SIAM J. MATRIX ANAL. APPL.
Vol. 29, No. 1, pp. 98–116

SCHWARZ ITERATIONS FOR SYMMETRIC POSITIVE SEMIDEFINITE PROBLEMS*

REINHARD NABBEN† AND DANIEL B. SZYLD‡

Abstract. Convergence properties of additive and multiplicative Schwarz iterations for solving linear systems of equations with a symmetric positive semidefinite matrix are analyzed. The analysis presented applies to matrices whose principal submatrices are nonsingular, i.e., positive definite. These matrices appear in discretizations of some elliptic partial differential equations, e.g., those with Neumann or periodic boundary conditions.

Key words. linear systems, additive Schwarz, multiplicative Schwarz, domain decomposition methods, symmetric positive semidefinite systems, singular matrices, comparison theorems, overlap, coarse grid correction

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

DOI. 10.1137/050644203

1. Introduction.

Domain decomposition methods, including additive and multiplicative Schwarz, are widely used for the numerical solution of partial differential equations; see, e.g., [38], [41], [44]. Advantages of these methods include enhancement of parallelism and a localized treatment. One can find algebraic descriptions of them, e.g., in [14], [20], [47], especially for symmetric positive definite problems.

In this paper, we adopt the algebraic representation of additive and multiplicative Schwarz developed in a series of papers [1], [18], [19], [34], [35], where analysis of convergence and properties for several variants of the methods are provided, both for symmetric positive definite and for nonsingular M-matrices. Recently, convergence properties were studied for singular systems arising in the solution of Markov chains, i.e., singular M-matrices with all principal submatrices being nonsingular [7], [32]. In particular, this theory applies to singular matrices with a one-dimensional nullspace, and to those representing irreducible Markov chains; see, e.g., [42]. We also mention the recent work on multiplicative Schwarz iterations for positive semidefinite operators [26], [28].

In this paper, we extend the theory to the symmetric positive semidefinite case, with particular emphasis on the singular case (the analysis of the symmetric positive definite case is known; see, e.g., [1], [21, Ch. 11], [41], [44]). We study in particular the case when all principal submatrices are nonsingular, i.e., positive definite. This situation arises in practice, e.g., in the discretization of certain elliptic differential equations such as $-\Delta u + u = f$ with Neumann or periodic boundary conditions; see, e.g., [5]. We show that in this case, the additive and multiplicative Schwarz iterations are convergent and we characterize the convergence factor γ for such methods (sections 4 and 5). We use the theory of matrix splittings (see section 3) to obtain these convergence properties. We remark that we do not use splittings to produce new

*Received by the editors November 3, 2005; accepted for publication (in revised form) by M. Benzi June 9, 2006; published electronically December 21, 2006.
http://www.siam.org/journals/simax/29-1/64420.html
†Institute für Mathematik, Technische Universität Berlin, D-10623 Berlin, Germany (nabben@math.tu-berlin.de).
‡Department of Mathematics, Temple University (038-16), 1805 N. Broad St., Philadelphia, PA 19122-6094 (szyld@temple.edu). Supported in part by the U.S. National Science Foundation under grant DMS-0207525, and by the U.S. Department of Energy under grant DE-FG02-05ER25672.

stationary iterative methods. What we do is recast the Schwarz iteration matrices as coming from specific splittings, and we use this setup as an analytical tool to obtain convergence results.

The convergence theory we develop implies that the corresponding preconditioned matrices have zero as an isolated point in the spectrum. The rest of the spectrum is contained in a circle centered at one with radius $\gamma < 1$. When considering additive and multiplicative Schwarz preconditioners for singular systems, one needs to use Krylov subspace methods which are sometimes tailored for this case; see, e.g., [17], [23], [39], and the references given therein.

We believe that our purely algebraic approach is much simpler than that of [26], [28], and in addition, it can be applied to problems which may not have a variational formulation. Of course our approach is only valid for the finite dimensional case. We also consider the case of inexact local solvers (section 6), and the influence of the amount of overlap and the number of blocks in the convergence rate (sections 7 and 8). Finally, we study the convergence of two-level methods, i.e., methods where a coarse grid correction is considered as well (section 9).

2. The algebraic representation and notation. We first briefly describe the additive and multiplicative Schwarz methods and give some auxiliary results. Additional notation and background are also given in the next section.

Let $\mathcal{R}(A)$ be the range of A. Consider the linear system in $\mathbb{R}^n$ of the form

$$Ax = b, \quad b \in \mathcal{R}(A). \tag{2.1}$$

In this paper we consider the case where A is symmetric positive semidefinite, and we denote this by $A \succeq O$. We assume that every principal submatrix of A is nonsingular, i.e., a symmetric positive definite matrix, and if A_i is such a submatrix, we denote this by $A_i \succ O$. This situation occurs, for instance, when the null-space of A, $\mathcal{N}(A)$, is unidimensional and any generator of it has no zero entries; cf. [5].

We consider p subspaces V_i, with $\dim V_i = n_i, i = 1, \ldots, p$, which are spanned by columns of the identity I over $\mathbb{R}^n$ and such that

$$\sum_{i=1}^{n} V_i = \mathbb{R}^n =: V. \tag{2.2}$$

Note that the subspaces V_i may overlap. Between the subspaces V_i and the space V we consider the following mappings:

$$R_i : V \to V_i, \qquad R_i^T : V_i \to V,$$

where $\operatorname{rank}(R_i^T) = n_i$. R_i is called the restriction operator while R_i^T is called the prolongation operator. We also use the matrices

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

where $A_i := R_i A R_i^T$ is a permutation of a principal submatrix of A, which because of our assumption is nonsingular. Note that P_i is a projection.

With these projections the damped additive Schwarz method used as an iterative method to solve (2.1) can be described as

$$\begin{aligned} x^{k+1} &= x^k + \theta \sum_{i=1}^{p} R_i^T A_i^{-1} R_i (b - Ax^k) \\ &= \left(I - \theta \sum_{i=1}^{p} R_i^T A_i^{-1} R_i A \right) x^k + \left(\theta \sum_{i=1}^{p} R_i^T A_i^{-1} R_i \right) b, \end{aligned} \tag{2.3}$$

where $0 < \theta \leq 1$ is a damping parameter; see [8], [11], [12], [13], [20], [21, Ch. 11], [41], [44]. The iteration matrix is then given by

$$T_{AS,\theta} = I - \theta \sum_{i=1}^{p} R_i^T A_i^{-1} R_i A = I - \theta \sum_{i=1}^{p} P_i, \tag{2.4}$$

or, using the notation

$$M_{AS}^{-1} = \sum_{i=1}^{p} R_i^T A_i^{-1} R_i, \tag{2.5}$$

then, the iteration matrix (2.4) can be written as

$$T_{AS,\theta} = I - \theta M_{AS}^{-1} A.$$

Later on, in Theorem 4.2, we show that the matrix on the right-hand side in (2.5) is nonsingular, and therefore it makes sense to denote it as M_{AS}^{-1}. Furthermore, for each $\theta > 0$ one can define a splitting of A for which the iteration matrix is precisely (2.4). One such splitting is $A = \frac{1}{\theta} M_{AS} - (\frac{1}{\theta} M_{AS} - A)$. When A is singular, such splitting however is not unique; see [2].

Very often in practice the additive Schwarz method is used for preconditioning a Krylov subspace method. In the symmetric cases considered here the method of choice is the conjugate gradient method; for a study of this method for singular systems, see [23]. While the matrix A may be singular, the preconditioning matrix M is usually assumed to be symmetric positive definite. The additive Schwarz preconditioner is M_{AS}^{-1} and the preconditioned matrix is then

$$M_{AS}^{-1} A = \sum_{i=1}^{p} P_i = I - T_{AS,1}.$$

The multiplicative Schwarz method can be written as the iteration

$$x^{k+1} = T_{MS} x^k + c, \qquad k = 0, 1, \ldots, \tag{2.6}$$

with the iteration matrix

$$T_{MS} = (I - P_p)(I - P_{p-1}) \cdots (I - P_1) = \prod_{i=p}^{1} (I - P_i), \tag{2.7}$$

and a certain vector c. The corresponding preconditioned matrix in this case is $I - T_{MS}$.

Remark 2.1. Observe that for any vector $y \in \mathcal{N}(A)$, i.e., such that $Ay = 0$, one has $Ty = y$ for both iteration matrices $T = T_{AS,\theta}$ of (2.4), or $T = T_{MS}$ of (2.7). This implies in particular that we need to require in our iterations, such as (2.3), that $x_0 \notin \mathcal{N}(A)$.

We outline our strategy to prove the convergence of the iterations (2.3) and (2.6). We need to show that the powers of the iteration matrices (2.4) and (2.7) converge to a limit; see Definition 3.1 below. One sufficient condition for this to hold is that there is a splitting of A of the form $A = M - N$ with M nonsingular such that $M^{-1}N$ is the iteration matrix, and we show that this splitting is P-regular (see

Definition 3.3 below), which implies convergence; see Theorem 3.2 below. We also use certain comparison theorems to relate the convergence of different versions of these iterations. We present a context for these analytical tools in section 3. In the rest of this section, we repeat the algebraic characterization of the Schwarz methods used, e.g., in [1], which is the basis to produce the above-mentioned splittings.

As already mentioned, we assume that the rows of R_i are rows of the $n \times n$ identity matrix I, e.g., of the form

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

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

$$R_i = [I_i|O]\,\pi_i \tag{2.8}$$

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

$$\pi_i A \pi_i^T = \begin{bmatrix} A_i & K_i \\ K_i^T & A_{\neg i} \end{bmatrix}, \tag{2.9}$$

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

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

with $I_{\neg i}$ the identity on $\mathbb{R}^{n-n_i}$.

For each $i = 1, \ldots, p$, we define

$$E_i := R_i^T R_i \in \mathbb{R}^{n\times n}. \tag{2.10}$$

These diagonal matrices have ones on the diagonal in every row where R_i^T has nonzeros. We further need sets S_i defined by

$$S_i := \{j \in \{1, \ldots, n\} : (E_i)_{j,j} = 1\}.$$

Then

$$\bigcup_{i=1}^{p} S_i = S = \{1, 2, \ldots, n\}; \tag{2.11}$$

i.e., each index 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. In other words, in the case of overlapping subspaces, 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.

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 (2.8) as

$$M_i = \pi_i^T \begin{bmatrix} A_i & O \\ O & D_{\neg i} \end{bmatrix} \pi_i, \tag{2.12}$$

where under our assumptions on $A \succeq O$, we have that $D_{\neg i} = \mathrm{diag}(A_{\neg i}) \succ O$, and thus M_i is invertible.

With the definitions (2.10) and (2.12) we obtain the following equality which we will use throughout the paper:

$$E_i M_i^{-1} A = R_i^T A_i^{-1} R_i A = P_i, \quad i = 1, \ldots, p. \tag{2.13}$$

3. Convergent matrices, splittings, and comparison theorems. In this section we present some more definitions and results which we use in the rest of the paper.

DEFINITION 3.1. *A matrix T is called convergent if* $\lim_{k\to\infty} T^k$ *exists. This is equivalent to the following three conditions:*

(1) $\rho(T) \le 1$.
(2) $\mathrm{rank}(I - T) = \mathrm{rank}(I - T)^2$.
(3) *If* $|\lambda| = 1$ *for an eigenvalue* λ *of* T*, then* $\lambda = 1$.

Condition 2 states that the index of the matrix $I - T$ is one, or in this case that $ind_1 T = 1$ [3]. Several equivalent conditions can be found in [43]. One of them is the following:

$$ind_1 T = 1 \Leftrightarrow \mathcal{R}(I - T) \cap \mathcal{N}(I - T) = \{0\}, \tag{3.1}$$

i.e., that the intersection of the range and the null-space of $I - T$ is trivial.

If $\rho(T) = 1$ for a convergent matrix then the asymptotic rate of convergence is given by

$$\gamma(T) := \max\{|\lambda| \ : \ \lambda \in \sigma(T), |\lambda| < 1\}. \tag{3.2}$$

When A is singular, and we have a nonsingular matrix M, and a convergent matrix T such that $A = M(I-T)$, then $P = \lim_{k\to\infty} T^k$ is a projection onto $\mathcal{N}(A) = \mathcal{N}(I-T)$. In fact $P = I-(I-T)(I-T)^D$, where $(I-T)^D$ denote the Drazin inverse of $(I-T)$. Furthermore, if we let $c = M^{-1}b$, and consider the iteration $x^{k+1} = Tx^k + c$, $x_0 \notin \mathcal{N}(A)$ (cf. (2.3)), then $\lim_{k\to\infty} x^k = (I-T)^D c + (I-P)x^0$; see, e.g., [3, Ch. 7.6].

A useful result in the analysis of convergent iteration matrices is the following, due to Keller [24].

THEOREM 3.2. *Let A be symmetric and let M be nonsingular such that $M + M^T - A$ is positive definite. Then $T = I - M^{-1}A$ is convergent if and only if A is positive semidefinite.*

Note than when M is symmetric this theorem says that if $2M - A \succ O$, then T is convergent if and only if $A \succeq O$.

DEFINITION 3.3. *A splitting $A = M - N$ is called* P-regular *if $M + M^T - A \succ O$* [36]*, and* strong P-regular *if in addition $N \succ O$* [33].

With this definition, Theorem 3.2 indicates that a sufficient condition for convergence of T is that $A = M - N$ is a P-regular splitting of a positive semidefinite matrix. Weaker sufficient conditions, and also necessary conditions not requiring the nonsingularity of M, can be found in the recent paper [27].

The following result is a new sufficient condition for convergence, which we use later in the paper.

LEMMA 3.4. *Let A be symmetric positive semidefinite and let $A = M - N$ with M symmetric positive definite. If*

$$A^{\frac{1}{2}} M^{-1} A^{\frac{1}{2}} \prec 2I,$$

then $T = I - M^{-1}A$ is convergent and $A = M - N$ is a P-regular splitting.

Proof. We have $A^{\frac{1}{2}}M^{-1}A^{\frac{1}{2}} \prec 2I$. Thus

$$\sigma(A^{\frac{1}{2}}M^{-1}A^{\frac{1}{2}}) \subset [0,2).$$

Since

$$\sigma(A^{\frac{1}{2}}M^{-1}A^{\frac{1}{2}}) = \sigma(M^{-1}A) = \sigma(AM^{-1}) = \sigma(AM^{-\frac{1}{2}}M^{-\frac{1}{2}}) = \sigma(M^{-\frac{1}{2}}AM^{-\frac{1}{2}}),$$

we have that

$$2I - M^{-\frac{1}{2}}AM^{-\frac{1}{2}} \succ 0.$$

Hence,

$$M^{\frac{1}{2}}(2I - M^{-\frac{1}{2}}AM^{-\frac{1}{2}})M^{\frac{1}{2}} \succ 0$$

and therefore,

$$2M - A \succ 0;$$

i.e., we have a P-regular splitting. Using Theorem 3.2 we obtain that $T = I - M^{-1}A$ is convergent. ∎

The use of P-regular splittings as sufficient conditions for convergence of classical stationary iterative methods for symmetric matrices mimics the use of regular or weak regular splittings as sufficient conditions for the convergence of classical stationary iterative methods for monotone matrices; see, e.g., the classic books [3], [37], [45]. In this case, the rate of convergence of the iterative method is given by the spectral radius of the iteration matrix. Thus, the rate of convergence of two iterative methods for monotone matrices can be compared by looking at the corresponding spectral radii. Many comparison theorems using different hypothesis on the splittings have appeared in the literature; see, e.g., [9], [10], [16], [29], [33], [45], [46], and other references therein. When the iteration matrices have spectral radius equal to one, as is usually the case for singular linear systems, the convergence rate is given by (3.2). Comparison theorems for these can be found in [30], [31]. Here we present a new comparison theorem, which we use in our context.

We first present the following result due to Weyl; see [22, Theorem 4.3.7]. Let $M \succeq O$, and denote its eigenvalues by $\lambda_1(M) \geq \lambda_2(M), \ldots, \lambda_n(M) \geq 0$.

PROPOSITION 3.5. *Let M_1 and M_2 be two symmetric positive semidefinite matrices. If $M_1 \succeq M_2$ then $\lambda_i(M_1) \geq \lambda_i(M_2)$ for all i.*

Of course, this proposition is valid when M is positive definite as well.

THEOREM 3.6. *Let A be symmetric positive semidefinite. Let M_1 and M_2 be symmetric positive definite and let $N_1 := M_1 - A$ and $N_2 := M_2 - A$. If*

$$M_1^{-1} \succeq M_2^{-1},$$

then

$$\lambda_i(M_1^{-1}N_1) \leq \lambda_i(M_2^{-1}N_2)$$

for all i. If additionally N_1 and N_2 are positive semidefinite, then

$$\gamma(M_1^{-1}N_1) \leq \gamma(M_2^{-1}N_2).$$

Proof. We first note that

$$\sigma(M_k^{-1}A) = \sigma(M_k^{-1}A^{\frac{1}{2}}A^{\frac{1}{2}}) = \sigma(A^{\frac{1}{2}}M_k^{-1}A^{\frac{1}{2}}), \quad k = 1, 2.$$

With Proposition 3.5 we obtain for each i that

$$\lambda_i(M_1^{-1}A) = \lambda_i(A^{\frac{1}{2}}M_1^{-1}A^{\frac{1}{2}}) \geq \lambda_i(A^{\frac{1}{2}}M_2^{-1}A^{\frac{1}{2}}) = \lambda_i(M_2^{-1}A). \tag{3.3}$$

Since $M_k^{-1}N_k = I - M_k^{-1}A$, $k = 1, 2$, (3.3) indicates that for each i,

$$\lambda_i(M_1^{-1}N_1) \leq \lambda_i(M_2^{-1}N_2).$$

If N_1 and N_2 are positive semidefinite then all eigenvalues of $M_1^{-1}N_1$ and $M_2^{-1}N_2$ are nonnegative, and therefore

$$\gamma(M_1^{-1}N_1) \leq \gamma(M_2^{-1}N_2). \qquad \square$$

4. Convergence of additive Schwarz. We begin with an auxiliary result, the proof of which follows by a straightforward calculation.

LEMMA 4.1. *Let A be symmetric positive semidefinite. Then*

$$A^{\frac{1}{2}}R_i^T(R_iAR_i^T)^{-1}R_iA^{\frac{1}{2}}$$

is an orthogonal projection. Thus, $I - A^{\frac{1}{2}}R_i^T(R_iAR_i^T)^{-1}R_iA^{\frac{1}{2}}$ is also an orthogonal projection and as a consequence

$$A^{\frac{1}{2}}R_i^T(R_iAR_i^T)^{-1}R_iA^{\frac{1}{2}} \preceq I, \tag{4.1}$$

and

$$\sigma(A^{\frac{1}{2}}R_i^T(R_iAR_i^T)^{-1}R_iA^{\frac{1}{2}}) = \{0, 1\}.$$

THEOREM 4.2. *Let A be symmetric positive semidefinite such that each principal submatrix is positive definite. Let $b \in \mathcal{R}(A)$ and $x_0 \notin \mathcal{N}(A)$. If $0 < \theta < 2/p$, then the additive Schwarz iteration defined by (2.4) is convergent and the splitting defined by $M = \frac{1}{\theta}M_{AS}$ is P-regular.*

Proof. First, as is done in [21] for the nonsingular case, we prove that the matrix

$$\sum_{i=1}^{p} R_i^T(R_iAR_i^T)^{-1}R_i$$

is nonsingular. To that end, let the vector x be such that

$$\sum_{i=1}^{p} R_i^T(R_iAR_i^T)^{-1}R_ix = 0.$$

Hence

$$x^T\sum_{i=1}^{p} R_i^T(R_iAR_i^T)^{-1}R_ix = 0,$$

and thus

$$\sum_{i=1}^{p}(A_i^{-\frac{1}{2}}R_ix)^TA_i^{-\frac{1}{2}}R_ix = \sum_{i=1}^{p}||A_i^{-\frac{1}{2}}R_ix||_2^2 = 0,$$

which implies $R_i x = 0$ for $i = 1, \ldots, p$. By our assumption (2.2) this implies that $x = 0$.

Using Lemma 4.1 we have that (4.1) holds. Summing up, we have

$$A^{\frac{1}{2}} \left(\sum_{i=1}^{p} R_i^T (R_i A R_i^T)^{-1} R_i \right) A^{\frac{1}{2}} \preceq pI, \tag{4.2}$$

and since $\theta < 2/p$, we have $A^{\frac{1}{2}} \theta M_{AS}^{-1} A^{\frac{1}{2}} \prec 2I$. We can now use Lemma 3.4, and this completes the proof. ◻

As is done in [21, Ch. 11.2.4] in the symmetric positive definite case, a careful look at the sum in (4.2) indicates that we can replace the number of subdomains p with the number of colors q of the graph of A. Thus $A^{\frac{1}{2}} M_{AS}^{-1} A^{\frac{1}{2}} \prec qI$, and if $\theta < 2/q$, we have convergence.

Remark 4.3. If we further restrict the value of the damping parameter to $\theta < 1/p$ (or $\theta < 1/q$), we have that the splitting defined by $\frac{1}{\theta} M_{AS}$ is strong P-regular. This follows since in this case $A^{\frac{1}{2}} \theta M_{AS}^{-1} A^{\frac{1}{2}} \prec I$, which implies $\frac{1}{\theta} M_{AS} \succ A$.

We note that the result in Theorem 4.2 applies in particular to the symmetric positive definite case. Thus, in our formulation we have doubled the interval of admissible damping factors for convergence of the damped additive Schwarz method, since the usual restriction is that $\theta < 1/q$; see [18], [21, Ch. 11.2.4]. We mention also that simple examples show that this method may not be convergent for $\theta = 1$.

From Theorem 4.2 it follows that the only eigenvalue of T in the unit circle is $\lambda = 1$, and since we showed that M_{AS} is nonsingular, the corresponding eigenvector is a generator of the one-dimensional $\mathcal{N}(A)$. It follows then (see, e.g., [22, section 4.2]), that the convergence factor (3.2) of the additive Schwarz iteration can be characterized as

$$\begin{aligned} \gamma(T_{AS,\theta}) &= \max_{\substack{z \perp \mathcal{N}(A) \\ z^T z = 1}} z^T T_{AS,\theta} z \\ &= \max_{\substack{z \perp \mathcal{N}(A) \\ (z,z)=1}} \left(1 - \theta \sum_{i=1}^{p} (R_i^T A_i^{-1} R_i z, Az) \right) \\ &= 1 - \theta \left(\min_{\substack{z \perp \mathcal{N}(A) \\ (z,z)=1}} \sum_{i=1}^{p} (R_i^T A_i^{-1} R_i z, Az) \right). \end{aligned} \tag{4.3}$$

We note that on the subspace $\mathcal{N}(A)^\perp$, the matrix A is positive definite. Let us call $\hat{A} = A|_{\mathcal{N}(A)^\perp}$, and we can thus replace A with $\hat{A}$ in (4.3). Furthermore, since $\hat{A}^{1/2}$ is invertible, we can write $w = \hat{A}^{1/2} z$, and write (4.3) as

$$\gamma(T_{AS,\theta}) = 1 - \theta \left(\min_{\substack{\hat{A}^{-1/2} w \perp \mathcal{N}(A) \\ (w, \hat{A}^{-1} w)=1}} \sum_{i=1}^{p} w^T \hat{A}^{-1/2} R_i^T A_i^{-1} R_i \hat{A}^{1/2} w \right). \tag{4.4}$$

We point out that the characterization (4.4) is also valid for the case of A symmetric positive definite, in which case we have $\hat{A} = A$.

5. Convergence of multiplicative Schwarz. We begin with an important auxiliary result.

LEMMA 5.1. *Let A be a symmetric positive semidefinite matrix such that each principal submatrix is positive definite. Let $x,\ y \in \mathbb{R}^n$, such that*

$$y = (I - E_i M_i^{-1} A)x, \tag{5.1}$$

where E_i is defined in (2.10) *and M_i in* (2.12)*. Then the following holds:*

$$y^T Ay - x^T Ax = -(y-x)^T E_i A E_i (y-x) \leq 0. \tag{5.2}$$

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 (2.10) and (2.8) we have that

$$E_i = \pi_i^T \begin{bmatrix} I_i & O \\ O & O \end{bmatrix} \pi_i. \tag{5.3}$$

Consider now (5.1), whence we immediately have that

$$y_2 = x_2, \tag{5.4}$$

and using (2.12) and (2.9), we also get

$$A_i y_1 = -A_{12} x_2, \tag{5.5}$$

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

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

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 $A \succeq O$, $E_i A E_i$ is semidefinite as well, and the right-hand side of (5.2) is nonpositive. ∎

THEOREM 5.2. *Let A be a symmetric positive semidefinite matrix such that each principal submatrix is positive definite. Let $b \in \mathcal{R}(A)$ and $x_0 \notin \mathcal{N}(A)$. Then the multiplicative Schwarz iteration defined by* (2.6) *is convergent.*

Proof. We need to prove that the iteration matrix $T = T_{MS}$ is convergent; i.e., we need to prove conditions (1), (2), and (3) of Definition 3.1.

(1) Starting with $z = x^{(1)} \notin \mathcal{N}(A)$, let $x^{(i+1)} = (I - P_i)x^{(i)}$. Thus $x^{(p+1)} = Tx^{(1)}$. Using (5.2) repeatedly, and canceling terms, we obtain

$$\begin{aligned} z^T T^T A T z - z^T A z &= -\sum_{i=1}^{p} (x^{(i+1)} - x^{(i)})^T E_i A E_i (x^{(i+1)} - x^{(i)}) \\ &= -\sum_{i=1}^{p} ((x^{(i+1)} - x^{(i)})^T E_i) E_i A E_i (E_i (x^{(i+1)} - x^{(i)})). \end{aligned} \tag{5.6}$$

Since $E_i A E_i$ is positive definite it follows that the right-hand side of (5.6) 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 5.1 to obtain (5.4). 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 (5.6) must be negative, and we obtain

$$z^T T^T A T z - z^T A z < 0.$$

Thus we have that for all $\lambda \in \sigma(T)$ with corresponding eigenvector $y \notin \mathcal{N}(A)$

$$\lambda^2 y^T A y - y^T A y < 0. \tag{5.7}$$

Hence $\lambda^2 - 1 < 0$. Thus

$$|\lambda| < 1.$$

If $\lambda \in \sigma(T)$ but the corresponding eigenvector $y \in \mathcal{N}(A)$, we easily obtain from the definition of T that $\lambda = 1$. Hence, $\rho(T) \leq 1$.

(2) By (3.1), it suffices to prove that $\mathcal{N}(I - T) \cap \mathcal{R}(I - T) = \{0\}$. Here we have that $\mathcal{N}(A) = \mathcal{N}(I - T)$. This holds since $y \notin \mathcal{N}(A)$ implies $Ty \neq y$ by part (1), i.e., $y \notin \mathcal{N}(I - T)$. On the other hand $y \in \mathcal{N}(A)$ implies $y \in \mathcal{N}(I - T)$, using the definition of T; cf. Remark 2.1. Hence, we need to prove that

$$\mathcal{N}(A) \cap \mathcal{R}(I - T) = \{0\}. \tag{5.8}$$

Let $x \in \mathcal{N}(A) \cap \mathcal{R}(I - T)$. Then there exists a y with $(I - T)y = x$, i.e., $y = Ty + x$. Since $x \in \mathcal{N}(A)$ we obtain

$$A(I - T)y = Ax = 0, \quad \text{and thus} \quad y^T A y - y^T A T y = 0.$$

Using $y = Ty + x$ we get

$$y^T A y - y^T T^T A T y + x^T A T y = y^T A y - y^T T^T A T y = 0.$$

Part (1) of this proof now implies $y \in \mathcal{N}(A)$; cf. (5.7). Therefore, by Remark 2.1, $x = (I - T)y = 0$, which completes this part of the proof.

(3) As proved above we have $\lambda < 1$ for all $\lambda \in \sigma(T)$ with corresponding eigenvector $y \notin \mathcal{N}(A)$. Thus if $|\lambda| = 1$ for some eigenvalue λ of T then the corresponding eigenvector y must be in the null-space of A. Hence $Ay = 0$. But then $Ty = y$ and thus $\lambda = 1$. ◻

We mention that we need to prove explicitly (5.8) since we do not have an explicit representation of a nonsingular matrix M_{MS} such that $M_{MS}^{-1} A = I - T_{MS}$. The existence of such a matrix, i.e., of a splitting induced by T_{MS} [2] is only obtained after the theorem is proved. Any splitting induced by such a matrix M_{MS} is thus P-regular.

We also comment on the fact that in some cases one may want to have a symmetric operator, and in such a case, the natural multiplicative operator is

$$T_{SMS} = (I - P_1)(I - P_2) \cdots (I - P_{p-1})(I - P_p)(I - P_{p-1}) \cdots (I - P_1). \tag{5.9}$$

It follows that Theorem 6.1 applies to this case as well, and that a posteriori, there exists a nonsingular matrix M_{SMS} such that $M_{SMS}^{-1} A = I - T_{SMS}$. We can characterize the convergence factor (3.2) of this symmetric multiplicative Schwarz iteration as

$$\gamma = \gamma(T_{SMS}) = \max_{\substack{z \perp \mathcal{N}(A) \\ z^T z = 1}} (z, T_{SMS} z). \tag{5.10}$$

6. Inexact local solvers. In this section we study the effect of varying how exactly (or inexactly) the local problems are solved. The convergence of these very practical versions of the methods is based on the same ideas used to prove that of the standard Schwarz iterations in sections 4 and 5. The influence of different levels of inexactness is analyzed using our comparison theorem, Theorem 3.6.

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., [6], [41], [44]. The expression $\tilde{A}_i^{-1} z_i$ often represents an approximation to the solution of the system $A_i z_i = v_i$ using some steps of an (inner) iterative method. By replacing A_i with $\tilde{A}_i$ in (2.4) one obtains the damped additive Schwarz iterations with inexact local solvers, and its iteration matrix is then

$$\tilde{T}_{AS,\theta} = I - \theta \sum_{i=1}^{p} R_i^T \tilde{A}_i^{-1} R_i A. \tag{6.1}$$

The iteration matrices $T_{AS,\theta}$ and $\tilde{T}_{AS,\theta}$ in (2.4) and (6.1) are induced by splittings $A = M_\theta - N_\theta$ and $A = \tilde{M}_\theta - \tilde{N}_\theta$ where

$$M_\theta^{-1} = \theta \sum_{i=1}^{p} R_i^T A_i^{-1} R_i = \theta \sum_{i=1}^{p} E_i M_i^{-1} \succ O, \tag{6.2}$$

$$\tilde{M}_\theta^{-1} = \theta \sum_{i=1}^{p} R_i^T \tilde{A}_i^{-1} R_i = \theta \sum_{i=1}^{p} E_i \tilde{M}_i^{-1} \succ O. \tag{6.3}$$

Here

$$\tilde{M}_i = \pi_i^T \begin{bmatrix} \tilde{A}_i & O \\ O & D_{\neg i} \end{bmatrix} \pi_i, \quad \text{and thus} \quad \tilde{M}_i^{-1} = \pi_i^T \begin{bmatrix} \tilde{A}_i^{-1} & O \\ O & D_{\neg i}^{-1} \end{bmatrix} \pi_i. \tag{6.4}$$

The fact that the matrix (6.3) is nonsingular follows in the same manner as in the proof that (6.2) is nonsingular in Theorem 4.2.

In the case considered in this paper we assume, as is generally done (see, e.g., [21, Ch. 11.2.4]), that the inexact local solvers correspond to symmetric positive definite matrices and satisfy

$$\tilde{A}_i \succeq A_i. \tag{6.5}$$

For examples of splittings for which the inequality (6.5) holds, see, e.g., [33]. A situation worth mentioning where (6.5) holds is when A_i is semidefinite and the inexact local solver is definite. This process is usually called regularization; see, e.g., [15], [25].

THEOREM 6.1. *Let A be a symmetric positive semidefinite matrix such that each principal submatrix is positive definite. Let $b \in \mathcal{R}(A)$ and $x_0 \notin \mathcal{N}(A)$. Let $\tilde{A}_i$ and $\bar{A}_i$ be inexact local solvers of A_i satisfying $\tilde{A}_i \succeq \bar{A}_i \succeq A_i$. Let $\bar{T}_{AS,\theta}$ be obtained by replacing $\tilde{A}_i$ by $\bar{A}_i$ in (6.1), $i = 1, \ldots, p$. Let the damping factor $0 < \theta < 2/p$. Then the inexact additive Schwarz iterations defined by (6.1) and $\bar{T}_{AS,\theta}$ are convergent, and the splittings induced by these iteration matrices are P-regular. With the stronger hypothesis that $0 < \theta < 1/p$, we also have that $\gamma(T_{AS,\theta}) \le \gamma(\bar{T}_{AS,\theta}) \le \gamma(\tilde{T}_{AS,\theta})$, and the splittings induced by these iteration matrices are strongly P-regular.*

Proof. Since $\tilde{A}_i \succeq A_i$ we have

$$\tilde{A}_i^{-1} \preceq A_i^{-1}, \tag{6.6}$$

and thus, using Lemma 4.1

$$A^{\frac{1}{2}} R_i^T \tilde{A}_i^{-1} R_i A^{\frac{1}{2}} \preceq A^{\frac{1}{2}} R_i^T A_i^{-1} R_i A^{\frac{1}{2}} \preceq I.$$

Similar inequalities are obtained with $\bar{A}_i$. The rest of the convergence proof proceeds in the same manner as that of Theorem 4.2.

Consider the matrices (6.2) and (6.3) which are symmetric positive definite using M_i as in (2.12) and $\tilde{M}_i$ as in (6.4). From (6.6), we have that $M_\theta^{-1} \succeq \tilde{M}_\theta^{-1} \succ O$. This implies $M_\theta \preceq \tilde{M}_\theta$ and $N_\theta \preceq \tilde{N}_\theta$. By Remark 4.3, we have that $N_\theta \succ O$, i.e., that the splittings are strong P-regular. The same results are obtained in the case of $\bar{A}_i$. The theorem follows from Theorem 3.6. ☐

As was the case with Theorem 4.2, we can replace p in the restriction on the damping parameter with q, the number of colors; i.e., we guarantee convergence of additive Schwarz with inexact local solvers for $\theta < 2/q$. Since Theorem 6.1 applies in particular to the symmetric positive definite case, we have again double the interval of admissible damping factors for the additive Schwarz iteration with inexact local solvers; cf. [1].

Remark 6.2. An alternative proof of the second part of Theorem 6.1 can be obtained by considering the two convergence factors, $\gamma(T_{AS,\theta})$ given by (4.4) for the exact case, and the second given by

$$\gamma(\tilde{T}_{AS,\theta}) = 1 - \theta \left(\min_{\substack{\hat{A}^{-1/2} w \perp \mathcal{N}(A) \\ (w, \hat{A}^{-1} w) = 1}} \sum_{0=1}^{p} w^T \hat{A}^{-1/2} R_i^T \tilde{A}_i^{-1} R_i \hat{A}^{1/2} w \right) \tag{6.7}$$

for the inexact case. Since $\sigma(\hat{A}^{-1/2} R_i^T A_i^{-1} R_i \hat{A}^{1/2}) = \{0\} \cup \sigma(A_i^{-1})$ and $\sigma(\hat{A}^{-1/2} R_i^T \tilde{A}_i^{-1} R_i \hat{A}^{1/2}) = \{0\} \cup \sigma(\tilde{A}_i^{-1})$, and since $-\tilde{A}_i^{-1} \succeq -A_i^{-1}$, we have that

$$-w^T \hat{A}^{-1/2} R_i^T \tilde{A}_i^{-1} R_i \hat{A}^{1/2} w \;\geq\; -w^T \hat{A}^{-1/2} R_i^T A_i^{-1} R_i \hat{A}^{1/2} w, \quad i = 1, \ldots, p,$$

which implies that $\gamma(\tilde{T}_{AS,\theta}) \geq \gamma(T_{AS,\theta})$.

For simplicity, in Theorem 6.1, 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.

The implication of Theorem 6.1 is that by replacing the local solvers A_i with the approximate counterparts $\tilde{A}_i$, the additive Schwarz iteration is expected to take more iterations. In practice, a solve with $\tilde{A}_i$ should be sufficiently less expensive so that the overall method is cheaper.

Next we consider the multiplicative Schwarz method with inexact local solvers on the subdomains. Here we assume that the approximations $\tilde{A}_i$ satisfy

$$\tilde{A}_i + \tilde{A}_i^T - A_i \succ 0. \tag{6.8}$$

This assumption implies that

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

Using (6.4), the inexact multiplicative Schwarz iteration matrix is given by

$$\tilde{T} = (I - E_p \tilde{M}_p^{-1} A)(I - E_{p-1} \tilde{M}_{p-1}^{-1} A) \cdots (I - E_1 \tilde{M}_1^{-1} A). \tag{6.9}$$

LEMMA 6.3. *Let A be a symmetric positive semidefinite 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* (6.4) *with $\tilde{A}_i$ satisfying* (6.8). *Then the following identity holds:*

$$-(y-x)^T E_i(\tilde{M}_i^T + \tilde{M}_i - A)E_i(y-x) \le 0. \tag{6.10}$$

Proof. The proof proceeds as that of Lemma 5.1. We have that (5.4) holds, but instead of (5.5) we have $\tilde{A}_i y_1 = (\tilde{A}_i - A_i)x_1 - A_{12}x_2$. We then obtain

$$\begin{aligned} y^T Ay - x^T Ax &= x_2^T A_{21}(y_1 - x_1) + (y_1^T - x_1^T)A_{12}x_2 + y_1^T A_i y_1 - x_1^T A_i x_1 \\ &= (x_1^T(\tilde{A}_i - A_i)^T - y_1^T\tilde{A}_i^T)(y_1 - x_1) \\ &\quad + (y_1^T - x_1^T)((\tilde{A}_i - A_i)x_1 - \tilde{A}_i y_1) + y_1^T A_i y_1 - x_1^T A_i x_1 \\ &= (-x_1^T A_i - (y_1^T - x_1^T)\tilde{A}_i^T)(y_1 - x_1) \\ &\quad + (y_1^T - x_1^T)(-A_i x_1 - \tilde{A}_i(y_1 - x_1)) + y_1^T A_i y_1 - x_1^T A_i x_1 \\ &= -(y_1^T - x_1^T)(\tilde{A}_i + \tilde{A}_i^T - A_i)(y_1 - x_1) \\ &= -(y-x)^T E_i(\tilde{M}_i^T + \tilde{M}_i - A)E_i(y-x) \le 0, \end{aligned}$$

where the last inequality follows from (6.8) and the form of the matrices $\tilde{M}_i$ in (6.4). ◻

THEOREM 6.4. *Let A be a symmetric positive semidefinite matrix such that each principal submatrix is positive definite. Let $b \in \mathcal{R}(A)$ and $x_0 \notin \mathcal{N}(A)$. Then the multiplicative Schwarz iteration with iteration matrix* (6.9) *with $\tilde{M}_i$ defined in* (6.4) *and with inexact local solvers satisfying* (6.8) *converges to the solution of $Ax = b$.*

Proof. We need to prove that the iteration matrix $\tilde{T}$ is convergent; i.e., we need to prove conditions (1), (2), and (3) of Definition 3.1. The proof is similar to the proof of Theorem 5.2. The only difference appears in proving condition (1). Here we use Lemma 6.3 and obtain

$$z^T\tilde{T}^T A\tilde{T}z - z^T Az < 0$$

for all $z \notin \mathcal{N}(A)$, and the rest of the proof follows. ◻

A symmetric version of multiplicative Schwarz with inexact local solvers can also be constructed in a way similar to (5.9), and its convergence factor can be characterized in a way similar to (5.10).

We mention that a comparison analogous to that of the second part of Theorem 6.1 is not valid for multiplicative Schwarz, not even in the definite case. A counterexample can be found in [40].

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

Let us consider two sets of subblocks (subdomains) of the matrix A, as defined by the sets (2.11), such that one has more overlap than the other; i.e., let

$$\hat{S}_i \supseteq S_i, \quad i = 1, \ldots, p, \tag{7.1}$$

with $\bigcup_{i=1}^p \hat{S}_i = \bigcup_{i=1}^p S_i = S$. 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 (2.10). The relation (7.1) implies that

$$I \succeq \hat{E}_i \succeq E_i \succeq O. \tag{7.2}$$

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 (2.12) define

$$\hat{M}_i = \hat{\pi}_i^T \begin{bmatrix} \hat{A}_i & O \\ O & \hat{D}_{\neg i} \end{bmatrix} \hat{\pi}_i, \tag{7.3}$$

where $\hat{D}_{\neg i} = \text{diag}(\hat{A}_{\neg i}) \succ O$, and $\hat{A}_{\neg i}$ is the $(n - \hat{n}_i) \times (n - \hat{n}_i)$ complementary principal submatrix of A as in (2.9). As in (2.13), 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 $\hat{\pi}_i = \pi_i$. Thus, A_i is a principal submatrix of $\hat{A}_i$, and $\hat{M}_i$ has the same diagonal as M_i.

We will apply to these the following result for symmetric positive definite matrices which can be found, e.g., in [21].

LEMMA 7.1. *Let A be a symmetric positive definite matrix and the form of the matrices $\tilde{M}_i$ in* (6.4). *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 $R_i^T A_i^{-1} R_i \preceq A^{-1}$.*

We consider the case of damped additive Schwarz with iteration matrix (2.4), and the iteration matrix corresponding to the larger overlap is

$$\hat{T}_{AS,\theta} = I - \theta \sum_{i=1}^{p} \hat{R}_i^T \hat{A}_i^{-1} \hat{R}_i A. \tag{7.4}$$

THEOREM 7.2. *Let A be a symmetric positive semidefinite matrix such that each principal submatrix is positive definite. Let $b \in \mathcal{R}(A)$ and $x_0 \notin \mathcal{N}(A)$. Consider two sets of subblocks of A defined by* (7.1), *and the two corresponding additive Schwarz iterations* (2.4) *and* (7.4). *Let the damping factor $\theta \leq 1/p$, which implies in particular that the additive Schwarz methods are convergent. Then, $\gamma(\hat{T}_\theta) \leq \gamma(T_\theta)$.*

Proof. As mentioned above assume that all the principal submatrices of A of order less than n are nonsingular. 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 7.1 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 Remark 4.3, these splittings are strong P-regular, and the theorem follows from Theorem 3.6. ∎

We note that an alternative proof similar to that in Remark 6.2 can be applied here, using the relation $\hat{R}_i^T \hat{A}_i^{-1} \hat{R}_i = \hat{Q}_i \succeq Q_i = R_i^T A_i^{-1} R_i$ just proved.

Theorem 7.2 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., [4], [41]. Of course, the faster convergence rate brings an associated increased cost of the local solvers, 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.

We should mention that with an increase of overlap, the number of colors of the graph may decrease, so that the damping factor may need to be revised. In all cases, the maximum restriction is $\theta < 1/p$.

A comparison analogous to that of Theorem 7.2 is not valid for multiplicative Schwarz, not even in the definite case. A counterexample can be found in [40].

8. 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? We show that for the additive Schwarz method the more subblocks (subdomains), the slower the convergence. In a limiting case, if we have a single variable in each block and there is no overlap, this is the classic Jacobi method, and our results indicate that this has asymptotically slower convergence than any sets of blocks for additive Schwarz.

As in the situations described in sections 6 and 7, the slower convergence may be partially compensated by less expensive local solvers, since they are of smaller dimension.

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

$$S_{i_j} \subset S_i, \quad j = 1, \ldots, k_i, \tag{8.1}$$

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

$$E_{i_j} \preceq E_i, \quad j = 1, \ldots, k_i, \quad \text{and} \quad \sum_{j=1}^{k_i} E_{i_j} = E_i, \; i = 1, \ldots, p. \tag{8.2}$$

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., (7.3)), 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

$$\bar{T}_\theta = I - \theta \sum_{i=1}^{p} \sum_{j=1}^{k_i} E_{i_j} M_{i_j}^{-1} A \tag{8.3}$$

(cf. (2.4)), and an induced strong P-splitting (assuming the proper restriction on θ) $A = \bar{M}_\theta - \bar{N}_\theta$ is given by

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

THEOREM 8.1. *Let A be a symmetric positive semidefinite matrix such that each principal submatrix is positive definite. Let $b \in \mathcal{R}(A)$ and $x_0 \notin \mathcal{N}(A)$. Consider two sets of subblocks of A defined by* (2.11) *and* (8.1), *respectively, and the two corresponding additive Schwarz iterations defined by* (2.4) *and* (8.3). *Let $k = \max_i k_i$, and let the damping factors be $\theta \leq 1/p$, and $\bar{\theta} = \theta/k \leq 1/(kp)$. This implies that in particular the additive Schwarz methods are convergent. Then, $\gamma(T_\theta) \leq \gamma(\bar{T}_{\bar{\theta}})$.*

Proof. As in the proof of Theorem 7.2 we have, using Lemma 7.1, that

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

Therefore, $\sum_{j=1}^{k_i} Q_{i_j} \preceq k_i Q_i$, 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 Theorem 3.6 and the fact that these are strong P-regular splittings, as shown in Remark 4.3. □

As in the previous sections a comparison analogous to that of Theorem 8.1 is not valid for multiplicative Schwarz, not even in the definite case. Again, a counterexample can be found in [40].

9. Two-level schemes. We consider now two-level schemes, i.e., those in which an additional step is taken, corresponding to a coarse grid correction. In the nonsingular case, this additional step makes Schwarz methods optimal in the sense that the condition number of the preconditioned matrix $M^{-1}A$ is independent of the mesh size; see, e.g., [38], [41], [44]. In our setting, for the coarse grid correction consider an additional subspace V_0 of V, and the corresponding projection $P_0 = R_0^T A_0^{-1} R_0 A = R_0^T (R_0 A R_0^T)^{-1} R_0 A$. There are several cases we consider here: additive Schwarz with coarse grid correction, with iteration matrix given by

$$T_{ASc,\theta} = T_{AS,\theta} - \theta R_0^T A_0^{-1} R_0 A = I - \theta \sum_{i=0}^{p} R_i^T A_i^{-1} R_i A = I - \theta \sum_{i=0}^{p} P_i; \tag{9.1}$$

multiplicative Schwarz with coarse grid correction, with iteration matrix given by

$$T_{MSc} = T_{MS}(I - P_0) = \prod_{i=p}^{0} (I - P_i),$$

or in the symmetrized case by $T_{SMSc} = (I - P_0) T_{SMS} (I - P_0)$; multiplicative Schwarz additively corrected, known as the two-level hybrid I Schwarz method, with iteration matrix given by

$$H_{I,\theta} = I - \theta P_0 - \theta(I - T_{MS}) = I - \theta(G_0 + M_{MS}^{-1})A,$$

where $G_0 = R_0^T A_0^{-1} R_0$; and the two-level hybrid II Schwarz method, which is additive Schwarz multiplicatively corrected, with iteration matrix given by

$$H_{II,\theta} = T_{AS,\theta}(I - P_0).$$

We begin our analysis with the additive Schwarz iteration with coarse grid correction. By comparing the iteration matrices in (9.1) and (2.4), one can see that

Theorem 4.2 is valid in this case as well, with the exception that the damping factor θ needs to be less than $2/(p+1)$. Therefore we have that the matrix $T_{ASc,\theta}$ is a convergent matrix, and that the induced splitting defined by $M_{ASc,\theta}^{-1} = \theta \sum_{i=0}^{p} R_i^T A_i^{-1} R_i$ is P-regular. We can also show that coarse grid correction does not increase (and may decrease) the convergence factor of the iterations.

THEOREM 9.1. *Let A be a symmetric positive semidefinite matrix such that each principal submatrix is positive definite. Then $\gamma(T_{ASc,\theta}) \leq \gamma(T_{AS,\theta})$.*

Proof. We use the fact that $G_0 = R_0^T A_0^{-1} R_0 \succeq 0$ to conclude that

$$M_{ASc,\theta}^{-1} = \theta(M_{AS}^{-1} + G_0) \succeq \theta M_{AS}^{-1} \ .$$

The theorem now follows by the application of Theorem 3.6. □

A characterization similar to (4.4) applies to this two-level method, with one more term in the sum. Thus, an alternative proof of this theorem using this characterization can be done in a manner similar to that in Remark 6.2.

Next, we consider the multiplicative Schwarz iterations with coarse grid correction. It is not hard to see that Theorem 5.2 applies to this case as well, so that T_{MSc} and T_{SMSc} are convergent.

We conclude by mentioning that the coarse grid corrections can be applied to the methods with inexact solvers described in section 6 as well, and since the analysis is very similar, we do not repeat it.

Acknowledgments. We thank Michele Benzi and the referees for their comments on an earlier version of this paper, which helped improve our presentation.

REFERENCES

[1] M. Benzi, A. Frommer, R. Nabben, and D. B. Szyld, *Algebraic theory of multiplicative Schwarz methods*, Numer. Math., 89 (2001), pp. 605–639.

[2] M. Benzi and D. B. Szyld, *Existence and uniqueness of splittings for stationary iterative methods with applications to alternating methods*, Numer. Math., 76 (1997), pp. 309–321.

[3] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Classics Appl. Math. 9, SIAM, Philadelphia, 1994.

[4] P. E. Bjørstad and O. B. Widlund, *To overlap or not to overlap: A note on a domain decomposition method for elliptic problems*, SIAM J. Sci. Statist. Comput., 10 (1989), pp. 1053–1061.

[5] P. Bochev and R. B. Lehoucq, *On the finite element solution of the pure Neumann problem*, SIAM Rev., 47 (2005), pp. 50–66.

[6] J. H. Bramble, J. E. Pasciak, and A. T. Vassilev, *Analysis of non-overlapping domain decomposition algorithms with inexact solves*, Math. Comp., 67 (1998), pp. 1–19.

[7] R. Bru, F. Pedroche, and D. B. Szyld, *Additive Schwarz iterations for Markov chains*, SIAM J. Matrix Anal. Appl., 27 (2005), pp. 445–458.

[8] T. F. Chan and T. P. Mathew, *Domain decomposition algorithms*, in Acta Numerica 1994, Acta Numer., Cambridge University Press, Cambridge, UK, 1994, pp. 61–143.

[9] J.-J. Climent and C. Perea, *Some comparison theorems for weak nonnegative splittings of bounded operators*, Linear Algebra Appl., 275–276 (1998), pp. 77–106.

[10] G. Csordas and R. S. Varga, *Comparisons of regular splittings of matrices*, Numer. Math., 44 (1984), pp. 23–35.

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

[12] M. Dryja and O. B. Widlund, *An additive variant of the Schwarz alternating method for the case of many subregions*, Technical Report 339, Ultracomputer Note 131, Department of Computer Science, Courant Institute, New York University, New York, NY, 1987.

[13] M. Dryja and O. B. Widlund, *Some domain decomposition algorithms for elliptic problems*, in Iterative Methods for Large Linear Systems, Academic Press, San Diego, 1989, pp. 273–291.

[14] M. Dryja and O. B. Widlund, *Towards a unified theory of domain decomposition algorithms for elliptic problems*, in Proceedings of the Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. F. Chan, R. Glowinski, J. Périaux, and O. B. Widlund, eds., SIAM, Philadelphia, 1990, pp. 3–21.

[15] M. Dryja and O. B. Widlund, *Schwarz methods of Neumann-Neumann type for three-dimensional elliptic finite element problems*, Comm. Pure Appl. Math., 48 (1995), pp. 121–155.

[16] L. Elsner, *Comparisons of weak regular splittings and multisplitting methods*, Numer. Math., 56 (1989), pp. 283–289.

[17] R. W. Freund and M. Hochbruck, *On the use of two QMR algorithms for solving singular systems and applications in Markov Chain modelling*, Numer. Linear Algebra Appl., 1 (1994), pp. 403–420.

[18] A. Frommer and D. B. Szyld, *Weighted max norms, splittings, and overlapping additive Schwarz iterations*, Numer. Math., 83 (1999), pp. 259–278.

[19] A. Frommer and D. B. Szyld, *An algebraic convergence theory for restricted additive Schwarz methods using weighted max norms*, SIAM J. Numer. Anal., 39 (2001), pp. 463–479.

[20] M. Griebel and P. Oswald, *On the abstract theory of additive and multiplicative Schwarz algorithms*, Numer. Math., 70 (1995), pp. 163–180.

[21] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations*, Springer-Verlag, New York, 1994.

[22] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.

[23] W. J. Kammerer and M. Z. Nashed, *On the convergence of the conjugate gradient method for singular linear operator equations*, SIAM J. Numer. Anal., 9 (1972), pp. 165–181.

[24] H. B. Keller, *On the solution of singular and semidefinite linear systems by iteration*, SIAM J. Numer. Anal., 2 (1965), pp. 281–290.

[25] A. Klawonn and O. B. Widlund, *A domain decomposition method with Lagrange multipliers and inexact solvers for linear elasticity*, SIAM J. Sci. Comput., 22 (2000), pp. 1199–1219.

[26] Y.-J. Lee, J. Wu, J. Xu, and L. Zikatanov, *A Sharp Convergence Estimate of the Method of Subspace Corrections for Singular Systems*, Technical Report AM259, Department of Mathematics, Pennsylvania State University, State College, PA, 2002.

[27] Y.-J. Lee, J. Wu, J. Xu, and L. Zikatanov, *On the convergence of iterative methods for semidefinite linear systems*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 634–641.

[28] Y.-J. Lee, J. Xu, and L. Zikatanov, *Successive subspace correction method for singular system of equations*, in Proceedings of the Fourteenth International Conference on Domain Decomposition Methods, I. Herrera, D. E. Keyes, O. B. Widlund, and R. Yates, eds., UNAM Press, Mexico City, Mexico, 2003, pp. 315–321. Also available online at http://www.ddm.org.

[29] I. Marek and D. B. Szyld, *Comparison theorems for weak splittings of bounded operators*, Numer. Math., 58 (1990), pp. 387–397.

[30] I. Marek and D. B. Szyld, *Comparison theorems for the convergence factor of iterative methods for singular matrices*, Linear Algebra Appl., 316 (2000), pp. 67–87.

[31] I. Marek and D. B. Szyld, *Comparison of convergence of general stationary iterative methods for singular matrices*, SIAM J. Matrix Anal. Appl., 24 (2002), pp. 68–77.

[32] I. Marek and D. B. Szyld, *Algebraic Schwarz methods for the numerical solution of Markov chains*, Linear Algebra Appl., 386 (2004), pp. 67–81.

[33] R. Nabben, *A note on comparison theorems of splittings and multisplittings of Hermitian positive definite matrices*, Linear Algebra Appl., 233 (1996), pp. 67–80.

[34] R. Nabben, *Comparisons between additive and multiplicative Schwarz iterations in domain decomposition methods*, Numer. Math., 95 (2003), pp. 145–162.

[35] R. Nabben and D. B. Szyld, *Convergence theory of restricted multiplicative Schwarz methods*, SIAM J. Numer. Anal., 40 (2003), pp. 2318–2336.

[36] J. M. Ortega, *Numerical Analysis: A Second Course*, Classics Appl. Math. 3, SIAM, Philadelphia, 1990.

[37] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Classics Appl. Math. 30, SIAM, Philadelphia, 2000.

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

[39] L. Reichel and Q. Ye, *Breakdown-free GMRES for singular systems*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 1001–1021.

[40] M. Schnitker, *Eine algebraische Konvergenztheorie der Schwarz-Verfahen für symmetrisch positiv definite Matrizen*, Examensarbeit, Universität Bielefeld, Bielefeld, Germany, 2002.

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

[42] W. J. Stewart, *Introduction to the Numerical Solution of Markov Chains*, Princeton University Press, Princeton, NJ, 1994.

[43] D. B. Szyld, *Equivalence of convergence conditions for iterative methods for singular equations*, Numer. Linear Algebra Appl., 1 (1994), pp. 151–154.

[44] A. Toselli and O. Widlund, *Domain Decomposition Methods—Algorithms and Theory*, Springer Series in Computational Mathematics 34, Springer-Verlag, Berlin, 2005.

[45] R. S. Varga, *Matrix Iterative Analysis. Second Revised and Expanded Edition*, Springer-Verlag, Berlin, 2000.

[46] Z. I. Woźnicki, *Nonnegative splitting theory*, Japan J. Indust. Appl. Math., 11 (1994), pp. 289–342.

[47] J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Rev., 34 (1992), pp. 581–613.