Advances in Engineering Software 41 (2010) 13–21

Contents lists available at ScienceDirect

Advances in Engineering Software

journal homepage: www.elsevier.com/locate/advengsoft

Alternating two-stage methods for consistent linear systems with applications to the parallel solution of Markov chains

Héctor Migallón [a], Violeta Migallón [b,*], Jose Penadés [b]

[a] Departamento de Física y Arquitectura de Computadores, Universidad Miguel Hernández, E-03202 Elche, Alicante, Spain
[b] Departamento de Ciencia de la Computación e Inteligencia Artificial, Universidad de Alicante, E-03071 Alicante, Spain

ARTICLE INFO

Article history:
Received 9 October 2007
Accepted 12 January 2008
Available online 25 February 2009

Keywords:
Alternating two-stage methods
Parallel algorithms
Linear systems
Block methods
Markov chains

ABSTRACT

Two-stage methods in which the inner iterations are accomplished by an alternating method are developed. Convergence of these methods is shown in the context of solving singular and nonsingular linear systems. These methods are suitable for parallel computation. Experiments related to finding stationary probability distribution of Markov chains are performed. These experiments demonstrate that the parallel implementation of these methods can solve singular systems of linear equations in substantially less time than the sequential counterparts.

1. Introduction

Consider the $n \times n$ linear system

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

where A is a matrix such that b is in $\mathscr{R}(A)$, the range of A.

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

$$x^{(l+1)} = M^{-1}Nx^{(l)} + M^{-1}b, \quad l = 0, 1, \ldots, \tag{2}$$

where $M^{-1}N$ is called the iteration matrix of the method. On the other hand, a two-stage method consists of approximating the linear system (2) by using another iterative procedure (inner iterations). That is, consider the splitting $M = F - G$ and perform, at each outer step $l, s(l)$ inner iterations of the iterative procedure induced by this splitting. Thus, the resulting method is

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

cf. [1]. Two-stage iterative methods have been studied, e.g., in [2–5]. In this paper, a two-stage iterative process is developed for the solution of the linear system (1), where at each outer iteration $l, l = 0, 1, \ldots$, the linear system (2) is approximated by using an alternating iterative procedure. More specifically, let $M = P - Q = R - S$ be two splittings of the matrix M. In order to approximate the linear system (2), for each $l, l = 0, 1, \ldots$, we perform $s(l)$ inner iterations of the general class of iterative methods of the form

$$z^{(k+\frac{1}{2})} = P^{-1}Qz^{(k)} + P^{-1}(Nx^{(l)} + b),$$
$$z^{(k+1)} = R^{-1}Sz^{(k+\frac{1}{2})} + R^{-1}(Nx^{(l)} + b), \quad k = 0, 1, \ldots, s(l) - 1$$

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

$$z^{(k+1)} = R^{-1}SP^{-1}Qz^{(k)} + R^{-1}(SP^{-1} + I)(Nx^{(l)} + b),$$
$$k = 0, 1, \ldots, s(l) - 1.$$

Thus, the alternating two-stage method can be written as follows

$$x^{(l+1)} = (R^{-1}SP^{-1}Q)^{s(l)}x^{(l)} + \sum_{j=0}^{s(l)-1}(R^{-1}SP^{-1}Q)^jR^{-1}(SP^{-1} + I)(Nx^{(l)} + b),$$
$$l = 0, 1, \ldots. \tag{4}$$

In a similar manner as the two-stage methods, we say that an alternating two-stage method is stationary when $s(l) = s$, for all l, while an alternating two-stage method is non-stationary if the number of inner iterations changes with the outer iteration l.

Clearly, given an initial vector $x^{(0)}$, the alternating two-stage iterative method (4) produces the sequence of vectors

$$x^{(l+1)} = T^{(l)}x^{(l)} + c_{s(l)}, \quad l = 0, 1, \ldots, \tag{5}$$

where

$$T^{(l)} = (R^{-1}SP^{-1}Q)^{s(l)} + \sum_{j=0}^{s(l)-1}(R^{-1}SP^{-1}Q)^jR^{-1}(SP^{-1} + I)N, \tag{6}$$

and $c_{s(l)} = \sum_{j=0}^{s(l)-1}(R^{-1}SP^{-1}Q)^jR^{-1}(SP^{-1} + I)b$.

* Corresponding author. Tel.: +34 965 903900; fax: +34 965 903902.
E-mail address: violeta@dccia.ua.es (V. Migallón).

0965-9978/$ - see front matter © 2009 Civil-Comp Ltd. and Elsevier Ltd. All rights reserved.
doi:10.1016/j.advengsoft.2008.12.021

In order to analyze the convergence of the alternating two-stage method (5) and taking into account that $A = M - N$ and $M = P - Q = R - S$, the iteration matrices $T^{(l)}, l = 0, 1, \ldots,$ defined in (6), are written as follows:

$$\begin{aligned} T^{(l)} &= (R^{-1}SP^{-1}Q)^{s(l)} + \sum_{j=0}^{s(l)-1} (R^{-1}SP^{-1}Q)^j R^{-1}(SP^{-1}+I)N \\ &= (R^{-1}SP^{-1}Q)^{s(l)} + \sum_{j=0}^{s(l)-1} (R^{-1}SP^{-1}Q)^j R^{-1}(SP^{-1}+I)(P-Q)M^{-1}N \\ &= (R^{-1}SP^{-1}Q)^{s(l)} + \sum_{j=0}^{s(l)-1} (R^{-1}SP^{-1}Q)^j R^{-1}S(I-P^{-1}Q)M^{-1}N \\ &\quad + \sum_{j=0}^{s(l)-1} (R^{-1}SP^{-1}Q)^j (I-R^{-1}S)M^{-1}N = (R^{-1}SP^{-1}Q)^{s(l)} \\ &\quad + (I-(R^{-1}SP^{-1}Q)^{s(l)})M^{-1}N, \quad l = 0, 1, \ldots \end{aligned} \tag{7}$$

In this paper, our study concentrates on these alternating two-stage methods. Specifically, in Section 3, we give convergence results of these methods for nonsingular linear systems, when the matrix A of the linear system is monotone, H-matrix or Hermitian positive definite. In Section 4, we also prove the convergence of these methods for consistent singular linear systems, when M-matrices or symmetric positive semidefinite matrices are considered. In Section 5, we explore the use of parallel implementation of these alternating two-stage methods for the solution of Markov chains. Previously, in Section 2, we present some definitions and preliminaries that are used later in the paper.

2. Notation and preliminaries

The notation and terminology adopted in this paper are along the lines of those used by Berman and Plemmons [6]. We say that a vector x is nonnegative (positive), denoted $x \geqslant 0 (x > 0)$, if all of its entries are nonnegative (positive). Similarly, a matrix B is said to be nonnegative, denoted $B \geqslant O$ (where O is the zero matrix), if all its entries are nonnegative. Given a matrix $A = (a_{ij})$, we define the matrix $|A| = (|a_{ij}|)$. It follows that $|A| \geqslant O$ and that $|AB| \leqslant |A||B|$ for any two matrices A and B of compatible size. By $\rho(A)$ we denote the spectral radius of the square matrix A. A general matrix A is called an M-matrix if A can be expressed as $A = sI - B$, with $B \geqslant O$, $s > 0$, and $\rho(B) \leqslant s$. The M-matrix A is singular when $s = \rho(B)$. The M-matrix A is 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 nonsingular matrix $A \in Z^{n\times n}$ is an M-matrix if and only if A is a monotone matrix ($A^{-1} \geqslant O$). For any matrix $A = (a_{ij}) \in \Re^{n\times n}$, we define its comparison matrix $\langle A\rangle = (\alpha_{ij})$ by $\alpha_{ii} = |a_{ii}|, \alpha_{ij} = -|a_{ij}|, i \neq j$. A nonsingular matrix A is said to be an H-matrix if $\langle A\rangle$ is an M-matrix.

Lemma 1 [7,8]. *Let $A, B \in \Re^{n\times n}$.*

(a) *If A is an H-matrix, then $|A^{-1}| \leqslant \langle A\rangle^{-1}$.*
(b) *If $|A| \leqslant B$ then $\rho(A) \leqslant \rho(B)$.*

Definition 2 [6,2,9]). Let $A \in \Re^{n\times n}$. A splitting $A = M - N$ is called

(a) regular if $M^{-1} \geqslant O$ and $N \geqslant O$,
(b) weak regular if $M^{-1} \geqslant O$ and $M^{-1}N \geqslant O$,
(c) *H-splitting* if $\langle M\rangle - |N|$ is a nonsingular M-matrix, and
(d) *H-compatible* splitting if $\langle A\rangle = \langle M\rangle - |N|$.

Lemma 3 [3]. *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 3, 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].

Lemma 4 [6,2]. *Let $A = M - N$ be a splitting.*

(a) *If the splitting is weak regular, then $\rho(M^{-1}N) < 1$ if and only if $A^{-1} \geqslant O$.*
(b) *If the splitting is an H-splitting, then A and M are H-matrices and $\rho(M^{-1}N) \leqslant \rho(\langle M\rangle^{-1}|N|) < 1$.*
(c) *If the splitting is an H-compatible splitting and A is an H-matrix, then it is an H-splitting and thus convergent.*

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^HAx is positive (nonnegative), for all complex $x \neq 0$. When A is Hermitian, this is equivalent to requiring that $x^HAx > 0 (x^HAx \geqslant 0)$, 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.

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 *ind*T, is the smallest nonnegative integer k such that $\mathscr{R}(T^{k+1}) = \mathscr{R}(T^k)$. By ind_1T we denote the *index* associated with the value one, i.e., $ind_1T = ind(I - T)$. Note that when $\rho(T) = 1$, $ind_1T \leqslant 1$ if and only if $ind_1T = 1$. 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 $\mathscr{N}(T)$ 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.

Theorem 5 [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) $\mathscr{N}(I - T) \oplus \mathscr{R}(I - T) = \Re^n$.

Condition (b) is equivalent to the existence of the group inverse $(I - T)^{\#}$, and it is also equivalent to having $ind_1T = 1$; see, e.g., [6]. We review in what follows the definition of some generalized inverses.

Definition 6 [6]. *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.

We would like to note that the group inverse $A^{\#}$ of a matrix A, if it exists, is unique. When A is nonsingular, each generalized inverse coincides with A^{-1}.

Theorem 7 [6]. *Let $T \in \Re^{n\times n}$, with $T \geqslant O$, and let C be a $\{1,2\}$-inverse of $I-T$ with $\mathscr{R}(C)$ complementary to $\mathscr{N}(I-T)$, such that C is nonnegative on $\mathscr{R}(I-T)$, i.e., the matrix C satisfies the following conditions.*

(i) $I-T=(I-T)C(I-T)$,
(ii) $C=C(I-T)C$,
(iii) $\mathscr{N}(I-T)\oplus\mathscr{R}(C)=\Re^n$,
(iv) *If* $x\in\mathscr{R}(I-T), x\geqslant 0$ *then* $Cx\geqslant 0$.

Then, $\rho(T)\leqslant 1$, and $ind_1(T)\leqslant 1$.

Lemma 8 [6]. *Let $T\in\Re^{n\times n}$ be semiconvergent. Then*

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

Definition 9 [6]. *A general M-matrix A is said to have property c if for some representation of $A=sI-B, s>0, B\geqslant O$, the matrix $s^{-1}B$ is semiconvergent.*

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

Theorem 10 [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 $\rho(T)\leqslant 1$, and $\mathscr{N}(I-T)\oplus\mathscr{R}(I-T)=\Re^n$.*

Theorem 11 [6]. *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.*

3. Convergence for nonsingular linear systems

Consider that A is a nonsingular matrix, and let x^* be the exact solution of (1) and let $\epsilon^{(l+1)}=x^{(l+1)}-x^*$ be the error at the $l+1$ iteration of the alternating two-stage iterative method (4). It is easy to prove that x^* is a fixed point of (5). Thus

$$\epsilon^{(l+1)} = T^{(l)}\epsilon^{(l)} = \cdots = T^{(l)}T^{(l-1)}\cdots T^{(0)}\epsilon^{(0)}, \quad l=0,1,\ldots.$$

Thus, the sequence of error vectors $\{\epsilon^{(l)}\}_{l=0}^{\infty}$ generated by the iteration (5) converges to the vector 0 if, and only if, $\lim_{l\to\infty}T^{(l)}T^{(l-1)}\cdots T^{(0)}=O$.

Lemma 12 [12]. *Let $A^{(l)}, l=0,1,\ldots,$ be a sequence of nonnegative matrices in $\Re^{n\times n}$. If there exist a real number $0\leqslant\theta<1$, and a vector $v>0$ in $\Re^n$, such that $A^{(l)}v\leqslant\theta v$, for all $l=0,1,\ldots,$ then $\rho(V^j)\leqslant\theta^j<1$, where $V_j=A^{(j)}\cdots A^{(1)}A^{(0)}$, and $\lim_{j\to\infty}V^j=O$.*

Theorem 13 [10]. *Let A be a nonsingular matrix such that $A^{-1}\geqslant O$. 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 $T=B^{-1}C$.*

Theorem 14 [2]. *Let $A=M-N$ be a convergent regular splitting, and let $M=F-G$ be a convergent weak regular splitting. Then, the two-stage iterative method (3) converges to the solution of the linear system (1), for any initial vector $x^{(0)}$ and for any sequence of inner iterations $s(l)\geqslant 1, l=0,1,\ldots$.*

Theorem 15. *Let $A^{-1}\geqslant O$. Consider the splitting $A=M-N$ is regular and the splittings $M=P-Q=R-S$ are weak regular. Then, the alternating two-stage method (5) converges to the solution of the linear system (1), for any initial vector $x^{(0)}$ and for any sequence of inner iterations $s(l)\geqslant 1, l=0,1,\ldots$.*

Proof. Since $A=M-N$ is a regular splitting, $M^{-1}\geqslant O$, then from Theorem 13 there exists a unique pair of matrices B, C, such that $R^{-1}SP^{-1}Q=B^{-1}C$ and $M=B-C$ is a weak regular splitting. That is, the iteration matrices defined in (7) can be written as $T^{(l)}=(B^{-1}C)^{s(l)}+(I-(B^{-1}C)^{s(l)})M^{-1}N, l=0,1,\ldots$

Thus, $T^{(l)}, l=0,1,\ldots,$ are the iteration matrices of a non-stationary two-stage method for the matrix A, with the regular splitting $A=M-N$, and $M=B-C$ being a weak regular splitting. Therefore, using Lemma 4(a) and Theorem 14, the proof is complete. □

Now we study the convergence of the alternating two-stage method (5) when A is an H-matrix and therefore not necessarily a monotone matrix. In the following theorem the fact that A is an H-matrix follows from Lemma 4(b).

Theorem 16. *Let $A=M-N$ be an H-splitting and let $M=P-Q=R-S$ be H-compatible splittings. Then, the alternating two-stage method (5) converges to the solution of the linear system (1), for any initial vector $x^{(0)}$ and any sequence of inner iterations $s(l)\geqslant 1, l=0,1,\ldots$.*

Proof. By Lemma 4(b) and (c), the matrices P and R are H-matrices. We use Lemma 1(a) to obtain the following bounds from (6).

$$|T^{(l)}| \leqslant (|R^{-1}||S||P^{-1}||Q|)^{s(l)} + \sum_{j=0}^{s(l)-1}(|R^{-1}||S||P^{-1}||Q|)^j|R^{-1}|(|S||P^{-1}|+I)|N|$$
$$\leqslant (\langle R\rangle^{-1}|S|\langle P\rangle^{-1}|Q|)^{s(l)} + \sum_{j=0}^{s(l)-1}(\langle R\rangle^{-1}|S|\langle P\rangle^{-1}|Q|)^j\langle R\rangle^{-1}(|S|\langle P\rangle^{-1}+I)|N|. \quad (8)$$

Let us denote by $\widehat{T}^{(l)}$, the matrix in (8). Clearly $\widehat{T}^{(l)}\geqslant O$. Moreover, this is the iteration matrix of an alternating two-stage method for the matrix $\langle M\rangle-|N|$ with the regular splittings $\langle M\rangle-|N|$ and $\langle M\rangle=\langle P\rangle-|Q|=\langle R\rangle-|S|$. Therefore, from (7), $\widehat{T}^{(l)}=(\langle R\rangle^{-1}|S|\langle P\rangle^{-1}|Q|)^{s(l)}+(I-(\langle R\rangle^{-1}|S|\langle P\rangle^{-1}|Q|)^{s(l)})\langle M\rangle^{-1}|N|$ is obtained.

It follows from Theorem 13 that there is a unique pair of matrices B, C, such that $\langle R\rangle^{-1}|S|\langle P\rangle^{-1}|Q|=B^{-1}C$ and $\langle M\rangle=B-C$ is a weak regular splitting. Thus,

$$\widehat{T}^{(l)} = (B^{-1}C)^{s(l)}+(I-(B^{-1}C)^{s(l)})\langle M\rangle^{-1}|N|$$
$$= I-\left((I-(B^{-1}C)^{s(l)})\langle M\rangle^{-1}\right)(\langle M\rangle-|N|). \quad (9)$$

Consider any fixed vector $e=(1,1,\ldots,1)^T>0$ and $x=(\langle M\rangle-|N|)^{-1}e>0$. Since $B^{-1}e>0$, and $\langle M\rangle^{-1}=(I-B^{-1}C)^{-1}B^{-1}=\sum_{j=0}^{\infty}(B^{-1}C)^jB^{-1}$, from (9) it follows $\widehat{T}^{(l)}x=x-\sum_{j=0}^{s(l)-1}(B^{-1}C)^jB^{-1}e\leqslant x-B^{-1}e<x$. Therefore, there exists $0\leqslant\theta<1$ such that $\widehat{T}^{(l)}x\leqslant\theta x, l=1,2,\ldots$, and from Lemmata 12 and 1(b) the proof is complete. □

We will now deal with the convergence of the alternating two-stage method when A is a Hermitian positive definite matrix.

Theorem 17 [13]. *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 18 [4]. *Let A be a Hermitian positive definite matrix. Consider $A=M-N$ such that M is Hermitian and N is positive semidefinite. Let $M=F-G$ be a P-regular splitting. Assume further that the sequence of inner iterations $\{s(l)\}_{l=0}^{\infty}$ remains bounded. Then, the two-stage iterative method (3) converges to the solution of the linear system (1), for any initial vector $x^{(0)}$.*

Theorem 19. *Let A be a Hermitian positive definite matrix. Consider $A=M-N$ such that M is Hermitian and N is positive semidefinite. Let $M=P-Q=R-S$ be P-regular splittings. Assume further that the* *sequence of inner iterations* $\{s(l)\}_{l=0}^{\infty}$ *remains bounded. Then, the alternating two-stage method (5) converges to the solution of the linear system (1), for any initial vector* $x^{(0)}$.

Proof. By hypotheses M is positive definite and therefore nonsingular. From (7) it follows that the iteration matrix of the corresponding alternating two-stage method (5) can be written as follows:

$$T^{(l)} = (R^{-1}SP^{-1}Q)^{s(l)} + (I - (R^{-1}SP^{-1}Q)^{s(l)})M^{-1}N.$$

Moreover, from Theorem 17, there is a unique pair of matrices B, C, such that $R^{-1}SP^{-1}Q = B^{-1}C$ and $M = B - C$ is a P-regular splitting. Thus, $T^{(l)}$ is the iteration matrix of a non-stationary two-stage method for the matrix $A = M - N$, with M Hermitian and N positive semidefinite and $M = B - C$ being P-regular. Therefore, using Theorem 18, the proof is complete. □

4. Convergence for consistent singular linear systems

Under certain conditions, discussed in this section, the alternating two-stage iterative method (5) can be extended to the case where the matrix of the linear system (1) is singular but the equations are consistent. In this case, the alternating two-stage iterative methods may be used to approximate a solution to the problem. In other words, if x^* is a solution of (1), and $e^{(l+1)} = x^{(l+1)} - x^*$, then $e^{(l+1)} = T^{(l)}e^{(l)}$, for $l = 0, 1, \ldots$ Thus, to study the convergence of the alternating two-stage method (5) we need to show that $T^{(l)}T^{(l-1)}\cdots T^{(0)}e^{(0)}$ converges to an element in $\mathcal{N}(A)$, the null space of A, when l tends to infinity.

Theorem 20 [14]. *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* $\|A^{(l)}(I - A^{(l)})(I - A^{(l)})^{\#}\| \leqslant \theta < 1$, $l = 0, 1, \ldots,$ *then* $\lim_{i\to\infty} A^{(l)}A^{(l-1)}\cdots A^{(0)} = P$, *where P is a projection matrix onto the subspace S.*

We proceed now to discuss the convergence of the alternating two-stage method when the coefficient matrix of the linear system is an M-matrix with *property c*.

Theorem 21. *Let A be an M-matrix with property c. Let the splitting* $A = M - N$ *be regular, and the splittings* $M = P - Q = R - S$ *be weak regular. Then, the matrices* $T^{(l)}, l = 0, 1, \ldots,$ *defined in (6), satisfy* $\rho(T^{(l)}) \leqslant 1$ *and* $\mathrm{ind}_1 T^{(l)} \leqslant 1$.

Proof. From (7), it follows that $I - T^{(l)} = (I - (R^{-1}SP^{-1}Q)^{s(l)})(I - M^{-1}N), l = 0, 1, \ldots$. Since $M^{-1} \geqslant O$ and the splittings $M = P - Q = R - S$, are weak regular, then we have $(R^{-1}SP^{-1}Q)^{s(l)} \geqslant O$, and from Theorem 13, $\rho((R^{-1}SP^{-1}Q)^{s(l)}) < 1$. Therefore, $(I - (R^{-1}SP^{-1}Q)^{s(l)})^{-1}$ exists and is a nonnegative matrix.

Consider the matrix $C = (I - M^{-1}N)^{\#}(I - (R^{-1}SP^{-1}Q)^{s(l)})^{-1}$, where $(I - M^{-1}N)^{\#}$ is the group generalized inverse of $(I - M^{-1}N)$. Its existence follows from Theorem 10. Therefore, in order to conclude the proof we are going to show that matrix C satisfies conditions (i)–(iv) of Theorem 7. Clearly, using Definition 6, the matrix C satisfies conditions (i) and (ii). Furthermore, it is easy to show that $\mathcal{R}(C) = \mathcal{R}((I - M^{-1}N)^{\#}) = \mathcal{R}(I - M^{-1}N)$ and

$$\mathcal{N}(I - T^{(l)}) = \mathcal{N}(I - M^{-1}N) = \mathcal{N}(A). \tag{10}$$

Moreover, from Theorem 10, it follows that $\mathcal{R}(I - M^{-1}N)$ and $\mathcal{N}(I - M^{-1}N)$ are complementary, and then (iii) is shown.Finally, let $x \in \mathcal{R}(I - T^{(l)}), x \geqslant 0$, then $(I - (R^{-1}SP^{-1}Q)^{s(l)})^{-1}x \in \mathcal{R}(I - M^{-1}N)$ and also $(I - (R^{-1}SP^{-1}Q)^{s(l)})^{-1}x \geqslant 0$. Since $M^{-1}N \geqslant O$ and $(I - M^{-1}N)^{\#}$ exists, it follows from Plemmons [15, Theorem 2] that $(I - M^{-1}N)^{\#}$ is nonnegative on $\mathcal{R}(I - M^{-1}N)$. Then $Cx \geqslant 0$ and the proof is complete. □

Theorem 22. *Let A be an M-matrix with property c. Let the splitting* $A = M - N$ *be regular, and the splittings* $M = P - Q = R - S$ *be weak regular. Assume further that the diagonal entries of the matrices* $P^{-1}Q$ *and* $R^{-1}S$, *are positive. Then, the matrices* $T^{(l)}, l = 0, 1, \ldots,$ *defined in (6), are semiconvergent.*

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

$$T^{(l)} = (R^{-1}SP^{-1}Q)^{s(l)} + \sum_{j=0}^{s(l)-1}(R^{-1}SP^{-1}Q)^{j}R^{-1}(SP^{-1} + I)N$$

are nonnegative and have positive diagonal entries. Moreover, from Theorem 21, the matrices $T^{(l)}, l = 0, 1, \ldots,$ satisfy condition (b) of Theorem 5 Therefore, using the result in [16, Theorem 2], the proof is complete. □

We would like to point out that in Theorem 22 we have assumed that the matrices $P^{-1}Q$ and $R^{-1}S$, 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 5 holds, we use a standard device by shifting the matrix, so that the value 1 is the only eigenvalue on the unit circle; see e.g., [6].

Theorem 23. *Let A be an M-matrix with property c. Let the splitting* $A = M - N$ *be regular, and the splittings* $M = P - Q = R - S$ *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 (6), are semiconvergent.*

Proof. Since $I - T_{\delta}^{(l)} = \delta(I - T^{(l)}), l = 0, 1, \ldots,$ from Theorem 21 it follows, for each $\delta \in (0, 1)$, that $\rho(T_{\delta}^{(l)}) \leqslant 1$ and $\mathcal{N}(I - T_{\delta}^{(l)}) \oplus \mathcal{R}(I - T_{\delta}^{(l)}) = \Re^n, l = 0, 1, \ldots$. Moreover, by the hypotheses on the splittings and from (6), $T^{(l)} \geqslant O$. Thus (see e.g., [6, Exercise 6.4.3]), $T_{\delta}^{(l)}$ has only the eigenvalue one on the unit circle, and from Theorem 5 it follows that $T_{\delta}^{(l)}$ is semiconvergent for all $\delta \in (0, 1)$. □

Therefore, if need be, Eq. (5) can be replaced in the alternating two-stage method by

$$x^{(l+1)} = \delta(T^{(l)}x^{(l)} + c_{s(l)}) + (1 - \delta)x^{(l-1)}, \quad l = 0, 1, \ldots. \tag{11}$$

Theorem 24. *Let A be an M-matrix with property c. Let the splitting* $A = M - N$ *be regular, and the splittings* $M = P - Q = R - S$ *be weak regular. Assume that the sequence of inner iterations* $\{s(l)\}_{l=0}^{\infty}$ *satisfies* $s(l) = s, l = 0, 1, \ldots$. *Then the following two results hold.*

(a) *If the diagonal entries of the matrices* $P^{-1}Q$ *and* $R^{-1}S$, *are positive, the stationary alternating two-stage method (5) converges to a solution of the consistent linear system* $Ax = b$, *for any initial vector* $x^{(0)}$.
(b) *The stationary alternating two-stage method (5) with the modification (11), converges to a solution of the consistent linear system* $Ax = b$, *for any initial vector* $x^{(0)}$.

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

$$T^{(l)} = T = (R^{-1}SP^{-1}Q)^{s} + \sum_{j=0}^{s-1}(R^{-1}SP^{-1}Q)^{j}R^{-1}(SP^{-1} + I)N,$$

cf. (6). Let x^* be a solution of (1), and $e^{(l)} = x^{(l)} - x^*$, then $e^{(l)} = Te^{(l-1)} = T^{l}e^{(0)}$, for $l = 1, 2, \ldots$. In the case (a), from Theorem 22, T is semiconvergent, and from Lemma 8 it follows that

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

Therefore, from (10) the semiconvergence is proved. The proof of part (b) is analogous, but using, in this case, Theorem 23. □

Theorem 25. *Let A be an M-matrix with property c. Let the splitting $A = M - N$ be regular, and the splittings $M = P - Q = R - S$ 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 (6). Assume further that the sequence of inner iterations $\{s(l)\}_{l=0}^{\infty}$ remains bounded. Then, the alternating two-stage iterative method (5) converges to a solution of the consistent linear system (1), for any initial vector $x^{(0)}$.*

Proof. The proof is an immediate consequence of Theorems 21 and 20. □

Next we study the symmetric positive semidefinite case. Firstly, we give an auxiliary result needed later.

Lemma 26 [13]. *Let A be a symmetric positive definite matrix and let $A = B - C$ be a P-regular splitting. Given $s \geqslant 1$, the unique splitting induced by $(B^{-1}C)^s$ is also a P-regular splitting.*

Theorem 27. *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 $M = P - Q = R - S$ be P-regular splittings. Assume that the sequence of inner iterations $\{s(l)\}_{l=0}^{\infty}$ satisfies $s(l) = s, l = 0,1,\ldots.$ Then the stationary alternating two-stage method (5) converges to a solution of the consistent linear system $Ax = b$, for any initial vector $x^{(0)}$.*

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

$$T^{(l)} = T = (R^{-1}SP^{-1}Q)^s + (I-(R^{-1}SP^{-1}Q)^s)M^{-1}N,$$

cf. (7). Moreover, from Theorem 17, there is a pair of matrices B, C, such that $R^{-1}SP^{-1}Q = B^{-1}C, M = B - C$ is a P-regular splitting and $\rho(B^{-1}C) < 1$. Therefore, $I-(B^{-1}C)^s$ is a nonsingular matrix. Thus, from Lemmas 3 and 26 it follows that the splitting induced by $(B^{-1}C)^s$, namely $M = \widehat{B} - \widehat{C}$, with $\widehat{B} = M(I-(B^{-1}C)^s)^{-1}$, is P-regular. Then, $T^{(l)} = T = \widehat{B}^{-1}\widehat{C} + (I-\widehat{B}^{-1}\widehat{C})M^{-1}N = \widehat{B}^{-1}(\widehat{C} + (\widehat{B}-\widehat{C})M^{-1}N) = \widehat{B}^{-1}(\widehat{C}+N)$. Thus, the splitting $A = \widehat{B} - (\widehat{C}+N)$ is a (non-unique) splitting induced by T. Since $\widehat{B}^T + \widehat{C}$ is positive definite and N is positive semidefinite, $\widehat{B}^T + \widehat{C} + N$ is positive definite and thus this splitting is P-regular. Therefore, from Theorem 11 it follows that T is a semiconvergent matrix and the proof is completed. □

Theorem 28. *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 $M = P - Q = R - S$ be P-regular splitting. 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 (6). Assume further that the sequence of inner iterations $\{s(l)\}_{l=0}^{\infty}$ remains bounded. Then, the alternating two-stage iterative method (5) converges to a solution of the consistent linear system (1), for any initial vector $x^{(0)}$.*

Proof. The proof is an immediate consequence of Theorems 20 and 27. □

5. Numerical experiments

The numerical experiments corresponding to the methods described in this paper have been performed on several parallel computing platforms and similar results have been obtained on all of them. Here, the results obtained in two different parallel operating systems are presented. One of them is a distributed multiprocessor IBM RS/6000 SP with 8 nodes. These nodes are 120 MHz Power2 Super Chip and are connected through a high performance switch with latency time of 40 microseconds and a bandwidth of 110 Mbytes per second. The second system used is an Ethernet network of 6 Pentiums IV running Linux and connected through a switch with a bandwidth of 1 Gbit per second. The parallel environment has been managed using the MPI library of parallel routines [17]. Moreover, the BLAS routines [18] for vector computations and the SPARSKIT routines [19] for handling sparse matrices, were used. The algorithms have been applied to the solution of singular linear systems arising from Markov chain modeling. Concretely, these algorithms 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., $B^Te = e$, where $e = (1,1,\ldots,1)^T$. This implies that $\rho(B) = \rho(B^T) = 1$; see e.g., [6]. The vector of probabilities is normalized so that $x^Te = 1$. In this case, the system to be solved is

$$(I-B)x = 0. \tag{12}$$

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 (5) with the modification (11), if need be, is guaranteed when the splittings are chosen as in Theorems 24 and 25. In these experiments, alternating block iterative methods for the solution of the linear system (12) are used. In the methods used for the solution of (12), the variables are partitioned into r groups, i.e., $x = [x_1^T, x_2^T, \ldots, x_r^T]^T, x_i \in \Re^{n_i}, i = 1,\ldots,r, \sum_{i=1}^r n_i = n$. Thus, the matrix $A = I - B$ is partitioned into $r \times r$ blocks as follows:

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ A_{21} & A_{22} & \cdots & A_{2r} \\ \vdots & \vdots & & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rr} \end{bmatrix} \tag{13}$$

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 processors p. Thus, we have assembled blocks from (13) into p groups, each group being assigned to one processor. There are r_ℓ blocks assigned to processor $\ell, \ell = 1,\ldots,p$, and thus $\sum_{\ell=1}^p r_\ell = r$. In order to describe the different versions of the iterative methods tested, we describe the Block Jacobi (BJ) and the Symmetric Block Gauss–Seidel (SBGS) algorithms, for solving the singular linear system (12), for a generic number of blocks q in (13). Taking into account that we concentrate on Markov chains, we assume that the solution x is normalized so that $x^Te = 1$, where e is a vector of all components equal to one. In fact, in the algorithms studied below, such normalization is assumed at every iteration.

Algorithm 1 (*Block Jacobi – BJ*). Given an initial vector $x^{(0)^T} = [x_1^{(0)^T}, x_2^{(0)^T}, \ldots, x_q^{(0)^T}]$

For $l = 1,2,\ldots,$ until convergence
  For $i = 1$ to q

$$\text{Solve (or approximate) } A_{ii}x_i^{(l)} = -\sum_{j=1 j\neq i}^{q} A_{ij}x_j^{(l-1)}. \tag{14}$$

The linear systems (14) can be solved independently of each other. Thus, this algorithm is inherently parallel. This parallelism is best exploited if the number of blocks matches the number of processors.

In order to ensure the regularity of the Block Jacobi splitting, when B is the transition matrix of a finite homogeneous Markov chain, one can suppose that each column of $N = A - \mathrm{diag}(A_{11},$

$\ldots, A_{rr}$) must have one nonzero 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 nonzero entry; see, e.g., [14].

Algorithm 2 (*Symmetric Block Gauss–Seidel – SBGS*). Given an initial vector $x^{(0)^T} = [x_1^{(0)^T}, x_2^{(0)^T}, \ldots, x_q^{(0)^T}]$

For $l = 0, 1, \ldots,$ until convergence
  For $i = 1$ to q
    Solve (or approximate)

$$A_{ii}x_i^{(l+\frac{1}{2})} = -\sum_{j=1}^{i-1} A_{ij}x_j^{(l+\frac{1}{2})} - \sum_{j=i+1}^{q} A_{ij}x_j^{(l)}, \tag{15}$$

  For $i = q$ to 1
    Solve (or approximate)

$$A_{ii}x_i^{(l+1)} = -\sum_{j=i+1}^{q} A_{ij}x_j^{(l+1)} - \sum_{j=1}^{i-1} A_{ij}x_j^{(l+\frac{1}{2})}. \tag{16}$$

It is clear, from the right hand sides of (15) and (16) that this algorithm is inherently sequential.

We are ready to describe the different parallel methods explored in this paper. We assume that there are p processors. As mentioned earlier, when each solution of (14) in Algorithm 1 is approximated by an alternating iterative method, this is called an alternating two-stage method. In particular, each solution of (14) could be approximated by the Symmetric Block Gauss–Seidel method.

Algorithm 3. Divide the r blocks of (13) into p groups, each assigned to a different processor.

1. Perform parallel BJ (with $q = p$ in Algorithm 1), i.e., each processor approximates the solution of one linear system (14). (This is the outer iteration).
2. Each solution of (14) in Step 1 is approximated using t steps of SBGS (with $q = r_\ell$ in Algorithm 2).
3. Each solution of (15) and (16) in Step 2 is approximated by a fixed number m of Gauss–Seidel (GS) iterations.

In Step 3 of Algorithm 3, each solution of (15) and (16) is approximated using GS iterations. In the following algorithm these solutions are obtained by a direct method.

Algorithm 4. Divide the r blocks of (13) into p groups, each assigned to a different processor.

1. Perform parallel BJ (with $q = p$ in Algorithm 1), i.e., each processor approximates the solution of one linear system (14). (This is the outer iteration).
2. Each solution of (14) in Step 1 is approximated using t steps of SBGS (with $q = r_\ell$ in Algorithm 2).
3. Each solution of (15) and (16) in Step 2 is obtained using LU factorizations.

In our experiments, we have used as the global stopping criterion (i.e., in the outer iteration) a test for the error to be less than a prescribed tolerance. In other words, we test if $\|Ax^{(l)}\|_2 < \varepsilon$.

As is well known, BJ, as well as SBGS, and the nested Algorithms 3 and 4, may not converge, especially for singular linear systems. In order to guarantee convergence, we use the customary device of shifting the iteration matrix from T to $T_\delta = \delta T + (1-\delta)I$, where $0 < \delta < 1, T = M^{-1}N$ and $A = M - N$ is a splitting that represents the iteration of the corresponding algorithm. 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 all experiments reported in this section we have used $\varepsilon = 10^{-6}$ and $\delta = 0.95$. We actually run our codes with two models in [20]. One of the specific models chosen is a biological model described in [21]: the 2D epidemic model of Ridler–Rowe. The matrix we use is of order 263,169 and has 1,050,625 nonzeros, and we label it TwoD. The other model is a multi-class, finite-buffer, priority queuing network model with applicability to telecommunications modelling described, e.g., in [22]. The matrix we use is of order 130,068 and has 875,896 nonzeros, and we label it QNATM.

In order to run the algorithms considered in this paper, three different permutation and partition methods were used to obtain the block structure (13), namely TPABLO [23], the near-complete decomposability test of the Markov chain analyzer (MARCA) [24], and the equal partitioning that forms (approximately) equal size blocks. However, for these models (they are not nearly completely decomposable, [25]), the use of partitions such as MARCA or TPABLO provides a substantial increase in computational time. Concretely, since the matrices are not nearly completely decomposable, MARCA does not find more than one strongly connected component after removing some elements according to a drop tolerance; therefore, MARCA permutation is the identity. On the other hand, our runs with TPABLO permutations using different parameters (blocking, threshold and minimum/maximum block sizes, see [23] for an explanations of these parameters) obtain more that one diagonal partition only when a maximum block size is indicated; after the corresponding permutation, the best computational time is between two/five times the one obtained for the algorithm that uses equal size blocks obtained from the original matrix.

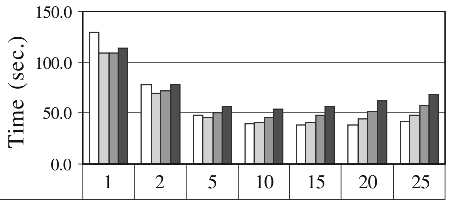
Therefore, we concentrate our experiments on the performance of the algorithms using equal size blocks obtained from the original matrix. These equal size blocks were obtained in the following way: firstly, we consider a block diagonal structure depending on the number of processors (Table 1 reports the order of these diagonal blocks, for the QNATM and TwoD matrices), and finally we construct, at each diagonal block, blocks of a predetermined and constant size $n_i = \eta$. Obviously, if the order of a diagonal block is not a multiple of η, there will be an extra block of order less than η.

In Fig. 1 we report results corresponding to Algorithm 3 on both parallel environments, for the QNATM matrix of order 130,068 varying the number of inner iterations m and using 4 processors. In this figure, the diagonal blocks have been divided into blocks of order $\eta = 50$. The best results are obtained for $m = 1$ or $m = 2$. It can also be observed in Fig. 1 that for a fixed number of processors and for each m the computational time starts to decrease as the number of SBGS steps, t, increases up to an optimal value of t after which the time starts to increase (in Fig. 1a, $t = 15$ for $m = 1$ and $t = 10$ for $m > 1$, and in Fig. 1b, the optimal value is $t = 20$ for $m = 1$ and $t = 15$ for $m > 1$). This behaviour is characteristic of two-stage methods, and it appears in all the results presented here.

Although this optimal value is hard to predict, a good choice for the value of t is one which balances the realization of more inner updates with the decrease of the global iterations (and its associ-

Table 1
Block size of diagonal blocks.

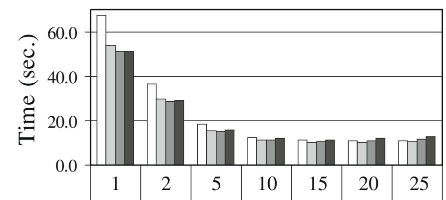
Matrix: QNATM of order 130,068					
$p = 4$	Blocks size	32,517	32,517	32,517	32,517
	Nonzeros in diag. blocks	215,254	211,532	209,317	214,097
Matrix: TwoD of order 263,169					
$p = 2$	Blocks size	131,584	131,585		
	Nonzeros in diag. blocks	524,798	524,802		
$p = 4$	Blocks size	65,792	65,792	65,792	65,793
	Nonzeros in diag. blocks	261,888	261,886	261,887	261,890



	1	2	5	10	15	20	25
Alg. 3 (m=1)	129.7	78.3	47.9	39.0	37.8	38.8	41.4
Alg. 3 (m=2)	109.1	69.2	45.7	40.3	41.2	43.9	47.5
Alg. 3 (m=3)	109.4	72.1	50.3	46.1	47.8	51.8	57.5
Alg. 3 (m=4)	114.1	77.6	56.8	53.4	56.8	61.8	68.8

Values of t

(a) IBM RS/6000 SP.

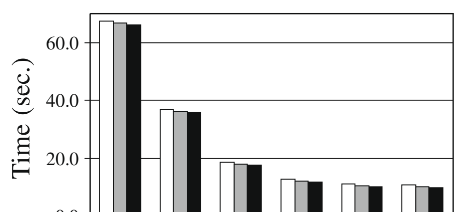


	1	2	5	10	15	20	25
Alg. 3 (m=1)	67.5	36.7	18.4	12.6	11.1	10.7	11.0
Alg. 3 (m=2)	53.8	29.7	15.4	11.2	10.3	10.3	10.7
Alg. 3 (m=3)	51.2	28.7	15.2	11.4	10.6	10.9	11.6
Alg. 3 (m=4)	51.0	28.9	15.8	12.0	11.3	11.9	12.7

Values of t

(b) Linux Cluster.

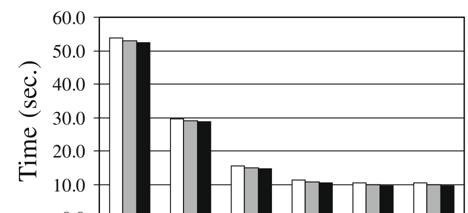
Fig. 1. Algorithm 3. Matrix QNATM of order 130,068. Four processors.



	1	2	5	10	15	20
Alg. 3 ($\eta = 50$)	67.5	36.7	18.4	12.6	11.1	10.7
Alg. 3 ($\eta = 100$)	66.6	36.0	17.8	12.1	10.5	10.1
Alg. 3 ($\eta = 150$)	66.2	35.7	17.5	11.8	10.2	9.9

Values of t

(a) Linux Cluster, m = 1 .



	1	2	5	10	15	20
Alg. 3 ($\eta = 50$)	53.8	29.7	15.4	11.2	10.3	10.3
Alg. 3 ($\eta = 100$)	52.9	29.0	14.9	10.7	9.8	9.7
Alg. 3 ($\eta = 150$)	52.4	28.7	14.6	10.4	9.5	9.4

Values of t

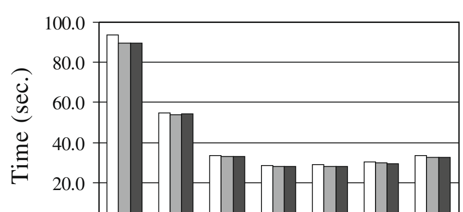
(b) Linux Cluster, m = 2 .

Fig. 2. Algorithm 3. Matrix QNATM of order 130,068. Four processors.

ated computational cost). As we can see, the choice for the value of t depends on the characteristic of the parallel computer system used. Thus, according to the characteristic of the IBM RS/6000 SP, that is, minor speed of process but greater speed in the interconnection network than the cluster of Pentiums, the optimal number of local steps must decrease (see e.g., Fig. 1). Our experience indicates that, for our problem models, good choices of the value of t are between 10 and 15 on the IBM RS/6000 SP, and between 15 and 20 when the cluster of Pentiums is used.

As can be seen in Fig. 2 there are no significant differences when another block size η is used. Nevertheless, we find that a good choice for the size of these blocks is $\eta = 150$. Hence, unless indicated otherwise, the results presented in the rest of this section have been performed with $\eta = 150$.

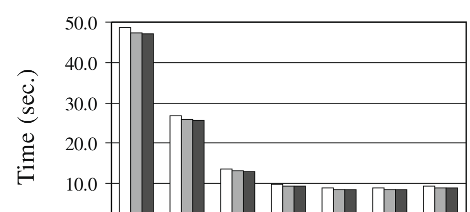
In Fig. 3, we illustrate the performance of Algorithm 4 on both parallel environments. The results reported in this figure are representative of other runs we have performed. For all values of t the timings are always better than those of Algorithm 3. That is, the



	1	2	5	10	15	20	25
Alg. 4 ($\eta = 50$)	93.6	54.7	33.7	28.6	28.8	30.5	33.4
Alg. 4 ($\eta = 100$)	89.6	54.0	33.1	28.0	28.1	29.8	32.5
Alg. 4 ($\eta = 150$)	89.7	54.1	33.1	27.9	28.0	29.6	32.4

Values of t

(a) IBM RS/6000 SP.

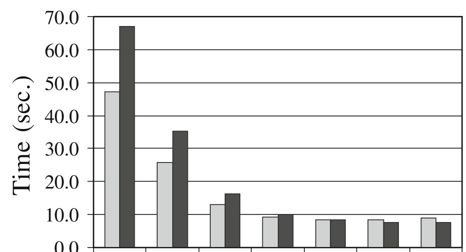


	1	2	5	10	15	20	25
Alg. 4 ($\eta = 50$)	48.6	26.7	13.6	9.8	9.0	9.0	9.5
Alg. 4 ($\eta = 100$)	47.4	25.8	13.1	9.3	8.5	8.5	8.9
Alg. 4 ($\eta = 150$)	47.2	25.8	13.1	9.3	8.5	8.4	8.8

Values of t

(b) Linux Cluster.

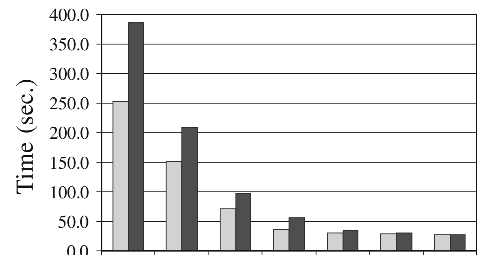
Fig. 3. Algorithm 4. Matrix QNATM of order 130,068. Four processors.



	1	2	5	10	15	20	25
Alg. 4 (SBGS)	47.2	25.8	13.1	9.3	8.5	8.4	8.8
Alg. 4 (BGS)	67.1	35.3	16.2	10.1	8.3	7.7	7.5

Values of t

(a) Linux Cluster. Matrix QNATM.

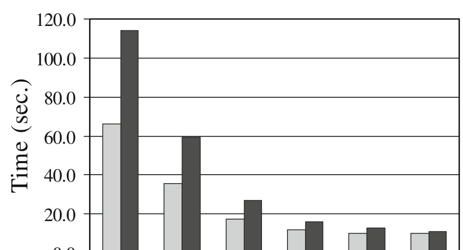


	1	2	5	10	15	20	25
Alg. 4 (SBGS)	252.8	151.9	71.2	36.0	30.7	28.5	27.9
Alg. 4 (BGS)	386.9	209.3	97.2	56.0	34.9	30.4	27.7

Values of t

(b) Linux Cluster. Matrix TwoD.

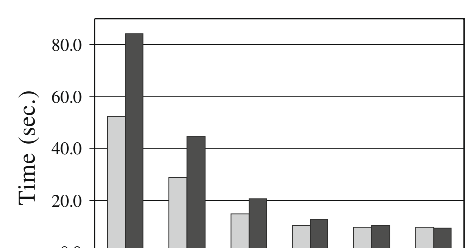
Fig. 4. Algorithm 4. Comparison of using alternating iterations. Four processors.



	1	2	5	10	15	20
Alg. 3 (SBGS)	66.2	35.7	17.5	11.8	10.2	9.9
Alg. 3 (BGS)	114.0	59.5	26.9	16.1	12.8	11.1

Values of t

(a) Linux Cluster, m = 1 .



	1	2	5	10	15	20
Alg. 3 (SBGS)	52.4	28.7	14.6	10.4	9.5	9.4
Alg. 3 (BGS)	84.4	44.4	20.5	12.8	10.4	9.3

Values of t

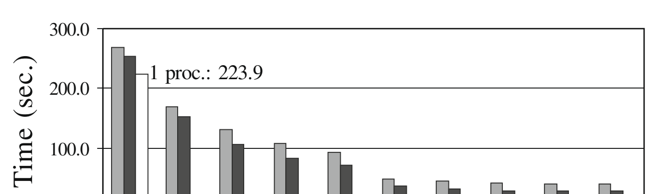
(b) Linux Cluster, m = 2 .

Fig. 5. Algorithm 3. Comparison of using alternating iterations. Matrix QNATM. Four processors.

use of LU factorizations to solve (15) and (16) instead of using Gauss–Seidel iterations provides a substantial reduction in computational time when solving these singular linear systems.

On the other hand, if in Step 2 of Algorithms 3 and 4 we use BGS iterations instead of SBGS iterations to approximate each solution of (14) we are in the presence of a two-stage method. Fig. 4 compares the use of two-stage and alternating two-stage methods for Algorithm 4. In Fig. 4a we report results corresponding to the QNATM matrix and Fig. 4b corresponds to the TwoD matrix. We note that for small t values, the use of SBGS alternating iterations is always preferable, and for high t values both methods have similar performance. This conclusion can be extended to Algorithm 3 as we can see from the results in Fig. 5.

In Fig. 6 we present results comparing the execution time of Algorithm 4 setting $p = 1, 2, 4$, that is, with 1, 2 or 4 processors, for the matrix TwoD and using a block size of $\eta = 150$. We note that when Algorithm 4 is used with a different number of processors, p, a different computational method is obtained, in terms of performed operations. In fact, Algorithm 4 with $p = 1$ corresponds to a SBGS algorithm in which the diagonal blocks (of order η) are solved by using LU factorizations; note that, in this case (i.e., $p = 1$), the parameter t used in Algorithm 4 is not relevant, that



	1	2	3	4	5	10	15	20	25	30
2 processors	269.2	169.4	130.8	107.7	93.1	48.9	44.0	42.0	40.3	39.8
4 processors	252.8	151.9	106.1	83.2	71.2	36.0	30.7	28.5	27.9	28.1

Values of t

Fig. 6. Algorithm 4. Matrix TWOD of order 263,169. Linux cluster.

is, for all values of t the same method is obtained. Fig. 6 shows a significant reduction of the execution time for Algorithm 4 when the number of processors increases, specially when a reasonable value of t is considered.

6. Conclusion

In this paper, a convergence theory of two-stage iterative methods for the solution of both nonsingular linear systems and consistent singular linear systems has been developed. Convergence of these methods, for nonsingular linear systems, is shown for monotone matrices, H-matrices and Hermitian positive definite matrices. Furthermore, for consistent singular linear systems, convergence theorems when the matrix of the linear system is either M-matrix or symmetric positive semidefinite are given. Although the theory has been demonstrated in a sequential context, it can be extended without difficulty to a parallel environment, as presented in the numerical experiments. The experiments performed for singular systems, arising from Markov chain models, show that the use of two-stage methods with alternating methods and direct methods in inner levels, provide a substantial reduction in computational time.

Acknowledgements

This research was supported by the Spanish Ministry of Science and Education under Grant No. TIN2008-06570-C04-04, and by Universidad de Alicante under grant number VIGROB-020.

We thank the referees for their comments and questions, which led to improvements of our presentation.

References

[1] Nichols NK. On the convergence of two-stage iterative processes for solving linear equations. SIAM J Numer Anal 1973;10:460–9.
[2] Frommer A, Szyld DB. H-splittings and two-stage iterative methods. Numer Math 1992;63:345–56.
[3] Lanzkron PJ, Rose DJ, Szyld DB. Convergence of nested classical iterative methods for linear systems. Numer Math 1991;58:685–702.
[4] Migallón V, Penadés J. Convergence of two-stage iterative methods for Hermitian positive definite matrices. Appl Math Lett 1997;10(3):79–83.
[5] Dayar T, Stewart WJ. Comparison of partitioning techniques for two-level iterative methods on large, sparse Markov chains. SIAM J Sci Comput 2000;21(5):1691–705.
[6] Berman A, Plemmons RJ. Nonnegative matrices in the mathematical sciences. 3rd ed. New York: Academic Press; 1979. Reprinted by SIAM, Philadelphia, 1994.
[7] Neumaier A. New techniques for the analysis of linear interval equations. Linear Algebra Appl 1984;58:273–325.
[8] Ortega JM, Rheinboldt WC. Iterative solution of nonlinear equations in several variables. New York and London: Academic Press; 1970.
[9] Varga RS. Matrix iterative analysis. Englewood Cliffs, New Jersey: Prentice-Hall; 1962.
[10] Benzi M, Szyld DB. Existence and uniqueness of splittings for stationary iterative methods with applications to alternating methods. Numer Math 1997;76:309–21.
[11] Neumann M, Plemmons RJ. Convergence of parallel multisplitting iterative methods for M-matrices. Linear Algebra Appl 1987;88/89:559–73.
[12] Robert F, Charnay M, Musy F. Itérations chaotiques série-parallèle pour des équations non-linéaires de point fixe. Appl Math 1975;20:1–38.
[13] Castel MJ, Migallón V, Penadés J. Convergence of non-stationary parallel multisplitting methods for Hermitian positive definite matrices. Math Comput 1998;67(221):209–20.
[14] Migallón V, Penadés J, Szyld DB. Block two-stage methods for singular systems and Markov chains. Numer Linear Algebra Appl 1996;3(5):413–26.
[15] Plemmons RJ. M-matrices leading to semiconvergent splittings. Linear Algebra Appl 1976;15:243–52.
[16] Alefeld G, Schneider N. On square roots of M-matrices. Linear Algebra Appl 1982;42:119–32.
[17] Gropp W, Lusk E, Skjellum A. Using MPI: portable parallel programming. Cambridge, MA: MIT Press; 1994.
[18] BLAS: Basic Linear Algebra Subroutine. <http://www.netlib.org/index.html>.
[19] Saad Y. SPARSKIT: a basic tool kit for sparse matrix computation. <http://www-users.cs.umn.edu/~saad/software/SPARSKIT/sparskit.html>.
[20] Stewart WJ. MARCA models: a collection of Markov chain models. <http://www.csc.ncsu.edu/faculty/WStewart/MARCA_Models/MARCAModels.html>.
[21] Pollett PK, Stewart DE. An efficient procedure for computing quasi-stationary distributions of Markov chains with sparse transition structure. Adv Appl Probab 1994;26:68–79.
[22] Philippe B, Saad Y, Stewart WJ. Numerical methods in Markov chain modeling. Oper Res 1992;40(6):1156–79.
[23] Choi H, Szyld DB. Application of threshold partitioning of sparse matrices to Markov chains. In: Proceedings of the IEEE international computer performance and dependability symposium; 1996. p. 158–65.
[24] Stewart WJ. MARCA: Markov chain analyzer. A software package for Markov modelling. In: Stewart WJ, editor. Numerical solution of Markov chains. NY: M. Dekker, Inc.; 1991. p. 137–62.
[25] Stewart WJ. Introduction to the numerical solution of Markov chains. Princeton, New Jersey: Princeton University Press; 1994.