ELSEVIER

Parallel Computing 25 (1999) 681–701

PARALLEL COMPUTING

www.elsevier.com/locate/parco

A class of asynchronous parallel multisplitting blockwise relaxation methods [1]

Zhong-Zhi Bai [2]

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

Received 22 April 1997; received in revised form 9 December 1998

Abstract

By the principle of using sufficiently the delayed information and based on the technique of successively accelerated overrelaxation (AOR), we set up a class of asynchronous multisplitting blockwise relaxation methods for solving the large sparse blocked system of linear equations, which comes from the discretizations of many differential equations. These new methods are efficient blockwise variants of the asynchronous parallel matrix multisplitting relaxed iterations discussed by Bai et al. (Parallel Computing 21 (1995) 565–582), and they are very smart for implementations on the MIMD multiprocessor systems. Under reasonable restrictions on the relaxation parameters as well as the multiple splittings, we establish the convergence theories of this class of new methods when the coefficient matrices of the blocked systems of linear equations are block H-matrices of different types. A lot of numerical experiments show that our new methods are applicable and efficient, and have better numerical behaviours than their pointwise alternatives investigated by Bai et al.

Keywords: Blocked system of linear equations; Block matrix multisplitting; Blockwise relaxation; Asynchronous parallel method; Convergence theory; Block H-matrix

1. Introduction

The matrix multisplitting methodology was originally invented by O'Leary and White (see Ref. [16]) in 1985 for solving parallely the large sparse systems of linear

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

[2] Tel.: +86 10 626 30992; fax: +86 10 625 42285; email: bzz@lsec.cc.ac.cn

PII: S 0 1 6 7 - 8 1 9 1 (9 9) 0 0 0 1 5 - 0

equations on the multiprocessor systems. Then, many authors have further improved and extended the method models, and deeply investigated the corresponding convergence theories from various angles. For example, Neumann and Plemmons [15] developed some more refined convergence results for one of the cases considered in Ref. [16], Elsner [11] established the comparison theorems about the asymptotic convergence rate of this case, Frommer and Mayer [13] discussed the successive overrelaxation (SOR) method in the sense of multisplitting, White [24,25] studied the convergence properties of the above matrix multisplitting methods for the symmetric positive definite matrix class, as well as matrix multisplitting methods as preconditioners, respectively, and Bai [4] extended these matrix multisplitting methods to the system of nonlinear equations. For details one can refer to [2,3,6–16], [22–25] and references therein.

Since the finite element or the finite difference discretizations of many partial differential equations usually result in the large sparse systems of linear equations of regularly blocked structures, recently, [1,5] further generalized the matrix multisplitting concept of O'Leary and White [16] to a blocked form and proposed a class of parallel matrix multisplitting blockwise relaxation methods. This class of methods, besides enjoying all the advantages of the existing pointwise parallel matrix multisplitting methods discussed in Refs. [16,22], possesses better convergence properties and robuster numerical behaviours. Therefore, it is very smart to implement on the SIMD multiprocessor systems, or in other words, in the synchronous parallel computational environments. Under reasonable constraints on both the relaxation parameters and the multiple splittings, the convergence and the convergence rate of this class of methods were investigated in detail in Refs. [1,5] when the coefficient matrices of the blocked systems of linear equations are block H-matrices and block M-matrices, respectively.

To suit the requirements of the MIMD multiprocessor systems, this paper will emphatically discuss efficient asynchronous variants of the parallel matrix multisplitting blockwise relaxation methods in Refs. [1,5]. In light of the principle of using sufficiently the delayed information and with the utilization of the technique of successively accelerated overrelaxation (AOR), we set up a class of asynchronous parallel blockwise relaxation methods for solving the blocked system of linear equations in the sense of block matrix multisplitting (BMM). Because these methods simultaneously take into account the special structures of the linear systems of equations and the concrete characteristics of the high-speed multiprocessor systems, they have a lot of good properties such as simple and convenient operations, flexible and freed communications and so on. Moreover, besides using new information as soon as available, these methods increase the efficient numerical computations and decrease the unnecessary data communications. Therefore, they can greatly execute the parallel computing efficiency of the MIMD-systems. Since at least two arbitrary parameters are involved in this class of novel methods, the suitable adjustments of these parameters can not only considerably improve the convergence properties of the new methods, but also generate a series of applicable and efficient asynchronous multisplitting blockwise relaxation methods, e.g., the asynchronous multisplitting blockwise Jacobi method, the asynchronous multisplitting blockwise Gauss-Seidel method and the asynchronous multisplitting blockwise SOR method, etc.

Under reasonable restrictions on the relaxation parameters and the multiple splittings, we establish the convergence theory of this class of asynchronous multisplitting blockwise relaxation methods when the coefficient matrices of the blocked systems of linear equations are block H-matrices of different types. Lastly, a lot of numerical experiments show that these new methods are applicable and efficient, and they have better convergence behaviours than their pointwise alternatives.

2. Establishments of the methods

First of all, we recall the mathematical descriptions of the blocked system of linear equations and the BMM introduced in Ref. [1].

Let $N(\leqslant n)$ and $n_i(\leqslant n)(i=1,2,\ldots,N)$ be given positive integers satisfying $\sum_{i=1}^{N} n_i = n$, and denote

$$V_n(n_1,\ldots,n_N) = \left\{x \in \mathbb{R}^n \mid x = (x_1^{\mathrm{T}},\ldots,x_N^{\mathrm{T}})^{\mathrm{T}}, x_i \in \mathbb{R}^{n_i}\right\},$$
$$\mathbb{L}_n(n_1,\ldots,n_N) = \left\{A \in \mathbb{R}^{n\times n} \mid A = (A_{ij})_{N\times N}, A_{ij} \in \mathbb{R}^{n_i\times n_j}\right\}.$$

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 $V_n(n_1,\ldots,n_N)$. Then, the blocked system of linear equations to be solved can be expressed as the form

$$Ax = b, \quad A \in \mathbb{L}_n \text{ nonsingular}, \quad x, b \in V_n, \tag{2.1}$$

where $A \in \mathbb{L}_n$ and $b \in V_n$ are general known coefficient matrix and right-hand vector, respectively, and x is the unknown vector. The partition of the linear system (2.1) 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 Ref. [17].

Given a positive integer $\alpha(\alpha \leqslant N)$, we separate the number set $\{1,2,\ldots,N\}$ into α nonempty subsets $J_k(k=1,2,\ldots,\alpha)$ such that

$$J_k \subseteq \{1,2,\ldots,N\} \quad \text{and} \quad \bigcup_{k=1}^{\alpha} J_k = \{1,2,\ldots,N\}.$$

Note that there may be overlappings among the subsets $J_1, J_2, \ldots, J_\alpha$. Corresponding to this separation, we introduce matrices

$$D = \mathrm{Diag}(A_{11},\ldots,A_{NN}) \in \mathbb{L}_n,$$
$$L_k = (\mathscr{L}_{ij}^{(k)}) \in \mathbb{L}_n, \quad \mathscr{L}_{ij}^{(k)} = \begin{cases} L_{ij}^{(k)} & \text{for } i,j \in J_k \text{ and } i > j, \\ 0 & \text{otherwise,} \end{cases}$$
$$U_k = (\mathscr{U}_{ij}^{(k)}) \in \mathbb{L}_n, \quad \mathscr{U}_{ij}^{(k)} = \begin{cases} U_{ij}^{(k)} & \text{for } i \neq j, \\ 0 & \text{otherwise,} \end{cases}$$

$$E_k = \mathrm{Diag}(E_{11}^{(k)}, \ldots, E_{NN}^{(k)}) \in \mathbb{L}_n, \quad E_{ii}^{(k)} = \begin{cases} E_{ii}^{(k)} & \text{for } i \in J_k, \\ 0 & \text{otherwise,} \end{cases}$$
$$i, j = 1, 2, \ldots, N; \quad k = 1, 2, \ldots, \alpha.$$

Here, $L_{ij}^{(k)}$ and $U_{ij}^{(k)}$ are $n_i \times n_j$ sub-matrices, and $E_{ii}^{(k)}$ are $n_i \times n_i$ sub-matrices. Obviously, D is a blocked diagonal matrix, $L_k (k = 1, 2, \ldots, \alpha)$ are blocked strictly lower triangular matrices, $U_k (k = 1, 2, \ldots, \alpha)$ are blocked zero-diagonal matrices, and $E_k (k = 1, 2, \ldots, \alpha)$ are blocked diagonal matrices. If they satisfy

1. D is nonsingular;
2. $A = D - L_k - U_k, \quad k = 1, 2, \ldots, \alpha;$
3. $\sum_{k=1}^{\alpha} E_k = I,$

then we call the collection of triples $(D - L_k, U_k, E_k)(k = 1, 2, \ldots, \alpha)$ a BMM of the blocked matrix $A \in \mathbb{L}_n$. Here, I denotes the identity matrix of order $n \times n$.

Note that differently from the (pointwise) matrix multisplitting (PMM) in the literature, e.g., [2,3,6–9,11–13,15,16,22–25], the BMM allows that the weighting matrices $E_k (k = 1, 2, \ldots, \alpha)$ are blocked diagonal matrices, and may have nonpositive elements. This property can be utilized to design new methods so that their iteration matrices have smaller condition numbers. However, when $N = n$ and $n_i = 1 (i = 1, 2, \ldots, N)$, the BMM naturally reduces to the PMM.

Assume that the referred multiprocessor system is constructed by α processors. To describe the new asynchronous multisplitting blockwise relaxation methods, we need to introduce the following notations:

(1) for $k \in \{1, 2, \ldots, \alpha\}$ and $p \in N_0 := \{0, 1, 2, \ldots\}$, $J_k(p)$ is used to denote a subset (may be empty set $\emptyset$) of the set J_k, and $J^{(k)} = \{J_k(p)\}_{p \in N_0}$;

(2) for $i \in \{1, 2, \ldots, N\}$ and $p \in N_0$, $\mathbb{N}_i(p) := \{k \mid i \in J_k(p), \ k = 1, 2, \ldots, \alpha\}$;

(3) for $k \in \{1, 2, \ldots, \alpha\}$, $\{s_i^{(k)}(p)\}_{p \in N_0} (i = 1, 2, \ldots, N)$ are used to denote N infinite nonnegative integer sequences, and $S^{(k)} = \{s_1^{(k)}(p), s_2^{(k)}(p), \ldots, s_N^{(k)}(p)\}_{p \in N_0}$;

$J^{(k)}$ and $S^{(k)} (k = 1, 2, \ldots, \alpha)$ have the following properties:

(a) for $k \in \{1, 2, \ldots, \alpha\}$ and $i \in \{1, 2, \ldots, N\}$, the set $\{p \in N_0 \mid i \in J_k(p)\}$ is infinite;

(b) for $p \in N_0$, $\cup_{k=1}^{\alpha} J_k(p) \neq \emptyset$;

(c) for $k \in \{1, 2, \ldots, \alpha\}$, $i \in \{1, 2, \ldots, N\}$ and $p \in N_0$, $s_i^{(k)}(p) \leqslant p$;

(d) for $k \in \{1, 2, \ldots, \alpha\}$ and $i \in \{1, 2, \ldots, N\}$, $\lim_{p \to \infty} s_i^{(k)}(p) = \infty$.

Now, based on the above prerequisites, we can establish the following asynchronous multisplitting blockwise AOR method for solving the blocked system of linear equations (2.1).

Method I. Let $x^0 \in V_n$ be an initial vector and assume that we have got the approximate solutions $x^0, x^1, \ldots, x^p$ of the blocked system of linear equations (2.1). Then the $(p+1)$th approximate solution

$$x^{p+1} = \left((x_1^{p+1})^{\mathrm{T}}, (x_2^{p+1})^{\mathrm{T}}, \ldots, (x_N^{p+1})^{\mathrm{T}} \right)^{\mathrm{T}} \in V_n$$

can be got in an element by element manner from

$$x_i^{p+1} = \sum_{k=1}^{\alpha} E_{ii}^{(k)} x_i^{p,k}, \quad i = 1, 2, \ldots, N, \tag{2.2}$$

where for each $k \in \{1, 2, \ldots, \alpha\}$, $x_i^{p,k}$ is successively determined either by

$$x_i^{p,k} = A_{ii}^{-1}\left\{ \gamma \sum_{\substack{j<i \\ j\in J_k(p)}} L_{ij}^{(k)} x_j^{p,k} + (\omega-\gamma) \sum_{\substack{j<i \\ j\in J_k}} L_{ij}^{(k)} x_j^{s_j^{(k)}(p)} \right.$$
$$\left. + \gamma \sum_{\substack{j<i \\ j\in J_k \setminus J_k(p)}} L_{ij}^{(k)} x_j^{s_j^{(k)}(p)} + \omega \sum_{j\neq i} U_{ij}^{(k)} x_j^{s_j^{(k)}(p)} + \omega b_i \right\} + (1-\omega) x_i^{s_i^{(k)}(p)} \tag{2.3}$$

if $i \in J_k(p)$, or by

$$x_i^{p,k} = x_i^p \tag{2.4}$$

if $i \notin J_k(p)$, while $\gamma \in [0, \infty)$ and $\omega \in (0, \infty)$ are the relaxation and acceleration factors, respectively.

The description of Method I evidently shows that when we implement this method on an MIMD multiprocessor system, each processor can update any subset of the components of the global approximation, in which the subscripts of the components belong to some corresponding subset J_k, $k \in \{1, 2, \ldots, \alpha\}$, or retrive any subset of the components of the global approximation residing in the host processor, at any time. This therefore avoids loss of time and efficiency in processor utilization, and makes the multiprocessor system attain its maximum efficiency. Moreover, compared with the asynchronous parallel multisplitting pointwise relaxation methods discussed in Ref. [6] as well as in Refs. [8,9,12,13,15,16,22–25], Method I is much rich in block matrix and vector operations. This then not only considerably improves the convergence conditions and properties of this method, but also increases the efficient numerical computations and decreases the unnecessary information exchanges. Therefore, it can possibly result in better computing results.

Noticing that two arbitrary parameters γ and ω are included in Method I, we can reasonably adjust these parameters so that this method possesses better convergence behaviour. Moreover, suitable choices of these two parameters can result in a series of applicable and efficient asynchronous multisplitting blockwise relaxation methods, e.g., the asynchronous multisplitting blockwise Jacobi method ($\gamma = 0, \omega = 1$), the asynchronous multisplitting blockwise Gauss–Seidel method ($\gamma = 1$, $\omega = 1$), and the asynchronous multisplitting blockwise SOR method ($\gamma = \omega > 0$).

On the other hand, Method I also covers a lot of important cases of existing and novel matrix multisplitting relaxation methods. For example, when $N = n$, $n_i = 1 (i = 1, 2, \ldots, N)$, it naturally becomes the asynchronous matrix multisplitting AOR method discussed in Ref. [6] for solving the (pointed) system of linear equations, which, in particular, covers the synchronous parallel multisplitting relaxation methods presented in Refs. [8,13,22], cf., also [2,3,11,12,15,16,24,25]; when

$$J_k(p) = J_k = \{1, 2, \ldots, N\}, \quad s_i^{(k)}(p) = p,$$
$$\forall i \in \{1, 2, \ldots, N\}, \quad \forall k \in \{1, 2, \ldots, \alpha\}, \quad \forall p \in N_0,$$

it turns to the synchronous matrix multisplitting blockwise AOR method studied in Ref. [1] for solving the blocked system of linear equations (2.1), cf., also [5]; and when

$$J_k(p) = J_k, \quad s_i^{(k)}(p) \equiv s_k(p) \in N_0 \quad (s_k(p) \text{ a positive integer sequence}),$$
$$\forall i \in \{1, 2, \ldots, N\}, \quad \forall k \in \{1, 2, \ldots, \alpha\}, \quad \forall p \in N_0,$$

it presents a novel asynchronous multisplitting blockwise AOR method for solving the blocked system of linear equations (2.1), which is just a blocked variant of the asynchronous matrix multisplitting AOR method investigated in Ref. [23], which, in particular, includes the parallel chaotic iteration method models proposed in Ref. [9,10].

The above illustrations show why Method I has rather generality and attractive computational properties.

If we supplement $x_j^{p,k} = x_j^{s_j^{(k)}(p)}$ for all $j \in J_k \setminus J_k(p)$, $k \in \{1, 2, \ldots, \alpha\}$, $p \in N_0$, and define the projection operator $\mathscr{P}_i : V_n \to \mathbb{R}^{n_i} (i = 1, 2, \ldots, N)$ by

$$\mathscr{P}_i(x) = x_i, \quad \forall x = (x_1^{\mathrm{T}}, x_2^{\mathrm{T}}, \ldots, x_N^{\mathrm{T}})^{\mathrm{T}} \in V_n, \tag{2.5}$$

then Method I can be equivalently expressed as the concise form

$$x_i^{p+1} = \sum_{k \in \mathbb{N}_i(p)} E_{ii}^{(k)} \mathscr{P}_i(y^{p,k}) + \sum_{k \notin \mathbb{N}_i(p)} E_{ii}^{(k)} x_i^p, \quad i = 1, 2, \ldots, N, \tag{2.6}$$

where

$$y^{p,k} = (D - \gamma L_k)^{-1}((1 - \omega)D + (\omega - \gamma)L_k + \omega U_k)x^{s^{(k)}(p)} + (D - \gamma L_k)^{-1}(\omega b), \tag{2.7}$$

and

$$x^{s^{(k)}(p)} = \left(\left(x_1^{s_1^{(k)}(p)} \right)^{\mathrm{T}}, \left(x_2^{s_2^{(k)}(p)} \right)^{\mathrm{T}}, \ldots, \left(x_N^{s_N^{(k)}(p)} \right)^{\mathrm{T}} \right)^{\mathrm{T}}. \tag{2.8}$$

Furthermore, if Method I is relaxed by another parameter $\beta \in (0, \infty)$, we can extend it to be the following extrapolated one:

Method II. Let $x^0 \in V_n$ be an initial vector and assume that we have got the approximate solutions $x^0, x^1, \ldots, x^p$ of the blocked system of linear equations (2.1). Then the $(p + 1)$th approximate solution

$$x^{p+1} = \left((x_1^{p+1})^{\mathrm{T}}, (x_2^{p+1})^{\mathrm{T}}, \ldots, (x_N^{p+1})^{\mathrm{T}} \right)^{\mathrm{T}} \in V_n$$

can be got in an element by element manner from

$$x_i^{p+1} = \sum_{k=1}^{\alpha} E_{ii}^{(k)} \bar{x}_i^{p,k}, \quad i = 1, 2, \ldots, N \tag{2.9}$$

with

$$\bar{x}_i^{p,k} = \begin{cases} \beta x_i^{p,k} + (1-\beta) x_i^{s_i^{(k)}(p)} & \text{for } i \in J_k(p), \\ x_i^p & \text{for } i \notin J_k(p), \end{cases} \quad k = 1,2,\ldots,\alpha, \tag{2.10}$$

where for each $k \in \{1,2,\ldots,\alpha\}$, $x_i^{p,k}$ is successively determined either by Eq. (2.3) for $i \in J_k(p)$ or by Eq. (2.4) for $i \notin J_k(p)$, while $\gamma \in [0,\infty)$ is the relaxation factor and $\omega, \beta \in (0,\infty)$ are the acceleration factors.

Evidently, besides having all the advantages of Method I, Method II is more flexible and extensive than Method I since it involves three arbitrary parameters, which can be chosen arbitrarily in practical implementations. Note that when $\beta = 1$, Method II naturally reduces to Method I.

Analogously, suitable combinations of Eqs. (2.9) and (2.10) as well as Eqs. (2.3) and (2.4) give the following unified form of Method II:

$$x_i^{p+1} = \sum_{k \in \mathbb{N}_i(p)} E_{ii}^{(k)} \mathscr{P}_i(\bar{y}^{p,k}) + \sum_{k \notin \mathbb{N}_i(p)} E_{ii}^{(k)} x_i^p, \quad i = 1,2,\ldots,N \tag{2.11}$$

with

$$\bar{y}^{p,k} = \beta y^{p,k} + (1-\beta) x^{s^{(k)}(p)}, \quad k \in \mathbb{N}_i(p), \; i = 1,2,\ldots,N, \tag{2.12}$$

where $y^{p,k}$ and $x^{s^{(k)}(p)}$ are defined by Eqs. (2.7) and (2.8), respectively, while the projection operators $\mathscr{P}_i : V_n \to \mathbb{R}^{n_i} (i = 1,2,\ldots,N)$ are defined by Eq. (2.5).

3. Preliminaries

The partial orderings $<$, $\leqslant$ and the absolute values $|\bullet|$ in $\mathbb{R}^m$ and $\mathbb{R}^{m\times m} (m = n, N, n_i (i = 1,2,\ldots,N))$ are defined according to the elements. For a matrix $G = (g_{ij}) \in \mathbb{R}^{m\times m}$, we call it an M-matrix if $g_{ij} \leqslant 0 (i \neq j, i,j = 1(1)m)$, G^{-1} exists and $G^{-1} \geqslant 0$. Here, for simplicity, we denote $i = 1,2,\ldots,m$ by $i = 1(1)m$, and this notation will be used in the sequel. Let $D_G = \mathrm{Diag}(g_{11},\ldots,g_{mm})$ and $B_G = D_G - G$. Then G is an M-matrix if and only if D_G is positive diagonal and $\rho(J_G) < 1$, where $\rho(\bullet)$ denotes the spectral radius of a matrix and $J_G = D_G^{-1} B_G$. We refer to [20,21] for further details.

Define two matrix sets

$$\mathbb{L}_{n,I}(n_1,\ldots,n_N) = \{M = (M_{ij}) \in \mathbb{L}_n \mid M_{ii} \in \mathbb{R}^{n_i \times n_i} \text{ nonsingular}, \; i = 1(1)N\},$$

$$\mathbb{L}_{n,I}^d(n_1,\ldots,n_N) = \{M = \mathrm{Diag}(M_{11},\ldots,M_{NN}) \mid M_{ii} \in \mathbb{R}^{n_i \times n_i} \text{ nonsingular}, \; i = 1(1)N\}.$$

Again, we do not write the parameters $(n_1,\ldots,n_N)$ when they are clear from the context. In the following, we will review the concepts of the block H-matrices and several related conclusions. To simplify the statements we denote by $\|\bullet\|$ any consistent matrix norm satisfying $\|I\| = 1$.

Definition 3.1 (*see Ref.* [1,18,19]). Let $M \in \mathbb{L}_{n,I}$. Then its type-I and type-II block comparison matrices $\langle M\rangle = (\langle M\rangle_{ij}) \in \mathbb{R}^{N\times N}$ and $\langle\langle M\rangle\rangle = (\langle\langle M\rangle\rangle_{ij}) \in \mathbb{R}^{N\times N}$ are respectively defined by

$$\langle M\rangle_{ij} = \begin{cases} \|M_{ij}^{-1}\|^{-1} & \text{for } i = j, \\ -\|M_{ij}\| & \text{for } i \neq j, \end{cases} \quad i,j = 1,2,\ldots,N$$

and

$$\langle\langle M\rangle\rangle_{ij} = \begin{cases} 1 & \text{for } i = j, \\ -\|M_{ii}^{-1}M_{ij}\| & \text{for } i \neq j, \end{cases} \quad i,j = 1,2,\ldots,N.$$

For two matrices $L \in \mathbb{L}_n$ and $M \in \mathbb{L}_{n,I}$, if we write

$$D(L) = \mathrm{Diag}(L_{11},\ldots,L_{NN}), \quad B(L) = D(L) - L,$$

$$J(M) = D(M)^{-1}B(M), \quad \mu_1(M) = \rho(J_{\langle M\rangle}), \quad \mu_2(M) = \rho(I - \langle\langle M\rangle\rangle),$$

then from [1] we know that

$$\langle I - J(M)\rangle = \langle\langle I - J(M)\rangle\rangle = \langle\langle M\rangle\rangle, \quad \mu_2(M) \leqslant \mu_1(M).$$

Definition 3.2 (*see Ref.* [1]). Let $M \in \mathbb{L}_{n,I}$. If there exist $P, Q \in \mathbb{L}_{n,I}^d$ such that $\langle PMQ\rangle$ is an M-matrix, then M is called a type-I block H-matrix, simply denoted as $H_{\mathrm{B}}^{(\mathrm{I})}(P,Q)$-matrix, according to the nonsingular blocked diagonal matrices P and Q. If there exist $P, Q \in \mathbb{L}_{n,I}^d$ such that $\langle\langle PMQ\rangle\rangle$ is an M-matrix, then M is called a type-II block H-matrix, simply denoted as $H_{\mathrm{B}}^{(\mathrm{II})}(P,Q)$-matrix, according to the nonsingular blocked diagonal matrices P and Q. Evidently, an $H_B^{(\mathrm{I})}(P,Q)$-matrix is certainly an $H_B^{(\mathrm{II})}(P,Q)$-matrix, but conversely, it is not true.

Definition 3.3 (*see Ref.* [1]). Let $M \in \mathbb{L}_n$. We call $[M] = (\|M_{ij}\|) \in \mathbb{R}^{N\times N}$ the block absolute value of the blocked matrix M. The block absolute value $[x] \in \mathbb{R}^N$ of a blocked vector $x \in V_n$ is defined in an analogous way.

These kinds of block absolute values have the following important properties.

Lemma 3.1 (see Ref. [1]). *Let $L, M \in \mathbb{L}_n$, $x, y \in V_n$ and $\gamma \in \mathbb{R}^1$. Then*
1. $|[L] - [M]| \leqslant [L + M] \leqslant [L] + [M] \quad (|[x] - [y]| \leqslant [x + y] \leqslant [x] + [y])$;
2. $[LM] \leqslant [L][M] \quad ([Mx] \leqslant [M][x])$;
3. $[\gamma M] \leqslant |\gamma|[M] \quad ([\gamma x] \leqslant |\gamma|[x])$;
4. $\rho(M) \leqslant \rho(|M|) \leqslant \rho([M])$ (*here*, $\|\bullet\|$ *is either* $\|\bullet\|_\infty$ *or* $\|\bullet\|_1$).

Lemma 3.2 (see Ref. [1]). *Let $M \in \mathbb{L}_{n,I}$ be an $H_{\mathrm{B}}^{(\mathrm{I})}(P,Q)$-matrix. Then*
1. *M is nonsingular*;
2. $[(PMQ)^{-1}] \leqslant \langle PMQ\rangle^{-1}$;
3. $\mu_1(PMQ) < 1$.

Lemma 3.3 (see Ref. [1]). *Let $M \in \mathbb{L}_{n,I}$ be an $H_{\mathrm{B}}^{(\mathrm{II})}(P,Q)$-matrix. Then*
1. *M is nonsingular*;
2. $[(PMQ)^{-1}] \leqslant \langle\langle PMQ\rangle\rangle^{-1}[D(PMQ)^{-1}]$;
3. $\mu_2(PMQ) < 1$.

Definition 3.4 (*see Ref.* [1]). Let $M = (M_{ij}) \in \mathbb{L}_{n,I}$ and write

$$\Omega_{\mathrm{B}}^{(\mathrm{I})}(M) = \{F = (F_{ij}) \in \mathbb{L}_{n,I} \mid \|F_{ii}^{-1}\| = \|M_{ii}^{-1}\|, \|F_{ij}\| = \|M_{ij}\|, i \neq j, i,j = 1(1)N\},$$

$$\Omega_{\mathrm{B}}^{(\mathrm{II})}(M) = \{F = (F_{ij}) \in \mathbb{L}_{n,I} \mid \|F_{ii}^{-1}F_{ij}\| = \|M_{ii}^{-1}M_{ij}\|, i,j = 1(1)N\}.$$

Then $\Omega_{\mathrm{B}}^{(\mathrm{I})}(M)$ and $\Omega_{\mathrm{B}}^{(\mathrm{II})}(M)$ are called the type-I and the type-II equimodular matrix sets associated with the matrix $M \in \mathbb{L}_{n,I}$.

We recall from [13,22] that for a matrix $\mathscr{A} \in \mathbb{R}^{m\times m}$, its PMM means a collection of triples $(\mathscr{D} - \mathscr{L}_k, \mathscr{U}_k, \mathscr{E}_k)(k = 1,2,\ldots,\alpha)$ for which $\mathscr{D} = \mathrm{Diag}(\mathscr{A})$, $\mathscr{L}_k \in \mathbb{R}^{m\times m}(k = 1,2,\ldots,\alpha)$ are strictly lower triangular matrices, $\mathscr{U}_k \in \mathbb{R}^{m\times m}(k = 1,2,\ldots,\alpha)$ are zero-diagonal matrices and $\mathscr{E}_k \in \mathbb{R}^{m\times m}(k = 1,2,\ldots,\alpha)$ are nonnegative diagonal matrices such that

(a) $\mathscr{D}$ is nonsingular;
(b) $\mathscr{A} = \mathscr{D} - \mathscr{L}_k - \mathscr{U}_k$, $k = 1,2,\ldots,\alpha$;
(c) $\sum_{k=1}^{\alpha} \mathscr{E}_k = I$ (the $m \times m$ identity matrix).

Denote e_j the jth unit vector in $\mathbb{R}^m$ for $j = 1,2,\ldots,m$, and let $\mathbb{N}_j(p)$, $s_j^{(k)}(p)(j = 1(1)m, p \in N_0)$ be defined as in section two for $k = 1,2,\ldots,\alpha$. To establish the convergence theories of Methods I and II, we state the following crucial conclusion, which is a simplified variant of Theorem 1 in Ref. [6].

Lemma 3.4 (see Ref. [6]). *Let $\mathscr{A} \in \mathbb{R}^{N\times N}$ be an M-matrix and $(\mathscr{D} - \mathscr{L}_k, \mathscr{U}_k, \mathscr{E}_k)$ $(k = 1,2,\ldots,\alpha)$ be its multisplitting with $\mathscr{B} = \mathscr{D} - \mathscr{A}$, and $\mathscr{L}_k \geqslant 0$ and $\mathscr{U}_k \geqslant 0$, $k = 1,2,\ldots,\alpha$. For any starting vector $\epsilon^0 \in \mathbb{R}^N$, define the sequence $\{\epsilon^p\}_{p\in N_0}$ recursively by*

$$\epsilon^{p+1} = (\epsilon_1^{p+1}, \epsilon_2^{p+1}, \ldots, \epsilon_N^{p+1})^{\mathrm{T}}, \quad \epsilon^{s^{(k)}(p)} = \left(\epsilon_1^{s_1^{(k)}(p)}, \epsilon_2^{s_2^{(k)}(p)}, \ldots, \epsilon_N^{s_N^{(k)}(p)}\right)^{\mathrm{T}},$$

$$\epsilon_j^{p+1} = \sum_{k\in\mathbb{N}_j(p)} e_j^{\mathrm{T}}\mathscr{E}_k(\mathscr{D} - r\mathscr{L}_k)^{-1}(|1-w|\mathscr{D} + (w-r)\mathscr{L}_k + w\mathscr{U}_k)\epsilon^{s^{(k)}(p)} + \sum_{k\notin\mathbb{N}_j(p)} e_j^{\mathrm{T}}\mathscr{E}_k\epsilon^p,$$

$$j = 1,2,\ldots,N; \quad k = 1,2,\ldots,\alpha; \quad p = 0,1,2,\ldots,$$

where $r, w \in [0,\infty)$ are arbitrary parameters. Then the sequence $\{\epsilon^p\}_{p\in N_0}$ converges to zero if r and w satisfy $0 \leqslant r \leqslant w$ and $0 < w < 2/(1 + \rho(\mathscr{D}^{-1}\mathscr{B}))$.

In the subsequent discussions, the norm $\|\bullet\|$ is taken to be either $\|\bullet\|_\infty$ or $\|\bullet\|_1$.

4. Convergence theories

Assume $x^* \in V_n$ is the unique solution of the blocked system of linear equations (2.1) and $\{x^p\}_{p\in N_0}$ is an approximate sequence of x^*. Define the error sequence $\{\varepsilon^p\}_{p\in N_0}$ by $\varepsilon^p = x^p - x^*$ $(\forall p \in N_0)$ and denote

$$\varepsilon^{s^{(k)}(p)} = \left(\left(\varepsilon_1^{s_1^{(k)}(p)}\right)^{\mathrm{T}}, \left(\varepsilon_2^{s_2^{(k)}(p)}\right)^{\mathrm{T}}, \ldots, \left(\varepsilon_N^{s_N^{(k)}(p)}\right)^{\mathrm{T}}\right)^{\mathrm{T}},$$
$$k = 1,2,\ldots,\alpha; \quad p = 0,1,2,\ldots.$$

Then, from Eqs. (2.5)–(2.8) and Eqs. (2.11) and (2.12) the error sequences $\{\varepsilon^p\}_{p\in N_0}$ corresponding to the iterative sequences $\{x^p\}_{p\in N_0}$ generated by Methods I and II can be respectively expressed as

$$\varepsilon_i^{p+1} = \sum_{k\in\mathbb{N}_i(p)} E_{ii}^{(k)}\mathscr{P}_i\left(L_k(\gamma,\omega)\varepsilon^{s^{(k)}(p)}\right) + \sum_{k\notin\mathbb{N}_i(p)} E_{ii}^{(k)}\varepsilon_i^p, \quad i = 1,2,\ldots,N, \tag{4.1}$$

$$\varepsilon_i^{p+1} = \sum_{k\in\mathbb{N}_i(p)} E_{ii}^{(k)}\mathscr{P}_i\left(L_k(\gamma,\omega,\beta)\varepsilon^{s^{(k)}(p)}\right) + \sum_{k\notin\mathbb{N}_i(p)} E_{ii}^{(k)}\varepsilon_i^p, \quad i = 1,2,\ldots,N, \tag{4.2}$$

where

$$\begin{aligned} &L_k(\gamma,\omega) = (D-\gamma L_k)^{-1}((1-\omega)D + (\omega-\gamma)L_k + \omega U_k),\\ &L_k(\gamma,\omega,\beta) = \beta L_k(\gamma,\omega) + (1-\beta)I,\\ &\quad k = 1,2,\ldots,\alpha. \end{aligned} \tag{4.3}$$

Evidently, to prove the convergence of Methods I and II, we only need to prove that the sequences $\{\varepsilon^p\}_{p\in N_0}$ defined by Eqs. (4.1) and (4.2), respectively, converge to zero.

Theorem 4.1. *Let $A \in \mathbb{L}_{n,I}$ be an $H_{\mathrm{B}}^{(\mathrm{I})}(P,Q)$-matrix, $H \in \Omega_{\mathrm{B}}^{(\mathrm{I})}(PAQ)$ and $(\overline{D}-\overline{L}_k, \overline{U}_k, E_k)(k = 1,2,\ldots,\alpha)$ be a BMM of the blocked matrix H with $\sum_{k=1}^{\alpha}[E_k] \leqslant I$ and*

$$\langle H\rangle = \langle\overline{D}\rangle - [\overline{L}_k] - [\overline{U}_k] \equiv D_{\langle H\rangle} - B_{\langle H\rangle}, \quad k = 1,2,\ldots,\alpha. \tag{4.4}$$

Then, for any starting vector $x^0 \in V_n$,

(i) *the sequence $\{x^p\}_{p\in N_0}$ generated by Method I for solving the blocked system of linear equations*

$$Hx = b, \quad H \in \Omega_{\mathrm{B}}^{(\mathrm{I})}(PAQ), \quad x, b \in V_n \tag{4.5}$$

converges to its unique solution $x^* \in V_n$ provided the relaxation parameters γ and *ω satisfy*

$$0 \leqslant \gamma \leqslant \omega, \quad 0 < \omega < \frac{2}{1+\mu_1(PAQ)}; \tag{4.6}$$

(ii) *the sequence $\{x^p\}_{p\in N_0}$ generated by Method II for solving the blocked system of linear equations* (4.5) *converges to its unique solution $x^* \in V_n$ provided the relaxation parameters γ and ω satisfy* Eq. (4.6), *and β satisfies*

$$0 < \beta < \frac{2}{1 + \eta_1^{(\omega)}(PAQ)}, \quad \eta_1^{(\omega)}(PAQ) = |1 - \omega| + \omega\mu_1(PAQ). \tag{4.7}$$

Proof. Simply denote

$$\tilde{L}_k = \overline{D}^{-1}\overline{L}_k, \quad \tilde{U}_k = \overline{D}^{-1}\overline{U}_k, \quad k = 1, 2, \ldots, \alpha. \tag{4.8}$$

Then we have

$$L_k(\gamma, \omega) = (I - \gamma\tilde{L}_k)^{-1}\Big((1 - \omega)I + (\omega - \gamma)\tilde{L}_k + \omega\tilde{U}_k\Big), \quad k = 1, 2, \ldots, \alpha \tag{4.9}$$

from Eq. (4.3). Clearly, $(I - \gamma\tilde{L}_k)(k = 1, 2, \ldots, \alpha)$ are $H_{\mathrm{B}}^{(\mathrm{I})}(I, I)$-matrix. In light of Lemma 3.2(2), it holds that

$$[(I - \gamma\tilde{L}_k)^{-1}] \leqslant \langle I - \gamma\tilde{L}_k\rangle^{-1} = (I - \gamma[\tilde{L}_k])^{-1}, \quad k = 1, 2, \ldots, \alpha.$$

Therefore, applying Lemma 3.1(1)–(3) we get

$$\begin{aligned} [L_k(\gamma, \omega)] &\leqslant [(I - \gamma\tilde{L}_k)^{-1}]\Big(|1 - \omega|I + (\omega - \gamma)[\tilde{L}_k] + \omega[\tilde{U}_k]\Big) \\ &\leqslant (I - \gamma[\tilde{L}_k])^{-1}\Big(|1 - \omega|I + (\omega - \gamma)[\tilde{L}_k] + \omega[\tilde{U}_k]\Big) \end{aligned} \tag{4.10}$$

for $k = 1, 2, \ldots, \alpha$, from Eq. (4.9).

On the other hand, by Lemma 3.2(2) and Eq. (4.8) we see that

$$[\tilde{L}_k] \leqslant D_{\langle H\rangle}^{-1}[\overline{L}_k], \quad [\tilde{U}_k] \leqslant D_{\langle H\rangle}^{-1}[\overline{U}_k], \quad k = 1, 2, \ldots, \alpha, \tag{4.11}$$

and thereafter, it holds that

$$\begin{aligned} (I - \gamma[\tilde{L}_k])^{-1} &= \sum_{p\in N_0}(\gamma[\tilde{L}_k])^p \leqslant \sum_{p\in N_0}(\gamma D_{\langle H\rangle}^{-1}[\overline{L}_k])^p \\ &= (I - \gamma D_{\langle H\rangle}^{-1}[\overline{L}_k])^{-1} = (D_{\langle H\rangle} - \gamma[\overline{L}_k])^{-1}D_{\langle H\rangle} \end{aligned}$$

for $k = 1, 2, \ldots, \alpha$. Substituting these inequalities and Eq. (4.11) into Eq. (4.10), we obtain

$$\begin{aligned} [L_k(\gamma, \omega)] &\leqslant (D_{\langle H\rangle} - \gamma[\overline{L}_k])^{-1}D_{\langle H\rangle}\Big(|1 - \omega|I + (\omega - \gamma)D_{\langle H\rangle}^{-1}[\overline{L}_k] + \omega D_{\langle H\rangle}^{-1}[\overline{U}_k]\Big) \\ &= (D_{\langle H\rangle} - \gamma[\overline{L}_k])^{-1}\big(|1 - \omega|D_{\langle H\rangle} + (\omega - \gamma)[\overline{L}_k] + \omega[\overline{U}_k]\big) \end{aligned} \tag{4.12}$$

for $k = 1, 2, \ldots, \alpha$.

Write

$$\mathscr{L}_k(\gamma, \omega) = (D_{\langle H\rangle} - \gamma[\overline{L}_k])^{-1}\big(|1 - \omega|D_{\langle H\rangle} + (\omega - \gamma)[\overline{L}_k] + \omega[\overline{U}_k]\big), \quad k = 1, 2, \ldots, \alpha. \tag{4.13}$$

Then Eq. (4.12) is equivalent to

$$[L_k(\gamma, \omega)] \leqslant \mathscr{L}_k(\gamma, \omega), \quad k = 1, 2, \ldots, \alpha. \tag{4.14}$$

Now, to prove (i) we take the block absolute values on both sides of Eq. (4.1). This then gives

$$\begin{aligned}
[\varepsilon_i^{p+1}] &\leqslant \sum_{k\in\mathbb{N}_i(p)} [E_{ii}^{(k)}]\Big[\mathscr{P}_i\Big(L_k(\gamma,\omega)\varepsilon^{s^{(k)}(p)}\Big)\Big] + \sum_{k\notin\mathbb{N}_i(p)} [E_{ii}^{(k)}][\varepsilon_i^p] \\
&\leqslant \sum_{k\in\mathbb{N}_i(p)} [E_{ii}^{(k)}]e_i^{\mathrm{T}}[L_k(\gamma,\omega)]\Big[\varepsilon^{s^{(k)}(p)}\Big] + \sum_{k\notin\mathbb{N}_i(p)} [E_{ii}^{(k)}][\varepsilon_i^p] \\
&\leqslant \sum_{k\in\mathbb{N}_i(p)} [E_{ii}^{(k)}]e_i^{\mathrm{T}}\mathscr{L}_k(\gamma,\omega)\Big[\varepsilon^{s^{(k)}(p)}\Big] + \sum_{k\notin\mathbb{N}_i(p)} [E_{ii}^{(k)}][\varepsilon_i^p] \qquad (4.15)
\end{aligned}$$

for $i = 1, 2, \ldots, N$, by our making use of Eq. (4.14).

Denote

$$\begin{aligned}
&\mathscr{D} = D_{\langle H\rangle}, \quad \mathscr{B} = B_{\langle H\rangle}, \quad \mathscr{L}_k = [\overline{L}_k], \quad \mathscr{U}_k = [\overline{U}_k], \\
&\mathscr{E}_i^{(k)} = [E_{ii}^{(k)}], \quad \mathscr{E}^{(k)} = \mathrm{Diag}(\mathscr{E}_1^{(k)}, \mathscr{E}_2^{(k)}, \ldots, \mathscr{E}_N^{(k)}), \\
&\tilde{\varepsilon}_i^p = [\varepsilon_i^p], \quad \tilde{\varepsilon}^p = (\tilde{\varepsilon}_1^p, \tilde{\varepsilon}_2^p, \ldots, \tilde{\varepsilon}_N^p)^{\mathrm{T}}, \\
&\tilde{\varepsilon}^{s^{(k)}(p)} = (\tilde{\varepsilon}_1^{s_1^{(k)}(p)}, \tilde{\varepsilon}_2^{s_2^{(k)}(p)}, \ldots, \tilde{\varepsilon}_N^{s_N^{(k)}(p)})^{\mathrm{T}}, \\
&\quad i = 1, 2, \ldots, N; \quad k = 1, 2, \ldots, \alpha; \quad p = 0, 1, 2, \ldots
\end{aligned}$$

Then Eqs. (4.13) and (4.15) can be equivalently represented as

$$\mathscr{L}_k(\gamma,\omega) = (\mathscr{D} - \gamma\mathscr{L}_k)^{-1}(|1-\omega|\mathscr{D} + (\omega-\gamma)\mathscr{L}_k + \omega\mathscr{U}_k), \quad k = 1, 2, \ldots, \alpha,$$

$$\tilde{\varepsilon}_i^{p+1} \leqslant \sum_{k\in\mathbb{N}_i(p)} \mathscr{E}_i^{(k)} e_i^{\mathrm{T}}\mathscr{L}_k(\gamma,\omega)\tilde{\varepsilon}^{s^{(k)}(p)} + \sum_{k\notin\mathbb{N}_i(p)} \mathscr{E}_i^{(k)}\tilde{\varepsilon}_i^p, \quad i = 1, 2, \ldots, N.$$

Again, define $\mathscr{A}(\omega) = ((1 - |1-\omega|)/\omega)\mathscr{D} - \mathscr{B}$. Then $\mathscr{A}(\omega) \in \mathbb{R}^{N\times N}$ is evidently an M-matrix under the restriction (4.6). Moreover, $(\mathscr{D} - \mathscr{L}_k, \mathscr{U}_k, \mathscr{E}_k)(k = 1, 2, \ldots, \alpha)$ forms a PMM of the matrix $\mathscr{A}(\omega)$, and it holds that $\mathscr{L}_k \geqslant 0$ and $\mathscr{U}_k \geqslant 0$, $k = 1, 2, \ldots, \alpha$. In accordance with Lemma 3.4, the sequence $\{\epsilon^p\}_{p\in N_0}$ determined recursively by $\epsilon^0 = \tilde{\varepsilon}^0$ and

$$\epsilon_i^{p+1} = \sum_{k\in\mathbb{N}_i(p)} \mathscr{E}_i^{(k)} e_i^{\mathrm{T}}\mathscr{L}_k(\gamma,\omega)\epsilon^{s^{(k)}(p)} + \sum_{k\notin\mathbb{N}_i(p)} \mathscr{E}_i^{(k)}\epsilon_i^p,$$

$$i = 1, 2, \ldots, N; \quad p = 0, 1, 2, \ldots$$

converges to zero.

In addition, we can inductively verify that the sequence $\{\epsilon^p\}_{p\in N_0}$ majorizes the sequence $\{\tilde{\varepsilon}^p\}_{p\in N_0}$, i.e., $\tilde{\varepsilon}^p \leqslant \epsilon^p (\forall p \in N_0)$. In fact, the case of $p = 0$ is trivial. If we assume $\tilde{\varepsilon}^p \leqslant \epsilon^p$ for some $p \in N_0$, then it holds that $\tilde{\varepsilon}_j^{s_j^{(k)}(p)} \leqslant \epsilon_j^{s_j^{(k)}(p)}$ $(j = 1(1)N; k = 1, 2, \ldots, \alpha)$, and hence, we have

$$\begin{aligned}
\tilde{\varepsilon}_i^{p+1} &\leqslant \sum_{k\in\mathbb{N}_i(p)} \mathscr{E}_i^{(k)} e_i^{\mathrm{T}}\mathscr{L}_k(\gamma,\omega)\tilde{\varepsilon}^{s^{(k)}(p)} + \sum_{k\notin\mathbb{N}_i(p)} \mathscr{E}_i^{(k)}\tilde{\varepsilon}_i^p \\
&\leqslant \sum_{k\in\mathbb{N}_i(p)} \mathscr{E}_i^{(k)} e_i^{\mathrm{T}}\mathscr{L}_k(\gamma,\omega)\epsilon^{s^{(k)}(p)} + \sum_{k\notin\mathbb{N}_i(p)} \mathscr{E}_i^{(k)}\epsilon_i^p = \epsilon_i^{p+1}
\end{aligned}$$

for $i = 1, 2, \ldots, N$. This completes the induction process. Therefore, $\{\tilde{\varepsilon}^p\}_{p\in N_0}$, or in other words, $\{[\varepsilon^p]\}_{p\in N_0}$ converges to zero, which immediately gives $\varepsilon^p \to 0 (p \to \infty)$, or $x^p \to x^* (p \to \infty)$.

We now turn to (ii). Noticing from Eq. (4.3) that for $k = 1, 2, \ldots, \alpha$, it holds that

$$\begin{aligned} L_k(\gamma, \omega, \beta) &= (D - \gamma L_k)^{-1}((1 - \omega\beta)D + (\omega\beta - \gamma)L_k + \omega\beta U_k) \\ &= (D - \gamma L_k)^{-1}((1 - \sigma)D + (\sigma - \gamma)L_k + \sigma U_k) = L_k(\gamma, \sigma), \end{aligned}$$

with $\sigma = \omega\beta$, we can conclude from (i) and Eq. (4.2) that the error sequence $\{\varepsilon^p\}_{p\in N_0}$, defined by Eq. (4.2), corresponding to Method II converges to zero provided $|1 - \sigma| + \sigma\mu_1(PAQ) < 1$.

Considering $\eta_1^{(\omega)}(PAQ) < 1$ when γ and ω are within the region (4.6), we easily know that Eq. (4.7) is equivalent to $|1 - \beta| + \beta\eta_1^{(\omega)}(PAQ) < 1$. However, we have

$$\begin{aligned} |1 - \sigma| + \sigma\mu_1(PAQ) &= |1 - \sigma| - |\beta - \sigma| + \beta\eta_1^{(\omega)}(PAQ) \\ &\leqslant |1 - \beta| + \beta\eta_1^{(\omega)}(PAQ) < 1. \end{aligned}$$

Therefore, the conditions of Theorem 4.1 guarantee the validity of its conclusion (ii).

Theorem 4.2. *Let $A \in \mathbb{L}_{n,I}$ be an $H_{\mathrm{B}}^{(\mathrm{II})}(P, Q)$-matrix, $H \in \Omega_{\mathrm{B}}^{(\mathrm{II})}(PAQ)$ and $(\overline{D} - \overline{L}_k, \overline{U}_k, E_k)(k = 1, 2, \ldots, \alpha)$ be a BMM of the blocked matrix H with $\sum_{k=1}^{\alpha}[E_k] \leqslant I$ and*

$$\langle\langle H\rangle\rangle = I - [\overline{D}^{-1}\overline{L}_k] - [\overline{D}^{-1}\overline{U}_k] = I - B_{\langle\langle H\rangle\rangle}, \quad k = 1, 2, \ldots, \alpha.$$

Then, for any starting vector $x^0 \in V_n$,

(i) *the sequence $\{x^p\}_{p\in N_0}$ generated by Method I for solving the blocked system of linear equations*

$$Hx = b, \quad H \in \Omega_{\mathrm{B}}^{(\mathrm{II})}(PAQ), \quad x, b \in V_n \tag{4.16}$$

converges to its unique solution $x^* \in V_n$ provided the relaxation parameters γ and ω satisfy

$$0 \leqslant \gamma \leqslant \omega, \quad 0 < \omega < \frac{2}{1 + \mu_2(PAQ)}; \tag{4.17}$$

(ii) *the sequence $\{x^p\}_{p\in N_0}$ generated by Method II for solving the blocked system of linear equations* (4.16) *converges to its unique solution $x^* \in V_n$ provided the relaxation parameters γ and ω satisfy Eq.* (4.17), *and β satisfies*

$$0 < \beta < \frac{2}{1 + \eta_2^{(\omega)}(PAQ)}, \quad \eta_2^{(\omega)}(PAQ) = |1 - \omega| + \omega\mu_2(PAQ).$$

Proof. Under the conditions of the theorem, we easily know that $H \in \mathbb{L}_{n,I}$ is an $H_{\mathrm{B}}^{(\mathrm{II})}(I, I)$-matrix and $\langle\langle H\rangle\rangle = \langle\langle PAQ\rangle\rangle$. Hence, in light of Lemma 3.3(3) and from $J_{\langle\langle H\rangle\rangle} = \mathrm{B}_{\langle\langle H\rangle\rangle}$, it holds that

$$\begin{aligned} \mu_2(H) &= \rho(J_{\langle\langle H\rangle\rangle}) = \rho(B_{\langle\langle H\rangle\rangle}) = \rho(B_{\langle\langle PAQ\rangle\rangle}) \\ &= \rho(J_{\langle\langle PAQ\rangle\rangle}) = \mu_2(PAQ) < 1. \end{aligned}$$

The remainder of the proof is much analogous to that of Theorem 4.1, so it is omitted here.

We comment that Theorems 4.1 and 4.2 apply to matrix A which may not have a nonsingular $D(A)$, or such that it does not belong to $H_{\mathrm{B}}^{(\mathrm{I})}(I,I)$ or $H_{\mathrm{B}}^{(\mathrm{II})}(I,I)$ but such that there exist nonsingular block diagonal matrices P and Q, such that PAQ has these properties.

5. Numerical example

We consider the blocked system of linear equations

$$Ax = b, \quad A \in \mathbb{L}_n, \ \ b \in V_n, \tag{5.1}$$

with $n_1 = n_2 = \cdots = n_N \equiv N$ and $n = N^2$, where

$$\begin{aligned} A &= \begin{pmatrix} B & -I & & & \\ -I & B & -I & & \\ & \ddots & \ddots & \ddots & \\ & & -I & B & -I \\ & & & -I & B \end{pmatrix} \in \mathbb{R}^{n\times n}, \\ B &= \begin{pmatrix} 4 & -1 & & & \\ -1 & 4 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 4 & -1 \\ & & & -1 & 4 \end{pmatrix} \in \mathbb{R}^{N\times N}, \end{aligned} \tag{5.2}$$

and $b \in \mathbb{R}^n$ is chosen such that the solution of this blocked system of linear equations has all components equal to one. This blocked system of linear equations is yielded by discretizing some Dirichlet problem defined on the square domain $(0,1)\times(0,1)$ through using the finite difference method at N^2 equidistant nodes of a quadratic grid.

Take $\alpha = 2$ and two positive integers m_1 and m_2 satisfying $1 \leqslant m_2 \leqslant m_1 \leqslant N$. Then, corresponding to the number sets $J_1 = \{1, 2, \dots, m_1\}$ and $J_2 = \{m_2, m_2+1, \dots, N\}$, we determine a BMM $(D - L_k, U_k, E_k)(k = 1, 2)$ of the blocked matrix $A \in \mathbb{L}_n$ in accordance with the following way:

$$D = \mathrm{Diag}(B, \dots, B) \in \mathbb{L}_n,$$

$$L_1 = (\mathscr{L}_{ij}^{(1)}) \in \mathbb{L}_n, \quad \mathscr{L}_{ij}^{(1)} = \begin{cases} I & \text{for } j = i-1 \text{ and } 2 \leqslant i \leqslant m_1, \\ 0 & \text{otherwise}, \end{cases}$$

$$U_1 = (\mathscr{U}_{ij}^{(1)}) \in \mathbb{L}_n, \quad \mathscr{U}_{ij}^{(1)} = \begin{cases} I & \text{for } j = i-1 \text{ and } m_1+1 \leqslant i \leqslant N, \\ I & \text{for } j = i+1 \text{ and } 1 \leqslant i \leqslant N-1, \\ 0 & \text{otherwise}, \end{cases}$$

$$L_2 = (\mathscr{L}_{ij}^{(2)}) \in \mathbb{L}_n, \quad \mathscr{L}_{ij}^{(2)} = \begin{cases} I & \text{for } j = i-1 \text{ and } m_2 \leqslant i \leqslant N, \\ 0 & \text{otherwise}, \end{cases}$$

$$U_2 = (\mathscr{U}_{ij}^{(2)}) \in \mathbb{L}_n, \quad \mathscr{U}_{ij}^{(2)} = \begin{cases} I & \text{for } j = i-1 \text{ and } 2 \leqslant i \leqslant m_2 - 1, \\ I & \text{for } j = i+1 \text{ and } 1 \leqslant i \leqslant N-1, \\ 0 & \text{otherwise,} \end{cases}$$

$$E_k = \mathrm{Diag}(E_{11}^{(k)}, \ldots, E_{NN}^{(k)}) \in \mathbb{L}_n, \quad k = 1, 2,$$

$$E_{ii}^{(1)} = \begin{cases} I & \text{for } 1 \leqslant i < m_2, \\ \frac{1}{2} I & \text{for } m_2 \leqslant i \leqslant m_1, \\ 0 & \text{for } m_1 < i \leqslant N, \end{cases} \quad E_{ii}^{(2)} = \begin{cases} 0 & \text{for } 1 \leqslant i < m_2, \\ \frac{1}{2} I & \text{for } m_2 \leqslant i \leqslant m_1, \\ I & \text{for } m_1 < i \leqslant N. \end{cases}$$

In particular, we select the positive integer pair (m_1, m_2) to be

(a) $m_1 = \mathrm{Int}\left(\frac{2N}{3}\right), \quad m_2 = \mathrm{Int}\left(\frac{N}{3}\right)$; and

(b) $m_1 = \mathrm{Int}\left(\frac{4N}{5}\right), \quad m_2 = \mathrm{Int}\left(\frac{N}{5}\right)$,

respectively, then we can get two kinds of concrete cases of the weighting matrices E_1 and E_2, here, $\mathrm{Int}(\bullet)$ denotes the integer part of the corresponding real number.

We sequentially imitate our new methods on the SGI INDY workstation by the asynchronous multisplitting blockwise Jacobi method, the asynchronous multisplitting blockwise Gauss-Seidel method, the asynchronous multisplitting blockwise SOR method and the asynchronous multisplitting blockwise AOR method according to different N (hence n), or a fixed N (then n) but different relaxation parameters γ and ω. In all our computations, we start with an initial vector having all components equal to 0.5 and terminated once the current iterations x^p satisfy $\|Ax^p - b\|_1 \leqslant 10^{-4}$. We list the corresponding iteration numbers in Tables 1–6 to

Table 1
Asynchronous multisplitting blockwise Jacobi method

N	10	15	20	30	40	50	100	150	200	250
(a)	327	636	1109	2325	4107	6288	24 348	53 863	95 586	148 939

Table 2
Asynchronous multisplitting blockwise Gauss–Seidel method

N	10	15	20	30	40	50	100	150	200	250
(a)	176	336	576	1194	2094	3194	12 270	27 073	47 983	74 696
(b)	154	312	528	1127	1952	3003	11 652	25 953	45 912	71 520

Table 3
Asynchronous multisplitting blockwise SOR method ($N = 100$)

ω	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
(a)	10 059	8216	6657	5316	4155	3135	2235	1431	702*	∞
(b)	9546	7788	6300	5025	3918	2949	2091	1323	612*	∞

Table 4
Asynchronous multisplitting blockwise AOR method ($N = 100$)

γ	1.90	1.95								2.00
ω	2.00	1.60	1.65	1.70	1.75	1.80	1.85	1.90	2.0	1.90
(a)	5260	606	591	577	564	555	549*	552	4960	3504
(b)	2923	555	540	525	513	504	499*	509	4037	3555

Table 5
Asynchronous multisplitting blockwise SOR method ($N = 15$)

ω	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
(a)	276	228	185	147	114	84*	84*	121	211	1120
(b)	257	210	169	135	102	67*	77	109	178	446

Table 6
Asynchronous multisplitting blockwise AOR method ($N = 15$)

γ	1.75	1.7			1.65				1.6
ω	1.7	1.5	[1.55, 1.6]	1.65	1.5	1.55	1.6	1.65	1.65
(a)	89	75	76	77	75	72	70*	72	81
(b)	84	77		72	65	63*	63*	67	68

show the feasibility and efficiency of our new methods. For clearness, a star '*' is used to show the optimal number of iteration steps for a case in a numerical table. From these tables it is easy to see that suitable selections of the weighting matrices and reasonable matches of the relaxation parameters can usually produce better computing effects.

For comparisons, we also consider the blocked system of linear equations (5.1) and (5.2) in the point sense. Likewise, we take $\alpha = 2$ and two positive integers $\hat{m}_1$ and $\hat{m}_2$ such that $1 \leqslant \hat{m}_2 \leqslant \hat{m}_1 \leqslant n$. Then, corresponding to the number sets $\hat{J}_1 = \{1, 2, \ldots, \hat{m}_1\}$ and $\hat{J}_2 = \{\hat{m}_2, \hat{m}_2 + 1, \ldots, n\}$, we determine a PMM $(\hat{D} - \hat{L}_k, \hat{U}_k, \hat{E}_k)(k = 1, 2)$ of the pointed matrix $A \in \mathbb{R}^{n \times n}$ in accordance with the following way:

$$\hat{D} = \mathrm{diag}(4,4,\ldots,4) \in \mathbb{R}^{n\times n},$$

$$\hat{L}_1 = (\hat{\mathscr{L}}^{(1)}_{ij}) \in \mathbb{R}^{n\times n}, \quad \hat{\mathscr{L}}^{(1)}_{ij} = \begin{cases} 1 & \text{for } j = i-1, i = kN+s, \\ & \text{and } 0 \leqslant k \leqslant \mathrm{Int}(\hat{m}_1/N) - 1, 2 \leqslant s \leqslant N, \\ 1 & \text{for } j = i-1, i = \mathrm{Int}(\hat{m}_1/N)\cdot N + s, \\ & \text{and } 2 \leqslant s \leqslant \mathrm{mod}(\hat{m}_1, N), \\ 1 & \text{for } j = i-N, N+1 \leqslant i \leqslant \hat{m}_1, \\ 0 & \text{otherwise}, \end{cases}$$

$$\hat{U}_1 = (\hat{\mathscr{U}}^{(1)}_{ij}) \in \mathbb{R}^{n\times n}, \quad \hat{\mathscr{U}}^{(1)}_{ij} = \begin{cases} 1 & \text{for } j = i+1, i = kN+s, \\ & \text{and } 0 \leqslant k \leqslant N-1, 1 \leqslant s \leqslant N-1, \\ 1 & \text{for } j = i+N, 1 \leqslant i \leqslant n-N, \\ 1 & \text{for } j = i-1, i = kN+s, \\ & \mathrm{Int}(\hat{m}_1/N) \leqslant k \leqslant N-1, \\ & \text{and } \max\{2, \mathrm{mod}(\hat{m}_1, N)+1\} \leqslant s \leqslant N, \\ 1 & \text{for } j = i-N, \hat{m}_1 + 1 \leqslant i \leqslant n, \\ 0 & \text{otherwise}, \end{cases}$$

$$\hat{L}_2 = (\hat{\mathscr{L}}^{(2)}_{ij}) \in \mathbb{R}^{n\times n}, \quad \hat{\mathscr{L}}^{(2)}_{ij} = \begin{cases} 1 & \text{for } j = i-1, i = kN+s, \\ & \text{and } \mathrm{Int}(\hat{m}_2/N) + 1 \leqslant k \leqslant N-1, \\ & 2 \leqslant s \leqslant N, \\ 1 & \text{for } j = i-1, i = \mathrm{Int}(\hat{m}_2/N)\cdot N + s, \\ & \text{and } \mathrm{mod}(\hat{m}_2, N) + 1 \leqslant s \leqslant N, \\ 1 & \text{for } j = i-N, \hat{m}_2 + N \leqslant i \leqslant n, \\ 0 & \text{otherwise}, \end{cases}$$

$$\hat{U}_2 = (\hat{\mathscr{U}}^{(2)}_{ij}) \in \mathbb{R}^{n\times n}, \quad \hat{\mathscr{U}}^{(2)}_{ij} = \begin{cases} 1 & \text{for } j = i+1, i = kN+s, \\ & \text{and } 0 \leqslant k \leqslant N-1, 1 \leqslant s \leqslant N-1, \\ 1 & \text{for } j = i+N, 1 \leqslant i \leqslant n-N, \\ 1 & \text{for } j = i-1, i = kN+s, \\ & \text{and } 0 \leqslant k \leqslant \mathrm{Int}((\hat{m}_2 - 1)/N) - 1, \\ & 2 \leqslant s \leqslant N, \\ 1 & \text{for } j = i-1, i = kN+s, \\ & k = \mathrm{Int}((\hat{m}_2 - 1)/N), \\ & \text{and } 2 \leqslant s \leqslant \mathrm{mod}(\hat{m}_2 - 1, N), \\ 1 & \text{for } j = i-N, N+1 \leqslant i \leqslant \hat{m}_2 - 1, \\ 0 & \text{otherwise}, \end{cases}$$

$$\hat{E}_k = \mathrm{diag}(e_1^{(k)}, e_2^{(k)}, \ldots, e_n^{(k)}) \in \mathbb{R}^{n\times n}, \quad k = 1, 2,$$

$$e_i^{(1)} = \begin{cases} 1 & \text{for } 1 \leqslant i < \hat{m}_2, \\ 0.5 & \text{for } \hat{m}_2 \leqslant i \leqslant \hat{m}_1, \\ 0 & \text{for } \hat{m}_1 < i \leqslant n, \end{cases} \quad e_i^{(2)} = \begin{cases} 0 & \text{for } 1 \leqslant i < \hat{m}_2, \\ 0.5 & \text{for } \hat{m}_2 \leqslant i \leqslant \hat{m}_1, \\ 1 & \text{for } \hat{m}_1 < i \leqslant n. \end{cases}$$

In particular, if we select the positive integer pair $(\hat{m}_1, \hat{m}_2)$ to be

(a) $\hat{m}_1 = \mathrm{Int}\left(\frac{2N}{3}\right) \cdot N, \quad \hat{m}_2 = \mathrm{Int}\left(\frac{N}{3}\right) \cdot N;$ and

(b) $\hat{m}_1 = \mathrm{Int}\left(\frac{4N}{5}\right) \cdot N, \quad \hat{m}_2 = \mathrm{Int}\left(\frac{N}{5}\right) \cdot N,$

respectively, then we can get two kinds of concrete cases of the weighting matrices $\hat{E}_1$ and $\hat{E}_2$. Clearly, the PMMs defined here are just the same as the BMMs defined above, respectively. Their difference is just in the expressions; the former is described in the pointed manner, while the later is in the blocked manner.

Using the same starting vector and the same stopping criterion as above, we sequentially imitate the asynchronous multisplitting (pointwise) Jacobi method, the asynchronous multisplitting (pointwise) Gauss–Seidel method, the asynchronous multisplitting (pointwise) SOR method and the asynchronous multisplitting (pointwise) AOR method described in Ref. [6] also on the SGI INDY workstation according to different n (hence N), or a fixed n (then N) but different relaxation parameters γ and ω (here, we use γ instead of r there). The corresponding iteration numbers are listed in Tables 7–10. Comparing the computing results in Table 1 with Table 7, Table 2 with Table 8, Table 5 with Table 9 and Table 6 with Table 10, we easily see that the asynchronous multisplitting blockwise relaxation methods discussed in this paper substantially have better numerical behaviours than their point alternatives studied in Ref. [6]. That is to say, for the divergence cases of the asynchronous multisplitting pointwise relaxation methods, the corresponding asynchronous multisplitting blockwise relaxation methods give convergence, while for the convergence cases of the asynchronous multisplitting pointwise relaxation methods, the corresponding asynchronous multisplitting blockwise relaxation methods require considerably

Table 7
Asynchronous multisplitting Jacobi method

N	5	10	15	20	30	40	50	100	150	200
(a)	186	618	∞	∞	∞	∞	∞	∞	∞	∞
(b)	169	561	∞	∞	∞	∞	∞	∞	∞	∞

Table 8
Asynchronous multisplitting Gauss–Seidel method

N	5	10	15	20	30	40	50	100	150	200
(a)	101	322	634	∞	∞	∞	∞	∞	∞	∞
(b)	94	291	602	∞	∞	∞	∞	∞	∞	∞

Table 9
Asynchronous multisplitting SOR method ($N = 15$)

ω	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
(a)	519	424	345	275	214	160	109*	151	265	2092
(b)	495	404	328	262	203	149	94*	135	218	641

Table 10
Asynchronous multisplitting AOR method ($N = 15$)

γ	1.7					1.75		
ω	1.5	1.55	1.6	1.65	1.75	1.5	[1.55, 1.65]	1.7
(a)	121	118	113	110	115	107	106*	110
(b)	101	98	95	94*	106	106		97

less iterations for convergence. It is also clear from these tables that, compared with the asynchronous multisplitting pointwise relaxation methods in Ref. [6], the asynchronous multisplitting blockwise relaxation methods in this paper can solve very large systems of linear equations.

At last, we remark that the example of the symmetric system of linear equations and its corresponding BMM, given in this section, satisfies the conditions of both Theorems 4.1 and 4.2. Moreover, if, in this example, we only replace the symmetric tridiagonal matrix $B \in \mathbb{R}^{N\times N}$ by the nonsymmetric tridiagonal matrix

$$B = \begin{pmatrix} 4 & -1 & & & \\ -0.5 & 4 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -0.5 & 4 & -1 \\ & & & -0.5 & 4 \end{pmatrix} \in \mathbb{R}^{N\times N}$$

but keep the splitting and the weighting matrices as not changed, then we get another example about a nonsymmetric system of linear equations and its corresponding BMM, which also satisfies the conditions of both Theorems 4.1 and 4.2.

Acknowledgements

The author is very much indebted to the referees for a lot of valuable comments and constructive suggestions on the original manuscript of this paper. He also thanks heartily Professors G.R. Joubert and M. Shimasaki for their help.

References

[1] Z.Z. Bai, Parallel matrix multisplitting block relaxation iteration methods, Math. Numer. Sinica 17 (1995) 238–252.
[2] Z.Z. Bai, On the convergence of the generalized matrix multisplitting relaxed methods, Comm. Numer. Meth. Engrg. 11 (1995) 363–371.
[3] Z.Z. Bai, Comparisons of the convergence and divergence rates of the parallel matrix multisplitting iteration methods, Chinese J. Engrg. Math. 11 (1994) 99–102.
[4] Z.Z. Bai, Asynchronous multisplitting AOR method for a system of nonlinear algebraic equations, Int. J. Computer Math. 55 (1995) 223–233.
[5] Z.Z. Bai, D.R. Wang, On the convergence of the matrix multisplitting block relaxation methods, in: Proceedings of the National Forth Parallel Algorithm Conference, Aeronautical Industry Press, Beijing, 1993.
[6] Z.Z. Bai, D.R. Wang, D.J. Evans, Models of asynchronous parallel matrix multisplitting relaxed iterations, Parallel Comput. 21 (1995) 565–582.
[7] Z.Z. Bai, D.R. Wang, D.J. Evans, On the convergence of asynchronous nested matrix multisplitting methods for linear systems, to appear in J. Comp. Math.
[8] 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.
[9] R. Bru, L. Elsner, M. Neumann, Models of parallel chaotic iteration methods, Linear Algebra Appl. 103 (1988) 175–192.
[10] D. Chazan, W.L. Miranker, Chaotic relaxation, Linear Algebra Appl. 2 (1969) 199–222.
[11] L. Elsner, Comparisons of weak regular splittings and multisplitting methods, Numer. Math. 56 (1989) 283–289.
[12] D.J. Evans, D.R. Wang, Z.Z. Bai, Matrix multi-splitting multi-parameter relaxation methods, Int. J. Computer Math. 43 (1992) 173–188.
[13] A. Frommer, G. Mayer, Convergence of relaxed parallel multisplitting methods, Linear Algebra Appl. 119 (1989) 141–152.
[14] I. Marek, D.B. Szyld, Comparison theorems for weak splittings of bounded operators, Numer. Math. 58 (1990) 387–397.
[15] M. Neumann, R.J. Plemmons, Convergence of parallel multisplitting iterative methods for M-matrices, Linear Algebra Appl. 88 (1987) 559–573.
[16] D.P. O'Leary, R.E. White, Multi-splittings of matrices and parallel solutions of linear systems, SIAM J. Alg. Disc. Meth. 6 (1985) 630–640.
[17] J. O'Neil, D.B. Szyld, A block ordering method for sparse matrices, SIAM J. Sci. Stat. Comput. 11 (1990) 811–823.
[18] F. Robert, M. Charnay, F. Musy, Itérations chaotiques série-paralléle pour des équations non-linéaires de point fixe, Aplikace Matematiky 20 (1975) 1–38.
[19] F. Robert, Blocs-H-Matrices et Convergence des Méthodes Itératives Classiques par Blocs, Linear Algebra Appl. 2 (1969) 223–265.
[20] R.S. Varga, Matrix Iterative Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1962.
[21] R.S. Varga, On recurring theorems on diagonal dominance, Linear Algebra Appl. 13 (1976) 1–9.
[22] D.R. Wang, On the convergence of the parallel multisplitting AOR algorithm, Linear Algebra Appl. 154–156 (1991) 473–486.

[23] 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.
[24] R.E. White, Multisplitting of a symmetric positive definite matrix, SIAM J. Matrix Anal. Appl. 11 (1990) 69–82.
[25] R.E. White, Multisplitting with different weighting schemes, SIAM J. Matrix Anal. Appl. 10 (1989) 481–493.