Journal of Computational Mathematics, Vol.17, No.6, 1999, 575–588.

ON THE CONVERGENCE OF ASYNCHRONOUS NESTED MATRIX MULTISPLITTING METHODS FOR LINEAR SYSTEMS*1)

Zhong-zhi Bai
(*State Key Laboratory of Scientific/Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences P.O. Box 2719, Beijing 100080, China*)

De-ren Wang
(*Department of Mathematics, Shanghai University, Shanghai 201800, China*)

D.J. Evans
(*Parallel Algorithms Research Centre, Loughborough University of Technology Loughborough, U.K.*)

Abstract

A class of asynchronous nested matrix multisplitting methods for solving large-scale systems of linear equations are proposed, and their convergence characterizations are studied in detail when the coefficient matrices of the linear systems are monotone matrices and H-matrices, respectively.

Key words: Solution of linear systems, Asynchronous parallel iteration, Matrix multisplitting, Relaxation method, Convergence.

1. Introduction

There has been a lot of literature (see [1]–[6] and [12]) on the parallel iterative methods for the large-scale system of linear equations

$$Ax = b, \quad A \in L(R^n) \ \text{ nonsingular}, \quad x, b \in R^n \tag{1.1}$$

in the sense of matrix multisplitting since the pioneering work of O'Leary and White (see [1]) was published in 1985. One of the most recent result may be the studies on a class of asynchronous parallel matrix multisplitting relaxation methods proposed in [6]. These methods, just as was pointed out in [6], are suitable to the high speed multiprocessor systems (MIMD-systems). However, the method given in the paper requires each processor of the MIMD-system to solve a sub-system of linear equations at every iterative step. The computations of the solutions of these α sub-systems of linear equations then turn to the main tasks in concrete implementations of this

* Received May 26, 1997.
1)Project 19601036 supported by the National Natural Science Foundation of China.

asynchronous parallel matrix multisplitting relaxation method. Therefore, it deserves further investigation on both the method model and the convergence theory.

In this paper, through combining each iteration distributed on the corresponding processor with an inner iteration, which is used to solve its sub-system of linear equations, we construct a class of new asynchronous matrix multisplitting methods, which are called, following the customary, asynchronous nested matrix multisplitting methods. The convergence properties of these new methods are discussed in detail when the coefficient matrix $A \in L(R^n)$ is a monotone matrix as well as an H-matrix. This work can be thought of a further development of [6], and also a generalization of [9]–[10] to asynchronous matrix multisplitting methods.

For the convenience of the subsequent discussions, in the remainder of this section, we will restate the first asynchronous parallel matrix multisplitting method in [6].

We recall that a collection of triples (M_i, N_i, E_i) $(i = 1, 2, \cdots, \alpha)$ $(\alpha \leq n$, a given positive integer) is called a multisplitting of a matrix $A \in L(R^n)$ if $M_i, N_i, E_i \in L(R^n)$ $(i = 1, 2, \cdots, \alpha)$ with each E_i being nonnegatively diagonal, and satisfy: (1) $A = M_i - N_i (i = 1, 2, \cdots, \alpha)$; (2) $\det(M_i) \neq 0 (i = 1, 2, \cdots, \alpha)$; and (3) $\sum\limits_i E_i = I (I \in L(R^n)$ is the identity matrix).

Here, we have assumed that the MIMD-system considered is made up of α CPU's. Correspondingly, the following notations are also indespensable: (i) for $\forall p \in N_0 = \{0, 1, 2, \cdots\}$, $J = \{J(p)\}_{p \in N_0}$ is used to denote a sequence of nonempty subset of the set $\{1, 2, \cdots, \alpha\}$; (ii) $S = \{s_1(p), s_2(p), \cdots, s_\alpha(p)\}_{p \in N_0}$ are α infinite sequences. The sets J and S have the following properties: (a) for $\forall i \in \{1, 2, \cdots, \alpha\}$, the set $\{p \in N_0 | i \in J(p)\}$ is infinite; (b) for $\forall i \in \{1, 2, \cdots, \alpha\}, \forall p \in N_0$, it holds that $s_i(p) \leq p$; and (c) for $\forall i \in \{1, 2, \cdots, \alpha\}$, it holds that $\lim\limits_{p \to \infty} s_i(p) = \infty$.

With these preparations, the asynchronous parallel matrix multisplitting method in [6] can be described as follows:

ALGORITHM (see [6]): Suppose that we have got approximations $x^0, x^1, \cdots, x^p$ to the solution x^* of (1.1). Then the $(p+1)$-th approximation x^{p+1} of x^* can be calculated by

$$x^{p+1} = \sum_i E_i x^{i,p} \tag{1.2}$$

with $x^{i,p}$ being either x^p for $i \notin J(p)$ or the solution of the sub-system of linear equations

$$M_i x^{i,p} = N_i x^{s_i(p)} + b \tag{1.3}$$

for $i \in J(p)$.

2. Asynchronous Nested Matrix Multisplitting Methods

For the purpose of establishing our new methods, we first introduce the following concept: A collection $(M_i : F_i, G_i; N_i; E_i)$ $(i = 1, 2, \cdots, \alpha)$ is called a two-level multisplitting of a matrix $A \in L(R^n)$ if (M_i, N_i, E_i) $(i = 1, 2, \cdots, \alpha)$ is a multisplitting of it and $M_i = F_i - G_i, \det(F_i) \neq 0 (i = 1, 2, \cdots, \alpha)$. Based on this concept, by solving each of the sub-system of linear equations (1.3) with an inner iteration method again, we can set up the following asynchronous nested matrix multisplitting method for the system of linear equations (1.1):

METHOD I: Suppose that we have got approximations $x^0, x^1, \cdots, x^p$ to the solution x^* of (1.1). Then the $(p+1)$-th approximation x^{p+1} of x^* can be calculated by (1.2) and

$$x^{i,p} = \begin{cases} x^{i,p,m_{i,p}}, & \text{for } i \in J(p), \\ x^p, & \text{for } i \notin J(p), \end{cases} \tag{2.1}$$

where each $x^{i,p,m_{i,p}}$ is determined by the following formulae with the starting point $x^{i,p,0} = x^{s_i(p)}$:

$$x^{i,p,m+1} = F_i^{-1} G_i x^{i,p,m} + F_i^{-1}(N_i x^{s_i(p)} + b), \quad m = 0, 1, \cdots, m_{i,p} - 1, \tag{2.2}$$

while $\{m_{i,p}\}_{p \in N_0} (i = 1, 2, \cdots, \alpha)$ are infinite positive integer sequences, which may be determined either explicitly in advance or implicitly in the implementing process of the method.

Obviously, this method covers the ALGORITHM in the previous section cited from [6] as well as the methods proposed in [8]–[10].

By substituting (2.1)–(2.2) into (1.2), we can equivalently express Method I as

$$\begin{aligned} x^{p+1} = &\sum_{i \in J(p)} E_i \Big[\Big(F_i^{-1} G_i\Big)^{m_{i,p}} + \sum_{j=0}^{m_{i,p}-1} \Big(F_i^{-1} G_i\Big)^j F_i^{-1} N_i \Big] x^{s_i(p)} + \sum_{i \notin J(p)} E_i x^p \\ &+ \sum_{i \in J(p)} E_i \sum_{j=0}^{m_{i,p}-1} \Big(F_i^{-1} G_i\Big)^j F_i^{-1} b. \end{aligned} \tag{2.3}$$

As a matter of fact, there are various kinds of two-level multisplittings. For example, if in the two-level multisplitting $(M_i : F_i, G_i; N_i; E_i)$ $(i = 1, 2, \cdots, \alpha)$ of the matrix $A \in L(R^n)$, for each $i \in \{1, 2, \cdots, \alpha\}$, we particularly take $F_i = D_i - L_i$, $D_i = \mathrm{diag}(M_i)$ with $\det(D_i) \neq 0$ and $G_i = U_i$, where $L_i \in L(R^n)$ is strictly lower triangular and $U_i \in L(R^n)$ is zero-diagonal, satisfying $M_i = D_i - L_i - U_i (i = 1, 2, \cdots, \alpha)$, then a new two-level multisplitting $(M_i : D_i - L_i, U_i; N_i; E_i)$ $(i = 1, 2, \cdots, \alpha)$ of the matrix A is obtained. Based on this special two-level matrix multisplitting, Method I can be immediately formulated as the following form:

METHOD II: Suppose that we have got approximations $x^0, x^1, \cdots, x^p$ to the solution x^* of (1.1). Then the $(p+1)$-th approximation x^{p+1} of x^* can be calculated by (1.2) and (2.1), where each $x^{i,p,m_{i,p}}$ is determined by the following formulae

$$\begin{cases} x^{i,p,0} = x^{s_i(p)}, \\ x^{i,p,m+1} = (D_i - rL_i)^{-1}[(1-\omega)D_i + (\omega - r)L_i \\ \qquad +\omega U_i] x^{i,p,m} + (D_i - rL_i)^{-1} \omega (N_i x^{s_i(p)} + b), \\ m = 0, 1, \cdots, m_{i,p} - 1 \end{cases} \tag{2.4}$$

for each $i \in J(p)$. The meanings of the sequences $\{m_{i,p}\}_{p\in N_0}$ $(i = 1, 2, \cdots, \alpha)$ are the same as in Method I, while $r \in [0, \infty)$ is called a relaxation factor and $\omega \in (0, \infty)$ an acceleration factor.

Since the sub-system of linear equations (1.3) is solved for each $i \in \{1, 2, \cdots, \alpha\}$ by an accelerated overrelaxation (AOR) method, we call Method II as asynchronous nested matrix multisplitting AOR method (ANMM-AOR method). When the relaxation parameter pair (r, ω) is specially chosen to be (ω, ω), (1,1) and (0,1), etc., the corresponding methods resulted from (1.2), (2.1) and (2.4) are called as asynchronous nested matrix multisplitting SOR method (ANMM-SOR method), asynchronous nested matrix multisplitting Gauss-Seidel method (ANMM-GS method) and asynchronous nested matrix multisplitting Jacobi method (ANMM-J method) and so on, respectively.

Analogously, by substituting (2.4) and (2.1) into (1.2), and making use of the expressions

$$\mathcal{L}_i(r,\omega) = (D_i - rL_i)^{-1}[(1-\omega)D_i + (\omega - r)L_i + \omega U_i], \quad i = 1, 2, \cdots, \alpha, \tag{2.5}$$

Method II can be simply written as

$$\begin{aligned} x^{p+1} = \sum_{i\in J(p)} E_i\Big[\Big(\mathcal{L}_i(r,\omega)\Big)^{m_{i,p}} + \sum_{j=0}^{m_{i,p}-1}\Big(\mathcal{L}_i(r,\omega)\Big)^j \omega(D_i - rL_i)^{-1}N_i\Big]x^{s_i(p)} \\ + \sum_{i\notin J(p)} E_i x^p + \sum_{i\in J(p)} E_i \sum_{j=0}^{m_{i,p}-1}\Big(\mathcal{L}_i(r,\omega)\Big)^j (D_i - rL_i)^{-1}\omega b. \end{aligned} \tag{2.6}$$

In order to set up the convergence theories of the above two asynchronous nested matrix multisplitting methods, we need to define the infinite number sequence $\{m_l\}_{l\in N_0}$ in accordance with the following rule: m_0 is the least positive integer such that

$$\bigcup_{0\le s(p)\le p<m_0} J(p) = \{1, 2, \cdots, \alpha\},$$

in general, m_{l+1} is the least positive integer such that

$$\bigcup_{m_l\le s(p)\le p<m_{l+1}} J(p) = \{1, 2, \cdots, \alpha\}, l = 0, 1, 2, \cdots,$$

where $s(p) = \min_i s_i(p)$. Evidently, $s(p) \le p$. Since $\lim_{p\to\infty} s_i(p) = \infty$, we obviously have $\lim_{p\to\infty} s(p) = \infty$. For the meaning of the sequence $\{m_l\}_{l\in N_0}$, one can see [6] for detail.

3. Concepts and Lemmas

We adopt the notations and concepts used in [2]–[3], [6] and [10]–[12]. The following lemma, having been confirmed in [10], summarizes relations between different splittings and results on convergence properties of these splittings.

Lemma 1. *Let $A = B - C$ be a splitting.*

a) *If the splitting is regular or weak regular, then* $\rho(B^{-1}C) < 1$ *iff* $A^{-1} \geq 0$.

b) *If the splitting is an* M*-splitting, then* $\rho(B^{-1}C) < 1$ *iff* A *is an* M*-matrix.*

c) *If the splitting is an* H*-splitting, then* A *and* B *are* H*-matrices and it holds that* $\rho(B^{-1}C) \leq \rho(\langle B\rangle^{-1}|C|) < 1$.

d) *If the splitting is an* M*-splitting, then it is a regular splitting.*

e) *If the splitting is an* M*-splitting and* A *is an* M*-matrix, then it is an* H*-splitting and also an* H*-compatible splitting.*

f) *If the splitting is an* H*-compatible splitting and* A *is an* H*-matrix, then it is an* H*-splitting and thus convergent.*

Define nonnegative diagonal matrix sequences $\{I_p^{(1)}\}_{p\in N_0}$ and $\{I_p^{(2)}\}_{p\in N_0} \in L(R^n)$ by $I_p^{(1)} = \sum_{i\in J(p)} E_i$ and $I_p^{(2)} = \sum_{i\notin J(p)} E_i$ $(p = 0, 1, 2, \cdots)$, with E_i $(i = 1, 2, \cdots, \alpha)$ being the weighting matrices, i.e., $E_i \geq 0$ $(i = 1, 2, \cdots, \alpha)$ are diagonal and satisfy $\sum_i E_i = I$. Then in light of [6]–[7] we know that the following lemmas hold.

Lemma 2. *Let* $x \in R^n$ *be a positive vector*$(x > 0)$. *If the sequence* $\{\varepsilon^p\}_{p\in N_0}$ *satisfies*

$$|\varepsilon^{p+1}| \leq I_p^{(1)}x + I_p^{(2)}|\varepsilon^p|, \quad p = 0, 1, 2, \cdots.$$

Then for any nonnegative integer $q \leq p - 1$,

$$|\varepsilon^{p+1}| \leq \left(I - \prod_{j=p-q-1}^{p} I_j^{(2)}\right)x + \prod_{j=p-q-1}^{p} I_j^{(2)}|\varepsilon^{p-q-1}|.$$

Lemma 3. *Let* $m_{-1} = 0$ *and* $I^{(l)} = \prod_{p=m_{l-1}}^{m_l-1} I_p^{(2)}$ $(l = 0, 1, 2, \cdots)$. *Then for any positive vector* $x \in R^n$, *there exists* $\{\gamma^{(l)}\}_{l\in N_0} \in [0, \gamma] \subseteq [0, 1)$ *such that* $I^{(l)}x \leq \gamma^{(l)}x (l = 0, 1, 2, \cdots)$.

4. Convergence Analysis of Method I

Because the coefficient matrix in (1.1) is nonsingular, there exists a unique $x^* \in R^n$ such that $Ax^* = b$. Noticing the definition of Method I in section 2, according to (2.3) we easily know that the following relation holds:

$$\begin{aligned} x^* = \sum_{i\in J(p)} E_i\Big[\Big(F_i^{-1}G_i\Big)^{m_{i,p}} + \sum_{j=0}^{m_{i,p}-1}\Big(F_i^{-1}G_i\Big)^j F_i^{-1}N_i\Big]x^* + \sum_{i\notin J(p)} E_i x^* \\ + \sum_{i\in J(p)} E_i \sum_{j=0}^{m_{i,p}-1}\Big(F_i^{-1}G_i\Big)^j F_i^{-1}b. \end{aligned} \tag{4.1}$$

Let ε^p denote the error vector $\varepsilon^p = x^p - x^*$, by subtracting (4.1) from (2.3) we see that $\{\varepsilon^p\}_{p\in N_0}$ should satisfy

$$\varepsilon^{p+1} = \sum_{i\in J(p)} E_i T_{i,p}\varepsilon^{s_i(p)} + \sum_{i\notin J(p)} E_i\varepsilon^p, \tag{4.2}$$

where

$$T_{i,p} = \left(F_i^{-1}G_i\right)^{m_{i,p}} + \sum_{j=0}^{m_{i,p}-1}\left(F_i^{-1}G_i\right)^j F_i^{-1}N_i, \quad i=1,2,\cdots,\alpha, \quad p\in N_0. \tag{4.3}$$

Evidently, if $|\varepsilon^p| \to 0$ as $p\to\infty$, then we can conclude the convergence of Method I. In the remainder of this section, we will verify this fact by considering two cases of (1.1): $A\in L(R^n)$ is a monotone matrix and an H-matrix, respectively.

4.1 Monotone matrix case

Theorem 4.1. *Let $A\in L(R^n)$ be a monotone matrix, and $(M_i: F_i, G_i; N_i; E_i)(i=1,2,\cdots,\alpha)$ be a two-level multisplitting of it with $A = M_i - N_i$ $(i=1,2,\cdots,\alpha)$ being regular and $M_i = F_i - G_i$ $(i=1,2,\cdots,\alpha)$ weak regular. Then, for any starting vector $x^0\in R^n$, the iterative sequence $\{x^p\}_{p\in N_0}$ generated by Method I converges independently of the sequences $\{m_{i,p}\}_{p\in N_0}$ to the unique solution x^* of the system of linear equations* (1.1).

Proof. Since $A\in L(R^n)$ is a monotone matrix, we see that for any positive vector $u\in R^n$, there exists a positive vector $v\in R^n$ such that $Av=u$. Since $A=M_i-N_i(i=1,2,\cdots,\alpha)$ are regular splittings, $v - M_i^{-1}N_iv = M_i^{-1}Av = M_i^{-1}u > 0$ $(i=1,2,\cdots,\alpha)$. Moreover, as $M_i = F_i - G_i$ $(i=1,2,\cdots,\alpha)$ are weak regular splittings, we have

$$M_i^{-1} = (I - F_i^{-1}G_i)^{-1}F_i^{-1}, \quad i=1,2,\cdots,\alpha. \tag{4.4}$$

Now, from (4.3) we know that $T_{i,p}(i=1,2,\cdots,\alpha,\forall p\in N_0)$ are nonnegative and

$$\begin{aligned} T_{i,p} &= \left(F_i^{-1}G_i\right)^{m_{i,p}} + \sum_{j=0}^{m_{i,p}-1}\left(F_i^{-1}G_i\right)^j (I - F_i^{-1}G_i)M_i^{-1}N_i \\ &= I - \left[I - \left(F_i^{-1}G_i\right)^{m_{i,p}}\right]M_i^{-1}A. \end{aligned} \tag{4.5}$$

Through substituting (4.4) into (4.5) and making use of the identity

$$\left[I - \left(F_i^{-1}G_i\right)^{m_{i,p}}\right](I - F_i^{-1}G_i)^{-1} = \sum_{j=0}^{m_{i,p}-1}(F_i^{-1}G_i)^j,$$

we obtain

$$T_{i,p} = I - \sum_{j=0}^{m_{i,p}-1}(F_i^{-1}G_i)^jF_i^{-1}A.$$

Therefore, we have

$$\begin{aligned} T_{i,p}v &= v - \sum_{j=0}^{m_{i,p}-1}(F_i^{-1}G_i)^jF_i^{-1}u = v - F_i^{-1}u - \sum_{j=1}^{m_{i,p}-1}(F_i^{-1}G_i)^jF_i^{-1}u \\ &\le v - F_i^{-1}u, \quad i=1,2,\cdots,\alpha. \end{aligned}$$

Additionally, because of $F_i^{-1}u > 0$ $(i = 1, 2, \cdots, \alpha)$, we see that $v - F_i^{-1}u < v$ $(i = 1, 2, \cdots, \alpha)$. Hence, there exists a $\theta \in [0, 1)$ such that $v - F_i^{-1}u \leq \theta v$ $(i = 1, 2, \cdots, \alpha)$, which further implies that

$$T_{i,p}v \leq \theta v, \quad i = 1, 2, \cdots, \alpha, \quad \forall p \in N_0. \tag{4.6}$$

Based on (4.2)–(4.3) and (4.6), we see that once we generally suppose

$$|\varepsilon^t| \leq \Delta v, \qquad t = 0, 1, \cdots, p \tag{4.7}$$

for some $\Delta \in [0, \infty)$, it holds that

$$\begin{aligned} |\varepsilon^{p+1}| &\leq \sum_{i \in J(p)} E_i T_{i,p} |\varepsilon^{s_i(p)}| + \sum_{i \notin J(p)} E_i |\varepsilon^p| \leq \sum_{i \in J(p)} E_i T_{i,p} \Delta v + \sum_{i \notin J(p)} E_i |\varepsilon^p| \\ &\leq \theta \Delta \sum_{i \in J(p)} E_i v + \sum_{i \notin J(p)} E_i |\varepsilon^p|, \end{aligned}$$

or

$$|\varepsilon^{p+1}| \leq \theta \Delta I_p^{(1)} v + I_p^{(2)} |\varepsilon^p|, \tag{4.8}$$

where we have used the facts $s_i(p) \leq p$ and $|\varepsilon^{s_i(p)}| \leq \Delta v$ for $i = 1, 2, \cdots, \alpha$.

As a matter of fact, we can always admit that the initial error vector ε^0 satisfies

$$|\varepsilon^0| \leq \delta v \tag{4.9}$$

for some suitably chosen $\delta \in (0, \infty)$. Up to now, the proof of the theorem can be fulfilled in three parts by making use of (4.7)–(4.9).

Part I. $|\varepsilon^p| \leq \delta v$, $\forall p \in N_0$. Evidently, by induction this fact can be immediately verified beginning from (4.9) and making use of the observation (4.8).

Part II. $|\varepsilon^p| \leq \Delta_l v$, $\forall p \geq m_l$, where $\Delta_{-1} = \delta$, $\Delta_l = (\theta + (1 - \theta)\gamma^{(l)})\Delta_{l-1}$ $(l = 0, 1, 2, \cdots)$, and the sequence $\{\gamma^{(l)}\}_{l \in N_0}$ is defined in Lemma 3. In fact, for $l = 0$, by Lemma 2 and (4.7)–(4.8) with $\Delta = \delta$, we obtain

$$|\varepsilon^p| \leq \theta I_p^{(1)} \delta v + I_p^{(2)} |\varepsilon^{p-1}| \leq \left(I - \prod_{j=0}^{p-1} I_j^{(2)} \right) \theta \delta v + \prod_{j=0}^{p-1} I_j^{(2)} |\varepsilon^0|.$$

According to (4.9) and Lemma 3, we have

$$\begin{aligned} |\varepsilon^p| &\leq \left(I - \prod_{j=0}^{p-1} I_j^{(2)} \right) \theta \delta v + \prod_{j=0}^{p-1} I_j^{(2)} \delta v = \left(\theta I + (1 - \theta) \prod_{j=0}^{p-1} I_j^{(2)} \right) \delta v \\ &\leq (\theta I + (1 - \theta) I^{(0)}) \delta v \leq (\theta + (1 - \theta)\gamma^{(0)}) \delta v = \Delta_0 v. \end{aligned}$$

This shows that $|\varepsilon^p| \leq \Delta_l v (\forall p \geq m_l)$ is valid for $l = 0$. Now, suppose that $|\varepsilon^p| \leq \Delta_l v (\forall p \geq m_l)$ is true for $l \geq 1$. Then by using Lemmas 2 and 3 and starting from (4.7)–(4.8) with $\Delta = \Delta_l$, we get for $p \geq m_{l+1}$ that

$$|\varepsilon^p| \leq \theta I_p^{(1)} \Delta_l v + I_p^{(2)} |\varepsilon^{p-1}| \leq \left(I - \prod_{j=m_l}^{p-1} I_j^{(2)} \right) \theta \Delta_l v + \prod_{j=m_l}^{p-1} I_j^{(2)} |\varepsilon^{m_l}|$$

$$\leq \Big(I - \prod_{j=m_l}^{p-1} I_j^{(2)}\Big)\theta\Delta_l v + \prod_{j=m_l}^{p-1} I_j^{(2)}\Delta_l v.$$

Similar to the above derivation for $l = 0$, we can also conclude that $|\varepsilon^p| \leq \Delta_{l+1}v(\forall p \geq m_{l+1})$. By induction, we have proved the conclusion.

Part III. $|\varepsilon^p| \longrightarrow 0$ $(p \longrightarrow \infty)$. To test this fact, we let $\beta^{(l)} = \theta + (1-\theta)\gamma^{(l)}$ $(l = 0, 1, 2, \cdots)$. Clearly, for $l = 0, 1, 2, \cdots$, it holds that $\beta^{(l)} \in [0, \beta]$ with $\beta = \theta + (1-\theta)\gamma < 1$ and $\Delta_{l+1} = \beta^{(l+1)}\Delta_l$ with $\Delta_0 = \beta^{(0)}\delta$. Since

$$\Delta_{l+1} = \beta^{(l+1)}\Delta_l = \cdots = \prod_{j=0}^{l+1} \beta^{(j)}\delta \leq \beta^{l+2}\delta \longrightarrow 0 \quad (l \longrightarrow \infty),$$

by taking limits on both sides of the inequality in Part II, we immediately obtain $|\varepsilon^p| \longrightarrow 0$ $(p \longrightarrow \infty)$.

An important case of Theorem 4.1 is the following convergence theory about the ALGORITHM proposed in [6] for monotone matrix.

Theorem 4.2. *Let $A \in L(R^n)$ be a monotone matrix, and (M_i, N_i, E_i) $(i = 1, 2, \cdots, \alpha)$ be a multisplitting of it with $A = M_i - N_i (i = 1, 2, \cdots, \alpha)$ being weak regular. Then, for any starting vector $x^0 \in R^n$, the iterative sequence $\{x^p\}_{p \in N_0}$ generated by the ALGORITHM converges to the unique solution x^* of the system of linear equations* (1.1).

4.2 H-matrix case

Theorem 4.3. *Let $A \in L(R^n)$ be an H-matrix, and $(M_i : F_i, G_i; N_i; E_i)$ $(i = 1, 2, \cdots, \alpha)$ be a two-level multisplitting of it with both $A = M_i - N_i$ $(i = 1, 2, \cdots, \alpha)$ and $M_i = F_i - G_i$ $(i = 1, 2, \cdots, \alpha)$ being H-compatible splittings. Then, for any starting vector $x^0 \in R^n$, the iterative sequence $\{x^p\}_{p \in N_0}$ generated by Method I converges independently of the sequences $\{m_{i,p}\}_{p \in N_0}$ to the unique solution x^* of the system of linear equations* (1.1).

Proof. By Lemma 1 f), $A = M_i - N_i (i = 1, 2, \cdots, \alpha)$ are H-splittings. From Lemma 1 c), we see that $M_i (i = 1, 2, \cdots, \alpha)$ are H-matrices. Therefore,

$$\langle A\rangle = \langle M_i\rangle - |N_i|, \quad i = 1, 2, \cdots, \alpha \tag{4.10}$$

are M-splittings. Similarly, we know that $F_i (i = 1, 2, \cdots, \alpha)$ are H-matrices and

$$\langle M_i\rangle = \langle F_i\rangle - |G_i|, \quad i = 1, 2, \cdots, \alpha \tag{4.11}$$

are also M-splittings.

Now, from (4.3), for $i = 1, 2, \cdots, \alpha$ and $p \in N_0$, we can obtain the following estimates:

$$|T_{i,p}| \leq \Big(\langle F_i\rangle^{-1}|G_i|\Big)^{m_{i,p}} + \sum_{j=0}^{m_{i,p}-1} \Big(\langle F_i\rangle^{-1}|G_i|\Big)^j \langle F_i\rangle^{-1}|N_i|. \tag{4.12}$$

Define

$$\hat{T}_{i,p} = \Big(\langle F_i\rangle^{-1}|G_i|\Big)^{m_{i,p}} + \sum_{j=0}^{m_{i,p}-1} \Big(\langle F_i\rangle^{-1}|G_i|\Big)^j \langle F_i\rangle^{-1}|N_i|. \tag{4.13}$$

and consider the sequence $\{\hat{\varepsilon}^p\}_{p\in N_0}$ generated by

$$\hat{\varepsilon}^0 = |\varepsilon^0| \equiv |x^0 - x^*|, \quad \hat{\varepsilon}^{p+1} = \sum_{i\in J(p)} E_i \hat{T}_{i,p} \hat{\varepsilon}^{s_i(p)} + \sum_{i\notin J(p)} E_i \hat{\varepsilon}^p, \quad p = 0,1,2,\cdots. \tag{4.14}$$

By comparing (4.13) and (4.14) with (4.3) and (4.2), and considering (4.10)–(4.11), following the proof process of Theorem 4.1 we can immediately conclude that under the conditions of this theorem, $\hat{\varepsilon}^p \longrightarrow 0$ as $p \longrightarrow \infty$.

On the other hand, by induction we can prove

$$|\varepsilon^p| \le \hat{\varepsilon}^p, \quad p = 0,1,2,\cdots. \tag{4.15}$$

In fact, when $p = 0$ (4.15) is obviously true. Suppose that (4.15) is true for $p = 0,1,2,\cdots,t$. Then $s_i(t) \le t$ $(i = 1,2,\cdots,\alpha)$ directly give the estimates

$$|\varepsilon^{s_i(t)}| \le \hat{\varepsilon}^{s_i(t)}, \quad i = 1,2,\cdots,\alpha. \tag{4.16}$$

Now, for $p = t+1$, from (4.2) and (4.12)–(4.14) as well as (4.16), by direct calculations we get

$$|\varepsilon^{t+1}| \le \sum_{i\in J(t)} E_i |T_{i,t}||\varepsilon^{s_i(t)}| + \sum_{i\notin J(t)} E_i |\varepsilon^t| \le \sum_{i\in J(t)} E_i \hat{T}_{i,t} \hat{\varepsilon}^{s_i(t)} + \sum_{i\notin J(t)} E_i \hat{\varepsilon}^t = \hat{\varepsilon}^{t+1},$$

which shows (4.15) is also true for $p = t+1$.

5. Convergence Analysis of Method II

Analogous to section 4, we know that the error vector sequence $\{\varepsilon^p\}_{p\in N_0}$ corresponding to Method II satisfies

$$\varepsilon^{p+1} = \sum_{i\in J(p)} E_i T_{i,p}(r,\omega) \varepsilon^{s_i(p)} + \sum_{i\notin J(p)} E_i \varepsilon^p, \tag{5.1}$$

where

$$\begin{cases} T_{i,p}(r,\omega) = \Big(\mathcal{L}_i(r,\omega)\Big)^{m_{i,p}} + \sum\limits_{j=0}^{m_{i,p}-1} \Big(\mathcal{L}_i(r,\omega)\Big)^j \omega (D_i - rL_i)^{-1} N_i, \\ i = 1,2,\cdots,\alpha, \quad p \in N_0. \end{cases} \tag{5.2}$$

Clearly, to prove the convergence of Method II, we only need to verify $|\varepsilon^p| \to 0$ as $p \to \infty$. Because the test of this fact is similar to the corresponding one of Method I, here, we only use the convergence theory for the H-matrix case as an example to show its proving skeleton, while for the completion of the convergence theory of Method II, we also list its convergence theorem for the monotone matrix case but omit its proof.

Theorem 5.1. *Let $A \in L(R^n)$ be a monotone matrix, and $(M_i : D_i - L_i, U_i; N_i; E_i)$ $(i = 1,2,\cdots,\alpha)$ be a two-level multisplitting of it with $A = M_i - N_i$ $(i = 1,2,\cdots,\alpha)$ being regular and $D_i \ge 0$, $L_i \ge 0$, $U_i \ge 0$ $(i = 1,2,\cdots,\alpha)$. Then, for any starting vector $x^0 \in R^n$, the iterative sequence $\{x^p\}_{p\in N_0}$ generated by Method II converges independently of* *the sequences* $\{m_{i,p}\}_{p\in N_0}$ *to the unique solution* x^* *of the system of linear equations* (1.1) *provided the relaxation parameters* r *and* ω *satisfy* $0 \le r \le \omega, 0 < \omega \le 1$.

Theorem 5.2. *Let* $A \in L(R^n)$ *be an* H*-matrix with* $D = diag(A)$ *and* $B = D - A$. *Assume that* $(M_i : D_i - L_i, U_i; N_i; E_i)$ $(i = 1, 2, \cdots, \alpha)$ *is a two-level multisplitting of the matrix* A *with* $A = M_i - N_i$ $(i = 1, 2, \cdots, \alpha)$ *being* H*-compatible splittings,*

$$\langle M_i \rangle = |D_i| - |L_i| - |U_i|, \quad i = 1, 2, \cdots, \alpha \tag{5.3}$$

and

$$diag\ (M_i) = D_i = D, \quad i = 1, 2, \cdots, \alpha. \tag{5.4}$$

Then, for any starting vector $x^0 \in R^n$, *the iterative sequence* $\{x^p\}_{p\in N_0}$ *generated by Method II converges independently of the sequences* $\{m_{i,p}\}_{p\in N_0}$ *to the unique solution* x^* *of the system of linear equations* (1.1) *provided the relaxation parameters* r *and* ω *satisfy*

$$0 \le r \le \omega, \qquad 0 < \omega < 2/(1 + \rho(|D|^{-1}|B|)). \tag{5.5}$$

Proof. In light of Lemma 1 f) we know that M_i $(i = 1, 2, \cdots, \alpha)$ are H-matrices. Let

$$C_i = D_i - M_i, \qquad i = 1, 2, \cdots, \alpha, \tag{5.6}$$

by (5.3)–(5.4) we easily see that

$$|C_i| = |L_i| + |U_i|, \quad i = 1, 2, \cdots, \alpha. \tag{5.7}$$

Because $(D_i - rL_i)$ $(i = 1, 2, \cdots, \alpha)$ are H-matrices, we have

$$|(D_i - rL_i)^{-1}| \le \langle D_i - rL_i \rangle^{-1} = (|D_i| - r|L_i|)^{-1}, \quad i = 1, 2, \cdots, \alpha. \tag{5.8}$$

Moreover, for the matrices $\mathcal{L}_i(r, \omega)$ $(i = 1, 2, \cdots, \alpha)$ defined by (2.5), we have the following estimates:

$$\begin{aligned} |\mathcal{L}_i(r,\omega)| &\le |(D_i - rL_i)^{-1}|[|1-\omega||D_i| + (\omega - r)|L_i| + \omega|U_i|] \\ &\le (|D_i| - r|L_i|)^{-1}[|1-\omega||D_i| + (\omega - r)|L_i| + \omega|U_i|] := \hat{\mathcal{L}}_i(r,\omega). \end{aligned} \tag{5.9}$$

Presently, by (5.8)–(5.9) we can obtain from (5.2) that

$$\begin{aligned} |T_{i,p}(r,\omega)| &\le |\mathcal{L}_i(r,\omega)|^{m_{i,p}} + \sum_{j=0}^{m_{i,p}-1} |\mathcal{L}_i(r,\omega)|^j \omega(|D_i| - r|L_i|)^{-1}|N_i| \\ &\le \left(\hat{\mathcal{L}}_i(r,\omega)\right)^{m_{i,p}} + \sum_{j=0}^{m_{i,p}-1} \left(\hat{\mathcal{L}}_i(r,\omega)\right)^j \omega(|D_i| - r|L_i|)^{-1}|N_i| \end{aligned} \tag{5.10}$$

holds for each $i \in \{1, 2, \cdots, \alpha\}$.

Define

$$\hat{T}_{i,p}(r,\omega) = \left(\hat{\mathcal{L}}_i(r,\omega)\right)^{m_{i,p}} + \sum_{j=0}^{m_{i,p}-1} \left(\hat{\mathcal{L}}_i(r,\omega)\right)^j \omega(|D_i| - r|L_i|)^{-1}|N_i|, \tag{5.11}$$

and consider the sequence $\{\hat{\varepsilon}^p\}_{p\in N_0}$ yielded by

$$\hat{\varepsilon}^0 = |\varepsilon^0| \equiv |x^0 - x^*|, \quad \hat{\varepsilon}^{p+1} = \sum_{i\in J(p)} E_i \hat{T}_{i,p}(r,\omega)\hat{\varepsilon}^{s_i(p)} + \sum_{i\notin J(p)} E_i \hat{\varepsilon}^p. \tag{5.12}$$

For $i = 1, 2, \cdots, \alpha$, let

$$\begin{cases} \mathcal{A}(\omega) = \frac{1-|1-\omega|}{\omega}|D| - |B|, \\ \mathcal{M}_i(\omega) = \frac{1-|1-\omega|}{\omega}|D| - |C_i|, \\ \mathcal{N}_i(\omega) = |N_i|, \\ \mathcal{F}_i(r,\omega) = \frac{1}{\omega}(|D| - r|L_i|), \\ \mathcal{G}_i(r,\omega) = \frac{1}{\omega}[|1-\omega||D| + (\omega - r)|L_i| + \omega|U_i|]. \end{cases} \tag{5.13}$$

Then it obviously holds that

$$\begin{cases} \mathcal{A}(\omega) = \mathcal{M}_i(\omega) - \mathcal{N}_i(\omega), \\ \mathcal{M}_i(\omega) = \mathcal{F}_i(r,\omega) - \mathcal{G}_i(r,\omega), \\ \hat{\mathcal{L}}_i(r,\omega) = \mathcal{F}_i(r,\omega)^{-1}\mathcal{G}_i(r,\omega), \\ \hat{T}_{i,p}(r,\omega) = \Big(\hat{\mathcal{L}}_i(r,\omega)\Big)^{m_{i,p}} + \sum\limits_{j=0}^{m_{i,p}-1}\Big(\hat{\mathcal{L}}_i(r,\omega)\Big)^j \mathcal{F}_i(r,\omega)^{-1}\mathcal{N}_i(\omega). \end{cases} \tag{5.14}$$

Since $A \in L(R^n)$ is an H-matrix, we have $\rho(|D|^{-1}|B|) < 1$. Noticing the varying region (5.5) of the relaxation parameter pair (r,ω), we see that $\mathcal{A}(\omega) \in L(R^n)$ is an M-matrix. Furthermore, as

$$|C_i| = |D| - \langle M_i\rangle = |D| - (\langle A\rangle + |N_i|) = |B| - |N_i| \le |B|, \quad i = 1, 2, \cdots, \alpha,$$

in accordance with [13, 2.4.10] we know that $\mathcal{M}_i(\omega) \in L(R^n)$ $(i = 1, 2, \cdots, \alpha)$ are M-matrices, too. Therefore, both $\mathcal{A}(\omega) = \mathcal{M}_i(\omega) - \mathcal{N}_i(\omega)$ and $\mathcal{M}_i(\omega) = \mathcal{F}_i(r,\omega) - \mathcal{G}_i(r,\omega)$ are M-splittings for $i = 1, 2, \cdots, \alpha$ under the conditions of this theorem. By making use of Theorem 4.1, we know that $\hat{\varepsilon}^p \longrightarrow 0$ as $p \longrightarrow \infty$.

Similar to the proof of Theorem 4.3, we can also confirm that $\{\hat{\varepsilon}^p\}_{p\in N_0}$ is a majorizing sequence of $\{\varepsilon^p\}_{p\in N_0}$ defined by (5.1)–(5.2), i.e., $|\varepsilon^p| \le \hat{\varepsilon}^p$ $(p = 0, 1, 2, \cdots)$. Hence, we finally obtain $\varepsilon^p \longrightarrow 0 (p \longrightarrow \infty)$.

6. Numerical Results

For a given positive integer $\widetilde{n}$, let $n = \widetilde{n}^2$ and consider the system of linear equations (1.1) with

$$\begin{cases} A = BlockTridiag(-I, \widetilde{B}, -I) \in L(R^n), \\ \widetilde{B} = tridiag(-1, 4, -1) \in L(R^{\widetilde{n}}), \\ b = (10, 10, \cdots, 10)^T \in R^n. \end{cases}$$

This example naturally comes from the finite difference discretization of a Dirichlet problem on the unit square $[0,1] \times [0,1]$; see [11] and [13] for details. This system of linear equations is solved by the ANMM-AOR method and ANMM-SOR method.

In our computations, with $(2\alpha - 1)$ positive integers $n_1, n_2, \cdots, n_{2\alpha-1}$ satisfying $n_k = \frac{k\widetilde{n}}{2\alpha-1}$ $(k = 1, 2, \cdots, 2\alpha - 1)$ we let processor i solve the variables $x_j (j = \widetilde{n} n_{2i-3} + 1, \widetilde{n} n_{2i-3} + 2, \cdots, \widetilde{n} n_{2i})$. Here, we stipulate that $n_{-1} = 0$ and $n_{2\alpha} = \widetilde{n}$. The inner iteration numbers are taken to be $m_{i,p} \equiv mip$ $(i = 1, 2, \cdots, \alpha, p \in N_0)$, and the splitting and the weighting matrices are taken to be $N_i = M_i - A$ and

$$\begin{aligned}
M_i &= diag(\overbrace{4I, \cdots, 4I}^{\tilde{n} n_{2i-3}}, \overbrace{\widetilde{B}, \cdots, \widetilde{B}}^{\tilde{n}(n_{2i}-n_{2i-3})}, 4I, \cdots, 4I), \\
E_i &= diag(\overbrace{0, \cdots, 0}^{\tilde{n} n_{2i-3}}, \overbrace{\mu_{\tilde{n} n_{2i-3}+1} I, \cdots, \mu_{\tilde{n} n_{2i}} I}^{\tilde{n}(n_{2i}-n_{2i-3})}, 0, \cdots, 0), \\
L_i &= \text{the strictly lower triangular matrix of } (-M_i), \\
U_i &= \text{the strictly upper triangular matrix of } (-M_i),
\end{aligned}$$

respectively, where

$$\mu_j = \begin{cases} 0.5, & \text{if } \widetilde{n} n_{2i-3} + 1 \le j \le \widetilde{n} n_{2i-2}, \\ 1.0, & \text{if } \widetilde{n} n_{2i-2} + 1 \le j \le \widetilde{n} n_{2i-1}, \\ 0.5, & \text{if } \widetilde{n} n_{2i-1} + 1 \le j \le \widetilde{n} n_{2i}. \end{cases}$$

We remark that this system of linear equations and this two-level multisplitting of the matrix $A \in L(R^n)$ satisfy all the theoretical hypotheses made in the previous sections.

The parallel computer used is the SGI Power Challenge multiprocessor located in Oxford University Computing Laboratory. Computations are done corresponding to $n = 6400$, and various processor numbers α and relaxation parameter pairs (r, ω). All our computations are started from an initial vector having all components equal to -100, and terminated once the current iterations x^p obey $\dfrac{\|Ax^p - b\|_1}{\|Ax^0 - b\|_1} \le 10^{-7}$ or the stopping criterion is not satisfied after 8000 iteration steps. For $\alpha = 4$, the corresponding sequential CPU time (CPU) in seconds and parallel speed-up (SP) are listed in the following numerical tables. Here, the SP is defined to be the ratio of the sequential CPU to the corresponding parallel runnings. We remark that the parallel CPU time is not listed in the numerical tables since it can be easily obtained by dividing the sequential CPU by the corresponding parallel SP. From our computations we see that suitable choices of the relaxation parameters r and ω can greatly accelerate the convergence rates of the relaxation methods, and the asynchronous nested matrix multisplitting relaxation methods have better numerical behaviour than the ordinary asynchronous multisplitting relaxation methods. Moreover, the asynchronous nested matrix multisplitting AOR method has larger convergence domain than the asynchronous nested matrix multisplitting SOR method, and the convergence rate of the former is, in general, moderately faster than that of the later. Evidently, the numerical results further confirm the theories established in the previous sections, and also show that our new methods are feasible and efficient for solving the system of linear equations (1.1) on the high-speed multiprocessor systems.

Table I: ANMM-SOR method

ω		0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5	1.6
mip=1	CPU	235.1	184.8	178.7	165.3	168.8	118.1	∞	∞	∞
	SP	3.1	3.1	3.1	3.1	3.2	3.1	–	–	–
mip=2	CPU	203.3	159.9	160.5	153.9	146.7	128.3	189.7	∞	∞
	SP	3.0	3.1	3.1	3.2	3.1	3.1	3.2	–	–
mip=3	CPU	176.6	161.7	162.2	156.5	157.6	148.7	167.2	∞	∞
	SP	3.3	3.1	3.1	3.1	3.2	3.1	3.1	–	–
mip=4	CPU	178.7	171.3	175.1	159.8	166.1	161.3	171.3	163.0	171.6
	SP	3.1	3.1	3.0	3.1	3.2	3.1	3.2	3.2	3.1
mip=5	CPU	200.9	183.1	188.9	177.9	173.3	179.6	194.7	196.1	∞
	SP	3.1	3.1	3.1	3.1	3.2	3.1	3.2	3.2	–

Table II: ANMM-AOR method

r		1.0	1.1	1.2	1.3	1.4	1.1	1.2	1.2	1.1
ω		0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5	1.6
mip=1	CPU	220.2	176.4	149.8	119.4	111.3	119.3	∞	∞	∞
	SP	3.0	3.1	3.3	3.1	3.1	3.1	–	–	–
mip=2	CPU	164.7	139.2	121.4	116.4	102.5	117.4	115.7	151.2	146.1
	SP	3.1	3.1	3.1	3.1	2.9	3.1	3.2	3.2	3.1
mip=3	CPU	155.4	154.1	132.7	150.3	147.7	126.7	134.8	105.3	∞
	SP	3.0	3.1	3.1	3.1	3.1	3.1	3.2	3.2	–
mip=4	CPU	173.4	163.4	151.7	152.8	154.7	164.2	157.4	134.7	159.3
	SP	3.1	3.1	3.2	3.1	3.2	3.1	3.2	3.2	3.1
mip=5	CPU	189.1	182.5	177.2	178.3	175.4	181.1	183.7	172.3	193.2
	SP	3.1	3.1	3.1	3.1	3.2	3.1	3.2	3.2	3.1

Acknowledgement. The authors are very much indebted to the referees for their valuable suggestions about the original manuscript of this paper. They also thank Dr. Y.G. Huang for his help during the running of the numerical experiments and Professor J.C. Sun for his discussion on asynchronous parallel iterations. Part of this work was fulfilled during the first author's visit of Oxford University Computing Laboratory and Atlas Centre, Rutherford Appleton Laboratory in England during August 1997-August 1998, and he sincerely thanks his advisors, Professors A.J. Wathen and I.S. Duff, for their warm encouragements and helps.

References

[1] D.P. O'Leary, R.E. White, Multisplittings of matrices and parallel solution of linear systems, *SIAM J. Alg. Disc. Methods,* **6** (1985), 630–640.

[2] A. Frommer, G. Mayer, Convergence of relaxed parallel multisplitting methods, *Linear Algebra Appl.,* **119** (1989), 141–152.

[3] D.R. Wang, On the convergence of the parallel multisplitting AOR algorithm, *Linear Algebra Appl.,* **154-156** (1991), 473–486.

[4] L. Elsner, Comparison of weak regular splittings and multisplitting methods, *Numer. Math.,* **56** (1989), 283–289.

[5] M. Neumann, R.J. Plemmons, Convergence of parallel multisplitting iterative methods for M-matrices, *Linear Algebra Appl.,* **88/89** (1987), 559–573.

[6] D.R. Wang, Z.Z. Bai, D.J. Evans, A class of asynchronous parallel matrix multisplitting relaxation methods, *Parallel Algorithms Appl.,* **2** (1994), 173–192.

[7] D.R. Wang, Z.Z. Bai, D.J. Evans, A class of asynchronous parallel nonlinear multisplitting relaxation methods, *Parallel Algorithms Appl.,* **2** (1994), 209–228.

[8] P.J. Lanzkron, D.J. Rose, D.B. Szyld, Convergence of nested classical iterative methods for linear systems, *Numer. Math.,* **58** (1991), 685–702.

[9] D.B. Szyld, M.T. Jones, Two-stage and multisplitting methods for the parallel solution of linear systems, *SIAM J. Matrix Anal. Appl.,* **13** (1992), 671–679.

[10] A. Frommer, D.B. Szyld, H-splittings and two-stage iterative methods, *Numer. Math.,* **63** (1992), 345–356.

[11] R.S. Varga, Matrix Iterative Analysis, Prentice Hall, Englewood Cliffs, N. J., 1962.

[12] R. Bru, L. Elsner, M. Neumann, Models of parallel chaotic iteration methods, *Linear Algebra Appl.,* **103** (1998), 175–192.

[13] J.M. Ortega, W.C. Rheinboldt, Iterative Solutions of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

[14] D.J. Evans, D.R. Wang, An asynchronous parallel algorithm for solving a class of nonlinear simultaneous equations, *Parallel Comput.,* **17** (1991), 165–180.