Available at WWW.MATHEMATICSWEB.ORG POWERED BY SCIENCE DIRECT

Journal of Computational and Applied Mathematics 159 (2003) 1–11

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.com/locate/cam

On the convergence of parallel nonstationary multisplitting iteration methods☆

Zhong-Zhi Bai*

State Key Laboratory of Scientific/Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, P.O. Box 2719, Beijing 100080, People's Republic of China

Received 2 September 2002; received in revised form 16 November 2002

Abstract

The convergence properties of a variant of the parallel chaotic multisplitting iteration method, called the nonstationary multisplitting iteration method, for solving large sparse systems of linear equations are further discussed when the coefficient matrix is an H-matrix or a positive definite matrix, respectively. Moreover, when the coefficient matrix is a monotone matrix, the monotone convergence theory and the monotone comparison theorem about this method are established. This directly leads to several novel sufficient conditions for guaranteeing the convergence of this parallel nonstationary multisplitting iteration method.

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

Keywords: System of linear equations; Matrix multisplitting; Nonstationary iteration; Convergence

1. Introduction

The parallel matrix multisplitting iteration method solves the unique solution $x^* \in \mathbb{R}^n$ of the large sparse system of linear equations

$$Ax = b, \quad A \in L(\mathbb{R}^n) \qquad \text{nonsingular}, \quad x, b \in \mathbb{R}^n \tag{1}$$

on a multiprocessor system. Suppose the multiprocessor system consists of K processors, which are connected to a host processor which may be any of the K processors, and let (B_k, C_k, E_k)

☆ Subsidized by The Special Funds For Major State Basic Research Projects G1999032803.
* Corresponding author. Tel.: +86-10-62630992; fax: +86-10-62545820.
E-mail address: bzz@lsec.cc.ac.cn (Z.-Z. Bai).

0377-0427/03/$ - see front matter
doi:10.1016/S0377-0427(03)00555-7

$(k = 1,2,\ldots,K)$ be a multisplitting of the coefficient matrix $A \in L(\mathbb{R}^n)$, that is, the collection of triples (B_k, C_k, E_k) $(k = 1,2,\ldots,K)$ satisfies: (1) $A = B_k - C_k$; (2) B_k is nonsingular; and (3) E_k is *nonnegative* diagonal such that $\sum_{k=1}^{K} E_k = I$ (the identity matrix). Then, the multisplitting iteration method [6,1] can be written as

$$x^{p+1} = \sum_{k=1}^{K} E_k B_k^{-1} C_k x^p + \sum_{k=1}^{K} E_k B_k^{-1} b, \qquad p = 0,1,2,\ldots. \tag{2}$$

In practical implementations, at each major stage of the iteration (2) the kth processor computes only those entries of the local iteration $x^{p,k} = B_k^{-1} C_k x^p + B_k^{-1} b$, which correspond to the nonzero diagonal entries of E_k. The processor then scales these entries so as to be able to deliver the vector $E_k x^{p,k}$ to the host processor. The asymptotic and monotone convergence properties of this multisplitting iteration method were studied in [1,5–7], respectively.

The multisplitting iteration method (2) can attain maximum efficiency in practical implementation provided the multiple splittings $A = B_k - C_k (k = 1,2,\ldots,K)$ and the weighting matrices $E_k (k = 1,2,\ldots,K)$ are carefully chosen such that the workload carried by each processor is roughly equally distributed. When such a balance can be achieved, then the individual processors are ready to contribute towards their update of the global iteration x^{p+1} at the same time, which, in turn, minimizes idle time. However, there are applications in which the original physical properties lead to problem (1) which quite naturally divides into subproblems of unequal sizes.

To avoid loss of time and efficiency in processor utilization, Bru et al. [4] further improved the multisplitting iteration method (2) and suggested a parallel chaotic multisplitting iteration method. Moreover, they proved the convergence of this method when the coefficient matrix $A \in L(\mathbb{R}^n)$ is monotone and the multiple splittings are weak regular. Based on this work, many authors further developed new methods and studied their convergence properties from different angles. As customary, from now on the parallel chaotic multisplitting iteration method in [4] will be called the parallel nonstationary multisplitting iteration method so that it is distinguished from the chaotic asynchronous iteration methods. We remark that various methods of asynchronous matrix multisplitting iterations for solving the system of linear equations (1) were discussed in [1,3] and references therein.

In this paper, we will further investigate the convergence properties of the above parallel nonstationary multisplitting iteration method. After briefly stating a convergence theorem about the H-matrix class, we prove the convergence of a generalized variant of the parallel nonstationary multisplitting iteration method when the coefficient matrix $A \in L(\mathbb{R}^n)$ is symmetric positive definite and the multisplittings satisfy certain conditions. Then, for the monotone matrix class, we establish the monotone convergence theory as well as the monotone comparison theorem of the parallel nonstationary multisplitting iteration method. Therefore, the convergence theory of this class of parallel nonstationary multisplitting iteration method is further developed.

2. The parallel nonstationary multisplitting iteration methods

Let $N_0 = \{0,1,2,\ldots\}$, and $(B_{p,k}, C_{p,k}, E_{p,k})$ $(k = 1,2,\ldots,K)$, $p \in N_0$, be a sequence of multisplittings of the matrix $A \in L(\mathbb{R}^n)$. That is to say, for $\forall p \in N_0$ and $\forall k \in \{1,2,\ldots,K\}$, it holds that: (1) $A = B_{p,k} - C_{p,k}$; (2) $B_{p,k}$ is nonsingular; and (3) $E_{p,k}$ is an $n \times n$ diagonal matrix, satisfying

$\sum_{k=1}^{K} E_{p,k} = I$. Note that here we permit *negative* entries on the diagonal of $E_{p,k}$. Then we consider the following parallel nonstationary multisplitting iteration method for solving the system of linear equations (1).

Method 2.1 (*Parallel Non-Stationary Multisplitting Method*):

1. *Choose an arbitrary starting vector* $x^0 \in \mathbb{R}^n$. Set $p := 0$.
2. *For each* $k \in \{1, 2, \dots, K\}$, *set* $x^{p,k,0} := x^p$, *and take an integer* $\mu_{p,k} > 0$.
3. *For each* $k \in \{1, 2, \dots, K\}$ *and* $\mu = 1$ *to* $\mu_{p,k}$, *let* $x^{p,k,\mu}$ *be the solution of the linear system*: $B_{p,k} x = C_{p,k} x^{p,k,\mu-1} + b$.
4. *For each* $k \in \{1, 2, \dots, K\}$, *set* $x^{p+1,k} := x^{p,k,\mu_{p,k}}$.
5. $x^{p+1} = \sum_{k=1}^{K} E_{p,k} x^{p+1,k}$.
6. *If* $x^{p+1} = x^p$, *then stop. Otherwise, set* $p := p + 1$ *and return to Step* 2.

For each $k \in \{1, 2, \dots, K\}$ and each $p \in N_0$, we introduce the affine operator $F_{p,k} : \mathbb{R}^n \to \mathbb{R}^n$ as $F_{p,k}(x) = B_{p,k}^{-1} C_{p,k} x + B_{p,k}^{-1} b$. Furthermore, if for a nonnegative integer μ we define $F_{p,k}^{\mu} = I$ when $\mu = 0$ and $F_{p,k}^{\mu} = \overbrace{F_{p,k} \circ F_{p,k} \circ \cdots \circ F_{p,k}}^{\mu \text{ times}}$ when $\mu > 0$, where μ is the number of compositions of $F_{p,k}$ with itself, then Method 2.1 can be rewritten in the following concise form:

$$x^{p+1} = \sum_{k=1}^{K} E_{p,k} F_{p,k}^{\mu_{p,k}}(x^p), \quad p = 0, 1, 2, \dots \quad . \tag{3}$$

Evidently, for the original stationary multisplitting (B_k, C_k, E_k) ($k = 1, 2, \dots, K$) of the matrix $A \in L(\mathbb{R}^n)$, Method 2.1 becomes the parallel chaotic multisplitting iteration method studied in [4]. In particular, when $\mu_{p,k} \equiv 1$, it recovers the parallel matrix multisplitting method (2). Otherwise, if $(B_{p,k}, C_{p,k}, E_{p,k})$ ($k = 1, 2, \dots, K$) is a dynamic multisplitting of the matrix $A \in L(\mathbb{R}^n)$, Method 2.1 introduces a new parallel nonstationary multisplitting iteration method.

In the implementations of Method 2.1, each processor can carry out a varying number of local iterations until a mutual phase time is reached when all processors are ready to contribute towards the global iteration. Therefore, this method can achieve high parallel efficiency, even for the case of unbalanced workload distribution.

When the coefficient matrix $A \in L(\mathbb{R}^n)$ is monotone and the multiple splittings are weak regular, similar to [4] we can prove the following convergence theorem for Method 2.1.

Theorem 2.1. *Let* $A \in L(\mathbb{R}^n)$ *be a monotone matrix and* $(B_{p,k}, C_{p,k}, E_{p,k})$ $(k = 1, 2, \dots, K)$, $p \in N_0$, *be a sequence of multisplittings of matrix A. Assume that the weighting matrices* $E_{p,k} \geqslant 0$ $(k = 1, 2, \dots, K, p \in N_0)$, *all splittings* $A = B_{p,k} - C_{p,k}$ $(k = 1, 2, \dots, K, p \in N_0)$ *are weak regular and there exist monotone matrices* $\bar{B}_k (k = 1, 2, \dots, K)$ *such that* $B_{p,k}^{-1} \geqslant \bar{B}_k^{-1}$ $(k = 1, 2, \dots, K, p \in N_0)$. *Then, for any initial vector* $x^0 \in \mathbb{R}^n$, *the sequence* $\{x^p\}_{p \in N_0}$ *generated by Method* 2.1 *converges to* $x^* \in \mathbb{R}^n$ *whenever* $\mu_{p,k} \geqslant 1$ $(k = 1, 2, \dots, K, p \in N_0)$.

We remark that the assumption $\mu_{p,k} \geqslant 1$ $(k = 1, 2, \dots, K, p \in N_0)$ can be weakened as follows: $\mu_{p,k} \geqslant 0$ $(k = 1, 2, \dots, K, p \in N_0)$ and for infinitely many p's, $\mu_{p,k} \geqslant 1$, for all $k = 1, 2, \dots, K$. The difference between these two kinds of conditions is that the latter permits, if necessary, for any processor to skip its contribution to any major step of the iteration provided that infinitely often all processors contribute simultaneously towards a global iteration when the iteration index p tends to infinity.

More generally, analogously to [5], we can prove the convergence of Method 2.1 for the H-matrix class. The corresponding theorem is stated below.

Theorem 2.2. *Let $A \in L(\mathbb{R}^n)$ be an H-matrix and $(B_{p,k}, C_{p,k}, E_{p,k})$ $(k = 1, 2, \dots, K)$, $p \in N_0$, be a sequence of multisplittings of matrix A. Assume that the weighting matrices $E_{p,k} \geqslant 0$ $(k=1, 2, \dots, K, p \in N_0)$ and all splittings $A = B_{p,k} - C_{p,k}$ $(k = 1, 2, \dots, K, p \in N_0)$ satisfy $\operatorname{diag}(B_{p,k}) = \operatorname{diag}(A)$ and $\langle A \rangle = \langle B_{p,k} \rangle - |C_{p,k}|$. Then, for any initial vector $x^0 \in \mathbb{R}^n$, the sequence $\{x^p\}_{p \in N_0}$ generated by Method 2.1 converges to $x^* \in \mathbb{R}^n$ provided $\mu_{p,k} \geqslant 1$ $(k = 1, 2, \dots, K, p \in N_0)$.*

We point out that the same remark about Theorem 2.1 is valid for Theorem 2.2. In addition, for the definitions of the comparison matrix $\langle \bullet \rangle$, the absolute value matrix $|\bullet|$ and the related concepts, one can refer to [1,5] for details.

3. Convergence theory for the positive definite matrix

Let $\{A_p\}_{p \in N_0}$ be a sequence of symmetric matrices in $L(\mathbb{R}^n)$. Then we call $A_p (p \in N_0)$ positive definite uniformly in p if there exists a positive constant c, independent of p, such that $x^{\mathrm{T}} A_p x \geqslant c x^{\mathrm{T}} x$ holds for all $x \in \mathbb{R}^n$. When the coefficient matrix $A \in L(\mathbb{R}^n)$ is symmetric positive definite, we have the following convergence theorem for Method 2.1.

Theorem 3.1. *Let $A \in L(\mathbb{R}^n)$ be a symmetric positive definite matrix, and for every $p \in N_0$, $(B_{p,k}, C_{p,k}, E_{p,k})$ $(k = 1, 2, \dots, K)$ be a multisplitting of matrix A such that*:

(a) *$B_{p,k} + C_{p,k}$ $(k = 1, 2, \dots, K)$ are positive definite uniformly in p; and*
(b) *$f(\sum_{k=1}^{K} E_{p,k} x^{p+1,k}) \leqslant \max_{1 \leqslant k \leqslant K} f(x^{p+1,k})$, where $f(x) = \frac{1}{2} x^{\mathrm{T}} A x - x^{\mathrm{T}} b$.*

Let $\mu_{p,k}$ $(k = 1, 2, \dots, K, p \in N_0)$ be positive integers bounded uniformly from above. Then the sequence $\{x^p\}_{p \in N_0}$ generated by Method 2.1 converges to $x^* \in \mathbb{R}^n$ independently of the positive integer sequences $\{\mu_{p,k}\}_{p \in N_0}$ $(k = 1, 2, \dots, K)$.

Proof. Through straightforward deduction we can obtain the identity $f(x) - f(x^*) = \frac{1}{2}(x - x^*)^{\mathrm{T}} A (x - x^*)$. Therefore, x^* is the unique solution of the system of linear equations (1) if and only if it is the unique global minimum point of the quadratic function $f(x)$. Furthermore, let $A = B - C$ be a splitting of the matrix A, i.e., $B \in L(\mathbb{R}^n)$ is a nonsingular matrix, and define $\bar{x} = B^{-1} C x + B^{-1} b$. Then we can obtain the equality $f(x) - f(\bar{x}) = \frac{1}{2}(x - \bar{x})^{\mathrm{T}} (B + C)(x - \bar{x})$. It follows from this equality and Method 2.1 that for $\mu = 1, 2, \dots, \mu_{p,k}$ and $p \in N_0$ we have

$$f(x^{p,k,\mu-1}) - f(x^{p,k,\mu}) = \frac{1}{2}(x^{p,k,\mu-1} - x^{p,k,\mu})^{\mathrm{T}} (B_{p,k} + C_{p,k})(x^{p,k,\mu-1} - x^{p,k,\mu}).$$

In accordance with assumption (a) there exists a positive constant c, independent of p and k, such that $x^{\mathrm{T}}(B_{p,k}+C_{p,k})x \geqslant cx^{\mathrm{T}}x$. Therefore,

$$\begin{aligned} f(x^p) &- f(x^{p,k,\mu_{p,k}}) \\ &= \sum_{\mu=1}^{\mu_{p,k}} (f(x^{p,k,\mu-1}) - f(x^{p,k,\mu})) \\ &= \frac{1}{2}\sum_{\mu=1}^{\mu_{p,k}} (x^{p,k,\mu-1} - x^{p,k,\mu})^{\mathrm{T}}(B_{p,k}+C_{p,k})(x^{p,k,\mu-1} - x^{p,k,\mu}) \\ &\geqslant \frac{c}{2}\sum_{\mu=1}^{\mu_{p,k}} \|x^{p,k,\mu-1} - x^{p,k,\mu}\|_2^2. \end{aligned} \tag{4}$$

In addition, by further noticing that $\{\mu_{p,k}\}$ is uniformly bounded from above by a positive integer, say J, we can get

$$\begin{aligned} \|x^p &- x^{p,k,\mu_{p,k}}\|_2^2 \\ &= \left\| \sum_{\mu=1}^{\mu_{p,k}} (x^{p,k,\mu-1} - x^{p,k,\mu}) \right\|_2^2 \leqslant \left(\sum_{\mu=1}^{\mu_{p,k}} \|x^{p,k,\mu-1} - x^{p,k,\mu}\|_2 \right)^2 \\ &\leqslant \mu_{p,k} \sum_{\mu=1}^{\mu_{p,k}} \|x^{p,k,\mu-1} - x^{p,k,\mu}\|_2^2 \leqslant J \sum_{\mu=1}^{\mu_{p,k}} \|x^{p,k,\mu-1} - x^{p,k,\mu}\|_2^2. \end{aligned}$$

Substituting this estimate into (4) yields $f(x^p) - f(x^{p+1,k}) \geqslant (c/2J)\|x^p - x^{p+1,k}\|_2^2$. From this inequality and assumption (b) we know that for all $p \in N_0$,

$$\begin{aligned} f(x^p) &- f(x^{p+1}) \\ &= f(x^p) - f\left(\sum_{k=1}^{K} E_{p,k} x^{p+1,k} \right) \geqslant f(x^p) - \max_{1\leqslant k\leqslant K} f(x^{p+1,k}) \\ &= f(x^p) - f(x^{p+1,k_{p+1}}) \geqslant \frac{c}{2J}\|x^p - x^{p+1,k_{p+1}}\|_2^2 \end{aligned} \tag{5}$$

holds, where k_{p+1} is an index such that $f(x^{p+1,k_{p+1}}) = \max_{1\leqslant k\leqslant K} f(x^{p+1,k})$.

We now prove that the sequence $\{x^p\}_{p\in N_0}$ is bounded. Otherwise, suppose that the sequence $\{x^p\}_{p\in N_0}$ is unbounded. Then there exists at least one subsequence $\{x^{p_\ell}\}_{\ell\in N_0}$ such that $\|x^{p_\ell}\|_2 \to \infty$ as $\ell \to \infty$. Since the identity $f(x) - f(x^*) = \frac{1}{2}(x-x^*)^{\mathrm{T}}A(x-x^*)$ shows that $\{f(x^p)\}_{p\in N_0}$ is bounded from below by $f(x^*)$ and (5) implies that $\{f(x^p)\}_{p\in N_0}$ is monotonically decreasing, we know that the sequence $\{f(x^{p_\ell})\}_{\ell\in N_0}$ is convergent. In particular, $\{f(x^{p_\ell})\}_{\ell\in N_0}$ is bounded. Now, consider the corresponding normalized sequence $\{x^{p_\ell}/\|x^{p_\ell}\|_2\}_{\ell\in N_0}$, which is bounded and hence has an accumulation point $\tilde{x}$ such that $\|\tilde{x}\|_2 = 1$. Assume, without loss of generality, that $\{x^{p_\ell}/\|x^{p_\ell}\|_2\}_{\ell\in N_0}$ converges to $\tilde{x}$. By passing to the limit $\ell \to \infty$, from $f(x^{p_\ell}) = \frac{1}{2}(x^{p_\ell})^{\mathrm{T}}Ax^{p_\ell} - (x^{p_\ell})^{\mathrm{T}}b$ we immediately get $\tilde{x}^T A\tilde{x} = 0$. This obviously contradicts the symmetric positive definiteness of the matrix $A \in L(\mathbb{R}^n)$. Therefore, the sequence $\{x^p\}_{p\in N_0}$ must be bounded.

In the following, we will further demonstrate that the sequence $\{x^p\}_{p\in N_0}$ generated by Method 2.1 converges to the unique solution of the system of linear equations (1). To this end, we only need to verify that every accumulation point of the sequence $\{x^p\}_{p\in N_0}$ is a solution of the system of linear equations (1). Let $\hat{x}$ be an arbitrary accumulation point of the sequence $\{x^p\}_{p\in N_0}$, and $\{x^{p_\ell}\}_{\ell\in N_0}$ be a subsequence that converges to $\hat{x}$. Since $\{f(x^{p_\ell})\}_{\ell\in N_0}$ converges to $f(\hat{x})$ as $\ell \to \infty$ and $\{f(x^p)\}_{p\in N_0}$ is nonincreasing by (5), the entire sequence $\{f(x^p)\}_{p\in N_0}$ converges to $f(\hat{x})$, too. Let the positive integer $k_{p+1} \in \{1, 2, \ldots, K\}$ be defined as in (5). Then by taking a further subsequence if necessary, we may assume that there exists some index $\hat{k} \in \{1, 2, \ldots, K\}$ such that $k_{p_\ell+1} = \hat{k}$ for all $\ell \in N_0$. Then the sequence $\{x^{p_\ell} - x^{p_\ell+1,\hat{k}}\}_{\ell\in N_0}$ converges to zero by (5), and the sequence $\{x^{p_\ell+1,\hat{k}}\}_{\ell\in N_0}$ converges to $\hat{x}$ as $\ell \to \infty$. From (4) it further holds that for any $\mu \in \{1, 2, \ldots, \mu_{p_\ell,\hat{k}}\}$, the sequence $\{x^{p_\ell,\hat{k},\mu}\}_{\ell\in N_0}$ converges to $\hat{x}$ as $\ell \to \infty$. Because $x^{p_\ell+1,\hat{k}}$ is a solution of the linear system $B_{p_\ell,\hat{k}} x^{p_\ell+1,\hat{k},\mu_{p_\ell,\hat{k}}-1} = C_{p_\ell,\hat{k}} x^{p_\ell,\hat{k}} + b$, it follows that $\hat{x}$ solves the system of linear equations (1). The proof of this theorem is fulfilled. $\square$

We remark that Theorem 3.1 can be straightforwardly generalized to the complex matrix case. In addition, assumption (a) in Theorem 3.1 is a standard condition imposed to guarantee the convergence of the iterative methods for the system of linear equations, and assumption (b) can be satisfied by various choices of the weighting matrices $E_{p,k}$ $(k = 1, 2, \ldots, K, p \in N_0)$. One of the possibilities is given by the choices of $E_{p,k} = \alpha_{p,k} I$ $(k = 1, 2, \ldots, K, p \in N_0)$, where $\alpha_{p,k}$ $(k = 1, 2, \ldots, K, p \in N_0)$ are nonnegative real numbers satisfying $\sum_{k=1}^K \alpha_{p,k} = 1$ $(p \in N_0)$. In this case, condition (b) is automatically satisfied if either of the following three classes of restrictions are further imposed:

(1) for $\{k_p \mid p \in N_0\} \subseteq \{1, 2, \ldots, K\}$, $\alpha_{p,k} = 1$ if $k = k_p$ and $\alpha_{p,k} = 0$ if $k \neq k_p$, where the indices k_p $(p \in N_0)$ are chosen either randomly at every iteration, or in a certain predetermined order such as the cyclic rule, or based on the function values $f(x^{p,k})$ $(k = 1, 2, \ldots, K)$ such that, for $p \in N_0$, $f(x^{p,k_p}) = \min_{1\leqslant k\leqslant K} f(x^{p,k})$;
(2) $\alpha_{p,k}$ $(k = 1, 2, \ldots, K, p \in N_0)$ are the minimizers of the functions

$$g(\alpha_{p,1}, \alpha_{p,2}, \ldots, \alpha_{p,K}) = f\left(\sum_{k=1}^K \alpha_{p,k} x^{p,k}\right), \quad p \in N_0;$$

(3) $A \in L(\mathbb{R}^n)$ is a positive semidefinite matrix.

Moreover, in Theorem 3.1 we does not make the hypothesis that the weighting matrices $E_{p,k}$ $(k = 1, 2, \ldots, K, p \in N_0)$ are nonnegative, which used to be an elementary hypothesis for establishing the convergence theories of the parallel multisplitting iteration methods. That is to say, even if some diagonal elements of the matrices $E_{p,k}$ $(k = 1, 2, \ldots, K, p \in N_0)$ are negative, Method 2.1 still converges provided the conditions of Theorem 3.1 are satisfied. The following example further gives concrete illustration about Theorem 3.1.

Example 3.1. $A = I \in L(\mathbb{R}^2)$, $b = 0 \in \mathbb{R}^2$. Evidently, $x^* = 0$. For simplicity, we take $K = 2$ and $B_k = \mathrm{diag}(1/(1 - \sigma_k), 1/(1 - \delta_k))$, $k = 1, 2$, where $\sigma_k, \delta_k \in \mathbb{R}^1 \setminus \{1\}$, $k = 1, 2$. Then we get two splittings $A = B_k - C_k$, $k = 1,2$, where $C_k = \mathrm{diag}(\sigma_k/(1-\sigma_k), \delta_k/(1-\delta_k))$. By direct computations, we have

$$H_k = B_k^{-1}C_k = \mathrm{diag}\,(\sigma_k, \delta_k), \quad Q_k = B_k + C_k = \mathrm{diag}\left(\frac{1+\sigma_k}{1-\sigma_k}, \frac{1+\delta_k}{1-\delta_k}\right)$$

and $x^{p+1,k} = H_k^{\mu_{p,k}} x^p = (\sigma_k^{\mu_{p,k}}[x^p]_1, \delta_k^{\mu_{p,k}}[x^p]_2)^{\mathrm{T}}$, where $x^p = ([x^p]_1, [x^p]_2)^{\mathrm{T}}$. Therefore,

$$f(x^{p+1,k}) = \frac{1}{2}(x^{p+1,k})^{\mathrm{T}} x^{p+1,k} = \frac{1}{2}(|\sigma_k|^{2\mu_{p,k}}([x^p]_1)^2 + |\delta_k|^{2\mu_{p,k}}([x^p]_2)^2).$$

Now, consider the following parallel nonstationary multisplitting methods corresponding to different cases of the weighting matrices:

(i) $E_{p,1} = \mathrm{diag}(1,0)$, $E_{p,2} = \mathrm{diag}(0,1)$. We have $x^{p+1} = (\sigma_1^{\mu_{p,1}}[x^p]_1, \delta_2^{\mu_{p,2}}[x^p]_2)^{\mathrm{T}}$, and hence, $f(x^{p+1}) = \frac{1}{2}(|\sigma_1|^{2\mu_{p,1}}([x^p]_1)^2 + |\delta_2|^{2\mu_{p,2}}([x^p]_2)^2)$. If we let σ_k, δ_k be such that $|\sigma_1| > |\sigma_2|$ and $|\delta_2| > |\delta_1|$, and $\mu_{p,k}$ be such that $\mu_{p,1} = \mu_{p,2} = \mu_p$, then it holds that $f(x^{p+1,1}) < f(x^{p+1})$ and $f(x^{p+1,2}) < f(x^{p+1})$ when $[x^p]_1 \neq 0$ and $[x^p]_2 \neq 0$. This implies that $f(x^{p+1}) > \max_{1 \leqslant k \leqslant 2} f(x^{p+1,k})$, and hence, assumption (b) of Theorem 3.1 is not satisfied. However, if we let σ_k, δ_k be such that $|\sigma_1| \leqslant |\sigma_2|$ or $|\delta_2| \leqslant |\delta_1|$, and $\mu_{p,k}$ be such that $\mu_{p,1} = \mu_{p,2} = \mu_p$, then it holds that $f(x^{p+1}) \leqslant \max_{1 \leqslant k \leqslant 2} f(x^{p+1,k})$, and hence, assumption (b) of Theorem 3.1 is satisfied.

(i_1) If we further let $\sigma_k, \delta_k \in (-1,1)$, then Q_k $(k=1,2)$ are symmetric positive definite matrices. Therefore, assumption (a) of Theorem 3.1 is satisfied. Moreover, we know that $\{x^p\}_{p \in N_0}$ is convergent. This shows that if assumption (a) is satisfied, then Method 2.1 is convergent whether assumption (b) is satisfied or not.

(i_2) If we further let $|\sigma_1| > 1$, then Q_1 is not a positive definite matrix. Therefore, assumption (a) of Theorem 3.1 is not satisfied. Moreover, we know that $\{x^p\}_{p \in N_0}$ is divergent. This shows that if assumption (a) is not satisfied, then Method 2.1 is divergent whether assumption (b) is satisfied or not.

(i_3) If we further let $|\sigma_2| > 1$, then Q_2 is not a positive definite matrix. Therefore, assumption (a) of Theorem 3.1 is not satisfied. Moreover, we know that $\{x^p\}_{p \in N_0}$ is convergent provided $|\sigma_1| < 1 < |\sigma_2|$ and $|\delta_2| < 1$, and divergent provided $|\sigma_1| > 1$. This shows that if assumption (a) is not satisfied but assumption (b) is satisfied, then Method 2.1 is either convergent or divergent.

From (i_1)–(i_3) we easily know that assumptions (a) and (b) of Theorem 3.1 are only sufficient conditions for guaranteeing the convergence of Method 2.1, but not necessary ones.

(ii) $E_{p,1} = \mathrm{diag}(-\frac{3}{4}, \frac{3}{4})$, $E_{p,2} = \mathrm{diag}(\frac{7}{4}, \frac{1}{4})$. Evidently, we have

$$x^{p+1} = \left(\frac{1}{4}[-3\sigma_1^{\mu_{p,1}} + 7\sigma_2^{\mu_{p,2}}][x^p]_1, \frac{1}{4}[3\delta_1^{\mu_{p,1}} + \delta_2^{\mu_{p,2}}][x^p]_2\right)^{\mathrm{T}},$$

$$f(x^{p+1}) = \frac{1}{32}([-3\sigma_1^{\mu_{p,1}} + 7\sigma_2^{\mu_{p,2}}]^2([x^p]_1)^2 + [3\delta_1^{\mu_{p,1}} + \delta_2^{\mu_{p,2}}]^2([x^p]_2)^2).$$

If we let $\sigma_1 = \frac{2}{3}$, $\sigma_2 = \frac{1}{3}$, $\delta_1 = \delta_2 = \frac{1}{2}$ and $\mu_{p,k} = 1$, then it holds that

$$f(x^{p+1,k}) = \begin{cases} \frac{1}{2}(\frac{4}{9}([x^p]_1)^2 + \frac{1}{4}([x^p]_2)^2) & \text{if } k = 1, \\ \frac{1}{2}(\frac{1}{9}([x^p]_1)^2 + \frac{1}{4}([x^p]_2)^2) & \text{if } k = 2, \end{cases}$$

$$f(x^{p+1}) = \frac{1}{32}\left(\frac{1}{9}([x^p]_1)^2 + \frac{9}{4}([x^p]_2)^2\right).$$

Hence, $f(x^{p+1}) \leqslant \max_{1 \leqslant k \leqslant 2} f(x^{p+1,k})$, i.e., assumption (b) of Theorem 3.1 is satisfied. Clearly, $Q_k (k = 1, 2)$ are symmetric positive definite matrices, and therefore, assumption (a) of Theorem 3.1 is also satisfied. Moreover, by direct computations we have $x^{p+1} = (\frac{1}{12}[x^p]_1, \frac{1}{2}[x^p]_2)^{\mathrm{T}}$, and hence, $\{x^p\}_{p \in N_0}$ is convergent. This shows that even if some entries on the diagonal of the weighting matrices are negative, assumptions (a) and (b) of Theorem 3.1 still hold and Method 2.1 converges.

4. The monotone convergence theory

For simplicity but without loss of generality, in this section we only consider a special case of Method 2.1, for which $B_{p,k} = B_k$, $C_{p,k} = C_k$, $E_{p,k} = E_k$ and $\mu_{p,k} = \mu_k$. Analogous to (3), the method just mentioned can be expressed as

$$x^{p+1} = \sum_{k=1}^{K} E_k F_k^{\mu_k}(x^p) \quad \text{where } F_k(x) = B_k^{-1} C_k x + B_k^{-1} b. \tag{6}$$

In the following, we will discuss the monotone convergence properties of the parallel nonstationary multisplitting iteration method (6) and investigate the influence of the multiple splittings and the composition numbers upon the convergence behavior of this method by slight and technical modification of the theorems and proofs in [7,2]. For this purpose, we assume throughout this section that $A \in L(\mathbb{R}^n)$ is a monotone matrix, and (B_k, C_k, E_k) $(k = 1, 2, \ldots, K)$ is its multisplitting where $A = B_k - C_k$ $(k = 1, 2, \ldots, K)$ are weak regular splittings and $E_k \geqslant 0$ $(k = 1, 2, \ldots, K)$. In addition, we introduce matrices

$$R = \sum_{k=1}^{K} E_k \sum_{\mu=0}^{\mu_k - 1} (B_k^{-1} C_k)^{\mu_k} B_k^{-1}, \qquad H = \sum_{k=1}^{K} E_k (B_k^{-1} C_k)^{\mu_k}. \tag{7}$$

Evidently, it holds that $H = I - RA$ and (6) can be equivalently written as

$$x^{p+1} = Hx^p + Rb, \quad p = 0, 1, 2, \ldots. \tag{8}$$

Based upon (7) and (8), we can straightforwardly obtain the following two-sided monotone approximation properties of iteration (6).

Theorem 4.1. *Assume that x^0 and y^0 are initial vectors obeying $x^0 \leqslant y^0$ and $Ax^0 \leqslant b \leqslant Ay^0$, and $\{x^p\}_{p \in N_0}$ and $\{y^p\}_{p \in N_0}$ are sequences starting from x^0 and y^0, respectively, and generated by (8).*

Then

(1) $x^p \leqslant x^{p+1} \leqslant y^{p+1} \leqslant y^p,\ p \in N_0$;
(2) $\lim_{p\to\infty} x^p = x^* = \lim_{p\to\infty} y^p$; *and*
(3) *for any* $z^0 \in \mathbb{R}^n$ *obeying* $x^0 \leqslant z^0 \leqslant y^0$, *the sequence* $\{z^p\}_{p\in N_0}$ *starting from* z^0 *and generated by* (8) *satisfies* $x^p \leqslant z^p \leqslant y^p$ $(\forall p \in N_0)$. *Hence*, $\lim_{p\to\infty} z^p = x^*$.

Theorem 4.2. *Let the conditions of Theorem* 4.1 *be satisfied. If we additionally suppose that* $R^{-1}H \geqslant 0$, *then it holds that* $Ax^p \leqslant b \leqslant Ay^p$, $p \in N_0$, *where* $\{x^p\}$ *and* $\{y^p\}$ *are sequences generated by* (8) *starting from* x^0 *and* y^0, *respectively.*

With Theorems 4.1 and 4.2, we can further compare the convergence rates of the parallel nonstationary multisplitting iteration methods, resulting from different multiple splittings $A = B_k^{(m)} - C_k^{(m)}$ $(k=1,2,\dots,K)$, $m=1,2$, and different composition numbers $\mu_k^{(m)}$ $(k=1,2,\dots,K, m=1,2)$, for solving the system of linear equations (1) in the sense of monotonicity. To this end, corresponding to (7) we construct matrices

$$R^{(m)} = \sum_{k=1}^{K} E_k \sum_{\mu=0}^{\mu_k^{(m)}-1} (B_k^{(m)^{-1}} C_k^{(m)})^{\mu} B_k^{(m)^{-1}},$$

$$H^{(m)} = \sum_{k=1}^{K} E_k (B_k^{(m)^{-1}} C_k^{(m)})^{\mu_k^{(m)}}, \quad m = 1,2. \tag{9}$$

Analogously, we have $H^{(m)} = I - R^{(m)}A$, $m = 1,2$.

Now, we consider the comparison of the monotone convergence rates between the sequences $\{x^p\}$ and $\{y^p\}$, defined according to (8) by

$$x^{p+1} = H^{(1)}x^p + R^{(1)}b, \quad y^{p+1} = H^{(2)}y^p + R^{(2)}b, \quad p = 0,1,2,\dots. \tag{10}$$

Theorem 4.3. *Let* $A \in L(\mathbb{R}^n)$ *be a monotone matrix, and* $(B_k^{(m)}, C_k^{(m)}, E_k)$ $(k=1,2,\dots,K)$, $m=1,2$, *be its two multisplittings where* $A = B_k^{(m)} - C_k^{(m)}$ $(k=1,2,\dots,K, m=1,2)$ *are weak regular splittings and* $E_k \geqslant 0$ $(k=1,2,\dots,K)$. *Assume that* $x^0 = y^0$ *is an initial vector, and* $\{x^p\}$ *and* $\{y^p\}$ *are sequences defined by* (10). *If either* $R^{(1)^{-1}}H^{(1)} \geqslant 0$ *or* $R^{(2)^{-1}}H^{(2)} \geqslant 0$ *holds, then we have* (a) $x^p \geqslant y^p$ $(p \in N_0)$ *as* $Ax^0 \leqslant b$; (b) $x^p \leqslant y^p$ $(p \in N_0)$ *as* $Ax^0 \geqslant b$, *provided for* $k = 1,2,\dots,K$, $\mu_k^{(1)} \geqslant \mu_k^{(2)}$ *and*

$$(B_k^{(1)^{-1}} C_k^{(1)})^{\mu} B_k^{(1)^{-1}} \geqslant (B_k^{(2)^{-1}} C_k^{(2)})^{\mu} B_k^{(2)^{-1}}, \quad \mu = 0,1,2,\dots,\mu_k^{(2)}. \tag{11}$$

In particular, from [2] *we see that* (11) *holds if* $B_k^{(1)^{-1}} \geqslant B_k^{(2)^{-1}}$ *and either* $C_k^{(1)} B_k^{(1)^{-1}} \geqslant 0$ *or* $C_k^{(2)} B_k^{(2)^{-1}} \geqslant 0$.

Proof. Because the proof of (b) is much analogous to that of (a), we only prove (a) by induction. For $p = 0$, (a) is obviously true. Suppose that (a) has been demonstrated for all $p \leqslant \ell$. Since in accordance with Theorems 4.1 and 4.2, both sequences $\{x^p\}$ and $\{y^p\}$ are monotonously increasing, and satisfy $Ax^p \leqslant b$ and $Ay^p \leqslant b$, we see that there exists a nonnegative vector $u^\ell \in \mathbb{R}^n$ such that $Ax^\ell + u^\ell = b$. By (10) and (9) we have

$$
\begin{aligned}
x^{\ell+1} &= \sum_{k=1}^{K} E_k \left[(B_k^{(1)^{-1}} C_k^{(1)})^{\mu_k^{(1)}} x^\ell + \sum_{\mu=0}^{\mu_k^{(1)}-1} (B_k^{(1)^{-1}} C_k^{(1)})^{\mu} B_k^{(1)^{-1}} (Ax^\ell + u^\ell) \right] \\
&= x^\ell + \sum_{k=1}^{K} E_k \sum_{\mu=0}^{\mu_k^{(1)-1}} (B_k^{(1)^{-1}} C_k^{(1)})^{\mu} B_k^{(1)^{-1}} u^\ell \\
&\geqslant x^\ell + \sum_{k=1}^{K} E_k \sum_{\mu=0}^{\mu_k^{(2)}-1} (B_k^{(2)^{-1}} C_k^{(2)})^{\mu} B_k^{(2)^{-1}} u^\ell \\
&= \sum_{k=1}^{K} E_k \left[(B_k^{(2)^{-1}} C_k^{(2)})^{\mu_k^{(2)}} x^\ell + \sum_{\mu=0}^{\mu_k^{(2)}-1} (B_k^{(2)^{-1}} C_k^{(2)})^{\mu} B_k^{(2)^{-1}} (Ax^\ell + u^\ell) \right] \\
&= \sum_{k=1}^{K} E_k \left[(B_k^{(2)^{-1}} C_k^{(2)})^{\mu_k^{(2)}} x^\ell + \sum_{\mu=0}^{\mu_k^{(2)}-1} (B_k^{(2)^{-1}} C_k^{(2)})^{\mu} B_k^{(2)^{-1}} b \right] \geqslant y^{\ell+1}.
\end{aligned}
$$

Up to now, the induction is accomplished, and (a) is demonstrated. □

Theorem 4.3 immediately leads to the following comparison theorem between the multisplitting method and the single-splitting method.

Theorem 4.4. *Let $A = \underline{B} - \underline{C} = \bar{B} - \bar{C}$ be two regular splittings of matrix A. Assume that $\underline{x}^0 = x^0 = \bar{x}^0$ is an initial vector, $\{x^p\}_{p\in N_0}$ is the sequence defined by (6), and $\{\underline{x}^p\}_{p\in N_0}$ and $\{\bar{x}^p\}_{p\in N_0}$ are sequences defined, respectively, by*

$$
\underline{x}^{p+1} = (\underline{B}^{-1}\underline{C})^{\mu_{\min}} \underline{x}^p + \sum_{\mu=0}^{\mu_{\min}-1} (\underline{B}^{-1}\underline{C})^{\mu} \underline{B}^{-1} b,
$$

$$
\bar{x}^{p+1} = (\bar{B}^{-1}\bar{C})^{\mu_{\min}} \bar{x}^p + \sum_{\mu=0}^{\mu_{\min}-1} (\bar{B}^{-1}\bar{C})^{\mu} \bar{B}^{-1} b,
$$

where $\mu_{\min}$ is a positive integer satisfying $\mu_{\min} \leqslant \min_{1\leqslant k\leqslant K} \{\mu_k\}$. Then: (a) $\underline{x}^p \leqslant x^p \leqslant \bar{x}^p$ $(p \in N_0)$ *as $Ax^0 \leqslant b$;* (b) $\underline{x}^p \geqslant x^p \geqslant \bar{x}^p$ $(p \in N_0)$ *as $Ax^0 \geqslant b$, provided $\bar{B}^{-1} \geqslant B_k^{-1} \geqslant \underline{B}^{-1}$, $k = 1, 2, \ldots, K$.*

5. Conclusion and remarks

We have proved the convergence of the nonstationary multisplitting method for solving a system of linear equations when the coefficient matrix is symmetric positive definite. Although we have realized that the numerical behavior of the nonstationary multisplitting method is better than the synchronous multisplitting method and may be worse than the asynchronous multisplitting method, this theory is of both practical and theoretical importance as it does not need to assume that the weighting matrices are scalar and nonnegative ones, and affords one possible way to establish the convergence theory of the asynchronous multisplitting method for symmetric positive definite matrices.

References

[1] 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.

[2] Z.-Z. Bai, D.-R. Wang, The monotone convergence of the two-stage iterative method for solving large sparse systems of linear equations, Appl. Math. Lett. 10 (1) (1997) 113–117.

[3] Z.-Z. Bai, D.-R. Wang, D.J. Evans, A class of asynchronous matrix multi-splitting multi-parameter relaxation iterations, J. Comput. Math. 16 (1998) 221–238.

[4] R. Bru, L. Elsner, M. Neumann, Models of parallel chaotic iteration methods, Linear Algebra Appl. 103 (1988) 175–194.

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

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

[7] D.-R. Wang, Z.-Z. Bai, On the monotone convergence of matrix multisplitting iteration methods, in: E.X. Jiang (Ed.), Proceedings of the '92 Shanghai International Numerical Algebra and Its Application Conference, China Science and Technology Press, Beijing, 1994.