Adv Comput Math
DOI 10.1007/s10444-011-9262-8

Modified parallel multisplitting iterative methods for non-Hermitian positive definite systems

Chuan-Long Wang · Guo-Yan Meng · Xue-Rong Yong

Received: 1 September 2011 / Accepted: 14 November 2011

Abstract In this paper we present three modified parallel multisplitting iterative methods for solving non-Hermitian positive definite systems $Ax = b$. The first is a direct generalization of the standard parallel multisplitting iterative method for solving this class of systems. The other two are the iterative methods obtained by optimizing the weighting matrices based on the sparsity of the coefficient matrix A. In our multisplitting there is only one that is required to be convergent (in a standard method all the splittings must be convergent), which not only decreases the difficulty of constructing the multisplitting of the coefficient matrix A, but also releases the constraints to the weighting matrices (unlike the standard methods, they are not necessarily be known or given in advance). We then prove the convergence and derive the convergent rates of the algorithms by making use of the standard quadratic optimization technique. Finally, our numerical computations indicate that the methods derived are feasible and efficient.

Communicated by Charles A. Micchelli.

This work is supported by NSF of China (11071184) and NSF of Shanxi Province (2010011006).

C.-L. Wang (✉)
Department of Mathematics, Taiyuan Normal University, Taiyuan 030012, Shanxi Province, People's Republic of China
e-mail: clwang218@126.com

G.-Y. Meng
Department of Computer Science, Xinzhou Normal University, Xinzhou 034000, Shanxi Province, People's Republic of China

X.-R. Yong
Department of Mathematical Sciences, University of Puerto Rico at Mayaguez, Mayaguez, PR 00681-9018, USA

Published online: 13 December 2011

Keywords Optimal weighting matrices · Parallel multisplitting iterative method · Non-Hermitian matrix · Convergence

Mathematics Subject Classifications (2010) 65F10 · 15A06

1 Introduction and preliminaries

In solving a large sparse linear system of equations

$$Ax = b, \tag{1.1}$$

where $A = (a_{ij}) \in C^{n\times n}$ is nonsingular, and $b \in C^n$, O'Leary and White [19] seem to be the first to introduce the parallel algorithms by multisplitting A and derive the convergence properties. Their formulas can be written as

$$\begin{aligned} A &= M_i - N_i, && i = 1, 2, \cdots, m, \\ M_i x_i^{(k)} &= N_i x^{(k-1)} + b, && k = 1, 2, \cdots. \\ x^{(k)} &= \sum_{i=1}^{m} E_i x_i^{(k)}, \end{aligned}$$

where E_i, called the *weighting matrices*, are nonnegative diagonal matrices that satisfy the equation $\sum_{i=1}^{m} E_i = I$. The triples $(M_i, N_i, E_i)_{i=1}^m$ is said to be a multisplitting (m splittings) of A. Subsequently, many authors studied the methods for the cases that A is an M-matrix [16], an H-matrix [1–5, 8] and an Hermitian positive definite matrix (e.g., [9–12, 15, 21, 22]), respectively. However, we noticed that all the above parallel multisplititng iterative methods are based on the following constraints to A:

- all splittings $A = M_i - N_i$, $i = 1, 2, \cdots, m$, must be convergent;
- the weighting matrices, $E_i^{(k)}$ $(i = 1, 2, \ldots, m;\ k = 1, 2, \cdots)$ (e.g. [2, 11, 14]) must be given in advance.

When the coefficient matrix A of the system (1.1) is non-Hermitian positive definite, the first author of this paper and Bai [18] presented some sufficient conditions that guarantee the convergence of the single ($m = 1$) iterative methods and we would like to address that Bai et al. [6, 7] also discussed two alternative methods, called HSS and PSS methods, which converge unconditionally to the unique solution of the system of equations (1.1). But a common drawback of those methods is that, if A is Hermitian or skew-Hermitian, the corresponding equations must be solved at each iteration step. The research into a skew-Hermitian system of linear equations is also conducted in [13, 19, 20]). But until now we have not seen an article that discusses the convergence of a parallel multisplitting iterative algorithm for the non-Hermitian positive definite systems of linear equations. This motivated us to come up with the algorithms proposed here.

The aim of this paper is to investigate the convergent parallel multisplitting iterative algorithms for the non-Hermitian positive definite systems of linear equations. By making use of the standard quadratic optimization technique we choose the optimal weighting matrices at each step. And in computation, we need only one of the splittings $A = M_i - N_i, i = 1, 2, \cdots, m$, to be convergent and all the others can be constructed arbitrarily. Thus, we not only decrease the difficulty of constructing the multisplitting of A, but also release the constraints to the weighting matrices (they are not necessarily be known or given in advance). The proposed three parallel multisplitting iterative methods are as follows.

- a direct generalization of the traditional parallel multisplitting iterative methods for solving non-Hermitian positive-definite systems;
- a method based on combining a special multisplitting that has 'conjugate' property and the sparsity of the matrix A;
- a parallel multisplitting iterative method with optimal weighting matrices.

It is convenient to introduce some essential notations and preliminaries. As usual, we use $C^{n\times n}$ to denote the $n \times n$ complex matrix set and C^n the n-dimensional complex vector space. X^* represents the conjugate transpose of a matrix or a vector X and $\langle x, y\rangle$ stands for the angle between the vectors x and y. We use $||\cdot||_2$ to denote the Euclidean norm. If A is an $n \times n$ Hermitian positive definite (or semi-definite) matrix then it is written as $A \succ 0$ (or $\succeq 0$). A matrix $A \in C^{n\times n}$ is called positive definite, if for all nonzero $x \in C^n$, $Re(x^*Ax) > 0$ is always true. For a complex or real matrix A we let

$$H(A) = \frac{1}{2}(A + A^*), \qquad S(A) = \frac{1}{2}(A - A^*). \tag{1.2}$$

The width l of a sparse matrix A is defined as

$$l = \max\{|j - i|, a_{ij} \neq 0\}. \tag{1.3}$$

For a large sparse matrix A, we always assume that $l \ll n$. Also, $[\frac{n}{m}]$ represents the integer part of the number $\frac{n}{m}$. For the residual vectors $r^{(k)} = Ax^{(k)} - b$, $k = 1, 2, \cdots$, if they satisfy

$$\left\|r^{(k)}\right\|_2 \leq \left\|M^{-1}N\right\|_2^k \|r_0\|_2 , \tag{1.4}$$

then the convergent rate q is defined as

$$q = -\ln\left(\|M^{-1}N\|_2\right). \tag{1.5}$$

In the following Section 2 we describe three different parallel multisplitting iterative algorithms and then in Section 3 we provide the proofs of their convergence. Finally, We apply our algorithms to two concrete examples and then illustrate the advantages of the algorithms.

2 Parallel algorithms

In this section we present three parallel multisplitting iterative algorithms. The convergence theorems for the algorithms will be established in the next section. The first is a direct generation of the standard parallel multisplitting iterative algorithm. The next two are the parallel multisplitting iterative algorithms obtained by optimizing the weighting matrices.

(I) Let the multisplitting of A be given by [9, 10]

$$A = B_i - C_i, \quad i = 1, 2, \cdots, m, \tag{2.1}$$

where B_i $(i = 1, 2, \cdots, m)$ are Hermitian diagonal block matrices, that is

$$B_i = diag\left(B_{i,1}, B_{i,2}, \cdots, B_{i,m}\right), \tag{2.2}$$

where $B_{i,1}, B_{i,2}, \cdots, B_{i,m}$ are Hermitian, and the weighting matrices E_i satisfy

$$\sum_{i=1}^{m} E_i = I, \; E_i = diag\left(\alpha_{i1} I_1, \alpha_{i2} I_2, \cdots, \alpha_{im} I_m\right) \geq 0, \; i = 1, 2, \cdots, m. \tag{2.3}$$

Algorithm 2.1

Give an initial point $x^{(0)}$ and a tolerance $\epsilon > 0$, for $k = 1, 2, \cdots$ until the process converges, do

Step 1. Solve in parallel the m equations

$$M_i x_i^{(k)} = N_i x^{(k-1)} + b, \; i = 1, 2, \cdots, m.$$

Step 2. Compute $x^{(k)}$ by the formula

$$x^{(k)} = \sum_{i=1}^{m} E_i x_i^{(k)}.$$

Step 3. If $\| Ax^{(k)} - b \|_2 < \epsilon$, stop; Otherwise, $k \Leftarrow k + 1$ and go back to Step 1.

(II) Let the multisplitting of A be given by

$$A = M_i - N_i, \qquad i = 1, 2, \cdots, m, \tag{2.4}$$

where M_1 is a Hermitian positive definite matrix, and let

$$M_i^{-1} = M_1^{-1} - P_i, \qquad i = 2, \cdots, m, \tag{2.5}$$

$$P_i = \left(0, \cdots, \left(P^{(i)}_{[\frac{n}{m}]\times n}\right)^*, \cdots, 0\right)^*,$$

$$P^{(i)}_{[\frac{n}{m}]\times n} = \begin{pmatrix} * & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots \\ * & * & \cdots & * \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}_{[\frac{n}{m}]\times n}, \tag{2.6}$$

where, the nonzero entries can be from row $i[\frac{n}{m}]+1$ to row $(i+1)[\frac{n}{m}]-l-1$ and where the zero entries are from row $(i+1)[\frac{n}{m}]-l$ to row $(i+1)[\frac{n}{m}]$. Similarly, the weighting matrices satisfy

$$\sum_{i=1}^{m} E_i^{(k)} = I, \qquad E_i^{(k)} = \alpha_i^{(k)} I, \qquad i = 1, 2, \cdots, m, \quad k = 1, 2, \cdots, \tag{2.7}$$

we now describe our second algorithm.

Algorithm 2.2

Give an initial point $x^{(0)}$ and a tolerance $\epsilon > 0$, let the residual vector $r^{(0)} = Ax^{(0)} - b$. For $k = 1, 2, \cdots$ until the algorithm converges, do

Step 1. Compute in parallel

$$x_i^{(k)} = x^{(k-1)} - M_i^{-1} r^{(k-1)}, \qquad i = 1, 2, \cdots, m,$$

$$r_i^{(k)} = Ax_i^{(k)} - b, \qquad i = 1, 2, \cdots, m.$$

Step 2. Compute α in parallel for $i = 2, \cdots, m$

$$\alpha_i^{(k)} = \frac{-\left(r_i^{(k)} - r_1^{(k)}\right)^* M_1^{-1} r_1^{(k)}}{\left(r_i^{(k)} - r_1^{(k)}\right)^* M_1^{-1} \left(r_i^{(k)} - r_1^{(k)}\right)}, \tag{2.8}$$

Step 3. Compute $x^{(k)}$ and $r^{(k)}$ by the following formulas

$$x^{(k)} = \sum_{i=1}^{m} \alpha_i^{(k)} x_i^{(k)} = x_1^{(k)} + \sum_{i=2}^{m} \alpha_i^{(k)} \left(x_i^{(k)} - x_1^{(k)}\right),$$

$$r^{(k)} = \sum_{i=1}^{m} \alpha_i r_i^{(k)}.$$

Step 4. If $\|r^{(k)}\|_2 < \epsilon$, stop; Otherwise, $k \Leftarrow k+1$ and go back to Step 1.

(III) Let the multisplitting of A be given by

$$A = F_i - G_i, \qquad i = 1, 2, \cdots, m, \tag{2.9}$$

where F_1 is a Hermitian positive definite matrix, and

$$E_i^{(k)} = \alpha_i^{(k)} I, \qquad i = 1, 2, \cdots, m, \quad k = 1, 2, \cdots, \tag{2.10}$$

now we release the constraint $\sum_{i=1}^{m} E_i^{(k)} = I$ and describe our third algorithm below.

Algorithm 2.3

Give an initial point $x^{(0)}$ and a tolerance $\epsilon > 0$, let $r^{(0)} = Ax^{(0)} - b$. For $k = 0, 1, 2, \cdots$ until the process converges, do

Step 1. Compute in parallel

$$F_i x_i^{(k)} = G_i x^{(k-1)} + b, \ i = 1, 2, \cdots, m.$$

Step 2. Compute

$$x^{(k)} = \sum_{i=1}^{m} \alpha_i^{(k)} x_i^{(k)},$$

$$r^{(k)} = Ax^{(k)} - b,$$

where $\alpha^{(k)} = (\alpha_1^{(k)}, \alpha_2^{(k)}, \cdots, \alpha_m^{(k)})$ is the solution to the following quadratic programming

$$\min_{\alpha} \frac{1}{2} r^* M_1^{-1} r, \tag{2.11}$$

$$r = A \left(\sum_{i=1}^{m} \alpha_i x_i^{(k)} \right) - b. \tag{2.12}$$

Step 3. If $\|r^{(k)}\|_2 < \epsilon$, stop; Otherwise, $k \Leftarrow k + 1$ and go back to Step 1.

In fact, if we define $X(k) = (x_1^{(k)}, \cdots, x_m^{(k)})$, then the solution of (2.11) and (2.12) is given by

$$\alpha^{(k)} = \left(X(k)^T A^T M_1^{-1} A X(k) \right)^{-1} X(k)^T A^T M_1^{-1} b. \tag{2.13}$$

Remark We would like to address the releases of the constraints that are required in the algorithms appeared in the previous articles. Algorithm 2.1 is a direct generalization of symmetric splitting discussed in [9]. The advantage of the algorithm is that the weighting matrices do not need to be scaled. In Algorithm 2.2, the splitting has "conjugate" property. By making use of the sparsity of the coefficient matrix A, we obtained more general weighting matrices $E_i^{(k)} = \alpha_i^{(k)} I$, $i = 1, 2, \cdots, m$ (the matrices do not need to be non-negative or stationary). In Algorithm 2.3, The m splittings do not need to be convergent for all i. Instead, we need only one convergent splitting of A in computation. Also, the weighting matrices $E_i^{(k)} = \alpha_i^{(k)} I$, $i = 1, 2, \cdots, m$ do not need to be nonnegative or stationary and we remove the constraint $\sum_{i=1}^{m} E_i = I$ in computation.

3 Analysis of convergence

In this section we discuss the convergence of the multisplitting algorithms addressed in the previous section. Our idea is by combining the sparsity of the matrix A and the property of the solution to a quadratic programming.

Theorem 3.1 *Let $A = M - N$, $\det M \neq 0$ be a splitting and $T = NM^{-1}$. If $H(A) \succ 0$. Then $M \succ \frac{1}{2}(H(A) + S(A)^* H(A)^{-1} S(A))$ if and only if*

$$\left\| M^{-\frac{1}{2}} T M^{\frac{1}{2}} \right\|_2 < 1. \tag{3.1}$$

Proof Note that proving $\| M^{-\frac{1}{2}} T M^{\frac{1}{2}} \|_2 < 1$ is equivalent to showing that

$$M^{-\frac{1}{2}} N^* M^{-1} N M^{-\frac{1}{2}} \prec I,$$

which is also equivalent to claiming that

$$N^* M^{-1} N \prec M. \tag{3.2}$$

From $N = M - A$, we have

$$(M - A)^* M^{-1} (M - A) = M - A^* - A + A^* M^{-1} A.$$

By combining this identity with formula (3.2) we obtain their another equivalent condition:

$$A^* M^{-1} A \prec A^* + A,$$

and it is also equivalent to

$$M^{-1} \prec (A^*)^{-1} (A^* + A) A^{-1} = A^{-1} + (A^*)^{-1}. \tag{3.3}$$

To prove the theorem we rewrite A^{-1} in terms of $H(A)^{-\frac{1}{2}}$ and $S(A)$. Since

$$\begin{aligned} A^{-1} &= (H(A) + S(A))^{-1} \\ &= H(A)^{-\frac{1}{2}} \left(I + \widehat{S}(A)\right)^{-1} H(A)^{-\frac{1}{2}} \\ &= H(A)^{-\frac{1}{2}} \left(I + \widehat{S}(A)\right)^{-1} \left(I - \widehat{S}(A)\right)^{-1} \left(I - \widehat{S}(A)\right) H(A)^{-\frac{1}{2}} \\ &= H(A)^{-\frac{1}{2}} \left(I - \widehat{S}^2(A)\right)^{-1} \left(I - \widehat{S}(A)\right) H(A)^{-\frac{1}{2}} \\ &= H(A)^{-\frac{1}{2}} \left(I - \widehat{S}^2(A)\right)^{-1} H(A)^{-\frac{1}{2}} - H(A)^{-\frac{1}{2}} \left(I - \widehat{S}^2(A)\right)^{-1} \widehat{S}(A) H(A)^{-\frac{1}{2}}, \end{aligned}$$

where $\widehat{S}(A) = H(A)^{-\frac{1}{2}} S(A) H(A)^{-\frac{1}{2}}$, we have that

$$A^{-1} + (A^*)^{-1} = 2H(A)^{-\frac{1}{2}} (I - \widehat{S}^2(A))^{-1} H(A)^{-\frac{1}{2}}.$$

Combining this relation with (3.3) yields

$$M^{-1} \prec 2H(A)^{-\frac{1}{2}} \left(I - \widehat{S}^2(A)\right)^{-1} H(A)^{-\frac{1}{2}},$$

which is equivalent to

$$M \succ \frac{1}{2}(H(A) + S(A)^* H(A)^{-1} S(A)).$$

Observing the assumptions in the theorem yields (3.1). □

Theorem 3.2 *Let $(B_i, C_i, E_i)_{i=1}^m$ be the multisplitting given by the formulas (2.1)–(2.3) and $B = diag(B_{i,1}, B_{i,2}, \cdots, B_{i,m})$ be of the same forms as B_i ($i = 1, 2, \cdots, m$). If*

$$B_i \succeq B \succ \frac{1}{2}(H(A) + S(A)^* H(A)^{-1} S(A)), \qquad H(A) \succ 0,$$

$$\sum_{i=1}^m E_i = I,\ i = 1, 2, \cdots, m, \qquad E_i \geq 0,$$

then the sequence $\{x^{(k)}\}$ generated by Algorithm 2.1 *converges to the solution of* (1.1).

Proof From the proofs given in [9, 17] it is trivially seen that the multisplitting $(B_i, C_i, E_i)_{i=1}^m$ is convergent if and only if the splitting

$$A = \left(\sum_{i=1}^m E_i B_i^{-1}\right)^{-1} - \left(\sum_{i=1}^m E_i B_i^{-1}\right)^{-1} \sum_{i=1}^m E_i B_i^{-1} C_i$$

is convergent. Thus, it suffices to prove

$$\left(\sum_{i=1}^m E_i B_i^{-1}\right)^{-1} \succeq B. \tag{3.4}$$

From (2.2) and (2.3) we have that

$$\begin{aligned} E_i B_i^{-1} &= diag\left(\alpha_{i1} B_{i,1}^{-1}, \alpha_{i2} B_{i,2}^{-1}, \cdots, \alpha_{im} B_{i,m}^{-1}\right) \\ &\preceq diag\left(\alpha_{i1} B_{11}^{-1}, \alpha_{i2} B_{22}^{-1}, \cdots, \alpha_{im} B_{mm}^{-1}\right). \end{aligned}$$

Consequently,

$$\sum_{i=1}^m E_i B_i^{-1} \preceq diag\left(\sum_{i=1}^m \alpha_{i1} B_{11}^{-1}, \sum_{i=1}^m \alpha_{i2} B_{22}^{-1}, \cdots, \sum_{i=1}^m \alpha_{im} B_{mm}^{-1}\right) = B^{-1},$$

which is equivalent to (3.4). Now, from Theorem 3.1 we see that the splitting $A = \left(\sum_{i=1}^m E_i B_i^{-1}\right)^{-1} - \left(\sum_{i=1}^m E_i B_i^{-1}\right)^{-1} \sum_{i=1}^m E_i B_i^{-1} C_i$ is convergent. This proves the theorem. □

Lemma 3.3 *If the widths of A and B are l and d, respectively, then the width of AB is at most $l+d$.*

Proof Let $A=(a_{ij}),\ B=(b_{ij}),\ C=(c_{ij})=AB$. Then

$$c_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj}=a_{ii-l}b_{i-lj}+\cdots+a_{ii+l}b_{i+lj}.$$

If $i-l-j>d$, then $b_{i-lj}=\cdots=b_{i+lj}=0$, so $c_{ij}=0$.
If $i+l<j-d$, then $b_{i-lj}=\cdots=b_{i+lj}=0$, so $c_{ij}=0$.
Hence, if $|i-j|>l+d$, then $c_{ij}=0$. □

Lemma 3.4 ('Conjugate' property) *Let the width of the matrix A be l, and let $l,\ m<<n$, $[\frac{n}{m}]=\frac{n}{m}$. Assume $P_i=(0,\cdots,(P^{(i)}_{[\frac{n}{m}]\times n})^*,\cdots,0)^*$, where $P^{(i)}_{[\frac{n}{m}]\times n}$ is i-th block matrix with nonzero rows being from $i[\frac{n}{m}]+1$ to $(i+1)[\frac{n}{m}]-l-1$. Then*

$$P_i^*AP_j=0,\qquad \forall\, i\neq j. \tag{3.5}$$

Proof Since

$$P_i^*AP_j=\left(0,\cdots,\left(P^{(i)}_{[\frac{n}{m}]\times n}\right)^*,0,\cdots,0\right)A\left(0,\cdots,\left(P^{(j)}_{[\frac{n}{m}]\times n}\right)^*,\cdots,0\right)^*$$

and the nonzero rows of AP_j are from $j[\frac{n}{m}]-l+1$ to $(j+1)[\frac{n}{m}]-1$, the nonzero columns of P_i^* are at most from $i[\frac{n}{m}]+1$ to $(i+1)[\frac{n}{m}]-l-1$. This implies $P_i^*AP_j=0$ for all $i,\ j$ such that $i\neq j$. □

Remark If $[\frac{n}{m}]<\frac{n}{m}$, we can select $P_m=\left(0,\cdots,0,\left(P^{(m)}_{k\times n}\right)^*\right)^*$, where $k=n-(m-1)[\frac{n}{m}]$, then $P^{(m)}_{k\times n}$ has $[\frac{n}{m}]-l-1$ nonzero rows. Hence, for the case that $[\frac{n}{m}]\neq\frac{n}{m}$, Lemma 3.4 is still true.

Theorem 3.5 *Let the multisplitting $(M_i,N_i,E_i)_{i=1}^m$ be given by formulas (2.4)–(2.7). Assume that P_i $(i=1,2,\cdots,m)$ satisfy $P_i^*A^*M_1^{-1}AP_j=0$. If $M_1\succ\frac{1}{2}(H(A)+S(A)^*H(A)^{-1}S(A))$ and $H(A)\succ0$. Then $\{x^{(k)}\}$ generated by Algorithm 2.2 converges to the solution of (1.1). Furthermore, if $\left\langle M_1^{-\frac{1}{2}}r^{(k)},\ M_1^{-\frac{1}{2}}r_1^{(k)}\right\rangle\geq\theta$, then the convergent rate is given by*

$$q=-\left(\ln||M_1^{-\frac{1}{2}}N_1M_1^{-\frac{1}{2}}||_2+\ln|\cos\theta|\right), \tag{3.6}$$

Proof Let $r^{(k)}=Ax^{(k)}-b$. Then

$$x_i^{(k)}=x^{(k-1)}-M_i^{-1}r^{(k-1)},\ i=1,2,\cdots,m, \tag{3.7}$$

which implies

$$r_i^{(k)}=r^{(k-1)}-AM_i^{-1}r^{(k-1)}=\left(I-AM_i^{-1}\right)r^{(k-1)},\ i=1,2,\cdots,m. \tag{3.8}$$

On the other hand, since

$$r = \sum_{i=1}^{m} \alpha_i r_i^{(k)} = r_1^{(k)} + \sum_{i=2}^{m} \alpha_i \left(r_i^{(k)} - r_1^{(k)} \right),$$

we consider the following quadratic programming

$$\begin{aligned} r^* M_1^{-1} r &= \left(r_1^{(k)} + \sum_{i=2}^{m} \alpha_i \left(r_i^{(k)} - r_1^{(k)} \right) \right)^* M_1^{-1} \left(r_1^{(k)} + \sum_{i=2}^{m} \alpha_i \left(r_i^{(k)} - r_1^{(k)} \right) \right) \\ &= \left(r_1^{(k)} \right)^* M_1^{-1} r_1^{(k)} + 2 \sum_{i=2}^{m} \alpha_i \left(r_i^{(k)} - r_1^{(k)} \right)^* M_1^{-1} r_1^{(k)} \\ &\quad + \sum_{i=2}^{m} \alpha_i \left(r_i^{(k)} - r_1^{(k)} \right)^* M_1^{-1} \sum_{i=2}^{m} \alpha_i \left(r_i^{(k)} - r_1^{(k)} \right) \\ &= \left(r_1^{(k)} \right)^* M_1^{-1} r_1^{(k)} + 2 \sum_{i=2}^{m} \alpha_i \left(r_i^{(k)} - r_1^{(k)} \right)^* M_1^{-1} r_1^{(k)} \\ &\quad + \sum_{i,j=2}^{m} \alpha_i \alpha_j \left(r_i^{(k)} - r_1^{(k)} \right)^* M_1^{-1} \left(r_j^{(k)} - r_1^{(k)} \right), \end{aligned} \tag{3.9}$$

from (3.8), it follows

$$\begin{aligned} r_i^{(k)} - r_1^{(k)} &= A \left(M_1^{-1} - M_i^{-1} \right) r^{(k-1)} \\ &= A P_i r^{(k-1)}, \end{aligned}$$

and

$$\left(r_i^{(k)} - r_1^{(k)} \right)^* M_1^{-1} \left(r_j^{(k)} - r_1^{(k)} \right) = \left(r^{(k-1)} \right)^* P_i^* A^* M_1^{-1} A P_j r^{(k-1)} = 0.$$

Hence, (3.9) is reduced to

$$\begin{aligned} r^* M_1^{-1} r &= \left(r_1^{(k)} \right)^* M_1^{-1} r_1^{(k)} + 2 \sum_{i=2}^{m} \alpha_i \left(r_i^{(k)} - r_1^{(k)} \right)^* M_1^{-1} r_1^{(k)} \\ &\quad + \sum_{i=2}^{m} \alpha_i^2 \left(r_i^{(k)} - r_1^{(k)} \right)^* M_1^{-1} \left(r_i^{(k)} - r_1^{(k)} \right). \end{aligned} \tag{3.10}$$

Now, by differentiating $r^* M_1^{-1} r$ with respect to $\alpha_i, i = 1, 2, \cdots, m$, we obtain formula (2.8). That is, the critical point $(\cdots, \alpha_i^{(k)}, \cdots)$ that minimizes the quadratic form $r^* M_1^{-1} r$:

$$\alpha_i^{(k)} = \frac{-\left(r_i^{(k)} - r_1^{(k)} \right)^* M_1^{-1} r_1^{(k)}}{\left(r_i^{(k)} - r_1^{(k)} \right)^* M_1^{-1} \left(r_i^{(k)} - r_1^{(k)} \right)}, \quad i = 2, 3, \cdots, m. \tag{3.11}$$

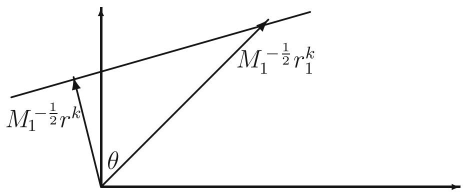

Fig. 1 The orthogonality between $M_1^{-\frac{1}{2}}r^{(k)}$ and $M_1^{-\frac{1}{2}}(r^{(k)}-r_1^{(k)})$

Next, we discuss the convergence and convergent rate of $\{x^{(k)}\}$. For $r^{(k)}$ and $r_1^{(k)}$, (3.11) yields

$$\left(r^{(k)}\right)^* M_1^{-1}\left(r^{(k)}-r_1^{(k)}\right)=0,$$

which implies that the vectors $M_1^{-\frac{1}{2}}r^{(k)}$ and $M_1^{-\frac{1}{2}}(r^{(k)}-r_1^{(k)})$ are orthogonal vectors shown in Fig. 1.

From the assumption $\left\langle M_1^{-\frac{1}{2}}r^{(k)}, M_1^{-\frac{1}{2}}r_1^{(k)}\right\rangle \geq \theta$, we see that

$$\left(r^{(k)}\right)^* M_1^{-1}r^{(k)} \leq \cos^2\theta \left(r_1^{(k)}\right)^* M_1^{-1}r_1^{(k)}. \tag{3.12}$$

Hence,

$$\begin{aligned}\left\|M_1^{-\frac{1}{2}}r^{(k)}\right\|_2 &\leq |\cos\theta| \left\|M_1^{-\frac{1}{2}}r_1^{(k)}\right\|_2\\ &= |\cos\theta| \left\|M_1^{-\frac{1}{2}}N_1M_1^{-1}r^{(k-1)}\right\|_2\\ &\leq |\cos\theta| \left\|M_1^{-\frac{1}{2}}N_1M_1^{-\frac{1}{2}}\right\|_2 \left\|M_1^{-\frac{1}{2}}r^{(k-1)}\right\|_2\\ &\vdots\\ &\leq |\cos^k\theta| \left\|M_1^{-\frac{1}{2}}N_1M_1^{-\frac{1}{2}}\right\|_2^k \left\|M_1^{-\frac{1}{2}}r^{(0)}\right\|_2\end{aligned}$$

From Theorem 3.1, we have that $\|M_1^{-\frac{1}{2}}N_1M_1^{-\frac{1}{2}}\|_2 < 1$. Thus,

$$\lim_{k\longrightarrow\infty}\left\|M_1^{-\frac{1}{2}}r^{(k)}\right\|_2 = 0, \tag{3.13}$$

which implies $\lim\limits_{k\longrightarrow\infty} r^{(k)}=0$. Furthermore, by the definition of the convergent rate (see (1.4) and (1.5)), we obtain (3.6). □

Theorem 3.6 *Let the multisplitting $(F_i, G_i, E_i)_{i=1}^m$ be given by formulas (2.9)–(2.10), and let $F_1 \succ \frac{1}{2}(H(A)+S(A)^*H(A)^{-1}S(A))$. Then the sequence $\{x^{(k)}\}$ that is generated by Algorithm 2.3 converges to the solution of (1.1). Furthermore, if $\langle F_1^{-\frac{1}{2}}r^{(k)}, F_1^{-\frac{1}{2}}r_1^{(k)}\rangle \geq \theta$, then the convergent rate is*

$$q=-\left(\ln\left\|F_1^{-\frac{1}{2}}G_1F_1^{-\frac{1}{2}}\right\|_2+\ln|\cos\theta|\right). \tag{3.14}$$

Proof Let $r^{(k)} = Ax^{(k)} - b$. Then combination of (2.11) and (2.12) implies (3.12). Thus, proceeding a similar derivation results in the following property

$$\lim_{k \longrightarrow \infty} r^{(k)} = 0 \tag{3.15}$$

and (3.14) (Their derivations are omitted), which proves the theorem. □

4 Numerical results

Example [6] Consider the generalized convection-diffusion equations in a two-dimensional case:

$$-\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} + q * exp(x+y) * x * \frac{\partial u}{\partial x} + q * exp(x+y) * y * \frac{du}{dy} = f \tag{4.1}$$

with the homogeneous Dirichlet boundary condition. We use the standard Ritz-Galerkin Finite Element Method and apply the conforming linear triangular elements to approximate its continuous solutions $u = x * y * (1-x) * (1-y)$ in the domain $\Omega = [0, 1] \times [0, 1]$, where the step-sizes along both x and y directions are selected to be the same $h = \frac{1}{128}$.

After discretization the matrix A of this equation is given by

$$A = \begin{bmatrix} A_{11} & B_{12} & & & \\ C_{21} & A_{22} & B_{23} & & \\ & \ddots & \ddots & \ddots & \\ & & C_{p-1p-2} & A_{p-1p-1} & B_{p-1p} \\ & & & C_{pp-1} & A_{pp} \end{bmatrix}_{p \times p},$$

where $A_{ii},\ i = 1, \cdots, p$ are n-by-n nonsymmetric matrices and $B_{ii+1}^T \neq C_{i+1i}$., and $np = 128^2$.

Let $q = 1$. Given $x^{(0)} = (0, \cdots, 0)^T$ and a tolerance $\varepsilon = 10^{-5}$, the iteration fails in computation for the iterative number that is up to 30,000.

Now, let $H(A) = D - L - L^T$, where $D = diag(H(A_{11}), \cdots, H(A_{pp}))$, and L the block strictly lower triangle matrix of $H(A)$. Since A is a sparse matrix with the block width $l = 1$, we choose

$$M_1 = D = diag(H(A_{11}), \cdots, H(A_{pp})).$$

the splittings $(M_i, N_i),\ i = 2, 3$ are constructed by (2.5) and (2.6), where the nonzero part of $P^{(i)}_{[\frac{p}{3}] \times p}$ is the corresponding part of the following matrix

$$\begin{bmatrix} H(A_{11}) & & & \\ & H(A_{22}) & & \\ & & \ddots & \\ & & & H(A_{pp}) \end{bmatrix}^{-1} \begin{bmatrix} 0 & B_{12} & & & \\ C_{21} & 0 & B_{23} & & \\ & \ddots & \ddots & \ddots & \\ & & C_{p-1p-2} & 0 & B_{p-1p} \\ & & & C_{pp-1} & 0 \end{bmatrix}.$$

Our numerical results of Algorithm 2.2, when $p = 64$, are given in Table 1.

When we use the following multisplitting iterative method

(1) $M_1 = (D - L)^T D^{-1}(D - L^T)$, $N_1 = M_1 - A$;
(2) the block SOR method (parameter $\omega = 1.8$);
(3) The block Jacobi method.
(4) Algorithm 2.3

The computation generates the following results (see Table 2).

If we replace the splitting (1) with the new splitting $M_1 = H(A)$, $N_1 = -S(A)$, and unchange the others, our algorithm generates much better results, illustrating in the following Table 3.

From the above numerical results we see that the multisplitting parallel algorithms for non-Hermitian linear systems are convergent. In addition, the multisplitting parallel algorithm has less iterative number than the single splitting iteration. The rate of convergence of the multisplitting parallel algorithm depends on the section of the main splitting.

Table 1 Comparisons of Algorithm 2.2 and the single splitting iterative methods when $p = 64$

Methods	$M_1 - N_1$	$M_2 - N_2$	$M_3 - N_3$	Algorithm 2.2
IT	16956	–	–	15454

Table 2 Comparisons of Algorithm 2.3 and the single splitting iterative methods

Methods	Block Jacobi	Block SOR	Splitting (1)	Algorithm 2.3
IT	–	1734	8762	629

Table 3 Comparisons of Algorithm 2.3 and the single splitting iterative methods

Methods	Block Jacobi	Block SOR	The new splitting	Algorithm 2.3
IT	–	1734	20	16

References

1. Bai, Z.-Z.: Parallel matrix multisplitting block relaxation iteration methods. Math. Numer. Sinica **17**, 238–252 (1995)
2. Bai, Z.-Z., Sun, J.-C., Wang, D.-R.: A unified framework for the construction of various matrix multislpitting iterative methods for large sparse system of linear equations. Comput. Math. Appl. **32**, 51–76 (1996)
3. Bai, Z.-Z., Evans, D.J.: Blockwise matrix multi-splitting multi-parameter block relaxation methods. Int. J. Comput. Math. **64**, 103–118 (1997)
4. Bai, Z.-Z.: On the convergence of additive and multiplicative splitting iterations for systems of linear equations. J. Comput. Appl. Math. **154**, 195–214 (2003)
5. Bai, Z.-Z.: On the convergence of parallel nonstationary multisplitting iteration methods. J. Comput. Appl. Math. **159**, 1–11 (2003)
6. Bai, Z.-Z., Golub, G.H., Ng, M.K.: Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems. SIAM J. Matrix Anal. Appl. **24**, 603–626 (2003)
7. Bai, Z.-Z., Golub, G.H., Lu, L.-Z., Yin, J.-F.: Block triangular and skew-Hermitian splitting methods for positive definite linear systems. SIAM J. Sci. Comput. **26**, 844–863 (2006)
8. Bai, Z.-Z., Wang, C.-L.: Convergence theorems for parallel multisplitting two-stage iterative methods for mildly nonlinear systems. Linear Algebra Appl. **362**, 237–250 (2003)
9. Cao, Z.-H., Liu, Z.-Y.: Symmetric multisplitting methods of a symmetric positive definite matrix. Linear Algebra Appl. **285**, 309–319 (1998)
10. Cao, Z.-H., Wang, Y.: On validity of m-step multisplitting preconditioners for linear systems. Appl. Math. Comput. **126**, 199–211 (2002)
11. Castel, M.J., Migallón, V., Penadés, J.: Convergence of non-stationary multisplitting methods for Hermitian positive definite matrices. Math. Comput. **67**, 209–220 (1998)
12. Climent, J.J., Perea, C.: Convergence and comparison theorems for multisplittings. Numer. Linear Algebra Appl. **6**, 93–107 (1999)
13. Krukier, L.A., Martynova, T.S., Bai, Z.-Z.: Product-type skew-Hermitian triangular splitting iteration methods for strongly non-Hermitian positive definite linear systems. J. Comput. Appl. Math. **232**, 3–16 (2009)
14. Migallón, V., Penadés, J., Szyld, D.-B.: Nonstationary multispltings with general weighting matrices. SIAM J. Matrix Anal. Appl. **22**, 1089–1094 (2001)
15. Nabben, R.: A note on comparison theorems for splittings and multisplittings of Hermitian positive definite matrices. Linear Algebra Appl. **233**, 67–83 (1996)
16. Neumann, M., Plemmons, R.J.: Convergence of parallel multisplitting iterative methods for M-matrices. Linear Algebra Appl. **88**, 559–573 (1987)
17. O'Leary, D.P., White, R.: Multi-splittings of matrices and parallel solutions of linear systems. SIAM J. Algebra Discrete Methods **6**, 630–640 (1985)
18. Wang, C.-L., Bai, Z.-Z.: Sufficient conditions of convergent splitting of non-Hermitian positive definite matrices. Linear Algebra Appl. **330**, 215–218 (2001)
19. Wang, L., Bai, Z.-Z.: Skew-Hermitian triangular splitting iteration methods for non-Hermitian positive definite linear systems of strong skew-Hermitian parts. BIT Numer. Math. **44**, 363–386 (2004)
20. Wang, L., Bai, Z.-Z.: Convergence conditions for splitting iteration methods for non-Hermitian linear systems. Linear Algebra Appl. **428**, 453–468 (2008)
21. White, R.E.: Multisplitting with different weighting schemes. SIAM J. Matrix Anal. Appl. **10**, 481–493 (1989)
22. White, R.E.: Multisplitting of a symmetric positive definite matrix. SIAM J. Matrix Anal. Appl. **11**, 69–82 (1990)