Sindh Univ. Res. Jour. (Sci.Ser.), Vol: 36 (1) 49 – 66 (2004)

A PARALLEL ALGORITHM FOR SOLVING A NONLINEAR PROBLEM APPLICABLE TO MIMD SYSTEMS

S. Memon and S. Jalbani

Institute of Mathematics and Computer Science, University of Sindh, Jamshoro, Sindh, Pakistan.

Abstract

In this paper we present asynchronous parallel algorithms for the solution of nonlinear systems of special form using the principles of system decomposition and matrix splitting , which are suitable for MIMD parallel processors. The convergence proof of the algorithms are given and some special cases discussed.

Keywords: Parallel algorithm, Nonlinear problem, MIMD systems

Introduction

The finite difference or finite element discretisation of many nonliear differential equations usually results in solving a system of nonlinear algebraic equations with the special form.,

$$B\,T(x) + A\,H(x) = b, \tag{1.1}$$

Where,

$T(x) = (T_i(x_i)) \in R^n, T_i : R^n \longrightarrow R^n$, and $H(x) = (H_i(x_i)) \in R^n, H_i : R^n \longrightarrow R^n$,

are diagonal continuous mapping, $A = (a_{ij}) \in L(R^n)$ is an M-matrix, $B = (b_{ij}) \in L(R^n)$ is a nonnegative matrix, $x \in R^n$ is an unknown vector, $b \in R^n$ is a known vector.

There are many papers [1-7] which discuss methods for solving (1.1) (including its special cases). The most relevant to (1.1) is R. E. white [7] designed parallel algorithms based on the Gauss seidel splitting which are applicable to SIMD systems and which provide an efficient reform to previous sequential algorithms as demonstrated by numerical tests.

However, this approach introduces a class of synchronous parallel algorithms and it is important to determine how to design asynchronous parallel algorithms for the rapidly developing multiprocessor (MIMD) systems of today.

In this paper, with the aid of Tarazi s general definition of asynchronous algorithms combined with the characteristics of problem (1.1) we design a class of asynchronous parallel algorithms with a certain generality which are applicable to MIMD SYSTEM, using

the approach of system decomposition and matrix splitting and under the same conditions as that of [7], we prove the existence and uniqueness of the solution of (1.1) and the convergence of asynchronous parallel algorithms in R^n. We prove that this class of asynchronous parallel algorithm is theoretically reliable. Furthermore, under several particular matrix splittings, we discuss the convergence of the corresponding asynchronous algorithms for the purpose of practical applications.

2. Establishment of asynchronous algorithms

consider the product space,

$$R^n = \prod_{i=1}^{\alpha} R^{n_i}, R^{n_i} \subset R^n, n_1 + n_2 + \ldots + n_\alpha = n.$$

It is clear that for each $x \in R^n$, it can be expressed as.

$$x = [x_1, x_2, \ldots, x_\alpha]^T, x_i \in R^{n_i}, i = 1, 2, \ldots \alpha.$$

Let the number set $S = \{1,2, \ldots, n\}$ be divided into α nonempty and disjointed subsets, denoted $s_1, s_2, \ldots, s_\alpha$ respectively, i.e.,

$$S = \{s_1, s_2, \ldots, s_\alpha\} = \{1,2, \ldots, \alpha\}$$

Thus , matrices A, B and vectors T, H can correspondingly be divided as,

$$A = (A_{k\ell}),\ A_{k\ell} = (a_{ij})_{n_k \times n_\ell}, i \in s_k, j \in s_\ell$$

$$B = (B_{k\ell}),\ B_{k\ell} = (b_{ij})_{n_k \times n_\ell},\ i \in s_k, j \in s_\ell$$

$$T(x) = \left(T_1(x_1), T_2(x_2), \ldots, T_\alpha(x_\alpha) \right)^T, \qquad H(x) = \left(H_1(x_1), H_2(x_2), \ldots, H_\alpha(x_\alpha) \right)^T$$

Though the above division, we can express the system of equations (1.1) as the form submatrices in which each subsystem is of the form,

$$A_{kk} H_k(X_k) + B_{kk} T_k(X_k) = b_k - \sum_{\ell \neq k} A_{k\ell} H_\ell(x_\ell) - \sum_{\ell \neq k} B_{k\ell} T_\ell(x_\ell), \forall k \in s \quad (2.1)$$

In order to construct an asynchronous parallel algorithm with certain generality, we assume that the submatrices A_{kk} are of regular splittings, that is,

$$A_{kk} = M_{kk} - R_{kk},$$

Where,

$$M^{-1}_{kk} \geq 0, R_{kk} \geq 0.$$

Moreover, the splitting of submatrices B_{kk} are,

$B_{kk} = U_{kk} + V_{kk}$, $U_{kk} \geq 0$, $V_{kk} \geq 0$.

Replacing A_{kk}, B_{kk} in (2 . 1) by (2 . 2) and (2 . 3), we obtain,

$$M_{kk} H_k (x_k) + U_{kk} T_k (x_k) = b_k - \sum_{\ell \neq k} A_{k\ell} H_\ell (X_\ell) - \sum_{\ell \neq k} B_{k\ell} T_\ell (X_\ell) + R_{kk} H_k (X_k)$$

$$- V_{kk} T_k (x_k), \forall k \in \check{S} \quad (2 . 4)$$

Thus, we can establish an asynchronous parallel algorithm based on (2.4). For this purpose, we need to introduce the following symbols:

(i) $\forall \rho \in N_O = \{ 0, 1, \ldots\}$, denote $J = \{J(\rho)\}_{\rho \in N_O}$ as the nonempty subset of the set $\{1, 2, \ldots, \alpha\}$,

(ii) denote $\check{S} = \{ s_1 (p), \ldots, s_\alpha (p)\}_{\rho \in N_O}$ as infinite sequences.

The sets J and $\check{S}$ have the following properties:

(a) $\forall k \in = \{ 1, 2, \ldots, \alpha\}$, the set $\{ \rho \in N_O \mid k \in J (\rho)\}$ is infinite.

(b) $\forall k \in = \{ 1, 2, \ldots, \alpha\}$, $\forall \rho \in N_O$, $s_k (\rho) \leq \rho$.

(c) $\forall k \in = \{ 1, 2, \ldots, \alpha\}$, $\lim_{p \to \infty \alpha} s_k (\rho) = \infty$

let $x^o \in R^n$ be any initial approximation of the solution of (1.1) and assume that we have obtained,

$\{x^o, x^1, \ldots, x^p\} \subset R^n$, by any algorithm, then the (ρ + 1)th approximation x^{p+1} can be calculated by the following formula,

$$(I) = \begin{cases} M_{kk} H_k (x_k^{p+1}) + U_{kk} T_k (x_k^{p+1}) = b_k - \sum_{\ell \neq k} A_{k\ell} H_\ell (x_\ell^{s_\ell (p)}) \\ \qquad - \sum_{\ell \neq k} B_{k\ell} T_\ell (x_\ell^{s_\ell (p)}) + R_{kk} H_k (x_k^{s_k (p)}) - V_{kk} T_k (x_k^{s_k (p)}), \; k \in J (\rho) \\ x_k^{p+1} = x_k^p, \qquad k \in J (\rho) \end{cases}$$

where $S_i (\rho)$ is defined by (ii) . J (ρ) in (I) embodies the definite part. It implies the index set of corrections at pth step. Then (I) is an asynchronous parallel algorithm to (1.1) this is an implicit algorithm because the left side of (I) is nonlinear mapping with a particular form for which the right side is known. Hence, we need to solve a system of nonlinear equations in order to obtain x_k^{p+1}. At this moment, in order to improve the efficiency, we can make use of

various methods which are applicable to the corresponding subsystem. Therefore, this class of algorithms is of great flexibility.

Particularly, when,

$J(\rho) = \{1, 2, \ldots, \alpha\}, \forall \rho \in N_0$

$S_i(\rho) = \rho, \forall i \in J(\rho), \forall \rho \in N_0$

Algorithm (I) becomes

$$(II)=\begin{cases} M_{kk} H_k (x_k^{p+1}) + U_{kk} T_k (x_k^{p+1}) = b_k - \sum_{\ell \neq k} A_{k\ell} H_\ell (X_\ell^P) - \sum_{\ell \neq k} B_{k\ell} T_\ell (x_\ell^P) \\ \qquad + R_{kk} H_k (x_k^p) - V_{kk} T_k (x^p{}_k), \forall k \in J(\rho) \end{cases}$$

This is a synchronous parallel algorithm. The nonlinear parallel algorithm presented in [4] is a special case of algorithm (II).

The working principle of algorithm (I), (II) is shown in the following figu (figure. 1, figure. 2).

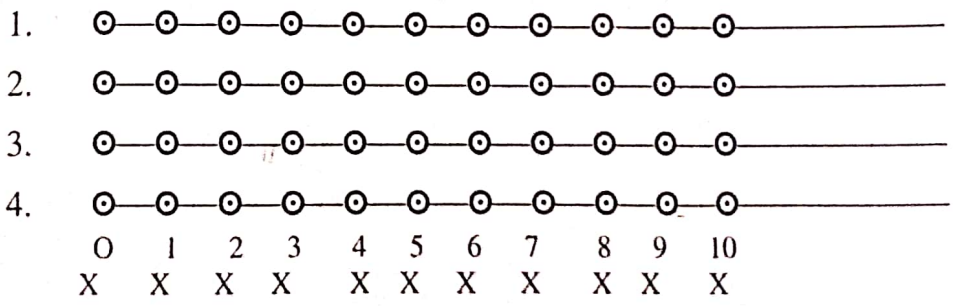


Figure .1

Suppose the problem is divided into four subsystems, and they are calculated in four processors, respectively. "." denotes the correct values and "0" denote the translation values, then by the definition of $J(\rho)$ we have ,

$J(0) = \{1\}, J(1) = \{2\}, J(2) = \{3\}, J(4) = \{1\}, J(4) = \{2\},$

$J(5) = \{1\}, J(6) = \{3\}, J(7) = \{2\}, J(8) = \{1\}, J(9) = \{4\}.$

From above we know that at the 10^{th} iteration the four subsystems all produce now information, that is, the four subsystems have been corrected at least once. At this moment, we have,

$U\ J(\rho) = \{1, 2, 3, 4, \}$

For algorithm (ii), we have,

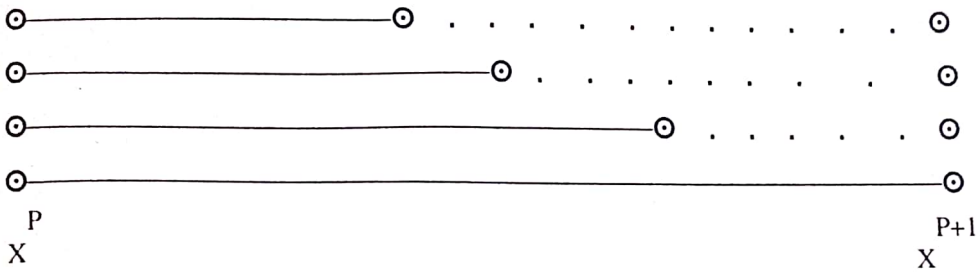


Figure. 2

Where the dotted lines denote the waiting time. Figure 1 and 2 reflect the different effect of the two algorithms.

According to the working principle in Figure 1, in order to guarantee the convergence of algorithm (I), we need to make the following assumptions to algorithm (I):

$\{m_\ell\}_{\ell \in N_o}$ is an infinite number sequence. It is generated by the following rule:

m_o is the least positive integer such that

$$U\ J(\rho) = \{1, 2, \ldots, \alpha\};$$

$$0 \le s(p) \le p < m_{\ell+1}$$

In general , $m_{\ell+1}$ is the least positive integer such that

$$U\ J(\rho) = \{1, 2, \ldots, \alpha\};$$

$$m_\ell \le s(p) \le p < m_{\ell+1}$$

where $S(\rho) = \min_i S_i(\rho)$.

The actual significance of the number sequence $\{m_\ell\}_{\ell \in N_o}$ is that when the times of correction of every subsystem reaches some predetermined number, the iterative index goes to $m_{\ell+1}$ from m_ℓ. For instance, as shown in Figure 1, we can suppose that every subsystem is corrected at least once, the iterative index goes to $m_{\ell+1}$ from m_ℓ and $m_{\ell+1} - m_\ell = 1$ for all ℓ.

3. Convergence of the algorithm

Let A be an M-matrix and $A_{kk} = M_{kk} - R_{kk}$ be the regular splitting of A_{kk}. Define the matrix,

$$M = (M_k\ell)_{n\ x\ n}$$
$$\alpha \quad \alpha$$

Where ,

$$M_k\ell = \begin{cases} 0, & k \neq \ell \\ M_{kk}, & k = \ell \end{cases}$$

It is clear that M is a block diagonal matrix. These matrices contain all the diagonals of A. we also define,

$$R = (R_k\ell)_{n\ x\ n}$$
$$\alpha \quad \alpha$$

Where ,

$$R_k\ell = \begin{cases} -A_k\ell, & \ell \neq k \\ R_{kk}, & \ell = k \end{cases}$$

It is clear that R is a nonnegative matrix.

From the above two definitions, we readily establish the following lemmas.

Lemma 1: Let A be an M-matrix and $A_{kk} = M_{kk} - R_{kk}$ be the regular splitting of A then A = M – R is also the regular splitting of A, and

$$M^{-1} = (\bar{M}_k\ell), = \begin{cases} 0, & k \neq \ell \\ M_{kk}^{-1} & k = 1 \end{cases}$$

The proof is direct from the definitions of M, R and the multiple role of matrices.

Lemma 2: Let A = M – R be the regular splitting of A, the $A^{-1} \geq 0$ if and only if $\rho(M^{-1}R) < 1$. This is a known conclusion see [2].

Lemma 3. Let $P \in L(R^n)$ be a nonnegative matrix and $\rho(P) < 1$, then for any $\epsilon > 0$ there exists a monotone norm $\|.\|_\epsilon$ such that,

$$\|P\|_\epsilon < \rho(P) + \epsilon$$

Proof: we only need to prove that there exists a nonsingular diagonal and nonnegative matrix E such that,

$$\|E^{-1}PE\|_\infty < \rho(P) + \epsilon,$$

that is,

$E^{-1} P E e < \rho (P) + \epsilon) e, e^T = [1, \ldots, 1]$

Note the equation,

$$(\rho (P) + \epsilon) e - E^{-1} P E e = (\rho (P) + \epsilon) E^{-1} (I - \frac{P}{\rho (P) + \epsilon}) E e.$$

Since $P \geq 0, \rho (\frac{P}{\rho (P) + \epsilon}) < 1, \forall \epsilon > 0$, then,

$$(I - \frac{P}{\rho (P) + \epsilon})^{-1} \geq 0$$

Let E be a nonsingular diagonal matrix such that,

$$E e = (I - \frac{P}{\rho (P) + \epsilon})^{-1} e, E e > 0,$$

then,

$$(\rho (P) + \epsilon) E^{-1} (I - \frac{P}{\rho (P) + \epsilon}) E e = (\rho (P) + \epsilon) E^{-1} e > 0$$

Therefore

$$(\rho (P) + \epsilon) e > E^{-1} P E e$$

Lemma 4: Let $A_1 \in L (R^n)$ be an M-Matrix, $A_1 = D_1 + B_1$, where the diagonals of B_1 are zero. Suppose that D_2, B_2 are any matrices such that,

$D_2 \geq 0, 0 \leq B_2 \leq B_1$

Then,

$A = D_1 + D_2 - (B_1 - B_2)$,

is an M-Matrix.

Proof: see [1].

We now introduce the definition of norm.

Let E be a positive diagonal matrix, then for all $x \in R^n$ we define

$$\| x \|_E = \| E^{-1} x \|_\infty ;$$

For any $x1 \in R^{n_i}$, $i = 1, 2, \ldots, \alpha$, we define,

$$\| x_i \|_i = \| E_{ii}^{-1} x_i \|_\infty, \| x \|_E = \max_i \| x_i \|_i.$$

We first prove the existence and the uniqueness of the solution of (1.1).

Theorem 1: Suppose that (1.1) satisfies:

I. A is an M-matrix, B is a nonnegative matrix and $b_{ii} > 0$, $\forall$ i,

II. H: $R^n \longrightarrow R^n$ is a diagonal, isotone and continuous mapping

III. T: $R^n \longrightarrow R^n$ is a diagonal and continuous mapping and any x, y $\in R^n$ satisfy,

$$T(x) - T(y) \geq \epsilon (x - y) > 0, \epsilon = \begin{bmatrix} \epsilon_1 & & 0 \\ & \ddots & \\ 0 & & \epsilon_n \end{bmatrix} > 0,$$

where x, y are comparable,

IV. $\dfrac{b_{ij}}{b_{ii}} \leq \dfrac{|a_{ij}|}{a_{ii}}$, $\forall$ i, j.

Then for any b, (1.1) exists a unique solution $x^* \in R^n$

Proof : First of all, we decompose the matrix A, B as follows:
$A = D - L - W$, $B = \bar{D} - \bar{L} - \bar{W}$,

where D, $\bar{D}$ are diagonal matrices, L , $\bar{L}$ are strictly lower triangular matrices,

W, $\bar{W}$ are strictly upper triangular matrices. Let

$F(x) = DH(x) + \bar{D}\,T(x)$, $G(x) = (L + W)\,H(x) + (\bar{L} + \bar{W})\,T(x)$.
For any comparable x, y $\in R^n$, calculate,

$F(x) - F(y) = D\,[H(x) - H(y)] + \bar{D}\,[T(x) - T(y)]$

$G(x) - G(y) = (L+W)\,[H(x) - H(y)] + (\bar{L} + \bar{W})\,[T(x) - T(y)]$.

We have,

$$|G(x) - G(y)| \leq |L + W|\,|H(x) - H(y)| + |\bar{L} + \bar{W}|\,|T(x) - T(y)| \qquad (3.1)$$

form assumptions (ii) , (iii) , we have ,

sign [H (x) - H (y)] = sign [T (x) – T (x) – T(y)],

then,

$$|F(x) - F(y)| = D\,|H(x) - H(y) - H(y)| + \bar{D}\,|T(x) - T(y)|,$$

$$|H(x) - H(y)| = D^{-1}\,|F(x) - F(y)| + D^{-1}D\,|T(x) - T(y)|.$$

Combining the above equations and (3 . 1), we have,

$$|G(x) - G(y)| \leq (L+W)D^{-1}|F(x) - F(y)| - (L+w)D^{-1}\bar{D}|T(x) - T(y)| + |\bar{L}+\bar{W}|T(x) - T(y)| \qquad (3.2)$$

Form assumptions (iv) we have ,

$$(L+W)D^{-1} \geq + |\bar{L}+\bar{W}|D^{-1}$$

Then from (3.2),

$$|G(x) - G(y)| \leq (L+W)D^{-1}|F(x) - F(y)| \qquad (3.3)$$

Moreover from assumptions (I) – (iii), for any given $z \in R^n$, the system of nonlinear equation, $F(x) = G(z)$

Always possesses a unique solution because F (x) is an M –function in R^n.

From the above fact, for any initial point $x^0 \in R^n$, the iterative process,

$$F(x^{k+1}) = G(x^k),\ k = 0, 1, 2, \ldots \qquad (3.4)$$

Is well defined.

Hence, for any iterative index k, from (3 . 3) we have,

$$|G(x^{k+1}) - H(x^k)| \leq (L+W)D^{-1}|F(x^{k+1}) - F(x^k)|$$

that is,

$$|F(x^{k+2}) - F(x^{k+1})| \leq (L+W)D^{-1}|F(x^{k+1}) - F(x^k)|$$

it follows that,

$$D^{-1}|F(x^{k+2}) - F(x^{k+1})| \leq D^{-1}(L+W)D^{-1}|F(x^{k+1}) - F(x^k)| \qquad (3.5)$$

Since $A=D-(L+W)$ is a regular splitting of A, then the spectral radius of

$P = D^{-1}(L+W)$ is less then 1, i .e,

$\rho(P) , < 1.$

Thus, from Lemma 3, there exists a monotone norm $\|\ .\|_E$ such that,

$\|P\|_E < 1.$

Taking norm on both sides of (3 . 5) we have,

$$\|D^{-1}|F(x^{k+2}) - F(x^{k+1})|\ \|_E \leq \|P\|_E \|D^{-1}|F(x^{k+1}) - F(x^k)|\ \|_E \leq \|P\|_E^{k+1}$$

$$\| D^{-1} | F(x^1) - F(x^0) | \|_E$$

Hence, the sequence $\{ D^{-1} F(x^k) \}_{K \in N_0}$ is a cauchy sequence. Since,

$$| F(x) - F(y) | \geq \bar{D} | T(x) - T(y) | \geq \bar{D} \epsilon | x - y |,$$

Then

$$\| D^{-1} \bar{D} \epsilon | x^{k+2} - x^{k+1} | \|_E \leq \| P \|_E^{k+1} \| D^{-1} | F(x) - F(x^0) | \|_E,$$

Thus sequence $\{x^k\}_{K \in N_0}$ is also a Cauchy sequence. Hence, there exists $x^* \in R^n$, such that,

$$\lim_{k \to \infty} x^k = x^*$$

By the continuity of F and G, we have
$f(x^*) = G(x^*)$.

From the definitions of F and G, it follows that x^* satisfies (1 . 1).
Finally, it is easy to prove that $x^* \in R^n$ is a unique solution of (1.1) from (3 . 3). ■

This theorem proved that under the conditions (I) – (iv), the solution of (1.1) exists and is unique in R^n. such conditions are satisfied for many practical problems, particularly, for systems of discrete equations of nonlinear differential equations.

Now we prove the convergence of the algorithm. For this purpose we introduce the following symbols.

$D = \text{diag}(A)$, $\bar{D} = \text{diag}(B)$,

The corresponding block forms are

$$D = (D_{kk})_{\substack{n \\ \alpha}} \times_{\substack{n \\ \alpha}}, \quad D_{kk} = \text{diag}(A_{kk}),$$

$$\bar{D} = (\bar{D}_{kk})_{\substack{n \\ \alpha}} \times_{\substack{n \\ \alpha}}, \quad \bar{D}_{kk} = \text{diag}(B_{kk}),$$

For convenience, we also intrduce the following symbols,

$$C_k = (C_k \ell), = (C_k \ell), = \begin{cases} -A_k \ell, & \ell \neq k \\ R_{kk}, & \ell = k \end{cases}$$

$$\bar{C}_k = (\bar{C}_k \ell), = (\bar{C}_k \ell), = \begin{cases} -B_k \ell, & \ell \neq k \\ V_{kk}, & \ell = k \end{cases}$$

Theorem 2: suppose the assumptions (I) - (iv) of Theorem 1 are satisfied.

Moreover ,

(v) For k = 1, . . . , α , there holds,

$(D_{kk} - A_{kk}) \geq R_{kk} \geq 0,\ (B_{kk} - \bar{D}_{kk}) \geq V_{kk} \geq 0$

$R_{kk}\, D_{kk}^{-1} \geq V_{kk}\, \bar{D}_{kk}^{-1},\ (U_{kk} - \bar{D}_{kk})\, \bar{D}_{kk}^{-1} \leq (D_{kk} - M_{kk})\, D_{kk}^{-1}$

Then for any $x^{o} \in R^{n}$ the asynchronous parallel algorithm defined in (I)

Section 2 is well posed and the sequence $\{ x^{p} \}_{p \in N_o}$ generated by algorithm (I)

converges to the unique solution $x^* \in R^{n}$ of (1.1).

Proof : For $k \in J(p)$, by Theorem 1, the corresponding subsystem,

$F_k (x_k) = G_k (z)_k,\ x_k,\ z_k \in R^{n_k}$

Where z_k is any given vector, there exists a unique solution $x_k^*,\ z_k^* \in R^{n_k}$

because F_k is an M-Function from the assumptions of the theorem. Hence , algorithm (I) is

well posed.

Since A = M – R is a regular splitting of A, A is an M-matrix and

D= diag (A), then,

$AD^{-1} = MD^{-1} - RD^{-1}$

Is also a regular splitting of $A\, D^{-1}$. From Lemma 2 and 3, we have that for any $\epsilon > 0$ there exists a diagonal matrix E > 0 and the corresponding norm $\| . \|_E$, such that.

$$\bar{q} = \| D M^{-1} R D^{-1} \|_E < 1.$$

Let,

$F(x) = D H(x) + D T(x),$

and x^o is any given initial approximation such that,

$$\| F(x^0) - F(x^*) \|_E = r.$$

Then we can verify that for any $P \in N_o$,

$$\| F(x^p) - F(x^*) \|_E \leq r. \qquad \textbf{(3 . 6)}$$

In fact, for $k \in J(p)$, algorithm (I) can be expressed as,

$$M_{kk} H_k (x_k^{p+1}) + U_{kk} T_k (x_k^{p+1}) = b_k + c_k H (v) - \bar{C}_k T (v), \qquad (3.7)$$

Where

$$V^T = [x_1^{s_1(p)}, \ldots, x_\alpha^{s_\alpha(p)}],$$

Since $x^* \in R^n$ is a unique solution of (1.1), then

$$M_{kk} H_k (x^*_k) + U_{kk} T_k (x^*_k) = b_k + c_k H (x^*) - \bar{C}_k T (x^*),$$

From **(3.7)** and **(3.8)**, we obtain,

$$M_{kk} [H_k (x_k^{p+1}) + H_k (x^*_k)] + U_{kk} [T_k (x_k^{p+1}) - T_k (x^*_k)] = c_k [H (v) - H (x^*)] - \bar{C}_k [T (v) - T (x^*)]$$

By definition of F, we have,

$$F_k (x_k^{p+1}) - F_k (x^*_k)] = C_k [H (v) - H (x^*)] - C_k [T(v) - T(x^*)] + (D_{kk} - M_{kk}) [H_k (x_k^{p+1}) - H_k (x^*_k)] + (D_{kk} - U_{kk}) [T_k (x_k^{p+1}) - T_k (x^*_k)]$$

Then

$$| F_k (x_k^{p+1}) - F_k (x^*_k)] \le C_k | H (v) - H (x^*)| - \bar{C}_k | T(v) - T(x^*) | + | D_{kk} - M_{kk} | | H_k (x_k^{p+1}) - H_k (x^*_k)| + | \bar{D}_{kk} - U_{kk} | | T_k (x_k^{p+1}) - T_k (x^*_k) | \qquad (3.9)$$

By the assumption $D_{kk} - M_{kk} \ge 0$, $\bar{D}_{kk} - U_{kk} \le 0$ and the fact for any $x, y \in R^n$

It holds,

$$|F(x) - F(y)| = D|H(x) - H(y)| + \bar{D}|T(x) - T(y)|,$$

$$|H(x) - H(y)| = D^{-1}| F(x) - F(y)| + D^{-1}\bar{D}| T(x) - T(y)|,$$

(3.9) becomes,

$$| F_k (x_k^{p+1}) - F_k (x^*_k)] \le C_k D^{-1} | F (v) - F (x^*)| - C_k D^{-1} \bar{D} | T(v) - T(x^*) | + \bar{C}_k | T (v) - T (x^*) | + (D_{kk} - M_{kk}) D_{kk}^{-1} - | F_k (x_k^{p+1}) - F_k (x^*_k) | - (D_{kk} - M_{kk}) D_{kk}^{-1} \bar{D}_{kk} | T_k (x_k^{p+1}) - T_k (x^*_k) | + | \bar{D}_{kk} - U_{kk} | T_k (x_k^{p+1}) T_k (x^*_k) | \qquad (3.10)$$

By the conditions,

$$\bar{C}_k \le C_k D^{-1} \bar{D}, (U_{kk} - \bar{D}_{kk}) \bar{D}_{kk}^{-1} \le (D_{kk} - M_{kk}) D_{kk}^{-1},$$

(3.10) can be simplified as,

$$| F_k (x_k^{p+1}) - F_k (x^*_k)] \le C_k D^{-1} | F(v) - F(x^*) | + (D_{kk} - M_{kk}) D_{kk}^{-1} | F_k (x_k^{p+1}) - F_k (x^*_k) |$$

or,

$$| F_k (x_k^{p+1}) - F_k (x^*_k)] \le D_{kk} M_{kk}^{-1} C_k D^{-1} | F(v) - F(x^*) |$$

since the matrix $D_{kk} M_{kk}^{-1} C_k D^{-1}$ is a submatrix of $D M^{-1} R D^{-1}$, then,

$$\| F_k (x_k^{p+1}) - F_k (x^*_k) \|_k \le q \| F(v) - F(x^*) \|_E \qquad (3.11)$$

Thus, we can prove the inequality (3.6) by induction. In fact for P=0 the conclusion is obviously true. Suppose that for $0 \le i \le P$, there holds,

$$\| F(x^I) - F(x^*) \|_E \le r.$$

When $i = p + 1$, for $k \in J(p)$, from (3 . 11) we have,

$$\| F_k (x_k^{p+1}) - F_k (x^*_k) \|_k \le \bar{q} \| F(v) - F(x^*) \|_E.$$

since for any $\ell \in \{ 1, 2 \ldots., \alpha \}$ S $\ell(p) \le P$, then,

$$\| F_\ell (x_\ell^{s\ell(p)}) - F_\ell (x^*_\ell) \|_\ell \le r$$

Since,

$$\| F_k (v) - F(x^*) \|_E = \max \| F_k (x_k^{s_k(p)}) - F_k (x^*_k) \|_k \le r,$$

then,

$$\| F_k (x_k^{p+1}) - F_k (x^*_k) \|_k \le \bar{q} r \qquad (3.12)$$

Moreover, for $k \bar{\in} J(p)$, $x_k^{p+1} = x_k^p$, then,

$$\| F_k (x_k^{p+1}) - F_k (x^*_k) \|_k = \| F_k (x_k^p) - F_k (x^*_k) \|_k \le r$$

Thus from (3.12), (3.13) for all $P \in N_0$, we have,

$$\| F(x^p) - F(x^*) \|_E \le r$$

From the above conclusion and assumption (iii), we have that for any $x, y \in R^n$

$$| F(x) - F(y) | \ge \bar{d} | T(x) - T(y) | \ge \bar{d} \epsilon_0 | x - y |,$$

Where $d = \min_i b_{ii}$, $\epsilon_0 = \min_i \epsilon_i$

Thus, for any $P \in N_0$,

$$\| x^p - x^* \|_E \le \frac{1}{\bar{d}. \epsilon_0} r = \bar{r}$$

Now, by induction, we can prove that for $P \geq m_\ell$

$$\| x^p - x^* \|_E \leq \bar{q}^{\ell+1} \bar{r}$$

First, for $l = 0$, by the definition of m_0, for $\forall\ p \geq m_0$ and $\forall\ k \in (1)\ \alpha$, there exists q: $0 \leq s(q) \leq q \leq p$, such that, $x_k^p = x_k^{q+1}$, $k \in J(q)$, where x_k^{q+1} is obtained from algorithm I, then,

$$\| F(v) - F(x^*) \|_E \leq r,$$

$$V^T = \begin{bmatrix} s_1(q) & s_2(q) & s_\alpha(q) \\ x_1 & , \quad x_2 & , \ldots\ x_\alpha \end{bmatrix}$$

and,

$$\| F_k(x_k^{q+1}) - F_k(x^*_k) \|_k \leq \bar{q}\, r,$$

$$\| x_k^{q+1} - x^* \|_k \leq \bar{q}\, \bar{r},\ \forall\ k \in \{1, 2, \ldots., \alpha\},$$

Thus

$$\| x^p - x^* \|_E \leq \bar{q}\ \bar{r}$$

By the induction assumption for all $p \geq m_\ell$.

$$\| x^p - x^* \|_E \leq \bar{q}^{\ell+1}\ \bar{r}.$$

We want to prove that for all $p \geq m_{\ell+1}$, the conclusion is true.

From the definition of $m_{\ell+1}$, for any $p \geq m_{\ell+1}$, $\forall\ k \in \{1, 2, \ldots, \alpha\}$

there exists q: $m_\ell \leq s(q) \leq q < p$, such that,

$x_k^p = x_k^{q+1}$, $k \in J(p)$,

Where x_k^{q+1} Is obtained from algorithm (I). Thus , by the inductive assumption, it is also valid for x_k^{p+1} that,

$$\| F_k(x_k^p) - F_k(x^*_k) \|_k = \| F_k(x_k^{q+1}) - F_k(x^*_k) \|_k \leq q \| F(v) - F(x^*) \|_E,$$

$$v^T = \begin{matrix} s_1(q) & & s_\alpha(q) \\ (x_1 & , , \ldots & x_\alpha) \end{matrix}$$

Since for all $k \in \{1, 2, \ldots \alpha\}$, $S(q) \leq S_k(q)$, then $m_\ell \leq s_k(q)$ thus

$$\| F_k(x_k^{s_1(q)}) - F_k(x^*_k) \|_k \leq \bar{q}^{\ell+1} r$$

or,

$$\| F(v) - F(x^*) \|_E \leq \bar{q}^{\ell+1} r$$

Combining (3.15) and the above inequality we obtain,

$$\| F_k (x_k^p) - F_k (x^*_k) \|_k \leq \bar{q}^{\ell+2} r$$

Thus from (3.14) we have,

$$\| x^p - x^* \|_E \leq \bar{q}^{\ell+2} \bar{r}$$

and the inductive process is completed. Therefore, under the assumption of the theorem, the sequence $\{x^p\}_{p\in N_0}$ generated by algorithm (I) satisfies,

$$\lim_{p \to \infty} x^p = x^*$$

The proof of the convergence of algorithm (I) is complete. ■

It is easy to see from the proof that the convergence rate is atleast linear.

Under the same conditions as theorem 2, the conclusion is also true for algorithm (II). The proof is implicit in the proof of Theorem 2 and we omit it here.

4. Special cases

Not the subsystem (2.1). if we take the splitting of A_{kk} as follows:

$$A_{kk} = D_{kk} - L_{kk} - R_{kk} = M_{kk} - R_{kk} \qquad (4.1)$$

Where $D_{kk} = \text{diag}\,(A_{kk})$, L_{kk} is a strictly lower triangular matrix, R_{kk} is a strictly upper triangular matrix. Since A is an M-matrix, then A_{kk} is also an M-matrix.
In (4.1), $R_{kk} \geq 0$ and for $M_{kk} = D_{kk} - L_{kk}$, $M_{kk} \geq 0$. Thus, (4.1) is a regular splitting of A_{kk}. If we take the corresponding splitting of B_{kk} as, $B_{kk} = (\bar{D}_{kk} - \bar{L}_{kk}) - \bar{R}_{kk} = U_{kk} + V_{kk}$ **(4.2)**

Then the corresponding algorithm (I) becomes,

$$(III) = \begin{cases} (D_{kk} - L_{kk})\, H_k (x_k^{p+1}) + (D_{kk} - L_{kk})\, T_k (x_k^{p+1}) = b_k + \sum_{\ell \neq k} (A_{k\ell} H_\ell (x_\ell^{s_\ell(p)}) \\ + \sum_{\ell \neq k} (-B_{k\ell}) T_\ell (x_\ell^{s_\ell(p)}) + R_{kk} H_k (x_k^{s_k(p)}) + R_{kk} T_k (x_k^{s_k(p)}) \qquad k \in J(\rho) \\ x_k^{p+1} = x_k^p, \qquad k \in J(\rho) \end{cases}$$

Algorithm (III) is obtained when the special splitting are taken for A_{kk}, B_{kk} and is a special case of algorithm (I) , since D_{kk} , L_{kk} , $\bar{R}_{kk}$, $\bar{L}_{kk}$, $\bar{R}_{kk}$, D_{kk} are easily known, then it is convenient for practical applications.

Moreover, under assumptions (I) - (iii) of Theorem 1, for the splitting (4.1), (4.2), the assumption (v) in Theorem 2 is superfluous. Thus, from Theorem 2, the convergence theorem of algorithm (III) is direct.

Theorem 3: suppose that assumptions (i) – (iv) of Theorem 1 are satisfied, then the asynchronous parallel algorithm (III) is well posed, and the sequence $\{x^p\}_{p \in N(0)}$ generated by algorithm (III) converges to the unique solution of (1.1).

The proof is the same as that of Theorem 2.

In section 2, we have discussed the decomposition of the system and the formation of subsystems without any rigid roles. Thus, in practical applications, we can make use of particular problems and make the best decomposition. In general, the quadrant interlocking splitting presented in [10] by Evans and Haghighi is valid to the theorems and algorithms of this paper. Certainly, we can take other decompositions. However, the decomposition (4.1) is the basic one, and it is easy to apply.

Here, we also give a synchronous from of algorithm (III).

$$(D_{kk} - L_{kk})\, H_k\,(x^{p+\frac{1}{2}}) + (\bar{D}_{kk} - \bar{L}_{kk})\, T_k\,(x^{p+\frac{1}{2}}) = b_k$$

$$-\sum_{\ell \neq k} (A_{k\ell}\, H_\ell\,(x_\ell^{s_\ell(p)})) - \sum_{\ell \neq k} (-B_{k\ell})\, T_\ell\,(x_\ell^{s_\ell(p)}) + R_{kk}\, H_k\,(x_k^{s_k(p)}) + R_{kk}\, T_k\,(x_k^{s_k(p)})$$

$$\text{(IV)} \quad \begin{cases} (D_{kk} - R_{kk})\, H_k\,(x^{p+1}) + (\bar{D}_{kk} - \bar{R}_{kk})\, T_k\,(x^{p+1}) = b_k \\ -\sum_{\ell \neq k} (A_{k\ell}\, H_\ell\,(x_\ell^{s_\ell(p)})) - \sum_{\ell \neq k} B_{k\ell}\, T_\ell\,(x_\ell^{s_\ell(p)}) + L_{kk}\, H_k\,(x_k^{p+\frac{1}{2}}) + \bar{L}_{kk}\, T_k\,(x_k^{p+\frac{1}{2}}) \\ x_k^{p+1} = x_k^{p} \end{cases} \quad \begin{matrix} k \in J(p) \\ k \in J(p) \end{matrix}$$

it is easily seen that algorithm (IV) is a modification of algorithm (III) and is a generalized version of the asynchronous parallelization of algorithm NSSSI presented in [9] by Shearer and Wofe. Under the conditions of Theorem 3, we have proved that algorithm (IV) is well posed and convergent.

Finally, it is safe to predict that the computational efficiency of these algorithms is certainly better than algorithms which are applicable to single processors for solving corresponding problems.

References

El–Tarazi, M. N. (1982) Some convergence results for asynchronous Algorithms. Numer. Math. **39**: 325 – 340.

Evans, D. J, R . S. Haghighi, (1982) Parallel Iterative methods for solving linear equations. Intern, J. Comp. Math., **10**: 247 – 286.

Leary, D.P O, R. E. White (1985) Multi-splittings of matrics and parallel solution of linear system, SIAM J. Alg. Disc. Meth., **6**: 630 – 640.

Ortega, J. M, W.C. Rheinboldt, (1970) Iterative solution of nonlinear equations in several Variables. New York-London: Academic press.

Shearer, J. M, M. A. Wofe, (1985) Some algorithms for the solution of a class of Nonlinear Algebraic Equations. Computing **35**: 63 – 72.

Varga, R.S. (1962) Matrix Iterative Analysis, Prentice-Hall, Englewood cliffs. N.J.,U.S. A.

White, R. E. (1982) An enthalpy formulation of the stefan problem. SIAM J. Numer. Anal., **19**: 1129 – 1157.

Whit, R. E. (1985) The binary alloy solidification problem: existence, uniquness and numerical approximation. SIAM. J. Numer. Anal., **22**: 205 – 244.

White, R .E. (1986,a) Parallel algorithms for nonlinear problem. SIAM J . Alg. Disc. Meth., **7**: 137 – 149.

White, R. E. (1986,b) A Nonlinear parallel algorithm with application to the stefan problem. SIAM J. Numer. Anal., **23**: 639 – 652 .