

# ADDITIVE SCHWARZ ITERATIONS FOR MARKOV CHAINS\*

RAFAEL BRU<sup>†</sup>, FRANCISCO PEDROCHE<sup>†</sup>, AND DANIEL B. SZYLD<sup>‡</sup>

**Abstract.** A convergence analysis is presented for additive Schwarz iterations when applied to consistent singular systems of equations of the form  $Ax = b$ . The theory applies to singular  $M$ -matrices with one-dimensional null space, and it is applicable in particular to systems representing ergodic Markov chains, and to certain discretizations of partial differential equations. Additive Schwarz can be seen as a generalization of Block Jacobi, where the set of indices defining the diagonal blocks have nonempty intersection; this is called the overlap. The presence of overlap is known to accelerate the convergence of the methods in the nonsingular case. By providing convergence results, as well as some characteristics of the induced splitting, we hope to encourage the use of this additional computational tool for the solution of Markov chains and other singular systems. We present several numerical examples showing that additive Schwarz performs better than block Jacobi. For completeness, a few numerical experiments with block Gauss-Seidel and multiplicative Schwarz are also included.

**Key words.** Markov chains, singular linear systems, additive Schwarz iterations, overlap.

**AMS subject classifications.** 65F10, 65F15, 15A48

**1. Introduction.** Stationary classical iterative methods such as block Jacobi and block Gauss-Seidel are extensively used nowadays for the numerical solution of Markov chains; see, e.g., [20]. Schwarz methods can be seen as generalizations of these as they introduce overlap between the blocks. It is the purpose of this paper to show theoretically and experimentally that additive Schwarz iterations can be a viable alternative.

Schwarz iterative methods were developed to solve linear systems of algebraic equations derived from discretizations of partial differential equations; see, e.g., [17], [18], [23]. Schwarz methods are usually used as preconditioners, though in many cases they have been used as stationary iterative methods associated with a physical domain decomposed in  $p$  overlapping subdomains, especially for nonsymmetric problems; see, e.g., [14]. In this paper we focus on additive Schwarz when used as a stationary iterative method for the solution of an  $n \times n$  consistent singular linear systems of the form

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

where we assume that the null space of  $A$  is one-dimensional. This situation applies in particular to matrices  $A = I - B$ , where  $B$  is irreducible and column stochastic, representing an ergodic Markov chain, and  $b = 0$ . In this case, the solution of (1.1) corresponds to the stationary probability distribution of the Markov chain [20]. More generally, we assume that  $A$  is a singular  $M$ -matrix with each principal submatrix being nonsingular. Other situations, such as discretization of certain differential equations with Neumann boundary conditions can also be represented with such matrices.

---

\*First submitted 7 October 2004. This version dated 22 March 2005.

<sup>†</sup>Institut de Matemàtica Multidisciplinar, Universitat Politècnica de València. Camí de Vera s/n. 46022 València. Spain. {rbru,pedroche@mat.upv.es}. Supported by Spanish DGI grants BFM2001-2641 and MTM2004-02998 and by the Oficina de Ciència i Tecnologia de la Presidència de La Generalitat Valenciana under project GRUPOS03/062.

<sup>‡</sup>Department of Mathematics, Temple University, Philadelphia, PA 19122-6094, U.S.A. szyld@math.temple.edu. Supported in part by the U.S. National Science Foundation under grant DMS-0207525.

The convergence of Schwarz iterations in the framework of an algebraic formulation was studied in [2], [6], [10], where theorems on convergence were presented for the case of  $A$  being symmetric positive definite, nonsingular  $M$ -matrix, and  $H$ -matrix. Recently, the convergence of multiplicative Schwarz iterations for singular systems was proved [13]. In the present paper, we prove the convergence of additive Schwarz iterations for singular systems and study certain properties of the induced splitting. When no overlap is present, additive Schwarz reduces to the classical block Jacobi method, while multiplicative Schwarz reduces to block Gauss-Seidel. We review these concepts later and we give definitions and theorems relating those matrices describing the overlap. Overlap has played a major role in the convergence of Schwarz methods for discretizations of differential equations. Here we present numerical evidence that overlap can indeed improve the convergence of iterative methods for singular systems, and specifically for ergodic Markov chains.

We mention that the use of Krylov subspace methods for these problems has not turned out to be so competitive; see, e.g., [16].

The paper is structured as follows. In section 2 we review the algebraic formulation of additive Schwarz iterations and give some preliminaries, especially concerning the overlap. In section 3 we give conditions for the convergence of this iterative method when used with Markov chains. In section 4 we prove, for certain important cases, the existence of a splitting induced by the iteration matrix of additive Schwarz. In section 5 we show some numerical experiments comparing the convergence factor of additive Schwarz and block Jacobi. We also provide execution times. For completeness we include a few numerical experiments with block Gauss-Seidel and multiplicative Schwarz.

**2. Overlapping blocks. Algebraic Representation.** To formally describe the overlapping blocks, let  $V_i$  be subspaces of  $V = \mathbb{R}^n$  of dimension  $n_i$ ,  $i = 1, \dots, p$ , ( $p > 1$ ) such that the sum of these subspaces span the whole space. These subspaces are not pairwise disjoint; on the contrary, their intersection is precisely the overlap,

and thus  $\sum_{i=1}^p n_i > n$ .

The restriction and prolongation operators map vectors from  $V$  to  $V_i$ , and vice-versa. The restriction operators used here are of the form

$$R_i = [I_i | O] \pi_i, \quad i = 1, \dots, p,$$

where  $I_i$  is the identity on  $\mathbb{R}^{n_i}$ , and  $\pi_i$  is a permutation matrix on  $\mathbb{R}^n$ . The prolongation operators considered here are  $R_i^T$ ,  $i = 1, \dots, p$ . Define the following diagonal matrices

$$E_i = R_i^T R_i, \quad i = 1, \dots, p, \quad (2.1)$$

which have nonzero diagonal elements (with value one) only in the columns which have a nonzero element the matrix  $R_i$ . We denote by

$$S_i = \{i_1, i_2, \dots, i_{n_i}\}, \quad i_1 < i_2 < \dots < i_{n_i}, \quad (2.2)$$

the set of indices corresponding to the columns of  $R_i$  which have nonzero elements, and by  $q$  the *measure of overlap*, i.e., the maximum over all possible rows, of the number of matrices  $E_i$  with a nonzero in the row. Thus

$$\sum_{i=1}^p E_i \leq qI; \quad (2.3)$$

and usually  $q \ll p$ .

The restriction of the matrix  $A$  to the subspace  $V_i$  is

$$A_i = R_i A R_i^T, \quad (2.4)$$

which is a symmetric permutation of an  $n_i \times n_i$  principal submatrix of  $A$ ,  $i = 1, \dots, p$ . We assume throughout the paper that the matrix  $A$  is such that  $A_i$  is nonsingular whenever  $n_i < n$ ,  $i = 1, \dots, p$ . We note that the indices on  $S_i$  in (2.2) are those corresponding to the rows and columns of  $A$  that contribute to  $A_i$ , whilst the arrangement of these rows and columns is given by  $R_i$  (or equivalently by  $\pi_i$ ). These matrices  $A_i$  are just the  $p$  overlapping blocks; see [2], [10] for more details.

With these definitions we can briefly describe the Schwarz iterations for the solution of (1.1); more detail can be found, e.g., in [2], [10]. The damped additive Schwarz iteration for the solution of (1.1) has the form

$$x^{k+1} = x^k + \theta \sum_{i=1}^p R_i^T A_i^{-1} R_i (b - A x^k),$$

where  $x^0$  is an initial vector and  $0 < \theta \leq 1$  is the *damping* factor. Thus, additive Schwarz iterations can be expressed as most classical methods by the iteration

$$x^{k+1} = T x^k + c, \quad k = 0, 1, \dots \quad (2.5)$$

The iteration matrix for this scheme is thus

$$T = T_\theta = I - \theta \sum_{i=1}^p R_i^T A_i^{-1} R_i A = I - \theta \sum_{i=1}^p P_i, \quad (2.6)$$

where  $P_i = R_i^T A_i^{-1} R_i A$  is a projection onto  $V_i$ . The iteration matrix for the multiplicative Schwarz iterations is

$$T = T_\mu = (I - P_p)(I - P_{p-1}) \cdots (I - P_1) = \prod_{i=p}^1 (I - P_i).$$

Following [10], we use a different algebraic representation of the iteration matrix (2.6). To that end, let  $A_{\neg i} = [O | I_{\neg i}] \pi_i A \pi_i^T [O | I_{\neg i}]^T$ , where  $I_{\neg i}$  is the  $(n - n_i) \times (n - n_i)$  identity matrix, and let

$$M_i = \pi_i^T \begin{bmatrix} A_i & O \\ O & \text{diag}(A_{\neg i}) \end{bmatrix} \pi_i. \quad (2.7)$$

We observe that this matrix  $M_i$  is just a copy of  $A$  with some coefficients changed to zero. In fact, we can express the (combinatorial) zero structure of  $M_i$  explicitly as

$$M_i(j, k) = 0 \quad \text{if } j \neq k \text{ and } j \notin S_i \text{ or } k \notin S_i, \quad (2.8)$$

for  $j, k = 1, \dots, n$ . Of course there might be more zeros in the other locations corresponding to  $A_i$ , but we do not use this fact.

It follows then from the form of the matrices (2.1) and (2.7) that

$$G_i := E_i M_i^{-1} = R_i^T A_i^{-1} R_i = \pi_i^T \begin{bmatrix} A_i^{-1} & O \\ O & O \end{bmatrix} \pi_i. \quad (2.9)$$

Using this equality, the iteration matrix  $T_\theta$  can be expressed as

$$T_\theta = I - \theta \sum_{i=1}^p E_i M_i^{-1} A. \quad (2.10)$$

Our proof of convergence of additive Schwarz iterations consists of showing that if  $\theta < 1/q$  then the limit  $\lim_{k \rightarrow \infty} T_\theta^k$  exists. We do this in section 3.

The existence of a splitting  $A = M - N$  so that  $T_\theta = M^{-1}N$  is based on showing that the matrix  $G = \sum_{i=1}^p E_i M_i^{-1}$  is nonsingular. As we shall see, this issue is not trivial, and one needs to consider some restrictions on the sets  $S_i$ . We do this in section 4.

**3. Convergence of additive Schwarz for Markov chains.** Convergence and comparison theorems for the classical stationary iterations (2.5) in the case of  $A$  being a nonsingular matrix are well established, and the rate of convergence of such iterative methods is given by the spectral radius of  $T$ ,  $\rho(T)$ ; see, e.g., [4], [24]. For the present case where  $A$  is singular the rate of convergence is given by the *convergence factor*  $\gamma(T)$  defined as

$$\gamma(T) = \max\{|\lambda|, \lambda \in \sigma(T), \lambda \neq 1\},$$

where  $\sigma(T)$  denotes the spectrum of  $T$ . Comparison theorems for splittings of singular matrices have appeared in [11], [12].

We consider  $A = I - B$ , with  $B$  an  $n \times n$  irreducible column stochastic matrix. We wish to show that  $\lim_{k \rightarrow \infty} T_\theta^k$  exists, i.e., it is *convergent*<sup>1</sup>. This limit if it exists is a projection onto the null space of the singular matrix  $A$ . In this case, the additive Schwarz iteration converges to a nontrivial solution of  $Ax = 0$  for any initial vector  $x^0 \notin \mathcal{R}(I - T)$ , the range of  $I - T$  [4].

As is well known [4], a matrix  $T$  with  $\rho(T) = 1$  is convergent if and only if each of the following conditions hold.

- (i)  $\lambda \in \sigma(T)$  with  $|\lambda| = 1$  implies  $\lambda = 1$ , i.e.,  $\gamma(T) < 1$ .
- (ii)  $\text{ind}_1 T = 1$ , i.e.,  $\text{rank}(I - T) = \text{rank}(I - T)^2$ .

When  $T \geq O$ , condition (i) can be replaced with  $T$  having positive diagonal entries [1]. Equivalent conditions for (ii) can be found in [22]. Here we will use the following lemma, whose proof can be found, e.g., in [5].

**LEMMA 3.1.** *Let  $T$  be a nonnegative square matrix such that  $Tv \leq v$  with  $v > 0$ . Then  $\rho(T) \leq 1$ . If furthermore  $\rho(T) = 1$ , then  $\text{ind}_1 T = 1$ .*

In our convergence proofs, we will show that the iteration matrix satisfies  $T \geq O$ , and that there is a positive vector  $v$  for which  $Tv = v$ , and thus conclude using this lemma that  $\rho(T) = 1$  and  $\text{ind}_1 T = 1$ , i.e., (ii) holds. We will also show that the diagonals of  $T$  are positive, and thus (i) also holds.

We classify splittings  $A = M - N$  of being of *nonnegative type* if the matrix  $M^{-1}N$  is nonnegative [13], as *weak regular* if both  $M^{-1}$  and  $M^{-1}N$  are nonnegative, and as *regular* if both  $M^{-1}$  and  $N$  are nonnegative [4], [24]. In [2], it was shown that if  $M_i$  is given by (2.7), then each splitting  $A = M_i - N_i$  is a regular splitting, and thus a splitting of nonnegative type. The same proof applies here.

We begin our analysis by studying the diagonals of certain matrices.

---

<sup>1</sup>Some authors, e.g., [4], call this *semiconvergent*

LEMMA 3.2. *Let  $M_i$  be given by (2.7) and let  $A = M_i - N_i$  then all entries in the diagonals of  $T_i = M_i^{-1}N_i$ ,  $i = 1, \dots, p$ , are zero, and thus  $\text{diag}(\sum_{i=1}^p E_i T_i) = O$ , where  $E_i$  is given by (2.1).*

*Proof.* Let us fix  $i$  as any of the indices from 1 to  $p$ . Recall the combinatorial zero structure of  $M_i$  (2.8) and note that from (2.7) we have that  $M_i^{-1}$  has zero blocks in the same locations than  $M_i$ . Therefore we have

$$M_i^{-1}(j, k) = 0 \quad \text{if } j \neq k \text{ and } j \notin S_i \text{ or } k \notin S_i. \quad (3.1)$$

The structure of  $N_i = M_i - A$ , taking into account (2.7), is given by

$$N_i(j, k) = \begin{cases} 0 & \text{if } j \in S_i \text{ and } k \in S_i, \\ 0 & \text{if } j = k, \\ |A(j, k)| & \text{otherwise.} \end{cases} \quad (3.2)$$

We study the  $(j, j)$  entry of  $T_i = M_i^{-1}N_i$ . To that end, denote  $a_{jk} = M_i^{-1}(j, k)$ ,  $b_{jk} = N_i(j, k)$ , and write

$$T_i(j, j) = \sum_{k=1}^n a_{jk} b_{kj}.$$

Let us assume first that  $j \neq k$ . If one of the indices  $j$  or  $k$  (or both) does not belong to  $S_i$  then from (3.1) we obtain  $a_{jk} = 0$ . If on the contrary, if  $k$  and  $j$  belong to  $S_i$  then from (3.2) we get  $b_{kj} = 0$ . In the case  $j = k$  we also have that  $b_{jj} = 0$ . Then we conclude that  $a_{jk} b_{kj} = 0$  for all  $j, k = 1, \dots, n$  and therefore  $T_i(j, j) = 0$  for all  $j = 1, \dots, n$ . Since  $i$  is an arbitrary index,  $i = 1, \dots, p$ , the lemma follows.  $\square$

We are ready to prove the main convergence theorem.

THEOREM 3.3. *Let  $A = I - B$ , where  $B$  is an irreducible  $n \times n$  column stochastic matrix. Let  $p > 1$  be a positive integer and  $A = M_i - N_i$  be splittings of nonnegative type. Then, if  $\theta < 1/q$  the iteration matrix of the additive Schwarz given by (2.10) is convergent and  $T_\theta \geq 0$ .*

*Proof.* Since  $B$  is irreducible and column stochastic, it has a positive Perron vector  $v$ , i.e.,  $Bv = v$ . We will show the following

- (a)  $T_\theta v = v > 0$
- (b)  $T_\theta \geq 0$
- (c)  $\text{diag}(T_\theta) > (0, 0, \dots, 0)$

As discussed after Lemma 3.1, we will have then that  $T_\theta$  satisfies (i) and (ii) and it is thus convergent.

To prove (a) we write  $Av = (I - B)v = v - Bv = 0$  and from (2.10) we conclude that  $T_\theta v = v > 0$ .

To prove (b) we use (2.3) and the hypothesis that  $\theta < 1/q$ , and write

$$I \geq \frac{1}{q} \sum_{i=1}^p E_i \geq \theta \sum_{i=1}^p E_i,$$

from which we have

$$T_\theta = I - \theta \sum_{i=1}^p E_i M_i^{-1} A \geq \theta \sum_{i=1}^p E_i - \theta \sum_{i=1}^p E_i M_i^{-1} A = \theta \sum_{i=1}^p E_i (I - M_i^{-1} A) \geq O,$$

where we have used that  $(I - M_i^{-1}A) = M_i^{-1}N_i \geq O$ , since by hypothesis the splittings are of nonnegative type.

Finally, to prove (c) we have that if  $\theta < 1/q$ ,  $\text{diag}(I) > \text{diag}(\theta \sum_{i=1}^p E_i)$ , and then

$$\begin{aligned} \text{diag}(T_\theta) &= \text{diag}(I - \theta \sum_{i=1}^p E_i M_i^{-1} A) > \text{diag}(\theta \sum_{i=1}^p E_i - \theta \sum_{i=1}^p E_i M_i^{-1} A) \\ &= \theta \text{diag}(\sum_{i=1}^p E_i M_i^{-1} N_i) = O, \end{aligned}$$

where we have used Lemma 3.2.  $\square$

We mention that Theorem 3.3 is valid for any singular  $M$ -matrix  $A$  with a one-dimensional null space spanned by a positive vector  $v$ . In particular it applies to ergodic Markov chains.

We also mention that there is an induced splitting  $A = M - N$  with  $T_\theta = M^{-1}N$  such that it is weak regular. The next section is devoted to show this existence.

**4. Nonsingularity and Induced Splitting.** In this section we show that for the case  $q = 2$ , the matrix  $G = \sum_{i=1}^p E_i M_i^{-1}$  is nonsingular, and use this fact to analyze the induced splitting. We begin by laying the groundwork with some definitions and preliminary results. We then analyze the case of  $A_i$ ,  $i = 1, \dots, p$ , being consecutive principal submatrices of  $A$ . This leads the way to the general case when each  $A_i$  is any symmetric permutation of a principal submatrix of  $A$ , i.e., of the form (2.4).

The structure of the overlap is quite simple when the blocks are consecutive overlapping diagonal blocks. To handle this situation we consider submatrices of  $A$  defined by the subscripts in the sets  $S_i$  given by (2.2). We denote by  $A(S_i)$  a submatrix of  $A$  with rows and columns given by a set  $S_i$ . In particular, when the subscripts in  $S_i$  are consecutive, i.e., of the form  $\{i_1, i_1 + 1, i_1 + 2, \dots, i_1 + k - 1\}$ , the submatrix  $A(S_i)$  is called a *consecutive principal submatrix* of  $A$ , and is denoted by  $A([i, j])$  where  $i = i_1$ ,  $j = i_1 + k - 1$ . Moreover, denoting by  $F_{qr}$  an *elementary permutation matrix* obtained by permuting the rows  $q$  and  $r$  of an identity matrix  $I$ , a permutation matrix given by

$$\pi = F_{k, j+k-1} \cdots F_{2, j+1} F_{1, j} \quad (4.1)$$

is called a *consecutive permutation matrix*.

The following two lemmas are a direct consequence of the above definitions and therefore their proofs are omitted.

LEMMA 4.1. *If  $\pi_i$  is a consecutive permutation matrix given by (4.1) with  $j = j_i$  and  $k = n_i$ , then  $A_i$  given by (2.4) is the consecutive principal submatrix of order  $n_i$  of  $A$ , given by*

$$A_i = A([j_i, j_i + n_i - 1]). \quad (4.2)$$

LEMMA 4.2. *For each  $i = 1, \dots, p$ , let the matrices  $\pi_i$  be consecutive permutation matrices given by (4.1) with  $j = j_i$  and  $k = n_i$ . Let  $m_i = j_i + n_i$ . Then the rows and columns indexed by  $\ell = 1, \dots, j_i - 1$  and  $\ell = m_i, \dots, n$  of the matrix  $G_i$  given by (2.9) with  $A_i = A([j_i, j_i + n_i - 1])$  are zero rows and columns.*

From Lemma 4.2 and the definition of consecutive principal submatrix we have

$$A_i^{-1} = G_i([j_i, j_i + n_i - 1])$$

and, as long as each  $A_i$  is given by (4.2), we obtain that the position of each consecutive principal submatrix  $A_i$  in  $A$  is the same as the position of each  $A_i^{-1}$  in  $G_i$ .

We note that given  $p$  subdomains that cover the whole main diagonal of  $A$  and provided that each  $A_i$  is a consecutive principal submatrix and  $j_1 = 1, j_{i+1} \geq j_i$ , it is easy to see that the size of the overlapping block between  $A_i$  and  $A_{i+1}$  is given by

$$p_i = n_i + j_i - j_{i+1}, \quad i = 1, 2, \dots, p-1. \quad (4.3)$$

**THEOREM 4.3.** *Let  $A \in \mathbb{R}^{n \times n}$  be such that any principal submatrix is a non-singular  $M$ -matrix. Let  $p \geq 2$  subdomains and let the matrices  $A_i$  with  $n_i < n$ ,  $i = 1, \dots, p$ , be defined by consecutive set of indices  $S_i$ , such that  $S_1 = \{1, \dots, n_1\}$ ,  $S_i = \{j_i, \dots, j_i + n_i - 1\}$ ,  $i = 2, \dots, p-1$ , and  $S_p = \{j_p, \dots, n\}$  verifying that  $S_i \cap S_{i+k} = \emptyset$ ,  $k \geq 2$ , and provided that each index  $j$  belongs to some set  $S_i$ , for  $j = 1, \dots, n$ . Then  $G = \sum_{i=1}^p E_i M_i^{-1}$  is nonsingular.*

*Proof.* Let  $v$  be a vector in the null space of  $G$ , i.e.,  $Gv = 0$ . We want to show that  $v = 0$ . We assume for now that  $p > 3$ , and comment on the cases  $p = 2, 3$  later. If there is no overlap, i.e.,  $q = 1$ , we have a block diagonal matrix, and the theorem holds. We consider the case where there is some overlap, i.e.,  $q = 2$ .

We partition each consecutive principal submatrix  $A_i$ ,  $i = 2, \dots, p-1$ , into  $3 \times 3$  block corresponding to the two overlap portions and one without overlap as follows

$$A_i = \begin{bmatrix} A_{i,11} & A_{i,12} & A_{i,13} \\ A_{i,21} & A_{i,22} & A_{i,23} \\ A_{i,31} & A_{i,32} & A_{i,33} \end{bmatrix}, \quad (4.4)$$

where  $A_{i,11}$  is  $p_{i-1} \times p_{i-1}$ ,  $A_{i,33}$  is  $p_i \times p_i$ , correspond to the two overlap portions, and  $A_{i,22}$  is square of order  $n_i - p_i - p_{i-1}$ .  $A_1$  and  $A_p$  are similarly partitioned as

$$A_1 = \begin{bmatrix} A_{1,22} & A_{1,23} \\ A_{1,32} & A_{1,33} \end{bmatrix}, \quad A_p = \begin{bmatrix} A_{p,11} & A_{p,12} \\ A_{p,21} & A_{p,22} \end{bmatrix}. \quad (4.5)$$

The important observation is that because of the overlap, we have that

$$A_{i,33} = A_{i+1,11}, \quad i = 1, \dots, p-1. \quad (4.6)$$

We partition  $v$  conformally to the partition of these blocks as

$$v^T = [v_{1,2}, v_{2,1}, v_{2,2}, v_{3,1}, v_{3,2}, \dots, v_{p-1,2}, v_{p,1}, v_{p,2}],$$

where  $v_{i,1} \in \mathbb{R}^{p_{i-1}}$ ,  $i = 2, \dots, p$ , are the overlapping variables.





With the nonsingularity of  $G$  established, we can show the existence of an induced splitting.

**THEOREM 4.5.** *Let  $A \in \mathbb{R}^{n \times n}$  be such that any principal submatrix is a nonsingular  $M$ -matrix. Let  $p \geq 2$  subdomains and let the matrices  $A_i$  be given by (2.4) with  $n_i < n$ , and defined by set of indices  $S_i$  verifying that  $S_i \cap S_{i+k} = \emptyset$ ,  $k \geq 2$ , and provided that each index  $j$  belongs to some set  $S_i$ , for  $j = 1, \dots, n$ . Then there exists a splitting  $A = M - N$  such that the iteration matrix  $T_\theta$  is given by  $T_\theta = M^{-1}N$ , where*

$$M^{-1} = \theta \sum_{i=1}^p E_i M_i^{-1}. \quad (4.16)$$

Furthermore, this splitting is a convergent weak regular splitting.

*Proof.* From Theorem 4.4 we have that  $G = \sum_{i=1}^p E_i M_i^{-1}$  is nonsingular. We can then define  $M = (1/\theta)G^{-1}$ , and  $N$  such that  $A = M - N$ . It follows that

$$T_\theta = I - M^{-1}A, \quad (4.17)$$

is the iteration matrix of the additive Schwarz iterations. For the second part, it is easy to see that all the conditions of Theorem 3.3 hold and therefore  $T_\theta = M^{-1}N \geq O$ . As we also have from (4.16) that  $M^{-1} \geq O$  we conclude that  $A = M - N$  is a weak regular splitting. Using Theorem 3.3 again,  $T_\theta$  is convergent.  $\square$

We comment here that the splitting defined by (4.16) is not the only splitting associated with  $T_\theta$ , there are infinitely many splittings  $A = M - N$  for which (4.17) holds [3].

**5. Illustrative Examples.** We begin with an example in which the induced splitting  $A = M - N$  is (convergent) weak regular for  $\theta = 0.3$  (in accordance with Theorem 4.5) but it is not a regular splitting.

**EXAMPLE 5.1.** Consider the matrix

$$A = \begin{bmatrix} 4/7 & -6/23 & -6/23 & -5/7 & -18/31 \\ -2/21 & 83/92 & -9/92 & 0 & -2/93 \\ -4/21 & -27/92 & 65/92 & 0 & -4/93 \\ -1/7 & -9/46 & -9/46 & 6/7 & -8/31 \\ -1/7 & -7/46 & -7/46 & -1/7 & 28/31 \end{bmatrix} \quad (5.1)$$

which is a singular  $M$ -matrix, and  $A = I - B$  with  $B$  an  $n \times n$  irreducible column stochastic matrix. Let us consider two subdomains such that  $p = q = 2$ . We use

$$R_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad R_2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

then, we have  $S_1 = \{1, 3, 5\}$ ,  $S_2 = \{2, 3, 4, 5\}$ , and

$$A_1 = \begin{bmatrix} 28/31 & -7/46 & -1/7 \\ -4/93 & 65/92 & -4/21 \\ -18/31 & -6/23 & 4/7 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} 65/92 & 0 & -4/93 & -27/92 \\ -9/46 & 6/7 & -8/31 & -9/46 \\ -7/46 & -1/7 & 28/31 & -7/46 \\ -9/92 & 0 & -2/93 & 83/92 \end{bmatrix}.$$

It follows that the splittings  $A = M_1 - N_1$  and  $A = M_2 - N_2$  are regular splittings, as expected, and one has

$$T_1 = M_1^{-1}N_1 = \begin{bmatrix} 0 & 9702/7157 & 0 & 15313/7157 & 0 \\ 184/1743 & 0 & 9/83 & 0 & 184/7719 \\ 0 & 5815/7157 & 0 & 30728/50099 & 0 \\ 1/6 & 21/92 & 21/92 & 0 & 28/93 \\ 0 & 3720/7157 & 0 & 8587/14314 & 0 \end{bmatrix},$$

$$T_2 = M_2^{-1}N_2 = \begin{bmatrix} 0 & 21/46 & 21/46 & 5/4 & 63/62 \\ 13861/91938 & 0 & 0 & 0 & 0 \\ 32231/91938 & 0 & 0 & 0 & 0 \\ 74/199 & 0 & 0 & 0 & 0 \\ 4619/15323 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Note that  $\text{diag}(T_i) = O$ , according to Lemma 3.2. For the splitting  $A = M - N$ , with  $M$  given by (4.16) with  $\theta = 0.3$ , we have

$$M = \begin{bmatrix} 89825/56742 & 1865/4439 & -10/23 & 3005/18914 & -30/31 \\ 9085/170226 & 155995/53268 & -15/92 & -415/56742 & -10/279 \\ -20/63 & -45/92 & 325/276 & 0 & -20/279 \\ 5120/28371 & -7955/8878 & -15/46 & 26340/9457 & -40/93 \\ -5/21 & -35/138 & -35/138 & -5/21 & 140/93 \end{bmatrix},$$

for which  $M^{-1} \geq O$ ,

$$N = \begin{bmatrix} 57401/56742 & 3023/4439 & -4/23 & 16515/18914 & -12/31 \\ 25297/170226 & 53969/26634 & -3/46 & -415/56742 & -4/279 \\ -8/63 & -9/46 & 65/138 & 0 & -8/279 \\ 9173/28371 & -3109/4439 & -3/23 & 18234/9457 & -16/93 \\ -2/21 & -7/69 & -7/69 & -2/21 & 56/93 \end{bmatrix},$$

and  $T_\theta = M^{-1}N \geq O$  with  $\sigma(T_\theta) = \{1, 2/5, 2/5, 2/5, 7/10\}$ .

For the matrix (5.1) we present in Table 5.2, the value of the convergence factor  $\gamma(T)$  for several cases of (damped) block Jacobi and (damped) additive Schwarz. The different variables in each of the blocks are given in Table 5.1. In all the cases considered, we have that the maximum overlap is  $q = 2$ . It can be readily appreciated that for the same damping factor  $\theta$  additive Schwarz has a lower convergence factor than block Jacobi.

Method	$p$	Variables in each block $S_i$
Block Jacobi (BJ1)	2	$S_1 = \{1, 3\}, S_2 = \{2, 4, 5\}$
Block Jacobi (BJ2)	2	$S_1 = \{1, 5\}, S_2 = \{2, 3, 4\}$
Additive Schwarz (AS1)	2	$S_1 = \{1, 3, 5\}, S_2 = \{2, 3, 4, 5\}$
Block Jacobi (BJ3)	3	$S_1 = \{1, 3\}, S_2 = \{2, 5\}, S_3 = \{4\}$
Additive Schwarz (AS2)	3	$S_1 = \{1, 3, 5\}, S_2 = \{2, 3, 5\}, S_3 = \{1, 4\}$

TABLE 5.1

EXAMPLE 5.2. We consider matrices representing a model of an interactive computer system, described as Example 1 in [16] or in [21]. We used the matrices of orders

$\theta$	BJ1	BJ2	AS1	BJ3	AS2
0.1	0.90003	0.90011	0.90000	0.8996	0.8892
0.2	0.80014	0.80051	0.80000	0.7992	0.7784
0.3	0.70036	0.70131	0.70000	0.6989	0.6675
0.4	0.60074	0.60271	0.60000	0.5987	0.5567
0.5	0.50139	0.50507	0.50000	0.4986	0.4459

TABLE 5.2

Convergence factor  $\gamma(T)$  for blocks described in Table 5.1. Matrix  $A$  is (5.1)

$n = 286, 1771,$  and  $5456$ , with  $1606, 11011,$  and  $35216$  nonzeros, respectively. These matrices have been used for numerical experiments by several authors, including [7], [8], [15], [16], [21], and can be obtained from [19]. We ran several experiments with varying blocks using (damped) block Jacobi and (damped) additive Schwarz. Our initial vector is  $x^0 = (1, 1, \dots, 1)$ . We report the convergence factor  $\gamma(T_\theta)$ , the number of iterations and the CPU time for convergence, i.e., for the 2-norm of the residual  $r = Ax^k$  to be reduced to less than  $10^{-10}$ . The numerical experiments are executed in only one processor of an IBM-1350 at the Universitat Politècnica de València, running Matlab 6. The times are in seconds. For comparison, we also report these three quantities for block Gauss-Seidel (GS) for multiplicative Schwarz (MS), i.e. using  $T_\mu$  in selected cases. For block Jacobi, the best results are for  $\theta = 0.90$ , and this is the one we report in Table 5.3. Table 5.4 shows the structure of the subdomains. These experiments illustrate the potential advantage of Schwarz methods, i.e., of the overlap, over the widely used classical iterative methods for Markov chains.

$n$		$p$	$q$	$\theta$	$\gamma(T_\theta)$	iter.	time	$\gamma(T_\mu)$	iter.	time
286	BJ	3	1	0.90	0.9992	13028	29.57			
286	BJ	5	1	0.90	0.9997	32439	86.79			
286	AS	3	2	0.49	0.9498	278	0.81			
286	AS	5	3	0.49	0.9914	1392	3.93			
286	AS	5	2	0.49	0.9316	164	0.40			
286	AS	3	2	0.81	0.9170	166	0.45			
286	AS	5	3	0.66	0.9885	1032	3.12			
286	AS	5	2	0.63	0.9120	127	0.31			
1771	BJ/GS	6	1	0.90	0.9978	2178	58.00	0.9951	1209	106.7
1771	AS/MS	6	2	0.49	0.9882	217	9.39	0.9522	107	9.90
1771	AS	6	2	0.90	0.9782	118	5.29			
5456	BJ/GS	10	1	0.90	0.9988	9592	1636.5	0.9972	4551	3292.
5456	AS/MS	10	2	0.49	0.9716	175	45.09	0.8876	45	32.00
5456	AS	10	2	0.90	0.9479	95	32.15			

TABLE 5.3

**6. Conclusion.** We have presented a convergence theory for additive Schwarz methods for linear systems whose matrix is a singular  $M$ -matrix with a one-dimensional null space. This applies in particular to ergodic Markov chains. The convergence theory is illustrated with several numerical experiments, where it is shown that using additive Schwarz, i.e., incorporating overlap variables common to more than one block, improves the asymptotic convergence factor, as well as execution times.

$n$		$p$	$q$	$S_i$
286	BJ	3	1	$\{1 : 100\}, \{101 : 200\}, \{201 : 286\}$
286	BJ	5	1	$\{1 : 66\}, \{67 : 139\}, \{140 : 216\}, \{217 : 273\}, \{274 : 286\}$
286	AS	3	2	$\{1 : 100\}, \{80 : 200\}, \{120 : 286\}$
286	AS	5	3	$\{1 : 30\}, \{25 : 70\}, \{25 : 80\}, \{40 : 150\}, \{100 : 286\}$
286	AS	5	2	$\{1 : 83\}, \{47 : 131\}, \{84 : 178\}, \{132 : 225\}, \{179 : 286\}$
286	AS	3	2	$\{1 : 100\}, \{80 : 200\}, \{120 : 286\}$
286	AS	5	3	$\{1 : 30\}, \{25 : 70\}, \{25 : 80\}, \{40 : 150\}, \{100 : 286\}$
286	AS	5	2	$\{1 : 83\}, \{47 : 131\}, \{84 : 178\}, \{132 : 225\}, \{179 : 286\}$
1771	BJ/GS	6	1	$\{1 : 300\}, \{301 : 601\}, \{602, 902\}, \{903 : 1203\},$ $\{1204 : 1503\}, \{1504 : 1771\}$
1771	AS/MS	6	2	$\{1 : 300\}, \{281 : 600\}, \{500 : 900\}, \{800 : 1200\},$ $\{1000 : 1500\}, \{1400 : 1771\}$
1771	AS	6	2	$\{1 : 300\}, \{281 : 600\}, \{500 : 900\}, \{800 : 1200\},$ $\{1000 : 1500\}, \{1400 : 1771\}$
5456	BJ/GS	10	1	$\{1 : 600\}, \{601 : 1200\}, \{1201 : 1800\}, \{1801 : 2400\},$ $\{2401 : 3000\}, \{3001 : 3600\}, \{3601 : 4200\},$ $\{4201 : 4800\}, \{4801 : 5400\}, \{5401 : 5456\}$
5456	AS/MS	10	2	$\{1 : 900\}, \{601 : 1500\}, \{1201 : 2100\}, \{1801 : 2700\},$ $\{2401 : 3300\}, \{3001 : 3900\}, \{3601 : 4500\},$ $\{4201 : 5100\}, \{4801 : 5455\}, \{5401 : 5456\}$
5456	AS	10	2	$\{1 : 900\}, \{601 : 1500\}, \{1201 : 2100\}, \{1801 : 2700\},$ $\{2401 : 3300\}, \{3001 : 3900\}, \{3601 : 4500\},$ $\{4201 : 5100\}, \{4801 : 5455\}, \{5401 : 5456\}$

TABLE 5.4

**Acknowledgment.** We thank Michele Benzi for his comments on an earlier manuscript.

## REFERENCES

- [1] G. Alefeld and N. Schneider. On square roots of  $M$ -matrices. *Linear Algebra and its Applications*, 42:119–132, 1982.
- [2] M. Benzi, A. Frommer, R. Nabben, and D. B. Szyld. Algebraic theory of multiplicative Schwarz methods, *Numerische Mathematik*, 89:605–639, 2001.
- [3] M. Benzi and D. B. Szyld. Existence and uniqueness of splittings for stationary iterative methods with applications to alternating methods. *Numerische Mathematik*, 76:309–321, 1997.
- [4] A. Berman and R. J. Plemmons. *Nonnegative matrices in the mathematical sciences*, Academic Press, New York, 1979. Reprinted and updated, SIAM, Philadelphia, 1994.
- [5] E. Bohl and I. Marek. A model of amplification. *Journal of Computational and Applied Mathematics*, 63:27–47, 1995.
- [6] R. Bru, F. Pedroche, and D. B. Szyld. Overlapping additive and multiplicative Schwarz iterations for  $H$ -matrices. *Linear Algebra and its Applications*, 393:91–105, 2004.
- [7] H. Choi and D. B. Szyld. Application of threshold partitioning of sparse matrices to Markov chains. In *Proceedings of the IEEE International Computer Performance and Dependability Symposium, IPDS'96, Urbana-Champaign, Illinois, September 4–6, 1996*, pages 158–165, Los Alamitos, California, 1996. IEEE Computer Society Press.
- [8] T. Dayar and W. J. Stewart. Comparison of partitioning techniques for two-level iterative methods on large, sparse Markov chains. *SIAM Journal on Scientific Computing*, 21:1691–1705, 2000.
- [9] A. Frommer and D. B. Szyld.  $H$ -Splittings and two-stage iterative methods, *Numerische Mathematik*, 63:345–356, 1992.

- [10] A. Frommer and D. B. Szyld. Weighted Max Norms, Splittings, and Overlapping Additive Schwarz Iterations, *Numerische Mathematik*, 83:259–278, 1999.
- [11] I. Marek and D. B. Szyld. Comparison theorems for the convergence factor of iterative methods for singular matrices, *Linear Algebra and its Applications*, 316:67–87, 2000.
- [12] I. Marek and D. B. Szyld. Comparison of convergence of general stationary iterative methods for singular matrices. *SIAM Journal on Matrix Analysis and Applications*, 24:68–77, 2002.
- [13] I. Marek and D. B. Szyld. Algebraic Schwarz methods for the numerical solution of Markov chains, *Linear Algebra and its Applications*, 386:67–81, 2004.
- [14] T. P. Mathew. Uniform convergence of the Schwarz alternating method for solving singularly perturbed advection-diffusion equations. *SIAM Journal on Numerical Analysis*, 35:1663–1683, 1998.
- [15] V. Migallón, J. Penadés, and D. B. Szyld. Experimental Study of Parallel Iterative Solutions of Markov Chains with Block Partitions. In B. Plateau, W. J. Stewart, and M. Silva, *Numerical Solution of Markov Chains (NSMC'99)*, 1999, pages 96–110. Prensas Universitarias de Zaragoza, Spain.
- [16] B. Philippe, Y. Saad, and W. J. Stewart. Numerical methods in Markov chain modeling. *Operations Research*, 40:1156–1179, 1992.
- [17] A. Quarteroni and A. Valli. *Domain Decomposition Methods for Partial Differential Equations*. Oxford Science Publications, Clarendon Press, Oxford, 1999.
- [18] B. F. Smith, P. E. Bjørstad, and W. D. Gropp. *Domain decomposition: Parallel multilevel methods for elliptic partial differential equations*. Cambridge University Press, Cambridge, 1996.
- [19] W. J. Stewart. MARCA Models: A collection of Markov chain models. Available online at [http://www.csc.ncsu.edu/faculty/WStewart/MARCA\\_Models/MARCA\\_Models.html](http://www.csc.ncsu.edu/faculty/WStewart/MARCA_Models/MARCA_Models.html).
- [20] W. J. Stewart. *Introduction to the Numerical Solution of Markov Chains*. Princeton University Press, Princeton, N.J., 1994.
- [21] W. J. Stewart and W. Wu. Numerical experiments with iteration and aggregation for Markov chains. *ORSA Journal on Computing*, 4:336–350, 1992.
- [22] D. B. Szyld. Equivalence of convergence conditions for iterative methods for singular equations. *Numerical Linear Algebra with Applications* 1:151–154, 1994.
- [23] A. Toselli and O. B. Widlund. *Domain Decomposition: Algorithms and Theory*. Springer, Heidelberg, 2004.
- [24] R. S. Varga. *Matrix Iterative Analysis*. Prentice-Hall, Englewood Cliffs, New Jersey, 1962. Second Edition, revised and expanded, Springer, Berlin, Heidelberg, New York, 2000.