BLOCK TWO-STAGE METHODS FOR SINGULAR SYSTEMS AND MARKOV CHAINS

Violeta Migallón, José Penadés,
and Daniel B. Szyld

Report 95-121
December 1995

This paper has been submitted for the Conference on Algebraic Multilevel Iteration Methods with Applicaitons (AMLI'96) to be held in Nijmegen, The Netherlands, June 13–15, 1996.

This report is available in the World Wide Web at `http://www.math.temple.edu/~szyld` or via anonymous ftp at `ftp.math.temple.edu` in directory `pub/szyld`.

BLOCK TWO-STAGE METHODS FOR SINGULAR SYSTEMS AND MARKOV CHAINS*

VIOLETA MIGALLÓN†, JOSÉ PENADÉS† AND DANIEL B. SZYLD‡

Abstract. The use of block two-stage methods for the iterative solution of consistent singular linear systems is studied. In particular, hypotheses are provided for the convergence of non-stationary methods, i.e., when the number of inner iterations may vary from block to block and from one outer iterations to another.

Key words. Iterative methods, linear systems, singular matrices, block methods, multisplitting, two-stage, non-stationary, Markov chains.

AMS(MOS) subject classification. 65F10, 65F15.

1. Introduction. We are interested in the block iterative solution of consistent $n \times n$ singular systems of linear equations of the form

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

where A is an M–matrix, i.e., when A can be expressed as $A = sI - B$, with $B \geq O$ (a nonnegative matrix), $s > 0$, and $\rho(B) \leq s$, where $\rho(B)$ denotes the spectral radius of B; see e.g., [5], [45]. The M–matrix A is singular when $s = \rho(B)$. By consistent, it is meant that b is in $\mathcal{R}(A)$, the range of A. In particular, 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, denoted $x \geq 0$, 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., [5], [41], [45]. The vector of probabilities is normalized so that $x^T e = 1$. In this case, the system (1) corresponds to $A = I - B$ (i.e., $s = 1$) and $b = 0$.

Let us assume that the set $\{1, 2, \ldots, n\}$ is partitioned into r disjoint sets of n_ℓ elements each, $1 \leq \ell \leq r$. Let us further assume that the equations and unknowns are rearranged (permuted) so that the $n \times n$ coefficient matrix A has the form

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

with the diagonal blocks $A_{\ell\ell}$ being square of order n_ℓ, $1 \leq \ell \leq r$, $\sum_{\ell=1}^{r} n_\ell = n$, and the vectors x and b are partitioned conformally. This partition may arise naturally due to the structure

* This work was initiated while the third author visited the Universidad de Alicante, supported by Conselleria de Cultura, Educació i Ciència de la Generalitat Valenciana, Spain, and was completed while the first two authors visited Temple University supported by Spanish DGICYT grant numbers PR95-012 and PR95-013.

† Departamento de Tecnología Informática y Computación, Universidad de Alicante, E-03080 Alicante, Spain (violeta@dtic.ua.es, jpenades@dtic.ua.es).

‡ Department of Mathematics, Temple University, Philadelphia, Pennsylvania 19122-2585, USA (szyld@math.temple.edu). Supported by the National Science Foundation grant DMS-9201728.

of the problem, or it may be obtained using some block partitioning algorithm; see e.g., [12], [36]. We know of instances, e.g., the one described by Ciardo, Gluckman, and Nicol [13] that the matrix A, or B, where $A = I - B$, is generated in a distributed array of processors, with groups of rows (or columns) computed in each processor. Thus, a partition of the form (2) is readily available.

Classical block methods, such as Block-Jacobi, or Block Gauss-Seidel, can be used for the solution of (1) partitioned as in (2). Description of these methods can be found, e.g., in the books by Berman and Plemmons [5], Stewart [41], or Varga [45]. A study of the convergence of such methods can be carried out without much difficulty by extending the results of the corresponding point methods for singular systems as analyzed, e.g., by Barker and Plemmons [2], Barker [3], Marek and Szyld [29], or Neumann and Plemmons [31]. For the solution of singular systems using Krylov-type methods, see the survey by Philippe, Saad and Stewart [37], and also Freund and Hochbruck [14].

At each step of these block iterative methods, linear systems of the form

$$A_{\ell\ell} v = g, \quad 1 \le \ell \le r, \tag{3}$$

need to be solved, where $A_{\ell\ell}$ are the diagonal blocks in (2). When the order of these diagonal blocks, n_ℓ, are large, it is natural to approximate their solution using an iterative method, and thus we are in the presence of a two-stage iterative method. This is in fact recommended by Stewart [41, Section 3.3]. Two-stage methods, sometimes called inner-outer iterations, were studied, e.g., by Nichols [33], Golub and Overton [20], [21], Lanzkron, Rose and Szyld [26], Frommer and Szyld [18], [19], and Bru, Migallón and Penadés [9]. Bru, Elsner and Neumann [7] recently studied two-stage methods for singular systems. When the number of iterations to approximate each of the systems (3) is the same for all ℓ, $1 \le \ell \le r$, and for each (outer) step, $i = 1, 2, \ldots$, it is said that the method is stationary, while a non-stationary block method is such that different (inner) iterations are performed in each block and/or in each iterative step, say $q(\ell, i)$ iterations $1 \le \ell \le r$, $i = 1, 2, \ldots$; see e.g., [9], [10], [18], [19], [30].

In this paper, we study the convergence of certain non-stationary two-stage methods for singular M–matrices. In particular, the convergence of the two-stage non-stationary Block-Jacobi method for Markov chains is analyzed.

For the two-stage Block-Jacobi method, let us consider the splittings $A_{\ell\ell} = B_\ell - C_\ell$, $1 \le \ell \le r$. Let the block diagonal matrix M consist of the diagonal blocks in (2), i.e.,

$$M = \text{ Diag } (A_{11}, \ldots, A_{\ell\ell}, \ldots, A_{rr}), \tag{4}$$

and consider the splitting $A = M - N$. We assume that the matrix M is nonsingular. We show in Section 4 when this assumption is satisfied for the Markov chains problem. Consider further the splittings $M = F_\ell - G_\ell$, $1 \le \ell \le r$, where F_ℓ is a block diagonal matrix having the $n_\ell \times n_\ell$ nonsingular matrix B_ℓ in the ℓth block, and identity matrices of the appropriate order in the other diagonal blocks. Moreover, let the $n \times n$ diagonal matrices E_ℓ have ones in the entries corresponding to the diagonal block $A_{\ell\ell}$ and zero otherwise. Note that this implies that

$$\sum_{\ell=1}^{r} E_\ell = I. \tag{5}$$

Thus, we have

$$F_\ell = \text{Diag }(I,\ldots,I,\ B_\ell, I,\ldots,I), \quad \text{and} \tag{6}$$

$$E_\ell = \text{Diag }(O,\ldots,O,I,O,\ldots,O). \tag{7}$$

With this notation, the following algorithm describes the non-stationary Block-Jacobi method for the approximate solution of (1).

ALGORITHM 1. (NON-STATIONARY TWO-STAGE MULTISPLITTING). Given the initial vector x_0, and a sequence of numbers of inner iterations $q(\ell,i)$, $1 \leq \ell \leq r$, $i = 1,2,\ldots$.

For $i = 1,2,\ldots$, until convergence.

For $\ell = 1$ to r

$y_{\ell,0} = x_{i-1}$

For $j = 1$ to $q(\ell,i)$

$$F_\ell y_{\ell,j} = G_\ell y_{\ell,j-1} + N x_{i-1} + b \tag{8}$$

$$x_i = \sum_{\ell=1}^{r} E_\ell y_{\ell,q(\ell,i)} \ . \tag{9}$$

This is a special case of Algorithm 4 in [10], where the matrix A is assumed to be nonsingular. This algorithm is much more general than the Block-Jacobi method, if, for example, M, F_ℓ, and/or the weighting matrices E_ℓ are different than (4), (6), (7), respectively, as long as the weighting matrices satisfy (5). In particular, this general formulation allows us to include in our convergence results, e.g., the more general Block-Jacobi type two-stage methods, see e.g., [19], or a method *with overlap*, i.e., where the matrices E_ℓ have some nonzeros in entries other than those of $A_{\ell\ell}$; cf. [17], [23].

Algorithm 1 extends the multisplitting algorithm introduced by O'Leary and White [35], and further studied and extended by many authors, e.g., by Frommer and Mayer [15], [16], Jones and Szyld [23], [43], Neumann and Plemmons [32], or White [46], [47]. It is easy to see that up to r different processors can be efficiently utilized in parallel, each computing the iterations (8). The fact that for each block, i.e., for each ℓ, $1 \leq \ell \leq r$, different number of iterations $q(\ell,i)$ are performed may be exploited to achieve better load balancing, especially if the order of the diagonal blocks, n_ℓ, vary over a large range; see e.g., [10].

The global iteration matrix of Algorithm 1 changes at each (outer) iteration and can be written as

$$T^{(i)} = \sum_{\ell=1}^{r} E_\ell \left[(F_\ell^{-1}G_\ell)^{q(\ell,i)} + \sum_{j=0}^{q(\ell,i)-1} (F_\ell^{-1}G_\ell)^j F_\ell^{-1} N \right], \quad i = 1,2,\ldots, \tag{10}$$

or equivalently, as

$$T^{(i)} = \sum_{\ell=1}^{r} E_\ell \left[(F_\ell^{-1}G_\ell)^{q(\ell,i)} + \left(I - (F_\ell^{-1}G_\ell)^{q(\ell,i)} \right) M^{-1} N \right], \quad i = 1,2,\ldots, \tag{11}$$

see e.g., [10]. In other words, if x^* is a solution of (1), and $e_i = x_i - x^*$, then $e_i = T^{(i)} e_{i-1}$, for $i = 1,2,\ldots$. Thus, to study the convergence of Algorithm 1 we need to show that $T^{(i)}T^{(i-1)}\cdots T^{(1)}e_0$ converges to an element in $\mathcal{N}(A)$, the null space of A, as $i \to \infty$, cf. Keller [24]. This is precisely our main result, shown, under certain hypotheses, in Section 3. In Section 4 we show how our results apply to the Markov chain problem. In the next section, we present some definitions and preliminaries used later in the paper.

2. Notation and Preliminaries. Let $T \in \mathbb{R}^{n\times n}$, by $\sigma(T)$ we denote the spectrum of the matrix T. We define $\gamma(T) = \max\{|\lambda| \ : \ \lambda \in \sigma(T), \ \lambda \neq 1\}$, i.e., $\gamma(T)$ is the maximum magnitude over all elements in $\sigma(T) \setminus \{1\}$. We say that two subspaces S_1 and S_2 on $\mathbb{R}^n$ are *complementary* if $S_1 \oplus S_2 = \mathbb{R}^n$, i.e., if $S_1 \cap S_2 = \{0\}$ and $S_1 + S_2 = \mathbb{R}^n$. The *index* of a square matrix T, denoted *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 to 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 $T \in \mathbb{R}^{n\times n}$, is *zero-convergent* if $\lim_{k\to\infty} T^k = O$. We say that T is *convergent* if $\lim_{k\to\infty} T^k$ exists. It is well known that a matrix T is zero-convergent if and only if $\rho(T) < 1$. If, on the other hand $\rho(T) = 1$, two different conditions need to be satisfied to guarantee convergence, as the following result shows; see e.g., [31], [34].

THEOREM 2.1. *Let $T \in \mathbb{R}^{n\times n}$ with $\rho(T) = 1$. The matrix T is convergent if and only if the following two statements hold.*
(a) $1 \in \sigma(T)$ *and* $\gamma(T) < 1$, *(b)* $\mathcal{N}(I-T) \oplus \mathcal{R}(I-T) = \mathbb{R}^n$.

Condition (a) of Theorem 2.1 means that $\rho(T) = 1$ is the only eigenvalue in the unit circle. Condition (b) is equivalent to having $ind_1 T = 1$; see e.g., [5], [42]. Alefeld and Schneider [1, Theorem 2], show that condition (b) together with the hypotheses that $T \geq O$, having positive diagonal entries and $\rho(T) = 1$, imply that T is convergent. Condition (b) is also equivalent to the existence of the group inverse $(I-T)^{\#}$. We review in what follows the definition of some generalized inverses; see e.g., [5], [11].

DEFINITION 2.2. *Let $A \in \mathbb{R}^{n\times n}$, 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 point out 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}. The following two results are (a part of) Theorem 6.4.12 and Lemma 7.6.11 of [5], respectively.

THEOREM 2.3. *Let $T \in \mathbb{R}^{n\times n}$, 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 four conditions.*

(i) $I - T = (I-T)C(I-T)$
(ii) $C = C(I-T)C$
(iii) $\mathcal{N}(I-T) \oplus \mathcal{R}(C) = \mathbb{R}^n$
(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 2.4. *Let $T \in \mathbb{R}^{n\times n}$ be convergent. Then* $\lim\limits_{k\to\infty} T^k = I - (I-T)(I-T)^{\#}$.

By $Z^{n\times n}$ we denote the set of all real $n \times n$ matrices which have all nonpositive off-diagonal entries. When $A = sI - B$ is a nonsingular M–matrix, then $\rho(B) < s$. Thus $s^{-1}B$ is zero-convergent. 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 convergent. Obviously, a nonsingular M–matrix always has *property c*. M–matrices with *property c* were introduced by Plemmons [38]. Several characterizations and sufficient conditions for matrices in $Z^{n\times n}$ to be M–matrices with *property c* are given by Neumann and Plemmons [31]; see also Berman and Plemmons [5]. For example, a matrix $A \in Z^{n\times n}$ having a positive vector x such that $Ax \geq 0$, is an M–matrix with *property c*. Some other characterizations use the concept of regular splittings.

A splitting $A = M - N$ is called a *regular splitting* if $M^{-1} \geq O$ and $N \geq O$. It is called a *weak regular splitting* if $M^{-1} \geq O$ and $M^{-1}N \geq O$; see e.g., [5], [45].

THEOREM 2.5. [31] *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 (a) $\rho(T) \leq 1$, and (b) $\mathcal{N}(I - T) \oplus \mathcal{R}(I - T) = \mathbb{R}^n$.*

In our convergence results, in Section 3, it is assumed that the matrix A of the system (1) is an M–matrix with *property c*. In particular, as shown, e.g., in [5], matrices representing finite homogeneous Markov chains satisfy this property, and thus, our results apply to that case; see further Section 4. Also, symmetric positive semidefinite matrices in $Z^{n\times n}$ are M–matrices with *property c*, and thus coefficient matrices for certain systems of linear equations resulting from finite difference methods for partial differential equations, such as the Discrete Neumann Problem in a rectangular region or the Poisson's equation with periodic boundary condition; see e.g., [5], [39].

A square matrix $A = [a_{ij}]$ is called *column diagonally dominant* if

$$|a_{jj}| \geq \sum_{i=1, i\neq j}^{n} |a_{ij}|, \quad 1 \leq j \leq n, \tag{12}$$

it is called *strictly column diagonally dominant* if strict inequality holds in (12), for $j = 1, 2, \ldots, n$, and *irreducibly column diagonally dominant* if A is irreducible and strict inequality of (12) holds for at least one j; see e.g., [5], [45]. The following results are part of Theorem 6.2.3 and 6.2.7 of [5].

LEMMA 2.6. *Let $A \in Z^{n\times n}$ satisfying one of the following conditions.*
(a) A is strictly column diagonally dominant and the diagonal entries are positive.
(b) A is irreducible and there exists a positive vector x with $Ax \geq 0$ and $Ax \neq 0$.
Then, A is a nonsingular M–matrix.

The product of zero-convergent matrices is not necessarily zero-convergent and may not tend to zero; see e.g., Johnson and Bru [22], or Robert, Charnay and Musy [40]. A sufficient condition to assure the convergence to zero of a product of different matrices is given in the following lemma by Bru and Fuster [8].

LEMMA 2.7. *Let $A^{(i)}$, $i = 1, 2, \ldots$, be a sequence of square complex matrices. If there exists a matrix norm $\|\cdot\|$ such that $\|A^{(i)}\| \leq \theta < 1$, $i = 1, 2, \ldots$, then $\lim_{i\to\infty} A^{(i)}A^{(i-1)}\cdots A^{(1)} = O$.*

The following result is a tool used in our convergence analysis. It is stated without proof in [27], [28], [29], [42], and used implicitely by Krieger [25].

THEOREM 2.8. *Let $T \in \mathbb{R}^{n\times n}$. T is convergent if and only if $T = P + Q$, where $P^2 = P$, $PQ = QP = O$, and $\rho(Q) < 1$. Moreover, P is a projection onto $\mathcal{N}(I - T)$.*

Proof. Suppose T is convergent. Let $P = \lim_{k\to\infty} T^k$. By Lemma 2.4, we have $P = I - (I - T)(I - T)^\#$. We then set $Q = T - P$. It easy to show, using the properties of the group inverse in Definition 2.2, that $P^2 = P$ and $PQ = QP = O$. Moreover, since $Q^k = (T - P)^k = T^k - P$, we have $\lim_{k\to\infty} Q^k = O$ and thus $\rho(Q) < 1$. On the other hand, as $(I - T)P = O$, it follows that P is a projection onto $\mathcal{N}(I - T)$. The converse is obvious. □

3. Convergence of Non-stationary Methods. We begin by showing that the iteration matrix at each (outer) step of Algorithm 1 has index 1 associated to the value one. This result generalizes Lemma 4.1 of Bru, Elsner and Neumann [7] to the case where we have r different splittings $M = F_\ell - G_\ell$, $1 \leq \ell \leq r$, and thus, to the non-stationary Block-Jacobi method.

THEOREM 3.1. *Let A be an M–matrix with property c. Let the splitting $A = M - N$ be regular, and the splittings $M = F_\ell - G_\ell$, $1 \leq \ell \leq r$, be weak regular. Then, the matrices $T^{(i)}$ defined in (10) satisfy $\rho(T^{(i)}) \leq 1$ and $ind_1 T^{(i)} \leq 1$ for all $i = 1, 2, \ldots$.*

Proof. From (11), it follows that $I - T^{(i)} = (I - H^{(i)})(I - M^{-1}N)$, $i = 1, 2, \ldots$, where $H^{(i)} = \sum_{\ell=1}^{r} E_\ell (F_\ell^{-1} G_\ell)^{q(\ell,i)}$, $i = 1, 2, \ldots$. Since $M^{-1} \geq O$ and the splittings $M = F_\ell - G_\ell$, $1 \leq \ell \leq r$, are weak regular, then we have that $H^{(i)} \geq O$, and also, as shown, e.g., in [6], that $\rho(H^{(i)}) < 1$. Therefore, $(I - H^{(i)})^{-1}$ exists and it is a nonnegative matrix.
To conclude our proof we exhibit a matrix C which satisfies conditions (i)-(iv) of Theorem 2.3. Let $C = (I - M^{-1}N)^{\#}(I - H^{(i)})^{-1}$, where $(I - M^{-1}N)^{\#}$ is the group generalized inverse of $(I - M^{-1}N)$. Its existence follows from Theorem 2.5. Clearly, using Definition 2.2, the matrix C satisfies conditions (i) and (ii). Furthermore it easy to show that $\mathcal{R}(C) = \mathcal{R}((I - M^{-1}N)^{\#}) = \mathcal{R}(I - M^{-1}N)$ and

$$\mathcal{N}(I - T^{(i)}) = \mathcal{N}(I - M^{-1}N) = \mathcal{N}(A). \tag{13}$$

Since, again by Theorem 2.5, $\mathcal{R}(I - M^{-1}N)$ and $\mathcal{N}(I - M^{-1}N)$ are complementary, (iii) is shown. Finally, let $x \in \mathcal{R}(I - T^{(i)})$, $x \geq 0$, then $(I - H^{(i)})^{-1}x \in \mathcal{R}(I - M^{-1}N)$ and also $(I - H^{(i)})^{-1}x \geq 0$. Since $M^{-1}N \geq O$ and $(I - M^{-1}N)^{\#}$ exists it follows from Plemmons [38, Theorem 2] that $(I - M^{-1}N)^{\#}$ is nonnegative on $\mathcal{R}(I - M^{-1}N)$. Then $Cx \geq 0$. □

Thus, we have shown that the iteration matrices $T^{(i)}$ satisfy condition (b) of Theorem 2.1, $i = 1, 2, \ldots$. A way to insure that condition (a) holds is to find, as done by Bru, Elsner and Neumann in [7], a matrix norm $\|\cdot\|$ such that $T^{(i)}$ is paracontracting with respect to this norm (i.e., for every vector x, $T^{(i)}x \neq x \Leftrightarrow \|T^{(i)}x\| < \|x\|$). However, in many cases the matrices are not paracontracting. For example, if condition (a) is not satisfied, that norm never exists and then another tool is needed.

THEOREM 3.2. *Let A be an M–matrix with property c. Let the splitting $A = M - N$ be regular, and the splittings $M = F_\ell - G_\ell$, $1 \leq \ell \leq r$, be weak regular. Assume further that the diagonal entries of $F_\ell^{-1} G_\ell$, $1 \leq \ell \leq r$, are positive. Then, the matrices $T^{(i)}$ defined in (10) are convergent, for all $i = 1, 2, \ldots$.*

Proof. By the hypotheses, $\sum_{\ell=1}^{r} E_\ell \left[\sum_{j=0}^{q(\ell,i)-1} (F_\ell^{-1} G_\ell)^j F_\ell^{-1} N \right]$ are nonnegative matrices and the diagonal entries of the matrices $\sum_{\ell=1}^{r} E_\ell (F_\ell^{-1} G_\ell)^{q(\ell,i)}$ are positive. Therefore, the matrices $T^{(i)}$, $i = 1, 2, \ldots$, are also nonnegative and have positive diagonal entries. Since, by Theorem 3.1, $T^{(i)}$, $i = 1, 2, \ldots$, satisfy condition (b) of Theorem 2.1, using the result in [1, Theorem 2], the proof is complete. □

In Theorem 3.2 we have assumed that the matrices $F_\ell^{-1} G_\ell$, $1 \leq \ell \leq r$, have positive diagonal entries. However, the iteration matrices of the classical Jacobi or Gauss-Seidel methods do not have this property. Then, to insure that condition (a) of Theorem 2.1 holds, we may use a standard device by shifting the matrix, so that the value 1 is the only eigenvalue on the unit circle; see e.g., [5], [25], [31]. We thus have the following result.

THEOREM 3.3. *Let A be an M−matrix with property c. Let the splitting $A = M - N$ be regular, and the splittings $M = F_\ell - G_\ell$, $1 \le \ell \le r$, be weak regular. Then, for each $\alpha \in (0,1)$, the matrices $T_\alpha^{(i)} = \alpha T^{(i)} + (1-\alpha)I$, $i = 1,2,\ldots$, with $T^{(i)}$ defined in (10), are convergent.*

Proof. From Theorem 3.1 and using that $I - T_\alpha^{(i)} = \alpha(I - T^{(i)})$, $i = 1,2,\ldots$, it follows, for each $\alpha \in (0,1)$, that $\rho(T_\alpha^{(i)}) \le 1$ and $\mathcal{N}(I - T_\alpha^{(i)}) \oplus \mathcal{R}(I - T_\alpha^{(i)}) = \mathbb{R}^n$, $i = 1,2,\ldots$. On the other hand, by the hypotheses on the splittings and from (10), $T^{(i)} \ge O$. Thus (see e.g., [5, Exercise 6.4.3]), $T_\alpha^{(i)}$ has only the eigenvalue one on the unit circle, and by Theorem 2.1, $T_\alpha^{(i)}$ is convergent for all $\alpha \in (0,1)$. □

Consequently, if need be, one can replace equation (9) in Algorithm 1 by

$$x_i = \alpha \sum_{\ell=1}^{r} E_\ell y_{\ell,q(\ell,i)} + (1-\alpha)x_{i-1}, \qquad 0 < \alpha < 1. \tag{14}$$

We can now prove the convergence of Algorithm 1 when $q(\ell,i) = q(\ell)$, $i = 1,2,\ldots$. This means that the number of inner iterations performed in each block can be different from each other, but stays fixed in each outer step. In practice, this allows us to counterweight the work in each processor, producing a good overall load balance. Note that one can find good values of $q(\ell,i) = q(\ell)$ by experimenting first with a few outer iterations, varying $q(\ell,i)$ for each i.

THEOREM 3.4. *Let A be an M−matrix with property c. Let the splitting $A = M - N$ be regular, and the splittings $M = F_\ell - G_\ell$, $1 \le \ell \le r$, be weak regular. Assume further that for all ℓ, $1 \le \ell \le r$, there exists an integer $q(\ell)$ such that $q(\ell,i) = q(\ell)$, $i = 1,2,\ldots$, $1 \le \ell \le r$. Then the following two results hold.*
(a) If the diagonal entries of $F_\ell^{-1}G_\ell$, $1 \le \ell \le r$, are positive, the non-stationary two-stage multisplitting Algorithm 1 converges to a solution of the consistent linear system $Ax = b$, for any initial vector x_0.
(b) The non-stationary two-stage multisplitting Algorithm 1 with the modification (14), converges to a solution of the consistent linear system $Ax = b$, for any initial vector x_0.

Proof. Since $q(\ell,i) = q(\ell)$, $i = 1,2,\ldots$, $1 \le \ell \le r$, then there is a single iteration matrix, i.e.,

$$T^{(i)} = T = \sum_{\ell=1}^{r} E_\ell \left[(F_\ell^{-1}G_\ell)^{q(\ell)} + \sum_{j=0}^{q(\ell)-1} (F_\ell^{-1}G_\ell)^j F_\ell^{-1} N \right],$$

cf. (10). Let x^* be a solution of (1), and $e_i = x_i - x^*$, then $e_i = Te_{i-1} = T^i e_0$, for $i = 1,2,\ldots$. We show that the sequence $\{e_i\}_{i=0}^\infty$ converges to a vector in $\mathcal{N}(A)$. Since, by Theorem 3.2, T is convergent, by Lemma 2.4, $\lim_{i\to\infty} e_i = \lim_{i\to\infty} T^i e_0 = [I - (I-T)(I-T)^\#]e_0 \in \mathcal{N}(I-T)$. Then, by (13) the convergence is proved. The proof of part (b) is analogous using Theorem 3.3. □

We now formulate a basic theorem to insure the convergence of a product of convergent matrices. This result generalizes Lemma 2.7.

THEOREM 3.5. *Let $A^{(i)}$, $i = 1,2,\ldots$, be a sequence of square complex matrices such that each group inverse $(I - A^{(i)})^\#$ exists. Suppose that there is a subspace S satisfying $\mathcal{N}(I - A^{(i)}) = S$, $i = 1,2,\ldots$. If there exists a matrix norm $\|\cdot\|$ such that the set $\{\|A^{(i)}\|\}_{i=1}^\infty$ remains bounded and $\|A^{(i)}(I - A^{(i)})(I - A^{(i)})^\#\| \le \theta < 1$, $i = 1,2,\ldots$, then*

$$\lim_{i\to\infty} A^{(i)}A^{(i-1)}\cdots A^{(1)} = P, \tag{15}$$

where P is a projection matrix on the subspace S.

Proof. Let v_0 be arbitrary, and let $v_i = A^{(i)}v_{i-1} = A^{(i)}A^{(i-1)}\cdots A^{(1)}v_0$, for $i = 1, 2, \ldots$. The first part of the proof consists of showing that the sequence $\{v_i\}_{i=0}^{\infty}$ converges to a vector in S. Let $P_i = I - (I - A^{(i)})(I - A^{(i)})^{\#}$ and $Q_i = A^{(i)}(I - A^{(i)})(I - A^{(i)})^{\#}$ be the matrices defined in Theorem 2.8 where $A^{(i)} = P_i + Q_i,\ i = 1, 2, \ldots$. It is easy to show that

$$P_iP_j = P_j \quad \text{and} \quad Q_iP_j = O, \quad i, j = 1, 2, \ldots. \tag{16}$$

Note that from Theorem 2.8 and from the hypotheses, P_i, $i = 1, 2, \ldots$, are projections matrices onto the same subspace S. Thus, from Theorem 2.8 and (16) we have that

$$A^{(i)}A^{(i-1)}\cdots A^{(1)} = P_1 + \sum_{k=1}^{i-1} P_{k+1}\left(Q_kQ_{k-1}\cdots Q_1\right) + Q_iQ_{i-1}\cdots Q_1. \tag{17}$$

By hypothesis

$$\|Q_i\| \leq \theta < 1,\ i = 1, 2, \ldots, \tag{18}$$

and using Lemma 2.7, we have that

$$\lim_{i\to\infty} Q_iQ_{i-1}\cdots Q_1 = O. \tag{19}$$

Then, by (17), if the limit of $\{v_i\}_{i=0}^{\infty}$ exists, it lies on S.
To conclude the first part of the proof we show that the sequence $\{v_i\}_{i=0}^{\infty}$ is a Cauchy sequence, and thus convergent. Let a, b be positive constants so that $\|v_0\| \leq b$ and $a \geq \sup\limits_{i=1,2,\ldots} \|P_i\| \geq 1$. From (19), it follows that for every $\epsilon > 0$, there exists an index i_0 such that

$$\|Q_iQ_{i-1}\cdots Q_1\| \leq \frac{\epsilon}{2ab}(1-\theta), \qquad \text{for all } i \geq i_0. \tag{20}$$

Let s, t be integers such that $s \geq i_0$, $t \geq i_0$ and, without loss of generality, let $s > t$. Thus, using the bounds (18) and (20), it follows that, for all $k \geq 1$

$$\begin{aligned} \|P_{t+k}Q_{t+k-1}\cdots Q_{t+1}Q_t\cdots Q_1\| &\leq \|P_{t+k}\|\ \|Q_{t+k-1}\|\cdots\|Q_{t+1}\|\ \|Q_t\cdots Q_1\| \\ &\leq a\theta^{k-1}\frac{\epsilon}{2ab}(1-\theta) = \frac{\epsilon(1-\theta)}{2b}\theta^{k-1}. \end{aligned}$$

Then, by the expression of the products of the matrices $A^{(i)}$ in (17)

$$\begin{aligned} \|v_s - v_t\| &= \|A^{(s)}A^{(s-1)}\cdots A^{(1)}v_0 - A^{(t)}A^{(t-1)}\cdots A^{(1)}v_0\| \\ &\leq \|P_{t+1}Q_tQ_{t-1}\cdots Q_1 + P_{t+2}Q_{t+1}Q_t\cdots Q_1 + \ldots + P_sQ_{s-1}Q_{s-2}\cdots Q_1 \\ &\qquad + Q_sQ_{s-1}\cdots Q_1 - Q_tQ_{t-1}\cdots Q_1\|\ \|v_0\| \\ &\leq \frac{\epsilon}{2}(1-\theta)\left[1 + \theta + \theta^2 + \ldots + \theta^{s-t-1} + \frac{\theta^{s-t}}{a} + \frac{1}{a}\right] \\ &\leq \frac{\epsilon}{2}(1-\theta^{s-t+1}) + \frac{\epsilon}{2}(1-\theta) \leq \epsilon. \end{aligned}$$

Hence, the sequence $\{v_i\}_{i=0}^{\infty}$ converges to an element in S, for any v_0.
Thus, we have that for each column of the identity e^k, $k = 1, \ldots, n$, $\lim\limits_{i\to\infty} A^{(i)}A^{(i-1)}\cdots A^{(1)}e^k$

exists and lies in S. Thus, the limit (15) exists and each of the columns of P lies in S, i.e., for any vector y, $Py \in S$. Furthermore, if $x \in S$, we have that $A^{(i)}x = x$, $i = 1, 2, \ldots$, and thus $Px = x$. Then $P^2 = P$, i.e., P is a projection. □

We point out that the upper bound $\|Q_i\| \leq \theta < 1$, $i = 1, 2, \ldots$, cannot be omitted, as we illustrate with the following example.

EXAMPLE 1. Consider the matrices

$$A^{(i)} = \begin{bmatrix} 1 & 0 \\ 0 & exp(-\frac{1}{i^2}) \end{bmatrix}, \quad i = 1, 2, \ldots.$$

It follows that

$$P_i = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \text{ and } Q_i = \begin{bmatrix} 0 & 0 \\ 0 & exp(-\frac{1}{i^2}) \end{bmatrix}, \quad i = 1, 2, \ldots,$$

with $\rho(Q_i) < 1$, $i = 1, 2, \ldots$. Since $\lim_{i \to \infty} \rho(Q_i) = 1$, there is no matrix norm $\| \cdot \|$ such that $\|Q_i\| \leq \theta < 1$, for all $i = 1, 2, \ldots$. All equations $v = A^{(i)}v$ have the unique normalized solution $v = (1, 0)^T$, but the sequence $v_0 = (1, 1)^T$, $v_i = A^{(i)}v_{i-1}$ converges to the vector $(1, exp(-\frac{\pi^2}{6}))^T$.

We use Theorem 3.5 to prove the convergence of Algorithm 1 for any bounded sequence $q(\ell, i)$, $i = 1, 2, \ldots$, $1 \leq \ell \leq r$. This condition is very realistic in practice, since there is always a maximum number of inner iterations for each block.

THEOREM 3.6. *Let A be an M–matrix with property c. Let the splitting $A = M - N$ be regular, and the splittings $M = F_\ell - G_\ell$, $1 \leq \ell \leq r$, be weak regular. Suppose that there exists a matrix norm $\| \cdot \|$ such that $\|T^{(i)}(I - T^{(i)})(I - T^{(i)})^\#\| < 1$, $i = 1, 2, \ldots$, where $T^{(i)}$ are defined in (10). Then, for any bounded sequence of number of inner iterations $q(\ell, i) \geq 1$, $i = 1, 2, \ldots$, $1 \leq \ell \leq r$, the non-stationary two-stage multisplitting Algorithm 1 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 3.1 and 3.5. □

We note that the convergence results of Theorem 3.4 can be regarded as corollaries of Theorem 3.6.

We point out that, in light of Theorem 3.6, from the proof of Theorem 3.5, one can find *all* solutions of (1) by finding one of them first, and then finding a basis of $\mathcal{N}(A)$ using the columns of the identity as initial vectors in Algorithm 1, modified with (14) if needed, and with $b = 0$.

We conclude the section with a few remarks.

If equation (8) is relaxed with some $0 < \omega < 1$, i.e., if it is replaced with

$$y_{\ell,j} = \omega F_\ell^{-1}(G_\ell y_{\ell,j-1} + Nx_{i-1} + b) + (1 - \omega)y_{\ell,j-1},$$

the induced splitting (see [4], [26]) is also weak regular, and thus theorems 3.1–3.4 and 3.6 hold as well. Hence the relaxed non-stationary two-stage multisplitting algorithm to solve the consistent linear system $Ax = b$ converges to a solution.

In theorems 3.1–3.4 and 3.6, the fact that A is an M–matrix with *property c* is only used to guarantee the existence of $(I - M^{-1}N)^\#$. The existence of this group inverse can be obtained by some other hypotheses; see e.g., [5]. Here we mention two of them. One is that

$A \in \mathbb{R}^{n\times n}$ is *range monotone*, i.e., if $Ax \geq 0$ and $x \in \mathcal{R}(A)$, then $x \geq 0$; this hypothesis was used in [7]. Another condition is that there exists a positive vector x such that $Ax \geq 0$.

On the other hand, in the mentioned theorems, we have the hypotheses that A is an M–matrix with *property c*, and that the splitting $A = M - N$ is regular. This implies, by Theorem 2.5, that $\rho(M^{-1}N) \leq 1$, but this fact is not used explicitly in the proofs of the theorems. We note that if one were to replace the hypothesis of *property c* by the apparently less restrictive hypothesis of the existence of $(I - M^{-1}N)^{\#}$, this assumption together with the regularity of the splitting $A = M - N$ actually implies that $\rho(M^{-1}N) \leq 1$; see e.g., [5].

4. Application to finite Markov chains. As mentioned earlier, 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 non-stationary two-stage multisplitting Algorithm 1 with the modification (14), if need be, is guaranteed when the splittings are chosen as in theorems 3.4 and 3.6. In particular, the two-stage non-stationary Block-Jacobi method can be used to find a stationary probability distribution, i.e., a (normalized) solution of $Ax = 0$. When the Markov chain is ergodic, i.e., when B is irreducible, there exists a unique stationary probability distribution vector. In this case, the mentioned algorithm produces this (normalized) vector. However, if the Markov chain is not ergodic, $\mathcal{N}(A)$ has dimension greater than one. All stationary probability distributions can be found by repeated use of the algorithm, with initial vectors being each column of the identity. This procedure was proposed by Tanabe [44] in the context of the conjugate gradient method, and by Marek and Szyld [29] using semi-iterative methods.

In what follows, we present conditions that guarantee that the Block-Jacobi splitting is a regular splitting, and thus satisfy the hypotheses of our convergence theorems.

THEOREM 4.1. *Let B be a transition matrix of a finite homogeneous Markov chain. Consider $A = I - B$ partitioned as in (2) and the Block-Jacobi splitting $A = M - N$ defined in (4). If each matrix $A_{\ell\ell}$, $1 \leq \ell \leq r$, is either strictly or irreducibly column diagonally dominant, then, the splitting $A = M - N$ is regular.*

Proof. Since B is nonnegative and each column sum is 1, if $A_{\ell\ell}$ is strictly column diagonally dominant then its diagonal entries are positive. Moreover, as $A_{\ell\ell} \in Z^{n\times n}$ from Lemma 2.6 (a), $A_{\ell\ell}$ is a nonsingular M–matrix.
On the other hand, if $A_{\ell\ell}$ is column diagonally dominant with at least one strict inequality (12), then $A_{\ell\ell}^T(1, 1, \ldots, 1)^T$ is a non zero nonnegative vector. Then since $A_{\ell\ell}$ is irreducible, from Lemma 2.6 (b), $A_{\ell\ell}$ is a nonsingular M–matrix.
Since, in addition, N is always a nonnegative matrix, it follows in both cases that the splitting $A = M - N$ is regular. □

It is easy to see that, for any matrix in $Z^{n\times n}$, the Block Gauss-Seidel splitting is regular if and only if the Block-Jacobi splitting is. Then, Theorem 4.1 remains valid for the Block Gauss-Seidel splitting.

We conclude by giving equivalent conditions for the two hypotheses of Theorem 4.1. First, each matrix $A_{\ell\ell}$, $1 \leq \ell \leq r$, is strictly column diagonally dominant if and only if each column of N has at least one nonzero entry. Second, the matrix $A_{\ell\ell}$, $1 \leq \ell \leq r$, is column diagonally dominant with at least one strict inequality (12) if and only if at least one column for each corresponding blocks in N, i.e, each of the matrices $\left[-A_{1\ell}^T, \cdots, -A_{\ell-1,\ell}^T, O^T, -A_{\ell+1,\ell}^T, \cdots, -A_{r\ell}^T\right]^T$, $1 \leq \ell \leq r$, has at least one nonzero entry.

REFERENCES

[1] Götz Alefeld and Hans Schneider. On square roots of M-matrices. *Linear Algebra and its Applications*, 42:119–132, 1982.

[2] George P. Barker and Robert J. Plemmons. Convergent iterations for computing stationary distributions of Markov chains. *SIAM Journal on Algebraic and Discrete Methods*, 7:390–398, 1986.

[3] V. A. Barker. Numerical solution of sparse singular systems of equations arising from ergodic Markov chains. *Communications in Statistics. Stochastic Models*, 5:355–381, 1989.

[4] Michele Benzi and Daniel B. Szyld. Existence and uniqueness of splittings for stationary iterative methods with applications to alternating methods. Technical Report 95-81, Department of Mathematics, Temple University, Philadelphia, August 1995.

[5] Abraham Berman and Robert J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, New York, third edition, 1979. Reprinted by SIAM, Philadelphia, 1994.

[6] Rafael Bru, Ludwig Elsner, and Michael Neumann. Models of parallel chaotic iteration methods. *Linear Algebra and its Applications*, 103:175–192, 1988.

[7] Rafael Bru, Ludwig Elsner, and Michael Neumann. Convergence of infinite products of matrices and inner-outer iteration schemes. *Electronic Transactions on Numerical Analysis*, 2:183–193, 1994.

[8] Rafael Bru and Robert Fuster. Parallel chaotic extrapolated Jacobi method. *Applied Mathematics Letters*, 3:65–69, 1990.

[9] Rafael Bru, Violeta Migallón, and José Penadés. Chaotic methods for the parallel solution of linear systems. *Computing Systems in Engineering*. To Appear.

[10] Rafael Bru, Violeta Migallón, José Penadés, and Daniel B. Szyld. Parallel, synchronous and asynchronous two-stage multisplitting methods. *Electronic Transactions on Numerical Analysis*, 3:24–38, 1995.

[11] Stephen L. Campbell and Carl D. Meyer, Jr. *Generalized Inverses of Linear Transformations*. Pitman, London – San Francisco – Melbourne, 1979. Reprinted by Dover, New York, 1991.

[12] Hwajeong Choi and Daniel B. Szyld. Theshold partitionings of sparse matrices and applications to Markov chains. In preparation.

[13] Gianfranco Ciardo, Joshua Gluckman, and David Nicol. Distributed state-space generation of discrete-state stochastic models, 1995. Preprint, Department of Computer Science, College of William and Mary, Williamsburg, Virginia, 1995.

[14] Roland W. Freund and Marlis Hochbruck. On the use of two QMR algorithms for solving singular systems and applications in Markov Chain modelling. *Numerical Linear Algebra with Applications*, 1:403–420, 1994.

[15] Andreas Frommer and Günter Mayer. Convergence of relaxed parallel multisplitting methods. *Linear Algebra and its Applications*, 119:141–152, 1989.

[16] Andreas Frommer and Günter Mayer. On the theory and practice of multisplitting methods in parallel computation. *Computing*, 49:63–74, 1992.

[17] Andreas Frommer and Bert Pohl. A comparison result for multisplittings and waveform relaxation methods. *Numerical Linear Algebra with Applications*, 2:335–346, 1995.

[18] Andreas Frommer and Daniel B. Szyld. H-splittings and two-stage iterative methods. *Numerische Mathematik*, 63:345–356, 1992.

[19] Andreas Frommer and Daniel B. Szyld. Asynchronous two-stage iterative methods. *Numerische Mathematik*, 69:141–153, 1994.

[20] Gene H. Golub and Michael L. Overton. Convergence of a two-stage Richardson iterative procedure for solving systems of linear equations. In G.A. Watson, editor, *Numerical Analysis (Proceedings of the Ninth Biennial Conference, Dundee, Scotland, 1981), Lecture Notes in Mathematics 912*, pages 128–139, New York, 1982. Springer Verlag.

[21] Gene H. Golub and Michael L. Overton. The convergence of inexact Chebyshev and Richardson iterative methods for solving linear systems. *Numerische Mathematik*, 53:571–593, 1988.

[22] Charles R. Johnson and Rafael Bru. The spectral radius of a product of nonnegative matrices. *Linear Algebra and its Applications*, 141:227–240, 1990.

[23] Mark T. Jones and Daniel B. Szyld. Two-stage multisplitting methods with overlapping blocks. To appear in *Numerical Linear Algebra with Applications*.

[24] Herbert B. Keller. On the solution of singular and semidefinite linear systems by iteration. *SIAM Journal on Numerical Analysis*, 2:281–290, 1965.

[25] Udo R. Krieger. On a two-level multigrid solution method for finite Markov chains. *Linear Algebra and its Applications*, 223/224:415–439, 1995.

[26] Paul J. Lanzkron, Donald J. Rose, and Daniel B. Szyld. Convergence of nested classical iterative methods for linear systems. *Numerische Mathematik*, 58:685–702, 1991.

[27] Ivo Marek. On square roots of M-operators. *Linear Algebra and its Applications*, 223/224:501–520, 1995.

[28] Ivo Marek and Daniel B. Szyld. Splittings of M-operators: Irreducibility and the index of the iteration operator. *Numerical Functional Analysis and Optimization*, 11:529–553, 1990.

[29] Ivo Marek and Daniel B. Szyld. Iterative and semi-iterative methods for computing stationary probability vectors of Markov operators. *Mathematics of Computation*, 61:719–731, 1993.

[30] José Mas, Violeta Migallón, José Penadés, and Daniel B. Szyld. A non–stationary parallel relaxed multisplitting method. *Linear Algebra and its Applications*, 237/238 (April 1996).

[31] Michael Neumann and Robert J. Plemmons. Convergent nonnegative matrices and iterative methods for consistent linear systems. *Numerische Mathematik*, 31:265–279, 1978.

[32] Michael Neumann and Robert J. Plemmons. Convergence of parallel multisplitting iterative methods for M-matrices. *Linear Algebra and its Applications*, 88/89:559–573, 1987.

[33] Nancy K. Nichols. On the convergence of two-stage iterative processes for solving linear equations. *SIAM Journal on Numerical Analysis*, 10:460–469, 1973.

[34] Rufus Oldenburger. Infinite powers of matrices and characteristic roots. *Duke Mathematical Journal*, 6:357–361, 1940.

[35] Dianne P. O'Leary and Robert E. White. Multi-splittings of matrices and parallel solution of linear systems. *SIAM Journal on Algebraic and Discrete Mathematics*, 6:630–640, 1985.

[36] James O'Neil and Daniel B. Szyld. A block ordering method for sparse matrices. *SIAM Journal on Scientific and Statistical Computing*, 11:811–823, 1990.

[37] Bernard Philippe, Yousef Saad, and William J. Stewart. Numerical methods in Markov chain modeling. *Operations Research*, 40:1156–1179, 1992.

[38] Robert J. Plemmons. M-matrices leading to semiconvergent splittings. *Linear Algebra and its Applications*, 15:243–252, 1976.

[39] Robert J. Plemmons. Regular splittings and the discrete Neumann problem. *Numerische Mathematik*, 25:153–161, 1976.

[40] F. Robert, M. Charnay, and F. Musy. Itérations chaotiques série-parallèle pour des équations non-linéaires de point fixe. *Aplikace Matematiky*, 20:1–38, 1975.

[41] William J. Stewart. *Introduction to the Numerical Solution of Markov Chains.* Princeton University Press, Princeton, New Jersey, 1994.

[42] Daniel B. Szyld. Equivalence of convergence conditions for iterative methods for singular equations. *Numerical Linear Algebra with Applications*, 1:151–154, 1994.

[43] Daniel B. Szyld and Mark T. Jones. Two-stage and multisplitting methods for the parallel solution of linear systems. *SIAM Journal on Matrix Analysis and Applications*, 13:671–679, 1992.

[44] Kunio Tanabe. The conjugate gradient method for computing all extremal stationary probability vectors of a stochastic matrix. *Annals of the Institute of Statistical Mathematics, Part B*, 37:173–187, 1985.

[45] Richard S. Varga. *Matrix Iterative Analysis.* Prentice-Hall, Englewood Cliffs, New Jersey, 1962.

[46] Robert E. White. Multisplittings and parallel iterative methods. *Computer Methods in Applied Mechanics and Engineering*, 64:567–577, 1987.

[47] Robert E. White. Multisplitting of a symmetric positive definite matrix. *SIAM Journal on Matrix Analysis and Applications*, 11:69–82, 1990.