ALGEBRAIC THEORY OF MULTIPLICATIVE SCHWARZ METHODS*

MICHELE BENZI†, REINHARD NABBEN‡, AND DANIEL B. SZYLD§

Abstract. 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. Some results on algebraic additive Schwarz are also included.

Key words. Linear systems, M-matrix, symmetric positive definite matrix, multiplicative Schwarz, weighted max norms, splittings, comparison theorems, overlap, coarse grid correction.

AMS subject classifications. 65F10, 65F35, 65M55.

1. Introduction.

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

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

by multiplicative Schwarz methods. Our aim is to apply the theory of matrix splittings to study the convergence of this class of methods, 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 sections 5 and 6 where 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., [8], [34], [38], 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, [9] 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 [6].

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 [44] and Griebel and Oswald [21], where A-norm bounds are obtained for the symmetric positive definite case. In other cases, e.g., in [28], the maximum principle is used to show convergence.

In this paper we concentrate on the case of algebraic multiplicative Schwarz, although we include some 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 [19] where algebraic additive Schwarz was considered, and it complements the heuristic study [6].

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 remarks 3.9 and 4.10 below. In the nonsymmetric case, we can prove convergence

*This version dated 10 February 2000.

†Computer Research and Applications (CIC-3), MS B256, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA (benzi@lanl.gov). This work was performed under the auspices of the Department of Energy through grant W-7405-ENG-36 with Los Alamos National Laboratory.

‡Fakultät für Mathematik, Universität Bielefeld, Postfach 10 01 31, 33501 Bielefeld, Germany (nabben@mathematik.uni-bielefeld.de).

§Department of Mathematics, Temple University, Philadelphia, Pennsylvania 19122-6094, USA (szyld@math.temple.edu). This work was supported by the National Science Foundation grant DMS-9973219.

without any condition on the coarse grid correction, and in fact convergence is shown without the need for a coarse grid correction; see remark 3.5.

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, $1 \le i \le p$. 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. 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 multiplicative 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 [6], [15], [17], [18], [36], [37], [40], [45]; see also [26]. 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 $B \ge O$ ($B > O$) if its entries are nonnegative (positive). We say that $B \ge C$ if $B - C \ge O$, 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., $A^{-1} \ge O$. It follows that if A and B are nonsingular M-matrices and $A \ge B$, then $A^{-1} \le B^{-1}$ [2], [42]. By $\rho(B)$ we denote the spectral radius of the matrix B. If $A \ge B \ge O$, then $\rho(A) \ge \rho(B)$; see, e.g., [2], [42].

A matrix B is symmetric positive definite (s.p.d.), denoted $B \succ O$, if it is symmetric and if for all vectors $u \ne 0$, $u^T B u > 0$, and positive semidefinite, denoted $B \succeq O$, if for all vectors $u \ne 0$, $u^T B u \ge 0$. 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 $A \succeq B \succeq O$, then $\rho(A) \ge \rho(B)$; see, e.g., [23].

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], [42], [43]. The splitting is P-regular if $M^T + N$ is positive definite [33]. 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 $C = B - A$.*

The following characterization of P-regular splittings from [19] will be useful in our analysis.

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 call 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

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

Following Hackbush [22, 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.

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

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

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. ◻

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 than the one used in [19], although we obtain the same characterization (13). All results in [19] 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 [19]. 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:

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

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., [24], [35]. Weighted max norms play a fundamental role in the study of asynchronous methods (see [20], [39]), 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. As usual the matrix norm is defined as $\|T\|_w = \sup_{\|x\|_w=1} \|Tx\|_w$. We point out that for $T \geq O$, $Tw < \gamma w$ implies $\|T\|_w < \gamma$ ($\gamma > 0$) [35].

LEMMA 2.5. *Let T, $\tilde{T}$ be nonnegative matrices. Assume that $T \leq \tilde{T}$. Then $||T||_w \leq ||\tilde{T}||_w$ for any vector $w > 0$.*

Proof. The result follows from the observation that

$$\max_i \frac{(Tw)_i}{w_i} \leq \max_i \frac{(\tilde{T}w)_i}{w_i}$$

and the fact that

$$\max_i \frac{(Tw)_i}{w_i} = ||Tw||_w = ||T||_w, \qquad \max_i \frac{(\tilde{T}w)_i}{w_i} = ||\tilde{T}w||_w = ||\tilde{T}||_w. \quad \square$$

As we shall see, in many cases we provide A-norm counterparts to results obtained with weighted max norms. One result for which this is not possible is Lemma 2.5 as the following example shows.

EXAMPLE 2.6. Let

$$A = \begin{bmatrix} a^2 & 0 \\ 0 & b^2 \end{bmatrix}, \quad \tilde{T} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \quad T \;=\; \begin{bmatrix} 1-\varepsilon & -\varepsilon \\ -\varepsilon & 2-\varepsilon \end{bmatrix},$$

with $0 < \varepsilon < 2/3$. It follows that $T \preceq \tilde{T}$. We can compute $||\tilde{T}||_A = ||A^{1/2}\tilde{T}A^{-1/2}||_2 = 2$, and

$$||T||_A = \left\| \begin{bmatrix} 1-\varepsilon & -\varepsilon a/b \\ -\varepsilon b/a & 2-\varepsilon \end{bmatrix} \right\|_2 \geq |\varepsilon a/b|.$$

Since a and b are arbitrary, we can have $\|T\|_A \geq \|\tilde{T}\|_A$.

In this paper we will use several comparison theorems. The first relates the weighted max norms of the iteration matrices and is from [19].

THEOREM 2.7. *Let $A^{-1} \geq O$, and let $A = M_1 - N_1 = M_2 - N_2$ be two weak regular splittings of A with*

$$M_1^{-1} \geq M_2^{-1}. \tag{15}$$

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

$$||M_1^{-1}N_1||_w \leq ||M_2^{-1}N_2||_w. \tag{16}$$

If the inequality (15) *is strict, then, the inequality in* (16) *is also strict.*

The following comparison theorem relates spectral radii of two iteration matrices, and is due to Woźnicki [43].

THEOREM 2.8. *Let $A^{-1} \geq O$. Assume that $A = M_1 - N_1 = M_2 - N_2$ are two nonnegative splittings with $M_1^{-1} \geq M_2^{-1}$. Then,*

$$\rho(M_1^{-1}N_1) \leq \rho(M_2^{-1}N_2) < 1. \tag{17}$$

The inequality (17) *is strict if $A^{-1} > O$ and $M_1^{-1} > M_2^{-1}$.*

We point out that the inequality (17) does not necessarily hold with the weaker hypothesis on the splittings of Theorem 2.7, as shown by Elsner [14].

The following counterpart of Theorem 2.8 in the s.p.d. case is from [31].

THEOREM 2.9. *Let $A \succ O$. Assume that $A = M_1 - N_1 = M_2 - N_2$ are two (strong P-regular) splittings with $O \preceq N_1 \preceq N_2$. Then,* (17) *holds. The first inequality* (17) *is strict if $O \preceq N_1 \prec N_2$.*

We conclude this section with a new comparison theorem, which is the counterpart to Theorem 2.7 using A-norms, where A is s.p.d. We first prove a new result.

LEMMA 2.10. *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 form the following identities:

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

The following theorem follows now directly from Lemma 2.10 and Theorem 2.9.

THEOREM 2.11. *Let $A \succ O$, and let $A = M_1 - N_1 = M_2 - N_2$ be two (strong P-regular) splittings of A with*

$$O \preceq N_1 \preceq N_2. \tag{18}$$

Then,

$$||M_1^{-1}N_1||_A \leq ||M_2^{-1}N_2||_A < 1. \tag{19}$$

If the second inequality in (18) *is strict, then, the first inequality in* (19) *is also strict.*

The hypothesis (18) cannot be weakened, i.e., we need to assume that the matrices N_1 and N_2 are positive semidefinite matrices. Examples in [31] show that Theorems 2.9 and 2.11 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*

$$T = (I - E_pM_p^{-1}A)(I - E_{p-1}M_{p-1}^{-1}A) \cdots (I - E_1M_1^{-1}A). \tag{20}$$

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

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

$$I - E_iM_i^{-1}A = I - E_i + E_i(I - M_i^{-1}A) \geq O, \quad 1 \leq i \leq p \tag{21}$$

and $I - M_i^{-1}A \geq O$ 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 $\bigcup_{i=1}^{p} S_i = \{1, 2, \ldots, n\}$, we conclude that $Tw < w$. It follows that $\|T\|_w < 1$ and therefore $\rho(T) < 1$. □

REMARK 3.2. 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 [32]; see also [5] and the extensive bibliography therein.

REMARK 3.3. In the special case where $E_i = I$ for all $i = 1, 2, \ldots, p$, we obtain an extension to the case of p splittings of Theorem 3.2 in [1]; see also the remarks at the end of section 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 3.4. *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 $\rho(T) \leq \|T\|_w < 1$. There exists a unique splitting $A = B - C$ such that $T = B^{-1}C$, and this splitting is nonnegative.*

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

$$I - P_i = I - E_i M_i^{-1} A, \quad 1 \leq i \leq p. \tag{22}$$

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, $\rho(T) \leq \|T\|_w < 1$ 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 $B^{-1}z \geq 0$ for all $z \geq 0$. Letting $v = A^{-1}z \geq 0$, all we need to show is that $(I - T)v \geq 0$, or $Tv \leq v$. 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 $\bar{P}_i = AR_i^T A_i^{-1} R_i$. To complete the proof we show that each factor $I - \bar{P}_i$ is nonnegative. In fact, $I - \bar{P}_i^T = I - R_i^T A_i^{-T} R_i A^T = I - E_i M_i^{-T} A^T \geq O$, just as in (21). □

REMARK 3.5. 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., [7], [38, section 5.4]. As can be observed, in the M-matrix case, our Theorem 3.4 (as well as Theorem 4.4 with inexact solves) provides convergence without a coarse grid correction. In section 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 3.4 in the s.p.d. case. To that end, we first prove the following lemma.

LEMMA 3.6. *Let A be a symmetric positive definite matrix. Let $x,\ y \in \mathbb{R}^n$, such that*

$$y = (I - E_i M_i^{-1} A)x, \tag{23}$$

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

$$||y||_A^2 - ||x||_A^2 = -(y-x)^T E_i A E_i (y-x) \leq 0. \tag{24}$$

Furthermore, $\|I - E_i M_i^{-1} A\|_A = 1$.

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

$$E_i = \pi_i^T \begin{bmatrix} I_i & O \\ O & O \end{bmatrix} \pi_i. \tag{25}$$

Consider now (23), whence we immediately have that

$$y_2 = x_2, \tag{26}$$

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

$$A_i y_1 = -A_{12} x_2, \tag{27}$$

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

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

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

THEOREM 3.7. *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*

$\rho(T) \leq \|T\|_A < 1$. *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 3.4 we have the relations (22) 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 (24) repeatedly, and cancelling terms, we obtain

$$||Tx^{(1)}||_A^2 - ||x^{(1)}||_A^2 = -\sum_{i=1}^{p}(x^{(i+1)} - x^{(i)})^T E_i A E_i (x^{(i+1)} - x^{(i)}). \tag{28}$$

Since E_iAE_i is positive semidefinite it follows that the right hand side of (28) 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 3.6 to obtain (26). 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 $E_iz = 0$. Since $\displaystyle\sum_{i=1}^{p} E_i \geq I$ there is no such common nonzero fixed point. Hence the right hand side of (28) 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 3.8. In Lemma 3.6 and in Theorem 3.7 it was not required that the matrix (14) define a P-regular splitting. Nevertheless, the product of the operators (22) 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.

Remark 3.9. In the usual theory of convergence of multiplicative Schwarz methods for symmetric positive definite matrices, assumptions are made on the decomposition of S into subsets and on the matrices; see, e.g., (4.2) and (4.3) in [44], or similar conditions in [8], [22, Ch. 11], [34], [38]. These assumptions include certain bounds, governed by parameters. It is desirable that these parameters be independent of the mesh. The conditions of theorems 3.4 and 3.7 are very general and there is no dependency on any parameter.

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

$$T = T_\theta = I - \theta\sum_{i=1}^{p} R_i^T A_i^{-1} R_i A, \tag{29}$$

where $0 < \theta \leq 1$ is a damping parameter; see [8], [10], [11], [12], [21], [22, Ch. 11], [38]. 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 [19], [22, 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 $A_iy_i = z_i$ 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., [4], and the above mentioned references. By replacing A_i with $\tilde{A}_i$ in (29) one obtains the damped additive Schwarz iterations with inexact local solves, and its iteration matrix is then

$$\tilde{T}_\theta = I - \theta\sum_{i=1}^{p} R_i^T \tilde{A}_i^{-1} R_i A. \tag{30}$$

In the M-matrix case we assume, as in [19], that the inexact solves correspond to monotone matrices and satisfy

$$\tilde{A}_i \geq A_i. \tag{31}$$

Notice that this is equivalent to the condition that the splittings

$$A_i = \tilde{A}_i - (\tilde{A}_i - A_i) \quad \text{be regular splittings.} \tag{32}$$

In the s.p.d. case we assume, as is generally done (see, e.g., [22, Ch. 11]), that the inexact solves correspond to s.p.d. matrices and satisfy

$$\tilde{A}_i \succeq A_i. \tag{33}$$

This assumption implies that

$$A_i = \tilde{A}_i - (\tilde{A}_i - A_i) \quad \text{are } P\text{-regular splittings.} \tag{34}$$

Conditions (31) and (33) 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) [29]. In these cases, the inequality (31) holds, or equivalently, we have (weak) regular splittings [29], [41]. For examples of splittings for which the inequality (33) holds see [31]. Another situation worth mentioning where (33) holds is when A_i is semidefinite and the inexact solver is definite. This process is usually called regularization; see, e.g., [13], [27].

In [19] it is shown that the damped additive Schwarz iterations with inexact local solves converge in the M-matrix case under the condition (31) and $\theta \le 1/q$. Furthermore, it is shown that the induced splittings corresponding to (29) and (30) $A = M_\theta - N_\theta = \tilde{M}_\theta - \tilde{N}_\theta$ 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 (29) and (30) 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* (30)*, $i = 1, \ldots, p$. Moreover, $\rho(T_\theta) \le \rho(\bar{T}_\theta) \le \rho(\tilde{T}_\theta)$. The splittings induced by these iteration matrices are nonnegative.*

Proof. Observe that

$$M_\theta^{-1} = \theta \sum_{i=1}^{p} R_i^T A_i^{-1} R_i = \theta \sum_{i=1}^{p} E_i M_i^{-1} \ge O, \tag{35}$$

$$\tilde{M}_\theta^{-1} = \theta \sum_{i=1}^{p} R_i^T \tilde{A}_i^{-1} R_i = \theta \sum_{i=1}^{p} E_i \tilde{M}_i^{-1} \ge O, \tag{36}$$

where

$$\tilde{M}_i = \pi_i^T \begin{bmatrix} \tilde{A}_i & O \\ O & D_{\neg i} \end{bmatrix} \pi_i, \quad \text{and thus} \quad \tilde{M}_i^{-1} = \pi_i^T \begin{bmatrix} \tilde{A}_i^{-1} & O \\ O & D_{\neg i}^{-1} \end{bmatrix} \pi_i. \tag{37}$$

Since (31) implies $A_i^{-1} \ge \tilde{A}_i^{-1}$, we have

$$M_i^{-1} \ge \tilde{M}_i^{-1}, \quad \text{for} \quad i = 1, \ldots, p, \tag{38}$$

and consequently $M_\theta^{-1} \ge \tilde{M}_\theta^{-1}$. It was shown in [19] 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 [19] it follows that $\hat{T}_\theta \geq O$, hence $I - AM_\theta^{-1} \geq O$ and the induced splitting is nonnegative. Thus, using Theorem 2.7, we have that if $w > 0$ is such that $Aw > 0$, $\|T_\theta\|_w \leq \|\tilde{T}_\theta\|_w$. Also, using Theorem 2.8, we have that $\rho(T_\theta) \leq \rho(\tilde{T}_\theta)$. The other inequalities follow in the same manner. □

When A is s.p.d. and the inexact solves satisfy (33), convergence holds if $\theta < 1/q$, as shown, e.g., in [22, Ch. 11]. Furthermore, the induced splitting defined by $\tilde{M}_\theta$ is P-regular; see [19]. 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. Furthermore, the more inexact the local solves are, the slower the convergence.

Theorem 4.2. *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* (30)*, $i = 1, \ldots, p$. Moreover, $\rho(T_\theta) \leq \rho(\bar{T}_\theta) \leq \rho(\tilde{T}_\theta)$. The splittings induced by these iteration matrices are strong P-regular.*

Proof. Consider the matrices (35) and (36) which are symmetric positive definite using M_i as in (14) and

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

Since (33) 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 2.9 and 2.11 once we establish $N_\theta \succeq O$, i.e., that the splittings are strong P-regular. To that end, we use the result from [22] that $A \preceq \tilde{q}M$, where M is M_θ for the value $\theta = 1$. Thus, since $\theta < 1/\tilde{q}$, we have $N_\theta = M_\theta - A = \frac{1}{\theta}M - A \succeq O$. □

Remark 4.3. For simplicity, in theorems 4.1 and 4.2, 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

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

cf. (20). We first prove convergence in the M-matrix case, and proceed with comparisons varying the amount of inexactness of the local solves.

Theorem 4.4. *Let A be a nonsingular M-matrix. Then the multiplicative Schwarz iteration with iteration matrix* (40) *and with inexact solves satisfying* (31) *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 $\rho(\tilde{T}) \leq \|\tilde{T}\|_w < 1$. 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 3.4. All we need to show is that each splitting $A = \tilde{M}_i - \tilde{N}_i$, with $\tilde{M}_i$ as in (37) is regular. Since $\tilde{A}_i$ is monotone, it follows from (37) 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 (31), (12), and the fact that A is an M-matrix, is nonnegative. □

Remark 4.5. Theorem 4.4 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. (31). This is the same assumption used in [19], and it implies that the splittings $A = \tilde{M}_i - \tilde{N}_i$ are weak regular.

Theorem 4.6. *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 $w > 0$, $||T||_w \leq ||\bar{T}||_w \leq ||\tilde{T}||_w < 1$, where $\bar{T}$ is obtained by replacing $\tilde{A}_i$ by $\bar{A}_i$ in* (40)*, $i = 1, \ldots, p$.*

Proof. We show that $\rho(T) \leq \rho(\tilde{T})$; the inequalities for $\rho(\bar{T})$ are proved in the same way. All we need to do is to compare each factor $I - E_iM_i^{-1}A$ of T with the corresponding factor $I - E_i\tilde{M}_i^{-1}A$ of $\tilde{T}$. From

(41) $$I - E_iM_i^{-1}A \leq I - E_i\tilde{M}_i^{-1}A$$

and the fact that each of these factors is nonnegative it follows that $O \leq T \leq \tilde{T}$ (see [23, p. 491]). Therefore $\rho(T) \leq \rho(\tilde{T})$, and by Lemma 2.5, $||T||_w \leq ||\tilde{T}||_w$. To see that (41) holds, note that

$$E_i(M_i^{-1} - \tilde{M}_i^{-1})A = \pi_i^T \begin{bmatrix} I - \tilde{A}_i^{-1}A_i & O \\ O & O \end{bmatrix} \pi_i$$

is a nonnegative matrix since the splittings $A_i = \tilde{A}_i - (\tilde{A}_i - A_i)$ are regular. □

REMARK 4.7. 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 4.4 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 4.4 for the s.p.d. case. Consider inexact solves $\tilde{A}_i$ so that (34) holds. Note that we do not require (33) to hold here. First we present a result similar to Lemma 3.6, cf. [30].

LEMMA 4.8. *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* (39) *with $\tilde{A}_i$ satisfying* (34). *Then the following identity holds:*

$$||y||_A^2 - ||x||_A^2 = -(y-x)^T E_i(\tilde{M}_i^T + \tilde{M}_i - A)E_i(y-x) \leq 0. \tag{42}$$

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

Proof. The proof proceeds as that of Lemma 3.6. We have that (26) holds, but instead of (27) 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 3.6. □

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 3.7.

THEOREM 4.9. *Let A be a symmetric positive definite matrix . Then the multiplicative Schwarz iteration with iteration matrix* (40) *with $\tilde{M}_i$ is defined in* (39) *and with inexact solves satisfying* (34) *converges to the solution of $Ax = b$ for any choice of the initial guess x^0. In fact, we have $\rho(\tilde{T}) \leq \|\tilde{T}\|_A < 1$. 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 4.10. Our convergence theorem 4.9 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 [38]. We only require that $\tilde{A}_i^T + \tilde{A}_i - A_i \succ O$.

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. We restrict our analysis in this section to the M-matrix case.

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

$$\hat{S}_i \supseteq S_i, \quad i = 1, \ldots, n, \tag{43}$$

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 (43) implies that

$$I \geq \hat{E}_i \geq E_i \geq O. \tag{44}$$

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

$$\hat{M}_i = \hat{\pi}_i^T \begin{bmatrix} \hat{A}_i & O \\ O & \hat{D}_{\neg i} \end{bmatrix} \hat{\pi}_i, \tag{45}$$

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

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

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 $\hat{\pi}_i = \pi_i$. The $\hat{n}_i \times \hat{n}_i$ matrix $\hat{A}_i$ is of the form

$$\hat{A}_i = \begin{bmatrix} A_i & \hat{K}_i \\ \hat{L}_i & \hat{B}_i \end{bmatrix},$$

where $\hat{B}_i$ is an $(\hat{n}_i - n_i) \times (\hat{n}_i - n_i)$ principal submatrix of A (and of $A_{\neg i}$), and therefore a nonsingular M-matrix. Here

$$\hat{K}_i \leq O \quad \text{and} \quad \hat{L}_i \leq O \tag{46}$$

are submatrices of K_i and L_i, respectively; see (5). Let $D_{\star i}$ be the diagonal of $\hat{B}_i$ and define the matrix

$$B_i = \begin{bmatrix} A_i & O \\ O & \hat{D}_{\star i} \end{bmatrix},$$

then we have

$$M_i = \hat{\pi}_i^T \begin{bmatrix} B_i & O \\ O & \hat{D}_{\neg i} \end{bmatrix} \hat{\pi}_i. \tag{47}$$

Comparing (45), (47), and using (46) and the fact that $\hat{B}_i$ is an M-matrix, it follows that

$$\hat{M}_i \leq M_i. \tag{48}$$

We consider first the case of damped additive Schwarz with iteration matrix (29), and the iteration matrix corresponding to the larger overlap is

$$\hat{T}_\theta = I - \theta \sum_{i=1}^{p} \hat{R}_i^T \hat{A}_i^{-1} \hat{R}_i A. \tag{49}$$

Before we state our first comparison result, let us mention that in general one expects the increase of overlap defined by (43) 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* (43)*, and the two corresponding additive Schwarz iterations* (29) *and* (49)*. 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 (48) that

$$\hat{M}_i^{-1} \geq M_i^{-1}, \tag{50}$$

and together with (44) 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 2.7 and Theorem 2.8. ◻

This theorem indicates 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., [3], [38]. In a way similar to that described in Remark 4.7, 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 5.2. Results similar to Theorem 5.1 were shown for (additive) multisplitting methods in [16] and [25]. 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 (20) and the corresponding one with the larger overlap, i.e.,

$$\hat{T} = (I - \hat{E}_p \hat{M}_p^{-1} A)(I - \hat{E}_{p-1} \hat{M}_{p-1}^{-1} A) \cdots (I - \hat{E}_1 \hat{M}_1^{-1} A). \tag{51}$$

Convergence follows in the M-matrix case from Theorem 3.4.

THEOREM 5.3. *Let A be a nonsingular M-matrix. Consider two sets of subblocks of A defined by* (43)*, and the two corresponding multiplicative Schwarz iterations* (20) *and* (51)*. Then, $\rho(\hat{T}) \leq \rho(T)$, and for any vector $w > 0$, $||\hat{T}||_w \leq ||T||_w$.*

Proof. The proof proceeds as in the proof of Theorem 4.6 by showing that $\hat{T} \leq T$, by comparing each factor of (20) and (51). A direct computation shows that

$$\hat{E}_i(\hat{M}_i^{-1} - M_i^{-1})A = \hat{\pi}_i^T \begin{bmatrix} I - B_i^{-1} \hat{A}_i & O \\ O & O \end{bmatrix} \hat{\pi}_i, \tag{52}$$

a nonnegative matrix. Hence $\hat{E}_i M_i^{-1} A \leq \hat{E}_i \hat{M}_i^{-1} A$ and from (44) we get $E_i M_i^{-1} A \leq \hat{E}_i \hat{M}_i^{-1} A$. Therefore

$$I - \hat{E}_i \hat{M}_i^{-1} A \leq I - E_i M_i^{-1} A$$

and the result follows. ◻

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 a limiting case, if each block is a single variable, this is slower. This result is consistent with the classical comparison theorem of Varga [42], 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 sublocks, i.e., we have

$$S_{i_j} \subset S_i, \quad j = 1, \ldots, k_i, \tag{53}$$

$\bigcup_{j=1}^{k_i} S_{i_j} = S_i$, 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,

$$E_{i_j} \leq E_i, \quad j = 1, \ldots, k_i, \quad \text{and} \quad \sum_{j=1}^{k_i} E_{i_j} = E_i, \; i = 1, \ldots, p. \tag{54}$$

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., (45)), 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

$$\bar{T}_\theta = I - \theta \sum_{i=1}^{p} \sum_{j=1}^{k_i} E_{i_j} M_{i_j}^{-1} A, \tag{55}$$

cf. (29), 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 (53), 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* (53)*, respectively, and the two corresponding additive Schwarz iterations* (29) *and* (55)*. Let the damping factor $\theta \leq 1/q$, which implies that the additive Schwarz methods are convergent. Then, $\|T_\theta\|_w \leq \|\bar{T}_\theta\|_w$, where $w > 0$ is such that $Aw > 0$. Furthermore, $\rho(T_\theta) \leq \rho(\bar{T}_\theta)$.*

Proof. In the same way that the inclusion (43) implies the inequality (48) and in turn the inequality (50), here the inclusion (53) implies that

$$M_{i_j}^{-1} \leq M_i^{-1}, \quad j = 1, \ldots, k_i, \; i = 1, \ldots, p. \tag{56}$$

Combining (54) with (56) we have that

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

and thus, $\bar{M}_\theta^{-1} \leq M_\theta^{-1}$, which implies the result, using theorems 2.7 and 2.8. □

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

$$\tilde{T} = \prod_{i=p}^{1} \prod_{j=k_i}^{1} (I - P_{i_j}), \tag{57}$$

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 6.2. *Let A be a nonsingular M-matrix. Consider two sets of subblocks of A defined by* (8) *and* (53)*, respectively, and the two corresponding multiplicative Schwarz iterations* (3) *and* (57)*. Then $\rho(T) \leq \rho(\tilde{T})$, and $||T||_w \leq ||\tilde{T}||_w$ for any vector $w > 0$.*

Proof. The idea is, again, to compare the iteration matrices factor by factor. Using the fact that each $I - P_i$ is a projection, we can rewrite T (the iteration matrix for p blocks) so that it has exactly the same number of factors as $\tilde{T}$:

$$T = \prod_{i=p}^{1} (I - P_i)^{k_i}. \tag{58}$$

The result now follows from the inequalities $(I - P_i) \leq (I - P_{i_j})$, which can be established with essentially the same argument as the one used in the proof of Theorem 5.3. □

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 section 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

$$H = (I - G_0 A)T \tag{59}$$

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. As before, associated with this matrix R_0, we define matrices E_0 and M_0 are 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* (59)*, 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 3.4 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 3.4. Furthermore,

$$F^{-1} = B^{-1} + G_0(I - AB^{-1}) \geq B^{-1} \geq O, \tag{60}$$

where $A = B - (B - A)$ is the (unique) nonnegative splitting induced by T. By virtue of (60) and Theorem 2.8 we conclude that $\rho(H) \leq \rho(T)$, and using Theorem 2.7, $\|H\|_w \leq \|T\|_w$. □

THEOREM 7.2. *Let A be a symmetric positive definite matrix. Let T and H be the iteration matrices defined in* (2) *and* (59)*, respectively. Then $\rho(H) \leq ||H||_A \leq ||T||_A < 1$. Furthermore, the splitting induced by H is P-regular.*

Proof. From Theorem 3.7, we have $\|T\|_A < 1$, and from Lemma 3.6 we have that $||I - G_0 A||_A = 1$. Hence

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

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 [19], 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 [38]. In the other case, the additive Schwarz method is multiplicatively corrected, leading to the *two-level hybrid II Schwarz method*; see [38, pp. 47–48]. We begin with the multiplicative Schwarz method with additive correction. In this method, the iteration matrix is of the form

$$H_\theta = I - \theta(G_0 + B^{-1})A \tag{61}$$

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 3.4 and 3.7.

THEOREM 7.3. *Let A be a nonsingular M-matrix. If $0 < \theta \leq 1/2$, the two-level hybrid I Schwarz method, with iteration matrix* (61)*, converges to the solution of $Ax = b$ for any choice of the initial guess x^0. In fact, for any $w = Ae > 0$ with $e > 0$, we have $\rho(H_\theta) \leq \|H_\theta\|_w < 1$. Furthermore, if $BE_0 + M_0$ is invertible, the splitting induced by H_θ is nonnegative.*

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

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

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

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

This matrix is invertible, and

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

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 (25), 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* (61)*, 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 3.6, Theorem 3.7, 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 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

$$H_\theta = (I - G_0 A) T_\theta \tag{62}$$

where T_θ is given by (29). 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 [22] and [19].

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* (62), *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

$$H_\theta w = (I - G_0 A) w_1 \le w_1 < w,$$

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 2.7 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 2.8, 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* (62), *converges to the solution of $Ax = b$ for any choice of the initial guess x^0. Furthermore, $||H_\theta||_A \le ||T_\theta||_A < 1$, and the splitting induced by H_θ is P-regular.*

Proof. With Lemma 3.5 we have

$$||H_\theta||_A = ||(I - G_0 A) T_\theta||_A \le ||I - G_0 A||_A ||T_\theta||_A \le ||T_\theta||_A < 1.$$

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 $(I - G_0 A)T$ 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 $I - G_0 A$, 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 section 5, we only consider the M-matrix case. Let the iteration matrices be given by

$$(I - G_0 A)T \quad \text{and} \quad (I - \hat{G}_0 A)T,$$

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, $G_0 = R_0^T(R_0 A R_0^T)^{-1} R_0$ and $\hat{G}_0 = \hat{R}_0^T(\hat{R}_0 A \hat{R}_0^T)^{-1} \hat{R}_0$ 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 5.3.

Acknowledgements. The authors thank Andreas Frommer for his contribution in the proof of Lemma 3.1. His generosity is warmly appreciated. They also thank Olof Widlund for informative discussions. Parts of this work were performed while the first author was visiting the Department of Mathematics at Temple University, and while the third author was a visitor at Los Alamos National Laboratory. The support and hospitality of both institutions are gratefully acknowledged.

REFERENCES

[1] 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.

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

[3] P. E. Bjørstad and O. B. Widlund. 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, 1989.

[4] J. H. Bramble, J. E. Pasciak, and A. T. Vassilev. Analysis of non-overlapping domain decomposition algorithms with inexact solves. *Mathematics of Computation*, 67:1–19, 1998.

[5] R. Bru, V. Migallón, J. Penadés, and D. B. Szyld. Parallel, synchronous and asynchronous two-stage multisplitting methods. *Electronic Transactions on Numerical Analysis*, 3:24–38, 1995.

[6] X.-C. Cai and Y. Saad. Overlapping domain decomposition algorithms for general sparse matrices. *Numerical Linear Algebra with Applications*, 3:221–237, 1996.

[7] X.-C. Cai and O. B. Widlund. Multiplicative Schwarz algorithms for some nonsymmetric and indefinite problems. *SIAM Journal on Numerical Analysis*, 30:936–952, 1993.

[8] T. F. Chan and T. P. Mathew. Domain decomposition methods. *Acta Numerica*, pages 61–143, 1994.

[9] H. Dağ and F. L. Alvarado. Toward improved uses of the conjugate gradient method for power system applications. *IEEE Transactions on Power Systems*, 12:1306–1314, 1997.

[10] M. Dryja. An additive Schwarz algorithm for two- and three-dimensional finite element problems. In T. F. Chan, R. Glowinski, G. A. Meurant, J. Pèriaux, and O. B. Widlund, editors, *Second International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 168–172, Philadelphia, 1989. SIAM.

[11] M. Dryja and O. B. Widlund. 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, 1987.

[12] M. Dryja and O. B. Widlund. Some domain decomposition algorithms for elliptic problems. In *Iterative Methods for Large Linear Systems*, pages 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.

[13] M. Dryja and O. B. Widlund. Schwarz methods of Neumann-Neumann type for three-dimensional elliptic finite element problems. *Communications in Pure and Applied Mathematics*, 48:121–155, 1995.

[14] L. Elsner. Comparison of weak regular splittings and multisplitting methods. *Numerische Mathematik*, 56:283–289, 1989.

[15] D. J. Evans, S. Jianping, and K. Li-shan. The convergence factor of the parallel Schwarz overrelaxation method for linear systems. *Parallel Computing*, 6:313–324, 1988.

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

[17] A. Frommer and H. Schwandt. A unified representation and theory of algebraic additive Schwarz and multisplitting methods. *SIAM Journal on Matrix Analysis and Applications*, 18:893–912, 1997.

[18] A. Frommer, H. Schwandt, and D. B. Szyld. Asynchronous weighted additive Schwarz methods. *Electronic Transactions on Numerical Analysis*, 5:48–61, 1997.

[19] A. Frommer and D. B. Szyld. Weighted max norms, splittings, and overlapping additive Schwarz iterations. *Numerische Mathematik*, 83:259–278, 1999.

[20] A. Frommer and D. B. Szyld. On asynchronous iterations. *Journal of Computational and Applied Mathematics*, to appear in 2000.

[21] M. Griebel and P. Oswald. On the abstract theory of additive and multiplicative Schwarz algorithms. *Numerische Mathematik*, 70: 163–180, 1995.

[22] W. Hackbusch. *Iterative Solution of Large Sparse Systems of Equations.* Springer Verlag, New York – Berlin – Heidelberg, 1994.

[23] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge - New York - Melbourne, 1985.

[24] A. S. Householder. *The Theory of Matrices in Numerical Analysis.* Blaisdell, Waltham, Mass., 1964. Reprinted by Dover, New York, 1975.

[25] M. T. Jones and D. B. Szyld. Two-stage multisplitting methods with overlapping blocks. *Numerical Linear Algebra with Applications*, 3:113–124, 1996.

[26] S.-B. Kim, A. Hadjidimos, E. N. Houstis, and J. R. Rice. Multi-parameterized Schwarz alternating methods for elliptic boundary value problems. *Mathematics and Computers in Simulation*, 42:47–76, 1996.

[27] A. Klawonn and O. B. Widlund, A domain decomposition method with Lagrange multipliers for linear elasticity. Technical Report 780, Department of Computer Science, Courant Institute, New York University, 1999.

[28] 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.

[29] J. A. Meijerink and H. van der Vorst. An iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix. *Mathematics of Computation*, 31:148–162, 1977.

[30] R. Nabben. Chaotische parallele Iterationsverfahren für positiv definite Matrizen. Diplomarbeit, Universität Bielefeld, 1989.

[31] R. Nabben. A note on comparison theorems for splittings and multisplittings of Hermitian positive definite matrices. *Linear Algebra and its Applications*, 233:67–80, 1996.

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

[33] J. M. Ortega. *Numerical Analysis, A Second Course.* Academic Press, New York, 1972. Reprinted by SIAM, Philadelphia, 1990.

[34] A. Quarteroni and A. Valli. *Domain Decomposition Methods for Partial Differential Equations.* Oxford Science Publications, Clarendon Press, Oxford, 1999.

[35] W. C. Rheinboldt and J. S. Vandergraft. 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, 1973.

[36] G. Rodrigue. Inner/outer iterative methods and numerical Schwarz algorithms. *Parallel Computing*, 2:205–218, 1985.

[37] G. Rodrigue, K. Li-shan, and L. Yu-hui. Convergence and comparison analysis of some numerical Schwarz methods. *Numerische Mathematik*, 56:123–138, 1989.

[38] B. F. Smith, P. E. Bjørstad, and W. D. Gropp. *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations.* Cambridge University Press, Cambridge - New York - Melbourne, 1996.

[39] D. B. Szyld. Different models of parallel asynchronous iterations with overlapping blocks. *Computational and Applied Mathematics*, 17:101–115, 1998.

[40] W.-P. Tang. Generalized Schwarz splittings. *SIAM Journal on Scientific and Statistical Computing*, 13:573–595, 1992.

[41] R. S. Varga. Factorization and normalized iterative methods. In R. E. Langer, editor, *Boundary Problems in Differential Equations*, pages 121–142, Madison, 1960. The University of Wisconsin Press.

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

[43] Z. I. Woźnicki. Nonnegative splitting theory. *Japan Journal of Industrial and Applied Mathematics*, 11:289–342, 1994.

[44] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Review*, 34:581–613, 1992.

[45] L. Yu-hui and G. Rodrigue. Convergence rate of discrete Schwarz alternating algorithms for solving elliptic P.D.E.'s. In K. Li-shan, editor, *Parallel Algorithms and Domain Decomposition*, pages 249–264, Wuhan University Press, Wuhan, China, 1987.