_i = x_i^r \text{ if } d_i^* > 0, \quad x_i = x_i^r \text{ if } d_i^* < 0,$$

while it can be seen from Lemma 1 and (1.1) that X^* is the set of solutions of the linear system

$$My + q \;=\; d^*, \quad 0 \le y \le c, \quad y_i = 0 \text{ if } d_i^* > 0, \quad y_i = c_i \text{ if } d_i^* < 0.$$

Hence, by a well-known Lipschitz continuity property of the solutions of linear systems ([CGST86], [MaS87], [Rob73]) there exists, for each r, a $y^r \in X^*$ satisfying

$$\|x^r - y^r\| \le \theta(\|d^r - d^*\| + \sum_{d_i^*>0} x_i^r + \sum_{d_i^*<0} (c_i - x_i^r)),$$

where θ is some constant that depends on M only. Since $y^r \in X^*$, this in turn implies $\phi(x^r) \le \theta(\|d^r - d^*\| + \sum_{d_i^*>0} x_i^r + \sum_{d_i^*<0} (c_i - x_i^r))$. By (3.6a)-(3.6c), $\phi(x^r) \to 0$.

Q.E.D.

Notice that the proof of part (a) also shows that $f(x^r) \to f(x^*)$.

Under an additional regularity assumption on (P), we can show by using Lemmas 6 and 7 that $\{x^r\}$ converges at a <u>geometric</u> rate.

Proposition 1. Suppose that (P) satisfies the <u>strong complementarity</u> condition, i.e.,

$$d_i^* \quad = 0 \qquad \text{if and only if} \qquad 0 < x_i^* < c_i \ \forall\, x^* \in X^*.$$

Then each sequence of iterates $\{x^r\}$ generated by the matrix splitting algorithm (1.4)-(1.5) converges to an element of X^* at a geometric rate.

Proof. Since $Mx^r + q \to d^*$ (cf. Lemma 7 (a)) and $x^r - x^{r-1} \to 0$ (cf. Lemma 6 (a)), we have that $Bx^r + Cx^{r-1} + q = B(x^r - x^{r-1}) + Mx^r + q \to d^*$. This together with the fact $x^r = [\, x^r - (Bx^r + Cx^{r-1} + q)\,]^+$ (cf. (3.3)) implies that there exists an integer s such that, for all $r \geq s$,

$$x_i^r \ = \ 0, \ \forall\, i \text{ such that } d_i^* > 0, \qquad x_i^r \ = \ c_i, \ \forall\, i \text{ such that } d_i^* < 0. \tag{3.7}$$

On the other hand, for each $i \in I^*$ (i.e. $d_i^* = 0$), we have from the strong complementarity condition that $0 < x_i^* < c_i$ for all $x^* \in X^*$. Since X^* is a closed polyhedral set, this implies that there exists an $\varepsilon > 0$ such that $\varepsilon \leq x_i^* \leq c_i - \varepsilon$ for all $x^* \in X^*$ and all $i \in I^*$. Since the distance between x^r and X^* tends to 0 (cf. Lemma 7 (b)), this in turn implies that there exists an integer $t \geq s$ such that $0 < x_i^r < c_i$ for all $i \in I^*$ and all $r \geq t$, so that (since $x^r = [\, x^r - (Bx^r + Cx^{r-1} + q)\,]^+$ for all r)

$$B_{I^*}x^r + C_{I^*}x^{r-1} + q_{I^*} \ = \ 0, \qquad \forall\, r \geq t.$$

Consider any $x^* \in X^*$. Then, by Lemma 1, $Mx^* + q = d^*$, which together with the fact $M = B + C$ and $d_i^* = 0$ for all $i \in I^*$ yields $B_{I^*}x^* + C_{I^*}x^* + q_{I^*} = 0$. Subtracting this equation from the above equation yields $B_{I^*}(x^r - x^*) + C_{I^*}(x^{r-1} - x^*) = 0$ or, equivalently, $B_{I^*}z^r = -C_{I^*}z^{r-1}$ for all $r \geq t$, where we let $z^r = x^r - x^*$. From (3.7) and the fact $x_i^* = 0$ ($x_i^* = c_i$) if $d_i^* > 0$ ($d_i^* < 0$) we also have that $z_i^r = 0$ for all $i \notin I^*$ and all $r \geq t$. Hence

$$B_{I^*I^*}z_{I^*}^r \ = \ -\,C_{I^*I^*}z_{I^*}^{r-1}, \qquad \forall\, r \geq t,$$

or, equivalently (using $C_{I^*I^*} = M_{I^*I^*} - B_{I^*I^*}$ and multiplying both sides by $(B_{I^*I^*})^{-1}$),

$$z_{I^*}^r \ = \ (I - (B_{I^*I^*})^{-1}M_{I^*I^*})z_{I^*}^{r-1}, \qquad \forall\, r \geq t.$$

Upon multiplying both sides by $M_{I^*I^*}$, we obtain

$$M_{I^*I^*}z^r_{I^*} \;=\; (I - M_{I^*I^*}(B_{I^*I^*})^{-1})\, M_{I^*I^*}z^{r-1}_{I^*}, \qquad \forall\, r \geq t,$$

so that, by Lemma 4 (a), $\{M_{I^*I^*}z^r_{I^*}\}$ converges to zero at a geometric rate. Since $z^r_{I^*} - z^{r-1}_{I^*} = -(B_{I^*I^*})^{-1}M_{I^*I^*}\, z^{r-1}_{I^*}$ for all $r \geq t$ (cf. equation above), this shows that $\{z^r_{I^*}\}$ is a Cauchy sequence and that it converges at a geometric rate. Since $x^r = z^r + x^*$ for all r and $z_i^r = 0$ for all $i \notin I^*$ and all $r \geq t$, this in turn shows that $\{x^r\}$ converges at a geometric rate and, by Lemma 7 (b), the point to which $\{x^r\}$ converges is an element of X^*. **Q.E.D.**

Now let us map out the directions for the most intricate part of our proof. We know that $\{x^r\}$ approaches X^*, but we do not know if it is bounded. From the proof of Proposition 1 we see that if, for every i, either (i) $\{x_i^r\}$ stays fixed to one of the two boundary points 0 and c_i or (ii) $\{x_i^r\}$ stays strictly between 0 and c_i, then $\{x^r\}$ converges at a geometric rate. Hence the difficulty lies with those coordinates of x^r that bounce around the boundary of the feasible set, possibly causing one of the remaining coordinates to sail off to infinity. To resolve this difficulty, we will show that these coordinates perturb the movement of the remaining coordinates only (additively) by their own maximum deviation from the boundary. This fact, shown in Lemma 8 below, is based on the contraction property of the algorithmic mapping for the unconstrained case (cf. Lemma 4) and Lemma 5. Then those coordinates of x^r that start out far from the boundary will stay far from the boundary (cf. geometric convergence for the unconstrained case), unless one of the remaining coordinates also moves far from the boundary, so that, eventually, each coordinate of x^r either stays close to the boundary or stays far from the boundary. Those coordinates that stay close to the boundary are clearly bounded; those coordinates that stay far from the boundary are also bounded because perturbation by the other coordinates is bounded and, within themselves, the convergence is geometric (since they are effectively unconstrained). We now proceed with the actual proof.

Let

$$\beta \;=\; \max_{J \subseteq I^*} \sqrt{\mathrm{Card}(\tilde{J})}\; \Big\{ \big(\tau_J \|(B_{JJ})^{-1}\|\, \|M_{JJ}\|/(1-\rho_J) + \Delta + 1\big)\, \|(B_{JJ})^{-1}B_{J\tilde{J}}\| + \tau_J \|(B_{JJ})^{-1}\|\, \|M_{J\tilde{J}}\|/(1-\rho_J) \Big\}.$$

The following lemma, based on Lemmas 1, 4 and 5, shows that those coordinates of x^r that stay away from zero are influenced by the remaining coordinates only through the distance, scaled by β, of these remaining coordinates from zero. This result allows us to separate the effect of these two sets of coordinates on each other.

Lemma 8 (Coordinate Separation). Consider any $J \subseteq I^*$. If for some two integers $s \geq t \geq 0$ we have $0 < x_i^r < c_i$ for all $i \in J$ and all $r = t+1, t+2, \ldots, s$, then, for any $x^* \in X^*$,

$$\|x_J^s - x_J^*\| \;\leq\; \Delta\|x_J^t - x_J^*\| + \beta \max_{r\in\{t,\ldots,s\}} \|x_{\bar{J}}^r - x_{\bar{J}}^*\|_\infty .$$

Proof. The claim clearly holds if $s = t$ (since $\Delta \geq 1$). Suppose that $s > t$. Since $x_i^r > 0$ for all $i \in J$ and all $r = t+1,\ldots,s$, it follows from the fact $x^{r+1} = [x^{r+1} - (Bx^{r+1} + Cx^r + q)]^+$ for all r (cf. (3.3)) that

$$B_J x^{r+1} + C_J x^r + q_J \quad = \quad 0, \qquad\qquad r = t,\ldots,s-1,$$

or equivalently (using $M_J = B_J + C_J$),

$$B_J(x^{r+1} - x^r) + M_J x^r + q_J \quad = \quad 0, \qquad\qquad r = t,\ldots,s-1.$$

Since $J \subseteq I^*$, we also have (using Lemma 1 and the definition of I^*)

$$M_J x^* + q_J \quad = \quad 0.$$

Combining the above two equalities and multiplying by $(B_{JJ})^{-1}$ yields

$$(B_{JJ})^{-1} B_J(x^{r+1} - x^r) + (B_{JJ})^{-1} M_J(x^r - x^*) = 0, \qquad\qquad r = t,\ldots,s-1.$$

This in turn implies, after some rearrangement of terms, that

$$\begin{aligned} x_J^{r+1} - x_J^* \;=\; & (I - (B_{JJ})^{-1} M_{JJ})(x_J^r - x_J^*) - (B_{JJ})^{-1} M_{J\bar{J}}(x_{\bar{J}}^r - x_{\bar{J}}^*) \\ & - (B_{JJ})^{-1} B_{J\bar{J}}(x_{\bar{J}}^{r+1} - x_{\bar{J}}^r), \qquad r = t,\ldots,s-1. \end{aligned}$$

By successively applying the above recursion for $r = t, \ldots, s-1$, we obtain

$$\begin{aligned} x_J^s - x_J^* &= (G)^h(x_J^t - x_J^*) - \sum_{k=0}^{h-1} (G)^{h-k-1}(B_{JJ})^{-1}M_{J\tilde{J}}(x_{\tilde{J}}^{t+k} - x_{\tilde{J}}^*) \\ &\quad - \sum_{k=0}^{h-1} (G)^{h-k-1}(B_{JJ})^{-1}B_{J\tilde{J}}(x_{\tilde{J}}^{t+k+1} - x_{\tilde{J}}^{t+k}), \end{aligned} \tag{3.8}$$

where we denote $G = I - (B_{JJ})^{-1}M_{JJ}$ and $h = s - t$. Now we estimate the last sum in Eq. (3.8). Let $y^k = (B_{JJ})^{-1}B_{J\tilde{J}}(x_{\tilde{J}}^{t+k} - x_{\tilde{J}}^*)$. Then the last sum in Eq. (3.8) can be rewritten as $\sum_{k=0}^{h-1} (G)^{h-k-1}(y^{k+1} - y^k)$. By rearranging the terms within the summation sign, we obtain an alternative form for this sum:

$$\begin{aligned} \sum_{k=0}^{h-1} (G)^{h-k-1}(y^{k+1} - y^k) &= \sum_{k=0}^{h-1} (G)^{h-k-1}y^{k+1} - \sum_{k=0}^{h-1} (G)^{h-k-1}y^k \\ &= \sum_{k=1}^{h-1} (G)^{h-k-1}(G)y^k + y^h - (G)^{h-1}y^0 - \sum_{k=1}^{h-1} (G)^{h-k-1}y^k \\ &= \sum_{k=1}^{h-1} (G)^{h-k-1}(G - I)y^k + y^h - (G)^{h-1}y^0. \end{aligned}$$

Since $G - I = -(B_{JJ})^{-1}M_{JJ}$, this together with (3.8) implies that

$$\begin{aligned} x_J^s - x_J^* &= (G)^h(x_J^t - x_J^*) - \sum_{k=0}^{h-1} (G)^{h-k-1}(B_{JJ})^{-1}M_{J\tilde{J}}(x_{\tilde{J}}^{t+k} - x_{\tilde{J}}^*) \\ &\quad + \sum_{k=1}^{h-1} (G)^{h-k-1}(B_{JJ})^{-1}M_{JJ}y^k + y^h - (G)^{h-1}y^0. \end{aligned}$$

Let $H = I - M_{JJ}(B_{JJ})^{-1}$. Then $G = (B_{JJ})^{-1}HB_{JJ}$, so that $(G)^{h-k-1} = (B_{JJ})^{-1}(H)^{h-k-1}B_{JJ}$ for all k. This together with the above equation yields

$$\begin{aligned} x_J^s - x_J^* &= (G)^h(x_J^t - x_J^*) - \sum_{k=0}^{h-1} (B_{JJ})^{-1}(H)^{h-k-1}M_{J\tilde{J}}(x_{\tilde{J}}^{t+k} - x_{\tilde{J}}^*) \\ &\quad + \sum_{k=1}^{h-1} (B_{JJ})^{-1}(H)^{h-k-1}M_{JJ}y^k + y^h - (G)^{h-1}y^0. \end{aligned}$$

Also, since $J \subseteq I^*$, we have from Lemma 4 that $\|(H)^{h-k-1}z\| \le \tau_J (\rho_J)^{h-k-1}\|z\|$ for any $z \in \mathrm{Span}(M_{JJ})$ and $\|(G)^h z\| \le \Delta\|z\|$ for any z. This, together with the above equation and the fact $\mathrm{Span}(M_{J\bar{J}}) \subseteq \mathrm{Span}(M_{JJ})$ (cf. Lemma 5), implies

$$\begin{aligned}
\|x_J^s - x_J^*\| &\le \|(G)^h(x_J^t - x_J^*)\| + \sum_{k=0}^{h-1} \|(B_{JJ})^{-1}\| \, \|(H)^{h-k-1} M_{J\bar{J}}(x_{\bar{J}}^{t+k} - x_{\bar{J}}^*)\| \\
&\qquad + \sum_{k=1}^{h-1} \|(B_{JJ})^{-1}\| \, \|(H)^{h-k-1} M_{JJ} y^k\| + \|y^h\| + \|(G)^{h-1} y^0\| \\
&\le \Delta\|x_J^t - x_J^*\| + \sum_{k=0}^{h-1} \|(B_{JJ})^{-1}\| \, \tau_J (\rho_J)^{h-k-1} \|M_{J\bar{J}}(x_{\bar{J}}^{t+k} - x_{\bar{J}}^*)\| \\
&\qquad + \sum_{k=1}^{h-1} \|(B_{JJ})^{-1}\| \, \tau_J (\rho_J)^{h-k-1} \|M_{JJ} y^k\| + \|y^h\| + \Delta\|y^0\| \\
&\le \Delta\|x_J^t - x_J^*\| + \tau_J \|(B_{JJ})^{-1}\| \, \|M_{J\bar{J}}\| \sum_{k=0}^{h-1} (\rho_J)^{h-k-1} \max_{r \in \{t,\dots,s-1\}} \|x_{\bar{J}}^r - x_{\bar{J}}^*\| \\
&\quad + \tau_J \|(B_{JJ})^{-1}\| \, \|M_{JJ}\| \sum_{k=1}^{h-1} (\rho_J)^{h-k-1} \max_{k \in \{1,\dots,h-1\}} \|y^k\| + \|y^h\| + \Delta\|y^0\|. \\
&\le \Delta\|x_J^t - x_J^*\| + \tau_J \|(B_{JJ})^{-1}\| \, \|M_{J\bar{J}}\|(1 - \rho_J)^{-1} \max_{r \in \{t,\dots,s-1\}} \|x_{\bar{J}}^r - x_{\bar{J}}^*\| \\
&\quad + \tau_J \|(B_{JJ})^{-1}\| \, \|M_{JJ}\|(1 - \rho_J)^{-1} \max_{k \in \{1,\dots,h-1\}} \|y^k\| + \|y^h\| + \Delta\|y^0\|.
\end{aligned}$$

Since $y^k = (B_{JJ})^{-1}B_{J\bar{J}}(x_{\bar{J}}^{t+k} - x_{\bar{J}}^*)$, we also have $\|y^k\| \le \|(B_{JJ})^{-1}B_{J\bar{J}}\| \, \|x_{\bar{J}}^{t+k} - x_{\bar{J}}^*\|$ and the lemma is proven. **Q.E.D.**

By using Lemmas 6, 7 and 8, we can now prove our main result that $\{x^r\}$ converges. The basic idea of the proof is to show that those coordinates of x^r that are bounded sufficiently far away from zero are essentially unaffected by the rest. This then allows us to treat these coordinates as if they are unconstrained in sign and by using the contraction property of $\mathcal{A}_B$ on them, we conclude convergence for these coordinates. We define the following scalars for the subsequent analysis:

$$\begin{aligned}
\sigma_0 &= 1, \\
\sigma_k &= \Delta + 3 + \beta + (\beta + 1)\sigma_{k-1}, \qquad k = 1, 2, \dots, n.
\end{aligned}$$

[Notice that $\sigma_k \ge 1$ for all k and is monotonically increasing with k.]

Lemma 9. For any $\delta > 0$, there exists an $x^* \in X^*$ and an $\hat{r} > 0$ such that

$$\|x^r - x^*\|_\infty \le \sigma_n\delta + \delta, \qquad \forall\, r \ge \hat{r}. \tag{3.9}$$

Proof. To simplify the proof, we will assume that $c_i = \infty$ for all i. The case where $c_i < \infty$ for some i can be handled by making minor modifications to the proof. Furthermore, by using Lemmas 6 (a) and 7 (b), we will without loss of generality assume that

$$\phi(x^r) \le \delta, \qquad \forall\, r, \tag{3.10a}$$
$$\|x^{r+1} - x^r\| \le \delta, \qquad \forall\, r. \tag{3.10b}$$

Since $c_i = \infty$ for all i, we have that $x_i^* = 0$ for all $i \notin I^*$ and all $x^* \in X^*$, and it immediately follows from (3.10a) that

$$x_i^r \le \delta, \qquad \forall\, r, \qquad \forall\, i \notin I^*. \tag{3.11}$$

We first have the following lemma which states that Lemma 9 holds in the special case where the coordinates that start near the boundary of X remain near the boundary (also assuming that the remaining coordinates start far from the boundary).

Lemma 10. Fix any $k \in \{1,\ldots,n\}$. If for some nonempty $J \subseteq I^*$ and some two integers $t' > t$ we have

$$x_i^t > \sigma_k\delta, \qquad \forall\, i \in J, \tag{3.12a}$$
$$x_i^r \le \sigma_{k-1}\delta, \qquad \forall\, i \notin J, \qquad \forall\, r = t, t+1, \ldots, t'-1, \tag{3.12b}$$

then the following hold:

(a) $\quad x_i^{t'} > \sigma_{k-1}\delta, \qquad \forall\, i \in J.$

(b) There exists an $x^* \in X^*$ such that

$$\|x^r - x^*\|_\infty \le \sigma_k\delta, \qquad \forall\, r = t, t+1, \ldots, t'-1.$$

Proof. Let x^* be any element of X^* satisfying $\phi(x^t) = \|x^t - x^*\|$. Then we have from (3.10a) that

$$\|x^t - x^*\| \ \le\ \delta. \tag{3.13}$$

Also we have from (3.12b) that, for all $i \notin J$, $x_i^* \le x_i^t + \|x^t - x^*\| \le \sigma_{k-1}\delta + \|x^t - x^*\|$, which together with (3.13) implies $0 \le x_i^* \le \sigma_{k-1}\delta + \delta$. Since $0 \le x_i^r \le \sigma_{k-1}\delta$ for $r = t, t+1, \ldots, t'-1$ (cf. (3.12b)), this in turn implies that

$$|x_i^r - x_i^*| \ \le\ \sigma_{k-1}\delta + \delta, \qquad \forall\ i \notin J, \qquad \forall\ r = t, t+1, \ldots, t'-1. \tag{3.14}$$

Next we prove by induction that, for $r = t, t+1, \ldots, t'-1$,

$$x_i^r \ >\ \sigma_{k-1}\delta + \delta, \qquad \forall\ i \in J. \tag{3.15}$$

Eq. (3.15) clearly holds for $r = t$ (cf. (3.12a) and $\sigma_k \ge \sigma_{k-1} + 1$). Suppose that (3.15) holds for $r = t, t+1, \ldots, s$, for some $s \in \{t, t+1, \ldots, t'-2\}$. We will prove that it also holds for $r = s+1$. Since $0 < x_i^r < c_i$ for all $i \in J$ and all $r = t+1, \ldots, s$ (cf. (3.15) and $c_i = \infty$ for all i), we have from Lemma 8 that

$$\|x_J^s - x_J^*\| \ \le\ \Delta\|x_J^t - x_J^*\| \ +\ \beta \max_{r \in \{t, \ldots, s\}} \|x_{\bar{J}}^r - x_{\bar{J}}^*\|_\infty,$$

which together with (3.13) and (3.14) implies

$$\|x_J^s - x_J^*\| \ \le\ \Delta\delta \ +\ \beta\,(\sigma_{k-1}\delta + \delta). \tag{3.16}$$

Then we have that, for any $i \in J$,

$$\begin{aligned}
x_i^{s+1} &\ \ge\ x_i^t - \|x_J^t - x_J^{s+1}\| \\
&\ \ge\ x_i^t - (\|x_J^t - x_J^*\| + \|x_J^* - x_J^s\| + \|x_J^s - x_J^{s+1}\|) \\
&\ >\ \sigma_k\delta - (\delta + \|x_J^* - x_J^s\| + \delta) \\
&\ \ge\ \sigma_k\delta - (\delta + (\Delta\delta + \beta\sigma_{k-1}\delta + \beta\delta) + \delta) \\
&\ =\ \sigma_{k-1}\delta + \delta,
\end{aligned}$$

where the strict inequality follows from Eqs. (3.10b), (3.12a) and (3.13). This completes the induction and proves that (3.15) holds for $r = t, t+1, \ldots, t'-1$. Since (3.15) holds for $r = t, t+1, \ldots, t'-1$, it can be seen from the argument above that (3.16) holds for $s = t, t+1, \ldots, t'-1$, which combined with (3.14) (and using the facts $\beta > 1$ and $\|z\|_\infty \le \|z\|$ for all z) yields

$$\|x^r - x^*\|_\infty \le (\Delta + \beta\sigma_{k-1} + \beta)\delta, \qquad \forall\, r = t, t+1, \ldots, t'-1.$$

Since $\Delta + \beta\sigma_{k-1} + \beta \le \sigma_k$, this proves part (b). From (3.15) with $r = t'-1$, we have that, for all $i \in J$,

$$\begin{aligned} x_i^{t'} &\ge x_i^{t'-1} - \|x^{t'-1} - x^{t'}\| \\ &> \sigma_{k-1}\delta + \delta - \|x^{t'-1} - x^{t'}\|. \end{aligned}$$

Since $\|x^{t'-1} - x^{t'}\| \le \delta$ (cf. (3.10b)), this proves part (a). **Q.E.D.**

The following lemma extends Lemma 10 by removing the assumption that the coordinates that start near the boundary of X remain near the boundary (while still assuming that the remaining coordinates start far from the boundary):

Lemma 11. Fix any $k \in \{1,\ldots,n\}$. If for some $J \subseteq I^*$ with $\mathrm{Card}(J) \ge \mathrm{Card}(I^*) - k + 1$ and some integer t we have

$$\begin{aligned} x_i^t &> \sigma_k\delta, && \forall\, i \in J, \\ x_i^t &\le \sigma_{k-1}\delta, && \forall\, i \notin J, \end{aligned}$$

then there exists an $x^* \in X^*$ and a $\bar{t} \ge t$ satisfying

$$\|x^r - x^*\|_\infty \le \sigma_k\delta, \qquad \forall\, r \ge \bar{t}. \tag{3.17}$$

Proof. We prove by induction on k. By Lemma 10 (b), we see that the claim holds for $k = 1$. [Since in this case $J = I^*$, by (3.11), the condition (3.12b) is satisfied for all $t' \ge t$. Then Lemma 10 (b) yields that there exists an $x^* \in X^*$ such that $\|x^r - x^*\|_\infty \le \sigma_k\delta$, for all $r \ge t$.] Suppose that the claim holds for $k = 1, 2, \ldots, h-1$, for some $h \ge 2$. We will show that it also holds for $k = h$. Fix any $J \subseteq I^*$ with $\mathrm{Card}(J) \geq \mathrm{Card}(I^*) - h + 1$ and any integer t for which

$$x_i^t > \sigma_h \delta, \qquad \forall\ i \in J, \tag{3.18a}$$

$$x_i^t \leq \sigma_{h-1} \delta, \qquad \forall\ i \notin J. \tag{3.18b}$$

We consider two cases:

(i) $x_i^r \leq \sigma_{h-1}\delta$, for all $i \notin J$ and all $r \geq t$.
Since $x_i^t > \sigma_h \delta$, for all $i \in J$ (cf. (3.18a)), it immediately follows from Lemma 10 (b) that there exists an $x^* \in X^*$ such that

$$\|x^r - x^*\|_\infty \leq \sigma_h \delta, \qquad \forall\ r \geq t.$$

This shows that (3.17) holds for $k = h$ (with $\bar{t} = t$ and with the above choice of x^*).

(ii) There exists an $r > t$ and an $i \notin J$ such that $x_i^r > \sigma_{h-1}\delta$.
Let

$$t' = \text{Smallest } r\ (r > t) \text{ such that } x_i^r > \sigma_{h-1}\delta \text{ for some } i \notin J.$$

Then, by (3.18b), $x_i^r \leq \sigma_{h-1}\delta$ for all $i \notin J$ and all $r = t, t+1, \ldots, t'-1$. Since $x_i^t > \sigma_h\delta$, for all $i \in J$ (cf. (3.18a)), Lemma 10 (a) yields that

$$x_i^{t'} > \sigma_{h-1}\delta, \qquad \forall\ i \in J. \tag{3.19}$$

Consider the h+1 intervals

$$[0,\sigma_0\delta],\ (\sigma_0\delta,\sigma_1\delta],\ (\sigma_1\delta,\sigma_2\delta],\ \ldots,\ (\sigma_{h-2}\delta,\sigma_{h-1}\delta],\ (\sigma_{h-1}\delta,\infty).$$

We have from (3.19) and the fact $x_i^{t'} > \sigma_{h-1}\delta$ for some $i \notin J$ that the (h+1)-st interval contains at least $\mathrm{Card}(J) + 1$ elements from the set $\{x_1^{t'}, x_2^{t'}, \ldots, x_n^{t'}\}$. Also, (3.11) and $\sigma_0 = 1$ imply that the first interval contains at least $n - \mathrm{Card}(I^*)$ elements from the same set. Since $\mathrm{Card}(J) \geq \mathrm{Card}(I^*) - h + 1$, this leaves at most $h - 2$ elements from that set to go into the remaining $h-1$ intervals. Hence, by the Pigeon Hole principle, there must exist some $j \in \{1,2,\ldots,h-1\}$ such that

$$x_i^{t'} \notin (\sigma_{j-1}\delta, \sigma_j\delta], \qquad \forall\, i.$$

Let h' be the <u>largest</u> j for which this occurs. Then the interval $(\sigma_{h'}\delta, \infty)$ contains at least $\mathrm{Card}(J) + h - h'$ elements from the set $\{x_1^{t'}, x_2^{t'}, \ldots, x_n^{t'}\}$. Let J' be the index set for these elements, i.e., $J' = \{\, i \mid x_i^{t'} > \sigma_{h'}\delta \,\}$. Then we have

$$\begin{aligned} x_i^{t'} &> \sigma_{h'}\delta, && \forall\, i \in J', \\ x_i^{t'} &\le \sigma_{h'-1}\delta, && \forall\, i \notin J', \end{aligned}$$

and

$$\begin{aligned} \mathrm{Card}(J') &\ge \mathrm{Card}(J) + h - h' \\ &\ge \mathrm{Card}(I^*) + 1 - h'. \end{aligned}$$

Moreover, by (3.11), we have $J' \subseteq I^*$. Since $h' < h$, we can apply our induction hypothesis to h', J' and t' to conclude that there exists an $x^* \in X^*$ and a $\bar{t} \ge t'$ satisfying

$$\|x^r - x^*\|_\infty \le \sigma_{h'}\delta, \qquad \forall\, r \ge \bar{t}.$$

Since $\sigma_{h'} \le \sigma_h$, this shows that (3.17) holds for $k = h$ (with the given $\bar{t}$ and x^*).

This then completes the induction on k and proves the lemma. **Q.E.D.**

Now we use Lemma 11 to prove our claim. Fix any integer $\bar{r} \ge 1$. Consider the two possible cases: either (i) $x_i^r \le \sigma_n\delta$ for all i and all $r \ge \bar{r}$, or (ii) there exists a $t \ge \bar{r}$ and an i such that $x_i^t > \sigma_n\delta$. In case (i), let x^* be an element of X^* such that $\phi(x^{\bar{r}}) = \|x^{\bar{r}} - x^*\|$. Then we have from (3.10a) that, for all i,

$$0 \le x_i^* \le x_i^{\bar{r}} + \|x^{\bar{r}} - x^*\| \le \sigma_n\delta + \delta.$$

Since $0 \le x_i^r \le \sigma_n\delta$ for all i and all $r \ge \bar{r}$, this implies that

$$\|x^r - x^*\|_\infty \le \sigma_n\delta + \delta, \qquad \forall\, r \ge \bar{r}.$$

Hence (3.9) holds with $\hat{r} = \bar{r}$ and with the above choice of x^*. Now consider case (ii). In this case, by the Pigeon Hole principle, one of the following n intervals

$$(\sigma_0\delta, \sigma_1\delta], \ (\sigma_1\delta, \sigma_2\delta], \ \ldots, \ (\sigma_{n-1}\delta, \sigma_n\delta]$$

does not contain any element from $\{x_1^t, x_2^t, \ldots, x_n^t\}$, i.e., there exists $j \in \{1,2,\ldots,n\}$ such that

$$x_i^t \notin (\sigma_{j-1}\delta, \sigma_j\delta], \qquad \forall\, i.$$

Choose k to be the <u>largest</u> such j and let $J = \{\, i \mid x_i^t > \sigma_k\delta \,\}$. Then $\mathrm{Card}(J) \ge n - k + 1$ and

$$\begin{aligned} x_i^t &> \sigma_k\delta, && \forall\, i \in J, \\ x_i^t &\le \sigma_{k-1}\delta, && \forall\, i \notin J. \end{aligned}$$

Moreover, by (3.11) and $\sigma_k \ge 1$, we see that $J \subseteq I^*$. Hence the assumptions of Lemma 11 is satisfied by k, J and t, and it follows from Lemma 11 that there exists an $x^* \in X^*$ and a $\bar{t} \ge t$ satisfying

$$\|x^r - x^*\|_\infty \le \sigma_k\delta, \qquad \forall\, r \ge \bar{t}.$$

Since $\sigma_k \le \sigma_n$, this shows that (3.9) holds (with the given x^* and with $\hat{r} = \bar{t}$). **Q.E.D.**

The following main convergence result then follows as a corollary of Lemma 9.

Theorem 1. The matrix splitting algorithm (1.4)-(1.5) is well-defined and it generates a sequence of iterates $\{x^r\}$ converging to an element of X^*.

Proof. The algorithm is well-defined by Lemma 2 (a). Now, for any $\varepsilon > 0$, Lemma 9 shows that there exists an $x^* \in X^*$ and an $\hat{r} > 0$ such that

$$\|x^r - x^*\|_\infty < \varepsilon/2, \qquad \forall\, r \geq \hat{r}.$$

Hence, for all $r_1, r_2 > \hat{r}$, there holds

$$\begin{aligned}\|x^{r_1} - x^{r_2}\|_\infty &\leq \|x^{r_1} - x^*\|_\infty + \|x^* - x^{r_2}\|_\infty \\ &\leq \varepsilon/2 + \varepsilon/2 = \varepsilon.\end{aligned}$$

This implies that $\{x^r\}$ is a Cauchy sequence so that it converges. By Lemma 7 (b), it converges to an element of X^*. **Q.E.D.**

4. SOR Matrix Splitting Algorithms

In this section we consider three extensions of the basic algorithm (1.4)-(1.5). Firstly we consider one that adds an under/over-relaxation parameter to the algorithm. This extension is motivated by the block SOR methods of Cottle, Golub and Sacher [CGS78], of Cottle and Goheen [CoG78], and of Cottle and Pang [CoP82] (which introduced a mechanism for over-relaxation) and the methods of Mangasarian [Man77] and of Aganagic [Aga78] (which introduced a mechanism for underrelaxation). Secondly we consider a Gauss-Seidel extension of the basic algorithm. In this algorithm, only a subset of the coordinates are relaxed at each iteration while the other coordinates are held fixed. Lastly, we consider an SOR extension of the basic algorithm which allows non-cyclic order of relaxation. This third algorithm contains the previous two as special cases but is shown to be convergent only in a certain weak sense.

We first describe the under/over-relaxation extension. In the algorithm, we choose a splitting (B,C) of M and a relaxation parameter $\bar{\omega}$ satisfying

$$0 < \bar{\omega}, \qquad B - C + (1-\bar{\omega})M \text{ is positive definite.} \tag{4.1a}$$

We also choose a second relaxation parameter $\underline{\omega}$ satisfying

$$0 < \underline{\omega} \leq \min\{1,\bar{\omega}\}, \tag{4.1b}$$

and an $n\times n$ positive diagonal matrix D. Then, for any chosen $x^0 \in X$, we generate a sequence of vectors $\{x^0, x^1, \ldots\}$ in X by the formula

$$x^{r+1} = (1-\omega^r)x^r + \omega^r \hat{x}^r, \tag{4.2a}$$

where $\hat{x}^r$ is a solution of the equation

$$y = [\, y - D(By + Cx^r + q)\,]^+, \tag{4.2b}$$

and ω^r is any scalar in $[\underline{\omega}, \bar{\omega}]$ such that x^{r+1} given by (4.2a) is in X.

Notice that if (B,C) is a regular splitting and

$$0 < \bar{\omega} < 1 + 1/\|Q^{-1/2}MQ^{-1/2}\|,$$

where Q denotes the symmetric part of $B - C$, i.e. $Q = ((B-C) + (B-C)^T)/2$, then (4.1a) is satisfied. Hence, the above algorithm contains as a special case the algorithm (1.4)-(1.5) (let $\underline{\omega} = \bar{\omega} = 1$ and D be the $n\times n$ identity matrix). The relaxation parameter ω^r introduced in (4.2a) is useful mainly when $\omega^r > 1$ (i.e., over-relaxation [OrR70]), which in some cases can significantly improve the convergence. Nonetheless, the case of under-relaxation, i.e. $\omega^r < 1$, is also of some practical interest since, in this case, it is only required that $B - C$ be positive definite on the null space of M (instead of on the entire space) in order for (4.1a) to hold. The purpose for introducing the matrix D in (4.2b) is, from the point of view of convergence, largely cosmetic as the presence of D does not change the sequence of iterates generated. [To see this, note that since D is a diagonal matrix, y is a solution of (4.2b) if and only if $y \in X$ and $y_i = 0$ ($y_i = c_i$) for all i such that $D_{ii}(B_i y + C_i x^r + q_i) > 0$ (< 0). Since $D_{ii} > 0$ for all i, this set of conditions is equivalent to $y \in X$ and $y_i = 0$ ($y_i = c_i$) for all i such that $B_i y + C_i x^r + q_i > 0$ (< 0), which in turn is equivalent to $y = [\, y - (By + Cx^r + q)\,]^+$ or, by (1.4), $y = \mathcal{A}_B(x^r)$.] However, by choosing D to match the structure of B and C, we can in some cases simplify the form of the iteration (see Section 5 for examples). Note that since the sequence of iterates generated is independent of D, we can also allow D to be time-varying.

By modifying the argument used in Sections 2 and 3, we can show that the algorithm (4.1a)-(4.2b) is convergent:

Theorem 2. For any splitting (B,C) of M and any scalars $\underline{\omega}$, $\bar{\omega}$ satisfying (4.1a)-(4.1b), any n×n positive diagonal matrix D and any $x^0 \in X$, the sequence of iterates $\{x^r\}$ generated by (4.2a)-(4.2b) is well defined and converges to an element of X^*.

Proof (sketch). Since $2B = (B - C + (1-\bar{\omega})M) + \bar{\omega}M$ (cf. $M = B + C$) and $\bar{\omega}M$ is positive semi-definite, we have from (4.1a) that B is positive definite. The proof of Lemma 2 (a) then shows that $\mathcal{A}_B$ is a well-defined point-to-point mapping. For r = 0,1,..., let $\hat{x}^r$ be a solution of (4.2b). Then from the preceding discussion we have that

$$\hat{x}^r = \mathcal{A}_B(x^r), \qquad r = 0,1,\ldots,$$

so that $\{\hat{x}^r\}$ is well-defined. Since $x^{r+1} = (1-\omega^r)x^r + \omega^r\hat{x}^r$ for all r, $\{x^r\}$ is well-defined.

Now we show that $\{x^r\}$ is convergent. The proof of this is very similar to that of Theorem 1, with Lemmas 4, 6 and 8 replaced by more general versions of themselves that take into account the relaxation parameters. Firstly we have the following extension of Lemma 4:

Lemma 12. For any splitting (B,C) of M and any scalars $\underline{\omega}$, $\bar{\omega}$ satisfying (4.1a)-(4.1b), the following hold:

(a) For any nonempty $J \subseteq \{1,\ldots,n\}$, there exist $\rho_J \in (0,1)$ and $\tau_J > 0$ such that

$$\left\|\prod_{h=1}^{k}(I - \theta^h M_{JJ}(B_{JJ})^{-1})z\right\| \le \tau_J (\rho_J)^k \|z\|, \qquad \forall\, k \ge 1, \qquad \forall\, z \in \mathrm{Span}(M_{JJ}),$$

for all sequences of scalars $\{\theta^1, \theta^2, \ldots\}$ in the interval $[\underline{\omega},\bar{\omega}]$.

(b) There exists a $\Delta > 0$ such that, for any nonempty $J \subseteq \{1,\ldots,n\}$,

$$\left\|\prod_{h=1}^{k}(I - \theta^h (B_{JJ})^{-1} M_{JJ})z\right\| \le \Delta\|z\|, \qquad \forall\, k \ge 1, \qquad \forall\, z,$$

for all sequences of scalars $\{\theta^1, \theta^2, \ldots\}$ in the interval $[\underline{\omega},\bar{\omega}]$.

Proof. This proof is similar to that of Lemmas 3 and 4. Fix any nonempty $J \subseteq \{1,\ldots,n\}$ and consider the iteration

$$y^{r+1} = (1-\bar{\omega})y^r - \bar{\omega}(B_{JJ})^{-1}C_{JJ}y^r = (I - \bar{\omega}(B_{JJ})^{-1}M_{JJ})y^r, \qquad r = 0, 1, \ldots.$$

Let $g:\Re^{\mathrm{Card}(J)} \to \Re$ denote the function $g(y) = \langle y, M_{JJ}y\rangle/2$. By using $M_{JJ} = B_{JJ} + C_{JJ}$, we find that

$$g(y^{r+1}) = g(y^r) - \frac{1}{2\bar{\omega}}\langle y^{r+1}-y^r, ((1-\bar{\omega})M_{JJ} + B_{JJ} - C_{JJ})(y^{r+1}-y^r)\rangle.$$

Since $(1-\bar{\omega})M_{JJ} + B_{JJ} - C_{JJ}$ is positive definite and $g(y)$ is non-negative for all y, this implies that $y^{r+1} - y^r \to 0$. Then, by an argument analogous to that used in the proof of Lemma 3, we obtain that $(I - \bar{\omega}M_{JJ}(B_{JJ})^{-1})^k z \to 0$ as $k \to \infty$ for all $z \in \mathrm{Span}(M_{JJ})$. Part (a) then follows from the following fact proven in Appendix A (letting $A = I - \bar{\omega}M_{JJ}(B_{JJ})^{-1}$, $\mathcal{V} = \mathrm{Span}(M_{JJ})$, and $\delta = \underline{\omega}/\bar{\omega}$)

Fact 2. For any $m \times m$ matrix A and any linear subspace $\mathcal{V}$ of $\Re^m$, if $(A)^k z \to 0$ as $k \to \infty$ for all $z \in \mathcal{V}$, then, for each $\delta \in (0,1]$, there exist $\rho_\delta \in (0,1)$ and $\tau_\delta > 0$ such that

$$\left\| \prod_{h=1}^{k} ((1-\theta^h)I + \theta^h A)\, z \right\| \le \tau_\delta\, (\rho_\delta)^k \|z\|, \qquad \forall\, k \ge 1, \qquad \forall\, z \in \mathcal{V},$$

for all sequences of scalars $\{\theta^1, \theta^2, \ldots\}$ in the interval $[\delta, 1]$.

Part (b) then follows from an argument analogous to that used in the proof of Lemma 3.
Q.E.D.

By using Lemma 12, we obtain the following extension of Lemma 8:

Lemma 13. Consider any $J \subseteq I^*$. If for some two integers $s \ge t \ge 0$ we have $0 < \hat{x}_i^r < c_i$ for all $i \in J$ and all $r = t+1, t+2, \ldots, s$, then, for any $x^* \in X^*$,

$$\|x_J^s - x_J^*\| \;\le\; \Delta\|x_J^t - x_J^*\| + \beta \max_{r\in\{t,\ldots,s\}} \|x_{\bar J}^r - x_{\bar J}^*\|_\infty,$$

where Δ is as in Lemma 12 and β is a constant that depends on M, B, $\underline{\omega}$ and $\bar{\omega}$ only.

Proof. Since $\hat{x}^r = \mathcal{A}_B(x^r)$ for all r, by an argument analogous to that used in the proof of Lemma 8, we obtain that

$$\begin{aligned}\hat{x}_J^r - x_J^* &= (I - (B_{JJ})^{-1}M_{JJ})(x_J^r - x_J^*) - (B_{JJ})^{-1}M_{J\bar J}(x_{\bar J}^r - x_{\bar J}^*)\\ &\qquad - (B_{JJ})^{-1}B_{J\bar J}(\hat{x}_{\bar J}^r - x_{\bar J}^r), \qquad r = t,\ldots,s-1.\end{aligned}$$

Since $x^{r+1} - x^r = \omega^r(\hat{x}^r - x^r)$ for all r (cf. (4.2a)), this in turn implies that

$$\begin{aligned}x_J^{r+1} - x_J^* &= (I - \omega^r(B_{JJ})^{-1}M_{JJ})(x_J^r - x_J^*) - \omega^r(B_{JJ})^{-1}M_{J\bar J}(x_{\bar J}^r - x_{\bar J}^*)\\ &\qquad - (B_{JJ})^{-1}B_{J\bar J}(x_{\bar J}^{r+1} - x_{\bar J}^r), \quad r = t,\ldots,s-1.\end{aligned}$$

By successively applying the above recursion for r = t, ..., s−1, we obtain

$$\begin{aligned}x_J^s - x_J^* &= (G^{h-1})\cdots(G^0)(x_J^t - x_J^*)\\ &\quad - \sum_{k=0}^{h-1}(G^{h-1})\cdots(G^{k+1})\omega^{t+k}(B_{JJ})^{-1}M_{J\bar J}(x_{\bar J}^{t+k} - x_{\bar J}^*)\\ &\quad - \sum_{k=0}^{h-1}(G^{h-1})\cdots(G^{k+1})(y^{k+1} - y^k),\end{aligned}$$

where we denote $G^k = I - \omega^{t+k}(B_{JJ})^{-1}M_{JJ}$, $y^k = (B_{JJ})^{-1}B_{J\bar J}(x_{\bar J}^{t+k} - x_{\bar J}^*)$, and $h = s - t$. By rearranging terms within the last sum as in the proof of Lemma 8, we obtain the alternative expression:

$$\begin{aligned}x_J^s - x_J^* &= (G^{h-1})\cdots(G^0)(x_J^t - x_J^*)\\ &\quad - \sum_{k=0}^{h-1}(G^{h-1})\cdots(G^{k+1})\omega^{t+k}(B_{JJ})^{-1}M_{J\bar J}(x_{\bar J}^{t+k} - x_{\bar J}^*)\\ &\quad - \sum_{k=1}^{h-1}(G^{h-1})\cdots(G^{k+1})(G^k - I)y^k + y^h - (G^{h-1})\cdots(G^1)y^0.\end{aligned}$$

Let $H^k = I - \omega^{t+k}M_{JJ}(B_{JJ})^{-1}$. Then $G^k = (B_{JJ})^{-1}(H^k)B_{JJ}$, so that $(G^{h-1})\cdots(G^{k+1}) = (B_{JJ})^{-1}(H^{h-1})\cdots(H^{k+1})B_{JJ}$ for all k. This together with the above equation and the fact $G^k - I = -\omega^{t+k}(B_{JJ})^{-1}M_{JJ}$ yields

$$\begin{aligned} x_J{}^s - x_J{}^* \;=\; & (G^{h-1})\cdots(G^0)(x_J{}^t - x_J{}^*) \\ & - \sum_{k=0}^{h-1} \omega^{t+k}(B_{JJ})^{-1}(H^{h-1})\cdots(H^{k+1})M_{J\bar{J}}(x_{\bar{J}}{}^{t+k} - x_{\bar{J}}{}^*) \\ & + \sum_{k=1}^{h-1} \omega^{t+k}(B_{JJ})^{-1}(H^{h-1})\cdots(H^{k+1})M_{JJ}y^k + y^h - (G^{h-1})\cdots(G^1)y^0. \end{aligned}$$

The remainder of the proof then follows from Lemma 12 (using the fact $\underline{\omega} \le \omega^r \le \bar{\omega}$ for all r) and an argument analogous to that used in the proof of Lemma 8. **Q.E.D.**

We now show that Lemma 6 holds. Fix any integer $r \ge 0$. Let $x = x^r$, $y = \hat{x}^r$, and $\omega = \omega^r$. By using the equation $M = B + C$, we find that

$$\begin{aligned} f((1-\omega)x + \omega y) - f(x) \;=\; & -\frac{\omega}{2}\langle y - x, ((1-\omega)M + B - C)(y - x)\rangle \\ & + \omega\langle y - x, By + Cx + q\rangle. \end{aligned}$$

Since $\langle y - x, By + Cx + q\rangle \le 0$ (cf. $y = [\, y - (By + Cx + q)\,]^+$), this together with $0 < \omega \le \bar{\omega}$ and the positive semi-definite property of M yields

$$\begin{aligned} f((1-\omega)x + \omega y) - f(x) \;\le\; & -\frac{\omega}{2}\langle y - x, ((1-\omega)M + B - C)(y - x)\rangle \\ \le\; & -\frac{\omega}{2}\langle y - x, ((1-\bar{\omega})M + B - C)(y - x)\rangle. \end{aligned}$$

Since $(1-\bar{\omega})M + B - C$ is positive definite (cf. (4.1a)) and $\omega \ge \underline{\omega} > 0$, this implies $f((1-\omega)x + \omega y) - f(x) \le -\gamma\|y - x\|^2$, for some postive constant γ. Hence $f(x^{r+1}) - f(x^r) \le -\gamma\|x^r - \hat{x}^r\|$ for all r. Since f is bounded from below, this implies $x^r - \hat{x}^r \to 0$ and $x^r - [\, x^r - Mx^r - q\,]^+ \to 0$. Since $\|x^{r+1} - x^r\| \le \bar{\omega}\|x^r - \hat{x}^r\|$ (cf. (4.2a) and $0 < \omega^r \le \bar{\omega}$), we also have $x^{r+1} - x^r \to 0$, so that Lemma 6 holds. **Q.E.D.**

Since Lemma 7 depends only on Lemma 6, it follows that Lemma 7 also holds. Now, we have from (4.2a) that, for all r, $x^{r+1} - \hat{x}^r = (1-1/\omega^r)(x^{r+1} - x^r)$, so that

$$\|x^{r+1} - \hat{x}^r\| \le \max\{1,1/\underline{\omega}\}\|x^{r+1} - x^r\|.$$

Hence, by redefining the scalar σ_0 to be $1 + \max\{1,1/\underline{\omega}\}$ with the other scalars $\sigma_1, \ldots, \sigma_n$ recursively defined as before, the proof of Lemma 10 (with Lemma 8 replaced by Lemma 13) still goes through. Lemmas 9 and 11 then follow from Lemmas 6, 7 and 10.

Q.E.D.

Remark 4. We can also use <u>different relaxation parameters for different coordinates</u> provided that the relaxation parameters are <u>fixed</u>. More precisely, let us consider the following under-relaxed algorithm:

$$x^{r+1} = \begin{bmatrix} 1-\bar{\omega}_1 & & \\ & \ldots & \\ & & 1-\bar{\omega}_n \end{bmatrix} x^r + \begin{bmatrix} \bar{\omega}_1 & & \\ & \ldots & \\ & & \bar{\omega}_n \end{bmatrix} \mathcal{A}_B(x^r), \qquad r = 0, 1, \ldots,$$

where $x^0 \in X$, $\bar{\omega}_1, \ldots, \bar{\omega}_n$ are fixed scalars in the interval (0,1], and (B,C) is a splitting of M for which the matrix

$$2B \begin{bmatrix} 1/\bar{\omega}_1 & & \\ & \ldots & \\ & & 1/\bar{\omega}_n \end{bmatrix} - M$$

is positive definite. In the special case where $\bar{\omega}_1 = \ldots = \bar{\omega}_n$, this algorithm reduces to the algorithm (4.1a)-(4.2b) using fixed under-relaxation. By suitably modifying the proof of Theorem 2, it can be shown that this under-relaxed algorithm is convergent.

Now we consider a Gauss-Seidel type algorithm. Let $\{1,\ldots,n\}$ be partitioned into m nonempty, mutually disjoint subsets $I_1, I_2, \ldots, I_m$ (i.e., $I_i \cap I_j = \emptyset$ if $i \neq j$ and $I_1 \cup \ldots \cup I_m = \{1,\ldots,n\}$). For $j = 1,\ldots,m$, we choose a regular splitting $(B_{I_jI_j}, C_{I_jI_j})$ of $M_{I_jI_j}$ and define a corresponding mapping $\mathcal{A}_j: X \to X$ by

$$\mathcal{A}_j(x) = \{\, y \in \Re^n \mid y_{I_j} = [\, y_{I_j} - (B_{I_jI_j}y_{I_j} + C_{I_jI_j}x_{I_j} + M_{I_j\bar{I}_j}x_{\bar{I}_j} + q_{I_j})\,]_j^+, \quad y_{\bar{I}_j} = x_{\bar{I}_j} \,\}, \tag{4.3}$$

where $[\cdot]_j^+$ denotes the orthogonal projection onto the box $\times_{i \in I_j}[0,c_i]$. By Lemma 2 (a), $\mathcal{A}_j$ is a well-defined point-to-point mapping. The mapping $\mathcal{A}_j$ has the effect of applying a matrix splitting iteration to the subset of coordinates indexed by I_j, while the other coordinates are held fixed. The Gauss-Seidel matrix splitting (GS-MS) algorithm generates a sequence of iterates by applying cyclically the mappings $\mathcal{A}_1, \ldots, \mathcal{A}_m$:

GS-MS Algorithm. Choose an $x^0 \in X$. Generate a sequence of vectors $\{x^0, x^1, \ldots\}$ in X by the formula

$$x^{r+1} = (\mathcal{A}_m \circ \cdots \circ \mathcal{A}_2 \circ \mathcal{A}_1)(x^r), \qquad r = 0, 1, \ldots. \tag{4.4}$$

It is easily seen that in the special case where m = 1, this algorithm reduces to the algorithm (4.1a)-(4.2b) with relaxation parameters $\underline{\omega} = \bar{\omega} = 1$.

By extending the proof of Theorem 2, we can show that the GS-MS algorithm is convergent:

Theorem 3. The sequence of iterates generated by the GS-MS algorithm (4.3)-(4.4) converges to an element of X^*.

Proof (sketch). Similar to the proof of Theorem 2, it suffices to show that Lemmas 6 and 8 hold.

We first show that Lemma 6 holds. Since the I_j's are disjoint, we have from (4.4) that

$$x_{I_j}^{r+1} = \mathcal{A}_j(x_{I_1}^{r+1}, \ldots, x_{I_{j-1}}^{r+1}, x_{I_j}^{r}, \ldots, x_{I_m}^{r}), \qquad j = 1, \ldots, m, \quad \forall\, r.$$

Therefore, by (4.3), each $x_{I_j}^{r+1}$ satisfies

$$x_{I_j}^{r+1} = [x_{I_j}^{r+1} - (B_{I_jI_j}x_{I_j}^{r+1} + C_{I_jI_j}x_{I_j}^{r} + \sum_{k<j} M_{I_jI_k}x_{I_{k-1}}^{r+1} + \sum_{k>j} M_{I_jI_k}x_{I_k}^{r} + q_{I_j})]_j^+, \tag{4.5}$$

so that Lemma 2 (b) yields

$$f(x_{I_1}^{r+1},\ldots,x_{I_j}^{r+1},x_{I_{j+1}}^{r},\ldots,x_{I_m}^{r}) - f(x_{I_1}^{r+1},\ldots,x_{I_{j-1}}^{r+1},x_{I_j}^{r},\ldots,x_{I_m}^{r})$$
$$\leq -\langle x_{I_j}^{r+1} - x_{I_j}^{r}, (B_{I_jI_j} - C_{I_jI_j})(x_{I_j}^{r+1} - x_{I_j}^{r})\rangle/2.$$

Since $B_{I_jI_j} - C_{I_jI_j}$ is positive definite, this implies that there exists a $\gamma > 0$ such that

$$f(x_{I_1}^{r+1},\ldots,x_{I_j}^{r+1},x_{I_{j+1}}^{r},\ldots,x_{I_m}^{r}) \leq f(x_{I_1}^{r+1},\ldots,x_{I_{j-1}}^{r+1},x_{I_j}^{r},\ldots,x_{I_m}^{r}) - \gamma\|x_{I_j}^{r+1} - x_{I_j}^{r}\|^2,$$

for all r and all j. By applying this inequality recursively for all j, we obtain that

$$f(x^{r+1}) \leq f(x^r) - \gamma\sum_{j=1}^{m}\|x_{I_j}^{r+1} - x_{I_j}^{r}\|^2, \qquad \forall\ r.$$

Since f is bounded from below, this shows that $\sum_{j=1}^{m}\|x_{I_j}^{r+1} - x_{I_j}^{r}\|^2 \to 0$ or, equivalently, $x^{r+1} - x^r \to 0$. Now, from (4.5) we have that

$$\begin{aligned}
&\|x_{I_j}^{r+1} - [\,x_{I_j}^{r+1} - (M_{I_j}x^{r+1} + q_{I_j})\,]_j^+\| \\
&\quad = \|[\,x_{I_j}^{r+1} - (B_{I_jI_j}x_{I_j}^{r+1} + C_{I_jI_j}x_{I_j}^{r} + \sum_{k<j}M_{I_jI_k}x_{I_{k-1}}^{r+1} + \sum_{k>j}M_{I_jI_k}x_{I_k}^{r} + q_{I_j})]_j^+ \\
&\qquad - [\,x_{I_j}^{r+1} - (M_{I_j}x^{r+1} + q_{I_j})\,]_j^+\| \\
&\quad \leq \|-B_{I_jI_j}(x_{I_j}^{r+1} - x_{I_j}^{r}) + \sum_{k\geq j}M_{I_jI_k}(x_{I_k}^{r+1} - x_{I_k}^{r})\|, \qquad j = 1,\ldots,m,
\end{aligned}$$

where the inequality follows from the nonexpansive property of the projection mapping $[\cdot]_j^+$, i.e. $\|[y]_j^+ - [x]_j^+\| \leq \|y - x\|$ for all x, y. Since $x^{r+1} - x^r \to 0$, the above inequality shows that $x^{r+1} - [\,x^{r+1} - (Mx^{r+1} + q)\,]^+ \to 0$.

Now we show that Lemma 8 still holds (possibly with a different β). Consider any $J \subseteq I^*$ and suppose that for some two integers $s \geq t \geq 0$ we have $0 < x_i^r < c_i$ for all $i \in J$ and all $r = t+1, t+2, \ldots, s$. Let $J_j = J \cap I_j$ and $\hat{J}_j = \tilde{J} \cap I_j$. Then $I_j = J_j \cup \hat{J}_j$ for all j, and we have from (4.5) that, for any $r \in \{t,\ldots,s-1\}$,

$$
\begin{aligned}
0 &= B_{J_jI_j}x_{I_j}^{r+1} + C_{J_jI_j}x_{I_j}^{\ r} + \sum_{k<j} M_{J_jI_k}x_{I_k}^{r+1} + \sum_{k>j} M_{J_jI_k}x_{I_k}^{\ r} + q_{J_j} \\
&= B_{J_jJ_j}x_{J_j}^{r+1} + C_{J_jJ_j}x_{J_j}^{\ r} + \sum_{k<j} M_{J_jJ_k}x_{J_k}^{r+1} + \sum_{k>j} M_{J_jJ_k}x_{J_k}^{\ r} \\
&\quad + B_{J_j\hat{J}_j}x_{\hat{J}_j}^{r+1} + C_{J_j\hat{J}_j}x_{\hat{J}_j}^{\ r} + \sum_{k<j} M_{J_j\hat{J}_k}x_{\hat{J}_k}^{r+1} + \sum_{k>j} M_{J_j\hat{J}_k}x_{\hat{J}_k}^{\ r} + q_{J_j}, \qquad j = 1,\ldots,m.
\end{aligned}
$$

By rewriting the last four sums as $B_{J_j\hat{J}_j}(x_{\hat{J}_j}^{r+1} - x_{\hat{J}_j}^{\ r}) + \sum_{k<j} M_{J_j\hat{J}_k}(x_{\hat{J}_k}^{r+1} - x_{\hat{J}_k}^{\ r}) + \sum_{k=1}^{m} M_{J_j\hat{J}_k}x_{\hat{J}_k}^{\ r}$, we can express the above set of equations using a single matrix splitting:

$$
0 = \begin{bmatrix}
B_{J_1J_1} & & & \\
\vdots & \ddots & & \\
M_{J_jJ_1} \ \cdots & M_{J_jJ_{j-1}} & B_{J_jJ_j} & \\
\vdots & & \ddots & \\
M_{J_mJ_1} & \cdots & M_{J_mJ_{m-1}} & B_{J_mJ_m}
\end{bmatrix} x_J^{\ r+1} +
\begin{bmatrix}
C_{J_1J_1} & \cdots & M_{J_1J_{m-1}} & M_{J_1J_m} \\
 & \ddots & & \vdots \\
 & C_{J_jJ_j} & M_{J_jJ_{j+1}} \ \cdots & M_{J_jJ_m} \\
 & & \ddots & \vdots \\
 & & & C_{J_mJ_m}
\end{bmatrix} x_J^{\ r}
$$

$$
+ \begin{bmatrix}
B_{J_1\hat{J}_1} & & & \\
\vdots & \ddots & & \\
M_{J_j\hat{J}_1} \ \cdots & M_{J_j\hat{J}_{j-1}} & B_{J_j\hat{J}_j} & \\
\vdots & & \ddots & \\
M_{J_m\hat{J}_1} & \cdots & M_{J_m\hat{J}_{m-1}} & B_{J_m\hat{J}_m}
\end{bmatrix} (x_{\hat{J}}^{\ r+1} - x_{\hat{J}}^{\ r}) \quad + \quad M_{J\hat{J}}\, x_{\hat{J}}^{\ r} \quad + \quad q_J ,
$$

or equivalently,

$$
0 = Fx_J^{\ r+1} + Gx_J^{\ r} + H(x_{\hat{J}}^{\ r+1} - x_{\hat{J}}^{\ r}) + M_{J\hat{J}}\, x_{\hat{J}}^{\ r} + q_J,
$$

for suitably defined matrices F and H, with $G = M_{JJ} - F$. Fix any $x^* \in X^*$. By subtracting the identity $0 = M_Jx^* + q_J$ (cf. $J \subseteq I^*$ and Lemma 1) from the above equation and rearranging terms, we obtain

$$x_J^{r+1} - x_J^* = (I - (F)^{-1}M_{JJ})(x_J^r - x_J^*) - (F)^{-1}M_{J\bar{J}}(x_{\bar{J}}^r - x_{\bar{J}}^*) - (F)^{-1}H(x_{\bar{J}}^{r+1} - x_{\bar{J}}^r),$$
$$r = t,\ldots,s-1.$$

Now the matrix difference $F - G$ can be seen to have the form $L + E - L^T$, where L is certain strictly (block) lower triangular part of M_{JJ} and E is a block diagonal matrix whose j-th diagonal block is $B_{J_jJ_j} - C_{J_jJ_j}$. Therefore $\langle z, (F - G)z\rangle = \langle z, Ez\rangle > 0$ for all $z \neq 0$, where the strict inequality follows from the positive definite property of the $B_{J_jJ_j} - C_{J_jJ_j}$'s. This shows that (F,G) is a regular splitting of M_{JJ}. The rest of the proof then proceeds as in the proof of Lemma 8. **Q.E.D.**

Remark 5. We can also introduce under/over-relaxation in the GS-MS algorithm. More precisely, for each $j \in \{1,\ldots,m\}$, let $(B_{I_jI_j}, C_{I_jI_j})$ be a splitting of $M_{I_jI_j}$ and $\bar{\omega}_j$ be a scalar in (0,1] satisfying

$$B_{I_jI_j} - C_{I_jI_j} + (1-\bar{\omega}_j)M_{I_jI_j} \text{ is positive definite.}$$

We define $\mathcal{A}_j$ as in (4.3) (but with the above splitting) and let $\mathcal{R}_j$:X→X be the <u>under-relaxation</u> mapping corresponding to $\mathcal{A}_j$:

$$\mathcal{R}_j(x) = (1-\bar{\omega}_j)x + \bar{\omega}_j\mathcal{A}_j(x).$$

Then the under-relaxed GS-MS algorithm comprises applications of the mappings $\mathcal{R}_1, \ldots, \mathcal{R}_m$ in a cyclical manner:

$$x^{r+1} = (\mathcal{R}_m \circ \cdots \circ \mathcal{R}_2 \circ \mathcal{R}_1)(x^r), \qquad r = 0, 1, \ldots.$$

In the special case where $\bar{\omega}_1 = \ldots = \bar{\omega}_m = 1$, the above algorithm reduces to the GS-MS algorithm. We can furthermore introduce an over-relaxation mechanism at the end of each iteration:

$$x^{r+1} = (1-\omega^r)x^r + \omega^r(\mathcal{R}_m \circ \cdots \circ \mathcal{R}_2 \circ \mathcal{R}_1)(x^r), \qquad r = 0, 1, \ldots,$$

where each ω^r is chosen such that $x^{r+1} \in X$ and $\underline{\omega} \leq \omega^r \leq \bar{\omega}$. The relaxation parameters

$\underline{\omega}$ and $\bar{\omega}$ are chosen such that $0 < \underline{\omega} \le \min\{1,\bar{\omega}\}$ and $K + (1-\bar{\omega})M$ is positive definite, where K is the n×n block diagonal matrix whose diagonal blocks comprise the positive definite matrices $(B_{I_jI_j} - C_{I_jI_j} + (1-\bar{\omega}_j)M_{I_jI_j})/\bar{\omega}_j$, $j = 1,\ldots,m$. We can also introduce a positive diagonal matrix in the definition of $\mathcal{A}_j$ as in (4.2b). In the special case where $m = 1$ and $\bar{\omega}_1 = 1$, this latter algorithm reduces to the algorithm (4.1a)-(4.2b). Convergence of the above algorithms can be shown by modifying the proof of Theorems 2 and 3.

We can further extend the GS-MS algorithm to allow under/over-relaxation (during the updating of each subset of coordinates), non-disjoint subsets I_j, and non-cyclic order of relaxation. This leads to the following SOR type algorithm, which we call the SOR-MS algorithm: Let $I_1, \ldots, I_m$ be a finite collection of nonempty (not necessarily disjoint) subsets of $\{1,\ldots,n\}$ whose union equals $\{1,\ldots,n\}$. For each $j = 1,\ldots,m$, we choose a splitting $(B_{I_jI_j}, C_{I_jI_j})$ of $M_{I_jI_j}$ and a $\bar{\omega}_j > 0$ satisfying

$$B_{I_jI_j} - C_{I_jI_j} + (1-\bar{\omega}_j)M_{I_jI_j} \text{ is positive definite.} \tag{4.6a}$$

We also choose a second relaxation parameter $\underline{\omega}_j$ satisfying

$$0 < \underline{\omega}_j \le \min\{1,\bar{\omega}_j\}, \tag{4.6b}$$

and define $\mathcal{P}_j : X \to X$ to be the point-to-set mapping

$$\mathcal{P}_j(x) = \{\, z \mid z = (1-\omega)x + \omega\mathcal{A}_j(x),\ z \in X, \text{ for some } \underline{\omega}_j \le \omega \le \bar{\omega}_j \,\}, \tag{4.6c}$$

where $\mathcal{A}_j : X \to X$ is the point-to-point mapping given by (4.3). The SOR-MS algorithm generates a sequence of iterates by successively applying the mappings $\mathcal{P}_1, \ldots, \mathcal{P}_m$ (but not necessarily in any fixed order):

SOR-MS Algorithm. Choose an $x^0 \in X$. Generate a sequence of vectors $x^0, x^1, \ldots$ in X by the formula

$$x^{r+1} \in \mathcal{P}_{j^r}(x^r), \qquad r = 0, 1, \ldots, \tag{4.7}$$

where $j^0, j^1, \ldots$ is some sequence of indices in $\{1,\ldots,m\}$.

We will impose the following rule on the order of coordinate relaxation (e.g. [SaS73], [HeL78]):

Almost Cyclic Rule. There exists integer $\bar{r}$ such that $\{1,\ldots,m\} \subseteq \{j^{r+1}, j^{r+2}, \ldots, j^{r+\bar{r}}\}$ for all r.

The SOR-MS algorithm can be seen to contain all of the earlier algorithms as special cases. For example, if $m = 1$, then it reduces to the algorithm (4.1a)-(4.2b). If the I_j's are disjoint, $\underline{\omega}_j = \bar{\omega}_j = 1$ for all j, and $\{j^0, j^1, \ldots\} = \{1, \ldots, m, 1, \ldots, m, \ldots\}$, then it reduces to the GS-MS algorithm. It also contains other methods as special cases. For example, if the $M_{I_jI_j}$'s are positive definite and we choose $B_{I_jI_j} = M_{I_jI_j}$ for all j, then (4.6a) is equivalent to $\bar{\omega}_j < 2$ and the SOR-MS algorithm reduces to a block SOR method considered in [Tse88; §6.2]. If furthermore the I_j's are disjoint and $\{j^0, j^1, \ldots\} = \{1, \ldots, m, 1, \ldots, m, \ldots\}$, then it reduces to the block SOR methods considered in [CGS78], [CoG78]; and if $m = n$ and $I_j = \{j\}$ for all j, then it reduces to the point SOR methods of Herman and Lent [HeL78] and of Lent and Censor [LeC80]. For another example, if $I_j = \{1,\ldots,n\}$ for all j, then it reduces to a matrix splitting algorithm that alternates amongst m matrix splittings.

We have not been able to show that the SOR-MS algorithm is convergent in the sense of Theorems 1 to 3. However, by combining the second half of the proof of Theorem 2 with the first half of the proof of Theorem 3, we can show that it is convergent in the weaker sense of Lemma 7:

Theorem 4. Let $x^0, x^1, \ldots$ denote the iterates generated by the SOR-MS algorithm (4.3), (4.6a)-(4.7) under the Almost Cyclic rule. Then $Mx^r + q \to d^*$ and $\phi(x^r) \to 0$. Moreover, $f(x^r)$ tends to the optimal value of (P) and every limit point of $\{x^r\}$ is a solution of (P).

Although the above result is not the strongest possible, it nonetheless improves upon those existing. For example, it shows, for the first time, that the algorithms considered in [Tse88; §6.2], [HeL78], [LeC80], [CGS78] and [CoG78] generate iterates that approach the solution set X^*.

5. Application to Known Methods

In this section we apply the results developed in Sections 3 and 4 to a number of well-known methods and show, for the first time, that these methods are convergent without making any additional assumption on the problem. We also extend some of these methods to incorporate over-relaxation.

Example 1 (point SOR method). Suppose that M has positive diagonal entries. Consider the following well-known point SOR method [Hil57], [Cry71], [Man84] for solving (P)

$$x_i^{r+1} = [\, x_i^r - (\frac{\alpha}{M_{ii}})(\textstyle\sum_{j<i} M_{ij}x_j^{r+1} + \sum_{j\geq i} M_{ij}x_j^r + q_i)\,]_i^+, \quad i = 1,\ldots,n,$$

where α is a relaxation parameter in (0,2) and $[\cdot]_i^+$ denotes the orthogonal projection onto the interval $[0,c_i]$. [This method can be viewed alternatively as a (cyclic) coordinate descent method with inexact line search [Tse88; §6.2].] It is easily seen that this method is a special case of the algorithm (4.1a)-(4.2b) with $\underline{\omega} = \bar{\omega} = 1$ and the following choices of (B,C) and D:

$$B = \alpha^{-1}E + L, \qquad C = (1-\alpha^{-1})E + L^T, \qquad D = \alpha E^{-1},$$

where E and L are, respectively, the diagonal and the strictly lower triangular part of M. Since $B - C = (2\alpha^{-1}-1)E + L - L^T$, which is positive definite for all $\alpha \in (0,2)$, it follows from Theorem 2 that this method is convergent. This improves upon existing results (e.g., [Cry71], [Man84], [LiP87]), which require for convergence either M be strictly copositive or that a certain Slater condition hold (all of which lead to the compactness of X^*).

Example 2 (Gradient Projection algorithms). Consider the well-known gradient projection algorithm [Gol64], [LeP65] (also see [Ber82], [BeT89], [Lue73]) applied to solve (P):

$$x^{r+1} = [\, x^r - \alpha(Mx^r + q)\,]^+,$$

with α is a positive stepsize. It is easily seen that this is a special case of the algorithm (4.1a)-(4.2b) with $\underline{\omega} = \bar{\omega} = 1$ and the following choices of (B,C) and D:

$$B = \alpha^{-1}I, \qquad C = M - \alpha^{-1}I, \qquad D = \alpha I.$$

In this case $B - C$ can be seen to be positive definite for all $\alpha < 2/\|M\|$. Hence by Theorem 1, the algorithm is convergent for all $\alpha \in (0, 2/\|M\|)$. Aganagic [Aga78] proposed a modification of the above algorithm by adding a relaxation parameter $\omega \in (0,1]$:

$$x^{r+1} = (1-\omega)x^r + \omega[\, x^r - \alpha(Mx^r + q)\,]^+.$$

This algorithm is also a special case of the algorithm (4.1a)-(4.2b) with $\underline{\omega} = \bar{\omega} = \omega$ and with (B,C) and D given as above. Hence, by Theorem 2, this algorithm is also convergent for all $\alpha \in (0, 2/\|M\|)$. [This improves on the result of Aganagic which requires M to be positive definite for convergence. Furthermore, from Theorem 2 we see that over-relaxation (i.e. $\omega > 1$) is also permissible, as long as $\alpha\omega \in (0, 2/\|M\|)$.]

Example 3 (Mangasarian's algorithm). Consider the following iterative algorithm proposed by Mangasarian [Man77] (also see [Man84], [MaD87] for applications)

$$x^{r+1} = (1-\omega)x^r + \omega[\, x^r - \alpha E(Mx^r + q + K(x^{r+1} - x^r))\,]^+,$$

where $\omega \in (0,1]$, E is an $n \times n$ positive diagonal matrix, K is an $n \times n$ matrix, and α is a positive scalar. It can be seen that the above algorithm is a special case of the algorithm (4.1a)-(4.2b) with $\underline{\omega} = \bar{\omega} = \omega$ and the following choices of (B,C) and D:

$$B = (\alpha E)^{-1} + \omega K, \qquad C = M - (\alpha E)^{-1} - \omega K, \qquad D = \alpha E.$$

Since $B - C + (1-\omega)M = 2(\alpha E)^{-1} + 2\omega K - \omega M$, it follows from Theorem 2 that the above algorithm is well defined and convergent if $2(\alpha E)^{-1} + 2\omega K - \omega M$ is positive definite, which is exactly the condition given by Mangasarian [Man77; Eq. (6)]. [However, Mangasarian only showed that each limit point of the iterates generated by the algorithm is a solution (which does not imply that a limit point exists) and did <u>not</u> show that the algorithm itself is well-defined.]

Example 4 (block SOR method). Consider the following block SOR method of Cottle, Golub and Sacher [CGS78] and of Cottle and Goheen [CoG78] (also see [CoP82]): Partition the index set $\{1,\ldots,n\}$ into m nonempty, mutually disjoint subsets $I_1, \ldots, I_m$ and assume that $M_{I_jI_j}$ is <u>positive definite</u> for all j. Choose a relaxation parameter $\bar{\omega} \in (0,2)$. Then, for any given $x^0 \in X$, the method generates a sequence of iterates $\{x^0, x^1, \ldots\}$ whereby, given x^r, a new iterate x^{r+1} is generated as follows:

> Let $z^0 = x^r$. For $j = 1, \ldots, m$, compute $\hat{z}^j$ to be the (unique) solution to the following system of nonlinear equations ($[\cdot]_i^+$ denotes the orthogonal projection onto the interval $[0,c_i]$)
>
> $$\begin{aligned} z_i &= [\, z_i - (M_i z + q_i)\,]_i^+, && \forall\ i \in I_j, \\ z_i &= z_i^{j-1}, && \forall\ i \notin I_j, \end{aligned}$$
>
> and let $z^j = (1-\omega)z^{j-1} + \omega\hat{z}^j$, where ω is the largest scalar in $(0,\bar{\omega}]$ such that $z^j \in X$. Then set $x^{r+1} = z^m$.

[This method essentially replaces the strictly lower triangular (diagonal) part of M in the point SOR method by strictly lower triangular (diagonal) blocks.] In the case when $\bar{\omega} = 1$, this method can be seen to be a special case of the GS-MS algorithm (4.3)-(4.4) with

$$B_{I_jI_j} = M_{I_jI_j}, \qquad\qquad C_{I_jI_j} = 0,$$

so that by Theorem 3 it is convergent. If $0 < \bar{\omega} \le 1$, then by Remark 5 it is also convergent. [This improves upon the results of [CGS78] and [CoG78] which require M to be positive definite for convergence. It also obviates the need for the projection step employed in [CoP82] to ensure the existence of a limit point.] In the case when $1 < \bar{\omega} < 2$ however, the convergence of this method remains unresolved. It is known to be convergent only in the weak sense of Theorem 4.

Appendix A.

Consider any m×m matrix A and any linear subspace $\mathcal{V}$ of $\Re^m$ such that $(A)^k v \to 0$ as $k \to \infty$, for all $v \in \mathcal{V}$. Then the following hold:

(a) There exist $\rho \in (0,1)$ and $\tau > 0$ such that

$$\|(A)^k v\| \le \tau (\rho)^k \|v\|, \qquad \forall\, k \ge 1, \qquad \forall\, v \in \mathcal{V}.$$

(b) For any $\delta \in (0,1]$, there exist $\rho_\delta \in (0,1)$ and $\tau_\delta > 0$ such that

$$\Big\| \prod_{h=1}^{k} ((1-\omega^h)I + \omega^h A)\, v \Big\| \le \tau_\delta (\rho_\delta)^k \|v\|, \qquad \forall\, k \ge 1, \qquad \forall\, v \in \mathcal{V},$$

for all sequences of scalars $\{\omega^1, \omega^2, \ldots\}$ in the interval $[\delta, 1]$.

Proof. Let $\mathcal{V}$ be spanned by the vectors $v^1, v^2, \ldots, v^r$. Then we can write

$$A = P^{-1} J P,$$

for some invertible (complex) matrix P and J is the Jordan canonical form of A [OrR70]. Hence, for all i,

$$(P^{-1}JP)^k v^i = P^{-1}(J)^k P v^i \to 0, \qquad \text{as } k \to \infty.$$

Let $J_1, J_2, \ldots, J_s$ be the Jordan blocks of J, and let $\lambda_1, \lambda_2, \ldots, \lambda_s$ be the corresponding (complex) eigenvalues of J. [Hence J is block diagonal with J_j as its j-th diagonal block and each J_j is upper triangular with λ_j along the diagonals.] Let

$$Pv^i = \begin{bmatrix} P_1 v^i \\ P_2 v^i \\ \vdots \\ P_s v^i \end{bmatrix}$$

be the partition of Pv in accordance with the partition of J. Then since $(J)^k Pv^i \to 0$ as $k \to \infty$, for all i, this implies that

$$(J_j)^k P_j v^i \to 0 \qquad \text{as } k \to \infty, \qquad \forall\, i, \forall\, j.$$

Now if $|\lambda_j| \geq 1$, then we have that $P_j v^i = 0$ for all i. [To see this, suppose the contrary, so that $P_j v^i \neq 0$ for some i. Let μ denote the last nonzero component of $P_j v^i$. Then it is easily seen that the corresponding component of $(J_j)^k P_j v^i$ is $(\lambda_j)^k \mu$, which converges to zero by assumption. But this can happen only when $|\lambda_j| < 1$, a contradiction.] Therefore $P_j v^i = 0$ for those j for which $|\lambda_j| \geq 1$. Let

$$\tilde{J}_j = \begin{cases} J_j & \text{if } |\lambda_j| < 1, \\ 0 & \text{if } |\lambda_j| \geq 1. \end{cases}$$

Then

$$(J)^k Pv^i = \begin{bmatrix} (J_1)^k P_1 v^i \\ (J_2)^k P_2 v^i \\ \vdots \\ (J_s)^k P_s v^i \end{bmatrix} = \begin{bmatrix} (\tilde{J}_1)^k P_1 v^i \\ (\tilde{J}_2)^k P_2 v^i \\ \vdots \\ (\tilde{J}_s)^k P_s v^i \end{bmatrix}, \qquad \forall\, i.$$

Hence $(J)^k Pv = (\tilde{J})^k Pv$ for all $v \in \mathcal{V}$, where $\tilde{J}$ is the m×m block diagonal matrix whose j-th diagonal block is $\tilde{J}_j$. This in turn implies that

$$\begin{aligned} \|(A)^k v\| &= \|P^{-1}(J)^k Pv\| \\ &= \|P^{-1}(\tilde{J})^k P^k v\| \\ &\leq \|P^{-1}\|\, \|(\tilde{J})^k\|\, \|P\|\, \|v\|. \end{aligned}$$

Since the spectral radius of $\tilde{J}$ is strictly less than one, there exists $\rho \in (0,1)$ and $\tilde{\tau} > 0$ such that $\|(\tilde{J})^k\| \leq \tilde{\tau}\,(\rho)^k$ for all $k \geq 1$. Therefore (letting $\tau = \tilde{\tau}\, \|P^{-1}\|\, \|P\|$)

$$\|(A)^k v\| \le \|P^{-1}\| \tilde{\tau}(\rho)^k \|P\| \|v\| \le \tau(\rho)^k \|v\|, \qquad \forall\, k \ge 1.$$

This proves part (a).

Now we prove part (b). Fix any integer $k \ge 1$ and any sequence of scalars $\{\omega^1,\omega^2,\ldots,\omega^k\}$ in the interval $[\delta,1]$. Since $A = P^{-1}JP$, we have

$$\prod_{h=1}^{k}((1-\omega^h)I + \omega^h A) = \prod_{h=1}^{k} P^{-1}((1-\omega^h)I + \omega^h J)P = P^{-1}\Big(\prod_{h=1}^{k}((1-\omega^h)I + \omega^h J)\Big)P.$$

Let

$$\tilde{J}_j^{\,h} = \begin{cases} (1-\omega^h)I + \omega^h J_j & \text{if } |\lambda_j| < 1, \\ 0 & \text{if } |\lambda_j| \ge 1, \end{cases}$$

and let $\tilde{J}^h$ be the m×m block diagonal matrix whose j-th diagonal block is $\tilde{J}_j^{\,h}$. Since $P_j v = 0$, for all $v \in \mathcal{V}$ and all those j's for which $|\lambda_j| \ge 1$ (cf. proof of part (a)), we see that

$$\Big(\prod_{h=1}^{k}((1-\omega^h)I + \omega^h J_j)\Big)P_j v = \Big(\prod_{h=1}^{k} \tilde{J}_j^{\,h}\Big)P_j v, \qquad \forall\, v \in \mathcal{V}, \qquad \forall\, j.$$

It then follows that

$$\Big(\prod_{h=1}^{k}((1-\omega^h)I + \omega^h J)\Big)Pv = \Big(\prod_{h=1}^{k} \tilde{J}^h\Big)Pv, \qquad \forall\, v \in \mathcal{V}.$$

Since $\delta \le \omega^h \le 1$ for all h, the spectral radius of the $\tilde{J}^h$'s are bounded away from 1. Then, by an argument similar to that used for part (a), we obtain that there exist $\rho_\delta \in (0,1)$ and $\tau_\delta > 0$ (depending on A, P and δ only) such that

$$\|P^{-1}\Big(\prod_{h=1}^{k}((1-\omega^h)I+\omega^h J)\Big)Pv\| = \|P^{-1}\Big(\prod_{h=1}^{k}\tilde{J}^h\Big)Pv\|$$

$$\leq \tau_\delta\,(\rho_\delta)^k\,\|v\|, \qquad \forall\, v\in \mathcal{V}.$$

Q.E.D.

References

[Aga78] Aganagic, M., "Iterative Methods for Linear Complementarity Problems," Technical Report SOL 78-10, Systems Optimization Laboratory, Department of Operations Research, Stanford University, Stanford, CA (1978).

[Ber82] Bertsekas, D. P., *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, NY (1982).

[BHT87] Bertsekas, D. P., Hosein, P. A. and Tseng, P., "Relaxation Methods for Network Flow Problems with Convex Arc Costs," *SIAM J. Contr. & Optim.*, 25 (1987), 1219-1243.

[BeT89] Bertsekas, D. P. and Tsitsiklis, J. N., *Parallel and Distributed Computation: Numerical Methods*, Prentice-Hall, Englewood Cliffs, NJ (1989).

[CeG73] Cea, J. and Glowinski, R., "Sur des methodes d'optimisation par relaxation," *Revue Française d'Automatique, Informatique et Recherche Operationelle*, R-3 (1973), 5-31.

[CGST86] Cook, W., Gerards, A. M. H., Schrijver, A., Tardos, É., "Sensitivity Theorems in Integer Linear Programming," *Math. Prog.*, 34 (1986), 251-264.

[CoG78] Cottle, R. W. and Goheen, M. S., "A Special Class of Large Quadratic Programs," in (O. L. Mangasarian, R. R. Meyer and S. M. Robinson, eds.) *Nonlinear Programming 3*, Academic Press, New York, NY (1978), 361-390.

[CGS78] Cottle, R. W., Golub, G. H. and Sacher, R. S., "On the Solution of Large, Structured Linear Complementarity Problems: the Block Partitioned Case," *J. Appl. Math. and Optim.*, 4 (1978), 347-363.

[CoP82] Cottle, R. W. and Pang, J.-S., "On the Convergence of a Block Successive Over-Relaxation Method for a Class of Linear Complementarity Problems," *Math. Prog. Study 17*, North-Holland, Amsterdam (1982), 126-138.

[Cry71] Cryer, C. W., "The Solution of a Quadratic Programming Problem Using Systematic Overrelaxation," *SIAM J. Contr. & Optim.*, 9 (1971), 385-392.

[DeT84] Dembo, R. S. and Tulowitzki, U., "On the Minimization of Quadratic Functions Subject to Box Constraints," School of Organization and Management and Department of Computer Science, Yale University, New Haven, CT (1983; revised 1984).

[D'Es59] D'Esopo, D. A., "A Convex Programming Procedure," *Naval Res. Logist. Quart.*, 6 (1959), 33-42.

[Gol64] Goldstein, A. A., "Convex Programming in Hilbert Space," *Bull. Am. Math. Soc.*, 70 (1964), 709-710.

[Eav71] Eaves, B. C., "On Quadratic Programming," *Manag. Sci.*, 17 (1971), 698-711.

[FrW57] Frank, M. and Wolfe, P., "An Algorithm for Quadratic Programming," *Naval Res. Logist. Quart.*, 3 (1957), 95-110.

[HeL78] Herman, G. T. and Lent, A., "A Family of Iterative Quadratic Optimization Algorithms for Pairs of Inequalities, with Application in Diagnostic Radiology", *Math. Prog. Study 9*, North-Holland (1978), 15-29.

[Hil57] Hildreth, C., "A Quadratic Programming Procedure," *Naval Res. Logist. Quart.*, 4 (1957), 79-85; see also "Erratum," *Naval Res. Logistic Quart.*, 4 (1957), 361.

[KiS80] Kinderlehrer, D. and Stampacchia, G., *An Introduction to Variational Inequalities and Applications*, Academic Press, New York, NY (1980).

[LeC80] Lent, A. and Censor, Y., "Extensions of Hildreth's Row-Action Method for Quadratic Programming," *SIAM J. Contr. & Optim.*, 18 (1980), 444-454.

[LeP65] Levitin, E. S. and Poljak, B. T., "Constrained Minimization Methods," *Z. Vycisl. Mat. i Mat. Fiz.*, 6 (1965), 787-823. English translation in *USSR Comput. Math. Phys.*, 6 (1965), 1-50.

[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.

[Lue73] Luenberger, D. G., *Linear and Nonlinear Programming*, Addison-Wesley, Reading, MA (1973).

[Man77] Mangasarian, O. L., "Solution of Symmetric Linear Complementarity Problems by Iterative Methods," *J. Optim. Theory & Appl.*, 22 (1977), 465-485.

[Man84] Mangasarian, O. L., "Sparsity-Preserving SOR Algorithms for Separable Quadratic and Linear Programming," *Comput. & Oper. Res.*, 11 (1984), 105-112.

[MaD87] Mangasarian, O. L. and De Leone, R., "Parallel Successive Overrelaxation Methods for Symmetric Linear Complementarity Problems and Linear Programs," *J. Optim. Theory & Appl.*, 54 (1987), 437-446.

[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.

[OrR70] Ortega, J. M. and Rheinboldt, W. C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, NY (1970).

[Pan79] Pang, J.-S., "On Q-Matrices," *Math. Prog.*, 17 (1979), 243-247.

[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.

[Pan86] 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.

[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).

[SaS73] Sargent, R. W. H. and Sebastian, D. J., "On the Convergence of Sequential Minimization Algorithms," *J. Optim. Theory & Appl.*, 12 (1973), 567-575.

[Tse88] Tseng, P., "Dual Ascent Methods for Problems with Strictly Convex Costs and Linear Constraints: A Unified Approach," LIDS Report P-1792, M.I.T., Cambridge, MA (July 1988); to appear in *SIAM J. Contr. & Optim.*

[Tse89] Tseng, P., "Further Applications of a Splitting Algorithm to Decomposition in Variational Inequalities and Convex Programming," LIDS Report P-1866, M.I.T., Cambridge, MA (April 1989).