Numer. Math. 89: 605–639 (2001)
Digital Object Identifier (DOI) 10.1007/s002110100275

Numerische Mathematik

Algebraic theory of multiplicative Schwarz methods

Michele Benzi[1,*], **Andreas Frommer**[2], **Reinhard Nabben**[3], **Daniel B. Szyld**[4,**]

[1] Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322, USA; e-mail: benzi@mathcs.emory.edu
[2] Fachbereich Mathematik, Universität Wuppertal, 42097 Wuppertal, Germany; e-mail: frommer@math.uni-wuppertal.de
[3] Fakultät für Mathematik, Universität Bielefeld, Postfach 10 01 31, 33501 Bielefeld, Germany; e-mail: nabben@mathematik.uni-bielefeld.de
[4] Department of Mathematics, Temple University, Philadelphia, PA 19122-6094, USA; e-mail: szyld@math.temple.edu

Received February 21, 2000 / Revised version received July 12, 2000 / Published online April 5, 2001 –

Summary. The convergence of multiplicative Schwarz-type methods for solving linear systems when the coefficient matrix is either a nonsingular M-matrix or a symmetric positive definite matrix is studied using classical and new results from the theory of splittings. The effect on convergence of algorithmic parameters such as the number of subdomains, the amount of overlap, the result of inexact local solves and of "coarse grid" corrections (global coarse solves) is analyzed in an algebraic setting. Results on algebraic additive Schwarz are also included.

Mathematics Subject Classification (1991): 65F10, 65F35, 65M55

1. Introduction

We consider the solution of large sparse linear systems of the form

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

by multiplicative or additive Schwarz methods. Our aim is to apply the theory of matrix splittings to study the convergence of these classes of methods,

---

* This work was performed under the auspices of the U.S. Department of Energy through grant W-7405-ENG-36 with Los Alamos National Laboratory.
** This work was supported by the National Science Foundation grant DMS-9973219.
Correspondence to: D.B. Szyld

using properties of the coefficient matrix only. Specifically, we analyze two cases: the case where the coefficient matrix A is a nonsingular M-matrix, and when A is symmetric positive definite (s.p.d.). As we shall see, in several situations there is a nice common theory in the treatment of these two cases, using the appropriate splittings for each case. The exceptions are Sects. 5 and 6 where for multiplicative Schwarz only the M-matrix case is studied.

While several convergence results on Schwarz methods exist when the matrix A in (1) corresponds to the discretization of a differential equation (see, e.g., [9], [40], [44], and the extensive bibliography therein), there is a need to analyze these methods in a purely algebraic setting. As we show, there are instances where the tools developed here provide convergence analysis not available with the usual Sobolev space theory. We believe that the algebraic and analytical points of view complement each other. Furthermore, there are applications, such as electrical power networks and Leontief models in economics, where the matrix A does not come from a differential equation (and it is an M-matrix); see, respectively, [10] and [2]. Another case of interest is when the problem arises from the discretization of a differential equation but no geometric information about the underlying mesh is available to the solver. Additionally, an algebraic approach is useful for the case of unstructured meshes [7].

There are several papers with detailed abstract analysis (i.e., independent of the particular differential equations in question) of Schwarz methods, including those of Xu [51] and Griebel and Oswald [23], where A-norm bounds are obtained for the symmetric positive definite case; see also the nicely written survey [25]. In other cases, e.g., in [32], the maximum principle is used to show convergence.

In this paper we concentrate on the case of algebraic multiplicative Schwarz, although we include new results on additive Schwarz, and hybrid methods as well. We emphasize methods where overlap is used, i.e., when the same variable is present in more than one local solver. The present work can be seen as a continuation of [20] where algebraic additive Schwarz was considered, and it complements the heuristic study [7].

While we do not provide condition number estimates for a preconditioned system, our convergence results point out to the usefulness of the multilevel methods as solution methods as well as preconditioners for a wider class of problems. In the s.p.d. case we are able to prove convergence without the usual assumptions; see Remark 4.11 below. In the nonsymmetric case, we can prove convergence without any condition on the coarse grid correction, and in fact convergence is shown without the need for a coarse grid correction; see Remark 3.6.

Given an initial approximation x^0 to the solution of (1), the (one-level) multiplicative Schwarz method can be written as the stationary iteration

$$x^{k+1} = Tx^k + c, \qquad k = 0, 1, \ldots, \tag{2}$$

where

$$T = (I - P_p)(I - P_{p-1}) \cdots (I - P_1) = \prod_{i=p}^{1} (I - P_i) \tag{3}$$

and c is a certain vector. Here

$$P_i = R_i^T (R_i A R_i^T)^{-1} R_i A$$

where R_i is a matrix of dimension $n_i \times n$ with full row rank, $1 \leq i \leq p$. In the case of overlap we have $\sum_{i=1}^{p} n_i > n$. Note that each P_i, and hence each $I - P_i$, is a projection operator; i.e., $(I - P_i)^2 = I - P_i$. Each $I - P_i$ is singular and has spectral radius equal to 1. Yet, as we will see, the product T given by (3) has spectral radius strictly less than 1 under suitable assumptions. In fact, for an appropriate norm $\|I - P_i\| = 1$, $i = 1, \ldots, p$ (as is well known for A s.p.d.) but the product matrix T has norm less than 1.

The matrix R_i corresponds to the restriction operator from the whole space to a subdomain Ω_i (of dimension n_i) in the domain decomposition setting, and the matrix $A_i = R_i A R_i^T$ is the restriction of A to that subdomain. A solution using A_i is called a local solve, and this name carries to the purely algebraic case. Our approach consists in determining the unique splitting $A = B - C$ with B invertible and such that $T = B^{-1}C$, and to study the properties of that splitting; see Lemma 2.1 below. In this way we can exploit the rich theory of matrix splittings and prove convergence under appropriate conditions.

In this paper we emphasize the use of Schwarz methods as solvers rather than preconditioners. We note that when used as a preconditioner, particularly in the case of symmetric positive definite problems and the conjugate gradient method, the multiplicative Schwarz method is usually symmetrized; that is, the application of the p projections in (3) is followed by another sweep of projections applied in the reverse order. Many of the results and techniques of this paper can be applied to the symmetrized iterations.

There are a number of papers dealing with algebraic Schwarz methods, including [7], [16], [18], [19], [42], [43], [46], [53]; see also [29]. In many of these, only special cases, such as tridiagonal, or block-tridiagonal matrices, or matrices derived from a particular model problem, are studied. Our contribution is to provide convergence results for multiplicative Schwarz methods (with overlapping blocks) for general M-matrices and for s.p.d.

matrices. We present our convergence bounds in terms of matrix norms as well as spectral radii, and use both of these to compare the convergence of different versions. In particular, we analyze the effect on convergence of algorithmic parameters such as the number of blocks (or subdomains) p, the amount of overlap, inexact local solves, and the effect of adding coarse grid corrections (both multiplicatively and additively).

2. Auxiliary results

The purpose of this section is to introduce some notation and a few results that will be used extensively in the remainder of the paper. A matrix B is nonnegative (positive), denoted $B \geq O$ ($B > O$) if its entries are nonnegative (positive). We say that $B \geq C$ if $B - C \geq O$, and similarly with the strict inequality. These definitions carry over to vectors. A matrix A is a nonsingular M-matrix if its off-diagonal elements are nonpositive, and it is monotone, i.e., $A^{-1} \geq O$. It follows that if A and B are nonsingular M-matrices and $A \geq B$, then $A^{-1} \leq B^{-1}$ [2], [49]. By $\rho(B)$ we denote the spectral radius of the matrix B.

A matrix B is symmetric positive definite (s.p.d.), denoted $B \succ O$, if it is symmetric and if for all vectors $u \neq 0$, $u^T Bu > 0$, and positive semidefinite, denoted $B \succeq O$, if for all vectors $u \neq 0$, $u^T Bu \geq 0$. We say that $B \succeq C$ if $B - C \succeq O$, and similarly with the strict inequality. It follows that if A and B are s.p.d. and $A \succeq B$, then $A^{-1} \preceq B^{-1}$. If $A \succ O$ one can define the A-norm of a vector x as $\|x\|_A = (x^T Ax)^{1/2}$. This vector norm induces a matrix norm in the usual way.

We say that $A = M - N$ is a splitting if M is nonsingular. The splitting is regular if $M^{-1} \geq O$ and $N \geq O$; it is weak regular if $M^{-1} \geq O$ and $M^{-1}N \geq O$; and it is nonnegative if $M^{-1} \geq O$, $M^{-1}N \geq O$, and $NM^{-1} \geq O$ [2], [49], [50]. The splitting is P-regular if $M^T + N$ is positive definite [39]. Note that if A is symmetric $M^T + N = M^T + M - A$ is also symmetric. We say that a splitting is a *strong P-regular splitting* of A s.p.d., when $N \succeq O$. This implies that $M \succ O$ and that in particular it is a P-regular splitting. The following result, which can be found, e.g., in [1], shows that given an iteration matrix, there is a unique splitting for it.

Lemma 2.1. *Let A and T be square matrices such that A and $I - T$ are nonsingular. Then, there exists a unique pair of matrices B, C, such that B is nonsingular, $T = B^{-1}C$ and $A = B - C$. The matrices are $B = A(I - T)^{-1}$ and $C = B - A = A((I - T)^{-1} - I)$.*

The following characterization of P-regular splittings will be useful in our analysis; for a proof, see [20], or [52].

Lemma 2.2. *Let A be symmetric positive definite. Then $A = M - N$ is a P-regular splitting if and only if $\|M^{-1}N\|_A < 1$.*

In this paper we assume that the rows of R_i are rows of the $n \times n$ identity matrix I, e.g.,

$$R_i = \begin{bmatrix} 0\,0\,0\,0\,0\,0\,1\,0 \\ 0\,1\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,1\,0\,0\,0\,0 \end{bmatrix}.$$

This restriction operator is often called a Boolean gather operator, while its transpose R_i^T is called a Boolean scatter operator. Formally, such a matrix R_i can be expressed as

$$R_i = [I_i|O]\,\pi_i \tag{4}$$

with I_i the identity on $\mathbb{R}^{n_i}$ and π_i a permutation matrix on $\mathbb{R}^n$. In this case, it follows that A_i is an $n_i \times n_i$ principal submatrix of A. In fact, we can write

$$\pi_i A \pi_i^T = \begin{bmatrix} A_i & K_i \\ L_i & A_{\neg i} \end{bmatrix}, \tag{5}$$

where $A_{\neg i}$ is the principal submatrix of A "complementary" to A_i, i.e.

$$A_{\neg i} = [O|I_{\neg i}] \cdot \pi_i \cdot A \cdot \pi_i^T \cdot [O|I_{\neg i}]^T \tag{6}$$

with $I_{\neg i}$ the identity on $\mathbb{R}^{n-n_i}$. Recall that if A is an M-matrix, so are its principal submatrices, and thus both A_i and $A_{\neg i}$ are M-matrices [2]. Similarly if A is s.p.d., then, both A_i and $A_{\neg i}$ are s.p.d.

For each $i = 1, \ldots, p$, we construct diagonal matrices $E_i \in \mathbb{R}^{n \times n}$ associated with R_i from (4) as follows

$$E_i = R_i^T R_i. \tag{7}$$

These diagonal matrices have ones on the diagonal in every row where R_i^T has nonzeros. We further assume that if S_i is the set of indexes of the rows of the identity that are rows of R_i, then

$$\bigcup_{i=1}^{p} S_i = S = \{1, 2, \ldots, n\}. \tag{8}$$

In other words each variable is in at least one set S_i. This is equivalent to saying that $\sum_{i=1}^{p} E_i \geq I$, with equality if and only if there is no overlap. Note that in the case of overlapping blocks, we have here that each diagonal entry of $\sum_{i=1}^{p} E_i$ is greater than or equal to one, which implies nonsingularity. Only in the rows corresponding to overlap this matrix has an entry different from one. In the case of overlap, the maximum that these entries can attain is q, the measure of overlap defined below. We thus have that $\sum_{i=1}^{p} E_i \leq qI$.

Let us define a *measure of overlap* q of the decomposition (8) as the minimal number of sets V_k ($k = 1, \ldots, q$) such that

$$\bigcup_{k=1}^{q} V_k = \bigcup_{i=1}^{p} S_i = S = \{1, 2, \ldots, n\}, \tag{9}$$

where each S_i is a subset of some V_k, and if

$$S_i \subset V_k \text{ and } S_j \subset V_k \text{ for the same } k,\ i \neq j, \text{ then } S_i \cap S_j = \emptyset. \tag{10}$$

In other words, the measure of overlap is

$$q = \max_{j=1,\ldots,n} |\{i : j \in S_i\}|,$$

and obviously $q = 1$ implies that there is no overlap.

Following Hackbusch [24, Ch. 11], we define the *number of colors* $\tilde{q}$ of the decomposition (8) as the number of sets V_k satisfying (9), (10), and in addition, if $r \in S_i$, $s \in S_j$, then the matrix entries $a_{rs} = a_{sr} = 0$. It follows that $q \leq \tilde{q}$, and often this inequality is strict. Furthermore, q depends only on the partition of the variables, while $\tilde{q}$ also on the graph of the matrix A. As we shall see, these two quantities are used in the study of convergence of the additive Schwarz method. The measure of overlap is relevant in the M-matrix case, while the number of colors in the s.p.d. case.

We illustrate the concepts of measure of overlap and number of colors with two examples. Consider the 10×10 matrix

$$A = \begin{bmatrix} A_{11} & O & O & A_{1,4} & O \\ O & A_{2,2} & O & A_{2,4} & A_{2,5} \\ O & O & A_{3,3} & O & A_{3,5} \\ A_{4,1} & A_{4,2} & O & A_{4,4} & O \\ O & A_{5,2} & A_{5,3} & O & A_{5,5} \end{bmatrix},$$

where all diagonal blocks $A_{i,i}$, $i = 1, \ldots, 5$ are 2×2 matrices. Now let $S_i = \{2i - 1, 2i\}$, $i = 1, \ldots, 5$, i.e. there is no overlap. Hence, we have for the measure of overlap $q = 1$. We have only one set $V_1 = \bigcup_{i=1}^{5} S_i$. The number of colors is $\tilde{q} = 2$ with $V_1 = S_1 \cup S_2 \cup S_3, V_2 = S_4 \cup S_5$. If we take

$$S_1 = \{1, 2, 7, 8\},\ S_2 = \{3, 4\},\ S_3 = \{5, 6, 9, 10\},$$
$$S_4 = \{7, 8\},\ S_5 = \{9, 10\}$$

we have $q = 2$ with $V_1 = S_1 \cup S_2 \cup S_3$ and $V_2 = S_4 \cup S_5$. The number of colors is $\tilde{q} = 3$ with $V_1 = S_1 \cup S_3$, $V_2 = S_2$ and $V_3 = S_4 \cup S_5$.

If A is a nonsingular M-matrix, for each $i = 1, \ldots, p$, we construct a second set of matrices $M_i \in \mathbb{R}^{n\times n}$ associated with R_i from (4) as follows

$$M_i = \pi_i^T \begin{bmatrix} A_i & O \\ O & D_{\neg i} \end{bmatrix} \pi_i, \tag{11}$$

where

$$D_{\neg i} = \text{diag}(A_{\neg i}) \geq O \tag{12}$$

has positive entries along the diagonal and thus is invertible.

Proposition 2.3. *Let A be a nonsingular M-matrix. Let M_i be defined as in* (11). *Then the splittings $A = M_i - N_i$ are regular (and thus weak regular and nonnegative).*

Proof. Observe that

$$M_i^{-1} = \pi_i^T \begin{bmatrix} A_i^{-1} & O \\ O & D_{\neg i}^{-1} \end{bmatrix} \pi_i$$

is nonnegative. Thus, M_i is an M-matrix. Moreover $N_i = M_i - A$ is nonnegative, since it is a symmetric permutation of a matrix with a 2 by 2 block structure, the off-diagonal blocks being nonnegative and the diagonal blocks being either zero, or nonnegative with a zero diagonal. $\square$

With the definitions (7) and (11) we obtain the following equality which we will use throughout the paper

$$E_i M_i^{-1} = R_i^T A_i^{-1} R_i, \; i = 1, \ldots, p. \tag{13}$$

We note that the matrix M_i defined in (11) is different from the one used in [20], although we obtain the same characterization (13). All results in [20] hold *verbatim* for this different choice of M_i. In fact, we have a great deal of flexibility in choosing the matrices M_i, as long as the equality (13) holds. We will take advantage of this flexibility in sections 4–6 when analyzing the change in the convergence rate by varying the degree of overlap, the number of blocks (subdomains) and the level of inexactness of the local solves. For the analysis of the s.p.d. case, we choose a different set of matrices M_i satisfying (13), namely the choice made in [20]. We abuse the notation, but in each case it is clear from the context which matrix it is we are using.

Let A be s.p.d. For each $i = 1, \ldots, p$, we construct matrices $M_i \in \mathbb{R}^{n\times n}$ associated with R_i as follows

$$M_i = \pi_i^T \begin{bmatrix} A_i & O \\ O & A_{\neg i} \end{bmatrix} \pi_i. \tag{14}$$

It follows that M_i is s.p.d., and that it satisfies the identity (13).

Proposition 2.4. *Let A be a symmetric positive definite matrix. Let M_i be defined as in* (14). *Then, the splittings $A = M_i - N_i$ are P-regular.*

Proof. Since $A_i^T = A_i$ and $A_{\neg i}^T = A_{\neg i}$, we write

$$M_i^T + N_i = \pi_i^T \begin{bmatrix} A_i & -K_i \\ -L_i & A_{\neg i} \end{bmatrix} \pi_i.$$

The following identity shows that this matrix is s.p.d., and thus we have a P-regular splitting:

$$\pi_i^T \begin{bmatrix} I & \\ & -I \end{bmatrix} \begin{bmatrix} A_i & -K_i \\ -L_i & A_{\neg i} \end{bmatrix} \begin{bmatrix} I & \\ & -I \end{bmatrix} \pi_i = \pi_i^T \begin{bmatrix} A_i & K_i \\ L_i & A_{\neg i} \end{bmatrix} \pi_i = A.$$

□

Given a positive vector $w \in \mathbb{R}^n$, denoted $w > 0$, the weighted max-norm is defined for any $y \in \mathbb{R}^n$ as $\|y\|_w = \max_{j=1,\cdots,n} |\frac{1}{w_j} y_j|$; see, e.g., [26], [41]. As usual the matrix norm is defined as $\|T\|_w = \sup_{\|x\|_w=1} \|Tx\|_w$, and can be obtained as (see, e.g., [41])

$$\|T\|_w = \max_i \frac{(|T|w)_i}{w_i}. \tag{15}$$

We point out that for $T \geq O$, $Tw < \gamma w$ implies $\|T\|_w < \gamma$ ($\gamma > 0$) [41]. Weighted max norms play a fundamental role in the study of asynchronous methods (see [21], [45]), and are natural generalizations of the usual max norm. Most of our estimates hold for all positive vectors w of the form $w = A^{-1}e$, where e is any positive vector, i.e., for any positive vector w such that Aw is positive. In particular this would hold for $w = A^{-1}e$ and $e = (1, \ldots, 1)^T$, i.e., with w being the row sums of A^{-1}. Recall that for A a nonsingular M-matrix, $A^{-1} \geq O$, and that since A^{-1} is nonsingular, no row of it can be a zero row. This guarantees that $w = A^{-1}e > 0$. The same logic is used to conclude that $M^{-1}e > 0$ for any monotone matrix M, and this is also used in our proofs.

In this paper we will use several comparison theorems. The first relates the weighted max norms of the iteration matrices and can be found in [20], [37].

Theorem 2.5. *Let $A^{-1} \geq O$, and let $A = M - N = \bar{M} - \bar{N}$ be two weak regular splittings of A with*

$$M^{-1} \geq \bar{M}^{-1}. \tag{16}$$

Let $w > 0$ be such that $w = A^{-1}e$ for some $e > 0$. Then,

$$||M^{-1}N||_w \leq ||\bar{M}^{-1}\bar{N}||_w < 1. \tag{17}$$

If the inequality (16) *is strict, then the first inequality in* (17) *is also strict.*

We point out that with the same hypotheses of Theorem 2.5, an inequality of the form (17) does not necessarily hold for the spectral radii; see a counterexample in [15]. The following three lemmas are helpful in our comparisons of spectral radii. The first one is well known, and can be found, e.g., in [31].

Lemma 2.6. *Assume that $T \in \mathbb{R}^{n\times n}$ is nonnegative and that for some $\alpha \geq 0$ and for some nonzero vector $x \geq 0$, we have $Tx \geq \alpha x$. Then $\rho(T) \geq \alpha$. The inequality is strict if $Tx > \alpha x$.*

Lemma 2.7. *Assume that $A = M - N = \bar{M} - \bar{N}$ are two splittings of A, that $\bar{M}^{-1}\bar{N} \geq O$, and that $M^{-1}N$ has an eigenvector $x \geq 0$ with eigenvalue $\rho(M^{-1}N)$ such that $Ax \geq 0$. If (16) holds, then $\rho(M^{-1}N) \leq \rho(\bar{M}^{-1}\bar{N})$.*

Proof. We have $O \leq \bar{M}^{-1}\bar{N} = I - \bar{M}^{-1}A$. Therefore

$$(I - \bar{M}^{-1}A)x \geq (I - M^{-1}A)x = \rho(M^{-1}N)x \ ,$$

and the assertion follows from Lemma 2.6. □

Lemma 2.8. *Let A be monotone. Let $A = M - N$ be a splitting such that $M^{-1} \geq O$ and $NM^{-1} \geq O$ (sometimes called weak nonnegative of the second type). Then $\rho(M^{-1}N) < 1$ and there exists a non-zero vector $x \geq 0$ such that $M^{-1}Nx = \rho(M^{-1}N)x$ and $Ax \geq 0$.*

Proof. Note first that $A^T = M^T - N^T$ is a weak regular splitting of A^T with $A^{-T} \geq O$, and thus convergent [2], [49]. Therefore, $\rho((M^T)^{-1}N^T) = \rho((NM^{-1})^T) = \rho(NM^{-1}) < 1$. Moreover, since the spectrum of NM^{-1} is equal to that of $M^{-1}N$ we know that $\rho(NM^{-1}) = \rho(M^{-1}N)$ is an eigenvalue of $M^{-1}N$. Let $x \neq 0$ be an eigenvector of $M^{-1}N$ which is scaled in such a way that not all its components are negative. We now first prove the lemma by assuming not only $NM^{-1} \geq O$ but $NM^{-1} > O$. Then we have (denoting $\rho = \rho(M^{-1}N)$) that $M^{-1}Nx = \rho x$ and therefore

$$Nx = (NM^{-1})(Mx) = \rho Mx \ . \tag{18}$$

By the Perron-Frobenius theorem (see, e.g., [2], [49]), the positive matrix NM^{-1} has a positive eigenvector y belonging to the eigenvalue ρ and, up to scaling, this eigenvector is unique. Since $Mx \neq 0$ we therefore have $Mx = \alpha y$ for some $\alpha \neq 0$, and thus $x = \alpha M^{-1}y$. But $M^{-1}y > 0$, and since not all components of x are negative we see that $\alpha > 0$ and therefore $x > 0$ as well as $Mx > 0$. From (18) we have that

$$Ax = Mx - Nx = (1 - \rho)Mx$$

and since $\rho < 1$ this proves $Ax > 0$.

To complete the proof, assume now that $NM^{-1} \geq O$. Let $E \in \mathbb{R}^{n\times n}$ be the matrix with all entries 1 and take $\gamma > 0$ sufficiently small such that the series $\sum_{\nu=0}^{\infty}(\gamma EM^{-1})^{\nu}$ converges. In this case we have, since $EM^{-1} > O$,

$$O < \sum_{\nu=0}^{\infty}(\gamma EM^{-1})^{\nu} = (I - \gamma EM^{-1})^{-1}$$

as well as

$$O < B := (I - \gamma EM^{-1})^{-1}M^{-1} \, .$$

For all positive ε smaller than $\varepsilon_0 = \rho(BMA^{-1})$ we consider the splittings

$$A_\varepsilon = M - (N + \varepsilon BM) \, .$$

Then

$$(N + \varepsilon BM)M^{-1} = NM^{-1} + \varepsilon B > O$$

and

$$A_\varepsilon^{-1} = A^{-1}(I - \varepsilon BMA^{-1})^{-1} = A^{-1}\sum_{\nu=0}^{\infty}(\varepsilon BMA^{-1})^{\nu} \geq O \, .$$

By what we have already shown there exist positive vectors x_ε such that $M^{-1}(N+\varepsilon BM)x_\varepsilon = \rho(M^{-1}(N+\varepsilon BM))x_\varepsilon$ and $A_\varepsilon x_\varepsilon > 0$. We normalize these vectors to have norm 1 and put $\varepsilon_k = \frac{1}{k}\varepsilon_0$. Then the sequence x_{ε_k} admits a convergent subsequence with limit $x \geq 0$, $x \neq 0$. By continuity, this x satisfies $M^{-1}Nx = \rho x$ as well as $Ax \geq 0$. $\square$

The following theorem of Woźnicki [50] is now a direct consequence of Lemmas 2.7 and 2.8.

Theorem 2.9. *Let $A^{-1} \geq O$. Assume that $A = M - N = \bar{M} - \bar{N}$ are two nonnegative splittings with $M^{-1} \geq \bar{M}^{-1}$. Then,*

$$\rho(M^{-1}N) \leq \rho(\bar{M}^{-1}\bar{N}) < 1. \tag{19}$$

The inequality (19) *is strict if $A^{-1} > O$ and $M^{-1} > \bar{M}^{-1}$.*

The following counterpart of Theorem 2.9 in the s.p.d. case is from [35].

Theorem 2.10. *Let $A \succ O$. Assume that $A = M - N = \bar{M} - \bar{N}$ are two (strong P-regular) splittings with $O \preceq N \preceq \bar{N}$. Then,* (19) *holds. The first inequality* (19) *is strict if $O \preceq N \prec \bar{N}$.*

We conclude this section with a new comparison theorem, which is the counterpart to Theorem 2.5 using A-norms, where A is s.p.d. We first prove an intermediate result.

Lemma 2.11. *Let $A \succ O$, and let $A = M - N$ be a splitting of A such that M is symmetric. Then $\rho(M^{-1}N) = ||M^{-1}N||_A$.*

Proof. It follows from the following identities:

$$\begin{aligned}||M^{-1}N||_A = ||I - M^{-1}A||_A \;&=\; ||I - A^{1/2}M^{-1}A^{1/2}||_2\\ = \rho(I - A^{1/2}M^{-1}A^{1/2}) \;&=\; \rho(I - M^{-1}A) \;=\; \rho(M^{-1}N).\end{aligned}$$

□

The following theorem follows now directly from Lemma 2.11 and Theorem 2.10.

Theorem 2.12. *Let $A \succ O$, and let $A = M - N = \bar{M} - \bar{N}$ be two (strong P-regular) splittings of A with*

$$O \preceq N \preceq \bar{N}. \tag{20}$$

Then,

$$||M^{-1}N||_A \leq ||\bar{M}^{-1}\bar{N}||_A < 1. \tag{21}$$

If the second inequality in (20) *is strict, then, the first inequality in* (21) *is also strict.*

The hypothesis (20) cannot be weakened, i.e., we need to assume that the matrices N_1 and N_2 are positive semidefinite matrices. Examples in [35] show that Theorems 2.10 and 2.12 are not true if one only assumes P-regular splittings.

3. Convergence of the one-level method

In this section we prove convergence of the one-level scheme (2) under the assumption that the rows of R_i are rows of the $n \times n$ identity matrix I, i.e., that R_i has the form (4). Recall the definition of the sets S_i in (8). In general, the S_i are not disjoint. When they are, we have the multiplicative Schwarz method without overlap. The following important lemma covers both cases (overlapping and non-overlapping).

Lemma 3.1. *Let A be monotone, and let a collection of p triples (E_i, M_i, N_i) be given such that $O \leq E_i \leq I$, $\sum_{i=1}^{p} E_i \geq I$, and $A = M_i - N_i$ is a weak regular splitting for $1 \leq i \leq p$. Let*

$$T = (I - E_pM_p^{-1}A)(I - E_{p-1}M_{p-1}^{-1}A)\cdots(I - E_1M_1^{-1}A). \tag{22}$$

Then for any vector $w = A^{-1}e > 0$ with $e > 0$, $\rho(T) \leq \|T\|_w < 1$. Furthermore,

$$\|I - E_iM_i^{-1}A\|_w = 1, \quad i = 1, \ldots, p. \tag{23}$$

Proof. In order to show that $\|T\|_w < 1$, where $\|T\|_w$ denotes the maximum weighted norm of T with respect to a certain vector $w > 0$, we show that $T \geq O$ and $Tw < w$.

Clearly $T \geq O$ because for $i = 1, \ldots, p$,

$$\begin{aligned} I - E_i M_i^{-1} A &= I - E_i + E_i(I - M_i^{-1} A) \\ &= I - E_i + E_i M_i^{-1} N_i \geq O \end{aligned} \tag{24}$$

and $M_i^{-1} N_i \geq O$ since the splittings are weak regular.

Next, we show that $Tw < w$ with $w = A^{-1}e$ where $e > 0$. To begin with, note that

$$w_1 := (I - E_1 M_1^{-1} A)w = w - E_1 M_1^{-1} e \geq 0.$$

Hence, $0 \leq w_1 \leq w$, with strict inequality in the components corresponding to S_1. In other words, denoting with $(w_1)_i$ the ith component of w_1, we have

$$(w_1)_i \begin{cases} = w_i & \text{if } i \notin S_1; \\ < w_i & \text{if } i \in S_1. \end{cases}$$

Now let $w_2 := (I - E_2 M_2^{-1} A)w_1$, we claim that $w_2 \leq w$, and that in the components corresponding to S_2, the inequality is strict. Indeed,

$$0 \leq (I - E_2 M_2^{-1} A)w_1 = (I - E_2 M_2^{-1} A)(w_1 - w + w) \leq (I - E_2 M_2^{-1} A)w.$$

Now observe that

$$(w_2)_i \begin{cases} = (w_1)_i \leq w_i & \text{if } i \notin S_2; \\ < w_i & \text{if } i \in S_2, \end{cases}$$

since $i \in S_2$ implies that

$$(w_2)_i = [(I - E_2 M_2^{-1} A)(w_1 - w)]_i + (w - E_2 M_2^{-1} e)_i < w_i.$$

Similarly, one can show that for all $k \leq p - 1$,

$$(w_{k+1})_i \begin{cases} = (w_k)_i & \text{if } i \notin S_{k+1}; \\ < w_i & \text{if } i \in S_{k+1}. \end{cases}$$

Because $\displaystyle\bigcup_{i=1}^{p} S_i = \{1, 2, \ldots, n\}$, we conclude that $Tw < w$. It follows that $\|T\|_w < 1$ and therefore $\rho(T) < 1$. To complete the proof, observe that we have shown that for each $i = 1, \ldots, p$, and each $j = 1, \ldots n$

$$(I - E_i M_i^{-1} A w)_j \leq w_j, \tag{25}$$

and thus $\|I - E_i M_i^{-1} A\|_w \leq 1$. This upper bound is attained since we have shown the inequality in (25) is actually an equality for $j \notin S_i$, cf. (15). $\square$

Remark 3.2. Lemma 3.1 holds for any monotonic norm, i.e., a norm for which $0 \le v \le w$ implies $\|v\| \le \|w\|$. In fact, this is also the case for many other results in this paper. One of the exceptions is Theorem 4.7 in the next section, where the weighted max norm cannot be easily replaced.

Remark 3.3. The collection of triples $\{(E_i, M_i, N_i)\}_{i=1}^p$ can be thought of as a *multiplicative multisplitting* of A, in analogy with the standard (additive) multisplitting of a matrix in the sense of O'Leary and White [38]; see also [6] and the extensive bibliography therein, and [36] for further extensions.

Remark 3.4. In the special case where $E_i = I$ for all $i = 1, 2, \dots, p$, we obtain an extension to the case of p splittings of Theorem 3.2 in [1]; see also the remarks at the end of Sect. 3 in [1].

Lemma 3.1, together with the characterization (13) and Lemma 2.1, is the fundamental tool for proving the convergence of the multiplicative Schwarz method for nonsingular M-matrices.

Theorem 3.5. *Let A be a nonsingular M-matrix. Then the multiplicative Schwarz iteration* (2) *converges to the solution of $Ax = b$ for any choice of the initial guess x^0. In fact, for any $w = A^{-1}e > 0$ with $e > 0$, we have $\rho(T) \le \|T\|_w < 1$. Furthermore, there exists a unique splitting $A = B - C$ such that $T = B^{-1}C$, and this splitting is nonnegative.*

Proof. Let E_i be as in (7) and M_i as in (11). Observe that $O \le E_i \le I$, $1 \le i \le p$. The key to the proof is the characterization (13), from which we have

$$I - P_i = I - E_i M_i^{-1} A, \quad 1 \le i \le p. \tag{26}$$

Moreover, by Proposition 2.3, the splittings $A = M_i - N_i$ (with $N_i = M_i - A$) are regular. Hence, by Lemma 3.1, $\rho(T) \le \|T\|_w < 1$ for any $w = A^{-1}e > 0$ with $e > 0$, and the iteration (2) converges for any initial vector x^0. Furthermore, by Lemma 2.1, there exists a unique splitting $A = B - C$ such that $T = B^{-1}C$. To prove that the splitting is nonnegative we begin by showing that $B^{-1} = (I - T)A^{-1}$ is nonnegative or, equivalently, that $B^{-1}z \ge 0$ for all $z \ge 0$. Letting $v = A^{-1}z \ge 0$, all we need to show is that $(I - T)v \ge 0$, or $Tv \le v$. This is proved in the same way as Lemma 3.1. Hence, the unique splitting $A = B - C$ is weak regular. To show that it is nonnegative we need to show that $\bar{T} = I - AB^{-1}$ is also nonnegative. To see this, note that $\bar{T} = (I - \bar{P}_p)(I - \bar{P}_{p-1}) \cdots (I - \bar{P}_1)$, where $\bar{P}_i = AE_i M_i^{-1} = AR_i^T A_i^{-1} R_i$, in view of the representation (13). To complete the proof we show that each factor $I - \bar{P}_i$ is nonnegative. In fact,

$$I - \bar{P}_i^T = I - R_i^T A_i^{-T} R_i A^T = I - E_i M_i^{-T} A^T \ge O, \tag{27}$$

just as in (24). $\square$

Remark 3.6. In the analysis of multiplicative Schwarz for nonsymmetric problems using analytical tools, convergence is only obtained assuming the addition of a (multiplicative) coarse grid correction, and furthermore that the coarse grid be fine enough; see, e.g., [8], [44, Sect. 5.4]. As can be observed, in the M-matrix case, our Theorem 3.5 (as well as Theorem 4.5 with inexact solves) provides convergence without a coarse grid correction. In Sect. 7 we show convergence of the multiplicative Schwarz method with a coarse grid correction (both additive and multiplicative) without any restriction on how fine it is.

We turn now to the counterpart to the convergence Theorem 3.5 in the s.p.d. case. To that end, we first prove the following lemma.

Lemma 3.7. *Let A be a symmetric positive definite matrix. Let $x,\ y \in \mathbb{R}^n$, such that*

$$y = (I - E_i M_i^{-1} A)x, \tag{28}$$

where E_i is defined in (7) *and M_i in* (14)*. Then the following identity holds:*

$$||y||_A^2 - ||x||_A^2 = -(y-x)^T E_i A E_i (y-x) \leq 0. \tag{29}$$

Furthermore,

$$\|I - E_i M_i^{-1} A\|_A = 1, \quad i = 1, \ldots, p. \tag{30}$$

Proof. Consider $x = \pi_i^T (x_1^T, x_2^T)^T$ and $y = \pi_i^T (y_1^T, y_2^T)^T$, with $x_1,\ y_1 \in \mathbb{R}^{n_i}$. Further, from (7) and (4) we have that

$$E_i = \pi_i^T \begin{bmatrix} I_i & O \\ O & O \end{bmatrix} \pi_i. \tag{31}$$

Consider now (28), whence we immediately have that

$$y_2 = x_2, \tag{32}$$

and using (14) and (5), we also get

$$A_i y_1 = -A_{12} x_2, \tag{33}$$

where here we use the notation $A_{12} = K_i$, and similarly $A_{21} = L_i = A_{12}^T$. Using these identities we write

$$\begin{aligned}
y^T A y - x^T A x &= (y_1^T, y_2^T)\pi_i A \pi_i^T (y_1^T, y_2^T)^T - (x_1^T, x_2^T)\pi_i A \pi_i^T (x_1^T, x_2^T)^T \\
&= y_1^T A_i y_1 + y_2^T A_{21} y_1 + y_1^T A_{12} y_2 - x_1^T A_i x_1 - x_2^T A_{21} x_1 - x_1^T A_{12} x_2 \\
&= x_2^T A_{21}(y_1 - x_1) + (y_1^T - x_1^T) A_{12} x_2 + y_1^T A_i y_1 - x_1^T A_i x_1 \\
&= -y_1^T A_i (y_1 - x_1) - (y_1^T - x_1^T) A_i y_1 + y_1^T A_i y_1 - x_1^T A_i x_1 \\
&= -(y_1^T - x_1^T) A_i (y_1 - x_1) = -(y-x)^T E_i A E_i (y-x),
\end{aligned}$$

where the last equality follows from the identity

$$E_i A E_i = \pi_i^T \begin{bmatrix} A_i & O \\ O & O \end{bmatrix} \pi_i.$$

Since A is s.p.d., $E_i A E_i$ is semidefinite, and the right hand side of (29) is nonpositive. This implies that $\|I - E_i M_i^{-1} A\|_A = \|I - G_i A\|_A \leq 1$, with $G_i = R_i^T (R_i A R_i^T)^{-1} R_i$. To see that this upper bound on the norm is attained we write

$$\|(I - G_i A)x\|_A^2 = x^T A x - x^T A G_i A x.$$

Since G_i is semidefinite, let y be such that $y^T G_i y = 0$, e.g., y having zeros in the entries corresponding to the nonzero columns of R_i as in (4). Then, for $x = A^{-1} y$ we have that $\|(I - G_i A)x\|_A^2 = \|x\|_A^2$. $\square$

We note that the result (30) is well known; see, e.g., [24], [3]. Here we have shown a proof in terms of E_i and M_i simply for completeness, and to emphasize the similarity with (23).

Theorem 3.8. *Let A be a symmetric positive definite matrix. Then the multiplicative Schwarz iteration* (2) *converges to the solution of $Ax = b$ for any choice of the initial guess x^0. In fact, we have $\rho(T) \leq \|T\|_A < 1$. Furthermore, there exists a unique splitting $A = B - C$ such that $T = B^{-1}C$, and this splitting is P-regular.*

Proof. As in the proof of Theorem 3.5 we have the relations (26) following as a consequence of the equalities (13). Starting with $x^{(1)} \neq 0$ let $x^{(i+1)} = (I - P_i)x^{(i)}$. Thus $x^{(p+1)} = Tx^{(1)}$. Using (29) repeatedly, and canceling terms, we obtain

$$||Tx^{(1)}||_A^2 - ||x^{(1)}||_A^2 = -\sum_{i=1}^{p} (x^{(i+1)} - x^{(i)})^T E_i A E_i (x^{(i+1)} - x^{(i)}). \tag{34}$$

Since $E_i A E_i$ is positive semidefinite it follows that the right hand side of (34) is nonpositive. However, the right hand side is zero if and only if

$$E_i(x^{(i+1)} - x^{(i)}) = 0 \qquad \text{for all } i,\ i = 1, \ldots, p.$$

The other $n - n_i$ components of $x^{(i+1)} - x^{(i)}$ are also zero using the same argument as in Lemma 3.7 to obtain (32). But this implies $x^{(p+1)} = x^{(i+1)} = x^{(i)} = x^{(1)}, i = 1, \ldots, p$. Thus $x^{(1)}$ must be a common fixed point of $(I - P_i)$ for all $i = 1, \ldots, p$. However, the fixed points of the projections $(I - P_i)$ are just the vectors $z \in \mathbb{R}^n$ with $E_i z = 0$. Since $\sum_{i=1}^{p} E_i \geq I$ there is no such common nonzero fixed point. Hence the right hand side of (34) must be negative, and we obtain $\rho(T) \leq ||T||_A < 1$. Furthermore, by Lemma 2.1, there exists a unique splitting $A = B - C$ such that $T = B^{-1}C$. With Lemma 2.2 we obtain that this induced splitting is P-regular. $\square$

Remark 3.9. In Lemma 3.7 and in Theorem 3.8 it was not required that the matrix (14) define a P-regular splitting. Nevertheless, the product of the operators (26) produces a matrix with an induced splitting which is P-regular. In fact, we have that in the (unsymmetrized) multiplicative Schwarz method, $B^T + C$ is symmetric and positive definite.

The convergence result of Theorem 3.8 is not new; see, e.g., [9], [24, Ch. 11], [40], [44], [51]. Here we have given a different proof, as a counterpart to our new Theorem 3.5.

4. Inexact solves

In this section we study the effect of varying how exactly (or inexactly) the local problems are solved. We begin with some results for algebraic additive Schwarz. The additive Schwarz method for the solution of (1) is of the form (2), where

$$T = T_\theta = I - \theta \sum_{i=1}^{p} R_i^T A_i^{-1} R_i A, \tag{35}$$

where $0 < \theta \leq 1$ is a damping parameter; see [9], [11], [12], [13], [23], [24, Ch. 11], [44]. We emphasize that convergence of the damped additive Schwarz method is only guaranteed for $\theta \leq 1/q$ in the M-matrix case and for $\theta < 1/\tilde{q}$ in the s.p.d. case [20], [24, Ch. 11]. In fact, simple examples show that this method may not be convergent for $\theta = 1$.

Very often in practice, instead of solving the local problems $A_i y_i = z_i$ exactly, such linear systems are approximated by $\tilde{A}_i^{-1} z_i$ where $\tilde{A}_i$ is an approximation of A_i; see, e.g., [5], and the above mentioned references. By replacing A_i with $\tilde{A}_i$ in (35) one obtains the damped additive Schwarz iteration with inexact local solves, and its iteration matrix is then

$$\tilde{T}_\theta = I - \theta \sum_{i=1}^{p} R_i^T \tilde{A}_i^{-1} R_i A. \tag{36}$$

In the M-matrix case we assume, as in [20], that the inexact solves correspond to monotone matrices and satisfy

$$\tilde{A}_i \geq A_i. \tag{37}$$

Notice that this is equivalent to the condition that the splittings

$$A_i = \tilde{A}_i - (\tilde{A}_i - A_i) \quad \text{be regular splittings.} \tag{38}$$

In the s.p.d. case we assume, as is generally done (see, e.g., [24, Ch. 11]), that the inexact solves correspond to s.p.d. matrices and satisfy

$$\tilde{A}_i \succeq A_i. \tag{39}$$

This assumption implies that

$$A_i = \tilde{A}_i - (\tilde{A}_i - A_i) \quad \text{are } P\text{-regular splittings.} \tag{40}$$

Conditions (37) and (39) are easily satisfied. This is the case, e.g., if $\tilde{A}_i$ has a subset of the nonzeros of A_i (including the diagonal). This last case includes many standard splittings such as the diagonal, tridiagonal, or triangular part, as well as block versions of them. The other notable example is incomplete factorizations $\tilde{A}_i = L_i U_i$ where the nonzeros of the factors are in the locations of the nonzeros of A_i, and in particular ILU(0) [33]. In these cases, the inequality (37) holds, or equivalently, we have (weak) regular splittings [33], [48]. For examples of splittings for which the inequality (39) holds see [35]. Another situation worth mentioning where (39) holds is when A_i is semidefinite and the inexact solver is definite. This process is usually called regularization; see, e.g., [14], [30].

In [20] it is shown that the damped additive Schwarz iterations with inexact local solves converge in the M-matrix case under the condition (37) and $\theta \le 1/q$. Furthermore, it is shown that the induced splittings corresponding to (35) and (36) $A = M_\theta - N_\theta = \tilde{M}_\theta - \tilde{N}_\theta$ are weak regular. Here we show, under the same conditions, that the convergence rate is slower than in the exact case, and that the more inexact the local solves are, the slower the convergence. Furthermore, we show that the splittings induced by (35) and (36) are actually nonnegative, which allows us to compare spectral radii.

Theorem 4.1. *Let A be a nonsingular M-matrix. Let $\tilde{A}_i$ and $\bar{A}_i$ be inexact solves of A_i satisfying $\tilde{A}_i \ge \bar{A}_i \ge A_i$. Let the damping factor $\theta \le 1/q$, which implies that the damped additive Schwarz method is convergent. Then, $\|T_\theta\|_w \le \|\bar{T}_\theta\|_w \le \|\tilde{T}_\theta\|_w$, where $w > 0$ is such that $Aw > 0$, and $\bar{T}_\theta$ is obtained by replacing $\tilde{A}_i$ by $\bar{A}_i$ in* (36), *$i = 1, \ldots, p$. Moreover, $\rho(T_\theta) \le \rho(\bar{T}_\theta) \le \rho(\tilde{T}_\theta)$, and the splittings induced by these iteration matrices are nonnegative.*

Proof. Observe that

$$M_\theta^{-1} = \theta \sum_{i=1}^{p} R_i^T A_i^{-1} R_i = \theta \sum_{i=1}^{p} E_i M_i^{-1} \ge O, \tag{41}$$

$$\tilde{M}_\theta^{-1} = \theta \sum_{i=1}^{p} R_i^T \tilde{A}_i^{-1} R_i = \theta \sum_{i=1}^{p} E_i \tilde{M}_i^{-1} \ge O, \tag{42}$$

where

(43) $\tilde{M}_i = \pi_i^T \begin{bmatrix} \tilde{A}_i & O \\ O & D_{\neg i} \end{bmatrix} \pi_i$, and thus $\tilde{M}_i^{-1} = \pi_i^T \begin{bmatrix} \tilde{A}_i^{-1} & O \\ O & D_{\neg i}^{-1} \end{bmatrix} \pi_i$.

Since (37) implies $A_i^{-1} \geq \tilde{A}_i^{-1}$, we have

$$M_i^{-1} \geq \tilde{M}_i^{-1}, \text{ for } i = 1, \ldots, p, \tag{44}$$

and consequently $M_\theta^{-1} \geq \tilde{M}_\theta^{-1}$. It was shown in [20] that the unique splitting $A = M_\theta - N_\theta$ induced by T_θ is weak regular. The same is true of the splitting $A = \tilde{M}_\theta - \tilde{N}_\theta$. It is not difficult to show that these are actually nonnegative splittings. Consider the splitting induced by T_θ. All we need to show is that the matrix

$$\hat{T}_\theta = I - \theta \sum_{i=1}^{p} A R_i^T A_i^{-1} R_i$$

is nonnegative. Taking the transpose of this matrix and reasoning as in the proof of Theorem 3.4 in [20] it follows that $\hat{T}_\theta \geq O$, hence $I - AM_\theta^{-1} = N_\theta M_\theta^{-1} \geq O$ and the induced splitting is nonnegative. Thus, using Theorem 2.5, we have that if $w > 0$ is such that $Aw > 0$, $\|T_\theta\|_w \leq \|\tilde{T}_\theta\|_w$. Also, using Theorem 2.9, we have that $\rho(T_\theta) \leq \rho(\tilde{T}_\theta)$. The other inequalities follow in the same manner. $\square$

When A is s.p.d. and the inexact solves satisfy (39), convergence holds if $\theta < 1/\tilde{q}$, as shown, e.g., in [24, Ch. 11]. Furthermore, the induced splitting defined by $\tilde{M}_\theta$ is P-regular; see [20]. Here we show that under the same hypotheses the convergence using the inexact solves is slower as measured using either the spectral radii or the A-norm (these two quantities being equal in view of Lemma 2.11). Furthermore, the more inexact the local solves are, the slower the convergence.

We will use he following result for s.p.d. matrices which can be found, e.g., in [24].

Lemma 4.2. *Let A be a symmetric positive definite matrix, and $A_i = R_i A R_i^T$, R_i a restriction operator, so that A_i is a principal submatrix of A. Then $R_i^T A_i^{-1} R_i \preceq A^{-1}$.*

This result is used in [24, Lemma 11.2.7 (ii)], and in other references (e.g., [44]) to obtain directly the bound

$$A \preceq pM, \tag{45}$$

and further improve it to

$$A \preceq \tilde{q}M, \tag{46}$$

where M is M_θ for the value $\theta = 1$.

Theorem 4.3. *Let A be a symmetric positive definite matrix. Let $\tilde{A}_i$ and $\bar{A}_i$ be inexact solves of A_i satisfying $\tilde{A}_i \succeq \bar{A}_i \succeq A_i$. Let the damping factor $\theta < 1/\tilde{q}$, which implies that the damped additive Schwarz method is convergent. Then, $\|T_\theta\|_A \leq \|\bar{T}_\theta\|_A \leq \|\tilde{T}_\theta\|_A$, where $\bar{T}_\theta$ is obtained by replacing $\tilde{A}_i$ by $\bar{A}_i$ in* (36)*, $i = 1, \ldots, p$. Moreover, $\rho(T_\theta) \leq \rho(\bar{T}_\theta) \leq \rho(\tilde{T}_\theta)$, and the splittings induced by these iteration matrices are strong P-regular.*

Proof. Consider the matrices (41) and (42) which are symmetric positive definite using M_i as in (14) and

(47) $$\tilde{M}_i = \pi_i^T \begin{bmatrix} \tilde{A}_i & O \\ O & A_{\neg i} \end{bmatrix} \pi_i.$$

Since (39) implies $A_i^{-1} \succeq \tilde{A}_i^{-1}$, we have that $M_\theta^{-1} \succeq \tilde{M}_\theta^{-1} \succ O$. This implies $M_\theta \preceq \tilde{M}_\theta$ and $N_\theta \preceq \tilde{N}_\theta$. The theorem will follow from Theorems 2.10 and 2.12 once we establish $N_\theta \succeq O$, i.e., that the splittings are strong P-regular. To that end we use (46), and since $\theta < 1/\tilde{q}$, we have $N_\theta = M_\theta - A = \frac{1}{\theta} M - A \succeq O$. □

Remark 4.4. For simplicity, in Theorems 4.1 and 4.3, we assumed that the inexact versions use the same damping parameter θ. It is evident from the proofs that if the damping parameter for the inexact version is smaller, say $\tilde{\theta} < \theta$, the same conclusions hold.

We proceed now with similar results for multiplicative Schwarz with inexact solves. In this case, the iteration matrix is

(48) $$\tilde{T} = (I - E_p \tilde{M}_p^{-1} A)(I - E_{p-1} \tilde{M}_{p-1}^{-1} A) \cdots (I - E_1 \tilde{M}_1^{-1} A),$$

cf. (22). We first prove convergence in the M-matrix case, and proceed with comparisons varying the amount of inexactness of the local solves.

Theorem 4.5. *Let A be a nonsingular M-matrix. Then the multiplicative Schwarz iteration with iteration matrix* (48) *and with inexact solves satisfying* (37) *converges to the solution of $Ax = b$ for any choice of the initial guess x^0. In fact, for any $w = A^{-1}e > 0$ with $e > 0$, we have $\rho(\tilde{T}) \leq \|\tilde{T}\|_w < 1$. Furthermore, there exists a unique splitting $A = \tilde{B} - \tilde{C}$ such that $\tilde{T} = \tilde{B}^{-1}\tilde{C}$, and this splitting is nonnegative.*

Proof. The proof proceeds in the same manner as that of Theorem 3.5. All we need to show is that each splitting $A = \tilde{M}_i - \tilde{N}_i$, with $\tilde{M}_i$ as in (43) is regular. Since $\tilde{A}_i$ is monotone, it follows from (43) that $\tilde{M}_i^{-1} \geq O$. Now, $\tilde{N}_i = \tilde{M}_i - A$ and

$$\pi_i \tilde{N}_i \pi_i^T = \begin{bmatrix} \tilde{A}_i - A_i & -K_i \\ -L_i & D_{\neg i} - A_{\neg i} \end{bmatrix},$$

which, in view of (37), (12), and the fact that A is an M-matrix, is nonnegative. □

Remark 4.6. Theorem 4.5 holds with weaker hypotheses, namely, that the splittings $A_i = \tilde{A}_i - (\tilde{A}_i - A_i)$ are weak regular splittings, i.e., that $\tilde{A}_i^{-1}(\tilde{A}_i - A_i) \geq O$, cf. (37). This is the same assumption used in [20], and it implies that the splittings $A = \tilde{M}_i - \tilde{N}_i$ are weak regular.

Theorem 4.7. *Let A be a nonsingular M-matrix. Let $\tilde{A}_i$ and $\bar{A}_i$ be inexact solves of A_i satisfying $\tilde{A}_i \geq \bar{A}_i \geq A_i$. Then, $\rho(T) \leq \rho(\bar{T}) \leq \rho(\tilde{T})$, and for any $w > 0$ such that $Aw > 0$ we have $||T||_w \leq ||\bar{T}||_w \leq ||\tilde{T}||_w < 1$, where $\bar{T}$ is obtained by replacing $\tilde{A}_i$ by $\bar{A}_i$ in* (48)*, $i = 1, \ldots, p$.*

Proof. We start by establishing the inequalities for the spectral radii. We confine ourselves to show $\rho(T) \leq \rho(\bar{T})$; the inequalities for $\rho(\tilde{T})$ are proved in the same way. By Theorem 3.5 both iteration matrices, T and $\bar{T}$, arise from nonnegative splittings of A. Let $x \geq 0$, $x \neq 0$ be an eigenvector of T with eigenvalue $\rho(T)$. We will show that

$$\bar{T}x \geq Tx = \rho(T)x \tag{49}$$

so that by Lemma 2.6 we get the desired result $\rho(\bar{T}) \geq \rho(T)$. Let $x^0 = \bar{x}^0 = x$ and define $x^i := (I - E_i M_i^{-1} A)x^{i-1}$ and $\bar{x}^i := (I - E_i \bar{M}_i^{-1} A)\bar{x}^{i-1}$, $i = 1, \ldots, p$. Thus, $x^p = Tx$ and $\bar{x}^p = \bar{T}x$. To establish (49) we proceed by induction and show that

$$Ax^i \geq 0, \ i = 1, \ldots, p-1, \tag{50}$$

and

$$0 \leq x^i \leq \bar{x}^i, \ i = 1, \ldots, p. \tag{51}$$

We then have (49) since $x^p = Tx$ and $\bar{x}^p = \bar{T}x$; see (22) and (48).

For $i = 0$, (51) holds by assumption, while relation (50) is true by Lemma 2.8 (here it is crucial that the induced splittings are nonnegative).

Assume now that (50) and (51) are both true for some i. To obtain (50) for $i+1$, observe that $Ax^{i+1} = A(I - E_i M_i^{-1} A)x^i = (I - AE_i M_i^{-1})Ax^i$. By (27) we have $I - AE_i M_i^{-1} \geq O$ and $Ax^i \geq 0$ by the induction hypothesis, and thus (50) holds for $i+1$. To prove that (51) holds for $i+1$, we use (44), (50), and the induction hypothesis to obtain

$$x^{i+1} = (I - E_i M_i^{-1} A)x^i \leq (I - E_i \bar{M}_i^{-1} A)x^i \leq (I - E_i \bar{M}_i^{-1} A)\bar{x}^i = \bar{x}^{i+1}.$$

To establish the inequalities for the weighted max norms, one proceeds in precisely the same manner as before (using w instead of x) to show $\bar{T}w \geq Tw$. Since both matrices are nonnegative, we get $\|T\|_w \leq \|\bar{T}\|_w$. □

Remark 4.8. The purpose of using inexact local solves $\tilde{A}_i$ in lieu of A_i is to obtain convergence in less computational time. Theorems 4.1 and 4.5 indicate that, as to be expected, asymptotically the inexact methods have slower convergence rate. Nevertheless, they converge in less computational time if the saving from the inexact local solve is sufficiently large to offset the loss in convergence rate. This is often the case in practice.

We present now the counterpart to the convergence Theorem 4.5 for the s.p.d. case. Consider inexact solves $\tilde{A}_i$ so that (40) holds. Note that we do not require (39) to hold here. First we present a result similar to Lemma 3.7, cf. [34].

Lemma 4.9. *Let A be a symmetric positive definite matrix. Let $x,\ y \in \mathbb{R}^n$ such that $y = (I - E_i\tilde{M}_i^{-1}A)x$ where $\tilde{M}_i$ is defined in* (47) *with $\tilde{A}_i$ satisfying* (40). *Then the following identity holds:*

$$(52)\quad ||y||_A^2 - ||x||_A^2 = -(y-x)^T E_i(\tilde{M}_i^T + \tilde{M}_i - A)E_i(y-x) \le 0.$$

Furthermore, $\|I - E_i\tilde{M}_i^{-1}A\|_A = 1$, $i = 1, \dots, p$.

Proof. The proof proceeds as that of Lemma 3.7. We have that (32) holds, but instead of (33) we have $\tilde{A}_i y_1 = (\tilde{A}_i - A_i)x_1 - A_{12}x_2$. We then obtain

$$\begin{aligned}
y^TAy - x^TAx &= x_2^TA_{21}(y_1 - x_1) + (y_1^T - x_1^T)A_{12}x_2 + y_1^TA_iy_1 - x_1^TA_ix_1 \\
&= (x_1^T(\tilde{A}_i - A_i)^T - y_1^T\tilde{A}_i^T)(y_1 - x_1) + \\
&\quad\ (y_1^T - x_1^T)((\tilde{A}_i - A_i)x_1 - \tilde{A}_iy_1) + y_1^TA_iy_1 - x_1^TA_ix_1 \\
&= (-x_1^TA_i - (y_1^T - x_1^T)\tilde{A}_i^T)(y_1 - x_1) + \\
&\quad\ (y_1^T - x_1^T)(-A_ix_1 - \tilde{A}_i(y_1 - x_1)) + y_1^TA_iy_1 - x_1^TA_ix_1 \\
&= -(y_1^T - x_1^T)(\tilde{A}_i + \tilde{A}_i^T - A_i)(y_1 - x_1) \\
&= -(y-x)^TE_i(\tilde{M}_i^T + \tilde{M}_i - A)E_i(y-x)
\end{aligned}$$

The rest of the proof is almost identical to that of Lemma 3.7. □

The following theorem establishes the convergence of multiplicative Schwarz with inexact solves in the s.p.d. case, and its proof is almost identical to that of Theorem 3.8.

Theorem 4.10. *Let A be a symmetric positive definite matrix . Then the multiplicative Schwarz iteration with iteration matrix* (48) *with $\tilde{M}_i$ defined in* (47) *and with inexact solves satisfying* (40) *converges to the solution of $Ax = b$ for any choice of the initial guess x^0. In fact, we have $\rho(\tilde{T}) \le \|\tilde{T}\|_A < 1$. Furthermore, there exists a unique splitting $A = \tilde{B} - \tilde{C}$ such that $\tilde{T} = \tilde{B}^{-1}\tilde{C}$, and this splitting is P-regular.*

Remark 4.11. Our convergence Theorem 4.10 is quite general since the inexact solves $\tilde{A}_i$ need not be symmetric as is required in the standard treatment of Schwarz methods, e.g., in [44]. We only require that $\tilde{A}_i^T + \tilde{A}_i - A_i \succ O$.

The parallel between the results for M-matrices and those for s.p.d. matrices is not complete. We do not have at the moment a counterpart for Theorem 4.7.

5. Varying the amount of overlap

We study here how varying the amount of overlap between subblocks (subdomains) influences the convergence rate of both additive and multiplicative Schwarz.

Let us consider two sets of subblocks (subdomains) of the matrix A, as defined by the sets (8), such that one has more overlap than the other, i.e., let

$$\hat{S}_i \supseteq S_i, \quad i = 1, \ldots, p, \tag{53}$$

with $\bigcup_{i=1}^{p} \hat{S}_i = \bigcup_{i=1}^{p} S_i = S$. We make the natural assumption that the larger sets do not intersect with other sets from the same group of variables V_k, i.e., that the measure of overlap q does not change. Of course, each set $\hat{S}_i$ defines an $\hat{n}_i \times n$ matrix $\hat{R}_i$, where $\hat{n}_i$ is the cardinality of $\hat{S}_i$, and the corresponding $n \times n$ matrix $\hat{E}_i = \hat{R}_i^T \hat{R}_i$, as in (7). The relation (53) implies that

$$I \geq \hat{E}_i \geq E_i \geq O. \tag{54}$$

Similarly, if $\hat{\pi}_i$ is such that $\hat{R}_i = [I_i|O]\, \hat{\pi}_i$, with I_i the identity in $\mathbb{R}^{\hat{n}_i}$, we denote by $\hat{A}_i$ the corresponding principal submatrix of A, i.e.,

$$\hat{A}_i = \hat{R}_i A \hat{R}_i^T = [I_i|O] \cdot \hat{\pi}_i \cdot A \cdot \hat{\pi}_i^T \cdot [I_i|O]^T,$$

and, as in (11) define

$$\hat{M}_i = \hat{\pi}_i^T \begin{bmatrix} \hat{A}_i & O \\ O & \hat{D}_{\neg i} \end{bmatrix} \hat{\pi}_i, \tag{55}$$

where $\hat{D}_{\neg i} = \operatorname{diag}(\hat{A}_{\neg i}) \geq O$, and $\hat{A}_{\neg i}$ is the $(n - \hat{n}_i) \times (n - \hat{n}_i)$ complementary principal submatrix of A as in (6). As in (13), we have here also the fundamental identity

$$\hat{E}_i \hat{M}_i^{-1} = \hat{R}_i^T \hat{A}_i^{-1} \hat{R}_i, \quad i = 1, \ldots, p.$$

We want to compare $\hat{M}_i$ with M_i, although $\hat{A}_i$ and A_i are of different size. Without loss of generality, we can assume that the permutations π_i and $\hat{\pi}_i$ coincide on the set S_i, and that the indexes in S_i are the first n_i elements in $\hat{S}_i$. In fact, we can assume that $\hat{\pi}_i = \pi_i$. Thus, A_i is a principal submatrix of $\hat{A}_i$, and $\hat{M}_i$ has the same diagonal as M_i. Since both $\hat{A}_i$ and $\hat{M}_i$ are M-matrices, it follows that

$$\hat{M}_i \leq M_i,\ i = 1, \ldots, p. \tag{56}$$

We consider first the case of damped additive Schwarz with iteration matrix (35), and the iteration matrix corresponding to the larger overlap is

$$\hat{T}_\theta = I - \theta \sum_{i=1}^{p} \hat{R}_i^T \hat{A}_i^{-1} \hat{R}_i A. \tag{57}$$

Before we state our first comparison result, let us mention that in general one expects the increase of overlap defined by (53) to be such that the groupings of the sets is maintained, and thus the measure of overlap, q, to be the same. This is is not a constraint; if we have a different measure of overlap, and say, $\hat{q} \neq q$, we only need to change our hypothesis $\theta \leq 1/q$ to $\theta \leq 1/\max\{q, \hat{q}\}$.

Theorem 5.1. *Let A be a nonsingular M-matrix. Consider two sets of subblocks of A defined by* (53)*, and the two corresponding additive Schwarz iterations* (35) *and* (57)*. Let the damping factor $\theta \leq 1/q$, which implies that the additive Schwarz methods are convergent. Then, $\|\hat{T}_\theta\|_w \leq \|T_\theta\|_w$, where $w > 0$ is such that $Aw > 0$. Also, $\rho(\hat{T}_\theta) \leq \rho(T_\theta)$.*

Proof. Because M_i and $\hat{M}_i$ are both M-matrices, it follows from (56) that

$$\hat{M}_i^{-1} \geq M_i^{-1},\ i = 1, \ldots, p, \tag{58}$$

and together with (54) we have $\hat{E}_i \hat{M}_i^{-1} \geq E_i M_i^{-1}$. This implies that

$$\hat{M}_\theta^{-1} = \theta \sum_{i=1}^{p} \hat{R}_i^T \hat{A}_i^{-1} \hat{R}_i = \theta \sum_{i=1}^{p} \hat{E}_i \hat{M}_i^{-1} \geq \theta \sum_{i=1}^{p} E_i M_i^{-1} = M_\theta^{-1} \geq O,$$

where $A = \hat{M}_\theta - \hat{N}_\theta$ is the unique splitting such that $\hat{T}_\theta = \hat{M}_\theta^{-1} \hat{N}_\theta$; see Lemma 2.1. Since the splittings are nonnegative (see Theorem 4.1), the conclusions follow from Theorem 2.5 and Theorem 2.9. □

Theorem 5.2. *Let A be a symmetric positive definite matrix. Consider two sets of subblocks of A defined by* (53)*, and the two corresponding additive Schwarz iterations* (35) *and* (57)*. Let the damping factor $\theta \leq 1/\tilde{q}$, which implies that the additive Schwarz methods are convergent. Then, $\|\hat{T}_\theta\|_A \leq \|T_\theta\|_A$ and $\rho(\hat{T}_\theta) \leq \rho(T_\theta)$.*

Proof. Let $Q_i = E_i M_i^{-1} = R_i^T A_i^{-1} R_i$ and $\hat{Q}_i = \hat{E}_i \hat{M}_i^{-1} = \hat{R}_i^T \hat{A}_i^{-1} \hat{R}_i$. Since A_i is a principal submatrix of $\hat{A}_i$, by Lemma 4.2 we have that $\hat{Q}_i \succeq Q_i$. Therefore,

$$\hat{M}_\theta^{-1} = \theta \sum_{i=1}^{p} \hat{Q}_i \succeq \theta \sum_{i=1}^{p} Q_i = M_\theta^{-1} \succ O.$$

As shown in the proof of Theorem 4.3, these splittings are strong P-regular, and the theorem follows from Theorems 2.10 and 2.12. □

We remark that the proof of Theorem 5.2 does not really use the new representation (13), but it is of the same spirit as the proof of Theorem 5.1.

Theorems 5.1 and 5.2 indicate that the more overlap there is, the faster the convergence of the algebraic additive Schwarz method. As a special case, we have that overlap is better than no overlap. This is consistent with the analysis for grid-based methods; see, e.g., [4], [44]. In a way similar to that described in Remark 4.8, the faster convergence rate brings associated an increased cost of the local solves, since now they have matrices of larger dimension and more nonzeros. In the cited references a small amount of overlap is recommended, and the increase in cost is usually offset by faster convergence.

Remark 5.3. Results similar to Theorem 5.1 were shown for (additive) multisplitting methods in [17] and [28]; see also [22]. In these references, though, the weighting matrices had to be the same for both sets of splittings. Here we are able to prove this more general result since we do not require that $\sum_{i=1}^{p} E_i = \sum_{i=1}^{p} \hat{E}_i = I$, as in the multisplitting setting. Instead all we need is that these sums be invertible.

We consider now the algebraic multiplicative Schwarz iteration with (22) and the corresponding one with the larger overlap, i.e.,

$$\hat{T} = (I - \hat{E}_p \hat{M}_p^{-1} A)(I - \hat{E}_{p-1} \hat{M}_{p-1}^{-1} A) \cdots (I - \hat{E}_1 \hat{M}_1^{-1} A). \tag{59}$$

Convergence follows in the M-matrix case from Theorem 3.5.

Theorem 5.4. *Let A be a nonsingular M-matrix. Consider two sets of subblocks of A defined by* (53)*, and the two corresponding multiplicative Schwarz iterations* (22) *and* (59)*. Then, $\rho(\hat{T}) \le \rho(T)$, and for any vector $w > 0$ such that $Aw > 0$ we have $||\hat{T}||_w \le ||T||_w$.*

Proof. The proof proceeds exactly as in the proof of Theorem 4.7 using (58). □

6. Varying the number of blocks

We address here the following question. If we partition a block into smaller blocks, how is the convergence of the Schwarz method affected? In the M-matrix case, we show that for both additive and multiplicative Schwarz the more subblocks (subdomains) the slower the convergence. In the s.p.d. case, this is shown only for additive Schwarz. In a limiting case, if each block is a single variable, this is slower. This result is consistent with the classical comparison theorem of Varga [49], which for example shows that the point Jacobi (point Gauss-Seidel) method is asymptotically slower than block Jacobi (block Gauss-Seidel). As in the situations described in sections 4 and 5, the slower convergence may be partially compensated by less expensive local solves, since they are of smaller dimension.

Formally, consider each block of variables S_i partitioned into k_i subblocks, i.e., we have

$$S_{i_j} \subset S_i, \;\; j = 1, \ldots, k_i, \tag{60}$$

$\bigcup_{j=1}^{k_i} S_{i_j} = S_i$, and $S_{i_j} \cap S_{i_k} = \emptyset$ if $j \neq k$. Each set S_{i_j} has associated matrices R_{i_j} and $E_{i_j} = R_{i_j}^T R_{i_j}$. Since we have a partition,

$$E_{i_j} \leq E_i, \;\; j = 1, \ldots, k_i, \;\; \text{and} \;\; \sum_{j=1}^{k_i} E_{i_j} = E_i, \; i = 1, \ldots, p. \tag{61}$$

We define the matrices $A_{i_j} = R_{i_j} A R_{i_j}^T$, and M_{i_j} corresponding to the set S_{i_j} in the manner already familiar to the reader (see, e.g., (55)), so that

$$E_{i_j} M_{i_j}^{-1} = R_{i_j}^T A_{i_j}^{-1} R_{i_j}, \; j = 1, \ldots, k_i, \; i = 1, \ldots, p.$$

Given a fixed damping parameter θ, the iteration matrix of the refined partition is then

$$\bar{T}_\theta = I - \theta \sum_{i=1}^{p} \sum_{j=1}^{k_i} E_{i_j} M_{i_j}^{-1} A, \tag{62}$$

cf. (35), and the unique induced splitting $A = \bar{M}_\theta - \bar{N}_\theta$ (which is a weak regular splitting) is given by

$$\bar{M}_\theta^{-1} = \theta \sum_{i=1}^{p} \sum_{j=1}^{k_i} E_{i_j} M_{i_j}^{-1}.$$

We note that due to the inclusion (60), the measure of overlap q cannot increase.

Theorem 6.1. *Let A be a nonsingular M-matrix. Consider two sets of subblocks of A defined by* (8) *and* (60), *respectively, and the two corresponding additive Schwarz iterations* (35) *and* (62). *Let the damping factor $\theta \le 1/q$, which implies that the additive Schwarz methods are convergent. Then, $\|T_\theta\|_w \le \|\bar{T}_\theta\|_w$, where $w > 0$ is such that $Aw > 0$. Furthermore, $\rho(T_\theta) \le \rho(\bar{T}_\theta)$.*

Proof. In the same way that the inclusion (53) implies the inequality (56) and in turn the inequality (58), here the inclusion (60) implies that

$$M_{i_j}^{-1} \le M_i^{-1}, \;\; j = 1, \ldots, k_i, \; i = 1, \ldots, p. \tag{63}$$

Combining (61) with (63) we have that

$$\sum_{j=1}^{k_i} E_{i_j} M_{i_j}^{-1} \le \sum_{j=1}^{k_i} E_{i_j} M_i^{-1} = E_i M_i^{-1}$$

and thus, $\bar{M}_\theta^{-1} \le M_\theta^{-1}$, which implies the result, using Theorems 2.5 and 2.9. □

Theorem 6.2. *Let A be a symmetric positive definite matrix. Consider two sets of subblocks of A defined by* (8) *and* (60), *respectively, and the two corresponding additive Schwarz iterations* (35) *and* (62). *Let $k = \max_i k_i$, and let the damping factors be $\theta \le 1/\tilde{q}$, and $\bar{\theta} = \theta/k \le 1/(k\tilde{q})$. This implies that the additive Schwarz methods are convergent. Then, $\|T_\theta\|_A \le \|\bar{T}_{\bar{\theta}}\|_A$ and $\rho(T_\theta) \le \rho(\bar{T}_{\bar{\theta}})$.*

Proof. As in the proof of Theorem 5.2 we have, using Lemma 4.2, that

$$Q_{i_j} = E_{i_j} M_{i_j}^{-1} \preceq Q_i = E_i M_i^{-1}.$$

Therefore, $\displaystyle\sum_{j=1}^{k_i} Q_{i_j} \preceq k_i Q_i$, and

$$\bar{M}_\theta^{-1} = \theta \sum_{i=1}^{p} \sum_{j=1}^{k_i} Q_{i_j} \preceq k\theta \sum_{i=1}^{p} Q_i = k M_\theta^{-1},$$

which is equivalent to $\bar{M}_{\bar{\theta}}^{-1} = (1/k)\bar{M}_\theta^{-1} \preceq M_\theta^{-1}$. The theorem now follows using Theorems 2.10 and 2.12, and the fact that these are strong P-regular splittings, as shown in the proof of Theorem 4.3. □

We note that the fact that the bound for the damping factor $\bar{\theta}$ is lower than that for θ is consistent with the fact that we increase the number of regions in the same proportion. Nevertheless, the result of Theorem 6.2 holds for the same damping factor θ. This follows from the fact that due to (60), the number of colors does not increase. The proof has to be modified using the same arguments as in [24, Lemma 11.2.14] to improve the bound (45) to obtain (46).

Next, we consider the case of multiplicative Schwarz. Again, we can show that using more subblocks of smaller size results in slower asymptotic convergence rates. The iteration matrix for the multiplicative Schwarz method corresponding to the finer partition (more subblocks) is given by

$$\tilde{T} = \prod_{i=p}^{1} \prod_{j=k_i}^{1} (I - P_{i_j}), \tag{64}$$

where $P_{i_j} = E_{i_j} M_{i_j}^{-1} A = R_{i_j}^T A_{i_j}^{-1} R_{i_j} A$.

Theorem 6.3. *Let A be a nonsingular M-matrix. Consider two sets of subblocks of A defined by* (8) *and* (60)*, respectively, and the two corresponding multiplicative Schwarz iterations* (3) *and* (64)*. Then $\rho(T) \leq \rho(\tilde{T})$, and $||T||_w \leq ||\tilde{T}||_w$ for any vector $w > 0$ for which $Aw > 0$.*

Proof. Since each $P_i = E_i M_i^{-1} A = R_i^T A_i^{-1} R_i$ is a projection we have

$$I - P_i = (I - P_i)^2 = \ldots = (I - P_i)^{k_i}.$$

This allows us to represent T from (3) (or (22)) as a product with the same number of factors $\tilde{k} = \sum_{i=1}^{p} k_i$ as in the representation (64) for $\tilde{T}$, namely

$$T = \prod_{i=p}^{1} (I - P_i)^{k_i}. \tag{65}$$

We pair each of the $\tilde{k}$ factors $I - P_{i_j} = I - E_{i_j} M_{i_j}^{-1} A$ of $\tilde{T}$ in (64) with the corresponding factor $I - P_i = I - E_i M_i^{-1} A$ of T in (65). This pair of factors correspond to the set of indices S_{i_j} and S_i satisfying $S_{i_j} \subseteq S_i$. By (61) and (63) we have that $E_{i_j} M_{i_j}^{-1} \leq E_i M_i^{-1}$. We can therefore proceed in exactly the same manner as in the proofs of Theorems 4.7 and 5.4 to establish the desired results. $\square$

7. Two-level schemes

In this section we assume that all local solves are exact; however, analogous results hold for the case of inexact solves, provided that the conditions spelled out in Sect. 4 are satisfied. Suppose a “coarse grid” correction is added (multiplicatively) to the multiplicative Schwarz iteration (2). This results in a stationary method with an iteration matrix of the form

$$H = (I - G_0 A)T \tag{66}$$

where T is the iteration matrix of the multiplicative Schwarz method and $G_0 = R_0^T(R_0 A R_0^T)^{-1} R_0$. We assume here that R_0 is formed by some rows of the $(n \times n)$ identity matrix I, so that $R_0 A R_0^T$ is a principal submatrix of A. Typically, R_0 is defined in such a way that it has at least one row in common with each of the R_i matrices that define the multiplicative Schwarz iteration, $1 \leq i \leq p$. Thus, the number of rows in R_0 is no less than p, and should be much less than n. In particular, the coarse grid correction proposed in [47] and used, e.g., in [27], is of this form.

As before, associated with this matrix R_0, we define matrices E_0 and M_0 such that $E_0 M_0^{-1} A = G_0 A$, and $O \leq E_0 \leq I$. Note that if A is an M-matrix, $A = M_0 - (M_0 - A)$ is a regular splitting, and if A is s.p.d., $M_0 \succ O$. The (singular) matrix $I - G_0 A$ defines the global coarse solve, which follows the multiplicative Schwarz sweep. We are interested in comparing the convergence rate of the multiplicative Schwarz iteration with and without the coarse grid correction.

Theorem 7.1. *Let A be a nonsingular M-matrix. Let T and H be the iteration matrices defined in* (2) *and* (66)*, respectively. Then $\rho(H) \leq \rho(T)$, and for any vector $w = A^{-1}e > 0$ with $e > 0$, $\|H\|_w \leq \|T\|_w$. Furthermore, the splitting induced by H is nonnegative.*

Proof. It is clear from Theorem 3.5 that adding a coarse grid correction to the multiplicative Schwarz iteration preserves convergence: $\rho(H) < 1$. Hence, there exists a unique splitting $A = F - (F - A)$ such that $H = I - F^{-1}A$ and the splitting is nonnegative by Theorem 3.5. Furthermore,

$$F^{-1} = B^{-1} + G_0(I - AB^{-1}) \geq B^{-1} \geq O, \tag{67}$$

where $A = B - (B - A)$ is the (unique) nonnegative splitting induced by T. By virtue of (67) and Theorem 2.9 we conclude that $\rho(H) \leq \rho(T)$, and using Theorem 2.5, $\|H\|_w \leq \|T\|_w$. □

Theorem 7.2. *Let A be a symmetric positive definite matrix. Let T and H be the iteration matrices defined in* (2) *and* (66)*, respectively. Then $\rho(H) \leq ||H||_A \leq ||T||_A < 1$. Furthermore, the splitting induced by H is P-regular.*

Proof. From Theorem 3.8, we have $\|T\|_A < 1$, and from Lemma 3.7 we have that $||I - G_0 A||_A = 1$. Hence

$$||H||_A = ||(I - G_0 A)T||_A \leq ||I - G_0 A||_A ||T||_A = ||T||_A < 1.$$

The induced splitting is P-regular by Lemma 2.2. □

Hence, a coarse grid correction results in an asymptotic convergence rate which is at least as good as that of the multiplicative Schwarz iteration (2).

Theorems 7.1 and 7.2 refer to the case where the global coarse solve is *multiplicatively* applied to the multiplicative Schwarz iteration. In [20], the case of *additively* corrected additive Schwarz methods was studied. There remain two other situations to be analyzed, the so-called *hybrid* methods. In one case, the multiplicative Schwarz method is additively corrected; this is called the *two-level hybrid I Schwarz method* in [44]. In the other case, the additive Schwarz method is multiplicatively corrected, leading to the *two-level hybrid II Schwarz method*; see [44, pp. 47–48]. We begin with the multiplicative Schwarz method with additive correction. In this method, the iteration matrix is of the form

$$H_\theta = I - \theta(G_0 + B^{-1})A \tag{68}$$

where $\theta > 0$ is a damping parameter, $G_0 = R_0^T(R_0AR_0^T)^{-1}R_0$ and $A = B - C$ is the unique splitting induced by T, the iteration matrix of the multiplicative Schwarz method. If A is an M-matrix this splitting is nonnegative, and if A is s.p.d. this splitting is P-regular; see Theorems 3.5 and 3.8.

Theorem 7.3. *Let A be a nonsingular M-matrix. If $0 < \theta \le 1/2$, the two-level hybrid I Schwarz method, with iteration matrix* (68)*, converges to the solution of $Ax = b$ for any choice of the initial guess x^0. In fact, for any $w = A^{-1}e > 0$ with $e > 0$, we have $\rho(H_\theta) \le \|H_\theta\|_w < 1$. Furthermore, if $BE_0 + M_0$ is invertible, the splitting induced by H_θ is nonnegative.*

Proof. We first show that $H_\theta \ge O$. Indeed, letting $T_0 = I - M_0^{-1}A$, and since $O \le E_0 \le I$, we have

$$H_\theta = \theta(T + E_0T_0) + (1-\theta)I - \theta E_0,$$

a nonnegative matrix for $0 < \theta \le 1/2$. Then we use that $G_0 \ge O$ and that $B^{-1} \ge O$ and the fact that no row of B^{-1} can be all zeros to write for $w = A^{-1}e > 0$ with $e > 0$

$$H_\theta w = w - \theta(G_0e + B^{-1}e) < w,$$

concluding that $\|H_\theta\|_w < 1$.
For the nonnegativity of the splitting, assuming that $BE_0 + M_0$ is invertible, consider the matrix

$$M_\theta = \theta^{-1}M_0(BE_0 + M_0)^{-1}B.$$

This matrix is invertible, and

$$M_\theta^{-1} = \theta B^{-1}(BE_0 + M_0)M_0^{-1} = \theta(E_0M_0^{-1} + B^{-1}) \ge O.$$

Further, $H_\theta = I - M_\theta^{-1}A$; thus, $A = M_\theta - (M_\theta - A)$ is the unique splitting induced by H_θ. To complete the proof all we need to show is that $I - AM_\theta^{-1} \geq O$, for $0 < \theta \leq 1/2$. This follows in a way similar to the nonnegativity of H_θ using the fact that $A = B - C$ is a nonnegative splitting and $A = M_0 - (M_0 - A)$ is a regular splitting. □

The hypothesis that $BE_0 + M_0$ be nonsingular (not needed for convergence) is very mild. To see this, let $A_0 = R_0AR_0^T$, which is a principal submatrix of A and thus a nonsingular M-matrix. Let π_0 be the permutation so that (4) holds for $i = 0$. Then, we have the nonsingular matrix M_0 of the form (11). Since E_0 is of the form (31), BE_0 has n_0 nonzero rows which are rows of the monotone matrix B. Thus, the addition of the term BE_0 only affects the first n_0 rows of M_0 (once permuted), and $BE_0 + M_0$ is likely to continue to be nonsingular.

Theorem 7.4. *Let A be a symmetric positive definite matrix. If $0 < \theta \leq \frac{1}{2}$, the two-level hybrid I Schwarz method, with iteration matrix* (68), *converges to the solution of $Ax = b$ for any choice of the initial guess x^0. In fact, we have $\rho(H_\theta) \leq \|H_\theta\|_A < 1$. Furthermore, the splitting induced by H_θ is P-regular.*

Proof. We write for $\theta > 0$

$$\begin{aligned} ||H_\theta||_A &= ||\theta(I - G_0A) + \theta(I - B^{-1}A) + (1 - 2\theta)I||_A \\ &\leq \theta||I - G_0A||_A + \theta||I - B^{-1}A||_A + |1 - 2\theta| < 1, \end{aligned}$$

where the last inequality follows from Lemma 3.7, Theorem 3.8, and the assumption $\theta \leq 1/2$. With Lemma 2.2 the induced splitting is P-regular. □

Note that, because of the presence of the damping parameter θ, it is not generally possible to compare the asymptotic rate of convergence of the two-level hybrid I Schwarz method with that of the one-level multiplicative Schwarz method. In any case, the following simple example shows that, in general, one cannot expect the convergence rate to improve as a result of the addition of a global coarse solve.

Example 7.5. Let

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}, \quad R_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

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

Then the eigenvalues of H_θ are $1-2\theta$ (with multiplicity 2) and $1-\frac{8}{9}\theta$. This matrix is convergent for $0<\theta<1$, showing that, in general, the restriction on θ in the statement of Theorem 7.3 is a sufficient condition only. The matrix H_θ is nonnegative if and only if $0\le\theta\le\frac{1}{2}$. For $\theta=\frac{1}{2}$, the spectral radius is $\rho(H_{\frac{1}{2}})=\frac{5}{9}$. The minimum of $\rho(H_\theta)$ is attained for $\theta=\frac{9}{13}$, corresponding to $\rho(H_{\frac{9}{13}})=\frac{5}{13}\approx 0.3846$. The spectral radius of the one-level multiplicative Schwarz iteration matrix is $\rho(T)=\frac{1}{9}\approx 0.1111$. Thus, for this particular example, supplementing the one-level multiplicative Schwarz method with an additive global coarse solve results in a degradation of the asymptotic rate of convergence, for any value of the damping parameter θ.

For completeness, we take a look at the two-level hybrid II Schwarz method, i.e., additive Schwarz with a multiplicative coarse grid correction. The iteration matrix is now

$$H_\theta=(I-G_0A)T_\theta \tag{69}$$

where T_θ is given by (35). Here θ is the damping parameter; when A is a nonsingular M-matrix and $0<\theta\le 1/q$ (where q is the measure of overlap) we have $||T_\theta||_w<1$, with $w=A^{-1}e>0$ and $e>0$, and when A is s.p.d. and $\theta<1/\tilde{q}$ (where $\tilde{q}$ is the number of colors) we have $||T_\theta||_A<1$; see [24] and [20].

Theorem 7.6. *Let A be a nonsingular M-matrix. If $0<\theta\le 1/q$, the two-level hybrid II Schwarz method, with iteration matrix* (69)*, converges to the solution of $Ax=b$ for any choice of the initial guess x^0. Furthermore, $\rho(H_\theta)\le\rho(T_\theta)<1$, for any vector $w=A^{-1}e>0$ with $e>0$, $||H_\theta||_w\le ||T_\theta||_w<1$, and the splitting induced by H_θ is nonnegative.*

Proof. Letting $w_1=T_\theta w$, we have $w_1<w$. An argument identical to the one used in the proof of Lemma 3.1 shows that

$$H_\theta w=(I-G_0A)w_1\le w_1<w,$$

hence $||H_\theta||_w<1$ and the two-level hybrid II Schwarz method is convergent, provided that $0<\theta\le 1/q$. We already know that the unique splitting $A=M_\theta-N_\theta$ induced by T_θ is nonnegative. Let now $A=B-C$ be the unique splitting induced by H_θ. This splitting is weak regular, since $H_\theta\ge O$ and

$$B^{-1}=M_\theta^{-1}+G_0(I-AM_\theta^{-1})\ge M_\theta^{-1}\ge O.$$

It follows from Theorem 2.5 that $||H_\theta||_w\le||T_\theta||_w$. The splitting is actually nonnegative. Indeed,

$$I-AB^{-1}=(I-AG_0)(I-AM_\theta^{-1})\ge O.$$

Thus, by virtue of Theorem 2.9, we also have $\rho(H_\theta)\le\rho(T_\theta)$. □

Theorem 7.7. *Let A be a symmetric positive definite matrix. If $0 < \theta < 1/\tilde{q}$, the two-level hybrid II Schwarz method, with iteration matrix* (69), *converges to the solution of $Ax = b$ for any choice of the initial guess x^0. Furthermore, $||H_\theta||_A \leq ||T_\theta||_A < 1$, and the splitting induced by H_θ is P-regular.*

Proof. With Lemma 3.5 we have

$$||H_\theta||_A = ||(I - G_0A)T_\theta||_A \leq ||I - G_0A||_A||T_\theta||_A = ||T_\theta||_A < 1.$$

With Lemma 2.2 the induced splitting is P-regular. □

Remark 7.8. More generally, we could consider two-level methods where the iteration matrix is of the form $(I - G_0A)T$ and T represents one or more steps of a *smoother*. As long as T induces a nonnegative splitting of the nonsingular M-matrix A, or a P-regular splitting if A is s.p.d., one can show that the coarse grid correction, represented by the singular matrix $I - G_0A$, produces an asymptotic convergence rate which is at least as good as that achieved by the smoother alone.

Finally, we consider the case of two multiplicative two-level schemes which use different global coarse solves for the corrections, with one nested inside the other. As in Sect. 5, we only consider the M-matrix case. Let the iteration matrices be given by

$$(I - G_0A)T \quad \text{and} \quad (I - \hat{G}_0A)T,$$

respectively. Here G_0 and $\hat{G}_0$ correspond to subsets S_0 and $\hat{S}_0$ of S, with $S_0 \subset \hat{S}_0$. In other words, $G_0 = R_0^T(R_0AR_0^T)^{-1}R_0$ and $\hat{G}_0 = \hat{R}_0^T(\hat{R}_0A\hat{R}_0^T)^{-1}\hat{R}_0$ and every row of R_0 is also a row of $\hat{R}_0$. Then it is easy to see that the convergence rate, as measured either by the spectral radius or by the weighted maximum norm, is better for the method corresponding to the finer grid. The proof uses exactly the same argument as the one used to prove Theorem 5.4.

Acknowledgements. The authors thank Olof Widlund for informative discussions. Parts of this work were performed during several visits by some of us to Los Alamos National Laboratory, Universität Bielefeld, Universität Wuppertal, and Temple University, and during a meeting at Mathematisches Forschungsinstitut Oberwolfach. The support and hospitality of these institutions are gratefully acknowledged.

References

1. M. Benzi, D.B. Szyld (1997): Existence and uniqueness of splittings for stationary iterative methods with applications to alternating methods. Numerische Mathematik **76**, 309–321

2. A. Berman, R. J. Plemmons (1994): Nonnegative Matrices in the Mathematical Sciences. Academic Press, New York, third edition, 1979. Reprinted by SIAM, Philadelphia
3. P.E. Bjørstad, J. Mandel (1991): On the spectra of sums of orthogonal projections with applications to parallel computing. BIT **31**, 76–88
4. P.E. Bjørstad, O.B. Widlund (1989): To overlap or not to overlap: A note on a domain decomposition method for elliptic problems. SIAM Journal on Scientific and Statistical Computing **10**, 1053–1061
5. J.H. Bramble, J.E. Pasciak, A.T. Vassilev (1998): Analysis of non-overlapping domain decomposition algorithms with inexact solves. Mathematics of Computation **67**, 1–19
6. R. Bru, V. Migallón, J. Penadés, D. B. Szyld (1995): Parallel, synchronous and asynchronous two-stage multisplitting methods. Electronic Transactions on Numerical Analysis **3**, 24–38
7. X.-C. Cai, Y. Saad (1996): Overlapping domain decomposition algorithms for general sparse matrices. Numerical Linear Algebra with Applications **3**, 221–237
8. X.-C. Cai, O.B. Widlund (1993): Multiplicative Schwarz algorithms for some nonsymmetric and indefinite problems. SIAM Journal on Numerical Analysis **30**, 936–952
9. T.F. Chan, T.P. Mathew (1994): Domain decomposition methods. Acta Numerica **4**, 61–143
10. H. Dağ, F.L. Alvarado (1997): Toward improved uses of the conjugate gradient method for power system applications. IEEE Transactions on Power Systems **12**, 1306–1314
11. M. Dryja (1989): An additive Schwarz algorithm for two- and three-dimensional finite element problems. In: T.F. Chan, R. Glowinski, G.A. Meurant, J. Pèriaux, O. B. Widlund (eds) Second International Symposium on Domain Decomposition Methods for Partial Differential Equations, pp. 168–172, Philadelphia, SIAM
12. M. Dryja, O.B. Widlund (1987): An additive variant of the Schwarz alternating method for the case of many subregions. Technical Report 339, also Ultracomputer Note 131, Department of Computer Science, Courant Institute, New York University
13. M. Dryja, O.B. Widlund (1988): Some domain decomposition algorithms for elliptic problems. In: Iterative Methods for Large Linear Systems, pp. 273–291, San Diego, California, 1989. Academic Press. Proceeding of the Conference on Iterative Methods for Large Linear Systems held in Austin, Texas, October 19–21, 1988, to celebrate the sixty-fifth birthday of David M. Young, Jr.
14. M. Dryja, O.B. Widlund (1995) Schwarz methods of Neumann-Neumann type for three-dimensional elliptic finite element problems. Communications in Pure and Applied Mathematics **48**, 121–155
15. L. Elsner (1989): Comparison of weak regular splittings and multisplitting methods. Numerische Mathematik **56**, 283–289
16. D.J. Evans, S. Jianping, K. Li-shan (1988): The convergence factor of the parallel Schwarz overrelaxation method for linear systems. Parallel Computing **6**, 313–324
17. A. Frommer, B. Pohl (1995): A comparison result for multisplittings and waveform relaxation methods. Numerical Linear Algebra with Applications **2**, 335–346
18. A. Frommer, H. Schwandt (1997): A unified representation and theory of algebraic additive Schwarz and multisplitting methods. SIAM Journal on Matrix Analysis and Applications **18**, 893–912
19. A. Frommer, H. Schwandt, D.B. Szyld (1997): Asynchronous weighted additive Schwarz methods. Electronic Transactions on Numerical Analysis **5**, 48–61
20. A. Frommer, D.B. Szyld (1999): Weighted max norms, splittings, and overlapping additive Schwarz iterations. Numerische Mathematik **83**, 259–278
21. A. Frommer, D.B. Szyld (2000): On asynchronous iterations. Journal of Computational and Applied Mathematics **123**, 201–216

22. A. Frommer, D.B. Szyld (2001): An algebraic convergence theory for restricted additive Schwarz methods using weighted max norms. SIAM Journal on Numerical Analysis, to appear
23. M. Griebel, P. Oswald (1995): On the abstract theory of additive and multiplicative Schwarz algorithms. Numerische Mathematik **70**, 163–180
24. W. Hackbusch (1994) Iterative Solution of Large Sparse Systems of Equations. Springer Verlag, Berlin Heidelberg New York
25. M. Holst (1994): Algebraic Schwarz theory. Technical Report CRPC-994-10, Department of Applied Mathematics and CRPC, California Institute of Technology, Pasadena, California
26. A.S. Householder (1975): The Theory of Matrices in Numerical Analysis. Blaisdell, Waltham, Mass., 1964. Reprinted by Dover, New York
27. E.W. Jenkins, R.C. Berger, J.P. Hallberg, S.E. Howington, C.T. Kelley, J.H. Schmidt, A.K. Stagg, M.D. Tocci (2000): A two-level aggregation-based Newton-Krylov-Schwarz method for hydrology. Parallel computational fluid dynamics 1999. D.E. Keyes, A. Ecer, J. Periaux, N. Satofuka (eds.) Amsterdam: North Holland, pp. 257–264
28. M.T. Jones, D.B. Szyld (1996): Two-stage multisplitting methods with overlapping blocks. Numerical Linear Algebra with Applications **3**, 113–124
29. S.-B. Kim, A. Hadjidimos, E.N. Houstis, J.R. Rice (1996): Multi-parameterized Schwarz alternating methods for elliptic boundary value problems. Mathematics and Computers in Simulation **42**, 47–76
30. A. Klawonn, O.B. Widlund (2000): A domain decomposition method with Lagrange multipliers and inexact solvers for linear elasticity. SIAM Journal on Scientific Computing **22**, 1199–1219
31. I. Marek, D.B. Szyld (1990): Comparison theorems for weak splittings of bounded operators. Numerische Mathematik **58**, 387–397
32. T.P. Mathew (1998): Uniform convergence of the Schwarz alternating method for solving singularly perturbed advection-diffusion equations. SIAM Journal on Numerical Analysis **35**, 1663–1683
33. J.A. Meijerink, H. van der Vorst (1977): An iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix. Mathematics of Computation **31**, 148–162
34. R. Nabben (1989): Chaotische parallele Iterationsverfahren für positiv definite Matrizen. Diplomarbeit, Universität Bielefeld
35. R. Nabben (1996): A note on comparison theorems for splittings and multisplittings of Hermitian positive definite matrices. Linear Algebra and its Applications **233**, 67–80
36. R. Nabben, D.B. Szyld: Convergence theory of restricted multiplicative Schwarz methods. In preparation.
37. M. Neumann, R.J. Plemmons (1987): Convergence of parallel multisplitting methods for M-matrices. Linear Algebra and its Applications **88-89**, 559–574
38. D.P. O'Leary, R.E. White (1985): Multi-splittings of matrices and parallel solution of linear systems. SIAM Journal on Algebraic and Discrete Methods **6**, 630–640
39. J.M. Ortega (1972): Numerical Analysis, A Second Course. Academic Press, New York. Reprinted by SIAM, Philadelphia, 1990
40. A. Quarteroni, A. Valli (1999): Domain Decomposition Methods for Partial Differential Equations. Oxford Science Publications, Clarendon Press, Oxford
41. W.C. Rheinboldt, J.S. Vandergraft (1973): A simple approach to the Perron-Frobenius theory for positive operators on general partially-ordered finite-dimensional linear spaces. Mathematics of Computation **27**, 139–145

42. G. Rodrigue (1985): Inner/outer iterative methods and numerical Schwarz algorithms. Parallel Computing **2**, 205–218
43. G. Rodrigue, K. Li-shan, L. Yu-hui (1989): Convergence and comparison analysis of some numerical Schwarz methods. Numerische Mathematik **56**, 123–138
44. B.F. Smith, P.E. Bjørstad, W.D. Gropp (1996): Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations. Cambridge University Press, Cambridge - New York - Melbourne
45. D.B. Szyld (1998): Different models of parallel asynchronous iterations with overlapping blocks. Computational and Applied Mathematics **17**, 101–115
46. W.-P. Tang (1992): Generalized Schwarz splittings. SIAM Journal on Scientific and Statistical Computing, 13:573–595
47. P. Vaněk, J. Mandel, M. Brezina (1996): Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems. Computing **56**, 179–196
48. R.S. Varga (1960): Factorization and normalized iterative methods. In: R.E. Langer (ed.) Boundary Problems in Differential Equations, pp. 121–142, Madison, The University of Wisconsin Press
49. R.S. Varga (1962): Matrix Iterative Analysis. Prentice-Hall, Englewood Cliffs, New Jersey. Second Edition, revised and expanded, Springer, Berlin, Heidelberg, New York, 2000
50. Z.I. Woźnicki (1994): Nonnegative splitting theory. Japan Journal of Industrial and Applied Mathematics **11**, 289–342
51. J. Xu (1992): Iterative methods by space decomposition and subspace correction. SIAM Review **34**, 581–613
52. D.M. Young, Jr. (1971) Iterative Solution of Large Linear Systems. Academic Press, New York
53. L. Yu-hui, G. Rodrigue (1987): Convergence rate of discrete Schwarz alternating algorithms for solving elliptic P.D.E.'s. In: K. Li-shan (ed.) Parallel Algorithms and Domain Decomposition, pp. 249–264, Wuhan University Press, Wuhan, China