On the convergence of the parallel non-stationary multisplitting iteration methods*

Zhong-Zhi Bai

State Key Laboratory of Scientific/Engineering Computing

Institute of Computational Mathematics and Scientific/Engineering Computing

Chinese Academy of Sciences, P.O.Box 2719

Beijing 100080, P.R.China

Email: bzz@lsec.cc.ac.cn

*Project 19601036 supported by the NNSF of China

Abstract. The convergence properties of the non-stationary multisplitting iteration method, a variant of the parallel chaotic multisplitting iteration method for solving the large sparse system of linear equations presented in:

R. Bru, L. Elsner and M. Neumann: *Models of parallel chaotic iteration methods*, Linear Algebra and Its Applications, 103(1988),175-194)

are further discussed when the coefficient matrix is an H-matrix and 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 non-stationary multisplitting iteration method.

§1. Introduction

Consider the parallel solution of the large sparse system of linear equations

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

on a multiprocessor system.

The parallel multisplitting iteration method in

[OW] D.P. O'Leary and R.E. White: *Multi-splittings of matrices and parallel solution of linear systems*, **SIAM J. Alg. Disc. Methods, 6(1985), 630-640:**

Suppose the multiprocessor system have K processors, which are connected to a host processor that may be taken by any of the K processors.

Let $(B_k, C_k, E_k)(k = 1, 2, \cdots, K)$ be a multisplitting of the coefficient matrix $A \in L(R^n)$, that is, the collection of triples $(B_k, C_k, E_k)(k = 1, 2, \cdots, K)$ satisfies

(1) $A = B_k - C_k$**,** $k = 1, 2, \cdots, K$**;**

(2) $B_k(k = 1, 2, \cdots, K)$ **are nonsingular; and**

(3) $E_k(k = 1, 2, \cdots, K)$ **are nonnegative $n \times n$ diagonal matrices such that $\sum_{k=1}^{K} E_k = I$ (the identity matrix).**

Then the multisplitting iteration method in [OW] can be written as

$$x^{p+1} = \sum_{k=1}^{K} E_k B_k^{-1} C_k x^p + \sum_{k=1}^{K} E_k B_k^{-1} b, \qquad p = 0, 1, 2, \cdots. \tag{1.2}$$

• In practical implementations, at each major stage of the iteration (1.2) the k-th 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] D.P. O'Leary and R.E. White, Multi-splittings of matrices and parallel solution of linear systems, *SIAM J. Alg. Disc. Methods*, **6(1985), 630-640.**

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

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

[4] De-Ren Wang and Zhong-Zhi Bai, On the monotone convergence of matrix multisplitting iteration methods, In *Proceedings of '92 Shanghai International Numerical Algebra and Its Applications Conference*, **E.X. Jiang eds., China Science and Technology Press, Beijing, 1994.**

[5] Zhong-Zhi Bai, Jia-Chang Sun and De-Ren Wang, A unified framework for the construction of various matrix multisplitting iterative methods for large sparse system of linear equations, *Computers Math. Appl.*, **32:12(1996), 51-76.**

• The multisplitting iteration method (1.2) can attain max-

imum efficiency in practical implementation provided the multiple splittings $A = B_k - C_k (k = 1, 2, \cdots, K)$ and the weighting matrices $E_k (k = 1, 2, \cdots, K)$ are carefully chosen such that the workload carried by all processors 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 (1.1) which quite naturally divide into subproblems of unequal sizes.

To avoid loss of time and efficiency in processor utilization,

[BEN]: R. Bru, L. Elsner and M. Neumann, *Models of parallel chaotic iteration methods*, **Linear Algebra Appl., 103(1988), 175-194.**

further improved the multisplitting iteration method (1.2) and suggested a parallel chaotic multisplitting iteration method.

Moreover, they proved the convergence of this method when the coefficient matrix $A \in L(R^n)$ is monotone and the multiple splittings are weak regular.

This method was also called *the parallel non-stationary multisplitting iteration method* **by other authors.**

In this talk, we will report further investigatations on the convergence properties of the above parallel non-stationary 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 non-stationary multisplitting iteration method when the coefficient matrix $A \in L(R^n)$ 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 non-stationary multisplitting iteration method.**

- **Therefore, the convergence theory of this class of parallel non-stationary multisplitting iteration method is completed.**

§2. The parallel non-stationary multisplitting iteration methods

Let $N_0 = \{0, 1, 2, \cdots\}$, and $(B_{p,k}, C_{p,k}, E_{p,k})(k = 1, 2, \cdots, K)$, $p \in N_0$, be a sequence of multisplittings of the matrix $A \in L(R^n)$. That is to say, for $\forall p \in N_0$ and $\forall k \in \{1, 2, \cdots, K\}$, 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 non-stationary multisplitting iteration method for solving the system of linear equations (1.1).**

METHOD 2.1. *(Parallel Non-Stationary Multisplitting Method)*

Step 1. *Choose an arbitrary starting vector $x^0 \in R^n$. Set $p := 0$.*

Step 2. *For each $k \in \{1, 2, \cdots, K\}$, set $x^{p,k,0} := x^p$, and take a positive integer $\mu_{p,k}$.*

Step 3. *For each $k \in \{1, 2, \cdots, K\}$ and $\mu = 1$ to $\mu_{p,k}$, let $x^{p,k,\mu}$ be the solution of the linear system:*

$$B_{p,k} x = C_{p,k} x^{p,k,\mu-1} + b.$$

Step 4. *For each $k \in \{1, 2, \cdots, K\}$, set $x^{p+1,k} := x^{p,k,\mu_{p,k}}$.*

Step 5. $x^{p+1} = \sum\limits_{k=1}^{K} E_{p,k} x^{p+1,k}$.

Step 6. *If $x^{p+1} = x^p$, then stop. Otherwise, set $p := p + 1$ and return to Step 2.*

For each $k \in \{1, 2, \cdots, K\}$ and each $p \in N_0$, we introduce the affine operator $F_{p,k} : R^n \to R^n$ as follows:

$$F_{p,k}(x) = B_{p,k}^{-1} C_{p,k} x + B_{p,k}^{-1} b, \qquad \forall x \in R^n.$$

Furthermore, if for a nonnegative integer μ we define

$$F_{p,k}^{\mu} = \begin{cases} \overbrace{F_{p,k} \circ F_{p,k} \circ \cdots \circ F_{p,k}}^{\mu \text{ times}}, & \text{when } \mu > 0, \\ I, & \text{when } \mu = 0, \end{cases}$$

where μ is the number of compositions of $F_{p,k}$ with itself, then Method 2.1 can be rewritten as the following concise form:

$$x^{p+1} = \sum_{k=1}^{K} E_{p,k} F_{p,k}^{\mu_{p,k}}(x^p), \qquad p = 0, 1, 2, \cdots. \tag{2.1}$$

Evidently, for the original stationary multisplitting $(B_k, C_k, E_k)(k = 1, 2, \cdots, K)$ of the matrix $A \in L(R^n)$, Method 2.1 becomes the parallel chaotic multisplitting iteration method studied in Bru, Elsner and Neumann[BEN], that is,

Model A: $x^{p+1} = \sum\limits_{k=1}^{K} E_k F_k^{\mu_{p,k}}(x^p)$, $p = 0, 1, 2, \cdots$,

where for each $k \in \{1, 2, \cdots, K\}$, $F_k(x) = B_k^{-1} C_k x + B_k^{-1} b$ $(\forall x \in R^n)$.

In particular, when $\mu_{p,k} \equiv 1$ $(k = 1, 2, \cdots, K, p \in N_0)$, it recovers the parallel matrix multisplitting method (1.2) originated in O'Leary and White[OW]. Otherwise, if $(B_{p,k}, C_{p,k}, E_{p,k})(k = 1, 2, \cdots, K)$ is a dynamic multisplitting of the matrix $A \in L(R^n)$, Method 2.1 introduces a new parallel non-stationary 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 $A \in L(R^n)$ is monotone and the multiple splittings are weak regular, we can prove the following convergence theorem for Method 2.1.

THEOREM 2.1. *Let $A \in L(R^n)$ be a monotone matrix, $(B_{p,k}, C_{p,k}, E_{p,k})(k = 1, 2, \cdots, K)$, $p \in N_0$, be a sequence of multi-splittings of the matrix A.*
Assume that the weighting matrices $E_{p,k} \geq 0(k = 1, 2, \cdots, K, p \in N_0)$,
all the splittings $A = B_{p,k} - C_{p,k}(k = 1, 2, \cdots, K, p \in N_0)$ are weak regular,
there exist monotone matrices $\overline{B}_k(k = 1, 2, \cdots, K)$ such that $B_{p,k}^{-1} \geq \overline{B}_k^{-1}(k = 1, 2, \cdots, K, p \in N_0)$.
Then for any initial vector $x^0 \in R^n$, the sequence $\{x^p\}_{p \in N_0}$ generated by Method 2.1 converges to the unique solution $x^* \in R^n$ of the system of linear equations (1.1) whenever $\mu_{p,k} \geq 1(k = 1, 2, \cdots, K, p \in N_0)$.

We remark that the assumption $\mu_{p,k} \geq 1(k = 1, 2, \cdots, K, p \in N_0)$ can be weakened as follows: $\mu_{p,k} \geq 0(k = 1, 2, \cdots, K, p \in N_0)$ and for infinitely many p's, $\mu_{p,k} \geq 1$, for all $k = 1, 2, \cdots, K$.

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.

More generally, we can prove the convergence of Method 2.1 for the H-matrix class.

THEOREM 2.2. *Let* $A \in L(R^n)$ *be an H-matrix*
$(B_{p,k}, C_{p,k}, E_{p,k})(k = 1, 2, \cdots, K)$, $p \in N_0$, *be a sequence of multi-splittings of the matrix* A.
Assume that the weighting matrices $E_{p,k} \geq 0(k = 1, 2, \cdots, K, p \in N_0)$,
all the splittings $A = B_{p,k} - C_{p,k}(k = 1, 2, \cdots, K, p \in N_0)$ *satisfy*

$$\begin{cases} diag(B_{p,k}) = diag(A), \\ \langle A \rangle = \langle B_{p,k} \rangle - |C_{p,k}|, \end{cases} \quad k = 1, 2, \cdots, K, \quad p = 0, 1, 2, \cdots.$$

Then for any initial vector $x^0 \in R^n$, *the sequence* $\{x^p\}_{p \in N_0}$ *generated by Method 2.1 converges to the unique solution* $x^* \in R^n$ *of the system of linear equations (1.1) provided* $\mu_{p,k} \geq 1(k = 1, 2, \cdots, K, p \in N_0)$ *holds.*

We point out that the same remark about Theorem 2.1 is valid for Theorem 2.2.

§3. Convergence theory for the positive definite matrix

Let $\{A_p\}_{p\in N_0}$ be a sequence of matrices in $L(R^n)$. 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 $x^T A_p x \geq c x^T x$ holds for all $x \in R^n$.

When the coefficient matrix $A \in L(R^n)$ is symmetric positive definite, we have the following convergence theorem for Method 2.1.

THEOREM 3.1. *Let $A \in L(R^n)$ be a symmetric positive definite matrix,*

for every $p \in N_0$, $(B_{p,k}, C_{p,k}, E_{p,k})(k = 1, 2, \cdots, K)$ be a multisplitting of the matrix A such that

(a) $B_{p,k} + C_{p,k}(k = 1, 2, \cdots, K)$ are positive definite uniformly in p; and

(b) $f(\sum_{k=1}^{K} E_{p,k} x^{p+1,k}) \leq \max_{1\leq k\leq K} f(x^{p+1,k})$, where $f(x) = \frac{1}{2}x^T Ax - x^T b$.

Let $\mu_{p,k}(k = 1, 2, \cdots, K, p \in N_0)$ be positive integers bounded uniformly from above.

Then the sequence $\{x^p\}_{p\in N_0}$ generated by Method 2.1 converges, independently of the positive integer sequences $\{\mu_{p,k}\}_{p\in N_0}$ $(k = 1, 2, \cdots, K)$, to the unique solution $x^* \in R^n$ of the system of linear equations (1.1).

PROOF: Through straightforward deduction we can obtain the identity

$$f(x) - f(x^*) = \frac{1}{2}(x - x^*)^T A(x - x^*), \qquad \forall x \in R^n. \tag{3.1}$$

Therefore, x^* is the unique solution of the system of linear equations (1.1) 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., $B \in L(R^n)$ is a nonsingular matrix, and define $\overline{x} = B^{-1}Cx + B^{-1}b(\forall x \in R^n)$. Then we can obtain the equality

$$f(x) - f(\overline{x}) = \frac{1}{2}(x - \overline{x})^T(B + C)(x - \overline{x}). \tag{3.2}$$

From (3.2) and Method 2.1 we have

$$\begin{aligned} & f(x^{p,k,\mu-1}) - f(x^{p,k,\mu}) \\ & \quad = \frac{1}{2}(x^{p,k,\mu-1} - x^{p,k,\mu})^T(B_{p,k} + C_{p,k})(x^{p,k,\mu-1} - x^{p,k,\mu}), \\ & \qquad \mu = 1, 2, \cdots, \mu_{p,k}, \quad p \in N_0. \end{aligned}$$

In accordance with assumption (a) there exists a positive constant c, independent of p and k, such that

$$x^T(B_{p,k}+C_{p,k})x \geq cx^Tx, \qquad k = 1, 2, \cdots, K, \quad p \in N_0, \quad \forall x \in R^n.$$

Therefore,

$$f(x^p) - f(x^{p,k,\mu_{p,k}}) \geq \frac{c}{2}\sum_{\mu=1}^{\mu_{p,k}} \|x^{p,k,\mu-1} - x^{p,k,\mu}\|_2^2. \tag{3.3}$$

In addition, by further noticing that $\mu_{p,k}(k = 1, 2, \cdots, K, p \in N_0)$ are uniformly bounded from above by a positive integer, say J, we can get

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

Substituting this estimate into (3.3) yields

$$f(x^p)-f(x^{p+1,k}) \geq \frac{c}{2J}\|x^p-x^{p+1,k}\|_2^2, \qquad k = 1, 2, \cdots, K, \quad p \in N_0. \tag{3.4}$$

From assumption (b) and (3.4) we know that

$$\begin{aligned} f(x^p) - f(x^{p+1}) &= f(x^p) - f(\sum_{k=1}^{K} E_{p,k} x^{p+1,k}) \\ &\geq f(x^p) - \max_{1\leq k\leq K} f(x^{p+1,k}) \\ &= f(x^p) - f(x^{p+1,k_{p+1}}) \\ &\geq \frac{c}{2J}\|x^p - x^{p+1,k_{p+1}}\|_2^2 \end{aligned} \tag{3.5}$$

holds for all $p \in N_0$, where k_{p+1} is an index such that $f(x^{p+1,k_{p+1}}) = \max_{1\leq k\leq K} f(x^{p+1,k})$.

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 $\|x^{p_\ell}\|_2 \to \infty$ and $\left\{\frac{x^{p_\ell}}{\|x^{p_\ell}\|_2}\right\}_{\ell\in N_0} \to \widetilde{x}$, as $\ell \to \infty$, but $\widetilde{x}^T A \widetilde{x} = 0$. This obviously contradicts the symmetric positive definiteness of the matrix $A \in L(R^n)$. Therefore, the sequence $\{x^p\}_{p\in N_0}$ must be bounded.

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 (1.1). 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 (1.1). Let $\widehat{x}$ 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 $\widehat{x}$. Since $\{f(x^{p_\ell})\}_{\ell\in N_0}$ converges to $f(\widehat{x})$ as $\ell \to \infty$ and $\{f(x^p)\}_{p\in N_0}$ is nonincreasing by (3.5), the entire sequence $\{f(x^p)\}_{p\in N_0}$ converges to $f(\widehat{x})$, too. Let the positive integer $k_{p+1} \in \{1, 2, \cdots, K\}$ be defined as in (3.5). Then by taking a further subsequence if necessary, we may assume that there exists some index $\hat{k} \in \{1, 2, \cdots, K\}$ 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 (3.5), and the sequence $\{x^{p_\ell+1,\hat{k}}\}_{\ell\in N_0}$ converges to $\widehat{x}$ as $\ell \to \infty$. From (3.3) it further holds that for any $\mu \in \{1, 2, \cdots, \mu_{p_\ell,\hat{k}}\}$, the sequence $\{x^{p_\ell,\hat{k},\mu}\}_{\ell\in N_0}$ converges to $\widehat{x}$ 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 $\widehat{x}$ solves the system of linear equations (1.1).

Up to now, the proof of this theorem is fulfilled.

We remark that

- Theorem 3.1 can be straightforwardly generalized to the complex matrix case.

• 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) in Theorem 3.1 can be satisfied by various choices of the weighting matrices $E_{p,k}(k = 1, 2, \cdots, K, p \in N_0)$.

• One of the possibilities is given by the choices of $E_{p,k} = \alpha_{p,k} I (k = 1, 2, \cdots, K, p \in N_0)$, where $\alpha_{p,k}(k = 1, 2, \cdots, K, p \in N_0)$ are nonnegative real numbers satisfying $\sum\limits_{k=1}^{K} \alpha_{p,k} = 1 (p \in N_0)$.

• In this case, condition (b) is automatically satisfied if either of the following three classes of restrictions are further imposed:

(1) for a positive integer sequence $\{k_p\} \in \{1, 2, \cdots, K\}$,

$$\alpha_{p,k} = \begin{cases} 1, & \text{for } k = k_p, \\ 0, & \text{for } k \neq k_p, \end{cases} \qquad k = 1, 2, \cdots, K, \quad p \in N_0,$$

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 $f(x^{p,k})(k = 1, 2, \cdots, K)$ such that, for $p \in N_0$, $f(x^{p,k_p}) = \min_{1 \le k \le K} f(x^{p,k})$;

(2) $\alpha_{p,k}(k = 1, 2, \cdots, K, p \in N_0)$ are the minimizers of the functions $g(\alpha_{p,1}, \alpha_{p,2}, \cdots, \alpha_{p,K}) = f(\sum_{k=1}^{K} \alpha_{p,k} x^{p,k})$, $p \in N_0$; and

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

• Moreover, In Theorem 3.1 we does not make the hypothesis that the weighting matrices $E_{p,k}(k = 1, 2, \cdots, K, p \in N_0)$ 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 $E_{p,k}(k = 1, 2, \cdots, K, p \in N_0)$ 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: $A = I$, $b = 0$. Evidently, $x^* = 0$ is the unique solution of the system of linear equations. For simplicity, we take $K = 2$ and $B_k = diag(\frac{1}{1-\sigma_k}, \frac{1}{1-\delta_k})$, $k = 1, 2$, where $\sigma_k, \delta_k \in R^1 \setminus \{1\}$, $k = 1, 2$. Then we get two splittings $A = B_k - C_k$, $k = 1, 2$, where $C_k = diag(\frac{\sigma_k}{1-\sigma_k}, \frac{\delta_k}{1-\delta_k})$, $k = 1, 2$. By direct computations, we have

$$\begin{cases} H_k = B_k^{-1} C_k = diag(\sigma_k, \delta_k), \\ Q_k = B_k + C_k = diag(\dfrac{1+\sigma_k}{1-\sigma_k}, \dfrac{1+\delta_k}{1-\delta_k}), \end{cases} \qquad k = 1, 2$$

and

$$x^{p+1,k} = H_k^{\mu_{p,k}} x^p = (\sigma_k^{\mu_{p,k}} [x^p]_1, \delta_k^{\mu_{p,k}} [x^p]_2)^T, \qquad k = 1, 2, \quad p \in N_0,$$

where $x^p = ([x^p]_1, [x^p]_2)^T$. Therefore,

$$\begin{aligned} f(x^{p+1,k}) &= \frac{1}{2} (x^{p+1,k})^T x^{p+1,k} \\ &= \frac{1}{2} (|\sigma_k|^{2\mu_{p,k}} ([x^p]_1)^2 + |\delta_k|^{2\mu_{p,k}} ([x^p]_2)^2), \\ & \quad k = 1, 2, \quad p \in N_0. \end{aligned} \tag{3.7}$$

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

(i) $E_{p,1} = diag(1, 0)$, $E_{p,2} = diag(0, 1)$, $p \in N_0$. Evidently, we have

$$\begin{aligned} x^{p+1} &= E_{p,1} x^{p+1,1} + E_{p,2} x^{p+1,2} \\ &= (\sigma_1^{\mu_{p,1}} [x^p]_1, \delta_2^{\mu_{p,2}} [x^p]_2)^T, \qquad p \in N_0, \end{aligned} \tag{3.8}$$

and hence,

$$\begin{aligned}f(x^{p+1}) &= \frac{1}{2}(x^{p+1})^T x^{p+1}\\ &= \frac{1}{2}(|\sigma_1|^{2\mu_{p,1}}([x^p]_1)^2 + |\delta_2|^{2\mu_{p,2}}([x^p]_2)^2), \qquad p \in N_0.\end{aligned}$$

If we let $\sigma_k, \delta_k (k = 1, 2)$ be such that $|\sigma_1| > |\sigma_2|$ and $|\delta_2| > |\delta_1|$, and $\mu_{p,k} (k = 1, 2, p \in N_0)$ be such that $\mu_{p,1} = \mu_{p,2} = \mu_p (p \in N_0)$, then it holds that

$$\begin{aligned}f(x^{p+1,1}) &= \frac{1}{2}(|\sigma_1|^{2\mu_p}([x^p]_1)^2 + |\delta_1|^{2\mu_p}([x^p]_2)^2)\\ &< \frac{1}{2}(|\sigma_1|^{2\mu_p}([x^p]_1)^2 + |\delta_2|^{2\mu_p}([x^p]_2)^2) = f(x^{p+1}),\\ f(x^{p+1,2}) &= \frac{1}{2}(|\sigma_2|^{2\mu_p}([x^p]_1)^2 + |\delta_2|^{2\mu_p}([x^p]_2)^2)\\ &< \frac{1}{2}(|\sigma_1|^{2\mu_p}([x^p]_1)^2 + |\delta_2|^{2\mu_p}([x^p]_2)^2) = f(x^{p+1}),\end{aligned}$$

provided $[x^p]_1 \neq 0$ and $[x^p]_2 \neq 0$. This implies that $f(x^{p+1}) > \max_{1 \le k \le 2} f(x^{p+1,k})$, and hence, assumption (b) of Theorem 3.1 is not satisfied.

However, if we let $\sigma_k, \delta_k (k = 1, 2)$ be such that $|\sigma_1| \le |\sigma_2|$ or $|\delta_2| \le |\delta_1|$, and $\mu_{p,k} (k = 1, 2, p \in N_0)$ be such that $\mu_{p,1} = \mu_{p,2} = \mu_p (p \in N_0)$, then it holds that $f(x^{p+1}) \le \max_{1 \le k \le 2} f(x^{p+1,k})$, and hence, assumption (b) of Theorem 3.1 is satisfied.

(i_1) If we further let $\sigma_k, \delta_k \in (-1, 1)$, $k = 1, 2$, then $Q_k (k = 1, 2)$ are symmetric positive definite matrices. Therefore, assumption (a) of Theorem 3.1 is satisfied. Moreover, from (3.8) 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, from (3.8) 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, from (3.8) 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) $E_{p,1} = diag(-\frac{3}{4}, \frac{3}{4})$, $E_{p,2} = diag(\frac{7}{4}, \frac{1}{4})$, $p \in N_0$. Evidently, we have

$$x^{p+1} = (\frac{1}{4}[-3\sigma_1^{\mu_{p,1}}+7\sigma_2^{\mu_{p,2}}][x^p]_1, \frac{1}{4}[3\delta_1^{\mu_{p,1}}+\delta_2^{\mu_{p,2}}][x^p]_2)^T, \qquad p \in N_0,$$

and hence,

$$f(x^{p+1}) = \frac{1}{32}([-3\sigma_1^{\mu_{p,1}}+7\sigma_2^{\mu_{p,2}}]^2([x^p]_1)^2+[3\delta_1^{\mu_{p,1}}+\delta_2^{\mu_{p,2}}]^2([x^p]_2)^2), \quad p \in N_0.$$

If we let $\sigma_1 = \frac{2}{3}$, $\sigma_2 = \frac{1}{3}$, $\delta_1 = \delta_2 = \frac{1}{2}$ and $\mu_{p,k} = 1 (k = 1, 2, p \in N_0)$, then it holds that

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

and

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

It follows that $f(x^{p+1}) \leq \max_{1\leq k\leq 2} f(x^{p+1,k})$, 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

$$x^{p+1} = (\frac{1}{12}[x^p]_1, \frac{1}{2}[x^p]_2)^T, \qquad p \in N_0,$$

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

In the previous discussions, we have permitted the multiple splittings and the composition numbers to depend not only on the processor, but also on the index of the present global step of iteration, to allow for more generality.

In practice we may expect the multiple splittings and the local iteration numbers which each processor performs between two major steps of the method to be fixed and to depend only upon $A \in L(R^n)$ and the relative amount of workloads which are involved in computing the vectors $x^{p,k}$ for $k = 1, 2, \cdots, K$.

Hence, it is reasonable for us to consider only 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 $\mu_{p,k} = \mu_k$, $k = 1, 2, \cdots, K$. Analogous to (2.1), the method just mentioned can be expressed as

$$x^{p+1} = \sum_{k=1}^{K} E_k F_k^{\mu_k}(x^p), \qquad k = 1, 2, \cdots, K, \tag{4.1}$$

where $F_k(x) = B_k^{-1} C_k x + B_k^{-1} b (k = 1, 2, \cdots, K)$.

In the following, we will discuss the monotone convergence properties of the parallel non-stationary multisplitting iteration method (4.1) and investigate the influence of the multiple splittings and the composition numbers upon the convergence behaviour of this method.

For this purpose, we introduce matrices

$$
\begin{aligned}
R &= \sum_{k=1}^{K} E_k \sum_{\mu=0}^{\mu_k - 1} \left(B_k^{-1} C_k\right)^{\mu_k} B_k^{-1}, \\
H &= \sum_{k=1}^{K} E_k \left(B_k^{-1} C_k\right)^{\mu_k} .
\end{aligned} \tag{4.2}
$$

Evidently, it holds that

$$
H = I - RA \tag{4.3}
$$

and (4.1) can be equivalently written as

$$
x^{p+1} = Hx^p + Rb, \qquad p = 0, 1, 2, \cdots . \tag{4.4}
$$

Based upon (4.2) and (4.4), we can straightforwardly obtain the following two-sided monotone approximation properties of the iteration (4.1).

THEOREM 4.1. *Let $A \in L(R^n)$ be a monotone matrix, $(B_k, C_k, E_k)(k = 1, 2, \cdots, K)$ be its multisplitting with $A = B_k - C_k(k = 1, 2, \cdots, K)$ being weak regular splittings, $E_k \geq 0(k = 1, 2, \cdots, K)$. Assume that x^0 and y^0 are initial vectors obeying $x^0 \leq y^0$ and $Ax^0 \leq b \leq Ay^0$, $\{x^p\}_{p \in N_0}$ and $\{y^p\}_{p \in N_0}$ are sequences starting from x^0 and y^0, respectively, and generated by (4.4). Then*

(1) $x^p \leq x^{p+1} \leq y^{p+1} \leq y^p$, $p \in N_0$;

(2) $\lim_{p \to \infty} x^p = x^* = \lim_{p \to \infty} y^p$, where $x^* \in R^n$ is the unique solution of the system of linear equations (1.1); and

(3) for any $z^0 \in R^n$ obeying $x^0 \leq z^0 \leq y^0$, the sequence $\{z^p\}_{p \in N_0}$ starting from z^0 and generated by (4.4) satisfies $x^p \leq z^p \leq y^p$ $(\forall p \in N_0)$. 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 $R^{-1}H \geq 0$, then it holds that*

$$Ax^p \leq b \leq Ay^p, \qquad p \in N_0, \tag{4.5}$$

where $\{x^p\}$ and $\{y^p\}$ are sequences generated by (4.4) starting from x^0 and y^0, respectively.

With Theorem 4.1 and Theorem 4.2, we can further compare the convergence rates of the parallel non-stationary multisplitting iteration methods, resulted from different multiple splittings

$$A = B_k^{(m)} - C_k^{(m)}, \qquad k = 1, 2, \cdots, K, \quad m = 1, 2,$$

and different composition numbers $\mu_k^{(m)} (k = 1, 2, \cdots, K, m = 1, 2)$, for solving the system of linear equations (1.1) in the sense of monotonicity. To this end, corresponding to (4.2) we construct matrices

$$\begin{cases} R^{(m)} = \sum_{k=1}^{K} E_k \sum_{\mu=0}^{\mu_k^{(m)}-1} \left(B_k^{(m)^{-1}} C_k^{(m)}\right)^{\mu} B_k^{(m)^{-1}}, \\ H^{(m)} = \sum_{k=1}^{K} E_k \left(B_k^{(m)^{-1}} C_k^{(m)}\right)^{\mu_k^{(m)}}, \end{cases} \qquad m = 1, 2. \tag{4.6}$$

Analogous to (4.3), we have

$$H^{(m)} = I - R^{(m)}A, \qquad m = 1, 2. \tag{4.7}$$

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

$$\begin{cases} x^{p+1} = H^{(1)}x^p + R^{(1)}b, \\ y^{p+1} = H^{(2)}y^p + R^{(2)}b, \end{cases} \qquad p = 0, 1, 2, \cdots \tag{4.8}$$

according to (4.4).

Theorem 4.3. *Let $A \in L(R^n)$ be a monotone matrix $(B_k^{(m)}, C_k^{(m)}, E_k)(k = 1, 2, \cdots, K)$, $m = 1, 2$, be its two multisplittings with $A = B_k^{(m)} - C_k^{(m)}(k = 1, 2, \cdots, K, m = 1, 2)$ being weak regular splittings,*

$E_k \geq 0(k = 1, 2, \cdots, K)$.

Assume that $x^0 = y^0$ is an initial vector, and $\{x^p\}$ and $\{y^p\}$ are sequences defined by (4.8).

If either $R^{(1)^{-1}} H^{(1)} \geq 0$ or $R^{(2)^{-1}} H^{(2)} \geq 0$ holds, then we have

(a) $x^p \geq y^p (p = 0, 1, 2, \cdots)$, *as* $Ax^0 \leq b$;

(b) $x^p \leq y^p (p = 0, 1, 2, \cdots)$, *as* $Ax^0 \geq b$,

provided $\mu_k^{(1)} \geq \mu_k^{(2)}(k = 1, 2, \cdots, K)$ and

$$\begin{cases} \left(B_k^{(1)^{-1}} C_k^{(1)}\right)^{\mu} B_k^{(1)^{-1}} \geq \left(B_k^{(2)^{-1}} C_k^{(2)}\right)^{\mu} B_k^{(2)^{-1}}, \\ \mu = 0, 1, 2, \cdots, \mu_k^{(2)}; \quad k = 1, 2, \cdots, K. \end{cases} \tag{4.9}$$

In particular, from

Zhong-Zhi Bai and De-Ren 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*

we see that (4.9) holds in either of the following two cases:

(i) $B_k^{(1)^{-1}} \geq B_k^{(2)^{-1}}$ and $C_k^{(1)} B_k^{(1)^{-1}} \geq 0$, $k = 1, 2, \cdots, K$;

(ii) $B_k^{(1)^{-1}} \geq B_k^{(2)^{-1}}$ and $C_k^{(2)} B_k^{(2)^{-1}} \geq 0$, $k = 1, 2, \cdots, K$.

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

THEOREM 4.4. *Let $A \in L(R^n)$ be a monotone matrix, $(B_k, C_k, E_k)(k = 1, 2, \cdots, K)$ be its multisplitting with $A = B_k - C_k (k = 1, 2, \cdots, K)$ being weak regular splittings, $E_k \geq 0 (k = 1, 2, \cdots, K)$, $(\underline{B}, \underline{C})$ and $(\overline{B}, \overline{C})$ be its single splittings with $A = \underline{B} - \underline{C}$ and $A = \overline{B} - \overline{C}$ being regular splittings, respectively. Assume that $\underline{x}^0 = x^0 = \overline{x}^0$ is an initial vector, $\{x^p\}_{p \in N_0}$ is the sequence defined by (4.1), and $\{\underline{x}^p\}_{p \in N_0}$ and $\{\overline{x}^p\}_{p \in N_0}$ are sequences defined by*

$$\begin{cases} \underline{x}^{p+1} = \left(\underline{B}^{-1}\underline{C}\right)^{\mu_{\min}} \underline{x}^p + \sum\limits_{\mu=0}^{\mu_{\min}-1} \left(\underline{B}^{-1}\underline{C}\right)^{\mu} \underline{B}^{-1} b, \\ \overline{x}^{p+1} = \left(\overline{B}^{-1}\overline{C}\right)^{\mu_{\min}} \overline{x}^p + \sum\limits_{\mu=0}^{\mu_{\min}-1} \left(\overline{B}^{-1}\overline{C}\right)^{\mu} \overline{B}^{-1} b, \end{cases}$$

respectively, where $\mu_{\min}$ is a positive integer satisfying

$$\mu_{\min} \leq \min_{1 \leq k \leq K} \{\mu_k\}.$$

Then we have

(a) $\underline{x}^p \leq x^p \leq \overline{x}^p (p = 0, 1, 2, \cdots)$, *as* $Ax^0 \leq b$;

(b) $\underline{x}^p \geq x^p \geq \overline{x}^p (p = 0, 1, 2, \cdots)$, *as* $Ax^0 \geq b$,

provided $\overline{B}^{-1} \geq B_k^{-1} \geq \underline{B}^{-1}$, $k = 1, 2, \cdots, K$.