Journal of Computational and Applied Mathematics 110 (1999) 271–286

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.nl/locate/cam

A class of asynchronous multisplitting two-stage iterations for large sparse block systems of weakly nonlinear equations

Zhong-Zhi Bai[a,*,1], D.J. Evans[b], R.C. Calinescu[c]

[a] 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, People's Republic of China

[b] Department of Computing, The Nottingham Trent University, Burton Street, Nottingham NG1 4BU, England, UK

[c] Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford OX1 3QD, UK

Received 13 April 1998; received in revised form 6 May 1999

Abstract

For the block system of weakly nonlinear equations $Ax = G(x)$, where $A \in \mathbb{R}^{n\times n}$ is a large sparse block matrix and $G : \mathbb{R}^n \to \mathbb{R}^n$ is a block nonlinear mapping having certain smoothness properties, we present a class of asynchronous parallel multisplitting block two-stage iteration methods in this paper. These methods are actually the block variants and generalizations of the asynchronous multisplitting two-stage iteration methods studied by Bai and Huang (Journal of Computational and Applied Mathematics 93(1) (1998) 13–33), and they can achieve high parallel efficiency of the multiprocessor system, especially, when there is load imbalance. Under quite general conditions that $A \in \mathbb{R}^{n\times n}$ is a block H-matrix of different types and $G : \mathbb{R}^n \to \mathbb{R}^n$ is a block P-bounded mapping, we establish convergence theories of these asynchronous multisplitting block two-stage iteration methods. Numerical computations show that these new methods are very efficient for solving the block system of weakly nonlinear equations in the asynchronous parallel computing environment. © 1999 Elsevier Science B.V. All rights reserved.

MSC: 65H10; 65W05; CR: G1.3

Keywords: Block system of weakly nonlinear equations; Matrix multisplitting; Block two-stage iteration; Asynchronous parallel method; Block H-matrix; Convergence theory

1. Introduction

Consider the system of weakly nonlinear equations

$$Ax = G(x), \tag{1}$$

* Corresponding author. Tel.: +86-10-62630992; fax: +86-10-62542285.

E-mail addresses: bzz@lsec.cc.ac.cn (Z.-Z. Bai), evans@lboro.ac.uk (D.J. Evans), rcc@comlab.ox.ac.uk (R.C. Calinescu)

[1] Project 19601036 supported by the National Natural Science Foundation of China.

PII: S0377-0427(99)00236-8

where $A \in \mathbb{R}^{n\times n}$ is a large sparse nonsingular matrix, and $G : \mathbb{R}^n \to \mathbb{R}^n$ is a nonlinear function with certain smoothness properties. This system arises in many problems of science and engineering, and in particular in discretizations of certain nonlinear differential equations, e.g., of the form $\Delta u = \eta(u)$; see e.g. [26].

Since a general purpose method may be not always efficient for solving a special structure problem. Bai [3] established a class of sequential two-stage iteration methods for the system of weakly nonlinear equations by taking into account concrete properties of the involved matrix and mapping. Then, based on the matrix multisplitting technique introduced in O'Leary and White [24], Bai [4] presented efficient parallel generalizations of the above sequential two-stage iteration methods. Furthermore, to exploit the parallel efficiency of the high-speed multiprocessor systems as far as possible, Bai and Huang [9] recently proposed asynchronous multisplitting two-stage iteration methods for solving the system of weakly nonlinear equations (1). These asynchronous methods have the potentials of converging much faster than their synchronous counterparts in [4], especially, when there is load imbalance. When the matrix $A \in \mathbb{R}^{n\times n}$ is a pointwise H-matrix and the mapping $G : \mathbb{R}^n \to \mathbb{R}^n$ is a pointwise P-bounded mapping, the convergence of the afore-mentioned two-stage iteration methods were discussed in detail in [3,4,9], respectively, under suitable conditions imposed upon the multiple splittings of the matrix $A \in \mathbb{R}^{n\times n}$.

These multisplitting two-stage iteration methods are quite suitable for both the tightly coupled multiprocessor [2] and the multicomputer with a shared global memory. Regarding the multicomputer without a shared global memory, for which each processor has its own local memory and process cooperation occurs either through message passing or through memory shared between pairs of processors, Frommer and Szyld [21] presented another approach of multisplitting two-stage iteration methods, which are closely related to the additive Schwarz iteration methods. For further details we refer to Frommer and Schwandt [18] and Frommer et al. [19]. The generalizations of this class of asynchronous multisplitting iteration methods to the nonlinear fixed-point problems were studied by Bahi et al. [2] and Baz et al. [13].

In this paper, we are interested in block variants of the asynchronous multisplitting two-stage iteration methods in Bai and Huang [9] for the block system of weakly nonlinear equations (1) in which the matrix $A \in \mathbb{R}^{n\times n}$ is partitioned into $N \times N$ blocks $A_{\ell j} \in \mathbb{R}^{n_\ell \times n_j}$, with $\sum_{j=1}^{N} n_j = n$, i.e.,

$$A \in \mathbb{L}_n(n_1,\ldots,n_N) = \{A \in \mathbb{R}^{n\times n} \mid A = (A_{\ell j}), A_{\ell j} \in \mathbb{R}^{n_\ell \times n_j},\ 1 \leqslant \ell, j \leqslant N\}$$

and A_{jj} being nonsingular for $j=1,2,\ldots,N$, and the vectors x and $G(x)$ are partitioned into subvectors $x_j \in \mathbb{R}^{n_j}$ and $G_j(x) \in \mathbb{R}^{n_j}$, $j = 1,2,\ldots,N$, in a way conformally with the partition of A, i.e.,

$$x \in V_n(n_1,\ldots,n_N) = \{x \in \mathbb{R}^n \mid x = (x_1^{\mathrm{T}},\ldots,x_N^{\mathrm{T}})^{\mathrm{T}},\ x_j \in \mathbb{R}^{n_j},\ 1 \leqslant j \leqslant N\}.$$

When the context is clear we will simply use $\mathbb{L}_n$ for $\mathbb{L}_n(n_1,\ldots,n_N)$ and V_n for $V_n(n_1,\ldots,n_N)$, respectively. This partition may correspond to a partition of the underlying grid, or of the domain of the differential equation being studied, or it may originate from a partitioning algorithm of the sparse matrix A, as done, e.g., in [14,25]. We will discuss convergence properties of the above asynchronous multisplitting block two-stage iteration methods for rather general class of nonsingular matrices including block H-matrices of different types (see e.g. [5–7,17]), and nonlinear mappings

[2] A simple processor intercommunication pattern of this MIMD-system assumes that all processors work through a central switching mechanism to reach a shared global memory.

including block P-bounded mappings (see e.g. [12]). At last, some numerical results are given which show that the asynchronous multisplitting block two-stage iteration methods can achieve higher parallel efficiency than their synchronous counterparts as well as the asynchronous multisplitting point two-stage iteration methods in [9].

The rest of this paper is arranged as follows: We describe the asynchronous multisplitting block two-stage iteration methods paralleling the point versions in [9] in Section 2, and some preliminary results used in this paper is presented in Section 3. In Section 4, we analyze the global convergence of these asynchronous multisplitting block two-stage iteration methods, and finally, in Section 5 we give some numerical results run on the SGI power challenge multiprocessor computer.

2. The asynchronous multisplitting block two-stage methods

Assume the multiprocessor system consist of α processors and consider the (outer and inner) splittings [3] $A = B_i - C_i$, $B_i = M_i - N_i$, $i = 1, 2, \dots, \alpha$, and a set of diagonal nonnegative matrices E_i, $i = 1, 2, \dots, \alpha$, such that $\sum_{i=1}^{\alpha} E_i = I$ (the identity matrix). [4] Let $\tau_i(p)$ be a nonnegative integer that represents the index of an older global iterate which the ith processor uses to compute its pth local approximation, and $J(p)$ be a nonempty subset of the integer set $\{1, 2, \dots, \alpha\}$ that satisfies $i \in J(p)$ if and only if the ith processor starts its computation of a new iterate at the pth step. As is customary in the descriptions and analyses of asynchronous methods, we assume that the superscripts $\tau_i(p)$ and the subsets $J(p)$, $p \in N_0 = \{0, 1, 2, \dots\}$, satisfy the following conditions: (a) $\tau_i(p) \leqslant p$ for all $i \in \{1, 2, \dots, \alpha\}$ and $p \in N_0$; (b) $\lim_{p\to\infty} \tau_i(p) = \infty$ for all $i \in \{1, 2, \dots, \alpha\}$; and (c) The set $\{p \in N_0 \mid i \in J(p)\}$ is infinite for all $i \in \{1, 2, \dots, \alpha\}$. Condition (a) assumes that the currently unavailable information should not be used in the current computations; Condition (b) requires that every processor of the multiprocessor system must adopt new information to update its local variables continually; and Condition (c) demands that all processors of the multiprocessor system must proceed their local iterations without dead breakdown.

With these notations, the new asynchronous multisplitting block two-stage iteration method can be described as follows.

Method 1. (ASYNCHRONOUS MULTISPLITTING BLOCK TWO-STAGE METHOD)**.** Given an initial vector $x^0 \in V_n$. Suppose that we have got approximations $x^0, x^1, \dots, x^p$ to the solution $x^* \in V_n$ of the block system of weakly nonlinear equations (1). Then the next approximation x^{p+1} is obtained by

$$x^{p+1} = \sum_{i \in J(p)} E_i x^{p+1,i} + \sum_{i \notin J(p)} E_i x^p, \tag{2}$$

where for each $i \in J(p)$, $x^{p+1,i} = x^{p,\,i,\,s_i(p)}$ is computed from the recursive formula

$$x^{p,\,i,\,k+1} = M_i^{-1}(N_i x^{p,\,i,\,k} + C_i x^{\tau_i(p)} + G(x^{\tau_i(p)})), \quad k = 0, 1, \dots, s_i(p) - 1,$$

with the starting point $x^{p,\,i,\,0} = x^{\tau_i(p)}$.

---

[3] $A = B - C$ is called a splitting of the matrix A if the matrix B is nonsingular. One can refer to Section 3 for detail.
[4] Such a kind of matrices E_i $(i = 1, 2, \dots, \alpha)$ is called weighting matrices.

We remark that the mathematical description of Method 1 is much similar to its point counterpart in [9]. However, the philosophies behind these two methods are quite different because the operations of the former are understood in the blockwise sense while the operations of the latter in the pointwise sense.

A practical example of Method 1 can be given by letting

$$\begin{aligned} M_i &:= M_i(\gamma,\omega) = \frac{1}{\omega}(D_i - \gamma L_i), \\ N_i &:= N_i(\gamma,\omega) = \frac{1}{\omega}((1-\omega)D_i + (\omega-\gamma)L_i + \omega U_i), \end{aligned} \qquad i = 1,2,\ldots,\alpha,$$

where γ and ω ($\neq 0$) are relaxation parameters and for $i = 1,2,\ldots,\alpha$, $D_i = \mathrm{Diag}(B_i)$ are the block diagonal matrices of B_i, L_i are strictly block lower-triangular matrices and U_i are block zero-diagonal matrices satisfying $B_i = D_i - L_i - U_i (i = 1,2,\ldots,\alpha)$. The resulted method is called asynchronous multisplitting block two-stage AOR (accelerated overrelaxation) method.

Method 2. (Asynchronous Multisplitting Block Two-stage AOR Method)**.** Given an initial vector $x^0 \in V_n$. Suppose that we have got approximations $x^0, x^1, \ldots, x^p$ to the solution $x^* \in V_n$ of the block system of weakly nonlinear equations (1). Then the next approximation x^{p+1} is obtained by (2) where for each $i \in J(p)$, $x^{p+1,i} = x^{p,i,s_i(p)}$ is computed from the recursive formula

$$\begin{aligned} x^{p,i,k+1} &= (D_i - \gamma L_i)^{-1}\{((1-\omega)D_i + (\omega-\gamma)L_i + \omega U_i)x^{p,i,k} \\ &\quad + \omega(C_i x^{\tau_i(p)} + G(x^{\tau_i(p)}))\}, \quad k = 0,1,\ldots,s_i(p)-1 \end{aligned}$$

with the starting point $x^{p,i,0} = x^{\tau_i(p)}$.

If $G(x) \equiv b$ (a constant vector), then system (1) reduces to the system of linear equations $Ax = b$. For this special situation, Method 2 becomes the synchronous multisplitting block relaxation method in Bai [5] when $s_i(p) = 1$ and $\tau_i(p) = p$ ($i = 1,2,\ldots,\alpha$, $p \in N_0$), and it turns to the asynchronous multisplitting block relaxation method in Bai [7] when $s_i(p) = 1$ ($i = 1,2,\ldots,\alpha$, $p \in N_0$). We refer to Bai [8] for an improvement and generalization of this asynchronous multisplitting block relaxation method, and Bai [6] and Evans and Bai [17] for the two-sweep relaxed synchronous multisplitting block iteration method.

After direct calculations, we can briefly express Methods 1 and 2 in the matrix–vector forms

$$x^{p+1} = \sum_{i \in J(p)} E_i \left\{ (M_i^{-1}N_i)^{s_i(p)} x^{\tau_i(p)} + \sum_{k=0}^{s_i(p)-1} (M_i^{-1}N_i)^k M_i^{-1}(C_i x^{\tau_i(p)} + G(x^{\tau_i(p)})) \right\} + \sum_{i \notin J(p)} E_i x^p \quad (3)$$

and

$$\begin{aligned} x^{p+1} &= \sum_{i \in J(p)} E_i (M_i(\gamma,\omega)^{-1}N_i(\gamma,\omega))^{s_i(p)} x^{\tau_i(p)} + \sum_{i \notin J(p)} E_i x^p \\ &\quad + \sum_{i \in J(p)} E_i \sum_{k=0}^{s_i(p)-1} (M_i(\gamma,\omega)^{-1}N_i(\gamma,\omega))^k M_i(\gamma,\omega)^{-1}(C_i x^{\tau_i(p)} + G(x^{\tau_i(p)})), \end{aligned} \quad (4)$$

respectively.

3. Preliminaries

The representations $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n\times n}$ mean that the vector x and the matrix A are understood in the pointwise sense, while the representations $x \in V_n$ and $A \in \mathbb{L}_n$ mean that the vector x and the matrix A are understood in the blockwise sense. A vector $x \in \mathbb{R}^n$ is said nonnegative (positive), denoted $x \geqslant 0$ $(x > 0)$, if all its components are nonnegative (positive). For $x, y \in \mathbb{R}^n$, $x \geqslant y$ $(x > y)$ means that $x - y \geqslant 0$ $(x - y > 0)$, and $|x|$ denotes the point absolute value of x, which is defined to be the vector whose components are the absolute values of the corresponding components of x. These definitions carry over immediately to matrices. A nonsingular matrix $A \in \mathbb{R}^{n\times n}$ is called an M-matrix if it has non-positive off-diagonal entries and it is monotone (i.e., $A^{-1} \geqslant 0$). By $\rho(A)$ we denote the spectral radius of the square matrix A. Let $D_A = \operatorname{diag}(A)$ and $B_A = D_A - A$. Then A is an M-matrix if and only if D_A is positive diagonal, B_A is nonnegative and $\rho(D_A^{-1}B_A) < 1$. For the mathematical properties of nonnegative matrix, monotone matrix and M-matrix, we refer the reader to Varga [27] and Ortega and Rheinboldt [26].

We define the following subset of $\mathbb{L}_n$ used in the analysis of iterative methods for block H-matrices:

$$\mathbb{L}_{n,I}(n_1,\ldots,n_N) = \{A = (A_{\ell j}) \in \mathbb{L}_n \,|\, A_{\ell\ell} \in \mathbb{R}^{n_\ell \times n_\ell} \text{ nonsingular},\ \ell = 1,\ldots,N\}.$$

Again, we do not write the parameters $(n_1,\ldots,n_N)$, when they are clear from the context. For a matrix $A \in \mathbb{L}_n$, let $D(A) = \operatorname{Diag}(A_{11},\ldots,A_{NN})$, i.e., its block-diagonal part. Thus, $A \in \mathbb{L}_{n,I}$ if and only if $A \in \mathbb{L}_n$ and $D(A)$ is nonsingular.

For any matrix $A = (a_{\ell j}) \in \mathbb{R}^{n\times n}$, we define its comparison matrix $\wp(A) = (\wp(A)_{\ell j})$ by $\wp(A)_{\ell\ell} = |a_{\ell\ell}|$, $\wp(A)_{\ell j} = -|a_{\ell j}|, \ell \neq j$. Similarly, for $A \in \mathbb{L}_{n,I}$ we define its type-I and type-II comparison matrices $\langle A\rangle = (\langle A\rangle_{\ell j}) \in \mathbb{R}^{N\times N}$ and $\langle\langle A\rangle\rangle = (\langle\langle A\rangle\rangle_{\ell j}) \in \mathbb{R}^{N\times N}$ as $\langle A\rangle_{\ell\ell} = \|A_{\ell\ell}^{-1}\|^{-1}$, $\langle A\rangle_{\ell j} = -\|A_{\ell j}\|$, $\ell \neq j$, and $\langle\langle A\rangle\rangle_{\ell\ell} = 1$, $\langle\langle A\rangle\rangle_{\ell j} = -\|A_{\ell\ell}^{-1}A_{\ell j}\|$, $\ell \neq j$, $\ell, j = 1,2,\ldots,N$, respectively. We also define, for $A \in \mathbb{L}_n$, the block absolute value $[A] = (\|A_{\ell j}\|) \in \mathbb{R}^{N\times N}$. The definition for a vector $v \in V_n$ is analogous. Here $\|\cdot\|$ is some consistent matrix norm satisfying $\|I\| = 1$, which could be defined by $\|A_{\ell j}\| = \max_{\|x_j\|_j = 1} \|A_{\ell j}x_j\|_\ell$, $\ell, j = 1,2,\ldots,N$, with $\|\bullet\|_\ell$ some vector norm in $\mathbb{R}^{n_\ell}$, $\ell = 1,2,\ldots,N$. This block absolute value has the following properties.

Lemma 1 (Bai [5]). *Let $A, B \in \mathbb{L}_n$, $x, y \in V_n$ and $\gamma \in \mathbb{R}$. Then,*
(a) $|[A] - [B]| \leqslant [A + B] \leqslant [A] + [B]$, $|[x] - [y]| \leqslant [x + y] \leqslant [x] + [y]$,
(b) $[AB] \leqslant [A][B]$, $[Ax] \leqslant [A][x]$, *and*
(c) $[\gamma A] \leqslant |\gamma|[A]$, $[\gamma x] \leqslant |\gamma|[x]$.

$A \in \mathbb{R}^{n\times n}$ is said to be an H-matrix if $\wp(A)$ is an M-matrix $(A \in H_P)$. $A \in \mathbb{L}_{n,\mathrm{I}}$ is said to be a Type-I (Type-II) block H-matrix if $\langle A\rangle$ $(\langle\langle A\rangle\rangle)$ is an M-matrix in $\mathbb{R}^{N\times N}$ $(A \in H_B^{\mathrm{I}}$ $(A \in H_B^{\mathrm{II}}))$. It follows that $H_B^{\mathrm{I}} \subset H_B^{\mathrm{II}}$ with the inclusion being strict. For $A \in \mathbb{R}^{n\times n}$, the representation $A = M - N$ is called a splitting if M is nonsingular. It is called a convergent splitting if $\rho(M^{-1}N) < 1$. For $A \in \mathbb{L}_{n,I}$, a splitting $A = M - N$ is called an H_B^{I}-compatible (H_B^{II}-compatible) splitting if $\langle A\rangle = \langle M\rangle - [N]$ $(\langle\langle A\rangle\rangle = \langle\langle M\rangle\rangle - [D(M)^{-1}N])$. We refer the readers to [1,16,23,28] for different concepts and their motivations about block M-matrices and block H-matrices. The sequel convergence results about the asynchronous multisplitting block two-stage methods can be straightforwardly generalized to these variants of block H-matrices, with slight and technical modificatons. A mapping $G : V_n \to V_n$ is called block P-bounded if there exists a nonnegative matrix $P \in \mathbb{R}^{N\times N}$ such that $[G(x) -$

$G(y)]\leqslant P[x-y]$ holds for all $x, y \in V_n$. We point out that these concepts naturally reduce to the standard point ones in [15,20,22,27] when the parameters $(n_1,\ldots,n_N)$ are specialized to $(1,\ldots,1)$.

Lemma 2. (a) *If $A\in H_P$, then $|A^{-1}|\leqslant \wp(A)^{-1}$.*
(b) *If $A\in H_B^{\mathrm{I}}\subset \mathbb{L}_{n,I}$, then $[A^{-1}]\leqslant \langle A\rangle^{-1}$* [5].
(c) *If $A\in H_B^{\mathrm{II}}\subset \mathbb{L}_{n,I}$, then $[A^{-1}]\leqslant \langle\langle A\rangle\rangle^{-1}[D(A)^{-1}]$* [5].

We remark that when $n_1=n_2=\cdots=n_N=1$ and $N=n$, both H_B^{I}- and H_B^{II}-matrix classes reduce to the H_P-matrix class, and both statements (b) and (c) reduce to the statement (a), too.

Lemma 3 (Bai [5]). *Let $A=M-N$ be a splitting. If the splitting is H_B^{I}-compatible (H_B^{II}-compatible), then both A and $M\in H_B^{\mathrm{I}}$ ($\in H_B^{\mathrm{II}}$).*

Lemma 4 (Bai et al. [12]). *Let $A\in\mathbb{R}^{n\times n}$ be nonsingular. Then the block system of weakly nonlinear equations* (1) *has a unique solution provided either of the following two conditions holds*:
(a) *$A\in H_B^{\mathrm{I}}$, G is block P-bounded, and $\rho(\langle A\rangle^{-1}P)<1$.*
(b) *$A\in H_B^{\mathrm{II}}$, G is block P-bounded, and $\rho(\langle\langle A\rangle\rangle^{-1}[D(A)^{-1}]P)<1$.*

Lemma 5 (Bai et al. [11]). *Let $\{H_i^{(p)}\}_{p\in N_0}$ $(i=1,2,\ldots,\alpha)$ be sequences of nonnegative matrices in $\mathbb{R}^{N\times N}$, $\tilde{E}_i$ $(i=1,2,\ldots,\alpha)$ be nonnegative diagonal matrices in $\mathbb{R}^{N\times N}$ satisfying $\sum_{i=1}^{\alpha}\tilde{E}_i\leqslant I$, and $\{\varepsilon^p\}_{p\in N_0}$ be sequence in $\mathbb{R}^N$ defined by*

$$\varepsilon^{p+1}=\sum_{i\in J(p)}\tilde{E}_i\varepsilon^{\tau_i(p)}+\sum_{i\notin J(p)}\tilde{E}_i\varepsilon^{p},\qquad p=0,1,2,\ldots$$

with $\{J(p)\}_{p\in N_0}$ and $\{\tau_i(p)\}_{p\in N_0}$ $(i=1,2,\ldots,\alpha)$ being described in Section 2. Then $\lim_{p\to\infty}\varepsilon^p=0$ holds for any $\varepsilon^0\in\mathbb{R}^N$, provided there exist a constant $\theta\in[0,1)$ and a positive vector *$v\in\mathbb{R}^N$ such that $H_i^{(p)}v\leqslant\theta v$ $(i=1,2,\ldots,\alpha,\ p\in N_0)$.*

Lemma 5 presents a basic criterion for examining the convergence of an asynchronous matrix multisplitting iterative method for the monotone matrix class. An improvement and generalization of this result to an asynchronous multisplitting iterative method with dynamic multiple splittings and to a group of linear fixed-point equations with a common fixed-point was studied in Bai et al. [10].

4. The global convergence analyses

In this section, we will prove the global convergence of Methods 1 and 2 for *any* number of inner iterations when the matrix $A\in\mathbb{L}_n$ is a block matrix of different types and when the mapping $G:V_n\to V_n$ is a block P-bounded mapping.

Theorem 6. *Let $A\in H_B^{\mathrm{I}}$ $(H_B^{\mathrm{II}})\subset\mathbb{L}_{n,I}(n_1,\ldots,n_N)$. Let the splittings $A=B_i-C_i$ and $B_i=M_i-N_i$, $i=1,2,\ldots,\alpha$, be H_B^{I}-compatible (H_B^{II}-compatible and such that $D(M_i)=D(B_i)=D(A)$), and the weighting matrices E_i, $i=1,2,\ldots,\alpha$, satisfy $\sum_{i=1}^{\alpha}[E_i]\leqslant I$. Assume further that $G:V_n\to V_n$ is a block*

P-bounded mapping such that $\rho(\langle A\rangle^{-1}P) < 1(\rho(\langle\langle A\rangle\rangle^{-1}[D(A)^{-1}]P) < 1)$. Then, the asynchronous multisplitting block two-stage Method 1 converges to the unique solution of the system of weakly nonlinear equations (1), for any initial vector $x^0 \in V_n$ and any sequence of numbers of inner iterations $s_i(p) \geqslant 1,\ i = 1,2,\dots,\alpha,\ p \in N_0$.

Proof. Lemma 4 implies that (both in the Type-I and Type-II cases) there exists a unique vector $x^* \in V_n$ such that $Ax^* = G(x^*)$ under the assumptions of this theorem. Thus, if we let $\varepsilon^p = x^p - x^*$ be the error at the pth iteration of Method 1, then according to (3) $\{\varepsilon^p\}_{p\in N_0}$ satisfies

$$\varepsilon^{p+1} = \sum_{i\in J(p)} E_i \left\{ (M_i^{-1}N_i)^{s_i(p)}\varepsilon^{\tau_i(p)} + \sum_{k=0}^{s_i(p)-1} (M_i^{-1}N_i)^k M_i^{-1}(C_i\varepsilon^{\tau_i(p)} + G(x^{\tau_i(p)}) - G(x^*)) \right\} + \sum_{i\notin J(p)} E_i\varepsilon^p. \tag{5}$$

For the Type-I case, we note that the H_B^{I}-compatibility of the splittings imply that $M_i \in H_B^{\mathrm{I}}$, and thus by Lemma 2(b), we have $[M_i^{-1}] \leqslant \langle M_i\rangle^{-1},\ i = 1,2,\dots,\alpha$. Using these inequalities, by taking block absolute values on both sides of (5), applying Lemma 1 and the block P-bounded property of G, we obtain the inequality

$$[\varepsilon^{p+1}] \leqslant \sum_{i\in J(p)} [E_i]T_{i,1}^{(p)}[\varepsilon^{\tau_i(p)}] + \sum_{i\notin J(p)} [E_i][\varepsilon^p], \quad p \in N_0, \tag{6}$$

where

$$T_{i,1}^{(p)} = (\langle M_i\rangle^{-1}[N_i])^{s_i(p)} + \sum_{k=0}^{s_i(p)-1} (\langle M_i\rangle^{-1}[N_i])^k \langle M_i\rangle^{-1}([C_i] + P) \tag{7}$$

is a nonnegative matrix in $\mathbb{R}^{N\times N}$. Because

$$(\langle M_i\rangle^{-1}[N_i])^{s_i(p)} + \sum_{k=0}^{s_i(p)-1} (\langle M_i\rangle^{-1}[N_i])^k \langle M_i\rangle^{-1}[C_i] = I - \sum_{k=0}^{s_i(p)-1} (\langle M_i\rangle^{-1}[N_i])^k \langle M_i\rangle^{-1}\langle A\rangle,$$

we can rewrite (7) as

$$T_{i,1}^{(p)} = I - \sum_{k=0}^{s_i(p)-1} (\langle M_i\rangle^{-1}[N_i])^k \langle M_i\rangle^{-1}(\langle A\rangle - P). \tag{8}$$

Since $\langle A\rangle^{-1}$ and P are nonnegative matrices and by the hypothesis $\rho(\langle A\rangle^{-1}P) < 1$, the matrix $(\langle A\rangle - P)$ is a monotone matrix. Let $e = (1,1,\dots,1)^{\mathrm{T}} \in \mathbb{R}^N$ and $v = (\langle A\rangle - P)^{-1}e$. Then as $(\langle A\rangle - P)^{-1} \geqslant 0$ and no row of $(\langle A\rangle - P)^{-1}$ can have all null entries, we get $v > 0$. The same arguments result in $M_i^{-1}e > 0,\ i = 1,2,\dots,\alpha$. Then, from (8) we have

$$\begin{aligned} T_{i,1}^{(p)}v &= v - \sum_{k=0}^{s_i(p)-1} (\langle M_i\rangle^{-1}[N_i])^k \langle M_i\rangle^{-1}e \\ &= v - \langle M_i\rangle^{-1}e - \sum_{k=1}^{s_i(p)-1} (\langle M_i\rangle^{-1}[N_i])^k \langle M_i\rangle^{-1}e. \end{aligned}$$

Because the matrices $\langle M_i\rangle^{-1}$ and $[N_i]$ are nonnegative, it follows that $T_{i,\mathrm{I}}^{(p)}v \leqslant v - \langle M_i\rangle^{-1}e$. Moreover, since $v - \langle M_i\rangle^{-1}e < v$, there exist constants $\theta_i \in [0,1)$, $i=1,2,\ldots,\alpha$, such that $T_{i,\mathrm{I}}^{(p)}v \leqslant \theta_i v$, $i=1,2,\ldots,\alpha$, $p\in N_0$. Hence, by setting $\theta=\max_{1\leqslant i\leqslant\alpha}\{\theta_i\}$, we get

$$T_{i,\mathrm{I}}^{(p)}v \leqslant \theta v, \quad i=1,2,\ldots,\alpha,\ p\in N_0, \tag{9}$$

where $\theta\in[0,1)$ is a constant and $v\in\mathbb{R}^N$ is a positive vector. Note that $T_{i,\mathrm{I}}^{(p)}\geqslant 0$, $i=1,2,\ldots,\alpha$, $p\in N_0$.

Now, defining the sequence $\{\varepsilon^p\}_{p\in N_0}$ according to

$$\varepsilon^0=[\varepsilon^0], \quad \varepsilon^{p+1}=\sum_{i\in J(p)}\tilde{E}_i T_{i,\mathrm{I}}^{(p)}\varepsilon^{\tau_i(p)}+\sum_{i\notin J(p)}\tilde{E}_i\varepsilon^p, \quad p\in N_0,$$

where $\tilde{E}_i=[E_i]$, $i=1,2,\ldots,\alpha$, we can immediately deduce that $\{\varepsilon^p\}_{p\in N_0}$ is a majorizing sequence of $\{[\varepsilon^p]\}_{p\in N_0}$. That is to say, $[\varepsilon^p]\leqslant\varepsilon^p$ holds for all $p\in N_0$. By making use of Lemma 5 we immediately know that $\lim_{p\to\infty}\varepsilon^p=0$. Therefore, $\lim_{p\to\infty}[\varepsilon^p]=0$ and then, $\lim_{p\to\infty}\varepsilon^p=0$. This fulfils the proof for Type-I case.

For the Type-II case, let us denote by $\tilde{P}=[D(A)^{-1}]P$, and for $i=1,2,\ldots,\alpha$, $\tilde{B}_i=D(A)^{-1}B_i$, $\tilde{C}_i=D(A)^{-1}C_i$, $\tilde{M}_i=D(A)^{-1}M_i$ and $\tilde{N}_i=D(A)^{-1}N_i$. Observing that $\tilde{M}_i\in H_B^{\mathrm{I}}$, $i=1,2,\ldots,\alpha$, by Lemma 2(b), we have $[\tilde{M}_i^{-1}]\leqslant\langle\tilde{M}_i\rangle^{-1}$, $i=1,2,\ldots,\alpha$. By making use of these relations, taking block absolute values on both sides of (5) as before, after inserting $D(A)D(A)^{-1}$ in the appropriate places, we obtain

$$[\varepsilon^{p+1}]\leqslant\sum_{i\in J(p)}[E_i]T_{i,\mathrm{II}}^{(p)}[\varepsilon^{\tau_i(p)}]+\sum_{i\notin J(p)}[E_i][\varepsilon^p], \quad p\in N_0,$$

where

$$T_{i,\mathrm{II}}^{(p)}=(\langle\tilde{M}_i\rangle^{-1}[\tilde{N}_i])^{s_i(p)}+\sum_{k=0}^{s_i(p)-1}(\langle\tilde{M}_i\rangle^{-1}[\tilde{N}_i])^k\langle\tilde{M}_i\rangle^{-1}([\tilde{C}_i]+\tilde{P}).$$

This expression has the same form as (6), with matrices of the same structure as (7). Noticing that the splittings $D(A)^{-1}A=\tilde{B}_i-\tilde{C}_i$, and $\tilde{B}_i=\tilde{M}_i-\tilde{N}_i$ are H_B^{I}-compatible and they correspond to the system $D(A)^{-1}Ax=D(A)^{-1}G(x)$, which satisfy the hypotheses of the Type-I case, we therefore complete the proof. □

Theorem 7. *Let $A\in H_B^{\mathrm{I}}(H_B^{\mathrm{II}})\subset\mathbb{L}_{n,\mathrm{I}}(n_1,\ldots,n_N)$. Let the splittings $A=B_i-C_i$, $i=1,2,\ldots,\alpha$, be H_B^{I}-compatible (H_B^{II}-compatible) such that $D(B_i)=D(A)$, $i=1,2,\ldots,\alpha$, the splittings $B_i=D_i-L_i-U_i$, $i=1,2,\ldots,\alpha$, satisfy*

$$\langle B_i\rangle=\langle D_i\rangle-[L_i]-[U_i], \quad i=1,2,\ldots,\alpha$$

for Type-I case (and

$$\langle\langle B_i\rangle\rangle=I-[D_i^{-1}L_i]-[D_i^{-1}U_i], \quad i=1,2,\ldots,\alpha$$

for Type-II case), and the weighting matrices E_i, $i=1,2,\ldots,\alpha$, satisfy $\sum_{i=1}^{\alpha}[E_i]\leqslant I$. Assume further that $G:V_n\to V_n$ is a block P-bounded mapping such that $\rho(\langle A\rangle^{-1}P)<1(\rho(\langle\langle A\rangle\rangle^{-1}[D(A)^{-1}]P)<1)$.

Then, the asynchronous multisplitting block two-stage Method 2 *converges to the unique solution of the system of weakly nonlinear equations* (1), *for any initial vector* $x^0 \in V_n$ *and any sequence of numbers of inner iterations* $s_i(p) \geqslant 1,\ i = 1, 2, \ldots, \alpha,\ p \in N_0$, *provided the relaxation parameters* γ *and* ω *satisfy* $0 \leqslant \gamma \leqslant \omega$ *and*

$$0 < \omega < \frac{2}{1 + \rho(D_{\langle A\rangle}^{-1}(B_{\langle A\rangle} + P))} \left(0 < \omega < \frac{2}{1 + \rho(B_{\langle A\rangle} + [D(A)^{-1}]P)}\right).$$

Proof. Let us denote, again, by $\tilde{P} = [D(A)^{-1}]P$, and for $i = 1, 2, \ldots, \alpha,\ \tilde{B}_i = D(A)^{-1}B_i,\ \tilde{C}_i = D(A)^{-1}C_i,\ \tilde{L}_i = D(A)^{-1}L_i,\ \tilde{U}_i = D(A)^{-1}U_i$, and

$$\tilde{M}_i(\gamma, \omega) = \frac{1}{\omega}(I - \gamma\tilde{L}_i),$$
$$\tilde{N}_i(\gamma, \omega) = \frac{1}{\omega}((1 - \omega)I + (\omega - \gamma)\tilde{L}_i + \omega\tilde{U}_i).$$

Then we have

$$\begin{aligned} &\langle \tilde{M}_i(\gamma, \omega)\rangle = \langle\langle \tilde{M}_i(\gamma, \omega)\rangle\rangle = \frac{1}{\omega}(I - \gamma[\tilde{L}_i]) \equiv \mathscr{M}_i(\gamma, \omega), \\ &[\tilde{N}_i(\gamma, \omega)] \leqslant \frac{1}{\omega}(|1 - \omega|I + (\omega - \gamma)[\tilde{L}_i] + \omega[\tilde{U}_i]) \equiv \mathscr{N}_i(\gamma, \omega). \end{aligned} \tag{10}$$

Writing $\mathscr{C}_i = [\tilde{C}_i]$ and $\mathscr{P} = \tilde{P}$, following similar demonstration to the proof of Theorem 6 we know from (4) that the error $\{\varepsilon^p\}_{p \in N_0}$ of Method 2 satisfies

$$[\varepsilon^{p+1}] \leqslant \sum_{i \in J(p)} [E_i] T_i^{(p)} [\varepsilon^{\tau_i(p)}] + \sum_{i \notin J(p)} [E_i][\varepsilon^p], \quad p \in N_0 \tag{11}$$

for both Type-I and Type-II cases, where

$$\begin{aligned} T_i^{(p)} &= (\mathscr{M}_i(\gamma, \omega)^{-1}\mathscr{N}_i(\gamma, \omega))^{s_i(p)} + \sum_{k=0}^{s_i(p)-1} (\mathscr{M}_i(\gamma, \omega)^{-1}\mathscr{N}_i(\gamma, \omega))^k \mathscr{M}_i(\gamma, \omega)^{-1}(\mathscr{C}_i + \mathscr{P}) \\ &= I - \sum_{k=0}^{s_i(p)-1} (\mathscr{M}_i(\gamma, \omega)^{-1}\mathscr{N}_i(\gamma, \omega))^k \mathscr{M}_i(\gamma, \omega)^{-1}(\langle\langle A\rangle\rangle - \mathscr{P}). \end{aligned} \tag{12}$$

For Type-I case, we note that $\langle\langle A\rangle\rangle \geqslant I - D_{\langle A\rangle}^{-1}B_{\langle A\rangle}$ and for $i = 1, 2, \ldots, \alpha,\ [\tilde{L}_i] \leqslant D_{\langle A\rangle}^{-1}[L_i],\ [\tilde{U}_i] \leqslant D_{\langle A\rangle}^{-1}[U_i]$. Therefore,

$$\mathscr{M}_i(\gamma, \omega) \geqslant \frac{1}{\omega}(I - \gamma D_{\langle A\rangle}^{-1}[L_i]) \equiv \hat{\mathscr{M}}_i(\gamma, \omega),$$
$$\mathscr{N}_i(\gamma, \omega) \leqslant \frac{1}{\omega}(|1 - \omega|I + (\omega - \gamma)D_{\langle A\rangle}^{-1}[L_i] + \omega D_{\langle A\rangle}^{-1}[U_i]) \equiv \hat{\mathscr{N}}_i(\gamma, \omega).$$

Since for $i=1,2,\ldots,\alpha$, $\hat{\mathscr{M}}_i(\gamma,\omega)$ are M-matrices satisfying $\hat{\mathscr{M}}_i(\gamma,\omega)^{-1}\geqslant\mathscr{M}_i(\gamma,\omega)^{-1}$ and $\hat{\mathscr{N}}_i(\gamma,\omega)$ are nonnegative matrices, from (12) we have

$$\begin{aligned}T_i^{(p)} &\leqslant (\hat{\mathscr{M}}_i(\gamma,\omega)^{-1}\hat{\mathscr{N}}_i(\gamma,\omega))^{s_i(p)} + \sum_{k=0}^{s_i(p)-1}(\hat{\mathscr{M}}_i(\gamma,\omega)^{-1}\hat{\mathscr{N}}_i(\gamma,\omega))^k\hat{\mathscr{M}}_i(\gamma,\omega)^{-1}D_{\langle A\rangle}^{-1}(\hat{\mathscr{C}}_i+P)\\ &= I - \sum_{k=0}^{s_i(p)-1}(\hat{\mathscr{M}}_i(\gamma,\omega)^{-1}\hat{\mathscr{N}}_i(\gamma,\omega))^k\hat{\mathscr{M}}_i(\gamma,\omega)^{-1}D_{\langle A\rangle}^{-1}(\langle A\rangle - P)\\ &\equiv \hat{T}_i^{(p)},\end{aligned}$$

where $\hat{\mathscr{C}}_i=[C_i]$, $i=1,2,\ldots,\alpha$. Eq. (11) now gives

$$[\varepsilon^{p+1}]\leqslant\sum_{i\in J(p)}[E_i]\hat{T}_i^{(p)}[\varepsilon^{\tau_i(p)}]+\sum_{i\notin J(p)}[E_i][\varepsilon^p],\quad p\in N_0. \tag{13}$$

If we define

$$\begin{cases}\hat{\mathbb{A}}(\omega)=\dfrac{1-\omega-|1-\omega|}{\omega}I+D_{\langle A\rangle}^{-1}(\langle A\rangle-P),\\ \hat{\mathbb{B}}_i(\omega)=\dfrac{1-|1-\omega|}{\omega}I-D_{\langle A\rangle}^{-1}([L_i]+[U_i]),\quad i=1,2,\ldots,\alpha,\\ \hat{\mathbb{C}}_i(\omega)=D_{\langle A\rangle}^{-1}(\hat{\mathscr{C}}_i+P),\end{cases}$$

then it holds that

$$\hat{\mathbb{A}}(\omega)=\hat{\mathbb{B}}_i(\omega)-\hat{\mathbb{C}}_i(\omega),\quad \hat{\mathbb{B}}_i(\omega)=\hat{\mathscr{M}}_i(\gamma,\omega)-\hat{\mathscr{N}}_i(\gamma,\omega),\quad i=1,2,\ldots,\alpha.$$

Clearly, $\hat{\mathbb{C}}_i(\omega)\geqslant 0$, $i=1,2,\ldots,\alpha$; and $\hat{\mathbb{B}}_i(\omega)=\hat{\mathscr{M}}_i(\gamma,\omega)-\hat{\mathscr{N}}_i(\gamma,\omega)$, $i=1,2,\ldots,\alpha$, are M-splittings. In addition, the hypothesis $\rho(\langle A\rangle^{-1}P)<1$ immediately implies that $D_{\langle A\rangle}^{-1}(\langle A\rangle-P)$ is a monotone matrix. Hence, we see that $\hat{\mathbb{A}}(\omega)$ is an M-matrix. Moreover, the inequalities

$$\hat{\mathbb{A}}(\omega)\leqslant\hat{\mathbb{B}}_i(\omega)\leqslant I,\quad i=1,2,\ldots,\alpha$$

show that $\hat{\mathbb{B}}_i(\omega)$, $i=1,2,\ldots,\alpha$, are M-matrices. Up to now, similar to the proof of Theorem 6, we can demonstrate that there exist a constant $\theta\in[0,1)$ and a positive vector $v\in\mathbb{R}^n$ such that

$$\hat{T}_i^{(p)}v\leqslant\theta v,\quad i=1,2,\ldots,\alpha,\quad p\in N_0.$$

These estimates, together with (13) and Lemma 5, directly result in $\lim_{p\to\infty}\varepsilon^p=0$ for Type-I case.

The investigation of Type-II case is quite simple. It follows from (10) that

$$\mathscr{M}_i(\gamma,\omega)-\mathscr{N}_i(\gamma,\omega)=\frac{1-\omega-|1-\omega|}{\omega}I+\langle\langle\tilde{B}_i\rangle\rangle,\quad i=1,2,\ldots,\alpha.$$

By making use of these identities as well as (11) and (12), analogously to the proof of Theorem 6 we can also get $\lim_{p\to\infty}\varepsilon^p=0$ for Type-II case. □

5. Numerical results

Consider the system of weakly nonlinear equations (1) with $A=(A_{\ell j})\in\mathbb{L}_n(\tilde n,\dots,\tilde n)$,

$$A_{\ell j}=\begin{cases}\tilde B & \text{if } \ell=j,\\ -I & \text{if } |\ell-j|=1, \quad \ell,j=1,2,\dots,\tilde n,\\ 0 & \text{otherwise}\end{cases}$$

and with $G(x)\in V_n(\tilde n,\dots,\tilde n)$,

$$G_j(x)=(g_{(j-1)\tilde n+1}(x),g_{(j-1)\tilde n+2}(x),\dots,g_{j\tilde n}(x))^{\mathrm T},$$

where $\tilde B=\mathrm{tridiag}(-1,4+10h^2,-1)\in\mathbb{R}^{\tilde n\times\tilde n}$, $h=1/(\tilde n+1)^2$ and

$$\begin{cases}g_1(x)=h^2(|x_1|+\beta e^{(\beta-1)|x_1|}\sin x_1)-10,\\ g_j(x)=h^2(|x_j|+\beta e^{(\beta-1)|x_j|}\sin x_j\cos x_{j-1}), \quad j=2,3,\dots,n-1,\\ g_n(x)=h^2(|x_n|+\beta e^{(\beta-1)|x_n|}\sin x_n\cos x_{n-1})+10.\end{cases}$$

Evidently, we have $N=\tilde n$, $n_j=\tilde n$ $(j=1,2,\dots,\tilde n)$ and $n=\tilde n^2$. We solve this system of weakly nonlinear equations on the SGI Power Challenge multiprocessor having four processors.

The computations are done for the new asynchronous multisplitting block two-stage AOR method 2 (Block AMTS-AOR method), its corresponding point version in [9], i.e., the asynchronous multisplitting two-stage AOR method (Point AMTS-AOR method), and its corresponding synchronous version in [12], i.e., the synchronous multisplitting block two-stage AOR method (Block SMTS-AOR method). These three methods are implemented as PVM applications and tested on the afore-mentioned parallel computer. In our computations, the splitting matrices for the block methods are taken to be

$$B_i=(B_{\ell j}^{(i)})\in\mathbb{L}_n(\tilde n,\dots,\tilde n),$$

$$B_{\ell j}^{(i)}=\begin{cases}A_{\ell j} & \text{if } \ell,j\in J_i,\\ A_{\ell\ell} & \text{if } \ell,j\notin J_i \text{ and } \ell=j, \quad \ell,j=1,2,\dots,\tilde n,\\ 0 & \text{otherwise,}\end{cases}$$

$$D_i=\mathrm{Diag}(A_{11},A_{22},\dots,A_{\tilde n\tilde n}),$$

$L_i=$ the strictly block lower-triangular matrices of $(-B_i)$,

$U_i=$ the strictly block upper-triangular matrices of $(-B_i)$,

$$E_i=\mathrm{Diag}(\mu_{(i-1)n_0+1}I,\mu_{(i-1)n_0+2}I,\dots,\mu_{(i+1)n_0}I), \quad \mu_j\in[0,1],$$

where $n_0=\tilde n/(\alpha+1)$; $J_i=\{(i-1)n_0+1,(i-1)n_0+2,\dots,(i+1)n_0\}$; and for $j=1,2,\dots,2n_0$, $\mu_{(i-1)n_0+j}=1$ if $i=1$ and $1\leqslant j\leqslant n_0$ or if $i=\alpha$ and $n_0+1\leqslant j\leqslant 2n_0$, and $\mu_{(i-1)n_0+j}=0.5$ otherwise. The splitting matrices for the point methods can be defined in a similar way except that we take D_i,L_i and U_i to be point matrices, respectively, in the block case. In particular, we point out that these partitions allow that the ith processor of the multiprocessor system solves only the variables located in J_i.

Table 1
Point vs. Block AMTS-AOR method, $\beta = 0.0$

			$\alpha = 1$	$\alpha = 2$	$\alpha = 3$	$\alpha = 4$
Point	$\gamma = 1.0$	Time (ms)	43 432	35 479	31 233	20 509
AMTS-AOR	$\omega = 1.0$	Speedup	1.0	1.22	1.39	2.11
	$\gamma = 0.8$	Time (ms)	42 176	36 190	29 413	21 107
	$\omega = 0.7$	Speedup	1.0	1.17	1.43	2.0
Block	$\gamma = 1.0$	Time (ms)	14 184	9304	7515	6243
AMTS-AOR	$\omega = 1.0$	Speedup	1.0	1.52	1.88	2.27
	$\gamma = 0.9$	Time (ms)	15 043	8962	7403	6001
	$\omega = 1.1$	Speedup	1.0	1.68	2.03	2.51

Table 2
Point vs. Block AMTS-AOR method, $\beta = 1.0$

			$\alpha = 1$	$\alpha = 2$	$\alpha = 3$	$\alpha = 4$
Point	$\gamma = 1.0$	Time (ms)	24 958	17 034	13 467	11 026
AMTS-AOR	$\omega = 1.0$	Speedup	1.0	1.46	1.85	2.26
	$\gamma = 0.9$	Time (ms)	24 101	18 362	13 109	10 816
	$\omega = 1.0$	Speedup	1.0	1.31	1.84	2.23
Block	$\gamma = 1.0$	Time (ms)	13 562	8618	7097	6176
AMTS-AOR	$\omega = 1.0$	Speedup	1.0	1.57	1.91	2.19
	$\gamma = 0.8$	Time (ms)	13 225	8512	7203	5814
	$\omega = 0.8$	Speedup	1.0	1.55	1.84	2.27

Hence, the computation of a single outer iteration takes on one processor $2T_{seq}/(\alpha+1)$ time, where T_{seq} represents the sequential time of computing the outer iteration. Furthermore, since the communication overheads add to the execution time, $(\alpha+1)/2$ represents an upper bound for the expected speed-up of the implementations.

All computations are started from an initial vector having all components equal to -1.0, and terminated once the current iterations x^p obey

$$\frac{\|Ax^p - G(x^p)\|_1}{\|Ax^0 - G(x^0)\|_1} \leqslant 10^{-5}.$$

For $n = \tilde{n}^2 = 6400$, the corresponding timings and speed-ups are listed in Tables 1–4 and plotted in Figs. 1 and 2. Here, the speed-up is defined to be the ratio of the sequential computing time with the corresponding parallel running; and without particular description, the number of inner iterations s is assigned a randomly chosen value from the set $\{1,2,3,4\}$, i.e., $s = \mathrm{rand}(1..4)$, for each outer iteration and on each processor.

Tables 1 and 3 present some numerical computations of both the point and the block AMTS-AOR methods for the cases of $\beta = 0.0$ and $\beta = 1.0$, respectively. Simple comparisons show that the block AMTS-AOR method requires much less CPU to achieve the stopping criterion than the point

Table 3
Block SMTS-AOR method, $\beta = 0.0$

			$\alpha = 1$	$\alpha = 2$	$\alpha = 3$	$\alpha = 4$
$s = 1$	$\gamma = 1.0$	Time (ms)	33 617	25 507	23 196	15 684
	$\omega = 1.0$	Speedup	1.0	1.31	1.44	2.14
	$\gamma = 0.9$	Time (ms)	32 901	25 146	23 214	12 998
	$\omega = 1.0$	Speedup	1.0	1.31	1.42	2.53
$s = 2$	$\gamma = 1.0$	Time (ms)	21 436	15 662	14 893	9954
	$\omega = 1.0$	Speedup	1.0	1.36	1.43	2.15
	$\gamma = 0.9$	Time (ms)	20 404	15 385	14 504	9610
	$\omega = 0.9$	Speedup	1.0	1.33	1.41	2.12
$s = 3$	$\gamma = 1.0$	Time (ms)	17 406	11 639	11 093	7981
	$\omega = 1.0$	Speedup	1.0	1.49	1.56	2.18
	$\gamma = 0.9$	Time (ms)	17 283	10 956	11 205	7503
	$\omega = 1.0$	Speedup	1.0	1.58	1.54	2.30
$s = 4$	$\gamma = 1.0$	Time (ms)	15 439	10 120	8941	7077
	$\omega = 1.0$	Speedup	1.0	1.52	1.72	2.18
	$\gamma = 0.8$	Time (ms)	15 205	9872	8504	6812
	$\omega = 0.9$	Speedup	1.0	1.54	1.79	2.23

Table 4
Block SMTS-AOR method, $\beta = 1.0$

			$\alpha = 1$	$\alpha = 2$	$\alpha = 3$	$\alpha = 4$
$s = 1$	$\gamma = 1.0$	Time (ms)	29 336	26 365	21 044	15 358
	$\omega = 1.0$	Speedup	1.0	1.11	1.39	1.91
	$\gamma = 0.9$	Time (ms)	28 307	26 510	20 834	15 007
	$\omega = 0.9$	Speedup	1.0	1.07	1.36	1.89
$s = 2$	$\gamma = 1.0$	Time (ms)	19 410	15 461	11 283	9516
	$\omega = 1.0$	Speedup	1.0	1.25	1.72	2.03
	$\gamma = 0.9$	Time (ms)	17 120	14 902	10 871	9314
	$\omega = 1.0$	Speedup	1.0	1.15	1.57	1.84
$s = 3$	$\gamma = 1.0$	Time (ms)	16 260	11 349	9237	7262
	$\omega = 1.0$	Speedup	1.0	1.43	1.76	2.23
	$\gamma = 0.9$	Time (ms)	16 126	10 841	8964	7124
	$\omega = 1.0$	Speedup	1.0	1.49	1.80	2.26
$s = 4$	$\gamma = 1.0$	Time (ms)	14 526	10 679	8708	6720
	$\omega = 1.0$	Speedup	1.0	1.36	1.66	2.16
	$\gamma = 0.8$	Time (ms)	13 917	9763	8605	6347
	$\omega = 1.0$	Speedup	1.0	1.43	1.62	2.19

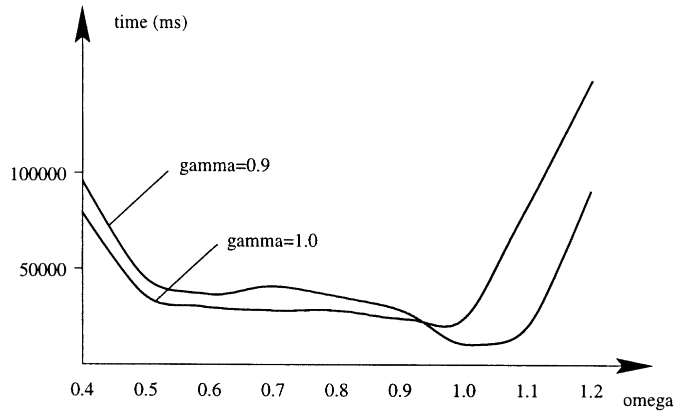

Fig. 1. The execution time for the Block AMTS-AOR method as a function of $\omega \in [0.4, 1.2]$. The other parameters are: $\beta = 1.0$, $s = \text{rand}(1..4)$, $\alpha = 4$.

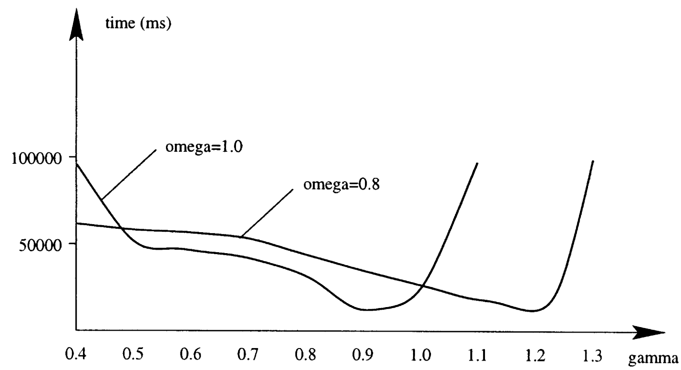

Fig. 2. The execution time for the Block AMTS-AOR method as a function of $\gamma \in [0.4, 1.3]$. The other parameters are: $\beta = 1.0$, $s = \text{rand}(1..4)$, $\alpha = 4$.

AMTS-AOR method, and its speed-up is also much higher than the latter one. Regarding the optimal processor number, four processor case is the best possible one in our experiments since it costs the least CPU and possesses the highest speedup.

Tables 2 and 4 present numerical performance of the block SMTS-AOR method for the cases of $\beta = 0.0$ and 1.0, respectively. The comparisons are focused on the numbers of the inner iterations and the processors, respectively. Roughly speaking, $s = 3, 4$ and $\alpha = 4$ give the best convergence of this method for almost all situations.

From Tables 1–4, it is clear that the block AMTS-AOR method has better numerical property than the block SMTS-AOR method in the senses of both CPU and speedup.

Through depicting the dependence curve of the CPU with respect to the relaxation parameters, Figs. 1 and 2 show that suitable match of the relaxation parameters can considerably accelerate the convergence speed of the asynchronous multisplitting block two-stage AOR method.

Concludingly, in all experiments, the Block AMTS-AOR method requires a significantly smaller number of outer iterations than the Point AMTS-AOR method as well as the Block SMTS-AOR method. For each of these three methods, the number of outer iterations remains almost constant when the number of processors increases. For the Block SMTS-AOR method, the execution time sharply decreases when the number of inner iterations s is increased from 1 to 3, but stabilizes and eventually raises if s is further increased. The initial improvement in the execution time is due to a decrease in the local computation time. We remember that the residual $(Ax - G(x))$, the most expensive computation in each outer iteration, only needs to be computed once per outer iteration. However, by increasing the number of inner iterations beyond 3 or 4, the convergence of the method is affected. When $s = 3, 4$, the Block SMTS-AOR method tends to be slower than the Block AMTS-AOR method, and it is expected that on a distributed memory parallel computer the Block AMTS-AOR method would be significantly better than the more restrictive Block SMTS-AOR method. The Point AMTS-AOR method is the least performant of the three methods. In addition, the numerical results clearly show that suitable choices of the relaxation parameters γ and ω can significantly improve the convergence properties of the parallel multisplitting two-stage AOR methods.

Acknowledgements

The authors are very much indebted to professor T. Yamamoto and the referees for their constructive comments and suggestions which considerably improved the original manuscript of this paper. They also thank professor De-Ren Wang for his valuable discussions on this topic.

References

[1] O. Axelsson, Preconditioning methods for block H-matrices, Computer Algorithms for Solving Linear Algebraic Equations. The State of the Art, NATO ASI Series F77, 1991, pp. 169–184.
[2] J. Bahi, J.C. Miellou, K. Rhofir, Asynchronous multisplitting methods for nonlinear fixed point problems, Numer. Algorithms 15 (1997) 315–345.
[3] Z.-Z. Bai, A class of two-stage iterative methods for systems of weakly nonlinear equations, Numer. Algorithms 14 (1997) 295–319.
[4] Z.-Z. Bai, Parallel multisplitting two-stage iterative methods for large sparse systems of weakly nonlinear equations, Numer. Algorithms 15 (1997) 347–372.
[5] Z.-Z. Bai, Parallel matrix multisplitting block relaxation methods, Mathematica Numerica Sinica 17 (1995) 238–252 (in Chinese).
[6] Z.-Z. Bai, A class of blockwise matrix multi-splitting multi-parameter block relaxation methods, in: Z.C. Shi, M. Mori (Eds.), Proceeding of the Third China–Japan Seminar on Numerical Mathematics, Science Press of China, Beijing/New York, 1998, pp. 1–10.
[7] Z.-Z. Bai, Asynchronous parallel matrix multisplitting block relaxation iteration methods, Numer. Math. J. Chinese Univ. (Chinese Series) 19 (1997) 28–39 (in Chinese).
[8] Z.-Z. Bai, A class of asynchronous parallel multisplitting blockwise relaxation methods, Parallel Computing, 1999, accepted for publication.
[9] Z.-Z. Bai, Y.-G. Huang, Asynchronous multisplitting two-stage iterations for systems of weakly nonlinear equations, J. Comput. Appl. Math. 93 (1) (1998) 13–33.
[10] Z.-Z. Bai, J.-C. Sun, D.-R. Wang, A unified framework for the construction of various matrix multisplitting iterative methods for large sparse system of linear equations, Comput. Math. Appl. 32 (12) (1996) 51–76.
[11] Z.-Z. Bai, D.-R. Wang, D.J. Evans, A class of asynchronous matrix multi-splitting multi-parameter relaxation iterations, J. Comput. Math. 16 (3) (1998) 221–238.
[12] Z.-Z. Bai, V. Migallón, J. Penadés, D.B. Szyld, Block and asynchronous two-stage methods for mildly nonlinear systems, Numerische Mathematik 82 (1) (1999) 1–20.

[13] D. El Baz, P. Spiteri, J.C. Miellou, Asynchronous multisplitting methods with flexible communication for pseudolinear P.D.E., in: D. Bainov (Ed.), Proceedings of the Eighth International Colloquium on Differential Equations, VSP, Utrecht, 1998, pp. 145–152.
[14] M. Benzi, H. Choi, D.B. Szyld, Threshold ordering for preconditioning nonsymmetric problems, in: G.H. Golub et al. (Eds.), Proceedings of the Scientific Computing Workshop'97 Hong Kong, Springer, Singapore, 1997, pp. 159–165.
[15] R. Bru, L. Elsner, M. Neumann, Models of parallel chaotic iteration methods, Linear Algebra Appl. 103 (1988) 175–192.
[16] L. Elsner, V. Mehrmann, Convergence of block iterative methods for linear systems arising in the numerical solution of Euler equation, Numerische Mathematik 59 (1991) 541–559.
[17] D.J. Evans, Z.-Z. Bai, Blockwise matrix multi-splitting multi-parameter block relaxation methods, Int. J. Comput. Math. 64 (1997) 103–118.
[18] A. Frommer, H. Schwandt, A unified representation and theory of algebraic additive Schwarz and multisplitting methods, SIAM J. Matrix Anal. Appl. 18 (1997) 893–912.
[19] A. Frommer, H. Schwandt, D.B. Szyld, Asynchronous weighted additive Schwarz methods, Electron Trans. Numer. Anal. 5 (1997) 48–61.
[20] A. Frommer, D.B. Szyld, H-splittings and two-stage iterative methods, Numerische Mathematik 63 (1992) 345–356.
[21] A. Frommer, D.B. Szyld, Asynchronous two-stage iterative methods, Numerische Mathematik 69 (1994) 141–153.
[22] M.T. Jones, D.B. Szyld, Two-stage multisplitting methods with overlapping blocks, Numer. Linear Algebra Appl. 3 (1996) 113–124.
[23] R. Nabben, On a class of matrices which arise in the numerical solution of Euler equations, Numerische Mathematik 63 (1992) 411–431.
[24] D.P. O'Leary, R.E. White, Multi-splittings of matrices and parallel solution of linear systems, SIAM J. Algebraic Discrete Meth. 6 (1985) 630–640.
[25] J. O'Neil, D.B. Szyld, A block ordering method for sparse matrices, SIAM J. Sci. Stat. Comput. 11 (1990) 811–823.
[26] J.M. Ortega, W.C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, San Diego, 1970.
[27] R.S. Varga, Matrix Iterative Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1962.
[28] Z.-Y. You, C.-L. Wang, The further generalization for block H-matrix, Pure Appl. Math. 11 (1995) 41–44.