Parallel Alternating Iterative Algorithms with and without Overlapping on Multicore Architectures

Héctor Migallón[a], Violeta Migallón[b], José Penadés[b]

[a]*Department of Physics and Computer Architectures, University Miguel Hernández, E-03202 Elche, Alicante, Spain*
[b]*Department of Computer Science and Artificial Intelligence, University of Alicante, E-03071 Spain*

Abstract

We consider the problem of solving large sparse linear systems where the coefficient matrix is possibly singular but the equations are consistent. Block two-stage methods in which the inner iterations are performed using alternating methods are studied. These methods are ideal for parallel processing and provide a very general setting to study parallel block methods including overlapping. Convergence properties of these methods are established when the matrix in question is either M-matrix or symmetric matrix. Different parallel versions of these methods and implementation strategies, with and without overlapping blocks, are explored. The reported experiments show the behavior and effectiveness of the designed parallel algorithms by exploiting the benefits of shared memory inside the nodes of current SMP supercomputers.

Keywords: Block two-stage methods, alternating iterations, overlapping, parallel computing, shared memory, distributed memory, Laplace's equation, Markov chains.

1. Introduction

Consider the problem of solving a linear system

$$Ax = b, \tag{1}$$

where A is an $n \times n$ matrix such that b is in $\mathcal{R}(A)$, the range of A.

Given a splitting $A = M - N$ (M nonsingular), a classical iterative method produces the following iteration scheme

$$Mx^{(l+1)} = Nx^{(l)} + b, \quad l = 0, 1, \ldots. \tag{2}$$

On the other hand, when the linear systems (2) are not solved exactly, but rather their solutions approximated by iterative methods, we are in the presence of a two-stage method (see e.g. [1], [2]). That is, consider the splitting $M = F - G$ and perform, at each outer step l, $q(l)$ inner iterations of the iterative procedure induced by this splitting. Thus, the two-stage method can be written as follows

$$x^{(l+1)} = (F^{-1}G)^{q(l)}x^{(l)} + \sum_{j=0}^{q(l)-1} (F^{-1}G)^j F^{-1}(Nx^{(l)} + b), \quad l = 0, 1, \ldots. \tag{3}$$

Without loss of generality, let us assume that the matrix A has the form

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ A_{21} & A_{22} & \cdots & A_{2r} \\ \vdots & \vdots & & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rr} \end{bmatrix}, \tag{4}$$

with the diagonal blocks A_{ii} being square of order n_i, $1 \leq i \leq r$, $\sum_{i=1}^{r} n_i = n$. Let $A = M - N$ be a splitting of A such that M is a block diagonal matrix $M = \text{Diag}\{M_1, \ldots, M_i, \ldots, M_r\}$, and let us consider the splittings $M_i = B_i - C_i$, $M_i = F_i - G_i$, $1 \leq i \leq r$. Let $M = P_i - Q_i = R_i - S_i$ be splittings of the matrix M such that

$$P_i = \text{Diag}\{I, \ldots, B_i, \ldots, I\}, \quad R_i = \text{Diag}\{I, \ldots, F_i, \ldots, I\}. \tag{5}$$

Moreover, let the $n \times n$ diagonal matrices E_i have ones in the entries corresponding to the diagonal block M_i and zero otherwise. In order to approximate the linear systems (2) we perform, at each outer iteration l, $q(i, l)$ inner iterations of the following alternating iterative scheme:

$$\begin{aligned} z_i^{(k+\frac{1}{2})} &= P_i^{-1}Q_i z_i^{(k)} + P_i^{-1}(Nx^{(l)} + b), \\ z_i^{(k+1)} &= R_i^{-1}S_i z_i^{(k+\frac{1}{2})} + R_i^{-1}(Nx^{(l)} + b), \quad k = 0, 1, \ldots, q(i, l) - 1, \end{aligned}$$

2

with $z^{(0)} = x^{(l)}$, or equivalently

$$z_i^{(k+1)} = R_i^{-1}S_iP_i^{-1}Q_iz^{(k)}+R_i^{-1}(S_iP_i^{-1}+I)(Nx^{(l)}+b), \quad k = 0, 1, \ldots, q(i,l)-1.$$

Thus, for $l = 0, 1, \ldots$, the alternating two-stage method can be written as follows,

$$
\begin{aligned}
x^{(l+1)} &= \sum_{i=1}^{r} E_iz_i^{q(i,l)} = \sum_{i=1}^{r} E_i[(R_i^{-1}S_iP_i^{-1}Q_i)^{q(i,l)}x^{(l)} \\
&+ \sum_{j=0}^{q(i,l)-1} (R_i^{-1}S_iP_i^{-1}Q_i)^jR_i^{-1}(S_iP_i^{-1}+I)(Nx^{(l)}+b)].
\end{aligned}
\tag{6}
$$

Note that the global iteration matrix of the alternating two-stage iterative method (6) can be written as $T^{(l)} = \sum_{i=1}^{r} E_iT_i^{(l)}$, with

$$T_i^{(l)} = (R_i^{-1}S_iP_i^{-1}Q_i)^{q(i,l)} + \sum_{j=0}^{q(i,l)-1} (R_i^{-1}S_iP_i^{-1}Q_i)^jR_i^{-1}(S_iP_i^{-1}+I)N, \tag{7}$$

or equivalently as

$$T^{(l)} = \sum_{i=1}^{r} E_i[(R_i^{-1}S_iP_i^{-1}Q_i)^{q(i,l)} + (I - (R_i^{-1}S_iP_i^{-1}Q_i)^{q(i,l)})M^{-1}N]. \tag{8}$$

With the above notation, the iterative scheme (6) describes an alternating two-stage Block-Jacobi type method but note that this method is much more general if, for example other matrices M, P_i, R_i and/or E_i are chosen. Particularly if $P_i = P$, $R_i = R$ for all $i = 1, \ldots, r$, this iteration scheme includes the alternating method described in [3] but this general formulation allows us to include overlapping setting the weighting diagonal nonnegative matrices E_i such that they add up to the identity. From a theoretical point of view, under certain hypotheses, the presence of overlap can reduce the convergence rate of the iterative solvers in the nonsingular case. Therefore, if the extra work required by the use of overlap is offset by a reduction in the number of iterations, probably the computation time will be reduced; see [4] and [5]. The experiments performed in [4] have been executed in only one processor using Matlab for the code implementation. However, to run

3

the experiments of [5], a parallel block iterative code was implemented with the Block Jacobi method as the outer iteration and the point Gauss-Seidel method as the inner iteration. The test matrix was generated from the discretization of the Laplace's equation using the standard five-point stencil and the experiments were performed on a parallel computer using 16 processors. As compared to the non-overlapping implementation, the parallel implementation with overlapping blocks achieved a time reduction about 5%, when the involved parameters in both algorithms were chosen near to the optimal values.

Despite the fact that the behavior of the convergence rate of the block-based iterative solvers with overlap is an open question, specially in the singular case, some numerical results given in [6] show that overlap can also improve the asymptotic convergence factor and the sequential execution time of iterative methods for singular systems, and specifically for ergodic Markov chains.

The use of quite general weighting matrices in (6) allows us the study of truly parallel methods (with or without overlap), i.e., methods in which each processor computes an approximation to the solution of a problem which is much smaller than the original problem. Recently convergence of (6) has been analyzed in the context of solving nonsingular linear systems obtaining similar convergence results to those obtained in [3]; see [7] and [8]. In this paper we give convergence results of these methods considering the general formulation for consistent linear systems. Concretely, in Section 3, we give convergence results of these methods when M-matrices or symmetric matrices are considered. The numerical experiments performed in Section 4 explore the behavior of these parallel algorithms for the solution of singular and nonsingular systems. Previously, in Section 2, we present some definitions and preliminaries that are used later in the paper. The conclusions are given in Section 5. This paper is based upon Migallón et al. [9], but the current paper includes the following additional research: new convergence results for symmetric positive semidefinite matrices are given and new parallel versions of these methods and implementation strategies, with and without overlapping blocks, are explored.

2. Notation and preliminaries

In this section we summarize some definitions and theoretical results used later in the paper. Concretely, main results about the existence and unique-

4

ness of splittings for stationary iterative methods are presented, the theoretical concepts of convergent and semiconvergent matrix are introduced along with the most important results that will be used in Section 3 to study the convergence of the alternating two-stage method when the coefficient matrix is both a singular M-matrix or a symmetric positive semidefinite matrix.

A general matrix A is called an M-matrix if A can be expressed as $A = sI - B$, with $B \geq O$, $s > 0$, and $\rho(B) \leq s$. The M-matrix A is singular when $s = \rho(B)$ and nonsingular when $s > \rho(B)$. Let $Z^{n \times n}$ denote the set of all real $n \times n$ matrices which have all non-positive off-diagonal entries. A splitting $A = M - N$ is called regular if $M^{-1} \geq O$ and $N \geq O$, and weak regular if $M^{-1} \geq O$ and $M^{-1}N \geq O$.

Lemma 1. [2] *Given a nonsingular matrix A and a matrix T such that $(I - T)^{-1}$ exists, there is a unique pair of matrices P, Q such that P is nonsingular, $T = P^{-1}Q$ and $A = P - Q$. The matrices are $P = A(I-T)^{-1}$ and $Q = P - A$.*

In the context of Lemma 1, it is said that the unique splitting $A = P - Q$ is induced by the iteration matrix T. We point out that when the matrix A is singular, the induced splitting is not unique; see e.g., [10].

Theorem 1. [10] *Let A be a nonsingular matrix such that $A^{-1} \geq O$. Let $A = M - N = P - Q$ be weak regular splittings. Consider the matrix $T = P^{-1}QM^{-1}N$, then $\rho(T) < 1$. Furthermore there is a unique pair of matrices B, C, such that $A = B - C$ is a weak regular splitting and $T = B^{-1}C$.*

Let $T \in \Re^{n \times n}$, by $\sigma(T)$ we denote the spectrum of the matrix T. We define $\gamma(T) = \max\{|\lambda| : \lambda \in \sigma(T), \lambda \neq 1\}$. We say that two subspaces S_1 and S_2 on $\Re^n$ are complementary if $S_1 \oplus S_2 = \Re^n$, i.e., if $S_1 \cap S_2 = \{0\}$ and $S_1 + S_2 = \Re^n$. The *index* of a square matrix T, denoted by *ind* T, is the smallest nonnegative integer k such that $\mathcal{R}(T^{k+1}) = \mathcal{R}(T^k)$. By $ind_1 T$ we denote the *index* associated with the value one, i.e., $ind_1 T = ind(I - T)$. Note that when $\rho(T) = 1$, $ind_1 T \leq 1$ if and only if $ind_1 T = 1$. We say that a matrix $T \in \Re^{n \times n}$, is convergent if $\lim_{k \to \infty} T^k = O$. It is well known that a matrix T is convergent if and only if $\rho(T) < 1$. By $\mathcal{N}(T)$ we denote the null space of T. We say that T is semiconvergent if $\lim_{k \to \infty} T^k$ exists, although it need not be the zero matrix. If, on the other hand, $\rho(T) = 1$, two different conditions need to be satisfied to guarantee semiconvergence, as the following result shows.

5

Theorem 2. [11] *Let $T \in \Re^{n \times n}$, with $\rho(T) = 1$. The matrix T is semiconvergent if and only if the following two statements hold.*

(a) $1 \in \sigma(T)$ *and* $\gamma(T) < 1$,

(b) $\mathcal{N}(I - T) \oplus \mathcal{R}(I - T) = \Re^n$.

Condition (b) is equivalent to the existence of the group inverse $(I - T)^{\#}$, and it is also equivalent to having $ind_1 T = 1$; see, e.g., [12].

Definition 1. [12] *Let $A \in \Re^{n \times n}$, and consider the following matrix equations.*
(1) $AXA = A$,
(2) $XAX = X$, and
(3) $AX = XA$.
A $\{1, 2\}$-inverse of A is a matrix X which satisfies conditions (1) and (2). If, in addition, X satisfies condition (3), X is said to be a group inverse of A.

Note that the group inverse $A^{\#}$ of a matrix A, if it exists, is unique and when A is nonsingular, each generalized inverse coincides with A^{-1}.

Theorem 3. [12] *Let $T \in \Re^{n \times n}$, with $T \geq O$, and let C be a $\{1, 2\}$-inverse of $I - T$ with $\mathcal{R}(C)$ complementary to $\mathcal{N}(I - T)$, such that C is nonnegative on $\mathcal{R}(I - T)$, i.e., the matrix C satisfies the following conditions.*
(i) $I - T = (I - T)C(I - T)$,
(ii) $C = C(I - T)C$,
(iii) $\mathcal{N}(I - T) \oplus \mathcal{R}(C) = \Re^n$,
(iv) If $x \in \mathcal{R}(I - T)$, $x \geq 0$ then $Cx \geq 0$.
Then, $\rho(T) \leq 1$, and $ind_1(T) \leq 1$.

Lemma 2. [12] *Let $T \in \Re^{n \times n}$ be semiconvergent. Then*

$$\lim_{k \to \infty} T^k = I - (I - T)(I - T)^{\#}.$$

Theorem 4. [13] *Let $A^{(l)}$, $l = 0, 1, \ldots$, be a sequence of square complex matrices such that each group inverse $(I - A^{(l)})^{\#}$ exists. Suppose that there is a subspace S satisfying $\mathcal{N}(I - A^{(l)}) = S$, $l = 0, 1, \ldots$. If there exists a matrix norm $\| \cdot \|$ such that the set $\{\|A^{(l)}\|\}_{l=0}^{\infty}$ remains bounded and $\|A^{(l)}(I - A^{(l)})(I - A^{(l)})^{\#}\| \leq \theta < 1$, $l = 0, 1, \ldots$, then $\lim_{i \to \infty} A^{(l)} A^{(l-1)} \cdots A^{(0)} = P$, where P is a projection matrix onto the subspace S.*

6

Definition 2. [12] *A general M-matrix A is said to have property c if for some representation of $A = sI - B$, $s > 0$, $B \geq O$, the matrix $s^{-1}B$ is semiconvergent.*

Obviously, a nonsingular M-matrix always has *property c*.

Theorem 5. [11] *Let $A \in Z^{n \times n}$. Let $A = M - N$ be a regular splitting, and let $T = M^{-1}N$. Then A is an M-matrix with property c if and only if $\rho(T) \leq 1$, and $\mathcal{N}(I - T) \oplus \mathcal{R}(I - T) = \Re^n$.*

The transpose and the conjugate transpose of a matrix $A \in \mathbb{C}^{n \times n}$ are denoted by A^T and A^H, respectively. Similarly, given a vector $x \in \mathbb{C}^n$, x^T and x^H denote the transpose and the conjugate transpose of x, respectively. A matrix $A \in \mathbb{C}^{n \times n}$ is said to be symmetric if $A = A^T$, and Hermitian if $A = A^H$. Clearly a real symmetric matrix is a particular case of a Hermitian matrix. A complex, not necessarily Hermitian matrix A, is called positive definite (positive semidefinite) if the real part of $x^H A x$ is positive (nonnegative), for all complex $x \neq 0$. When A is Hermitian, this is equivalent to requiring that $x^H A x > 0$ ($x^H A x \geq 0$), for all complex $x \neq 0$. A general matrix A is positive definite (positive semidefinite) if and only if the Hermitian matrix $A + A^H$ is positive definite (positive semidefinite). Given a matrix $A \in \mathbb{C}^{n \times n}$, the splitting $A = M - N$ is called P-regular if the matrix $M^H + N$ is positive definite. If a matrix A is symmetric positive definite it induces a vector norm $\|x\|_A = (x^T A x)^{\frac{1}{2}}$.

Lemma 3. [14] *Let A be a symmetric positive definite matrix and let $A = B - C$ be a P-regular splitting. Given $s \geq 1$, the unique splitting induced by $(B^{-1}C)^s$ is also a P-regular splitting.*

Theorem 6. [14] *Let A be a Hermitian positive definite matrix. Let $A = M - N = P - Q$ be P-regular splittings. Consider the matrix $T = P^{-1}QM^{-1}N$, then $\rho(T) < 1$. Moreover, the unique splitting $A = B - C$ induced by the iteration matrix T, such that $T = B^{-1}C$, is also P-regular.*

Theorem 7. [12] *Let $A = M - N$ be a P-regular splitting of a symmetric matrix A. Then the matrix $M^{-1}N$ is semiconvergent if and only if A is positive semidefinite.*

Theorem 8. [15] *Let A be a symmetric positive definite matrix. A splitting $A = M - N$ is P-regular if and only if $\|M^{-1}N\|_A < 1$.*

3. Convergence

In this section we analyze the convergence of the alternating two-stage method (6) in the context of singular M-matrices and symmetric positive semidefinite matrices. Additionally, these results are extended to the nonsingular case. First of all we proceed to study the convergence of the alternating two-stage method when A is an M-matrix with *property c*.

Theorem 9. *Let A be an M-matrix with property c. Let the splitting $A = M - N$ be regular, and the splittings $M = P_i - Q_i = R_i - S_i$, $1 \leq i \leq r$ be weak regular. Then, the matrices $T^{(l)}$, $l = 0, 1, \ldots$, defined in (7), satisfy $\rho(T^{(l)}) \leq 1$ and $ind_1 T^{(l)} \leq 1$.*

PROOF. From (8), it follows that $I - T^{(l)} = (I - H^{(l)})(I - M^{-1}N)$, $l = 0, 1, \ldots$, where $H^{(l)} = \sum_{i=1}^{r} E_i (R_i^{-1} S_i P_i^{-1} Q_i)^{q(i,l)}$. From [16, Theorem 2.1] and Theorem 1 it follows that $\rho(H^{(l)}) < 1$. Therefore $(I - H^{(l)})^{-1}$ exists. On the other hand the existence of $(I - M^{-1}N)^{\#}$ follows from Theorem 5. Let us consider the matrix $C = (I - M^{-1}N)^{\#}(I - H^{(l)})^{-1}$. Using Definition 1, the matrix C satisfies conditions (i) and (ii) of Theorem 3. Furthermore, $\mathcal{R}(C) = \mathcal{R}((I - M^{-1}N)^{\#}) = \mathcal{R}(I - M^{-1}N)$ and $\mathcal{N}(I - T^{(l)}) = \mathcal{N}(I - M^{-1}N) = \mathcal{N}(A)$. Moreover, from Theorem 5, it follows that $\mathcal{R}(I - M^{-1}N)$ and $\mathcal{N}(I - M^{-1}N)$ are complementary. Let $x \in \mathcal{R}(I - T^{(l)})$, $x \geq 0$, then following Theorem 3, to conclude the proof we need to show that $Cx \geq 0$. Since $M^{-1}N \geq O$ and $(I - M^{-1}N)^{\#}$ exists, it follows from [17, Theorem 2] that $(I - M^{-1}N)^{\#}$ is nonnegative on $\mathcal{R}(I - M^{-1}N)$. Taking into account that $(I - H^{(l)})^{-1}x \in \mathcal{R}(I - M^{-1}N)$ and $(I - H^{(l)})^{-1}x \geq 0$, the proof is complete.

Theorem 10. *Let A be an M-matrix with property c. Let the splitting $A = M - N$ be regular, and the splittings $M = P_i - Q_i = R_i - S_i$, $1 \leq i \leq r$ be weak regular. Assume further that the diagonal entries of the matrices $P_i^{-1} Q_i$ and $R_i^{-1} S_i$, are positive. Then, the matrices $T^{(l)}$, $l = 0, 1, \ldots$, defined in (7), are semiconvergent.*

PROOF. From the hypotheses it follows for all $l = 0, 1, \ldots$, that the matrices

$$T^{(l)} = \sum_{i=1}^{r} E_i [(R_i^{-1} S_i P_i^{-1} Q_i)^{q(i,l)} + \sum_{j=0}^{q(i,l)-1} (R_i^{-1} S_i P_i^{-1} Q_i)^j R_i^{-1}(S_i P_i^{-1} + I)N],$$

8

are nonnegative and have positive diagonal entries. Moreover, from Theorem 9, the matrices $T^{(l)}$, $l = 0, 1, \ldots$, satisfy condition (b) of Theorem 2. Therefore, using the result in [18, Theorem 2], the proof is complete.

Theorem 11. *Let A be an M-matrix with property c. Let the splitting $A = M - N$ be regular, and the splittings $M = P_i - Q_i = R_i - S_i$, $1 \leq i \leq r$ be weak regular. Then, for each $\delta \in (0, 1)$, the matrices $T_\delta^{(l)} = \delta T^{(l)} + (1 - \delta)I$, $l = 0, 1, \ldots$, with $T^{(l)}$ defined in (7), are semiconvergent.*

PROOF. Since $I - T_\delta^{(l)} = \delta(I - T^{(l)})$, $l = 0, 1, \ldots$, from Theorem 9 it follows, for each $\delta \in (0, 1)$, that $\rho(T_\delta^{(l)}) \leq 1$ and $\mathcal{N}(I - T_\delta^{(l)}) \oplus \mathcal{R}(I - T_\delta^{(l)}) = \Re^n$, $l = 0, 1, \ldots$. Moreover, $T^{(l)} \geq O$. Thus (see e.g., [12, Exercise 6.4.3]), $T_\delta^{(l)}$ has only the eigenvalue one on the unit circle, and from Theorem 2 it follows that $T_\delta^{(l)}$ is semiconvergent for all $\delta \in (0, 1)$.

Note that in Theorem 10 we have assumed that the matrices $P_i^{-1}Q_i$ and $R_i^{-1}S_i$, have positive diagonal entries. However, the iteration matrices of some classical alternating iterative methods do not have this property. In order to ensure that condition (a) of Theorem 2 holds, from Theorem 11 it follows that equation (6) can be replaced in the alternating two-stage method by

$$x^{(l+1)} = \delta(T^{(l)}x^{(l)} + c_l) + (1 - \delta)x^{(l)}, \quad l = 0, 1, \ldots. \tag{9}$$

Theorem 12. *Let A be an M-matrix with property c. Let the splitting $A = M - N$ be regular, and the splittings $M = P_i - Q_i = R_i - S_i$ be weak regular. Assume that the sequence of inner iterations $\{q(i, l)\}_{l=0}^{\infty}$ satisfies $q(i, l) = q(i)$, $l = 0, 1, \ldots$. Then the following two results hold.*
(a) If the diagonal entries of the matrices $P_i^{-1}Q_i$ and $R_i^{-1}S_i$, are positive, the alternating two-stage method (6) converges to a solution of the consistent linear system $Ax = b$, for any initial vector $x^{(0)}$.
(b) The alternating two-stage method (6) with the modification (9), converges to a solution of the consistent linear system $Ax = b$, for any initial vector $x^{(0)}$.

PROOF. Since $q(i, l) = q(i)$, $l = 0, 1, \ldots$, then there is a single iteration matrix, i.e.,

$$T^{(l)} = T = \sum_{i=1}^{r} E_i[(R_i^{-1}S_iP_i^{-1}Q_i)^{q(i)} + \sum_{j=0}^{q(i)-1} (R_i^{-1}S_iP_i^{-1}Q_i)^j R_i^{-1}(S_iP_i^{-1} + I)N].$$

9

Let x^* be a solution of (1), then $e^{(l)} = x^{(l)} - x^* = Te^{(l-1)} = T^l e^{(0)}$, for $l = 1, 2, \ldots$. In the case (a), from Theorem 10, T is semiconvergent, and from Lemma 2 it follows that

$$\lim_{l \to \infty} e^{(l)} = \lim_{l \to \infty} T^l e^{(0)} = [I - (I - T)(I - T)^{\#}]e^{(0)} \in \mathcal{N}(I - T) = \mathcal{N}(A).$$

Therefore, the proof of part (a) is complete. The proof of part (b) is analogous, but using Theorem 11.

Theorem 13. *Let A be an M-matrix with property c. Let the splitting $A = M - N$ be regular, and the splittings $M = P_i - Q_i = R_i - S_i, 1 \leq i \leq r$ be weak regular. Suppose that there exists a matrix norm $\| \cdot \|$ such that $\|T^{(l)}(I - T^{(l)})(I - T^{(l)})^{\#}\| < 1, \; l = 0, 1, \ldots$, where $T^{(l)}$ are defined in (7). Assume further that the sequence of inner iterations $q(i, l), \; l = 0, 1 \ldots, \; 1 \leq i \leq r$ remains bounded. Then, the alternating two-stage iterative method (6) converges to a solution of the consistent linear system $Ax = b$, for any initial vector $x^{(0)}$*

PROOF. The proof is an immediate consequence of Theorems 4 and 9.

Note that in the particular case in which A is a nonsingular M-matrix, the hypotheses of Theorem 13 can be lightened as follows.

Theorem 14. *Let A be a nonsingular M-matrix. Let the splitting $A = M - N$ be regular, and the splittings $M = P_i - Q_i = R_i - S_i, 1 \leq i \leq r$ be weak regular. Then, the alternating two-stage iterative method (6) converges to the solution of the nonsingular linear system $Ax = b$, for any initial vector $x^{(0)}$ and for any sequence of inner iterations $q(i, l), \; l = 0, 1 \ldots, \; 1 \leq i \leq r$.*

PROOF. Reasoning as in Theorem 15 of [3] the iteration matrices can be written as

$$T^{(l)} = \sum_{i=1}^{r} E_i[(B_i^{-1}C_i)^{q(i,l)} + (I - (B_i^{-1}C_i)^{q(i,l)})M^{-1}N], \; l = 0, 1, \ldots,$$

such that $B_i^{-1}C_i = R_i^{-1}S_i P_i^{-1}Q_i$, and $M = B_i - C_i, 1 \leq i \leq r$ are weak regular splittings. Therefore the matrices $T^{(l)}, l = 0, 1, \ldots$, can be seen as the iteration matrices of a two-stage multisplitting method. Therefore using the result of Theorem 3.2 in [19] the proof is complete.

10

Next we study the symmetric positive semidefinite case.

Theorem 15. *Let A be a symmetric positive semidefinite matrix. Let the splitting $A = M - N$ be such that M is a symmetric positive definite matrix and N is a positive semidefinite matrix. Let $M = P_i - Q_i = R_i - S_i, 1 \le i \le r$ be P-regular splittings and $E_i = \alpha_i I, 1 \le i \le r$, with $\alpha_i > 0$ and $\sum_{i=1}^{r} \alpha_i = 1$. Assume that the sequence of inner iterations $\{q(i,l)\}_{l=0}^{\infty}$ satisfies $q(i,l) = q(i), \; l = 0, 1, \dots$. Then the alternating two-stage method (6) converges to a solution of the consistent linear system $Ax = b$, for any initial vector $x^{(0)}$.*

PROOF. Since $q(i,l) = q(i), \; l = 0, 1, \dots$, then there is a single iteration matrix, i.e.,

$$T^{(l)} = T = \sum_{i=1}^{r} E_i[(R_i^{-1}S_iP_i^{-1}Q_i)^{q(i)} + (I - (R_i^{-1}S_iP_i^{-1}Q_i)^{q(i)})M^{-1}N].$$

Moreover, from Theorem 6, for each i, $1 \le i \le r$, there exists a pair of matrices B_i, C_i, such that $R_i^{-1}S_iP_i^{-1}Q_i = B_i^{-1}C_i$, $M = B_i - C_i$ is a P-regular splitting and $\rho(B_i^{-1}C_i) < 1$. Therefore, $I - (B_i^{-1}C_i)^{q(i)}$ is a nonsingular matrix. Thus, from Lemma 1 and Lemma 3 it follows that the splitting induced by $(B_i^{-1}C_i)^{q(i)}$, namely $M = \hat{B}_i - \hat{C}_i$, with $\hat{B}_i = M(I - (B_i^{-1}C_i)^{q(i)})^{-1}$, is P-regular.

On the other hand, since each matrix $\hat{B}_i$ is positive definite, the matrix $\sum_{i=1}^{r} E_i\hat{B}_i^{-1}$ is positive definite, and therefore nonsingular. Moreover $\rho(\sum_{i=1}^{r} E_i\hat{B}_i^{-1}\hat{C}_i) < 1$.

Let $\hat{B} = \left(\sum_{i=1}^{r} E_i\hat{B}_i^{-1}\right)^{-1}$ and $\hat{C} = \hat{B}\left(\sum_{i=1}^{r} E_i\hat{B}_i^{-1}\hat{C}_i\right)$. Then taking into account that $M = \hat{B} - \hat{C}$, it obtains

$$T^{(l)} = T = \sum_{i=1}^{r} E_i\hat{B}_i^{-1}\hat{C}_i + \left(I - \sum_{i=1}^{r} E_i\hat{B}_i^{-1}\hat{C}_i\right)M^{-1}N =$$

$$= \hat{B}^{-1}(\hat{C} + (\hat{B} - \hat{C})M^{-1}N) = \hat{B}^{-1}(\hat{C} + N).$$

Thus, the splitting $A = \hat{B} - (\hat{C} + N)$ is a (non-unique) splitting induced by T. Since $\hat{B}^T + \hat{C}$ is positive definite and N is positive semidefinite, $\hat{B}^T + \hat{C} + N$ is positive definite and thus this splitting is P-regular. Therefore, from Theorem 7 it follows that T is a semiconvergent matrix and the proof is completed.

11

Theorem 16. *Let A be a symmetric positive semidefinite matrix. Let the splitting $A = M - N$ be such that M is a symmetric positive definite matrix and N is a positive semidefinite matrix. Let $M = P_i - Q_i = R_i - S_i$ be P-regular splitting and $E_i = \alpha_i I, 1 \le i \le r$, with $\alpha_i > 0$ and $\sum_{i=1}^r \alpha_i = 1$. Suppose that there exists a matrix norm $\| \cdot \|$ such that $\|T^{(l)}(I - T^{(l)})(I - T^{(l)})^\# \| < 1$, $l = 0, 1, \ldots$, where $T^{(l)}$ are defined in (7). Assume further that the sequence of inner iterations $\{q(i,l)\}_{l=0}^\infty$ remains bounded. Then, the alternating two-stage iterative method (6) converges to a solution of the consistent linear system $Ax = b$, for any initial vector $x^{(0)}$.*

PROOF. The proof is an immediate consequence of Theorems 4 and 15.

We want to point out that in the particular case in which A is a symmetric positive definite matrix the linear system (1) is nonsingular and then the following results hold.

Corollary 1. *Let A be a symmetric positive definite matrix. Let the splitting $A = M - N$ be such that M is symmetric and N is a positive semidefinite matrix. Let $M = P_i - Q_i = R_i - S_i, 1 \le i \le r$ be P-regular splittings and $E_i = \alpha_i I, 1 \le i \le r$, with $\alpha_i > 0$ and $\sum_{i=1}^r \alpha_i = 1$. Assume that the sequence of inner iterations $\{q(i,l)\}_{l=0}^\infty$ satisfies $q(i,l) = q(i)$, $l = 0, 1, \ldots$. Then the alternating two-stage method (6) converges to the solution of the nonsingular linear system $Ax = b$, for any initial vector $x^{(0)}$.*

Corollary 2. *Let A be a symmetric positive definite matrix. Let the splitting $A = M - N$ be such that M is symmetric and N is a positive semidefinite matrix. Let $M = P_i - Q_i = R_i - S_i$ be P-regular splittings and $E_i = \alpha_i I, 1 \le i \le r$, with $\alpha_i > 0$ and $\sum_{i=1}^r \alpha_i = 1$. Suppose that there exists a matrix norm $\| \cdot \|$ such that $\|T^{(l)}(I - T^{(l)})(I - T^{(l)})^\# \| < 1$, $l = 0, 1, \ldots$, where $T^{(l)}$ are defined in (7). Assume further that the sequence of inner iterations $\{q(i,l)\}_{l=0}^\infty$ remains bounded. Then, the alternating two-stage iterative method (6) converges to the solution of the nonsingular linear system $Ax = b$, for any initial vector $x^{(0)}$.*

In [8], the authors give convergence results for the alternating two-stage method (6) similar to Corollary 1 but considering weighting matrices $E_i \ge O$, $1 \le i \le r$ and $\sum_{i=1}^r E_i = 1$. However, we point out that if the weighting matrices E_i are of this form, the alternating two-stage iterative method (6)

may not converge when A is symmetric positive semidefinite (or symmetric positive definite), even though the splittings satisfy Theorem 15 or 16 (Corollary 1 or 2); see Examples 1 and 2. Therefore, Theorems 3 and 4 of [8] are not true. This is due to they consider in the proof of these theorems that $\sum_{i=1}^{r} \|E_i\|_A = 1$ and this equality does not hold for the above general weighting matrices.

Example 1. Let A be the following symmetric positive semidefinite matrix.

$$A = \begin{bmatrix} 0.75 & 0.75 \\ 0.75 & 0.75 \end{bmatrix}.$$

Consider the splitting $A = M - N$, such that

$$M = \begin{bmatrix} 1 & 0.75 \\ 0.75 & 1 \end{bmatrix} = P_i - Q_i = R_i - S_i, \ i = 1, 2,$$

where

$$P_1 = R_1 = \begin{bmatrix} 1.75 & -2 \\ 2 & 3 \end{bmatrix}, \quad P_2 = R_2 = \begin{bmatrix} 3 & 2 \\ -2 & 1.75 \end{bmatrix}.$$

Note that M is a symmetric positive definite matrix, $N = M - A$ is a positive definite matrix and $M = P_i - Q_i = R_i - S_i, \ i = 1, 2$ are P-regular splitting. Then, setting $E_1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, $E_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, and $q(i, l) = q(i) = 1, \ i = 1, 2, \ l = 0, 1, \ldots$, the iteration matrix $T = \sum_{i=1}^{2} E_i[(R_i^{-1} S_i P_i^{-1} Q_i) + (I - (R_i^{-1} S_i P_i^{-1} Q_i))M^{-1}N] = \begin{bmatrix} 1.00986 & 0.00986 \\ 0.00986 & 1.00986 \end{bmatrix}$, has spectral radius equal to 1.0197, and the alternating two-stage iterative method (6) is not convergent.

Example 2. Let A be the following symmetric positive definite matrix

$$A = \begin{bmatrix} 0.75 & 0 \\ 0 & 0.5 \end{bmatrix}.$$

Consider the splitting $A = M - N$, such that

$$M = \begin{bmatrix} 0.75 & 0 \\ 0 & 0.75 \end{bmatrix} = P_i - Q_i = R_i - S_i, \ i = 1, 2,$$

where

$$P_1 = R_1 = \begin{bmatrix} 0.3934 & -2.0660 \\ 2.0660 & 7.6244 \end{bmatrix}, \quad P_2 = R_2 = \begin{bmatrix} 7.6244 & 2.0660 \\ -2.0660 & 0.3934 \end{bmatrix}.$$

13

Note that M is a symmetric matrix, $N = M - A$ is a positive semidefinite matrix and $M = P_i - Q_i = R_i - S_i$, $i = 1, 2$ are P-regular splitting. Then, setting $E_1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, $E_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, and $q(i, l) = q(i) = 1$, $i = 1, 2$, $l = 0, 1, \ldots$, the iteration matrix $T = \sum_{i=1}^{2} E_i[(R_i^{-1} S_i P_i^{-1} Q_i) + (I - (R_i^{-1} S_i P_i^{-1} Q_i))M^{-1}N] = \begin{bmatrix} 0.87499 & 0.16666 \\ 0.24999 & 0.91666 \end{bmatrix}$, has spectral radius equal to 1.101, and the alternating two-stage iterative method (6) is not convergent.

Theorem 17. *Let A be a symmetric positive semidefinite matrix. Let the splitting $A = M - N$ be such that $M = \mathrm{Diag}(M_1, \ldots, M_i, \ldots, M_r)$ is a symmetric positive definite matrix and N is a positive semidefinite matrix. Let $M = P_i - Q_i = R_i - S_i, 1 \le i \le r$ be the splittings defined in (5) such that the splittings $M_i = B_i - C_i = F_i - G_i, 1 \le i \le r$ are P-regular. Consider the diagonal matrices $E_i, 1 \le i \le r$ have ones in the entries corresponding to the diagonal block M_i and zero otherwise and assume that the sequence of inner iterations $\{q(i, l)\}_{l=0}^{\infty}$ satisfies $q(i, l) = q(i)$, $l = 0, 1, \ldots$. Then the alternating two-stage method (6) converges to a solution of the consistent linear system $Ax = b$, for any initial vector $x^{(0)}$.*

PROOF. Since $q(i, l) = q(i)$, $l = 0, 1, \ldots$, from (5) and (8) it follows that

$$T^{(l)} = T = H + (I - H)M^{-1}N,$$

with

$$H = \mathrm{Diag}((F_1^{-1} G_1 B_1^{-1} C_1)^{q(1)}, \ldots, (F_i^{-1} G_i B_i^{-1} C_i)^{q(i)}, \ldots, (F_r^{-1} G_r B_r^{-1} C_r)^{q(r)})$$

and $M^{-1} = \mathrm{Diag}(M_1^{-1}, \ldots, M_i^{-1}, \ldots, M_r^{-1})$. From Theorem 6, for i, $1 \le i \le r$, there exists a pair of matrices U_i, V_i, such that $F_i^{-1} G_i B_i^{-1} C_i = U_i^{-1} V_i$, $M_i = U_i - V_i$ is a P-regular splitting and $\rho(U_i^{-1} V_i) < 1$. Therefore, $I - (U_i^{-1} V_i)^{q(i)}$ is a nonsingular matrix. From Lemma 1 and Lemma 3 it follows that the splitting induced by $(U_i^{-1} V_i)^{q(i)}$, namely $M_i = \hat{U}_i - \hat{V}_i$, with $\hat{U}_i = M_i(I - (U_i^{-1} V_i)^{q(i)})^{-1}$, is P-regular. Let $\hat{U} = \mathrm{Diag}(\hat{U}_1, \ldots, \hat{U}_i, \ldots, \hat{U}_r)$ and $\hat{V} = \mathrm{Diag}(\hat{V}_1, \ldots, \hat{V}_i, \ldots, \hat{V}_r)$, then $T^{(l)} = T = \hat{U}^{-1} \hat{V} + (I - \hat{U}^{-1} \hat{V})M^{-1}N = \hat{U}^{-1}(\hat{V} + (\hat{U} - \hat{V})M^{-1}N) = \hat{U}^{-1}(\hat{V} + N)$. Thus, the splitting $A = \hat{U} - (\hat{V} + N)$ is a (non-unique) splitting induced by T. Since $\hat{U}^T + \hat{V}$ is positive definite and N is positive semidefinite, reasoning in the same way as in Theorem 15 the proof is completed.

14

Theorem 18. *Let A be a symmetric positive semidefinite matrix. Let the splitting $A = M - N$ be such that $M = \text{Diag}(M_1, \ldots, M_i, \ldots, M_r)$ is a symmetric positive definite matrix and N is a positive semidefinite matrix. Let $M = P_i - Q_i = R_i - S_i, 1 \le i \le r$ be the splittings defined in (5) such that the splittings $M_i = B_i - C_i = F_i - G_i, 1 \le i \le r$ are P-regular. Consider the diagonal matrices $E_i, 1 \le i \le r$ have ones in the entries corresponding to the diagonal block M_i and zero otherwise. Suppose that there exists a matrix norm $\| \cdot \|$ such that $\|T^{(l)}(I - T^{(l)})(I - T^{(l)})^{\#}\| < 1, \; l = 0, 1, \ldots,$ where $T^{(l)}$ are defined in (7). Assume further that the sequence of inner iterations $\{q(i, l)\}_{l=0}^{\infty}$ remains bounded. Then, the alternating two-stage iterative method (6) converges to a solution of the consistent linear system $Ax = b$, for any initial vector $x^{(0)}$.*

PROOF. The proof is an immediate consequence of Theorems 4 and 17.

Theorem 19. *Let A be a symmetric positive semidefinite matrix. Let the splitting $A = M - N$ be such that M is a symmetric positive definite matrix and N is a positive semidefinite matrix. Let $M = P_i - Q_i = R_i - S_i, 1 \le i \le r$ be P-regular splittings. Assume that the sequence of inner iterations $\{q(i, l)\}_{l=0}^{\infty}$ satisfies $q(i, l) = q(i), \; l = 0, 1, \ldots$. Given a fixed positive number $\theta < 1$, let $\eta = \theta/(\sum_{i=1}^{r} \|E_i\|_M)$. Let $\hat{q}$ be such that $\|(R_i^{-1}S_iP_i^{-1}Q_i)^q\|_M \le \eta$ for all $q \ge \hat{q}, \; i = 1, \ldots, r$. If $q(i) \ge \hat{q}$, then the alternating two-stage method (6) converges to a solution of the consistent linear system $Ax = b$, for any initial vector $x^{(0)}$.*

PROOF. Since $q(i, l) = q(i), \; l = 0, 1, \ldots$, then there is a single iteration matrix, i.e.,

$$T^{(l)} = T = \sum_{i=1}^{r} E_i[(R_i^{-1}S_iP_i^{-1}Q_i)^{q(i)} + (I - (R_i^{-1}S_iP_i^{-1}Q_i)^{q(i)})M^{-1}N].$$

Moreover, from Theorem 6, there is a pair of matrices B_i, C_i, such that $R_i^{-1}S_iP_i^{-1}Q_i = B_i^{-1}C_i$, and $M = B_i - C_i$ is a P-regular splitting. Since M is a symmetric positive definite matrix, by the hypotheses and Theorem 8 it obtains that $\| \sum_{i=1}^{r} E_i(B_i^{-1}C_i)^{q(i)}\|_M < 1$. Then, using Lemma 1 there exist B and C such that $\sum_{i=1}^{r} E_i(B_i^{-1}C_i)^{q(i)} = B^{-1}C$ and $M = B - C$. Since $\|B^{-1}C\|_M < 1$, this splitting is P-regular and it obtains

$$T^{(l)} = T = B^{-1}(C + (B - C)M^{-1}N) = B^{-1}(C + N).$$

15

Thus, the splitting $A = B - (C + N)$ is a (non-unique) splitting induced by T. Since $B^T + C$ is positive definite and N is positive semidefinite, reasoning in the same way as in Theorem 15 the proof is completed.

Note that, Theorem 15 requires the assumption that weighting matrices are multiples of identity. This assumption on the weighting matrices have little applicability for analysis of parallel processing. However, Theorem 19 does not have this restriction. In the Example 1, where there is no convergence using general weighting matrices, the splittings satisfies hypotheses of Theorem 19, and the smallest integer $\hat{q}$ for which $\|(R_i^{-1} S_i P_i^{-1} Q_i)^{\hat{q}}\|_M < \frac{1}{\sum_{i=1}^{2} \|E_i\|_M}$, is $\hat{q} = 11$. Thus for $q(i, l) = q(i) \geq 11 \ l = 0, 1, \ldots$, Theorem 19 assures the convergence of the alternating two-stage method (6). Note that, in this case, $q(i, l) = q(i) = 2$ is the smallest integer for which the alternating two-stage method (6) converges.

Theorem 20. *Let A be a symmetric positive semidefinite matrix. Let the splitting $A = M - N$ be such that M is a symmetric positive definite matrix and N is a positive semidefinite matrix. Let $M = P_i - Q_i = R_i - S_i$ be P-regular splitting. Given a fixed positive number $\theta < 1$, let $\eta = \theta/(\sum_{i=1}^{r} \|E_i\|_M)$. Let $\hat{q}$ be such that $\|(R_i^{-1} S_i P_i^{-1} Q_i)^q\|_M \leq \eta$ for all $q \geq \hat{q}$, $i = 1, \ldots, r$. Suppose that there exists a matrix norm $\| \cdot \|$ such that $\|T^{(l)}(I - T^{(l)})(I - T^{(l)})^{\#}\| < 1$, $l = 0, 1, \ldots$, where $T^{(l)}$ are defined in (7). If $q(i, l) \geq \hat{q}$, then, the alternating two-stage iterative method (6) converges to a solution of the consistent linear system $Ax = b$, for any initial vector $x^{(0)}$.*

PROOF. The proof is an immediate consequence of Theorems 4 and 19.

Note that in the particular case in which A is symmetric positive definite and therefore nonsingular, the hypotheses on the matrix M in Theorems 15–20 can be lightened in the sense that if $M = A + N$ is symmetric then M is a symmetric positive definite matrix.

4. Experimental setup and results

The alternating iterative methods treated here have been applied to the solution of singular and nonsingular linear systems. In the nonsingular case, the problem to be solved comes from the discretization of the Laplace's equation satisfying Dirichlet boundary conditions on the rectangle $\Omega = [0, a] \times [0, b]$. The discretization of the domain, using five point finite

differences, with $J \times K$ points equally spaced by h, yields a linear system $Ax = b$, where A is block tridiagonal, $A = \text{tridiag}[-I, C, -I]$, where I and C are $K \times K$ matrices, I is the identity, and $C = \text{tridiag}[-1, 4, -1]$. Note that A is a nonsingular M-matrix with $J \times J$ blocks of size $K \times K$. Then the convergence of the alternating two-stage method (6) is guaranteed when the splittings are chosen as in Theorem 14. Note that the Block Jacobi, the Block Gauss-Seidel and the Symmetric Block Gauss-Seidel splittings of an M-matrix are regular and weak regular splittings [12]. In the singular case, the test problems arise from Markov chain modelling. Concretely, these methods can be used to find the stationary probability distribution of a Markov chain, i.e., one is looking for a nonnegative vector x such that $Bx = x$, where B is a nonnegative column stochastic matrix, i.e., $B^T e = e$, where $e = (1, 1, \ldots, 1)^T$. This implies that $\rho(B) = \rho(B^T) = 1$; see e.g., [12]. The vector of probabilities is normalized so that $x^T e = 1$. In this case, the system to be solved is

$$(I - B)x = 0. \tag{10}$$

If B is a transition matrix of a Markov chain, the matrix $A = I - B$ is an M-matrix with *property c*, and thus the convergence of the alternating two-stage method (6) with the modification (9), if need be, is guaranteed when the splittings are chosen as in Theorems 12 and 13.

In the performed experiments , the variables are partitioned into r groups, i.e., $x = [x_1^T, x_2^T, \ldots, x_r^T]^T$, $x_i \in \Re^{n_i}$, $i = 1, \ldots, r$, $\sum_{i=1}^{r} n_i = n$. Therefore, the matrix A is partitioned into $r \times r$ blocks as it is described in (4), with the diagonal blocks A_{ii} being square of order n_i. In the experiments reported in this paper, the number of obtained groups, r, is larger than the number of processes p. Thus, we have assembled blocks from (4) into p groups, each group being assigned to one process. There are r_ℓ blocks assigned to process ℓ, $\ell = 1, \ldots, p$, and thus $\sum_{\ell=1}^{p} r_\ell = r$. In order to explain the tested iterative methods, we describe the Block Jacobi (BJ) and the Symmetric Block Gauss-Seidel (SBGS) Algorithms, for solving these linear systems, for a generic number of blocks q in (4).

By setting $b = [b_1^T, b_2^T, \ldots, b_r^T]^T = 0$, Algorithms 1 and 2 can be used for solving the Markov singular linear system (10). In this case, we will assume that the solution x is normalized so that $x^T e = 1$. In fact, in the algorithms studied below, such normalization is assumed at every iteration when singular linear systems are solved. On the other hand, in order to ensure the regularity of the Block Jacobi splitting, when B is the transition matrix of

17

Algorithm 1: (Block Jacobi - BJ)

Initialization $x^{(0)^T} = [x_1^{(0)^T}, x_2^{(0)^T}, \ldots, x_q^{(0)^T}], \ l = 0;$

repeat

    for $i = 1$ *to* q **do**

        Solve (or approximate) $A_{ii} x_i^{(l+1)} = -\sum_{\substack{j=1 \\ j \neq i}}^{q} A_{ij} x_j^{(l)} + b_i; (11)$

    end

until *convergence*;

Algorithm 2: (Symmetric Block Gauss-Seidel - SBGS)

Initialization $x^{(0)^T} = [x_1^{(0)^T}, x_2^{(0)^T}, \ldots, x_q^{(0)^T}], \ l = 0;$

repeat

    for $i = 1$ *to* q **do**

        Solve (or approximate)

        $A_{ii} x_i^{(l+\frac{1}{2})} = -\sum_{j=1}^{i-1} A_{ij} x_j^{(l+\frac{1}{2})} - \sum_{j=i+1}^{q} A_{ij} x_j^{(l)} + b_i;$ (12)

    end

until *convergence*;

repeat

    for $i = q$ *to* 1 **do**

        Solve (or approximate)

        $A_{ii} x_i^{(l+1)} = -\sum_{j=i+1}^{q} A_{ij} x_j^{(l+1)} - \sum_{j=1}^{i-1} A_{ij} x_j^{(l+\frac{1}{2})} + b_i;$ (13)

    end

until *convergence*;

---

a finite homogeneous Markov chain, one can suppose that each column of $N = A - \mathrm{Diag}(A_{11}, \ldots, A_{rr})$ must have one non-zero entry, or each block A_{ii} must be irreducible and at least one column, for each corresponding block in N, must have at least one non-zero entry; see, e.g., [13]. In Algorithm 1, the linear systems (11) can be solved independently of each other. Thus, this algorithm is inherently parallel. When each solution of (11) in Algorithm 1 is approximated by an alternating iterative method, this is called an alternating two-stage method. In particular, each solution of (11) could be approximated by the Symmetric Block Gauss-Seidel method. In order to describe the parallel algorithms explored in this paper, we assume that there are p processes with r_ℓ blocks assigned to process ℓ, $\ell = 1, \ldots, p$. In Step 3 of Algorithm 3 (PAGS), each solution of (12) and (13) is approximated using Gauss-Seidel iterations while in Algorithm 4 (PALU), these solutions are obtained using LU factorizations. We have used as the global stop-

---

Algorithm 3: Parallel Alternating Algorithm with GS inner iterations - PAGS

Divide the r blocks of (4) into p groups, each assigned to a different process ℓ;
1. Perform parallel BJ (with $q = p$ in Algorithm 1), i.e., each process approximates the solution of one linear system (11) (This is the outer iteration);
2. Each solution of (11) in Step 1 is approximated using t steps of SBGS (with $q = r_\ell$ in Algorithm 2);
3. Each solution of (12) and (13) in Step 2 is approximated by a fixed number m of Gauss-Seidel (GS) iterations;

---

ping criterion (i.e., in the outer iteration) $\|Ax^{(l)} - b\|_2 < \varepsilon$ and to guarantee convergence, for the singular linear systems, we use the customary device of shifting the iteration matrix from T to $T_\delta = \delta T + (1 - \delta)I$, where $0 < \delta < 1$. This shift is performed at the end of each outer iteration. This is also the point of the computation where the vector $x^{(l)}$ is normalized. In the experiments reported in this section we have used $\varepsilon = 10^{-6}$ and $\delta = 0.95$. We note that, the computation of both the stopping criterion and the normalization is also performed in parallel with the need of an AllReduce operation at the end. Additionally, the shift T_δ is also computed by all of the processes before an AllGatherv operation is accomplished.

19

Algorithm 4: Parallel Alternating Algorithm using LU factorizations - PALU

Divide the r blocks of (4) into p groups, each assigned to a different process ℓ;

1. Perform parallel BJ (with $q = p$ in Algorithm 1), i.e., each process approximates the solution of one linear system (11) (This is the outer iteration);

2. Each solution of (11) in Step 1 is approximated using t steps of SBGS (with $q = r_\ell$ in Algorithm 2);

3. Each solution of (12) and (13) in Step 2 is obtained using LU factorizations;

Matrix	Number of processes	Row-wise distribution
LAPLACE	4	25000×4
NCD	8	$((37828, 37827) \times 3, 37828 \times 2)$
QNATM1	4	523269×4
	8	$(261634, 261635) \times 4$
	12	174423×12
	16	$(130817 \times 3, 130818) \times 4$
QNATM2	4	1178117×4
	8	$(589058, 589059) \times 4$
	12	$(392706 \times 2, 392705) \times 4$
	16	$(294529 \times 3, 294530) \times 4$
	24	$(196353 \times 5, 196352) \times 4$

Table 1: Row-wise distribution for the test matrices.

We have implemented the algorithms described here on an HPC cluster of 26 nodes HP Proliant SL390s G7 connected through a network of low-latency QDR Infiniband-based. Each node consists of two Intel XEON X5660 hexacore at up to 2.8 GHz and 12MB cache per processor, with 48 GB of RAM. The operating system is CentOS Linux 5.6 for x86 64 bit.

We have run our codes using several test matrices. The singular models are taken from [20]. The first model is a multi-class, finite-buffer, priority queuing network model with applicability to telecommunications modelling. The matrices we use are of order 2093076 and 4712468 and we label them QNATM1 and QNATM2, respectively. The QNATM1 matrix has 14120056 nonzero entries and the QNATM2 matrix has 31796856 nonzero entries. The second model represents the system architecture of a time-shared, multiprogrammed, paged, virtual memory computer. The resulting nearly completely decomposable (NCD) matrix corresponds to 120 users. This matrix is of order 302621 and has 2074061 nonzero elements; see [21] for more details. In the nonsingular case, we discuss the results for a Laplace matrix of size 100000. This matrix, labelled as LAPLACE, has 1000×1000 blocks of size 100×100.

A sparse matrix format is considered in order to store the coefficient matrices and all block structures needed in the algorithms. Concretely, the Compressed Sparse Row (CSR) format was used, which is one of the most extensively used storage schemes for general sparse matrices, with minimal storage requirements.

We have used equal block sizes obtained from the original matrix. These equal block sizes were obtained in the following way: firstly, we consider a balanced block diagonal structure depending on the number of processes (see Table 1). That is a row-wise distribution is chosen, where each process gets, if possible, the same amount of consecutive rows. If the number of rows is not a multiple of the number of processes there are some processes with one more row. Hence, we construct, at each diagonal block, blocks of a predetermined and constant size $n_i = \eta$. If the order of a diagonal block is not a multiple of η, the last block is of order greater than η. Figure 1(a) shows this block structure for an intermediate process.

Algorithms 3 and 4 use a row domain decomposition, in which each subdomain consists of several consecutive rows of A without overlapping. Alternatively, we can consider that every subdomain has an overlapping in the following way: the number of rows of the original non-overlapped subdomains is incremented by $2s$ rows, s rows in the upper part and s rows in the lower part; the first and last subdomains are incremented by $2s$ rows in

(a) Without overlapping.
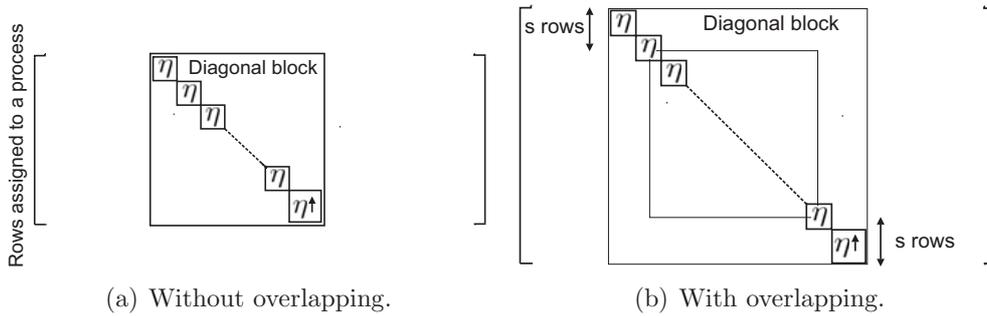
(b) With overlapping.

Figure 1: Row domain decomposition for an intermediate process.

(a) QNATM1 matrix.

	4	8	12	16
■ 1 core per node	84.3%	118.2%	111.8%	112.1%
■ 2 cores per node	122.1%	117.2%	113.3%	107.6%
☐ 4 cores per node	113.4%	108.8%	103.4%	97.8%
☐ 12 cores per node		67.6%		

(b) QNATM2 matrix.

	4	8	12	16	24
■ 1 core per node	88.8%	102.4%	84.7%	96.8%	79.9%
■ 2 cores per node	106.1%	96.8%	96.0%	92.5%	82.8%
☐ 4 cores per node	99.2%	91.7%	89.5%	86.4%	78.5%
☐ 12 cores per node		58.8%			50.8%
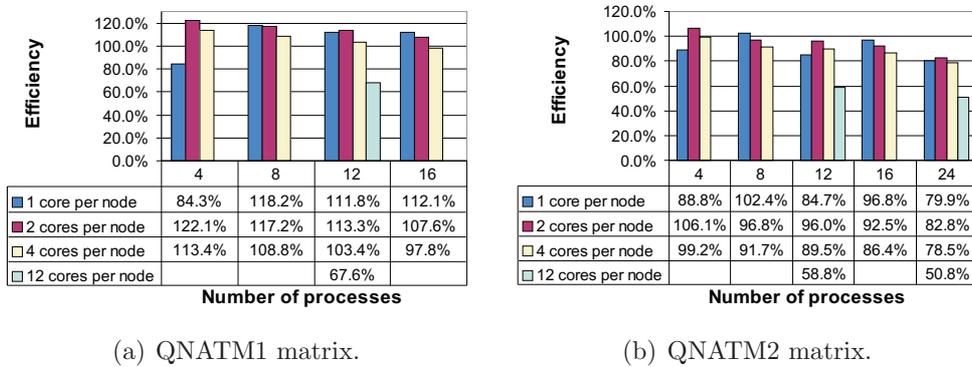
Figure 2: Efficiency PAGS Algorithm. QNATM matrices, $m = 2$, $t = 10$, $\eta = 50$.

the lower part and upper part, respectively. In each overlapping subdomain, calculations are performed independently of each other in the framework of a Block Jacobi iteration. Figure 1(b) shows this overlapping subdomain for an intermediate process. After each Block Jacobi iteration, the overlapping components of the current approximation are assembled in average way. In this case, the AllGatherv operation is performed over a vector of size $n+2sp$. After that operation, each process compacts this vector in another vector of size n taking into account that the overlapping components are assembled in average way. Moreover, using the overlapping pattern and that the iterative vectors are nonnegative, it is possible to compute the normalization in parallel as in the non-overlapped case. We have called the overlapped versions of Algorithms 3 and 4, OPAGS and OPALU Algorithms, respectively.

22

	t=1	t=2	t=5	t=10	t=15	t=20	t=25
PAGS m=1	34.12	23.05	16.58	15.00	15.11	15.75	16.68
PAGS m=2	29.84	21.25	16.45	15.76	16.43	17.52	18.93
PAGS m=3	31.15	23.04	18.58	18.25	19.41	20.92	22.75
PAGS m=4	33.77	25.66	21.37	21.34	22.66	24.60	27.01
PALU	24.03	15.80	10.98	9.97	10.29	10.92	11.83
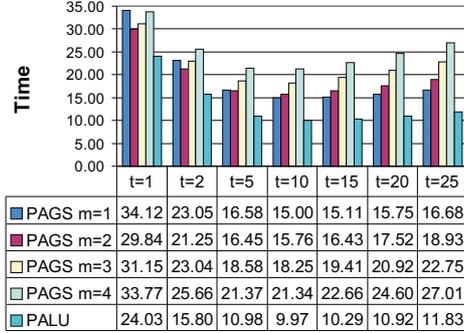
(a) QNATM1 matrix, 16 (16 × 1) nodes.

	t=1	t=2	t=5	t=10	t=15	t=20	t=25
PAGS m=1	110.5	68.81	44.44	38.16	40.69	41.85	43.91
PAGS m=2	98.73	64.56	45.00	40.68	41.59	43.81	46.72
PAGS m=3	98.01	66.26	48.40	45.17	46.83	49.94	53.80
PAGS m=4	101.5	70.59	53.41	50.99	53.51	57.38	61.87
PALU	81.39	50.75	33.03	28.80	28.89	30.45	32.55
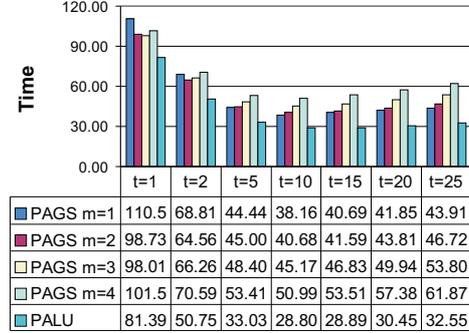
(b) QNATM2 matrix, 24 (24 × 1) nodes.

Figure 3: PAGS versus PALU Algorithms. QNATM matrices, $\eta = 50$. Distributed memory.

	PAGS t=15 - m=1	PAGS t=10 - m=2	PALU t=15	PALU t=10
8x1	29.23	29.90	20.77	20.51
4x2	29.43	30.15	24.49	24.31
2x4	33.12	32.47	30.10	29.64
1x8	44.59	39.86	38.44	38.12

(a) QNATM1 matrix, 8 processes.

	PAGS t=15 - m=1	PAGS t=10 - m=2	PALU t=10	PALU t=15
24x1	40.69	40.68	28.80	28.89
12x2	37.86	39.23	30.38	30.14
6x4	41.91	41.37	40.16	40.84
3x8	56.29	51.34	42.88	43.26
2x12	74.43	63.97	57.38	58.16

(b) QNATM2 matrix, 24 processes.

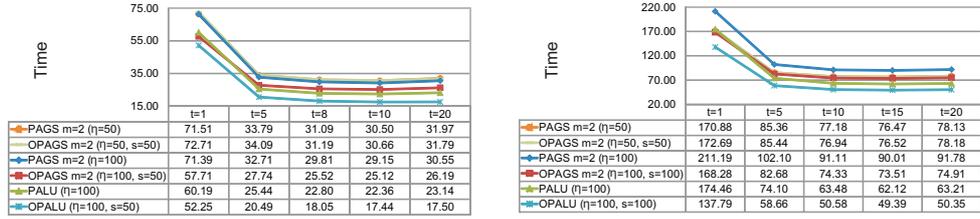Figure 4: PAGS versus PALU Algorithms. QNATM matrices, $\eta = 50$. Distributed memory versus shared memory.

	t=1	t=5	t=8	t=10	t=20
PAGS m=2 (η=50)	71.51	33.79	31.09	30.50	31.97
OPAGS m=2 (η=50, s=50)	72.71	34.09	31.19	30.66	31.79
PAGS m=2 (η=100)	71.39	32.71	29.81	29.15	30.55
OPAGS m=2 (η=100, s=50)	57.71	27.74	25.52	25.12	26.19
PALU (η=100)	60.19	25.44	22.80	22.36	23.14
OPALU (η=100, s=50)	52.25	20.49	18.05	17.44	17.50

(a) QNATM1 matrix, 12 (3 × 4) processes.

	t=1	t=5	t=10	t=15	t=20
PAGS m=2 (η=50)	170.88	85.36	77.18	76.47	78.13
OPAGS m=2 (η=50, s=50)	172.69	85.44	76.94	76.52	78.18
PAGS m=2 (η=100)	211.19	102.10	91.11	90.01	91.78
OPAGS m=2 (η=100, s=100)	168.28	82.68	74.33	73.51	74.91
PALU (η=100)	174.46	74.10	63.48	62.12	63.21
OPALU (η=100, s=100)	137.79	58.66	50.58	49.39	50.35

(b) QNATM2 matrix, 12 (4 × 3) processes.

Figure 5: Overlapping versus non-overlapping Algorithms. QNATM matrices. Distributed shared memory.

	t=1	t=3	t=5	t=8	t=10
PAGS m=2 (η=50)	8153	2794	1746	1178	997
PAGS m=2 (η=100)	8051	2761	1727	1167	988
PAGS m=2 (η=400)	7443	2569	1620	1107	944
PALU (η=50)	6618	2291	1454	1005	863
PALU (η=100)	6456	2239	1424	987	850
PALU (η=400)	2740	1034	738	596	554

(a) Number of iterations.

	t=1	t=3	t=5	t=8	t=10
PAGS m=2 (η=50)	18.05	14.41	14.14	14.74	15.41
PAGS m=2 (η=100)	17.91	14.19	13.92	14.51	15.16
PAGS m=2 (η=400)	15.79	12.54	12.38	13.04	13.72
PALU (η=50)	11.59	8.57	8.34	8.76	9.24
PALU (η=100)	11.15	8.29	8.09	8.53	9.02
PALU (η=400)	58.43	51.03	60.17	77.36	89.75

(b) Times.

Figure 6: PAGS versus PALU Algorithms. Laplace matrix. Distributed memory, 4 (4 × 1) nodes.

	t=1	t=3	t=5	t=8	t=10	t=20
PAGS m=1 (η=100)	21.14	14.73	13.67	13.43	13.57	15.55
OPAGS m=1 (η=100, s=50)	21.00	14.57	13.35	12.79	12.67	13.06
OPAGS m=1 (η=100, s=100)	20.91	14.54	13.28	12.66	12.49	12.49

(a) PAGS versus OPAGS Algorithms.

	t=1	t=3	t=5	t=8	t=10	t=20
PALU (η=100)	11.15	8.29	8.09	8.53	9.02	12.30
OPALU (η=100, s=50)	11.28	8.18	7.68	7.60	7.69	8.73
OPALU (η=100, s=100)	11.15	8.02	7.46	7.25	7.26	7.71

(b) PALU versus OPALU Algorithms.

Figure 7: Overlapping versus non-overlapping Algorithms. Laplace matrix. Distributed memory, 4 (4 × 1) nodes.

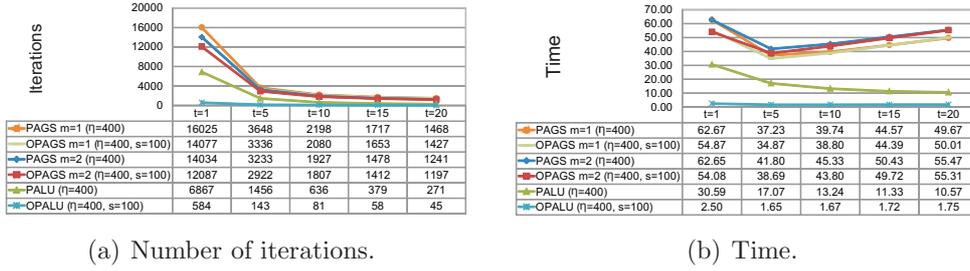|	t=1	t=5	t=10	t=15	t=20
PAGS m=1 (η=400)	16025	3648	2198	1717	1468
OPAGS m=1 (η=400, s=100)	14077	3336	2080	1653	1427
PAGS m=2 (η=400)	14034	3233	1927	1478	1241
OPAGS m=2 (η=400, s=100)	12087	2922	1807	1412	1197
PALU (η=400)	6867	1456	636	379	271
OPALU (η=400, s=100)	584	143	81	58	45

(a) Number of iterations.

	t=1	t=5	t=10	t=15	t=20
PAGS m=1 (η=400)	62.67	37.23	39.74	44.57	49.67
OPAGS m=1 (η=400, s=100)	54.87	34.87	38.80	44.39	50.01
PAGS m=2 (η=400)	62.65	41.80	45.33	50.43	55.47
OPAGS m=2 (η=400, s=100)	54.08	38.69	43.80	49.72	55.31
PALU (η=400)	30.59	17.07	13.24	11.33	10.57
OPALU (η=400, s=100)	2.50	1.65	1.67	1.72	1.75

(b) Time.

Figure 8: Overlapping versus non-overlapping Algorithms. NCD matrix. Distributed memory, 8 (8×1) nodes.

The parallel environment has been managed using MPI (Message Passing Interface) [22]. Moreover, the BLAS routines [23] for vector computations and the SPARSKIT routines [24] for handling sparse matrices have been used. The algorithms have been implemented and tested on distributed and shared memory, and using a hybrid distributed shared memory model. Concretely, each process is assigned to a core as follows: let p be the number of physical cores used, $p = d \times c$ indicates that d nodes of the parallel platform have been used and for each one of these nodes, c cores have been considered. Therefore, we use a philosophy of distributed shared memory using $p = d \times c$ processes or threads. Particularly, if $d = 1$, the algorithms are executed in shared memory using $p = c$ threads on a single node. Conversely, if $c = 1$, we are working on distributed memory using $p = d$ nodes. Figure 2 illustrates the efficiency of the PAGS Algorithm (Algorithm 3) for the QNATM matrices, using diagonal blocks of size $\eta = 50$ and varying the number of processes and the number of cores per node. In order to calculate the efficiency of these parallel algorithms, we have used as reference sequential algorithm the best sequential algorithm obtained from the best execution of Algorithm 3, using $p = 1$. Note that Algorithm 3 differs for each value of p, that is, different sequences of iterative vectors are obtained for several values of the number of processes p; in this sense, the number of iterations of the parallel executions of Figure 2(a) differs from the number of iterations of the sequential algorithm. Therefore superlinear efficiency can be obtained in certain cases.

Furthermore, the speed-up does not decrease significantly as the number of processes p is increased, i.e., we obtain a good scalability and efficiency of the algorithm. Generally, the best parallel results have been obtained using from 1 to 4 cores in each node. Moreover, it seems inappropriate to use 8

or more cores on a single node. Note that many processing cores sharing the same system bus and memory bandwidth limits the real performance advantage.

Figures 3 and 4 compare the behavior of the parallel PAGS and PALU Algorithms for the QNATM1 and QNATM2 matrices. We note that, for all tested models, the best parallel results for the PAGS Algorithm are obtained for $m = 1$ or $m = 2$, regardless of the number of processes. Note that the execution time of the PAGS and PALU Algorithms is reduced as the number of SBGS steps, t, increases up to an optimal value of t after which the time starts to increase. This behavior is characteristic of two-stage methods and although the optimal value of t is difficult to predict, a good choice of it is one which balances the realization of more inner updates with the decrease of the global iterations and its associated computational cost. For the QNATM models discussed here, the experiments lead us to conclude that a good choice of t for the three philosophies of implementation is $t = 10$. Figure 5 compares the execution time of the algorithms treated here with and without overlapping, for different sizes of blocks. For the QNATM matrices, we have obtained the best results for the OPALU Algorithm using blocks of size $\eta = 100$ and a relatively small overlapping compared with the size of the matrix ($s = 50$ for the QNATM1 matrix and $s = 100$ for the QNATM2 matrix). Our experience indicates that, for these matrices, good choices of t for both overlapping algorithms are between $10 - 15$, obtaining similar execution times. In Figures 6 and 7 we report results for the Laplace matrix. Similar results to those obtained for the QNATM matrices were obtained. Note that, for each value of t, the timings of the PALU Algorithm are generally lower than those of the PAGS Algorithm, obtaining the greatest saving time of the parallel PALU Algorithms in relation to the parallel PAGS Algorithms when only one core is used in each node, that is, using distributed memory. However if the size of blocks is relatively big with respect to the size of the matrix, the fill-in caused by the LU factorizations, in the sparse diagonal blocks, increases the execution time of the PALU Algorithm; see Figure 6 ($\eta = 400$). On the other hand, for the Laplace matrix, the saving time obtained with the overlapping algorithms in relation to the non-overlapping algorithms is poor when optimal values of t are considered (about 7% in Figure 7(a) and 10.4% in Figure 7(b)). This is due to the fact that the lower number of iterations achieved adding overlap is paid by a higher computational cost, in relation to the small global cost of the algorithms for the Laplace matrix. Let us highlight some observations about the results of these algorithms. In gen-

eral, usually larger overlap leads to faster convergence up to a certain point where increasing the overlap does not further improve the convergence rate. But, generally, the amount of overlap needed to minimize the total solving time is problem specific. As it can be seen in Figure 8, for the NCD matrix, the OPALU Algorithm with $\eta = 400$ and $s = 100$ reduces tremendously the number of iterations and the total time to solve the linear system in relation to the other algorithms analyzed in this paper.

5. Conclusions

In this paper we have studied the problem of solving large consistent linear systems by means of parallel alternating two-stage algorithms with and without overlapping. These algorithms have been applied to both singular and nonsingular large linear systems. In the nonsingular case, the problem to be solved comes from the discretization of the Laplace's equation while in the singular case the test problems arise from Markov chain modelling. The algorithms have been implemented and tested on distributed and shared memory, and using a distributed shared memory model, obtaining a good scalability and efficiency. Generally, the PALU algorithms behave better than the PAGS algorithms. On the other hand, the overlapping algorithms have sped up the convergence time of the non-overlapping algorithms. The amount of overlap needed to improve the convergence rate is problem specific and depends on the characteristics of the matrix and the block diagonal structure considered in the corresponding parallel algorithm.

Acknowledgements

References

[1] Frommer A, Szyld DB. H-splittings and two-stage iterative methods. Numer Math 1992;63:345–356.

[2] Lanzkron PJ, Rose DJ, Szyld DB. Convergence of nested classical iterative methods for linear systems. Numer Math 1991;58:685–702.

[3] Migallón H, Migallón V, Penadés J. Alternating two-stage methods for consistent linear systems with applications to the parallel solution of Markov chains. Adv Eng Softw 2010;41:13–21.

[4] Fritzsche D, Frommer A, Shank SD, Szyld DB. Overlapping blocks by growing a partition with applications to preconditioning. SIAM J Sci Comput 2013;35(1):A453–A473.

[5] Jones MT, Szyld DB. Two-stage multisplitting methods with overlapping blocks. Numer Linear Algebr Appl 1996;2:113–124.

[6] Bru R, Pedroche F, Szyld DB, Additive Schwarz iterations for Markov chains. SIAM J Matrix Anal Appl 2005;27(2):445–458.

[7] Wang G, Zhang N, Tan F. Parallel alternating two-stage methods for solving linear system. World Academy of Science, Engineering and Technology, International Science Index 2010;4(7):953–955.

[8] Wang G, Li X. Some results on parallel alternating two-stage methods. World Academy of Science, Engineering and Technology, International Science Index 2011;5(8):1244–1247.

[9] Migallón H, Migallón V, Penadés J. Block alternating two-stage methods for singular systems. In: Iványi P, Topping BHV, editors. Proceedings of the Ninth International Conference on Engineering Computational Technology. Stirlingshire, United Kingdom: Civil-Comp Press; 2014, Paper 30. doi:doi:10.4203/ccp.105.30.

[10] Benzi M, Szyld DB. Existence and uniqueness of splittings for stationary iterative methods with applications to alternating methods. Numer Math 1997;76:309–321.

[11] Neumann M, Plemmons RJ. Convergence of parallel multisplitting iterative methods for M-matrices. Linear Alg Appl 1987;88/89:559–573.

[12] Berman A, Plemmons RJ. Nonnegative Matrices in the Mathematical Sciences, 3rd ed. New York: Academic Press; 1979, reprinted by SIAM, Philadelphia, 1994.

[13] Migallón V, Penadés J, Szyld DB. Block two-stage methods for singular systems and Markov chains. Numer Linear Algebr Appl 1996;3(5):413–426.

[14] Castel MJ, Migallón V, Penadés J. Convergence of non-stationary parallel multisplitting methods for hermitian positive definite matrices. Math Comput 1998;67(221):209–220.

[15] Frommer A, Szyld DB. Weighted max norms, splittings, and overlapping additive Schwarz iterations. Numer Math 1999;3:259–278.

[16] Bru R, Elsner L, Neumann M. Models of parallel chaotic iteration methods. Linear Alg Appl 1988;103:175–192.

[17] Plemmons RJ. M-matrices leading to semiconvergent splittings. Linear Alg Appl 1976;15:243–252.

[18] Alefeld G, Schneider N. On square roots of M-matrices. Linear Alg Appl 1982;42:119–132.

[19] Bru R, Migallón V, Penadés J, Szyld DB. Parallel, synchronous and asynchronous two-stage multisplitting methods. Electron Trans Numer Anal 1995;3:24–38.

[20] Stewart WJ. Marca models: A collection of Markov chain models. Available online at http://www.csc.ncsu.edu/faculty/WStewart/MARCA_Models/MARCA_Models.html.

[21] Philippe B, Saad Y, Stewart WJ. Numerical methods in Markov chain modelling. Oper Res 1992;40:1156–1179.

[22] Dongarra J, Huss-Lederman S, Otto S, Snir M, Walkel D. MPI: The complete reference. 2nd ed. Cambridge, MA: The MIT Press; 1998.

[23] NETLIB. BLAS: Basic Linear Algebra Subprograms, http://www.netlib.org/index.html.

[24] Saad Y. SPARSKIT: A basic tool kit for sparse matrix computation. http://www-users.cs.umn.edu/~saad/software/SPARSKIT/sparskit.html.